

Teemu Kyttälä

**Solid-State Guitar
Amplifiers**

Copyright © 2008 Teemu Kyttälä. All rights reserved. No part of this publication may be reproduced in any form or by any means without the prior written permission of the author.

All trademarks are trademarks of their respective owners. The names are used in editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Solid-State Guitar Amplifiers

1st Edition

Acknowledgements

A heartfelt thanks goes out to Nigel Goodwin, Don Booth and Joseph Raymond who assisted in proofreading this book. They should not however be held responsible for any errors or oversights. Another heartfelt thanks goes out to Hugh Dean, whose contributions to diyAudio forum have more than once proven priceless. I also like to thank everybody who showed interest in this project.

I also like to express my gratitude to enthusiastic individuals such as Rod Elliott, R.G. Keen, Paul Kemble, Randall Aiken, John Broskie, Steve Bench, Duncan Munro, M.J. Renardson and Malcolm Hawksford. Your dedication and effort to present such vast databases of audio electronics-related information in the Internet has not been left unnoticed. My heartfelt thanks also goes out to the legion of people hosting audio- and electronics -related forums or schematic sharing sites.

...

I apologize that in the current state the grammatical accuracy of this book's content is still somewhat "lacking". English is not my native language and all proofreading has been done on voluntarily basis, which, unfortunately means that parts of the book have not been under the scope of the people who master the language better than I do. So, please, try to bear my "Ingrish". If you have comments, suggestions and corrections feel free to sent them to *teemu[spamfilter]kyttala@gmail.com*. They will be appreciated.

1	INTRODUCTION.....	9
1.1	FOREWORD	9
1.2	ABOUT THE SCHEMATIC NOTATION	10
1.3	COMMON TERMINOLOGY	11
1.4	ABOUT INTEGRATED CIRCUIT AMPLIFIERS.....	18
1.5	TUBES VERSUS TRANSISTORS	18
1.6	AMPLIFIER’S TONE – OR IS IT?.....	26
1.7	OPERATIONAL AMPLIFIERS VERSUS DISCRETE CIRCUITS.....	27
1.8	USE OF DSP IN AMPLIFIERS	28
1.9	DECIBEL REFERENCES	30
1.10	ABOUT VARIOUS TYPES OF DISTORTION	30
1.11	SOME HISTORY OF (TRANSISTORIZED) GUITAR AMPLIFIERS.....	31
1.12	THE DILEMMA OF “CLASSIC” AMPLIFIERS	40
1.13	THE IDEAL OF SIMPLICITY	41
2	A BASIC INTRODUCTORY TO POWER AMPLIFIERS	46
2.1	A QUICK OVERVIEW	46
2.2	THE EFFECT OF OUTPUT IMPEDANCE AND DAMPING FACTOR.....	48
2.3	THE BASIC BUILDING BLOCKS OF A POWER AMPLIFIER	49
2.4	BI-AMPING	51
2.5	REVEALING PERFORMANCE.....	52
3	A DETAILED EXAMINATION OF POWER AMPLIFIERS	55
3.1	VOLTAGE AMPLIFIER STAGES	55
3.2	OUTPUT STAGES.....	57
3.2.1	<i>Class-A</i>	57
3.2.2	<i>Complementary pairs and push-pull operation</i>	58
3.2.3	<i>Quasi complementary</i>	61
3.2.4	<i>Compound follower</i>	62
3.2.5	<i>Totem pole and topologies using output transformers</i>	64
3.3	INPUT STAGES	66
3.3.1	<i>Single transistor</i>	66
3.3.2	<i>Combined VAS and input</i>	68
3.3.3	<i>Current feedback or “singleton”</i>	68
3.3.4	<i>Long-tailed pair</i>	70
3.4	TECHNIQUES TO IMPROVE LINEARITY.....	73
3.4.1	<i>Balancing the long-tailed pair</i>	73
3.4.2	<i>Symmetric designs: Cascoded complementary LTP</i>	76
3.4.3	<i>Symmetric designs: Cascoded complementary CFB and “Semi-symmetry”</i>	79
3.4.4	<i>Balanced voltage amplifier stages</i>	80
3.5	OPERATIONAL AMPLIFIERS AS DRIVERS	81
3.6	NEGATIVE FEEDBACK LOOPS.....	85
3.6.1	<i>Basics</i>	85
3.6.2	<i>Current sensing and mixed-mode feedback</i>	87
3.6.3	<i>Resonance, Presence and damping controls</i>	90
3.6.4	<i>Feed-forward</i>	92
3.7	CROSSOVER DISTORTION COMPENSATION	94
3.8	CONSTANT CURRENT SOURCE LOADS	100
3.9	BOOTSTRAP CONSTANT CURRENT SOURCES	102
3.10	CASCODING AND RAIL SWITCHING	106

3.11	BRIDGED, PARALLEL AND GROUNDED OUTPUT CIRCUITS	109
3.12	CLASS-A+ AMPLIFIERS	112
3.13	AMPLIFIER STABILITY – PREVENTING OSCILLATION	113
3.14	PROTECTIVE CIRCUITS.....	117
3.14.1	<i>Protecting input stage from static voltage spikes</i>	117
3.14.2	<i>Turn-on delay</i>	118
3.14.3	<i>Current limiting and flyback protection</i>	119
3.14.4	<i>More about short circuit protection</i>	124
3.14.5	<i>Baker Clamps</i>	128
3.14.6	<i>Thermal switches</i>	130
3.14.7	<i>Thyristor latches and crowbars</i>	130
3.14.8	<i>Protecting speakers from DC</i>	133
3.14.9	<i>About fusing</i>	135
3.15	MOSFETS AND IGBTs.....	136
3.16	SIGNAL LIMITING WITHIN POWER AMPLIFIER CIRCUIT	139
3.17	DC SERVOS	142
3.18	IMPEDANCE MATCHING TRANSFORMERS AND 70-VOLT SYSTEMS.....	144
3.19	EXAMPLES OF HYBRID POWER AMPLIFIERS	146
3.19.1	<i>MosValve and its heirs</i>	146
3.19.2	<i>Peavey</i>	149
3.19.3	<i>Music Man</i>	150
3.19.4	<i>ADA</i>	151
3.19.5	<i>Ampeg SVT</i>	154
4	HEAT SINKING	156
4.1	BASICS OF HEAT GENERATION IN A SEMICONDUCTOR	156
4.2	HEAT SINK MATH	157
4.3	INTERFACE PROPERTIES	160
4.4	HEAT SINKS AND COOLING SOLUTIONS.....	161
5	POWER SUPPLIES	167
5.1	TRANSFORMERS	167
5.2	RECTIFIERS AND SUPPLY POLARITY.....	171
5.3	SUPPLY FILTERING	177
5.4	REGULATED SUPPLIES	181
5.5	INRUSH CURRENT AND HOW TO LIMIT IT.....	184
5.6	SURGE LIMITING WITH MOVs.....	185
5.7	POWER SUPPLY OVERVIEW	185
5.8	QUICK INTRODUCTORY TO SWITCHED MODE POWER SUPPLIES	187
5.9	SAFETY PRECAUTIONS.....	189
6	PREAMPLIFIER CIRCUITRY.....	194
6.1	WHAT IS A PREAMPLIFIER?.....	194
6.2	ABOUT TRANSFER CURVES AND CLIPPING CHARACTERISTICS.....	194
6.3	BASIC GAIN STAGES	196
6.3.1	<i>Common emitter</i>	196
6.3.2	<i>Stages using multiple transistors</i>	199
6.3.3	<i>Common source</i>	201
6.3.4	<i>SRPP</i>	206
6.3.5	<i>Discrete buffers</i>	207
6.3.6	<i>Operational amplifiers</i>	209
6.4	MAGNETIC PICKUPS AND INPUT STAGE.....	216

6.5	PIEZOELECTRIC PICKUPS	220
6.6	THE EFFECTS OF FREQUENCY RESPONSE.....	221
6.7	TONE CONTROLS	223
6.7.1	<i>Quick introduction to tone control terminology</i>	223
6.7.2	<i>Active tone controls</i>	224
6.7.3	<i>Scaling impedance and frequency</i>	224
6.7.4	<i>“Pass”-filters and other crude tone controls</i>	225
6.7.5	<i>Bingham</i>	227
6.7.6	<i>Baxandall circuit and its variations</i>	228
6.7.7	<i>“FMV” – tonestack: Fender, Marshall and Vox</i>	230
6.7.8	<i>“Big Muff” tone control and other ideas</i>	232
6.7.9	<i>Utilizing gyrators</i>	235
6.7.10	<i>T-Filters</i>	238
6.7.11	<i>About placement of the tone controls</i>	238
6.7.12	<i>Automatic frequency-compensated volume control</i>	239
6.8	VOLUME- AND GAIN CONTROLS	240
6.9	SPRING REVERBS	242
6.10	EFFECT LOOPS AND LINEOUT CIRCUITS	246
6.11	CLIPPING CIRCUITS.....	250
6.12	SOFT LIMITERS	257
6.13	TREMOLO CIRCUITS.....	264
6.14	CABINET SIMULATORS	267
6.15	PHASE SPLITTERS	268
6.16	EXAMPLES OF TUBE EMULATION CIRCUITS.....	268
6.16.1	<i>TransTube</i>	268
6.16.2	<i>Carvin SX series’ Classic Tube Emulation</i>	274
6.16.3	<i>FlexWave</i>	275
6.16.4	<i>Valve Reactor</i>	276
6.16.5	<i>Valvestate</i>	277
6.16.6	<i>SansAmp</i>	279
6.17	WHAT IS AN “ACOUSTIC AMPLIFIER”?.....	283
6.18	PRACTICAL DESIGN EXAMPLE	286
7	MISCELLANEOUS CIRCUITRY	293
7.1	LIGHT BULB CURRENT LIMITER	293
7.2	FANS	294
7.3	CLIPPING INDICATORS	296
7.4	SWITCHING CIRCUITS	298
7.4.1	<i>Jack switches</i>	298
7.4.2	<i>Switches in remote boxes</i>	299
7.4.3	<i>LDR switches</i>	299
7.4.4	<i>Bipolar transistor and FET-switches</i>	301
7.4.5	<i>Switching OpAmps</i>	305
7.4.6	<i>CMOS-switches, logic circuits and microcontrollers</i>	306
7.4.7	<i>Remote control circuits</i>	309
7.5	A LITTLE BIT ABOUT MIDI.....	311
7.6	SOME USEFUL INTERFACES FOR MICROPROCESSORS.....	312
7.7	RELAYS.....	315
8	EARTHING AND COMMON CIRCUITS	317
8.1	INTRODUCTORY TO GROUNDING AND ITS TERMINOLOGY	317
8.2	COMMON RETURN CURRENT ROUTING AND GROUNDING TOPOLOGIES	318

8.3	SAFETY GROUND CONSIDERATIONS.....	322
8.4	EARTHING AND GROUND LOOPS	324
8.5	“LOOP BREAKER” CIRCUITS	324
9	PRACTICAL COMPONENT CHOICES AND DESIGN ASPECTS.....	326
9.1	RESISTORS	326
9.2	CAPACITORS	327
9.3	POTENTIOMETERS AND TRIMMERS	330
9.4	SWITCHES AND JACKS	332
9.5	SEMICONDUCTORS	333
9.6	POWER SUPPLY PARTS	335
9.7	FUSING.....	336
9.8	PRINTED CIRCUIT BOARD DESIGN	337
9.9	COMPONENT LAYOUT AND HEAT DISSIPATION	344
9.10	ABOUT EPOXY.....	345
9.11	SPICE SIMULATIONS	346
9.12	HOW TO LEARN MORE?	348
10	SPEAKERS.....	350
10.1	POWER RATINGS AND OTHER MISUNDERSTANDINGS RELATED TO SPEAKERS	350
10.2	EFFICIENCY AND SPL RATINGS.....	350
10.3	IMPEDANCE AND FREQUENCY RESPONSE.....	353
10.4	DIRECTIVITY, DIFFRACTION AND DISPERSION	357
10.5	SPEAKER STRUCTURES	358
10.5.1	<i>Dynamic speakers</i>	358
10.5.2	<i>Field coil speakers</i>	360
10.5.3	<i>Piezoelectric and electrostatic speakers</i>	361
10.5.4	<i>Ribbon and planar speakers</i>	361
10.5.5	<i>Flat panel and NXT</i>	362
10.6	HOW A SPEAKER “DIES”	362
10.7	ELECTRICAL SPEAKER MODEL AND THIELE-SMALL PARAMETERS.....	364
11	CABINET DESIGN	368
11.1	MATERIALS.....	368
11.2	CONSTRUCTION TECHNIQUES	371
11.3	ABOUT ACOUSTIC DESIGN OF GUITAR CABINETS	373
11.4	RESONANCE	374
11.5	CABINET SHAPE: DIRECTIVITY, DIFFRACTION, BEAM BLOCKING ETC.....	376
11.6	OPEN-BACKED CABINETS	378
11.7	SEALED CABINETS.....	379
11.8	DUCTED/VENTED CABINETS	380
11.9	PASSIVE RADIATORS	386
11.10	SEALED ENCLOSURES REVISITED.....	386
11.11	HORN STRUCTURES	387
11.12	TRANSMISSION LINE ENCLOSURES.....	388
11.13	COMPOUND OR BAND-PASS CABINETS	389
11.14	ISOBARIC LOADING	390
11.15	ISOLATION CABINETS	390
11.16	EXAMPLES OF CABINET DIMENSIONS.....	391
11.17	TYPICAL SPEAKER WIRING CONFIGURATIONS.....	393
12	CHASSIS CONSTRUCTION.....	397

12.1	COMMON CHASSIS MATERIALS AND TYPES.....	397
12.2	BASICS OF METALWORKING	399
12.2.1	<i>Cutting metal</i>	399
12.2.2	<i>Bending</i>	399
12.2.3	<i>Drilling</i>	400
12.2.4	<i>Polishing and painting</i>	401
12.3	VENTILATION AND OTHER THERMAL-RELATED CONSIDERATIONS.....	402
12.4	EFFECTS OF MECHANICAL VIBRATION	404

1 Introduction

1.1 Foreword

Even these days a shroud of mystery covers many aspects of guitar amplifier design. Plenty of the material written about the topic is controversial or bluntly inaccurate. Even larger amounts of the material is focusing on rather unimportant topics – such as, which tube, transistor, diode, resistor material or capacitor brand provides the best tone. A huge amount of the material out there is written - and later interpreted and rewritten - by people who have little or no knowledge at all about even common electronics. On top of that, marketing departments of various companies are constantly and deliberately spreading a huge amount of propaganda in order to increase their sales and to convince the people that their new product is somehow exceptional, innovative and worth buying. Almost every design out there claims to be either “pristinely clean” or “capturing the tube tone and warmth” – whatever that is. There are so many reasons why so much discussion evolves around “mojo” topics that it would be vain to list them all here; unfortunately these topics also create room for false beliefs that give an impression that guitar amplifier designing is a black art of some kind. It is not. Amplifier designing is - and always will be – pure science based on electronics theory.

While there is plenty of material written about tube amplifier design solid-state amplifiers are mostly left ignored. The case is even worse when it comes to solid-state *guitar* amplifiers. Fortunately, solid-state and tube amplifier circuits share a lot of similarities. The theory is – for the most part – interchangeable. However, many principles used in designing tube amplifiers will not work with transistor circuitry and if they do, they might provide unsatisfactory results. Claims that a solid-state amplifier cannot sound as good as a tube amplifier is utterly untrue and mostly heard from the mouths of three kinds of people: 1) people that have very little or no knowledge about electronics, 2) tube snobs who persistently claim that all transistor amplifiers must sound bad, yet would stick a Fuzz Face (or similar) in front of their tube amplifier without even blinking an eye and 3) experts of tube amplifier circuitry that unfortunately possess a limited knowledge about solid-state design. While someone might master tube technology, the truth is that tube- and solid-state circuitry differ a lot from each other. One should not unreservedly believe statements about transistor circuitry presented by a tube expert – as one should not unreservedly believe statements about tube circuitry presented by a solid-state circuitry expert either. Though the knowledge about both technologies should go hand-in-hand, too often it doesn't. Very regrettable thing is that a great deal of highly talented tube guitar amplifier experts deliberately avoid educating themselves about solid-state as they see it as an inferior technology.

Yet, the fact remains that plenty of people start their electronics hobby by building solid-state audio circuits. Most of these people are repelled by the lethal voltages of tube circuits – or simply by the sheer cost of them. Many people would like to build a solid-state power amplifier but do not know how to do it. The popularity of chip amplifiers is a very good example of this phenomenon since IC-based circuits are easy to assemble even with little or no knowledge of circuit theory at all, what happens when these hobbyists wish to expand their knowledge is that they bump into utterly technical books filled with complex circuit designs and math formulas. While these

books have their purpose they certainly are not very helpful for a beginner. The amount of books that offer an insight to discrete power amplifier design - an insight that would be intelligible even for a complete novice - is certainly too small. Even more minimal is the amount of books that focuses into solid-state *guitar* amplifier electronics. This book is intended to fill this niche.

In this book I do not wish to present things with excessive technicality – there are better sources for that kind of information. Nor do I wish to make biased comparisons between solid state and tube technology. However, since the latter is an eternally hot topic (which always comes up when transistor amplifiers are mentioned) it is discussed shortly in section 1.5. My aim is to present various forms of basic circuitry used in guitar signal amplification - and the theory related to it – preferably in a form, which hopefully is understandable with common knowledge of circuit theory. One can then use this information as a base for learning more about concerned matter. If you don't know what a resistor or a capacitor is - or if you do not understand the basics of common circuit theory (such as Ohm's and Kirchhoff's laws) - I suggest you start with reading something other than this book.

Unlike some books, this one will not present a complete design example. I feel that this kind of approach would steal the base from the book since my idea essentially is to *teach* one to create such a design by him/herself as well as to make the designer ponder the aspects and choices related to the design process. Anyone can copy a design without learning anything in the process and even easier is to criticize a design and mark all theoretical content either good or bad according to it. I have spent a great deal of time in writing this book and gathering and assembling the information required to do so. If on top of that I'd had to create a design (*bear in mind that a design is not just a schematic*) the process would have become overwhelming and even more time consuming. Besides, I do not wish to give an impression that I favor some topologies over other – which is what people tend to think too easily if they are presented with something that should be “the ultimate design”.

Also, I feel obligated to include a compulsory warning here: Electricity can be dangerous. Electricity can cause injury or death! The circuits shown in this book are, to the best of the author's knowledge and belief, correct and safe. However, one must take great care when assembling electronic circuits – especially those that carry mains voltages. The author cannot accept any responsibility for any accidents that may occur. The use, application and transmission of electricity are subject to rules, regulations and guidance laid down in numerous regulations and standards. Most of these can be obtained for reference in your local library.

1.2 About the schematic notation

Some of the schematic symbol notation used in this book can be considered outdated and non-standard these days. However, I have stuck to notation procedures that are clear to decipher and natural to me. I will not apologize for it. Some may argue that a US style resistor symbol can be misinterpreted as a symbol of inductor and therefore using the European “block” symbol should be preferred. I agree that standards are good, but as my excuse, let it be mentioned that if I ever draw schematics by hand I simply find the US symbol faster and easier to draw. I have seen plenty of schematics where an inductor or transformer coil is actually drawn like the US symbol of resistor; this was actually very common in power electric notation of 1940's and 1950's. Yet

the alternative symbol notation methods have *never* caused me problems while reading schematics.

One may note that I have used dots to symbolize *all* interconnecting nodes and “jumpers” to symbolize schematic nodes that cross each other (perpendicularly). This kind of notation is also nowadays considered outdated. However, I feel sad that it is, since reading a *complex* schematic that is missing both of these notations is very difficult. The outdated way shows clearly, which nodes connect and which do not. Figuring this out from schematics that use alternative illustration methods is time-consuming and frustrating. The lack of these details, which make the schematic clearer - though harder to draw with modern CAD software – has cost me a lot more grief than trying to identify a “non-standard” symbol notation. However, I do not prefer the way of drawing dots that are nearly as large as the actual symbols or drawing one of the interconnecting nodes in a 45° angle in relation to other. Both of the aforementioned methods tend to make schematics unbelievably messy.

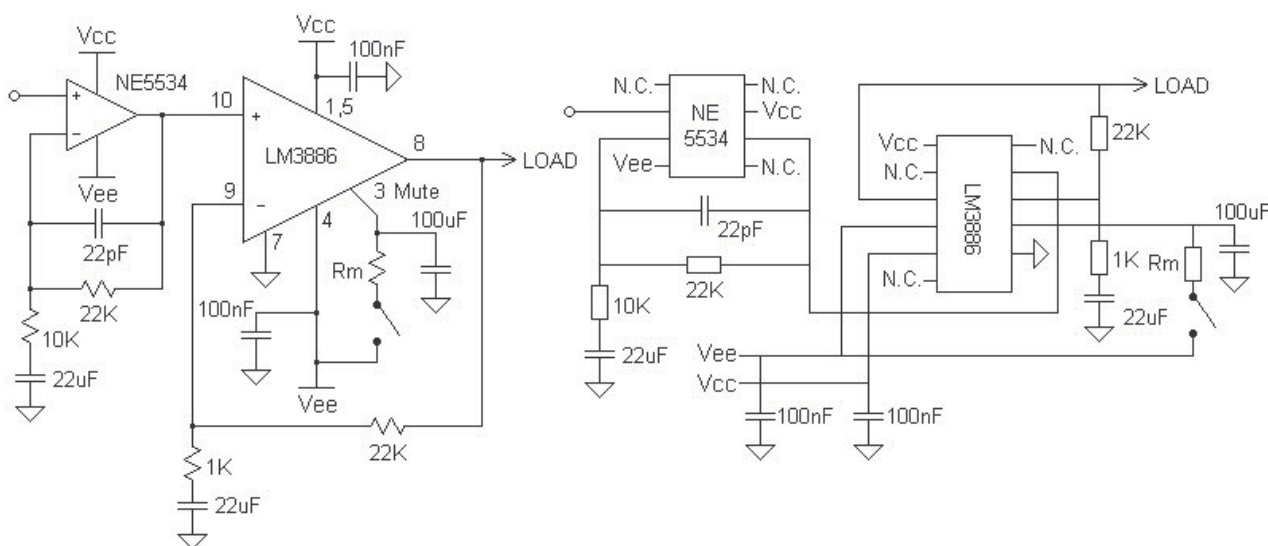


Figure 1.1 Which one of the schematics you consider more intelligible? How much does using “standard” symbol notation weigh in relation to using logical and clear layout?

The last horrid notation manner I have tried to avoid is drawing ICs as boxes with or without pin numbering. A box is the least informative symbol of a function unless it contains a description inside. *A model name is not a valid description!* If the circuit inside the box is an amplifier it can be drawn as a “triangle” instead since this is the standard symbol for amplifier in block diagrams. Boolean logic functions also have a dedicated set of symbols. I encourage everyone to learn this notation and use it. The pins of the symbol should have logical placement, i.e. positive and negative supplies located in adequate places as well as the inputs and outputs. If the IC has no other way to be drawn than a box then the important terminals should be marked with informative symbols or abbreviations such as IN, OUT, CLK IN, V+, V- and so on. These little details should save the observer from the trouble of searching the corresponding datasheet in order to even decipher the basics of the circuit depicted in the schematic.

1.3 Common terminology

Like every special field, the field of guitar amplifiers also has its own “jargon” that one must become familiar with in order to get an easier grasp of the basics. The subsequent part contains a small list of usual terms and a short explanation of them.

You will find a more comprehensive list of *electronics-related* abbreviations and terms from the end of this book. However, before delving into deeper theory I feel it is essential to get acquainted with the following:

Bias is a common term in electronics but became widely popularised by (tube) guitar amplifiers. The term bias means reference voltage, current or other input that is applied to a device or system to set the operating point. Generally, when referring to bias people are often referring to current applied to transistor's base or voltage applied to electron tube's grid. There is a clear distinction between idle (quiescent) current and bias current. Quiescent current is the current drawn by the whole circuit or a specific part of it (this is merely a question of selection) and it is (often) dictated by the bias current (or voltage) setting. However, the two terms do not mean the same thing! Quiescent current can be, for example, collector-emitter current flowing through the power transistor during idle state. It is dictated by the bias current setting but in this case the bias current itself flows through base-emitter junction and is usually significantly smaller.

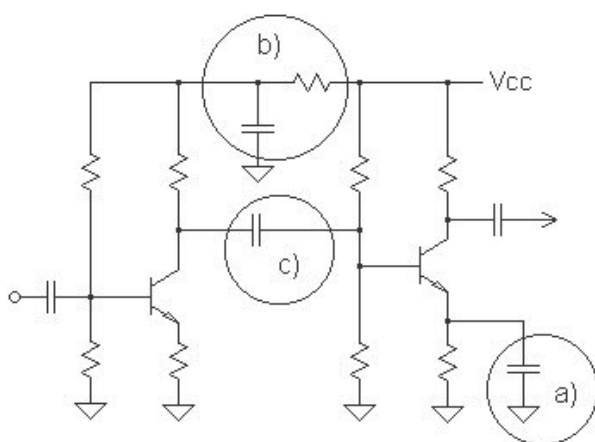


Figure 1.2 Sometimes the terminology for these circuit parts is ambiguous: a) Bypassing capacitor, b) Decoupling circuit consisting of a decoupling resistor and decoupling capacitor, and c) coupling capacitor performing AC coupling.

Bypassing, coupling and decoupling: Bypassing is a very literal expression that is used whenever a certain circuit or component bypasses the other. Such is the case, for example, in section a) of Figure 1.2 where a capacitor bypasses the emitter resistor. Another, less clear example is the capacitor in section b), which bypasses the whole first gain stage. This capacitor is a part of decoupling circuit. The term decoupling refers to preventing unintentional coupling between sub circuits, usually in order to prevent them from interacting with a way that might cause oscillation or otherwise impede performance. Commonly decoupling refers to process of isolating the circuit stages at the power supply level by filtering away the supply's AC component. Thus we now have three different names for the capacitor in section b): A filter capacitor, a bypassing capacitor and a decoupling capacitor. *Decoupling should not be confused with coupling*, which means intentionally connecting stages together, for example, by the means of “DC coupling” (connection on both DC and AC levels) or “AC coupling” (connection only at AC level). The capacitor in c) is a typical example of AC coupling. Replacing it with a short would DC couple the gain stages together.

Classes: When it comes to analog designs, amplifier circuits are classified as being either A, B or C class. A common fallacy is to interpret the operation class as a definition of the amplifier's quality – usually assuming that class-A amplifiers are the

best. In practice, this belief is quite erroneous because the class of operation does not relate to quality in any way.

A class-A amplifier conducts 360 degrees, in other words throughout the whole input cycle. Because class-A biasing ensures that the active devices never shut off during signal's "zero crossing point" a class-A amplifier exhibits no switching distortion. However, class-A amplifiers are extremely inefficient, which usually causes that distortion is created by other mechanisms. A typical class-A amplifier has a very low output power but an excessive power draw. This means that a huge amount of "wasted" power must be dissipated as heat.

A class-B amplifier conducts 180 degrees, in other words it only amplifies one of the signal's half waves. Typically two class-B amplifiers are run in a parallel configuration where each amplifier handles its own corresponding half wave. This arrangement is known as push-pull circuit. Class-B amplifiers are fairly efficient but inherently introduce a fair amount of switching distortion. A variation of class-B is the class-AB, which refers to extending the conduction time to anything between 181 to 359 degrees. This will *reduce* (but never remove) the amount of switching distortion at the expense of efficiency. Nevertheless, in a properly designed class-AB amplifier the amount of switching distortion is low enough to be perfectly *inaudible*. Most guitar amplifiers fall into this category.

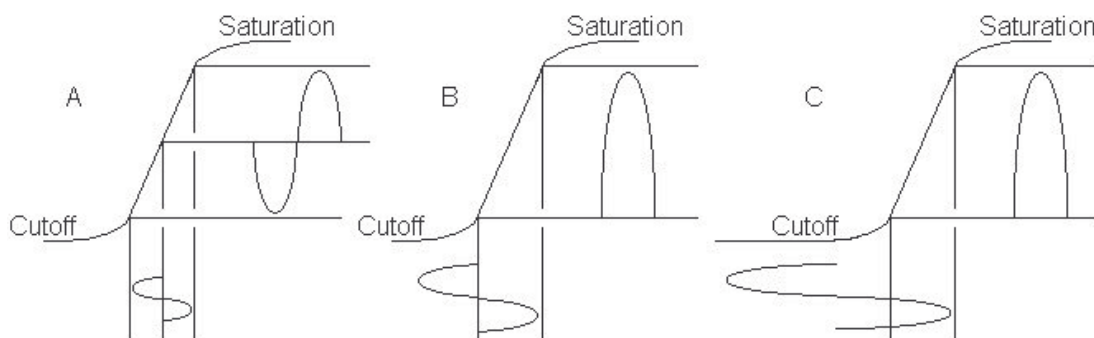


Figure 1.3 Operation classes for analog amplifiers.

Class-C amplifiers conduct less than 180 degrees. This means that they have a remarkable efficiency but too much switching distortion to work in audio applications alone. Various combinations of classes A, B and C are sometimes used as a base for some eccentric circuit topologies.

Classes D, E and F are definitions for switch mode amplifiers. Basically, the active devices in a switch mode amplifier are either on (saturated) or off (in cut off state). Yet, "D" does not stand for digital although the logic behind this misunderstanding is easy to understand. (D was simply the letter following C). Class-D amplifiers utilize specific modulation techniques like PWM, PDM or Delta-Sigma to control the switching. In other words, the analog input signal is first converted to a specific sequence of pulses that is somehow proportional to input signal amplitude. The pulses then control the output device(s). Switching happens at a frequency that is well above our hearing range and when the output signal from the amplifier is carefully filtered (right before the speaker load) it converts to an ordinary analog signal that is capable of driving conventional speaker systems. Filtering will greatly *reduce* artefacts of switching but again it cannot completely remove them. Since the active devices do not operate in their active region (pardon the pun) the efficiency of switching amplifiers is astonishingly high. However, at lower frequencies the efficiency is reduced due to

inductive filters that the circuitry requires. It must be noted that *switching amplifiers are not considered as “digital” unless they specifically utilize a digital modulation technique*, like pulse-code modulation. Classes E and F basically refer to alternative filtering embodiments.

Classes G and H refer to rail-switching topologies that are basically just special variants of ordinary amplifier classes like A and B. These circuits basically modulate the power supply voltage in a direct relation to the output signal, this in order to gain a greater efficiency. The definitions of class H and G are reverse in Europe and America.

At the time of writing, any other “classes” are just circuit topologies or trademarks of various companies and therefore do not refer to any real amplifier classes: Class-S is the name for Sandman’s feed-forward error correction scheme. Class-T is a trademark of TriPath Company and refers to their switching mode amplifiers. Class-Z is a trademark of Zetex and refers to a direct digital feedback technology. Similarly Crown has its class-I (or BCA), and Lab.grubben its class-TD or “Tracked Class D” topology.

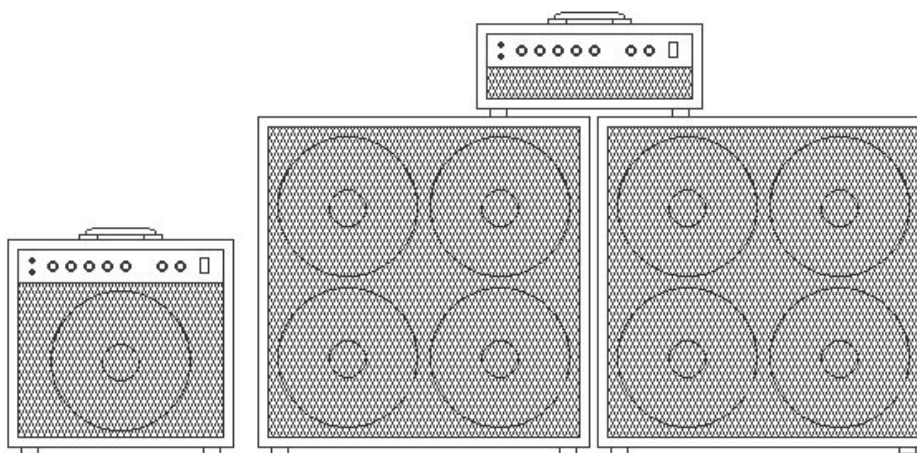


Figure 1.4 Left to right: Combo and head “piggybacked” on a full stack.

Combo, head, piggyback and stack – what are they? The term “combo” refers to a combination of an amplifier and a speaker system in the same casing or unit. A separate amplifier, on the other hand, is called a “head”, as it is usually placed on top of the speaker cabinets. “Piggybacking” head and a speaker cabinet this way, in turn, forms a “stack”. A “full stack” contains two or more speaker cabinets placed (“stacked”) on top of each other or side by side. A “half stack” contains only a single cabinet. In essence, the term “stack” is pretty vague and generally interpreted as referring to a combination of only two speaker cabinets. Thus a “double stack” has four cabinets in total (and in some cases two amplifiers as well - one serving as a “slave” driving the other half of the stacks).

So far things have been easy when considering semantic issues but that’s about to change: Is a “combo” with detachable amplifier unit (e.g. Ampeg Portaflex flip-top) really a combo or should it be interpreted as two individual units? How about when the speaker cabinet merely contains a carrying compartment for the amplifier unit? (This was common feature in several vintage amplifiers). How should a system consisting of an active (powered) speaker and a preamplifier head be called? ...Never mind various forms of rack unit configurations. It may be best that each one makes up their own mind on what each system should be called.

Driver: In electronics, the term “driver” commonly has two meanings: 1) an electronic component that controls another electronic component (e.g. voltage amplifier transistor “driving” a current amplifier stage) or 2) a loudspeaker (just the speaker, not the whole combination of a speaker and a cabinet). Due to matter of context it is quite impossible to get mixed up with these two different meanings, however, getting mixed up (and confused) with the various components belonging into the first category is very typical. In this book I have very much tried to refer to things with their own names (e.g. voltage amplifier instead of a driver), yet this is often simply impossible. Luckily, in many cases it is possible to make a division into various subcategories like drivers versus pre-drivers (pre-drivers being drivers controlling drivers). To avoid confusion these terms are discussed *within related context* in various chapters of this book.

FX Loop, or effect loop refers to various components involved in extending the amplifier’s internal signal path with external signal processing devices (pithily called “effects”). Since the external signal path is fed from a dedicated output of the amplifier and returned into a dedicated input an allegorical loop is formed (hence the name). The loop can be either in parallel or in series with the amplifier’s internal signal path. Usually the input and output of the loop have separate connectors but in some cases a single connector shares both duties. This special form is called an “insert”.

Hybrid, by definition, means combination of two or more different things aimed at achieving a particular objective. Generally, this term simply refers to a combination of alternative topologies within a single unit. Typical hybrid topologies are for example combinations of things like integrated and discrete circuits, digital and analog, tube and solid-state or BJT and FET. Understandably “hybrid” is therefore a very loose expression. At the time of writing, the term “hybrid” generally refers to circuits that combine tube and solid-state devices. Yet the general meaning of the term changes periodically since it tends to catch-up with the “hottest” technology. About four decades ago the term “hybrid amp” simply meant that the amplifier in question employed parts that were realized with integrated circuitry.

Impedance matching: This is another common (and commonly misused) electronics term. For *maximum power transfer*, the source and input impedances should be equal - in other words, they should *match* each other. This was very useful in the older days when power that was lost in a transfer process was very difficult to restore. However, in typical audio applications we do not need maximum power transfer. If low source impedance is connected to higher input impedance the power transfer is limited, but in turn the *voltage transfer* is higher and less prone to interference or other corruption. This is much more ideal configuration for audio signal transfer, say, from the guitar to the input stage, or from the input stage to the tone control circuit et cetera. This deliberate *mismatching* of impedances (usually by at least a ratio of 1:10) is called impedance (or voltage) *bridging*. With some twisted logic it may be considered as a “matching” since the basic idea is just choosing impedances that work well together. The opposite mismatching configuration (from high to low) would lead to signal corruption.

Modeling: Modeling or “emulating” circuits attempt to replicate the sound of electric circuits or electromechanical devices such as amplifiers and speaker systems. There is practically an immeasurable number of ways how a modeling circuit or device can be constructed but principally they can be divided into two main categories: Analog and

digital. There are also examples of electromechanical modelers - like the “Leslie” Doppler effect cabinet or a spring reverb. These, in turn, are nowadays modeled with conventional electronic circuitry. I can imagine that after few decades we similarly have modeling amplifiers that emulate the modeling amplifier designs of today. Modeling circuits are usually criticized for being “sterile” and “inaccurate” since they fail to capture all the nuances of the device they are imitating. This is often true but the critics also tend to miss a very important point: *When a designer resorts to modeling, instead of just replicating the original device, it is usually done to introduce some ideal properties that the original device didn't have.* Examples of such properties are increased versatility, cheaper price, lower power consumption, or simply being smaller, lighter or mechanically stronger than the original unit. Nearly every modeling device meets most of these design criteria extremely well.

Spring reverb: This is an electromechanical device that is used in creating a reverberation effect. A spring reverb includes a set of transducers and springs. One transducer converts current, that is proportional to input signal, to a vibration of the strings, which is then picked up by another transducer that in turn converts the vibration back to current.

Tone stack (and tone controls): Terms tone stack and tone controls are far too often used interchangeably. A tone stack is, as name implies, a circuit consisting of *stacked* filter circuits - active or passive - that control the tone of the amplifier. In such case the difference between the term tone controls and the term tone stack is merely one of usage. However, if the tone controls are spread all over the circuit, as opposed to stacking, it is very misleading to refer to the configuration using a term *stack*. Another term, or abbreviation, related to tone controls and stacks is “FMV”, which refers to *typical* stacked tone control circuit used in Fender, Marshall and Vox amplifiers - plus a countless number of their variants.

Tone vocabulary: While “tone vocabulary” is nice and is often quite a useful form of expression it is not precise or scientific language. People have different impressions about the meaning of subjective terms such as “lush”, “smooth”, “creamy”, “buttery”, “brittle”, “liquid” et cetera. There are plenty of terms that are even far too difficult to understand: What is “liquid tone” anyway? I have done my best to avoid using “tone vocabulary” in this book, however, sometimes it’s just plain impossible, so forgive me. A guitar amplifier designer must essentially deal with two worlds: The world of musicians, who deal with subjective and artistic issues (and who may in extremes consider bias and tone controls interchangeable) and the world of electric engineers, who deal with theory and science that leave no room for subjectivity. Understand that while it’s easy to say that a certain modification made the tone more “creamy”, it is much more professional and informative to state what the modification did circuit-wise; e.g. formed a low pass filter with a corner frequency of 5 kHz.

The following is a small list of some typical terminology:

Authoritative	“Pronounced” tone with enhanced but clear low-frequency content
Ballsy	Enhanced low-frequencies that you feel rather than hear
Barking	Literal expression: Distorted but dynamically strong tone that is suitable for rhythm guitar playing.
Boomy	“Smearing” of transients that makes bass reproduction sound muddled

Bright	Tone with enhanced high-frequency content
Brittle	Lacking in low end
Brown	Guitar sound of early Van Halen records. Sometimes associated with smooth breakup from clean to compression.
Buttery	Likely referring to combination of “smooth” and “organic”
Clean	Tone, which is <i>perceived</i> as unaltered, usually ranging from linear response (not clipped) to smooth compression. Highly subjective issue.
Clipped	Limited in dynamic range due to clipping off the high-amplitude signal content
Colored	A change in character, that over-emphasizes certain tones
Compressed	Limited in dynamic range due to clipping of high-amplitude signal content and enhancing the low-amplitude content
Creamy	See “Smooth”
Crunchy	The “Gritty” tone resulting from clipping distortion. Must be defined: e.g. “mild crunch”, “a lot of crunch” etc. Typical tone of 80’s metal and hard rock.
Dark	Attenuated high-frequency content
Distorted	Altered from original signal, usually by the means of distinctive clipping. Ranging from slight breakup to full “shredding” distortion.
Edgy	Aggressive character introduced by dialing in a “gritty” tone
Glassy	Highly subjective term. Sometimes paradoxically associated with either “harshness” or “smoothness”.
Gritty	The easily discernible increase in high-frequency harmonic content that results from clipping distortion
Harsh	Abrupt and nasty-sounding clipping tone, usually in conjunction with over-emphasized high-frequency content
Industrial	“Processed”, metallic or synthetic-like tone
Liquid	Likely refers to combination of “organic” and “smooth”
Lush	By definition, pleasing to senses
Mellow	See “Smooth”.
Organic	Rich and unpredictable in detail. Lively. Opposite to “sterile”.
Over-processed	Tone that is processed so much that it sounds unnatural and synthetic
Processed	Usually refers to an unnatural tone
Pronounced	A discernible tone with distinctive details
Shred or shredding	Literally associated with a metallic, buzzing sound of a shredder. Extremely distorted tone, which is typically rich in high frequency content. Sometimes with “authoritative” low-end. Closely related to music style (of modern metal).
Smearred	Distorted at small detail level. Lacking in definition
Smooth	Stripped from “harsh”-sounding high-frequency content
SRV	The abbreviation referring to guitar sound of Stevie Ray Vaughan
Sterile	Lack of character. Perhaps too “clean”.
Tight	Tone with fast attack, often “pronounced”
Warm	Enhanced low- and, in lesser extent, high frequencies. Sometimes associated with smooth compression that increases sustain
Woman tone	Eric Clapton’s thick, distorted guitar tone with reduced high-frequency content.

Then we have the more precise definitions for frequency bands:

Low (or sub) bass	20 – 80 Hz	1st – 2nd octaves	“Fullness”
Upper bass	80 – 320 Hz	3rd – 4th octaves	Root notes
Lower midrange	320 – 2.560 Hz	5th –7th octaves	“Intensity”
Upper midrange	2.560 – 5.120 Hz	8th octave	“Definition”
Treble	5.120 – 20 000 Hz	9th –10th octave	“Sparkle”

1.4 About integrated circuit amplifiers

Chip amplifiers are nice but I’m reluctant to deal with them. Even though they offer a far superior thermal coupling between internal devices, need no complex biasing circuitry and house features like short circuit protection and thermal cutoff, that would take a vast amount of board space to build externally, there’s something I dislike about chip amplifiers in general: An integrated circuit is a black box that allows no tweaking of the internal circuitry. One does not necessarily even have to understand how and why the devices work like they do. When using integrated circuitry one is mainly limited to following application notes supplied by the manufacturer and I think this steals the joy of learning and inventing something new - things that I consider an essential part of the electronics hobby.

Another aspect is the poor serviceability of IC-based amplifiers: For example, how can one repair a chip amplifier that uses a device that has become obsolete some time ago? The electronics market has seen far too many discontinued integrated circuits – and far too many devices that based an important part of their circuitry on such chips.

In my opinion, one should limit the use of integrated circuitry to common, all-around, building blocks, such as operational amplifiers, microprocessors, AD- and DA-converters, logic circuits, switches and timers. Such circuitry can simplify the design, increase its reliability and most importantly, it is usually available in standard packages with standardized pinouts. Even the most non-standard of the bunch – the micro controllers – have, in certain circuit families, existed unchanged as long as average personal computers – and most likely will continue to do so for some decades. However, integrated circuits with exotic special functions (dynamics processing etc.) are bound to get discontinued or obsolete and replacement parts will be difficult to find.

1.5 Tubes versus transistors

Nowadays tube amplifiers have a reputation of having the “ultimate tone” while their solid-state counterparts usually have a notorious reputation of sounding really bad. However, musicians *do* use transistor amplifiers: For example, during their career The Beatles used them on numerous occasions. Other famous musicians that have been using solid-state amplifiers include Frank Zappa, Carlos Santana, B.B. -and Albert King, Ike Turner, Billy Gibbons (ZZ Top), Brian May (Queen), Jerry Garcia (Grateful Dead), Andy Powell (Wishbone Ash), Paul Kantner (Jefferson Airplane), Francis Rossi (Status Quo), Chuck Berry, David Crosby, John Fogerty (Creedence Clearwater Revival), Robbie Krieger (The Doors), Jeff Beck, Neil Young, Jimmy Page (Led Zeppelin) and the late Darrell “Dimebag” Abbott (Pantera) and Kurt Cobain (Nirvana) - just to name a few. The list serves as a fine example that a good transistor sound is not a concept to be laughed at.

So what's the deal with mocking transistor amplifiers? More than several thousand of transistor guitar amplifier models have been manufactured so far and due to cheapness of semiconductor-based technology a lot of them have been built and designed very poorly. Besides that, one has to remember that even in very near past the semiconductors were considered somewhat hi-tech: The technology was new and no one seemed to know how to exploit its full potential. The manufacturing methods were poor as well. The quality of transistor amplifiers started to be acceptable only at the very end of the sixties and excellent at the end of the seventies. Some earlier designs might have had a good tone but the reliability of most pre-70's amplifiers was a joke. Transistor amplifiers suffered from "infant diseases", such as poor thermal stability, lack of short circuit protection and distortions created by mechanisms that were unknown to most 'til the very end of the 80's! In practice, only few people actually knew how to design a good transistor amplifier. Naturally, the solid-state craze of the sixties led to huge disappointment and people returned to using tube amplifiers. Transistor amplifiers had earned their bad reputation – the one they are still struggling to shake away.

Now, transistor amplifiers have "matured". At the same time solid-state technology and manufacturing of electronics has become extremely cheap. It's a shame that today, when there is a great chance for transistor amplifiers to shine with their quality, most of them are still designed and built poorly: A major portion of solid-state *guitar* amplifiers are still "bargain bin" and "entry level" models of questionable quality.

Vacuum tube technology, on the other hand, is not that much older but has been in use for a considerably longer period. Back in the 1930's, when the principles of semiconductors were invented, vacuum tube-based audio circuitry was already being used in movie theaters and large concert halls. Very soon the same technology was used for public addressing and in conjunction with the new electric guitars. At the end of the 1940's guitar amplifiers were already part of the popular equipment of wealthier bands. Certainly there has been a huge amount of lousy and mediocre tube amplifiers; we have just forgotten them a long time ago. The good circuits survived and unsurprisingly they haven't changed much since the 50's or 60's.

There are some (actually quite a lot) solid-state amplifiers that enjoy an equally good reputation as some of the best tube amplifiers. In no particular order: Fender Steel King and M-80, Pignose 7-100, Gibson GSS-100 and Les Paul solid-state "stack" (LP-1 & LP-2), Pearce Amplifier Systems G2R, Vox solid-state AC-30, some of the Thomas Organ's Vox models plus most UK transistor Vox amplifiers, Road Electronics amplifiers, most GMT and Gallien-Krueger amplifiers and some Peavey models, some Kustom models, the whole Norlin's Lab Series lineup, Randall's RG100-ES, RG80, RG75 and Warhead line-ups, Polytone Mini-Brute II and Mega-Brute, most Acoustic Control Corporation's amplifier models, Rickenbacker's Transonic series, most high-power solid-state Stadel and Hagström models, Pritchard amplifiers, Sunn Coliseum and Beta series, Crate Power Block and Roland's Jazz Chorus lineup.

Now, the readers of this chapter should not misunderstand me: I do not dislike tube amplifiers. What I find unappealing in them is their price. The special transformers that tube amplifiers require cost a lot and are difficult to acquire - except by ordering via Internet. The price issue is probably the single biggest reason why it seems like only good tube amplifiers are being made: Producing poor but expensive gear is bad

business. The other fact that I dislike is that most tube designs are *inefficient*. To add to the insult, most of them are *underpowered* as well. While people may go lyrical about “rounder” clipping characteristics of tubes - resulting into idiotic equations such as “one tube watt is equal to ten solid state watts” - I still do think that a 5 to 10 watt amplifier is underpowered and if it’s a tube amplifier then it is *expensive* as well. Similarly a 50 to 100 watt tube amplifier is inefficient, heavy *and very expensive*.

The “tube watt is louder than transistor watt” is a myth that needs to be straightened out: A watt is a watt, whether it is produced with a transistor, tube or steam engine. That in mind, the differences in loudness can usually be caused only by:

- Differences in the compared signals. There are several causes for this, such as clipping or unequal system gains and frequency responses.
- Differences in speaker system efficiencies throughout their effective bandwidth.

Last time I heard someone touting out this myth it was revealed that he was comparing a 5-watt tube amplifier with an efficient 10-inch driver to a 10-watt solid-state amplifier with a puny 6.5” driver. Power-loudness comparisons mean absolutely *nothing* unless they are performed using identical speaker systems! Similarly, power-loudness comparisons mean *nothing* unless they are performed with systems that have identical frequency responses! Not that many guitar amplifiers actually have. More than that, if comparisons are made it should be ensured that both systems actually deliver the same amount of power to the speaker. It should be realized that the dial of a volume control can’t be trusted as a measure of this since it really indicates nothing. In practice, it is almost impossible to compare two dissimilar amplifiers and get meaningful and scientifically valid results.

The fact that most marketing campaigns of solid-states amplifiers were (and are) also more than anxious to put a lot of “air” into the amplifier’s output power rating (sometimes even stating peak power instead of the average one) is another perfectly realistic explanation for the existence of the tube watt myth. In practice, the only *valid* reason explaining the differences in loudness is the fact that one of the signals is actually *distorted* in comparison to other. For example, if one of the compared amplifiers is clipping and the other one is operating at its linear region the one that is clipping will sound louder – this regardless of whether the clipping amplifier uses tube or solid-state technology. In this light, the comparison practically loses its meaning since it is a very subjective issue whether any distortion should be accepted in the signal or not. Most tube amplifier “fans” like it, but in the old days (pre 1960’s) the designers actually tried to get rid of it. Not all people like tube compression and to be honest, that compression is often unmistakably audible as well. Since distortion created by a clipping tube amplifier sometimes (but not always) has a “soft limiting” effect it is possible to harness an external soft limiter circuit to make a solid-state amplifier “clip” the signal as gracefully as tubes do – this kind of circuit can also have an on/off switch.

It must be noted though that soft clipping always narrows down the linear operating region and thus decreases the amount of “linear” output power that can be extracted from a given supply. Thus an amplifier that provides 100 watts of output power @ 1% THD when swinging rail-to-rail cannot employ an additional soft clipping scheme and emulate an inherently soft clipping 100 watts @ 1% THD amplifier. For that it needs more power reserve. In fact, when engaged the soft limiter might drop the amount of

output power at equal THD closer to 40 or 50 watts, depending on the softness of the utilized limiter setup. This explains why inherently soft clipping amplifiers - be they solid-state or tube - are louder when compared to ones that clip abruptly.

Reliability of tube amplifiers is another “myth”. In a properly designed tube amplifier the preamplifier tubes should last decades and the power amplifier tubes several years (depending on how actively one uses the amplifier of course). I do not regard the Vox AC30 or certain Marshall amplifiers that notoriously eat their power tubes in months as “properly designed” – regardless of how good they might sound. I do not consider tube amplifiers that quickly wear out inferiorly manufactured modern tubes as properly designed either - but that’s just me. A fact is that if you stress the amplifier it will fail a lot sooner (regardless of whether it uses tube or solid-state technology): Jimi Hendrix and Van Halen played their tube Marshalls “on ten”, they also had to hire technicians who fixed their amplifiers at least once a week! This procedure usually also included changing the *blown output transformers!* Not many people can afford to do something like this on a regular basis just to get that tone.

Anyway, most problems with amplifiers are actually mechanical faults. This includes broken solder joints, intermittency of jacks and potentiometers et cetera. Most of these problems increase when components susceptible to mechanical stress, (like jacks, switches, potentiometers and tube sockets), are mounted on PCB and often supported only by the solder joints. Yet these are all problems related to initial design and assembling quality – not the technology of the active devices. Two major reasons why solid-state Fender amplifiers from CBS era were so unreliable were that screws attaching power transistors to heat sink often were not fastened properly and soldering machines rarely got cleaned. Poor design choices create troublesome solid-state and tube amplifiers equally. For example, a certain Marshall derived its 5-volt supply for the digital circuitry from the filament winding. The supply was of course regulated but nevertheless, a high-voltage surge was enough to destroy the delicate digital circuit boards. A case where B+ shorted to filament winding (like when tube socket arced) was even more catastrophic. They could have prevented all of this by simply powering the digital circuitry from a separate power supply.

Another very common design mistake in modern tube amplifiers is using DC filament voltage to decrease hum. There is nothing bad in this less-noisy practice – unless it’s done in conjunction with an underrated bridge rectifier soldered to the PCB. Most bridge rectifiers should be equipped with a heat sink in order to meet the power dissipation ratio listed in the datasheet. Most people wouldn’t even dream of running output transistors without a heat sink but would do this for a bridge rectifier without giving it a second thought. The result: The rectifier gets hotter than expected, melts its solder joints, chars the printed circuit board and lifts the PCB traces. Usually capacitors next to it are heated severely as well, thus decreasing their life span. These are just a few examples of problems that exist in tube amplifiers but are more related to matter of poor design than the fact that tubes would be generally more unreliable devices.

The most important fact in the seemingly endless debate concerning solid-state versus tube sound is that tube amplifiers sound good only if they are designed properly. With the same principle, transistor amplifiers can also sound good. The principal mistake made by most people is to think that transistor and tube amplifiers are alike. They are not. Apart from fact the fact that both devices can amplify voltage and current a lot of differences exist and therefore most circuit topologies are not interchangeable.

Besides, there is no need for solid-state amplifiers to mimic tube amplifiers in order to sound good. Transistor circuits *can* mimic tube circuits, but if you expect it to happen simply you are mistaken.

Soft clipping characteristics of a vacuum tube (transfer curve) are easy to replicate with both analog and digital circuitry and the most perceivable character of tube power amplifiers, which is the varying frequency response into speaker load (refer to Figure 1.5), is easy to replicate as well since it is caused by no other reason than high output impedance. If the output impedance's effect to frequency response is compensated distinguishing any audible differences between tube and solid-state amplifiers becomes very hard and is basically limited to hearing the tube amplifier's output signal being intermodulated by mains frequency signal due to inferior regulation [1].

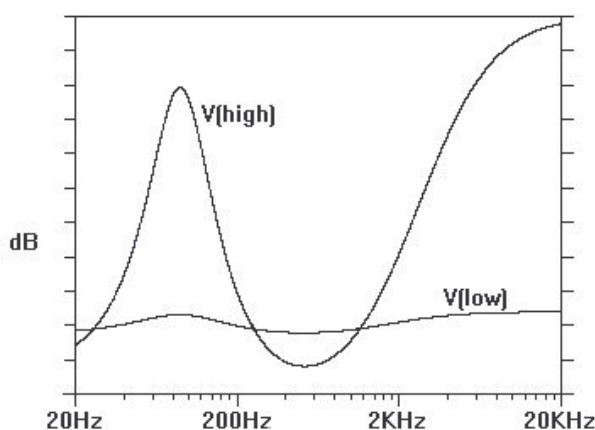


Figure 1.5 Reaction to speaker impedance in an amplifier having high (5 ohms) and low (0.2 ohms) output impedance.

Note: The Y-axis of the graph has no particular scale since it reflects the speaker's impedance at resonance. Even more, the resonant frequency is dependant on speaker and cabinet configuration and therefore the graph should be considered only as an example. The graph shows the voltage response to speaker impedance when both amplifiers are set to output equal power at 1 kHz sine wave signal. The driver used was Eminence Legend 875.

A topology called “mixed mode”- or “current feedback”, that raises the amplifier's output impedance, thus lowering its damping factor, has been known at least since the 1970's and has been used in various amplifiers. For instance, this topology is an essential part of the circuitry in Korg's Valve Reactor and Marshall's Valvestate amplifiers. If “vintage”, transformer-coupled transistor circuitry is utilized I wouldn't even use a term “mimicking tube circuits” since this topology is essentially identical. The result of using either topology is that the solid-state amplifier will start to behave like an ordinary tube amplifier and produces a higher voltage gain at higher load impedances. For reference, in a typical solid-state *current amplifier* output stage, the voltage gain remains nearly constant despite changes in load impedance.

The linearity and low high-order distortion figures of a triode tube are unquestionable but when combined with distortion produced by the output transformer the harmonic distortion pattern becomes pretty similar with transistor designs [26]. Band-pass filtering formed by saturating the output transformer is probably the hardest effect to mimic but usually solid-state distortion circuits already introduce a drastic bandwidth limit and “voicing” before and after the clipping stage – at least the ones that sound

good. Transformer saturation is a result of using underrated components (for example, 25 W output transformer instead of 50 W) so some tube amplifiers don't even exhibit this band-pass effect. For example, some older Hiwatt amplifiers used so generously rated Partridge "iron" that it made even HiFi amplifiers to pale in comparison. Gibson's older tube amplifiers were usually exactly the opposite and constantly utilized undersized output transformers.

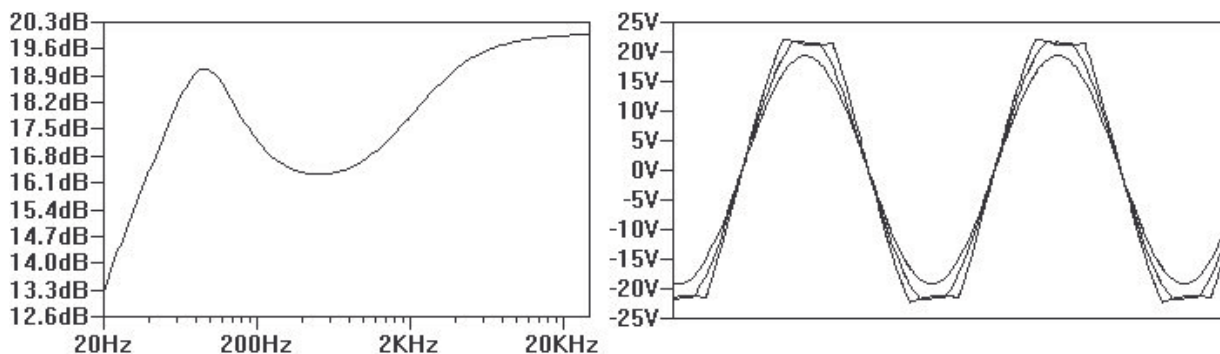


Figure 1.6 Frequency response and clipping behavior of a Music Man hybrid power amplifier.

Since many are likely interested in knowing more about certain "qualities" of the praised tube power amplifiers it might be worthwhile to take a look at the performance of the Music Man hybrid discussed in 3.19.3: Figure shows the frequency response of the amplifier when it's driving a 4-ohm Eminence Legend 875 speaker. (The same load is used in the following plots as well). It is clearly seen that the high output impedance together with the speaker's non-linear impedance causes a 3 dB emphasis at the resonant frequency. High frequencies are also emphasized. Like so many other designs, the amplifier in question employs some negative feedback to decrease the output impedance and tame this non-linearity. In fact, without feedback there would be a 16 dB peak at the resonant frequency and a 31 dB peak at 12 kHz. Due to enormous power demand these peaks alone would create serious amounts of distortion and likely even introduce instability. This is a common concern in all push-pull tube power amplifiers, not just this specific one. Thus we see that using negative feedback is often quite essential – even mandatory.

How about the hyped soft clipping? Well, we see that it's pretty non-existent: The amplifier clips as abruptly as an ordinary solid-state amplifier. Don't be fooled to think that the sudden and harsh clipping is due to use of solid-state technology; it's not. Even an all-tube amplifier may clip in a fairly similar manner – and sometimes even harsher depending on the circuit. For example, the all-tube output stage of the Univox U-1000 amplifier clips pretty much like an ordinary solid-state operational amplifier. It also suffers from poor clipping recovery that shows up as "rail sticking"! It should be noted that the concerned amplifier uses only a very small amount of negative feedback so we should ignore that typical "excuse" for poor performance.

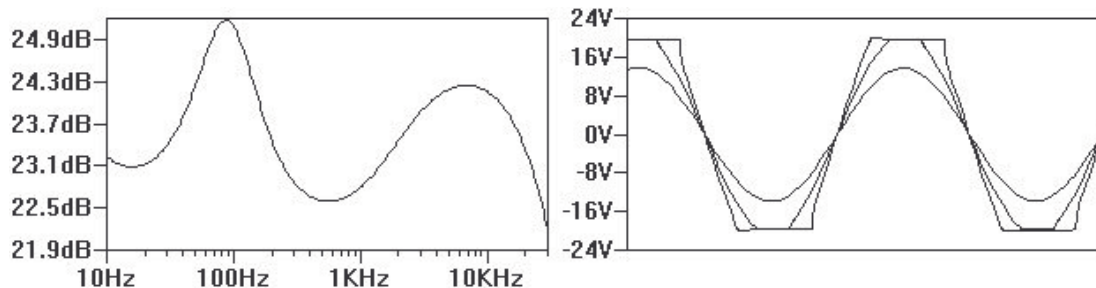


Figure 1.7 Frequency response and clipping behavior of Univox U-1000 all-tube amplifier.

Soft clipping mostly seems to be a device-related characteristic, which is shown by the comparison of Fender Bassman 5F6A and Marshall JTM45. The output stage of both amplifiers is identical with the exception of different power tubes and feedback takeoff point. Fender uses 6L6GC (which is the same tube used by aforementioned Music Man and Univox amplifiers) while Marshall uses EL34. The difference in clipping behavior is striking! Again this difference cannot be explained by the use of higher amount of negative feedback since Marshall employs more NFB than Fender: If you look at the decibel scale, the plots show how the effect of feedback flattens the frequency response of JTM45, thus “tightening” its tone.

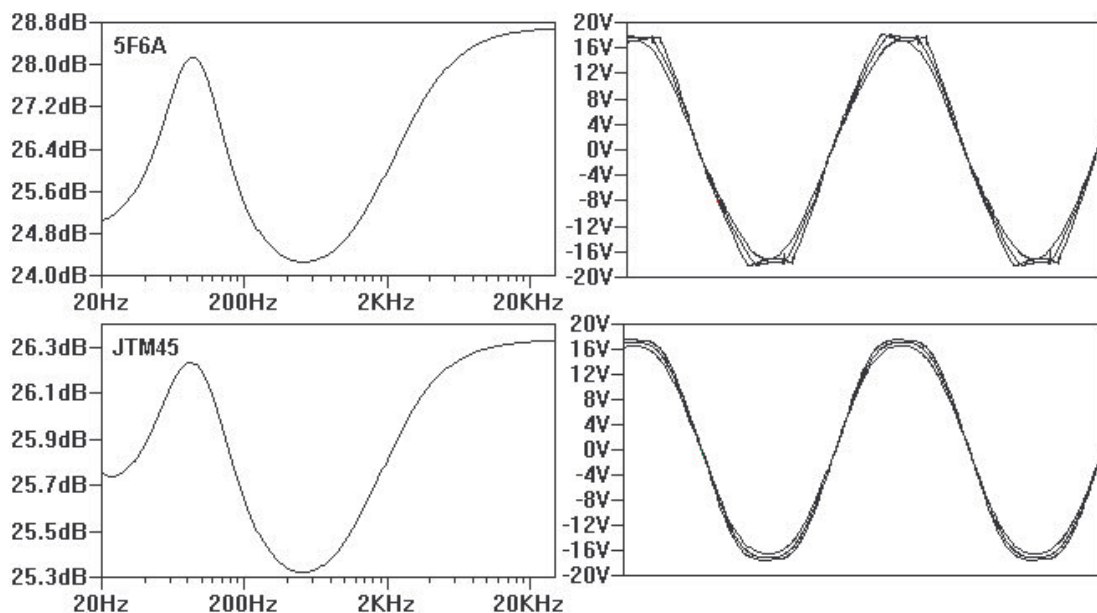


Figure 1.8 Frequency responses and clipping behavior of Fender Bassman 5F6A and its heir Marshall JTM45.

The flaunted harmonic content from a clipping push-pull tube power amplifier also looks very “unusual”: As the distortion increases, while more and more signal is getting clipped, we see that the even harmonics in comparison to odd ones barely increase in amplitude. In fact, the prominent harmonics are all odd ones. This is not against “textbook knowledge”, since any respectable book teaches that *push-pull stages cancel even harmonics*. However, this is against false beliefs and information spread by a vast number of people who have little idea of what they are talking about. The source of the common beliefs concerning tube amplifier harmonics is likely a misinterpretation of Russell O. Hamm’s famous study [2], which unfortunately failed to compare the differences of solid-state and tube amplifiers in a proper depth.

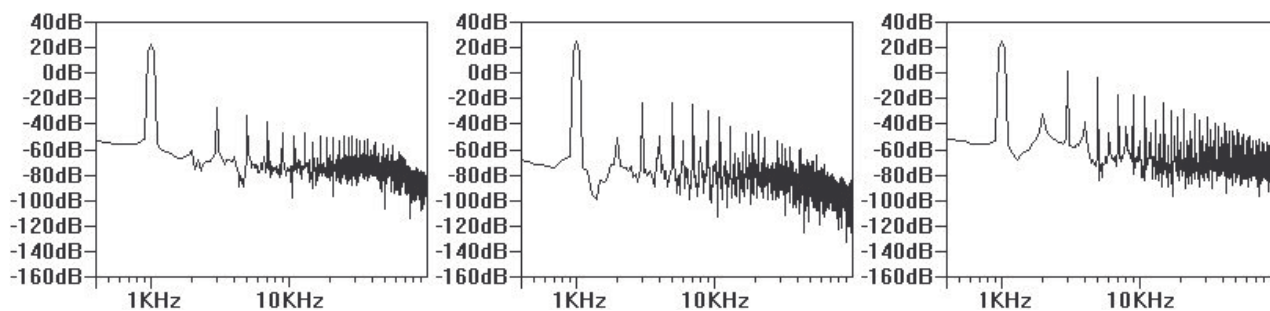


Figure 1.9 Harmonic spectrum of a clipping push-pull tube amplifier.

Note: Russell O. Hamm’s study was based on measurements of only four different microphone preamplifiers (“...four different commercially available preamplifiers,

using two or more stages of amplification. All the circuits use feedback, a couple are push-pull.”). Testing only four devices can hardly give universal results and even more: The test setup was highly biased since all of the measured tube circuits were single-ended while all of the measured transistor circuits were push-pull. Figures 10 and 11 (in the study), that depict square wave clipping of a transistor (and high-order harmonics that it creates), clearly reveal “ringing” oscillation at the onset of clipping. This means that Hamm compared an unstable circuit to a set of stable ones. The paper does not depict the tested circuits either, which leaves reader to speculate if there was something else behind the performance of the circuits than just the type of active devices they used. In essence, Hamm’s study presented rough generalizations that hold true only in some specific cases.

We have now debunked at least two “tube myths”: 1. That tubes *always* clip softly and 2. The claim that tube clipping *inherently* creates a dominant amount of even harmonics. We may add the detrimental effects of negative feedback into the great realm of myths as well. Essentially, the only universal characteristic of a typical tube amplifier is the naturally high output impedance and the effects created by it. This characteristic can be successfully replicated in solid-state amplifiers, even rather easily. These observations also correspond with the results of the study made by Bussey and Haigler in 1981 [1].

Yet, most “tube heads” will still question the fact that a solid-state amplifier could sound like a tube amplifier. They are partially right: A tube amplifier is a complex system and a 100 % accurate solid-state emulation of one would require an *even more complex system*. No one is willing to build one, nor is it even sane. DSP can get very close with a low component count but even that technology fails to capture all the details since there are just too many of them. Usually the memory and computing time of DSP systems is also dedicated to creating a *variety of decent models* rather than a single extremely detailed and accurate model. To model a single tube is quite easy, to model a set of them hooked up to an output transformer loaded by a speaker is a much more complicated task. The same thing applies “backwards” as well: Despite their claimed sonic merits a tube amplifier could never sound like an exact copy of a certain transistor amplifier. I believe that a tube design emulating, say, the Lab Series L5 or Roland JC-120, would have to be at least hundred times more complex than a solid-state design imitating, for example, a Marshall JCM800. Besides that, one has a problem of what to imitate. Due to large variations in tube parameters it is quite impossible to even find two JCM800s that would sound exactly the same.

What one needs to realize is that - as with most electronic circuits - it’s not about the device – it’s about the complete circuit built around it and hopefully this book will provide some further insight to that topic. The best practice to deal with the whole tube versus solid-state technology issue is to be very skeptical: Every time you read or hear a claim that a vacuum tube possesses a certain property you should stop and think whether that is specifically a property of the vacuum tube itself – or perhaps something caused by the different design approaches that vacuum tubes need. The effect of output impedance is a very good example of this.

There is also a great deal of misinformation and delusion concerning tonal effects of different tube types. It strikes me to hear people comment on how their Ampeg amplifiers have more midrange than the typical Fenders or Marshalls and then explain this characteristic by the fact that Ampegs used rather unconventional tubes. Wrong! This is again a thing that has nothing to do with the type or model of active devices.

Most Ampegs just used a “Baxandall”-type tone control circuit, which has fairly flat initial frequency response. Also, the midrange control of most Ampegs was an individual resonant circuit, which created a flat response at “5” and a significant mid-range *boost* when the control was turned fully clockwise. In comparison, a typical “Fender/Marshall/Vox”-type tone control initially introduces a *significant mid-range notch* – the notch is there even when you turn the mid-range control all the way up to “10”. It is no wonder that an Ampeg sounds “nasal” when conventional tone control settings are used.

A detail like this should be obvious to everyone who is even moderately skilled in the art of electronics so need I say more? Also, it never ceases to amaze me how many people think that the tone controls and frequency responses of all amplifiers are always identical. What is kind of sad is that instead of focusing on effects of things like speakers or tone controls, which can really make a huge difference tonal-wise, a typical “tube head” always assumes the tube choice is the cause for everything. I can see how easy it is to jump into this bandwagon and acclaim a “pro” status by speculating whether an Electro-Harmonix 6CA7 sounds as “mid-rangey” as an EL34 but unfortunately this really serves nobody and only adds more confusion. Bottom line is: If a tube would create such a major impact on frequency response and tone it would be safe to say it was defect. We don’t hear “transistor guys” saying how they switched the output MJE3055s to TIP3055s and got a mellower mid-range, do we? (Granted that there is a lot of similar and unnecessary discussion about tonal differences of OpAmps). Anyway, if a certain tube amplifier sounds different than the others you can be 99.9 % sure that you will find the explanation to this behavior from the circuit topology, speaker choice or cabinet design – not from the choice of tubes. Same thing applies to those OpAmps as well.

Anyway, I feel that this discussion has already exceeded a far too great length. The whole idea of creating some kind of a “competition” between solid-state and tube technologies is rather stupid and boring anyways. Both have merits and both should be respected.

Note: The author prefers not to use the term “valve” since it essentially depicts a device’s property to control a flow of certain matter through it. Since the term can be applied to both semiconductor technology and fluid mechanics (with equal sovereignty) it should be used only as a mere analogy – not as a synonym of a vacuum tube.

1.6 Amplifier’s tone – or is it?

Throughout the years, the use of a tube amplifier has been the first explanation that people come up with when they give praise to a certain tone. However, there are plenty of examples when the greatest contributing factor has been - in the words of Monty Python - something completely different: The legendary distorted guitar sounds on Link Wray’s “Rumble” were created by punching holes to speaker cones (note that many different versions of the song were recorded). In an interview Wray once told that this was the only idea he came up with to replicate the overdriven tones he had when playing live. Placing the amplifier into a staircase created the echo effect. Broken speaker was heard also on Jackie Brenston’s influential single “Rocket 88”. It is actually quite funny that most tones on the records that pioneered the use of distorted guitar sound were created with entirely different means than overdriving tube amplifiers; that just couldn’t be done in the studios since the high sound pressures

would have created a lot of unpleasant standing waves. Likely the first *record* containing authentic and “heavy” tube overdrive tone is Johnny Burnette Trio’s “The Train-Kept-A-Rollin’”. However, the distinctive tone came from the fact that the pickup of the guitar Paul Burlison used during the recording had its low-E string pole piece significantly raised. Onstage, the guitarist created the tone by pulling off one of the power tubes in his Fender Deluxe 5C3. (In push-pull circuits this creates a highly distorted, asymmetric output).

It is amazing how much amplifier’s tone can be shaped with simple matters of microphone- and speaker cabinet placement as well. For instance, the amplifier can be lifted off the floor for reduced bass and vice versa. To get a very large sound the amplifier can be placed facing the corner and miked from the behind. This is a very similar technique to that what blues legend Robert Johnson persisted on using while recording. In aforementioned configurations huge differences have been made even without considering aspects like microphone choice and placement. Most studio recording techniques work best if the amplifier doesn’t produce huge sound pressures, thus small, low-wattage combo amplifiers were (and are) often preferred: Just to name few examples, Frank Zappa used a Pignose 7-100, Brian May used the “Deacy” (which was simply a junked solid-state car amplifier fitted inside a HiFi speaker cabinet) and Iron Maiden utilized a Gallien-Krueger 250ML.

Today, when modeling technology has become an everyday part of signal processing, the responsibility of tone shaping is easily thrust on the shoulders of the guitar amplifier. However, the assumption that an amplifier would, or could, generate a desired tone is basically incorrect. A large portion of a certain tone is always a result of both playing technique and type of guitar. The playing technique can naturally be practiced. One should first learn to get the desired tone *acoustically* and only after that consider the contribution of the amplifier, effects and other equipment. The best bet is to start from the instrument: Different pickups have different tones – even when they belong into same category. The difference between single-coil and humbucker pickups is (usually) astounding – the difference between various pickup positions (like neck or bridge) is astounding as well. The type of bridge can have a huge impact on a guitar’s sustain. A Telecaster has hard time in sounding like a Les Paul and vice versa. The tone of both aforementioned guitars is far different from hollow body semi acoustic guitars like Gretsch’s. Furthermore, string gauge and pick thickness make a difference as well, heavier strings usually have a “fatter” tone. Thicker picks create a heavier, “compressed” tone while thinner picks provide a bright tone with a “snappy” attack. Of course, a lot depends on how you pick – and where you pick.

One final difference that must be mentioned is the choice of speakers and speaker cabinets: For instance, do you think that a Marshall head would sound like a stereotypical Marshall if it was played through an *open back* 1 x 12” cabinet? Would you get stereotypical Fender tones through a closed back 1960A cabinet?

1.7 Operational amplifiers versus discrete circuits

There is much controversy about the tone of operational amplifiers. My view about the debate is simple: As long as the design concept of maintaining the proper headroom (which unfortunately is often quite low) is followed, the operational amplifiers (used in low current stages) beat all discrete designs clearly – at least when it comes to acquiring clean tones. One can make a counterstatement noting that most operational amplifiers use a class-AB output stage and therefore suffer from crossover

distortion, however I think such statements are not very convincing. The class-AB stage of an operational amplifier uses such low current that it can be easily biased so close to class-A operation that the amount of crossover distortion is nearly non-existent. If you don't believe that, believe the distortion figures: A typical *discrete* common emitter stage has a THD figure of about 0.1 %, feedback paired stage 0.065 % and an emitter follower buffer about 0.059%. Even a *mediocre* operational amplifier has a THD figure of about 0.001 % and a good one about 0.0001 % - how much do you think the crossover distortion weighs in that?

The reasons for better distortion figures are replacing resistors with a linear constant current loading and - more importantly - using a large amount of linearizing negative feedback. Basically, the distortion is so minimal that it is hard to detect. In practice, it is very hard to actually detect *anything* in a manner that allows pinpointing a certain topology as the cause. R. M. Mottola, the technical and contributing editor of *American Lutherie*, the journal of the Guild of American Lutheries, once conducted a blind test that compared two bass preamplifiers; one was made out of an operational buffer amplifier circuitry and the other from a common source JFET. In those tests less than half of the test subjects were able to tell any difference between the two circuits - not to mention which one was preferable. Out of fourteen test subjects only two found the tone of the JFET amplifier preferable - another two preferred the operational amplifier circuitry. Five test subjects found the sound of both circuits identical. All test subjects were professional musicians. An article about the test was published in Journal of Musical Instrument Technology, issue 23, 2003.

Due to the high amount of negative feedback, operational amplifiers clip very harshly. Basically any active device - including a vacuum tube - would do this. The problem is that OpAmp's operation is based on using massive amounts of negative feedback. In practice one just has to deal with that by staying within supply voltage limits. A common beginner's mistake is to use OpAmp gain stages with huge gain figures like 20 to 50 when the reasonable amount would be a decade less. Almost every solid-state amplifier of these days has a multitude of OpAmps in the signal chain of preamplifier - even the famous Lab Series lineup used them. The designers of these amplifiers have understood that the best practice is to keep the concerned gain stages clean and use other methods for acquiring distortion and overdriven tones.

1.8 Use of DSP in amplifiers

When it comes to solid-state designs the author is a strong supporter of DSP. If a good resolution (bit rate) and high quality AD-DA converters are used it is impossible to distinguish a digitally processed signal from one processed with analog circuitry - this is of course assuming the algorithms used in the modeling process are good as well. One can describe digitally processed signals as "sterile". However, I'm more inclined to say that they have less distortion and a better noise figure. A digital processor circuit is easier to tweak than an analog circuit that does the same thing - very often all you need to do is to rewrite the code. Likely the circuit will also be much cheaper, have smaller physical size and be much more reliable. Today, the DSP technology is still a bit expensive but its price practically drops by a half every year - which is something one can't say, for example, about vacuum tube technology. The downside of DSP is the complicated servicing. Due to circuit nature one can't speak about old-fashioned repairs anymore. "Repair" of digital stuff practically means fitting in new replacement circuit boards - which naturally have to be ordered first. However, in a *competent* design the digital part of the circuitry is the least likely to blow or indicate

any problems. Mechanical faults plus aged or overheated components in high current circuitry are much more common reasons for problems.

With a digital audio processor and few converters it is, for example, possible to modify frequency response, transfer curve and signal delay. Basically this variety alone allows creating effects that realized in an analog design would require hundreds of components. Better yet, one can create hundreds of different effects just by adding some lines of code when in an analog circuitry one would have to construct hundreds of circuits. It has to be noted that creating good analog circuits is not that easy either. I suggest anyone to take a look at the schematic of, for example, Boss MT-2 Metal Zone to see what heights proper analog distortion circuit design often needs to take. One can't just throw in a few clipping diodes and think its okay like the designers tended to do in the 1960's – unless one's seeking that 1960's fuzz sound, of course. For reference, the MT-2 circuit uses 9 OpAmp stages and 11 transistors including FET switching circuits. I didn't even bother to count all resistors and capacitors – and remember: This circuit provides only one type of distortion tone.

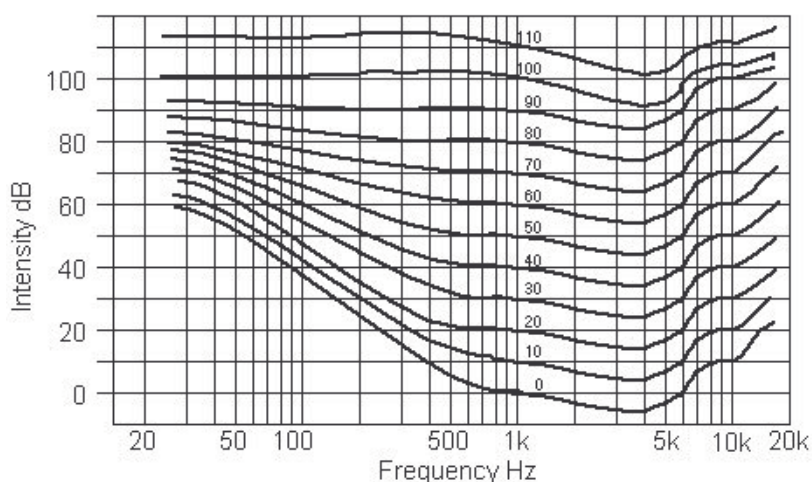


Figure 1.10 Equal loudness contours or Fletcher-Munson curve.

What most people also criticize is that an amplifier with a digital processor may sound good on bedroom levels but totally different when cranked. This is a stupid argument. Anyone familiar with the theory of human hearing and Fletcher-Munson (or Robinson-Dadson) curves knows well that human hearing is less sensitive to low and high frequency sounds on lower volume levels. A guitar effect tweaked to sound good in your bedroom likely has highly emphasized highs and lows and this becomes audible when the volume gets cranked. The emphasis may be in the order of several decibels, which will overdrive any amplifier – especially since most of it is at the low frequencies. If one has tweaked the effect (patch) to sound good while playing guitar at home and wearing headphones - and especially if the effect also employs a cabinet simulation or other frequency altering effects it is nearly impossible to make it sound good in a situation where a band is playing live and the guitarist uses a real guitar speaker cabinet. A professional musician already knows this and is familiar with techniques required to tweak the tone controls of the amplifier and the effect to suit up the performing environment and volume level. Another issue is how easy this tweaking is: Versatility is a good thing but can turn against itself if amplifiers and effects are too hard (and slow) to set up.

In case one is afraid to add complexity to the amplifier in the form of digital circuitry, the best practice is to design the amplifier to have a clean, “sterile” sound and a high-

power output stage that makes sure the amplifier stays clean at higher volumes as well. All digital circuitry can always be added “outboard” and basically the two greatest factors in determining the final tone of the amplifier - factors that no design based on maintaining the linearity of gain stages can help - are the frequency response of the tone controls and the frequency response and efficiency of the speaker system.

1.9 Decibel references

Different references used in dB calculations can sometimes be confusing so I felt it is beneficial to include a short description about most common ones. There are mainly two kinds of references for voltage, dBV and dBu (or dBv); the latter is specified for impedance of 600 ohms. Note the use of upper and lower case letters. The dBV means that the reference is 1 V across any impedance, while dBu and dBv use a reference of 0.775 V. “u” stands for unloaded so in its case the reference is not related to impedance. The reason for using dBv unit comes from a fact that dBv (and dBu) equal dBm when the impedance is 600 ohms. This was the input and output impedance used in old radio technology. To convert dBV to dBu add 2.2 dB.

The common references for power are dBm and dBW. The dBm means power relative to 1 milliwatt while dBW means power relative to 1 watt. Usually the unit should be specified as dBm (1mW) and dBW (1W). 0.775 V (RMS) signal to 600 ohms results into average power of 1mW, hence the reference level dBv.

Acoustics use a measure of sound pressure level (SPL). The reference level, dB (SPL), is 20 micropascals, which stands for quietest sound an average human can perceive.

1.10 About various types of distortion

The term distortion means alteration of a result signal in comparison with the original one. Roughly, distortion types can be separated into two categories: Linear and non-linear. Non-linear distortion always adds additional harmonics to the fundamental signal content. These can be presented by comparing the distorted signal to an ideal signal that has no other harmonic content than the fundamental frequency. Such an ideal signal is a sine wave. The additional harmonics can be even or odd depending on the multiplication ratio with the fundamental. Second harmonic has a frequency that is two times higher than the fundamental frequency, third harmonic has three times higher frequency and so on. Sub harmonics are harmonics that have a lower frequency than the fundamental one. The distortion modes can be divided into four categories:

1. Frequency distortion is taking place when all frequencies do not get amplified evenly. This form of distortion is “linear” meaning it doesn’t add additional harmonic content to the signal. In essence, anything related to deviation from flat frequency response can be considered as distortion. However, this type of distortion is very common and even sought after.
2. Intermodulation distortion is created by interaction of two signals that have a different frequency. For example, summing two waves also creates frequencies that are equal to sum and difference of the two fundamental frequencies. The process also creates the sum and difference frequencies of all harmonics. Their amplitude decays in relation to harmonic order. In the 1960’s another cause of intermodulation distortion was discovered by professor Timo Otala: The high frequency phase shift in feedback loops causes a delay on transient signals and the asymmetric summation of two, reverse phased transients creates frequency

components that did not exist in the signal before. The distortion type was named Transient Intermodulation Distortion (TID).

3. Phase distortion is occurring when the result signal has a distorted phase in relation to the original. This distortion type is also very common but mostly inaudible. However, one must understand that in certain applications any phase shift must be strictly avoided. For example, in video electronics reversed phase means inversion of the picture.
4. Amplitude distortion means that the signal levels of the result signal are distorted in comparison to the input signal. This distortion commonly occurs when a gain stage gets overdriven, thus causing the active device within the stage to saturate, cutoff or both. If we consider the process from the view of a single sine wave it is obvious that any distortion in wave shape will generate new harmonic frequencies. So-called crossover distortion is a form of amplitude distortion created by the requirement for a certain forward voltage needed to turn the active device on.

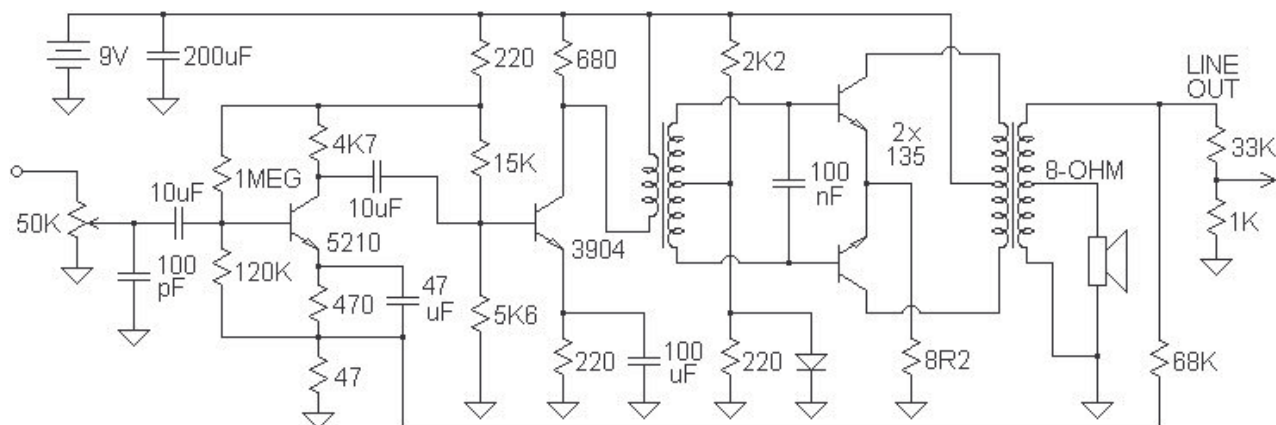
1.11 Some history of (transistorized) guitar amplifiers

The first solid-state power amplifiers were copied from tube designs and they used an output transformer that both coupled the output and worked as a collector load. Often a third (interstage) transformer, operating as a phase splitter, was incorporated as well. A demand for lower distortion and higher damping factor – and thus more linear output – however, soon pushed the designers to favor emitter follower circuits. With transistors this was much more ideal solution since it allowed omitting the bulky and expensive output transformer. However, since ideal class-B amplifiers require both NPN and PNP -type output devices - and NPN power devices were difficult to manufacture from Germanium, which was the top-notch semiconductor material of the time, - a transformer was still needed; this time to reverse the signal phase: If two drive signals in reverse phase were used, the power amplifier required only single polarity power devices. This so-called “totem pole” circuit arrangement was also a relic from vacuum tube times: A similar output transformerless design was used in driving capacitive (and fairly high) loads of coaxial cables. Due to the low emitter resistance of transistors it was even easier to push more output power from this design, yet due to the small amount of negative feedback it still retained a quite low damping factor. A schematic of the circuit topology is shown in Figure 3.9.

Note: A “totem pole” is a poor selection for a name since it was frequently used in describing any schematic block that contained stacked devices. In fact, the name tells us nothing about the circuit. A far better name would have been a stacked output stage with an input transformer.

One of the “breakthrough” innovations in solid-state power amplifier design happened in 1956 when Dr. Hung Chang Lin combined the Darlington- and Sziklai pair circuits together forming a “quasi-complementary” circuit. The topology allowed using identical, single polarity power transistors, which were driven by two complementary low-power transistors. The invention made it possible to reach high output powers, yet omit transformers completely from the signal path. In 1956 the circuit was published in Electronics magazine and in 1959 it was patented. Yet for almost a decade it was left unnoticed by most manufacturers. During those times tubes were still the number one choice for most audio applications. In 1963 H. J. Leak announced the “Stereo 30” amplifier, which was one of the first commercial *transformerless* transistor amplifiers and utilized the Lin’s idea. Soon after that the leading semiconductor material shifted from germanium to silicon, which was more reliable. The problem of finding well-

matched complementary devices, however, still persisted. It actually took a few years and a set of RCA application notes to popularize the Lin's idea but ever since then the design has been extensively used until very recent times when cheap and high-quality complementary power transistors have become widely available. Even now the quasi-complementary topology is widely used in integrated chip amplifiers due to their fabrication limitations. An example of the "Lin-topology" can be seen in Figure 3.6. Lin also invented lateral transistors and his studies with thermal effects of semiconductors led to utilization of diode-compensated bias configurations, which are now used in majority of transistor amplifiers.



Circuit design reminiscent of early transformer-coupled transistor amplifiers: Pignose 7-100.

Around 1965, transistor guitar amplifiers finally had their breakthrough. Kay had already introduced few of them in 1962, Gibson and Hagström in 1963 and Watkins Electric Music (WEM) and Vox in 1964. However, 1965 stands out as the year when almost every major company released their first transistor guitar amplifier. The list includes Selmer, Standel, Carvin, Peavey, Unicord, Thomas Organ, Dallas, Baldwin and likely several other smaller companies that have been long gone and forgotten. Fender, Kustom, Teneyck, Jordan and Silvertone followed the next year. For a while some of these amplifiers were the only equipment that famous acts would even consider using: For example, Jordan's transistor amplifiers were the loudest available. For the annoyance of some mod bands, that were trying to get higher and higher volume levels, they even beat the Marshall's tube amplifiers in loudness.

Since "super popular bands", that needed to play stadium concerts, had just appeared the first time in history the speaker systems of the "flagship" models were starting to get humongous: Gibson Les Paul LP-2 powered speaker cabinet had four 12-inch drivers, plus two 10-inch horns. Vox Super Beatle came with a cabinet housing four 12-inch speakers. ("Super Stack" had two of these cabinets). Baldwin's response to Vox was the 100-watt "Exterminator" with two 12-inch, two 10-inch and two 7-inch drivers – all in a single cabinet. Not to mention the horn cabinets or columns from Acoustic Control, Sunn and Kustom that appeared in the late 60's. Note that just a brief moment before all this, a 30-watt tube amplifier with two 12-inch speakers was considered more than adequate. To put things in perspective, back then the average PA used by almost any famous band was generally worse than the one you can find from an average club today.

Another major breakthrough happened at the end of the 60's (ca. 1967 – 1968) when RCA introduced a set of reliable, high-power transistors and simultaneously published a selection of application notes that introduced circuit topologies such as quasi-complementary output stage, differential input stage and VI limiting. These designs

finally popularized circuits that omitted bulky interstage transformers and were at least modestly reliable. Truthfully, the presented designs were still quite flawed (to put it mildly): The accuracy of transistor datasheets was a joke when compared to precise tube datasheets – and people had accustomed to design principles that included copying the application note circuit and assuming that the printed device information was accurate. The transistor technology was still far too new. Nevertheless, the basic operating principles were correct and the circuits influential and inspiring for countless audio designers. One can't underestimate the importance of cheap, and reliable new transistors that emerged either: For example, the “industry workhorse” 2N3055 was costing less than a dollar! After the introduction of RCA papers the amount of transistor-based power amplifiers skyrocketed.



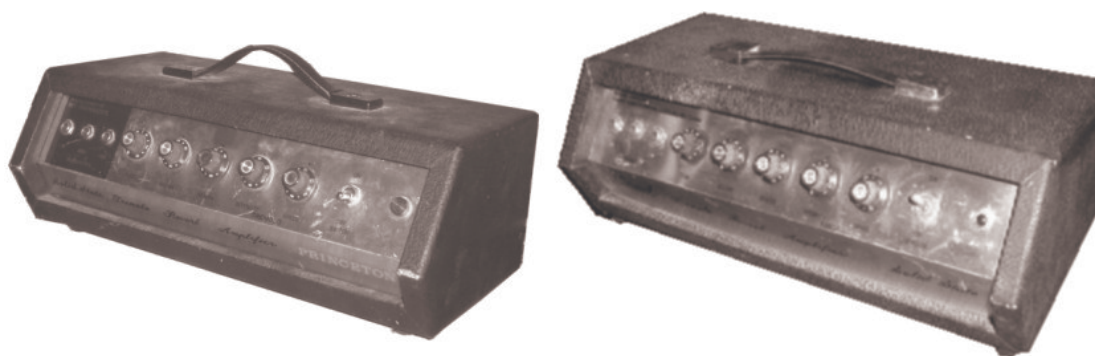
Some nearly forgotten brands from the sixties: Teneyck G3000 (1966) and Baldwin Professional Deluxe with multicolour push-button switches for “Supersound” circuit (ca. 1967). Teneyck amplifiers had a distinct similarity with Ampeg’s “Golden Glow” series from 1967 - 1968.

Thus the 70's were the real dawn of transistorized guitar amplifiers. Since most of them were entry level models, deliberately built cheaply and sold to people who could not afford a tube amplifier (that was usually higher quality) few of the designs stood out. The serious consumer HiFi market, on the other hand, had completely abandoned tube amplifiers, which led to the downfall of many major tube manufacturers and further broadened the gap between quality of transistor and tube designs. There were high quality designs, of course, but in such a small market as guitar amplifiers these were extremely rare – not to mention too expensive for most people. The reliable and linear solid-state designs from the era were mostly either HiFi or public address amplifiers.

During the years, solid-state audio and related technology has been rapidly developed by theory and innovations of famous designers such as (in no particular order) Paul Penfield, Bob Widlar, John Linsley-Hood, Erno Borbely, Malcolm Hawksford, Robert “Bob” Cordell, Marshall Leach, Bob Carver, Jean Hiraga, Rudolph Lender, Bob Pease, James Bongiorno, Douglas Self, Randy Slone, Daniel Meyer, Nelson Pass, Walt Jung, John Curl, Ron Mancini and numerous other people whose work has been credited for major companies.

One shouldn't forget the influence of instrument - and PA - amplifier designers who specifically worked with solid-state gear: *Robert Moog*, *Roger Meyer* and Electro-Harmonix designers *Mike Matthews* and *Bill Berko* had a huge impact on the development of various, revolutionary effect circuits. *Bob Rissi*, together with *Paul Spranger*, designed the first (and infamous) transistor Fender amplifiers; however, Rissi also designed the Rickenbacker's respected Transonic lineup. The latter aimed to

correct most of Fenders' mistakes and proved to be high quality and very innovative; the series was endorsed by Steppenwolf, Led Zeppelin and Jeff Beck also used these amplifiers on some occasions. Spranger later went to work for Altec.



Princeton PG200TRS Solid State Tremolo Reverb (left) and National GA920P Tremolo Reverb (right) amplifiers are by most of their features typical “bargain bin” SS amplifiers from the period between late 60’s and early 70’s: Both have low output power and transformerless quasi-complementary design with highly questionable thermal stability. Features like crude tone controls, spring reverb and tremolo circuit were of course a must. “Hip” catches included a slanted cabinet shape and an advertisement of using new FET transistors (only one was used in the entire amplifier). The piggyback amplifiers came with a matching 2 x 8” cabinet. The manufacturer for both Princeton and National was Japanese Panasonic. Selling same amplifiers under different brand names was (and still is) very common in the guitar amplifier market.

Bud Ross was the founder and head designer of Kustom and when it was sold he founded Road Electronics. The latter built powerful, high-quality amplifiers that were housed inside a durable flight-case construction. Unfortunately they are mostly unknown today. Road Electronics eventually merged with Rickenbacker that produced a “Road” series of amplifiers. Most of us likely remember Ross better from his later work with Ross pedals and amplifiers. Another “merger” with Rickenbacker was Jordan: The whole Jordan company was actually disbanded and its technical team moved to work for Rickenbacker. *George Cole* (the owner) became Rickenbacker’s Chief Engineer, *Bob Garcia* (Jordan’s Chief Engineer) became electronics factory manager and *Manny Gabler* (General Manager) became a salesman.

Robert “Bob” Teneyck worked for Ampeg developing their Gemini series of tube amplifiers in the early 60’s. Being a competent designer he held the patents for the company’s vibrato and tremolo circuits (US Patent 3,388,257). However, Teneyck also ventured to manufacture his own “Teneyck” transistor amplifiers between 1966 and 1967. Teneyck is also responsible for designing the first solid-state Sunn amplifier, the Orion, in 1969. The concerned amplifier was endorsed by none other than Jimi Hendrix. Unfortunately it was quite unreliable and therefore proved to be a commercial failure.

Seth Lover, the inventor of Humbucker pickup (especially the P.A.F -type), moved from Gibson to work for Fender and designed the solid-state Super Showman (in 1969). *Dan Pearce* was a co-designer in Moog’s “Lab Series” team and later founded Pearce Amplifier Systems that produced some high quality transistor amplifiers. Before starting his own business *Paul Rivera* worked for Fender. He didn’t design only tube gear but solid-state amplifiers like Yale, Montreaux and Studio Lead as well. Rivera also worked for Yamaha (G-100) and Pignose (150R Crossmix). *Hartley Peavey* and (former RCA designer) *Jack Sondermeyer* were the team behind Peavey amplifiers. First of them were designed by *Ron Matthews* (from Orradio) and *Jim Askew*, though. *Don Randall* did design work for Fender until he founded Randall.

And of course one can't forget *Bob Heil* (Heil Sound and inventor of "talk box" effect), *Norm Sundholm*, *Conrad Sundholm* and *Dick McCloud* (Sunn), *Dick Denney* (Vox), *Tommy Gumina* (Polytone), *Steven Marks* (Acoustic Control) and *Bob Gallien* (GMT and Gallien & Krueger).

Many former Acoustic Control employees also set up their own businesses: *Russ Allee* founded AMP together with *Roger Smith*, *Steven W. Rabe* founded SWR Engineering, *Aspen Pittman* founded Groove Tubes and ACC's consultant in speaker design, *Gene Cerwinski*, is the "Cerwin" of Cerwin-Vega. *Harvey Gerst* ventured to design a monstrous Concept 1 tube amplifier for Delta Products Corporation. It is rumored that the design was meant to drive Peavey out of business, however it didn't become successful and only about 100 units were built. Nevertheless, the Concept 1 sported many groundbreaking features: It had a built-in Ford 8-track tape-echo, 8 power tubes and 6 preamplifier tubes, variable damping control, a separate winding on the output transformer to handle feedback, "Studio Position" for lower output power (130 W average versus 10 W average @ 1kHz 1% THD) and five EQ pre-sets that mimicked frequency responses of Telecaster guitar and Marshall, Gibson and Fender amplifiers. The initial "Delta" frequency response was flat. At the time when the amplifier was introduced most of these features were considered incredibly revolutionary.



When introduced, the "tuck-n-roll", naugahide-covered Kustom amplifiers were by many ways "extraordinary".

There is a lot of controversy about who introduced the first all-solid-state guitar amplifier. The oldest reference I've seen so far is Paul Penfield's article "Transistorized Guitar Amplifier" in July issue of Radio & Television News magazine, 1956. This was not an actual unit, though, merely a description of how to build one. Nevertheless, magazines quite commonly included these do-it-yourself articles during the 1950's and the 1960's, mostly because analog electronics was a vastly more popular hobby than it is today. Semiconductor (and tube) manufacturers also liked (and contributed) them since they increased sales. No doubt they have had a huge impact on designers that worked for companies that built guitar amplifiers. The first reference to a real device I've seen so far dates back to 1960: Wandré's "Bikini" and Hofner's "Bat" guitars included an integrated transistor amplifier and a speaker. The circuitry was very similar to early low-power transistor radios containing germanium PNP transistors and a totem-pole push-pull output stage.

A Czechoslovakian company called Jolana introduced a similar product, the "Big Beat", in 1963. Out of the factory in Krnov came a transistor amplifier that you could bolt into one of the company's guitars using a set of special screws. The product was extremely innovative: For one, the amplifier had a built-in radio receiver so you could switch between guitar pickup or medium wave radio input. Maybe more importantly,

the 2-watt battery-powered amplifier was a *transformerless* “Lin-style” circuit, likely the first of its kind in a guitar amplifier.

The history of solid-state guitar amplifiers is very badly known and mostly undocumented. For a while I was under the impression that Hagström was the first company that manufactured a “separate” all-transistor guitar amplifier unit, the Model 1700 or “GA-85” (circa 1963). However, later research indicated that Kay had introduced a “Vanguard” series of transistor amplifiers already in 1962. While Kay amplifiers were pretty generic and low-power, the specifications of the Hagström GA-85 were by no means modest: It used germanium transistors, had a fully regulated power supply, output power of 85 watts (average) and housed two 12-inch Philips 5200 speakers and a Hammond Type 4 reverb tank. Also, Burns and Gibson transistor amplifiers had been introduced the same year as GA-85. Famous Vox T-60 is far from being the first transistor guitar/bass amplifier even though many people like to claim so. It was introduced in 1964.

I have compiled a small time line showing the development of guitar amplifiers – especially focusing on those ones that are solid-state. The first notes relating to transistor guitar amplifiers date to 1956 and one should focus on how gigantic leaps the technology has taken since then. For example, at the beginning of the 70’s transistor amplifiers were sold by almost every manufacturer while “boutique” hi-gain tube amplifiers presented a professional – albeit an expensive – alternative.

- 1927 • Harold Black invents the concept of using negative feedback to linearize an amplifier
- 1929 • Stromberg-Voisinet is the first company to market a functioning “electrified” string instrument and amplifier set
 - Various electric guitars with dynamic pickups hit the market. They are plugged into PA amplifiers
- 1932 • George D. Beauchamp’s patent application describes the operation of magnetic pickups and shows the “first guitar amplifier”
 - First Rickenbacker amplifier, “The Speaker”, emerges
 - First Electro amplifiers
- 1933 • Various amplifier brands sold in a set with (usually) a lap steel or Hawaiian guitar broaden the gap between dedicated guitar and plain PA amplifiers. Brands like Vivi-Tone, Volu-Tone, Vega, Dobro and Audio-Vox appear
- 1934 • First Epiphone (tube) amplifiers
- 1935 • First Gibson (tube) amplifier EH-150 (“Electric Hawaiian”)
- 1937 • Rickenbacker’s “Black Metal Box” is the first instrument amplifier that is housed in completely metallic enclosure (including the cabinet).
- 1946 • K & F Manufacturing amplifiers pave the way for Fender
- 1947 • First point-contact transistor is built in Bell Labs
 - Fender introduces their first amplifier: Model 26 “Woodie Deluxe”. Tube – of course.
- 1948 • Fender introduces “Broadcaster” (later renamed Telecaster) solid body electric guitar
 - Bell Labs team, lead by William Shockley, announces operation of the first transistor
- 1949 • *First 45 RPM, 7” records are introduced*

- 1950 • William Shockley develops bipolar junction transistor
- 1952 • *Age of Country & Western music*
- Gibson introduces Les Paul solid body
- Baxandall presents the negative feedback tone control circuit
- Bell Labs patents the Darlington transistor topology
- 1953 • Stadel is the first company to use J.B. Lansing speakers in guitar amplifiers
- First transistor radio is unveiled
- 1956 • *Elvis hits the charts*
- Lin introduces quasi-complementary topology
- Paul Penfield's article "Transistorized Guitar Amplifier" appears in July issue of Radio & Television News magazine
- 1957 • *Concept of stereo records is introduced*
- Stadel introduces "piggyback" amplifiers, front-mounted controls, closed-back cabinets and reflex structure
- First Vox (tube) amplifier (AC2/30) is released
- 1958 • *Link Wray hits the charts with distorted tones of "Rumble"*
- Vox releases AC30/4
- First all-solid-state organs are released
- 1960 • *The age of instrumental music and "British invasion": Beatles releases their debut album*
- Wandré's "Bikini" and Hofner's "Bat" guitars are equipped with integrated Davoli Krundaal CT642 *transistor* amplifier and speaker
- Fairchild introduces the "planar" manufacturing method that allows making cheap and "high quality" transistors
- Ferranti introduces first commercially acceptable Silicon transistors
- 1961 • Stadel releases the first "hybrid" amplifiers
- Based on Lin's inventions, Dick Tobey and Jack Dinsdale present a DIY transformerless class-B quasi-complementary amplifier in Wireless World magazine
- First *compact* reverb unit is introduced and immediately adopted by various guitar amplifier manufacturers (such as Fender)
- 1962 • Ken Bran and Dudley Craven (employees in Jim Marshall's store) begin to replicate Fender Bassman amplifiers giving birth to JTM45
- Kay's Vanguard line-up: Likely first all-solid-state guitar amplifiers.
- First fuzz box effect, Maestro FZ-1 Fuzz Tone, appears
- 1963 • First, poor quality, JFET transistors emerge
- Leak Stereo 30, first commercial amplifier that uses only mains transformer, is introduced
- Jolana "Big Beat", likely the first transistor amplifier for guitars having a "Lin-style" circuit.
- First Hagström transistor amplifier is introduced.
- First Burns transistor amplifier (likely the "Orbit" model). At the time this was more expensive than a new Vox AC-30.
- Likely the first Gibson transistor amplifiers emerge: These are "Starfire" series consisting of TR-1000RVT and TR-1000T.
- 1964 • First transistor amplifier from Vox, the T60 Bass.
- First WEM (Watkins Electric Music) transistor amplifiers.
- First commercial digital amplifier, Clive Sinclair's X10 DIY kit, is released (designed by Gordon Edge).

- 1965
 - *Rolling Stones “Satisfaction” launches a fuzz box craze*
 - Huge speaker stacks become ordinary arsenal of bands
 - First Selmer transistor amplifiers. Likely the “Taurus” was the first one. It was later changed cosmetically and renamed “Saturn 60”.
 - First Standel transistor amplifiers.
 - First Baldwin transistor amplifiers. (Baldwin had just bought Burns).
 - First Carvin transistor amplifiers. T-11 as the first one.
 - First Unicord/Univox transistor amplifier (BT505 Bass) appears. However, a true solid-state Univox guitar amplifier line-up is introduced as late as in 1971.
 - Thomas Organ takes over the US Vox supply and starts to build their own line up of amplifier models.
 - Hartley Peavey founds Peavey Electronics and introduces “Musician” and “Dyna-Bass” amplifiers designed by Ron Matthews.
 - Likely first Dallas transistor amplifiers appear.
 - Solid-state IC operational amplifiers become widely available
- 1966
 - Standel releases first active speakers (called “acoustic magnifiers”)
 - First Fender transistor amplifiers. These were designed by Bob Rissi and Paul Spranger
 - Bud Ross founds Kustom.
 - First Teneyck transistor amplifiers.
 - First Jordan transistor amplifiers.
 - First Sears Silvertone transistor amplifier models; manufactured by Danelectro.
 - Dallas Arbiter issues first Fuzz Faces
 - Mosrite introduces their lineup of transistor amplifiers and fails commercially with the Award BG-500 “The Ventures” model.
- 1967
 - “Brute 70” (OTL) amplifier kit is introduced by Popular Electronics magazine along with M/M/M Instrument Amplifier (OTL)
 - First all-transistor Magnatone amplifiers
 - Likely the first Premier transistor amplifiers are introduced.
- 1968
 - RCA introduces a new set of high-power transistors and publishes a collection of application notes concerning quasi-complementary and differential input stage topologies. These are highly influential for designers of transistor power amplifiers.
 - Acoustic Control Corporation releases its first amplifiers.
 - GMT 226A, designed by Bob Gallien, is the first instrument amplifier to use a stacked power transistor configuration (known as “cascade” or “beanstalk”). Next year Carlos Santana uses the amplifier at Woodstock.
 - Tommy Gumina founds Polytone
 - Rickenbacker releases “Transonic” series

- 1969
 - *Woodstock music festival*
 - The popular “Tiger” amplifier kit is introduced by Popular Electronics magazine
 - John Linsley-Hood publishes the famous 10 W Class A amplifier article (OTL)
 - Output transformerless power amplifiers become popular
 - Sunn introduces their first solid-state amplifier “Orion”, which is endorsed by Jimi Hendrix. The amplifier is a commercial failure and susceptible to field breakdowns.
 - Ovation and Lawrence release their first transistor amplifiers
 - Electro-Harmonix LPB-1 “Linear Power Booster” starts a craze for signal booster pedals
- 1970
 - *Black Sabbath releases their debut album*
 - Heathkit introduces kit amplifier models TA-17, TA-16 and TA-27
 - Birth of Kasino amplifiers. Kasino was a daughter brand of Kustom.
- 1971
 - Sunn is first to experiment with deliberately overdriving FETs
 - GMT 600B is considered by many to be the first instrument amplifier in plain metal enclosure (1937 Rickenbacker *was*)
 - First *successful* audio power amplifier IC, Sinclair’s IC12, is released. Its output power is puny 6 watts average.
- 1973
 - Russell O. Hamm’s article “Tubes Versus Transistors” is published
 - Marshall introduces their first solid-state amplifier line-up.
- 1974
 - First Music Man hybrid amplifiers are released (Sixty-Five series).
 - GMT 200G is (likely) the first channel-switching guitar amplifier. Mesa Boogie, Session, Fender and Marshall implemented the channel-switching feature to their amplifiers as late as in the early 80’s.
- 1975
 - Roland’s “Jazz Chorus” line-up is introduced. During the years the design goes through a number of changes.
- 1976
 - Radio Electronics introduces the concept of rail switching (class-G) and the next year Hitachi uses it in the Dynaharmony HMA 8300
 - Yamaha introduces first amplifiers that use MOSFET output devices
- 1977
 - Sunn Beta series amplifiers are first commercial products utilizing distortion stages based on overdriving logic gate ICs
- 1978
 - *Van Halen releases their debut album*
- 1979
 - Lab Series amplifiers, designed by Moog Synthesizer (division of Norlin), are introduced - year later the Lab 2 Series.
- 1981
 - Gallien-Krueger 200RB is the first rack-mount instrument amplifier
 - Gallien-Krueger 800RB is the first (bass) instrument “bi-amp”. It also introduces first removable AC cord and DI circuit
 - TUSC introduces the “Prestige” series hybrid amplifier with a DSP-based preamplifier (supporting memory patches) and a tube-based power amplifier stage.
- 1982
 - Following the success of Sony Walkman (1980) Scholz Research & Development introduces the compact Rockman headphone guitar amplifier
- 1983
 - *Metallica releases their debut album*
 - John Murphy designs a tube emulation circuit for Carvin amplifiers
- 1984
 - *Apple introduces personal computers*

- 1987 • Eric K. Pritchard patents his ideas of tube emulation. At the time they are implemented to PRS Harmonic Generator series of amplifiers. Pritchard later ventures to found his own company.
- 1989 • Tech 21 Sans Amp is introduced
- 1991 • *Nirvana releases their breakthrough album “Nevermind”*
- Marshall introduces the Valvestate lineup
- 1995 • Gallien-Krueger 2000RB is the first 1000W instrument amplifier
- 1996 • *Breakthrough of Internet*
- Line 6 is founded and releases AxSys
- 2000 • Line 6 releases POD launching a DSP craze
- 2005 • Vox introduces Valvetronix technology

1.12 The dilemma of “classic” amplifiers

The topic of famous solid-state amplifiers was slightly dealt with in 1.5. However, what is less well known is that behind the cosmetic exterior of the supposedly classic amplifier might lie something completely different. Take for example the Roland Jazz Chorus JC-120: There are at least eight or nine different versions of this respected amplifier – most of them quite different from another. The fact that earliest versions lacked features like brightness switch, external input jack or effect loop did not yet mean a drastic change in tone but what do you think is the effect of input impedance differences in the order of hundreds of kilo-ohms, resistor versus CCS loaded differential input stages - or perhaps IC-based preamplifier versus discrete preamplifier? How about totally different “distortion” circuits in almost each version - not to mention completely different component values and circuit topologies? And I used to wonder why people had such contradictory opinions about these amplifiers.

Roland has made vague attempts to explain the circuit changes: According to them the *real* model name is stated in cryptic letter code following the basic model name (e.g. JC-120-U, JC-120-JT). Unfortunately, the logic behind these model “names” is explained nowhere. In practice, JC-120s are mainly dissimilar amplifiers that just happen to share the same cabinet cosmetics - and not that many people realize this issue. I haven’t looked into this matter deeply enough to know whether even the speakers have been changed along the revisions as well. Now, which one of the many different circuits is *the* JC-120 that people fuzz about? If you ask me, the most “advanced” Jazz Chorus circuit I’ve seen actually dates back to 1979. Since then the circuit has been gradually simplified and its quality and reliability have decreased. They still are good amplifiers – but not as good as they used to be. Unfortunately the preamplifier circuit of the 1979 model was mainly based on ICs that are *obsolete* today.

How about another series from Roland: The “Cube”. This is another praised amplifier line-up but one should know that there are at least two “versions” of them as well – one from the 80’s and one from the late 90’s. Latter incorporates DSP. Don’t get me wrong, I’m not trying to mock Roland here; other manufactures fall as guilty on similar practice as well: For example, there are two almost completely dissimilar versions of Gibson’s “Thor”: One based on OpAmps, one using discrete circuitry – and two different versions of Gibson GSS-100. For the record, Gibson is “famous” from making frequent (minor and larger) design changes to its amplifier models during 50’s and 60’s. From certain periods of that era, there is no guarantee that any Gibson amplifier matches the published schematic.

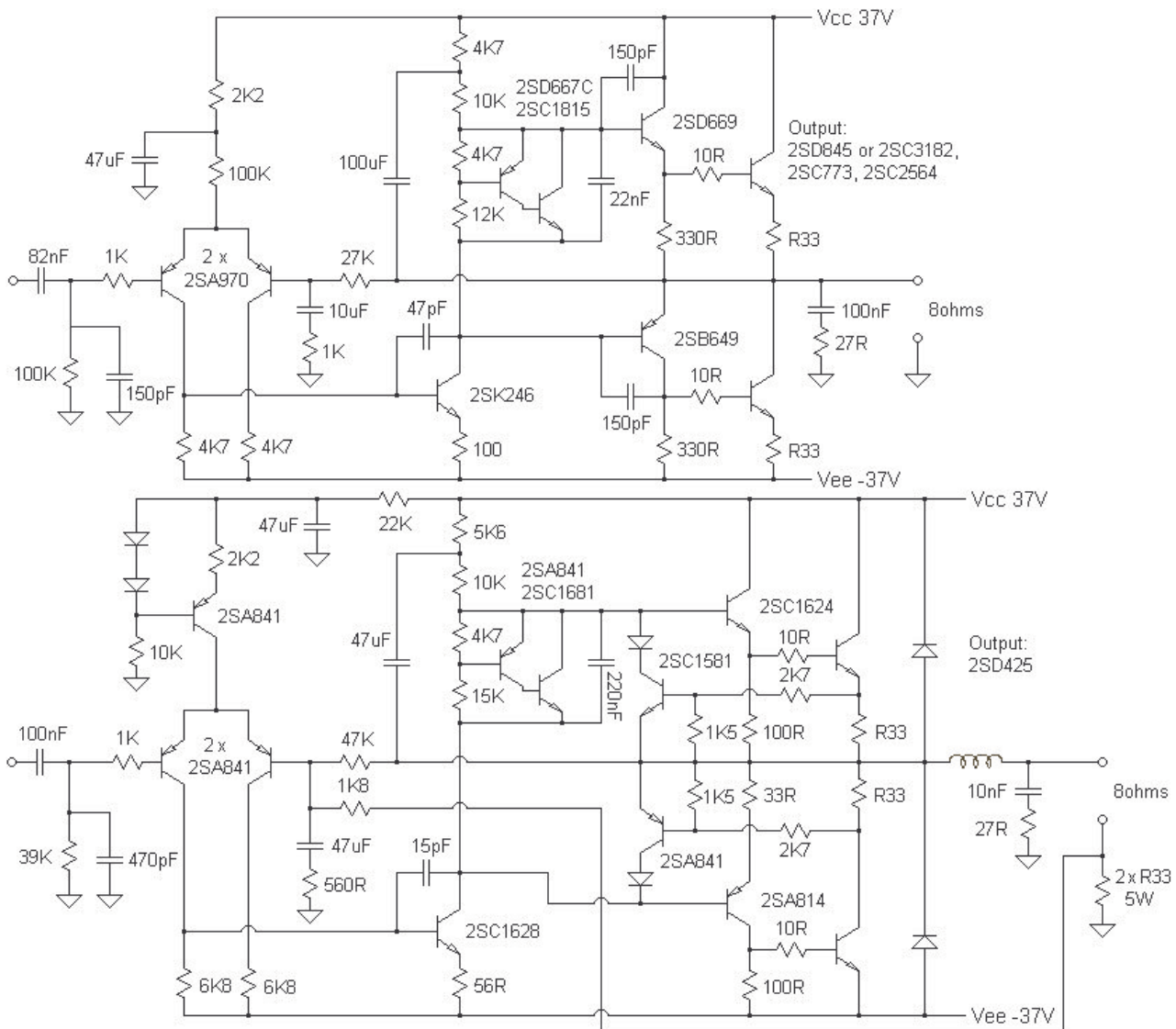


Figure 1.11 Power amplifier stages from two versions of Roland JC-120: Different transistors, different component values and partially different topologies. The same? No way. Notice the lack of mixed-mode feedback, constant current source loading and short circuit protection in the upper schematic.

All aforementioned are old amplifiers so their circuit diagrams are circulating around – with new amplifiers and designs one can never be sure what goes on. You have likely heard that Peavey Bandit is a good amplifier: Peavey has made about 10 different “Bandit” amplifiers using various topologies. Which one of them is *the* praised Bandit? This just goes to show that amplifiers having an identical model name are not necessarily identical at all and in worst cases the designs are very far from each other. There are revisions, which usually improve reliability or substitute parts that are no longer in stock - all this while still retaining the basic circuit topologies. ...And then there are “revisions” that change the design completely.

1.13 The ideal of simplicity

The quasi-complementary circuit mentioned earlier is a perfect example of how a good circuit sticks around and how focusing on details, that on the first glance may seem to have less importance, can make some drastic improvements. The improvements in modern transistor slew rates, component quality, and innovations concerning constant current circuitry, input stage and feedback topologies have made

more difference for the overall amplifier sound than the output topology it uses. Unless the amplifier is extremely badly designed, the differences between most emitter follower output topologies should be so small that they are undetectable by ear.

During the time designs tend to get bloated; there are many reasons for this but the most common one is that when an amplifier circuit is designed to perform better it usually gets more complex in the process. However, the simplicity is a virtue strived for and statements such as “there should be the smallest amount of components in the signal path as possible” or “complex designs make the sound muddy and increase distortion” are heard from mouths of numerous people. While such claims are mostly untrue, they can be surprisingly easy to believe if one examines the circuitry of some older amplifiers – especially if they are based on vacuum tube technology: The circuits are almost without exception very minimalist and contain only the least amount of “HiFi” circuitry, such as constant current sources or active tone control stages.

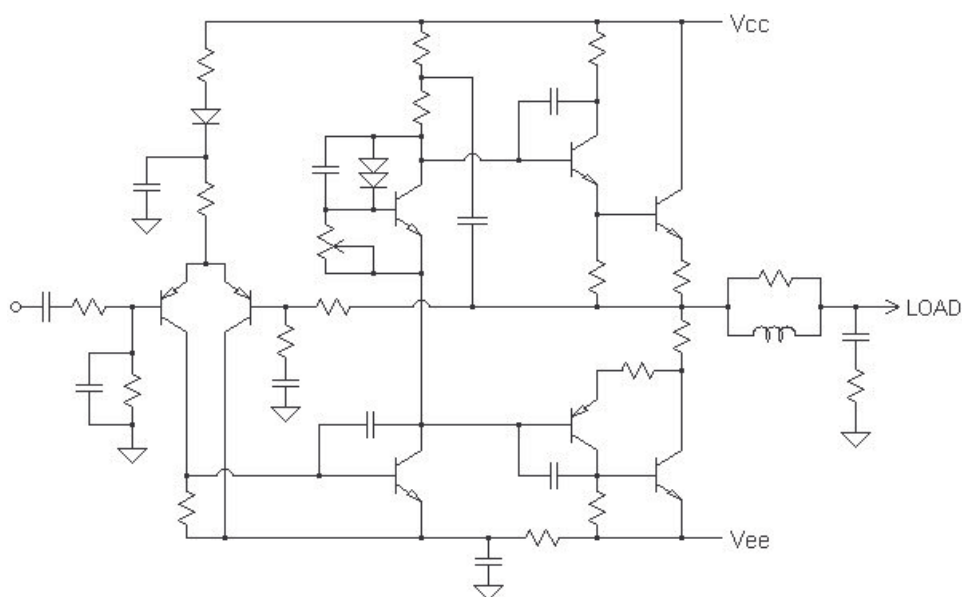


Figure 1.12 Circuit diagram of Harman Kardon Model 630 HiFi power amplifier.

A perfect example of minimalist amplifier design is Harman Kardon Model 630 (Figure 1.12). Despite its simplicity, even many High-End audio enthusiasts consider it as a great sounding amplifier. The design is very typical for amplifiers from the 1970's to 1980's. Except for varying component values, one could assume to find something similar inside almost any transistor amplifier of that era. When compared to what could be a typical modern amplifier design in Figure 1.13 the circuit seems surprisingly straightforward. Now, if you think that I have deliberately selected a very complex circuit as a reference I can safely assume you that this is not the case. Today, when compared to tight packed amplifiers full of SMD components, even this circuit can be considered as amazingly simple.

I conducted some SPICE simulations in order to find out if simple circuits truly have some benefits over complex ones. In the simulation two differential input stage power amplifiers were compared. The gain was set identical for both circuits - at least on 1 kHz sine input signal. The simple one used an “old” quasi-complementary output stage while the other, more complex design, used a Darlington complementary pair. The simple circuit had an imbalanced long-tailed pair input stage loaded by a simple

resistor. Complex one used a current mirror load to balance the LTP input, which in turn was fed by a constant current source. The voltage amplifier stage of the simple circuit was loaded by a bootstrap while the corresponding stage of the complex amplifier used an active constant current source load circuit. The signal paths and power supply circuits in both amplifiers were completely identical. Experiments also included deliberate “mismatching” of component values and internal properties of the transistors. For example, transistor parameters in both differential input and output circuits were deliberately mismatched - as were the values of the emitter and driver resistors. Identical mismatches were made for both circuits as much as the topology allowed. The purpose of this experiment was to simulate effects caused by real-life component differences. Tests were done using both simulation of a speaker load and purely resistive load. Both circuits are shown in Figure 1.14.

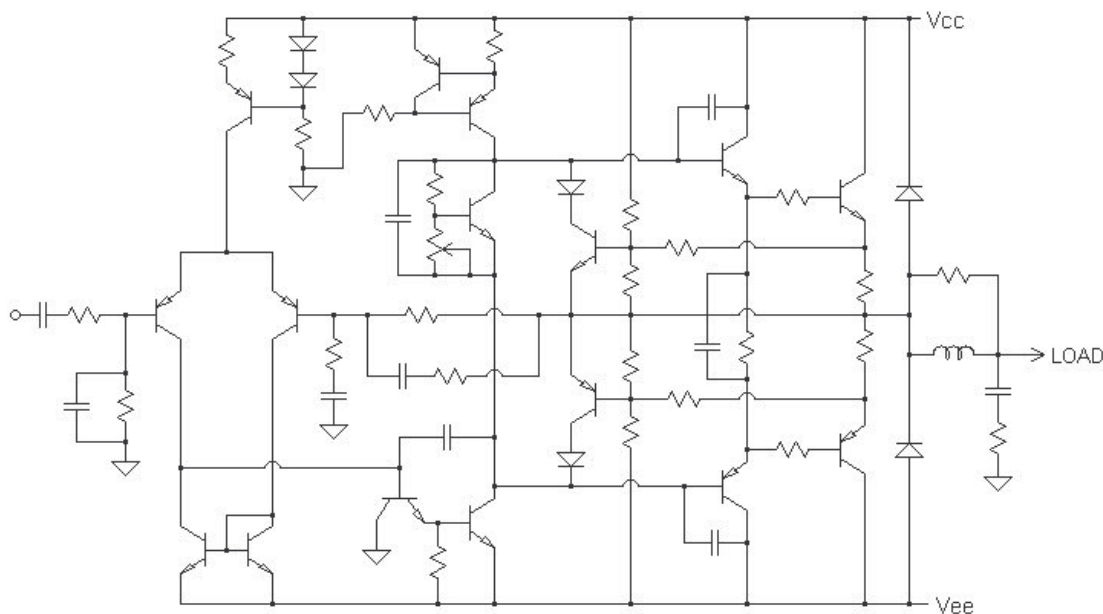


Figure 1.13 A typical modern power amplifier circuit.

In short, the results of the comparison were:

- The complex circuit had better power supply- and common mode rejection ratio. Amplitude of ripple was 40 dB higher in the simple circuit.
- The complex circuit was able to swing closer to rails.
- The simple circuit had higher DC offset in the output: A result of imbalance in the long-tailed pair input stage.
- The simple circuit had inferior distortion figures – especially the amount of second harmonic component was higher.
- The complex circuit recovered better from clipping.
- With a purely resistive load the simple circuit had a wider bandwidth towards high frequencies.
- Temperature variations had less effect on the complex circuit.
- Quite amazingly, subtle differences in component values and transistor models caused only little or no measurable difference in performance.
- The differences between output signals of both circuits were in every case undetectable by ear (even during clipping).

The experiment showed clearly that performance-wise the minimalist approach has little or no benefit at all. The use of simple circuits is also a big contributor to misconception that transistor amplifiers sound bad. The behavior of the complex

circuit was far more predictable in every aspect, which is always a design virtue. Then again, why did (and do) many (old) PA or guitar amplifiers use a simple approach? I can think of only one reason: Because the designers can get away with it. Remember there was no audible difference between the output signals of the simple and complex designs. If so, then why add something “irrelevant” to the circuit. If the predictability and stability of the circuit had become a problem then surely the designer would have chosen a more complex approach to solve it. During this book we will make a more thorough survey on reasons why and how simple circuits are improved and what other means can be used to shape the sound of an amplifier – usually in a lot more noticeable way.

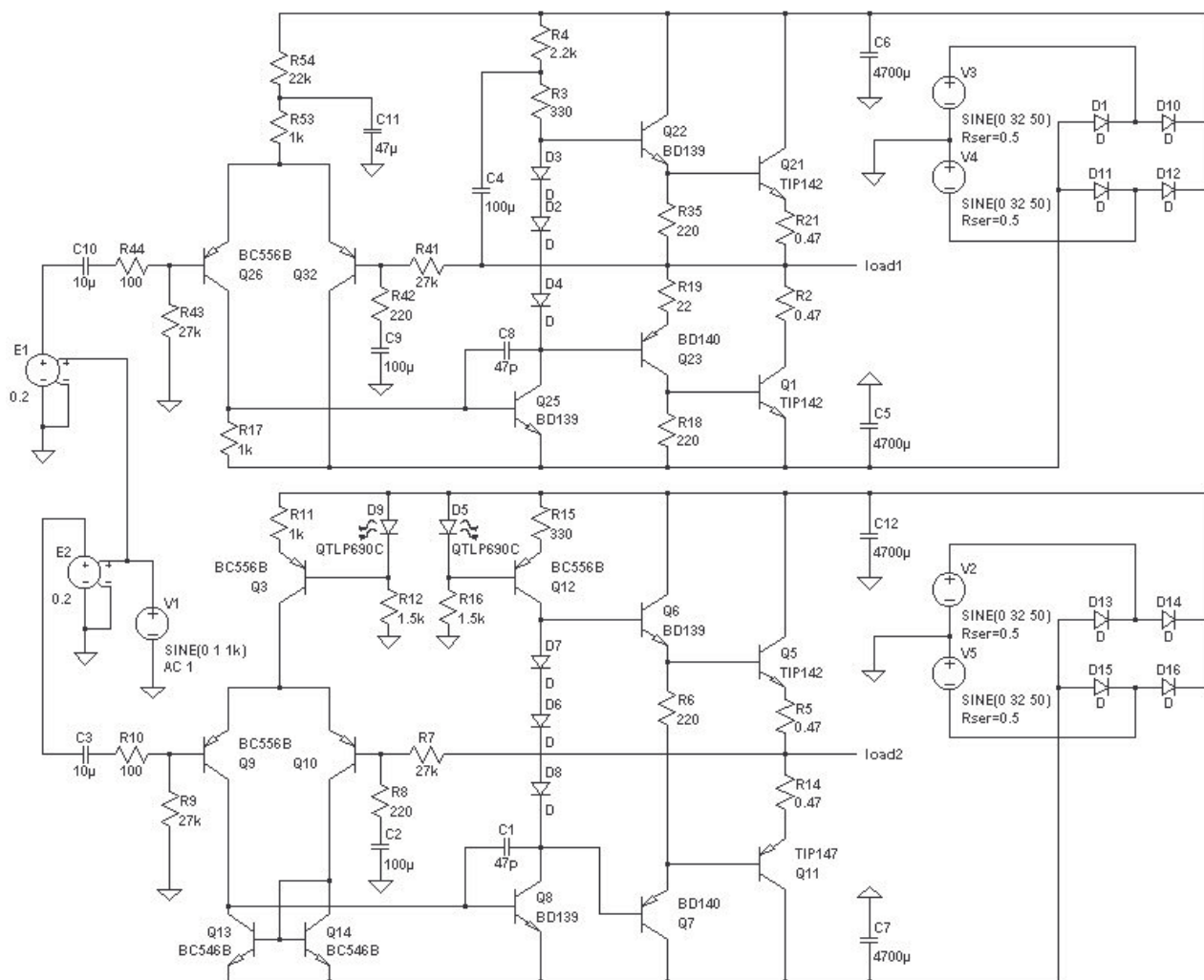


Figure 1.14 Circuits used in the SPICE simulation comparing simple and complex designs.

Essentially, we all expect an amplifier to possess certain features. This is a rather subjective issue, though: For some people simplicity is an appreciated feature, for others it's just a restriction. My opinion is that *simplicity is a virtue only when it does not affect such features as performance, versatility, reliability and tone*. In many cases a simple design cannot fulfill these demanding clauses. Features like better control of the frequency response, individual channels that can be toggled with a footswitch, reverb, effects loop - or even features we take as granted such as *linearity*, sufficient *loudness* and *reliability* – all tend to add to the complexity, parts count and cost of design. However, like the outcome proves this is not necessarily a bad thing at all. In essence, *complexity should not be feared* since that way of thinking may greatly reduce the versatility of the final product: In many cases it is worthwhile to break a (seemingly) complex circuit into subsections and evaluate the complexity again. A

modular building style goes hand-to-hand with this approach and usually lends itself to easier upgrading or modification. In commercial production it is also cheaper.

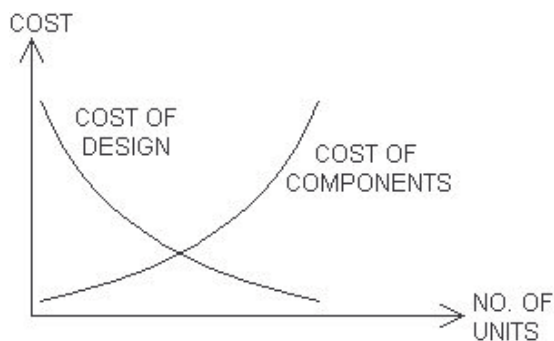


Figure 1.15 Cost of final product.

As a final thought on the subject: Many of the readers are likely do-it-yourself-oriented. Since “DIY” differs vastly from commercial production some facts should be considered: Since likely *only one or few* units are built the cost of components in relation to cost of design is rather small. The cost of design in this scenario is often very high – or better yet, *immeasurable*. Consequently, if you invest time (and money) in building an amplifier then you likely should build one that costs more but in time proves to be more versatile than a cheap and simple unit, which after the initial craze (from succeeding to build something that actually emits sound) is usually abandoned. This is a common case with low power designs that are generally categorized as “practice amplifiers”. Furthermore, if you do things correctly (and modularly) from the beginning you can easily modify or upgrade the design to suit your further needs. This is hardly ever possible with simple designs, which means you have to start the design and building process from the scratch again. However, a word of warning is in place: If you are unsure of your skills then a too ambitious project may prove to be less successful and thus disappointing.

2 A basic introductory to power amplifiers

2.1 A quick overview

Since power amplifiers may seem mysterious to some people a quick overview of their circuit structure is worthwhile before delving into more detailed examination. At first, it should be clearly understood that a power amplifier circuit does not (and can not) provide its output power out from a thin air - nor does it pass the input signal through in any way. *The input signal controls a circuitry that modulates the amount of current flow from the power supply to the load.* This means four very important things:

1. Both current- and voltage feeding capabilities of the power supply dictate the amount of output power. The output power cannot be higher than what the power supply can deliver.
2. The output power is also limited by the circuit's ability to exploit the power fed by the supply. (We refer to this factor as "efficiency").
3. The similarity of output and input signals is greatly dependant on the linearity of the amplifier circuitry. In other words, the amplifier circuitry *always* causes some difference (distortion) between controlling signal (input) and the result signal (output).
4. The output signal is (up to some level) modulated by all nonlinearities of supply voltage and current (for example, by the ripple voltage).

How the amplifier circuitry accomplishes the supply current modulation makes little or no difference from a simplified point of view since all power amplifier circuits basically operate the same way, which is; they utilize a certain number of active devices that alter the amount of current flow through them (think "valves"). The load is then connected in series with the active device(s). The circuitry can take the controlling signal either as a form of voltage or current and produce the desired gain either in the form of current, voltage or both.

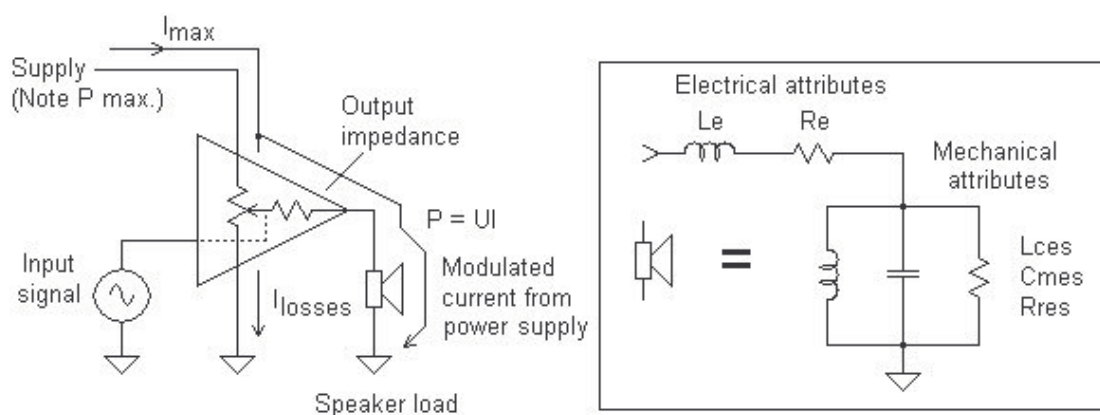


Figure 2.1 A conceptual schematic of power amplifier circuit.

The current flow and voltage potential that the circuitry can maintain at a certain load and in a certain determined manner will dictate the output power. Thus a typical method to measure the output power is to measure the *RMS voltage* of a sine wave output signal at a certain frequency and with a certain amount of distortion allowed. The measurement is done using a purely resistive load and since the load resistance is

known the power is easy to derive. A standard from the late 70's suggested using a measurement period of at least five minutes. This after all the channels of the amplifier had been preconditioned by driving them at 33 % of the rated output power for at least an hour. (The condition actually stresses most amplifiers more than running them at the full output power). Measurement was to be done with a 1 kHz sinusoidal input signal and the maximum amount of distortion allowed was 0.1 %.

However, the measurement techniques used by various manufacturers are quite random, which allows them to state various different output power figures that depended on such parameters as the input signal's frequency, amount of distortion in the output signal, and the time period that the measurement has lasted. Logically, when the amount of distortion is higher the measured power is also higher - thus some manufacturers can state that their amplifier has 100 watts of output power when it really has only 50 watts when measured with 0.1 % of total harmonic distortion (THD). For example, note that the DIN standard uses THD of 1 % instead of 0.1 %. Many guitar amplifiers are measured at THD of 5 % or 10 %. The output power may also depend on the signal frequency. Therefore the best practice is to state the relation of frequency response and power, say from 20 Hz to 30 kHz – usually with an acceptable THD value. This method will tell more about the actual power of the amplifier throughout its effective bandwidth. Some manufacturers quote the “peak power” rating. This is the output power derived from V_{PEAK} (instead of V_{RMS}) and is therefore two times higher than the average rating. Other peak ratings, such as the “musical power” or PMPO (“Peak Musical Power”), simply quote the amount of power that the amplifier can put out for some undefined, very small period of time. Since this period can be anything from few milliseconds to few microseconds such figure is essentially pretty useless. In practice, the only comparable measure is the average output power that the amplifier can sustain for a period of at least several seconds.

The procedure of quoting the power derived from RMS voltage has lead to a situation where people often incorrectly refer to the average output power as “Watts RMS”. Don't do this; it's erroneous! A good term is “long term almost undistorted average sine wave power into a resistive load” or just “average power”, although the latter form is a bit less informative.

The overall power conversion process is always inefficient by some factor and the losses are dissipated as heat. In highly inefficient circuits this will become a very big problem. The circuit's efficiency can be derived from a simple formula:

$$\text{Efficiency (\%)} = P_{OUTPUT} / P_{DRAWN} \times 10, \quad \text{(Equation 1)}$$

where P_{OUTPUT} = output power and P_{DRAWN} = power drawn from the mains. A rule of thumb for transistor class-AB circuits is to assume that they have an efficiency rating, which is always less than 74 % but higher than 50 %. Efficiency of a solid-state class-A amplifier is always less than 50 %, which is the theoretical maximum of class-A push-pull circuits. For tube amplifiers the rule of thumb ratings are even worse since a vast amount of power is wasted in the process of heating the filaments. Some digital amplifiers can have an efficiency rating that is higher than 80 %. However, this book will not discuss about digital or switching amplifiers.

2.2 The effect of output impedance and damping factor

An amplifier output can be considered similar to circuit shown in Figure 2.2. As can be observed, the output impedance is in series with the load impedance, thus the effect of output impedance will always attenuate the amplifier's output signal at some level. This is difficult to measure since the real life output node is the point in between the two impedances. In audio equipment the input impedance is usually several times higher than the output impedance of the preceding stage (impedance bridging). This is contrary to typical procedure of impedance matching, more common in video and RF systems.

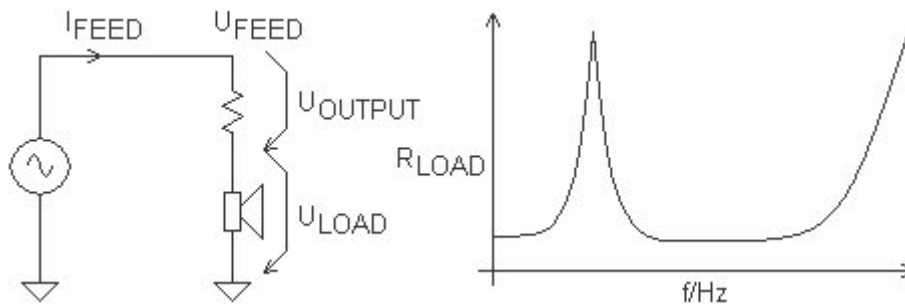


Figure 2.2 A conceptual schematic of power amplifier output.

Let's consider two amplifiers, one with an output impedance of 0.2 ohms, the other with an output impedance of 5 ohms. Both amplifiers drive a 4-ohm speaker load. Now, it would be easy to state which amplifier is better if the load resistance would be constant, however it is not: A typical speaker may easily present a load of 50 ohms at its resonant frequency. Also, at higher frequencies the speaker's impedance gradually rises due to inductance of the voice coil. Instead of constant 4 ohms, it is common that the load varies between 3 and about 50 ohms.

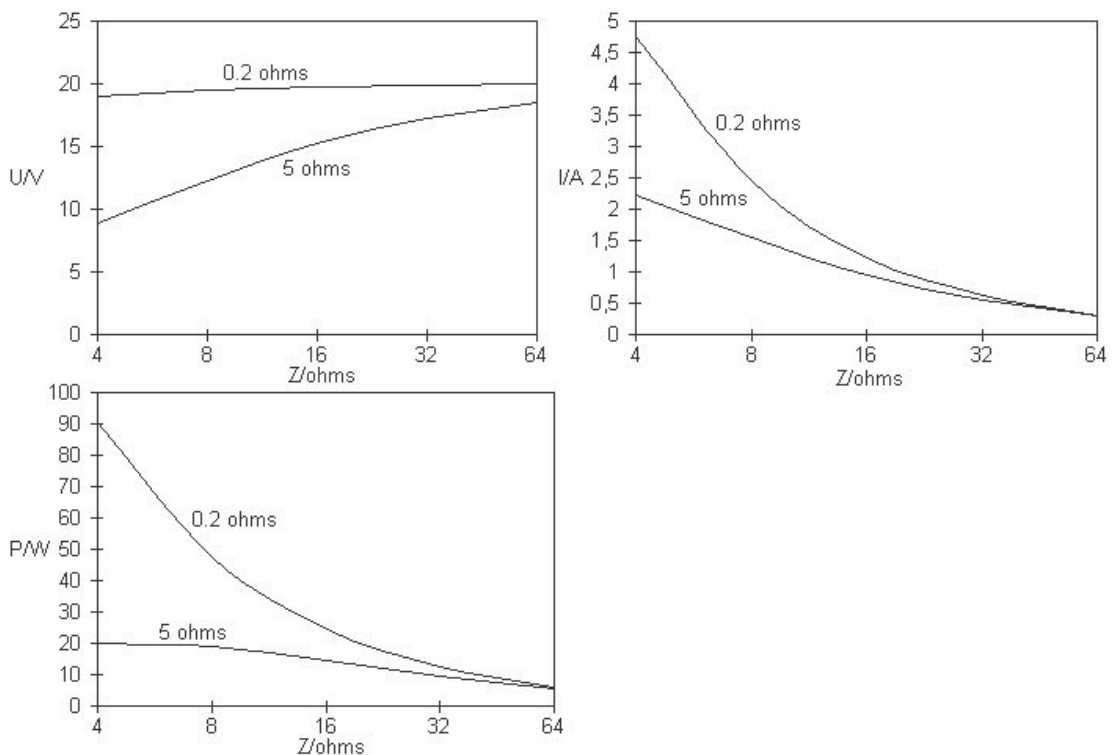


Figure 2.3 Voltage-, current- and power responses of example amplifiers with low and moderate output impedances.

Let's assume the amplifier has a constant voltage gain of 20 and is fed with $1 V_{PEAK}$ sine signal. Since the output impedance is in series with the load it effectively dictates the amplitude of the signal that is fed to the speaker load: An amplifier with low output impedance will have a quite linear *voltage* output regardless of the load impedance, while the amplifier with a higher output impedance will have quite logarithmic voltage output. With output *current* this is quite the opposite since high output impedance tends to balance the current draw. As a result, the power response of the amplifier with higher output impedance remains quite linear; the downside is the efficiency, or rather the lack of it: An amplifier with high output impedance is always less efficient than an amplifier with low output impedance, assuming all other aspects (such as current/voltage feeding capacity) remain equal. Figure 2.3 shows the corresponding graphs plotted.

The effect of low output impedance becomes more important when the amplifier's damping factor is considered. Damping factor is a variable that implies how effectively the amplifier can slow down the speaker cone movement caused by *inertia*. It can be calculated from a formula

$$DF = Z_{LOAD} / Z_{SOURCE} \quad \text{(Equation 2)}$$

Effectively, the impedance of the speaker cable in series with the load should be added to the formula as well. However, since this is usually an unknown factor (and quite insignificant as well) it is safe to ignore it. In HiFi design, damping factor below 20 is a sign of a "bad" amplifier. In guitar amplifier design, this is a highly subjective issue.

A vibrating speaker coil will generate a signal called back EMF (Electro Motive Force). Because back EMF is in opposite polarity with the speaker's motion it will attenuate it. Since damping factor varies with frequency the amplifier with low output impedance will provide a mentionable damping for higher load impedances, such as for a speaker in resonance. The effect of higher output impedance will definitively colour the sound since it "amplifies" the mechanical characteristics of the speaker itself. Some people regard the emphasized highs and enhanced bass introduced by free air resonance (typically around 60 – 120 Hz) as "warmth". However, due to lesser damping during resonance the amplifier may also suffer from "flabby" or "farting" bass tone and earlier cone break-up.

2.3 The basic building blocks of a power amplifier

A typical power amplifier usually consists of three parts, which are: 1) input stage, 2) voltage amplifier stage ("VAS") and 3) output stage ("OPS"). In simplest designs the first two parts are usually merged together. Such a circuit, consisting of three sub circuits, is typically referred to as "Lin topology" according to technical papers of the aforementioned author, which date back to the late 50's. Although old, the topology has remained nearly unbeatable and you can expect that about 98 % of all solid-state power amplifiers follow it.

A high gain voltage amplifier stage is the "heart" of most power amplifiers. Logically its job is to amplify the (usually) low amplitude input signal to a suitable level. Most VAS circuits work in class-A mode since they basically require only a small amount of current and therefore power losses over the active device can be retained

reasonably small. A basic VAS circuit found from 99 % of transistor power amplifiers is a simple common emitter amplifier.

The output stage is a current amplifier usually working in either class-A, class-B or class-AB mode. (A hefty sum of other output topologies - that usually require a complex circuitry switching additional transistors on and off due to loading - exists but for simplicity's sake they will not be presented here). The function of the output stage is to provide enough current gain so that voltage potential provided by VAS can exist over the low load impedance. (Remember that since $I = U / R$ and $P = UI$ both voltage and current are required to acquire the desired power). *The simplest current amplifier is an emitter follower.* Combining two complementary transistors the emitter followers can be connected in push-pull configuration where each transistor amplifies the current of its corresponding half wave. Such topology is known as class-B amplifier. Since in class-B amplifiers half of the supply voltage will not affect over the transistors during zero input signal (as is the case in class-A amplifiers) the efficiency of the circuit is fairly good. A typical configuration is to directly couple the bases of the output transistors to the collector of the VAS, thus the transistors do not require individual biasing.

Because transistor circuits have very low output impedance they can drive low impedance loads directly and no impedance matching in the form of a transformer is required. This is one of the main things that separate modern power amplifiers from typical tube circuits. Circuit topologies using output transformers can be realized with transistors as well - and as a matter of fact, the first transistor-based power amplifiers were copied from tube designs. Omitting the output transformer will, however, improve stability and reduce weight, cost and distortion.

A separate input stage is used for several functions; the most important of them is subtracting the (negative) feedback signal from the input signal – a procedure offering a notable amount of error correction. The other functions of this stage are to define the DC operating points and buffer the voltage amplifier stage. *The most common input stage topology today is undoubtedly the differential amplifier*, also known as a long-tailed pair or LTP in short. Before reasonably well-matched transistors became available at common markets (or other sophisticated matching methods had been developed) a complementary feedback pair (CFP) input stage was widely used instead of the LTP. These circuits will be discussed in detail later in this book.

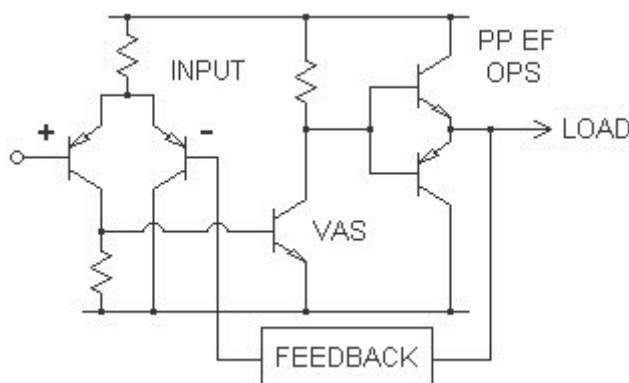


Figure 2.4 The basic power amplifier “Lin” topology with error-correcting differential input stage, common emitter voltage amplifier, push-pull emitter follower current amplifier and negative feedback path.

Transformer coupled topologies rely either on voltage-current conversion at the output stage or to interstage coupling between VAS and the output devices. In the first topology the transformer's primary side usually acts as collector load for the output devices, thus reflecting a "transformed" impedance of the speaker load. In the latter topology the primary of the transformer acts as a VAS load reflecting the impedance present at the bases of the output transistors.

The last but still very important "block" is the feedback network. Its task is to send the output signal to the VAS – at least in some form - thus it has an important part in error correcting as well as in bandwidth and gain limiting. Feedback can be either local, global or a mixture of both. Local feedback is utilized within a single gain stage while global feedback is usually fed across a multitude of stages. In modern power amplifiers typical feedback configurations include global feedback from output to VAS (or more often to the input stage preceding it), this is used mainly to limit the gain and set the DC operating points, and then there's some local "Miller compensation" feedback from the VAS transistor's collector to its base, this to limit the bandwidth, to improve stability and to improve linearity at higher frequencies.

Note: Named after John M. Miller, the "Miller effect" describes the effect where capacitance between input and output of a voltage amplifier (e.g. between base and collector of a transistor) is multiplied by a factor of $1 - A_v$, where A_v is the voltage gain of the corresponding stage. When A_v is negative (as in inverting amplifiers) the result from equation is greater than 1, thus capacitance is considerable increased.

2.4 Bi-Amping

In many conventional "full-range" amplifier configurations the amplified signal is divided into multiple frequency bands with the aid of a passive crossover circuit. The crossover feeds the speakers – each handling the reproduction of specific frequency bands. There are many deficiencies in this approach: First of all, the single amplifier must have quite high output power since it's singly responsible for all power amplification. Secondly, the passive crossover circuit introduces significant insertion losses, which has evidently more serious consequences at high power levels. Thirdly, the crossover presents a fairly complex load for the amplifier, which increases risk of unstable operation.

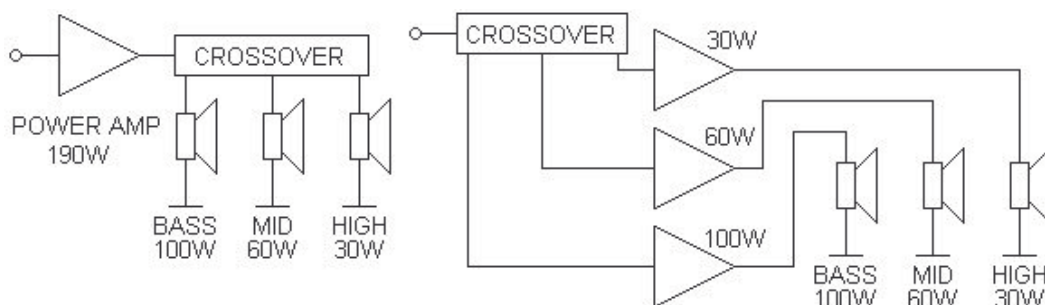


Figure 2.5 Amplifier concepts: Left: Conventional amplifier and right: Bi-Amp configuration.

"Bi-Amping" is a technique that strives to solve all these problems: The crossover circuit is placed before the power amplifier stage, thus the complex load that the crossover introduces to the amplifier is eliminated – along with the power losses caused by insertion loss. Practically, in such configuration, the crossover circuit can even be a part of the power amplifier circuit itself. As a downside, several individual amplifiers are required to handle the power amplification. However, this is not

necessarily as bad as it first seems: The output power requirement per each amplifier is significantly lower, which means the amplifiers are cheaper and easier to design.

A very common scenario in every amplifying configuration is that the huge demand of power during low frequency reproduction distorts the signal. Distortion at low frequencies is generally difficult to perceive but unfortunately in conventional configuration the condition (usually) distorts the signal throughout the bandwidth. Since in a Bi-Amp configuration individual amplifiers handle the amplifying of different bands the distortion in one of the amplifiers does not affect the behavior of others. This assures a clearer reproduction of the input signal.

Bi-Amping is still a fairly unusual concept for guitar amplifiers although earliest examples of using it date at least to early 50's. In acoustic (guitar) amplifiers – and some bass amplifiers – Bi-Amping is used quite frequently. One of the reasons for neglecting the Bi-Amp configuration is the narrow bandwidth on which the typical guitar amplifiers operate: The speakers cut heavily all frequencies that are above 4 – 6 kHz or below 60 – 120 Hz. Frequencies that do not belong inside this band are generally even considered unpleasant. However, the need for deep low end, that is typical for some modern music styles, has introduced the need for huge power reserve. In practice, there is no good reason why a guitar amplifier could not benefit from a Bi-Amping configuration that segregates the low end from middle and high frequencies.

2.5 Revealing performance

A good way to measure an amplifier's performance is to see how accurately it can produce a square wave signal. Square waves are rich in harmonics and thus provide an effective way to test both the amplitude response and phase shift. The test also reveals any overshoot, ringing or parasitic oscillations. A perfect amplifier would reproduce the square wave unaltered. In reality, no amplifier is capable of such performance.

Shown in section b) of Figure 2.6 is an example of distortion caused by too small slew rate. Slew rate represents the maximum rate of voltage (or current) change in a given time. To put it in common terms, slew rate is the speed of the amplifier and its unit is usually presented as V/ μ s, volts in microsecond. Slewing slows down signal changes thus it limits the effective bandwidth as the amplifier is too slow to appropriately respond to high frequency signal content. Since the slewing effect distorts the wave shape it also generates additional harmonic content. Slewing is one key reason why THD typically increases along with the frequency. For sinusoidal signals not to be subject to slew rate limitation the slew rate must exceed

$$2\pi \times f \times V_{PEAK} \tag{Equation 3}$$

where f = frequency and V_{PEAK} = peak value of waveform. For a high frequency amplifier a good slew rate is of course a mandatory. Yet, guitar amplifiers do not require an extraordinary bandwidth and higher slew rates usually go hand-in-hand with increased risk of instability. Thus there is no need for remarkable slewing rates in basic applications – on a contrary. However, an important point shown by the formula is that amplifiers with high output power *do* require higher slew rates due to generally higher requirement of voltage swing.

Plot c) shows other common forms of distortion: Overshoot and ringing. Both are caused by the internal capacitance of (output stage's emitter follower) transistors. In

most cases these symptoms manifest themselves during clipping condition or at the “crossover” region. Such conditions can also trigger continuous oscillation. Ways to fight these phenomena are discussed later.

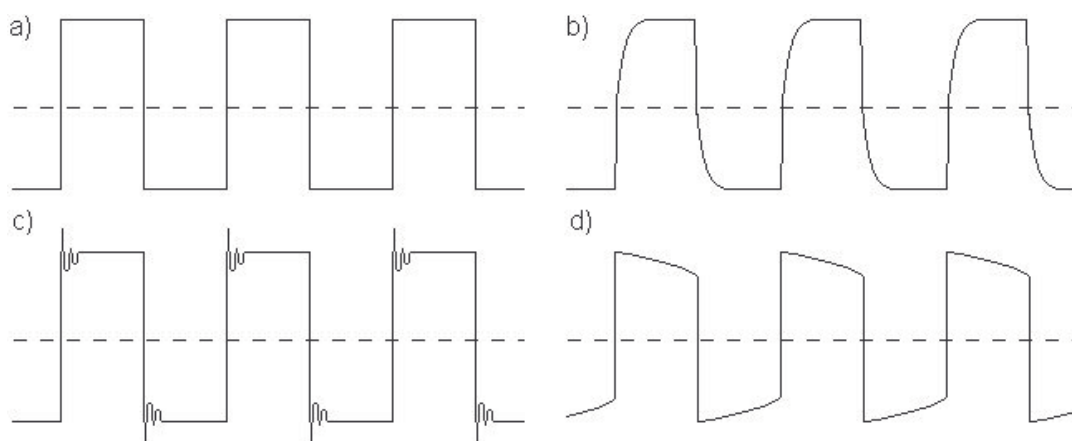


Figure 2.6 Common distortion types. a) Undistorted square wave, b) slew-induced distortion, c) settling with some parasitic oscillation (“ringing”) and d) hi-pass filtering or rail sag during clipping.

Hi-pass filtering reveals itself as amplitude distortion that gradually decreases the gain of “DC” components (that usually appear at the “top” of clipped waves). Another similar effect is created by supply voltage sag that lowers the clipping threshold. If these become severe enough the amplifier reproduces a square wave input as a saw tooth wave output. If coupled with the effect of inadequate slew rate the wave shape turns into a triangle. This is one of the aspects where “musicality” comes into the picture: Both slewing and hi-pass filtering round up square waves and “tidy” the harsh clipping common to solid-state circuitry. Unlike HiFi amplifiers, a guitar amplifier is not expected to provide an accurate reproduction of the input signal (unless that is specifically wanted) and here the designer can truly use his ears to provide a suitable compromise between accurate reproduction, bandwidth and “tone”.

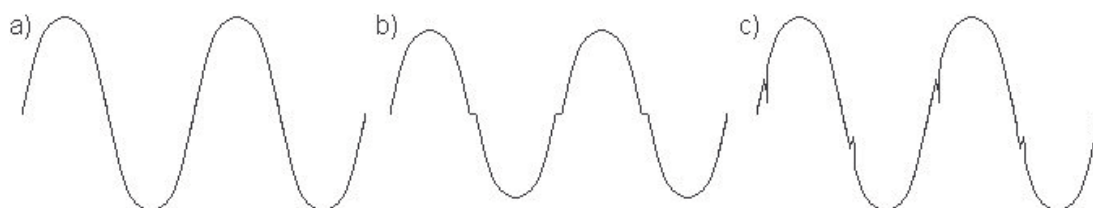


Figure 2.7 Crossover distortion. a) Undistorted sine wave and b) crossover distortion created by class-B topology and c) crossover distortion created by class-AB topology.

Crossover distortion is another common trouble of push-pull circuits that do not work in class-A. For transistor to exit cutoff region and begin to conduct, its driving signal must exceed a certain forward voltage. Due to this fact the gain of the circuit is reduced when the driving signal approaches its zero crossing points. The variation of gain will create a “kink” into the output. Since the kink is nearly constant in amplitude crossover distortion becomes more evident when output signal amplitude is reduced. A way to battle the crossover distortion is to extend the conduction time of the output devices: When done properly this will reduce the amount of crossover distortion so that it becomes *inaudible*. Insufficient (or too high) an amount of bias current (or voltage) will however leave little “ruptures” to the crossover regions. Since practical amplifier circuits run in class-AB instead of class-B the distortion type shown in c) is more common than type b).

Typical amplifiers with silicon transistors manifest the insufficient idle current with an annoying “buzz” during low frequency content and “metallic” high frequency tone. Crossover distortion is easy to recognize once you heard it once: This is about the nastiest distortion that a solid-state circuit can exhibit – especially since it, as opposed to clipping distortion, plagues more the “quieter” signal content. Naturally, the “harshness” is dependant on the transfer curve characteristics of the amplifying device. When compared to silicon, germanium transistors generally possess a “smoother” transfer characteristic (as does a vacuum tube) and therefore they are considered to give a more pleasant crossover distortion tone. Germanium also has a lower forward voltage. This does not mean that one should construct the output stage from unreliable, leaky, noisy, rare and expensive germanium transistors that on top of all this are also extremely sensitive to thermal effects and usually have poor power dissipation figures. I will state it again: Properly designed circuit will have so small amount of crossover distortion that it is inaudible.

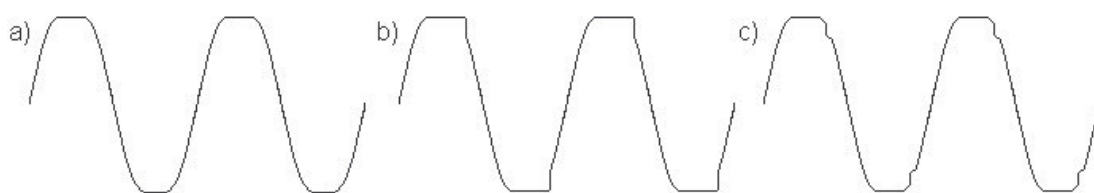


Figure 2.8 Clipping waveforms: a) Conventional clipping, b) “rail sticking” and c) “hysteresis”.

Last inherent form of amplitude distortion that is very tough to avoid is clipping. Clipping always introduces a wealth of additional harmonics and unfortunately it is in most cases far from the “pleasant” figure shown in a). A troublesome case of clipping is “rail sticking”: In this, the clipping condition is “extended”. Basically, every amplifier has a finite clipping recovery time but in badly designed circuits this time is so long that it becomes apparent. “Rail sticking” may introduce ripple to the output signal if the condition lasts for long. However, more troublesome is the full supply rail DC voltage signal fed to the speaker – this is destructive even in relatively short periods. There are examples of some extremely badly designed amplifiers that never recovered from clipping once it happened, thus they had a permanent and destructive “rail stick”.

Type c) is another common form of clipping: Here some chunks of the waveform have been “bitten away”. Very often the symptom manifests itself as clipping that has some delay or hysteresis either in the “beginning” or in the “end” of the clipping state. This introduces a wealth of high order harmonics that are very unpleasant. There are many ways how a circuit can act during clipping – most of them are dependent on the capacitive properties of the circuit and the output transistors. The load impedance has a huge effect on this behavior too.

3 A detailed examination of power amplifiers

3.1 Voltage amplifier stages

Voltage amplifier stage (in short VAS) has an important function of amplifying the low amplitude input signal to a level that (according to Ohm's law) can create a sufficient current flow through specific load impedance. HiFi amplifiers are typically tuned for -10 dB line level input voltage. This means nominal amplitude of 0.3162 V so the amplifier's gain is usually set to about 20 (26 dB). Most guitar power amplifiers tend to have even higher gain figures. The *open loop gain* of an amplifier (see 3.6) could virtually be as much as infinity but in reality it is pretty much limited by the stability of the circuitry itself. Since there is no need for gigantic amount of gain, the gain is limited to a practical amount with negative feedback, which also improves linearity. The amount of gain with the feedback loop enclosed is referred to as *closed loop gain*. Amplifier's headroom, on the other hand, is limited by the supply voltage- (and current) and therefore it usually is not worthwhile to have a closed loop gain figure much higher than

$$Headroom = (0,5 \times V_{SUPPLY}) / V_{INPUT} \tag{Equation 4}$$

unless a deliberate overdriving of the gain stage is sought after. I do not recommend serious overdriving of power amplifier's gain stages since it will - at worst - create a severe imbalance between supply rails and therefore introduce a significant DC component to the output signal. Note: In the formula V_{SUPPLY} means the voltage potential measured between the supply rails.

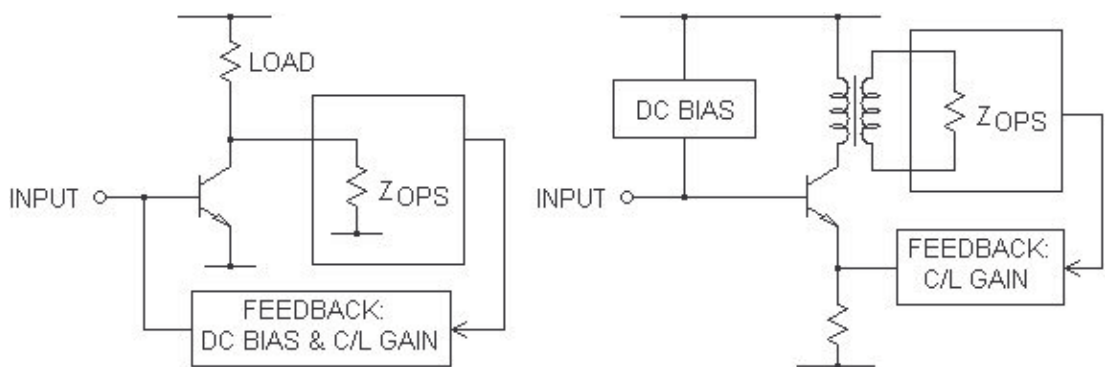


Figure 3.1 Typical voltage amplifier configurations. Left: Directly coupled common emitter and right: Interstage transformer coupled common emitter.

Figure 3.1 shows the most basic voltage amplifier configurations of which variations do exist but not with essential differences in the corresponding operation. The circuit shown on the left is undoubtedly the most common form of VAS nowadays: A common emitter circuit is directly coupled to the output stage, which usually is an emitter follower buffer. As a result, the VAS stage is (in theory) isolated from the output of the amplifier and loaded only by the collector load and impedance presented by the output stage's input. Since the output stage is an emitter follower the latter is (at least in theory) very high and thus rendered nearly insignificant. The DC levels are *feedback* biased and the amount of feedback from the OPS also defines the closed loop gain. Nowadays the feedback is usually fed to the input stage that precedes the VAS. Nevertheless, the basic principle of operation is pretty much identical.

The topology at right uses an interstage transformer to reflect the secondary impedance to the primary side, thus loading the common emitter circuit with it. In practice, this is not a very ideal solution: The transformer introduces large phase shifts that limit the amount of negative feedback that can be taken from the OPS. The results of this are higher output impedance and poor linearity because distortion created in OPS is poorly compensated. The transformer also creates a band pass filter and colors the signal. The reason why this topology was commonly utilized (in the 60's) was the fact that it was *at the time* the most practical *known* topology for coupling the output signal from VAS. (Back then output stages commonly used the totem pole configuration that consisted only of single polarity transistors). In some cases the poor performance of interstage coupling between VAS and OPS might be ideal for a guitar amplifier, though.

Virtually the VAS could be any gain inflicting circuit, such as an operational amplifier, but in most cases a simple common emitter circuit is proven to be all that is needed. One has to remember that the operation of this circuit is highly linearized by the use of negative feedback thus the amount of distortion originating from the VAS is very small – assuming the stage is working correctly, of course. Therefore the greatest differences in VAS gain stages, as can be expected, are not related to the circuit topology itself *but to way how the circuit is loaded*. Since modern designs are built on the basis of high open loop gain the most common way is to load the VAS with a constant current source load, which in theory can represent almost infinite impedance.

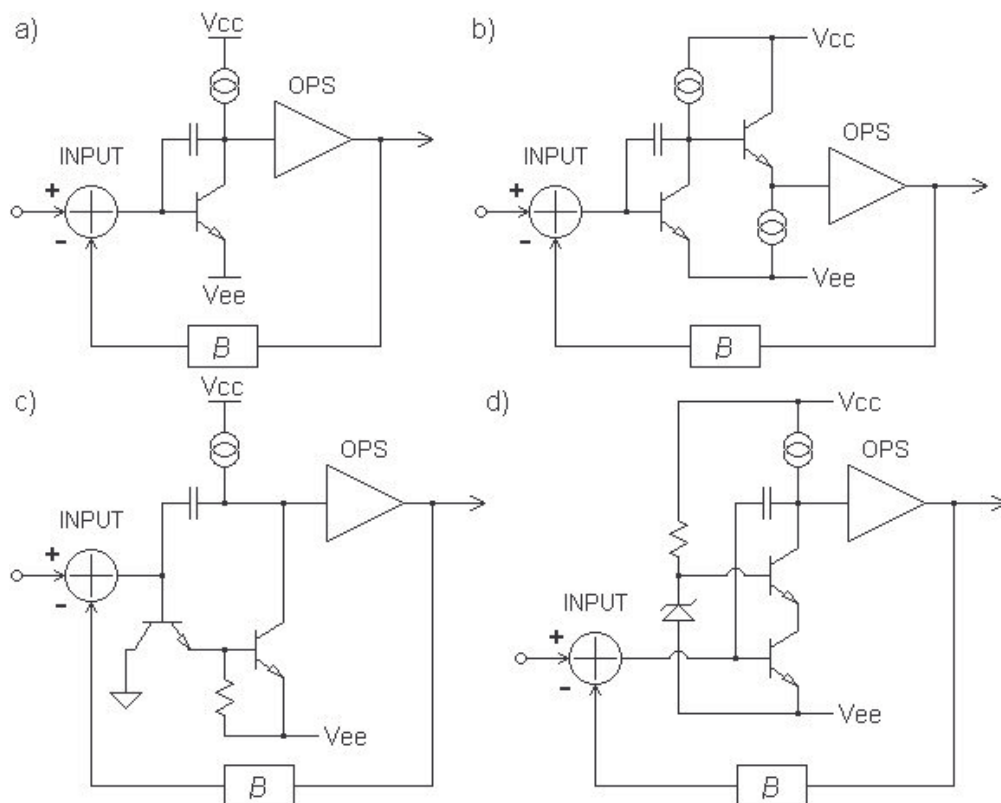


Figure 3.2 Most usual VAS configurations. a) Common emitter with current source load, b) buffered common emitter, c) Darlington (or “beta enhanced”) common emitter and d) cascode.

The Figure 3.2 presents few typical VAS topologies. Circuit a) is the one most commonly seen: A simple, Miller compensated and current-source-loaded emitter follower that is directly coupled to both individual input stage and OPS. Besides the local HF compensation, feedback is mostly controlled by the input stage, which we

will ignore for a moment. The constant current source load can be either active or passive (i.e. a bootstrap). This is indeed the basis of modern VAS topologies and as a rough rule of thumb can be therefore considered as the reference circuit for variations of the concerned topology.

Distortion in VAS is caused by the non-linear transfer characteristic of the common emitter amplifier. To effectively linearize these characteristics with negative feedback, VAS should always provide a high open loop gain. An attempt to increase gain by increasing collector load is futile: This merely decreases collector current causing a reduction in transconductance of the stage. The common approach has therefore been the use of constant current loading. This technique is discussed later. Another way to increase the gain is to increase the current gain (“beta”) of the transistor. Effectively, the simplest method to do this is to use a Darlington configuration (figure c). Note that adding another transistor to the circuit always introduces some phase shift therefore compromising the stability. The gain can also be increased with the use of cascode configuration (figure d) that raises collector impedance. Since the lower transistor is no longer exposed to full rail voltage the configuration allows the use of higher beta transistors, which usually tend to have smaller maximum V_{CE} ratings. In practice, the results of using either cascode or beta-enhanced configurations are similar [3]. However, due to its high output impedance, cascode performs quite poorly when it’s driving nonlinear impedances directly - and unfortunately an example of such is the input of a class-AB emitter follower output stage. In practice, the output of cascode VAS should always be buffered.

Which brings us to configuration b) that uses a buffer to isolate the nonlinear input impedance of the output stage from the VAS. As known, the base/input impedance of an emitter follower is mostly a product of current gain multiplied with the emitter load, therefore in a configuration that is not buffered, loading of the VAS – as well as its transconductance (“gain”) - changes in proportion to (speaker) load impedance, granted that this causes only a very minimal effect. If the VAS is buffered it is effectively loaded by a much more linear impedance of the buffer’s input. Also, we get rid of the effects that are caused by capacitance at the input of the output stage. This property is very useful especially when the output stage uses MOSFETs. Buffering the VAS also minimizes the effects of beta mismatches in output devices.

3.2 Output stages

3.2.1 Class-A

The simplest form of a current amplifier is an emitter follower, also known as buffer amplifier. This very basic class-A circuit can be constructed by directly coupling the base of the buffer transistor to the VAS. When current feedback is taken from the output node and fed to VAS the circuit will try to stabilize the DC operating points regardless of the loading – thus the circuit is essentially *feedback biased*. However, this topology is quite inefficient: During the positive half wave the transistor swings output voltage up but on the negative half wave an emitter resistor is required to pull voltage down. In order to obtain the maximum output power the resistance of the emitter resistor has to be equal or smaller than the load impedance. This causes a very high current to flow through it. Since voltage potential equal to half V_{CC} is constantly affecting over the resistor it has to dissipate a lot of power. Efficiency rating of such an amplifier is hardly ever higher than 10 %, which means that for 10 watts of output power the circuit has to draw 100 watts and dissipate 90 W!

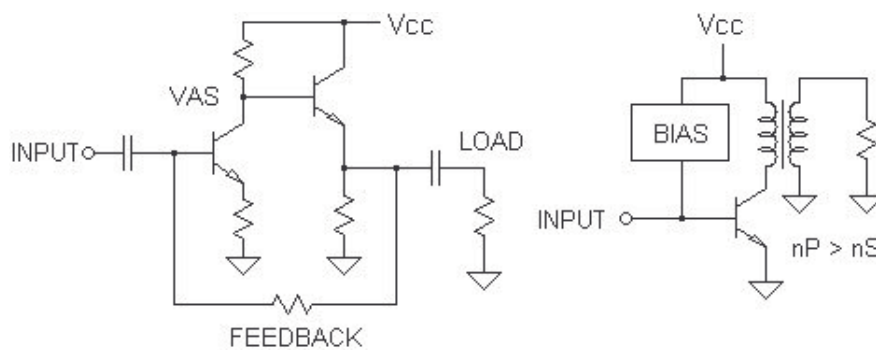


Figure 3.3 Class-A current amplifiers.

The other class-A single transistor amplifier relies on an output transformer performing voltage-current conversion. The transistor is connected in common emitter configuration so it acts as a voltage amplifier loaded by impedance reflected from the load at secondary side. The sole idea of the circuit is to convert a low-current high-voltage swing to high-current low-voltage swing – at the same time the transformer performs impedance matching. The problem is the high-voltage potential needed to create a reasonable amount of power in the secondary side. Up to the late 70's this topology was still commonly used in earphone amplifiers and transistor radios.

For the sake of their poor efficiency, further discussions of class-A power amplifier circuits are not included in any extensive examination. Using one as a power amplifier for an instrument would be mere stupidity in any other application than driving headphones or highly efficient speaker systems. Yes, I know that there are single-ended class-A guitar amplifiers but do you *really* want to tackle the ways to dissipate 45 watts of wasted power just to get an amplifier that has less output power than your ordinary portable radio? In applications with modest power demand, such as in spring reverb unit driving, a class-A circuit can work fine - but even then a class-AB amplifier can easily surpass it. Integrated low-power chip amplifiers, like LM386, or efficient operational amplifiers (e.g. NE5532 or LM833N) most likely will provide a more reliable and less-distorted performance coupled with a greater efficiency.

3.2.2 Complementary pairs and push-pull operation

To improve efficiency two emitter follower circuits were merged together whilst exploiting the possibilities of transistor bipolarity. The resulting circuit, presented in Figure 3.4, is called push-pull output stage, with a similar fashion to its tube counterpart. During positive input signals the NPN transistor pushes current to ground and during negative input signals the PNP transistor pulls current from ground. According to quiescent current of output devices the circuit works either in class-A or in class-B. If push-pull output stage is operated in class-A, the transistors never shut off and this way a maximum efficiency of 50 % can be reached. In class-B operation the theoretical efficiency maximum is 75 %.

Because in single supply amplifiers the output DC voltage potential between emitter resistors is half of the supply voltage, the (speaker) load has to be AC coupled. With a dual supply the output voltage is tied to a potential between the rails (which is commonly zero volts). However, even in such case the output's DC potential is zero volts only in theory. Nevertheless, coupling can be omitted if the DC offset voltage is reasonably small – meaning a maximum of few millivolts. Emitter resistors (also known as “ballast resistors”) are used to even out differences between internal emitter

resistances of transistors, thus their current sharing is improved. Ballast resistors are usually pretty low value, i.e. between 0.1 to 2 ohms. The resistors also improve thermal stability, prevent the complementary transistors from directly loading each other and dictate the amount of required bias current.

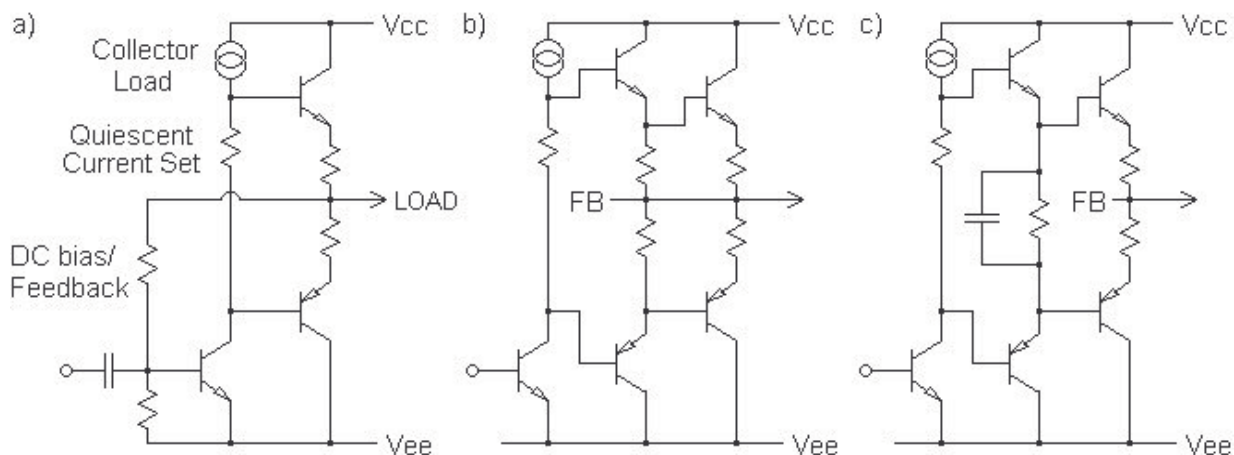


Figure 3.4 Complementary push-pull amplifiers.

Let's examine the typical topologies: The complementary circuit (a) is the basic topology behind all class-B, AB and C amplifiers. While it's admirably simple, it also has some efficiency issues: The driver transistor of VAS pulls down the voltage on the base of the PNP power transistor while the base of the NPN power transistor is pulled up by a resistor (collector load). The resistance together with the capacitance of the base-emitter junction is causing delay on transient signals, thus creating distortion. If the resistance is decreased, distortion will decrease as well but at the same time the power loss over the VAS transistor increases: Because almost the whole supply voltage potential may affect over the transistor even a low current will cause a power loss of several watts.

The Darlington complementary circuit (b and c) was developed as an answer. It uses two discrete transistors in a Darlington pair configuration to form one power transistor. Since the gain is higher (being product of gains of the individual devices), the first transistor in the pair requires less base drive current and the collector current of the voltage amplifier transistor can be retained lower. However, the introduction of another base-emitter junction means that the Darlington transistor requires a base-emitter voltage of around 1.3 V to turn on. This in turn causes a higher quiescent current demand to compensate crossover distortion. Some resistance (usually from 100 ohms to 1 kilo-ohm) between drivers' emitters is "required" to improve transient response and the way to employ this resistance is basically the only difference between circuits utilizing this topology. Typical configurations are the ones depicted in schematics b) and c).

In one of his books, Douglas Self explains that the configuration c) allows the bases of the output transistors to be reverse biased for faster turn-off and that the purpose of the capacitor is to increase the speed of switching [3]. The explanation is correct but unfortunately fails to highlight the importance of the capacitor: The reverse biasing effect that it introduces is the greatest factor in increasing the switching speed. The reverse biasing allows quick draining of the bases' charge. This is very important since it *prevents transistors from conducting simultaneously at higher frequencies* where higher currents are required for charging and discharging. If the aforementioned

cross conduction is too excessive the devices will dissipate a lot of power and fail (sooner or later). A typical value for the capacitor is about 1 μF .

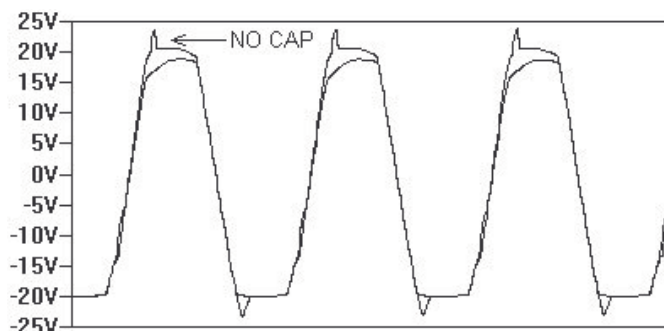


Figure 3.5 Hard clipped sine wave with and without the “speeding up” capacitor between emitters of driver transistors.

Cross conduction typically begins at frequencies that are above 50 kHz. This explains why many amplifiers can (or seem to) work adequately even without the “speed up” capacitor: Since losses increase in proportion to frequency the output signal must have a very high frequency before it becomes too “deadly” for the output stage. Occurrence of such signal is extremely rare, assuming the amplifier is well designed. Also, due to higher base capacitance, amplifiers with parallel output devices are more prone to cross conduction than amplifiers using only a single pair of output devices. However, one must acknowledge that small amount of cross conduction already takes place at lower frequencies and this creates unnecessary losses. For example, a 60-watt output stage with the capacitor may dissipate about 30 watts when pushing full output power at 50 kHz. Without the capacitor the losses may be as high as 60 watts – all that simply due to cross conduction. So, although both amplifiers would seem to work well, the one with the speed up capacitor usually runs significantly cooler.

It is also a common belief that the speed up capacitor improves recovery from clipping – this claim has been partially questioned by some people: For example, Marshall Leach even proposes that the capacitor is useless in dual-supply circuits [46]. It is remarkable that even though the Darlington complementary circuit is fairly old, theories concerning best “arrangements” for it are still for the most part contradictory. Concerning cross conduction, the configuration with the speed up capacitor seems to be the most *reliable* one but how about its linearity? Simple SPICE simulation revealed that linearity-wise the capacitor actually *degrades* the operation (in comparison to circuits b and c with the capacitor omitted) *when driven to clipping using a purely resistive load*. When the resistive load was changed to a more accurate simulation of a realistic speaker load the circuit with the capacitor began to clip in a distinctly different manner. However, none of the tests revealed any difference in *clipping recovery speed*. One must note that a real-life application and use of different semiconductors and component values will likely give completely different results. Switching speed is nevertheless proportional to driver’s emitter resistance and a lot of the non-linearity, caused by the output stage, *is removed by negative feedback*. (Note: FB in the schematic marks the feedback loop, which *is not* omitted).

At the time when Darlington complementary circuit was invented it was too expensive to implement into most commercial designs due to lack of decent PNP power transistors, therefore other topologies were favoured. Nowadays Darlington circuits are pretty common. Unfortunately, without sufficient quiescent current they tend to introduce a lot of high-order harmonic distortion, which makes them seemingly inferior in comparison to any other emitter follower output stage topology. A *correct*

bias, however, decreases the high-order harmonics of crossover distortion in conjunction with the signal levels [43]. The best property of the concerned topology is nevertheless its inherent stability: Due to low capacitance of the Darlington configuration the complementary emitter follower circuit is less vulnerable to parasitic oscillation (“ringing”) and “overshoots” than other presented output topologies.

3.2.3 Quasi complementary

The quasi-complementary circuit is the high-power successor of the symmetric complementary circuit. It is similarly based on the principle of using two transistors to form one power transistor with a high current gain but in this case the transistors conducting the highest currents share the same polarity. While the NPN pair is in a typical Darlington configuration (in which both transistors have the same polarity), the PNP transistor is not; instead it utilizes a topology called Sziklai pair, also known as compound transistor. In practice, the first transistor acts like a phase inverter. Current gains of compound transistors are similar to Darlington pairs but the advantage over them is that since there is only one base-emitter-junction a base-emitter voltage of only 0.65 V is sufficient for turning the transistor on. A schematic of the circuit is shown in Figure 3.6. Note the correct configuration of *emitter* resistors, which becomes important if several output transistors are used in parallel.

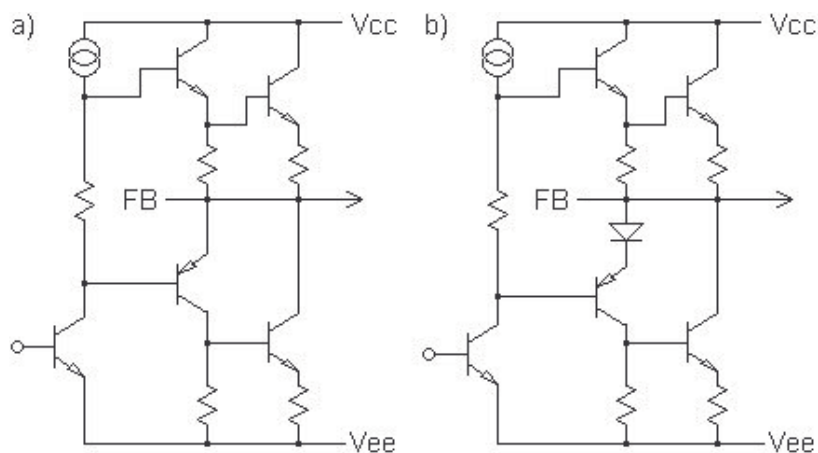


Figure 3.6 Quasi-complementary output topologies: a) standard and b) with Baxandall compensation.

While the quasi-complementary circuit provides a better matching between the power devices its operation near the crossover region is highly asymmetric. The reason for asymmetry lies in loading of the VAS: During positive excursions the VAS is loaded by two base-emitter junctions but during negative excursions only by one. This means that especially near the crossover point the gain factor becomes asymmetric. There is also difference in cut off characteristics of output devices that are essentially either emitter- or collector loaded [34]. The linearity can be improved by utilization of a Baxandall-diode, as in figure b) [3]. Essentially this adds up another diode junction to the lower transistor pair, thus rendering its operation closer to that of a Darlington circuit.

Parallel resistance and capacitance with the Baxandall diode can even further dampen the transients introduced by crossing over to positive half wave. Usually the parallel resistance will help with oscillation (ringing) problems that tend to occur when the amplifier is operating near its crossover or clipping regions. *The tendency to this oscillation is inherent for the concerned topology* and cannot be cured by increasing high frequency negative feedback in the feedback loop or VAS. A Zobel network

(discussed later) is usually helpless against it as well and using one is *essentially a bad solution* for removing any oscillation problems anyway. Since the parallel resistor works as a “snubber”, its resistance value strongly depends on the diode’s characteristics; a wrong value may even accentuate the oscillation. Typical values are 10 – 100 nanofarads for the capacitor and 50 – 200 ohms for the resistor. Since the main purpose of the combination of Baxandall diode and snubber is to *linearize* the amplifier it *is not* a foolproof method to dampen oscillation: The circuit mainly increases the threshold where the ringing starts and if the oscillation is persistent it is usually an indication that the collector resistance of the PNP driver transistor is too low.

3.2.4 Compound follower

The compound follower circuit is an all Sziklai design. The topology is also known as complementary-feedback circuit or compound pair. Some people claim that this topology has the best distortion figures, which is actually quite hard to believe: Although distortion of two base-emitter junctions with feedback is lower, the switching distortion, however, is notably higher – in similar fashion to Sziklai pair utilized in quasi-complementary circuit. The Sziklai circuits are also very prone to oscillation and the risk of it increases even more if additional power transistors are connected in parallel.

The respected – even overrated - status of compound follower circuit is likely explained by these facts: First of all, out of all possible driver-output transistor configurations, the circuit requires the least amount of bias voltage to compensate crossover distortion. Secondly, there is less need for an extremely precise bias setting than in a case of Darlington output (including quasi-complementary), thus the circuit can cope with natural variation of bias levels without introducing significant amounts of distortion. Since a compound pair circuit has only one V_{BE} drop that needs to be thermally tracked and compensated - *and that is the one of the driver transistor* - the requirements for thermal compensation scheme are less strict.

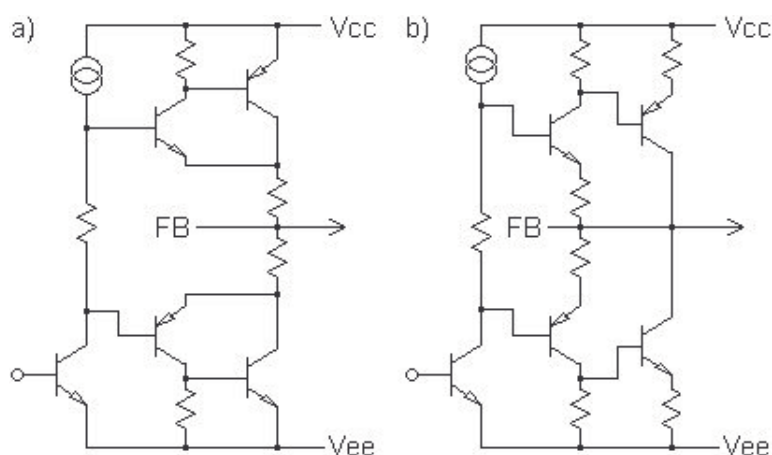


Figure 3.7 Typical compound follower output configurations.

Figure 3.7 shows two examples of compound follower circuits. In configurations that use parallel output devices, the scheme in figure b) should be preferred since it reduces the risk of oscillation and due to use of individual emitter resistors balances the current sharing of the output transistors. Unfortunately, if the collector of the output transistor does not connect the emitter of the driver before the emitter resistor (such as in figure a) thermal stability becomes dependant on the output devices. At

worse, the gain of the driver may even multiply the thermal effects. In both cases driver stage resistances should be kept as small as possible.

Today the concerned topology seems to be way too overrated: The compound followers have the poorest ability to swing close to rails when compared to rest of the emitter follower output topologies. A badly designed circuit will also have unimaginably terrible clipping characteristics; those usually involve “sticking to rails” quite heavily. Already mentioned was the inherently poorest stability out of all output stage configurations. While a properly designed compound follower circuit may perform well in a circuit simulator, in reality it is still extremely sensitive to layout issues. Practically, designing (and especially building) a stable compound follower output stage is very difficult. If the circuit uses many parallel output devices its realization may prove too difficult even for an expert. Since the output stage is one of the most critical sub circuits of a power amplifier this fact should not be underestimated.

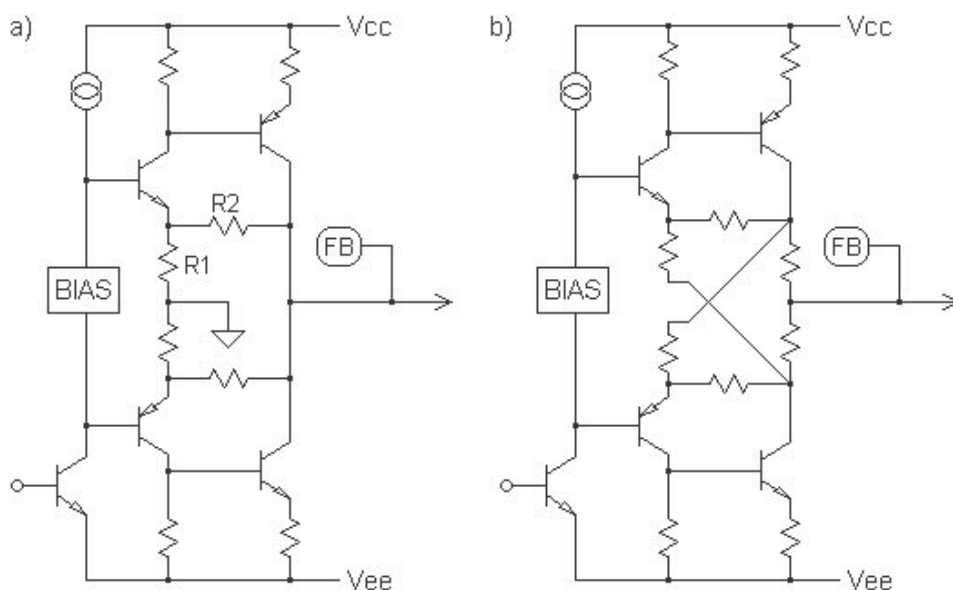


Figure 3.8 Compound follower modifications: a) Decreasing the amount of degenerative feedback introduces voltage gain. b) Error correction by current sampling.

A more unusual, although not uncommon, compound follower configuration is the one introducing voltage gain: As shown in the schematic, with the aid of shunting resistor R1, a resistor divider breaks the usual 100 % degenerative feedback path of the compound pair. Thus the gain is effectively set by the divider ratio $R2/R1+1$. The configuration has some intrinsic benefits: Since there is more gain available, the peak-to-peak drive voltage excursion of the VAS can be lowered, which often simplifies the design and reduces distortion. The circuit also has an inherent short circuit protection, since the driver's emitter resistor limits the amount of current that is delivered to output device's base. However, there are some considerable drawbacks as well: The more the amount of degenerative feedback is decreased, the closer matching between the output devices is required. One should also note that the local feedback of the output stage is reduced in proportion to increased gain. This makes the circuit much less linear than a conventional compound pair arrangement is. While people might live with the non-linearity (and even compensate it by increasing negative feedback) the most catastrophic results come from the circuit's inherent property to amplify its own thermal instability. Typical to Sziklai configurations, this circuit also suffers from the usual issues with high frequency oscillation.

Also moderately rare is the error-correcting configuration depicted in circuit b. In this arrangement the drivers sample the current that flows through the “opposite” output devices. When used properly, the technique decreases the amount of output stage distortion considerably.

3.2.5 Totem pole and topologies using output transformers

The totem pole circuit is the last output transformerless current amplifier topology. A “better” name for it would have undoubtedly been “an interstage transformer coupled push-pull emitter follower”. This very old-fashioned circuit is copied from a tube design; this time from an OTL circuit that was often used in driving high impedance cables. The example circuit uses a single supply, although use of dual supply was common as well. A very typical configuration was to power the class-AB buffer stage from dual rails (thus eliminating the need for AC coupling) and the class-A VAS from a single rail. A lot of variation exists in the class-A VAS circuit topologies – lot less in the circuitry of the buffer stage. The output stage may seem somewhat peculiar in comparison to modern circuits since it uses only single polarity transistors. Basically it still operates like a push-pull amplifier. An interstage transformer is used for both phase inversion and signal coupling. A network of voltage dividing resistors set the idle current for the output devices.

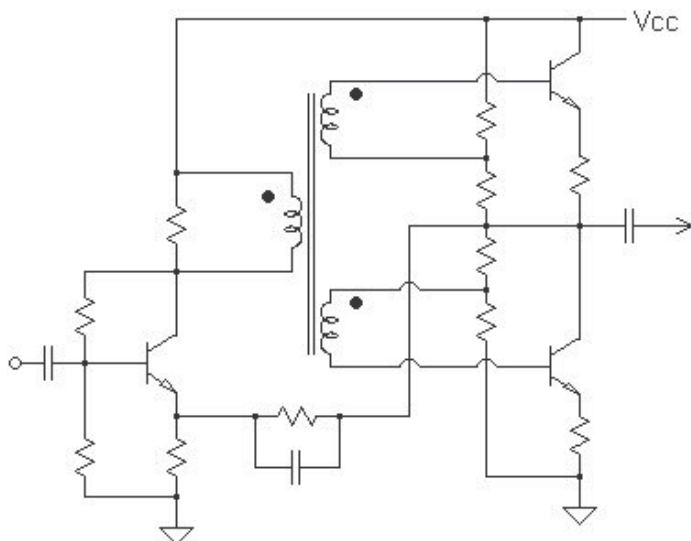


Figure 3.9 An example of the “totem pole” output stage topology.

At first glance the circuit seems like a combination of common emitter- and emitter-follower circuits, however, looks do deceive: Both circuits in fact work as emitter-followers. The basic operating principle is similar to that of quasi-complementary topology but the Sziklai pair configuration is substituted with a phase inverter formed by the inverted polarity of the secondary winding. The output of the “lower” transistor is referenced to output while the output of the “upper” transistor is referenced to ground. Viewed from the perspective of ground reference, the voltage swing of the upper transistor thus seems higher. However, both transistors actually have the same voltage swing – except that it has an opposite phase. The configuration effectively lowers the output impedance enough to avoid the requirement for impedance matching with an output transformer. Since the amount of negative feedback to VAS must be kept low due to phase shifts introduced by the transformer coupling the output impedance still remains fairly high, therefore the circuit retains the “warm, tube-like tone” as well. Furthermore, capacitively bypassing the “upper” bias voltage node to the output and “lower” bias voltage node to V_{EE}/GND can act as a bootstrap providing

even more gain. Instead of plain resistors, various temperature-sensing biasing schemes should be (and are) commonly used.

Today, this topology is very rare but it was widely used before the appearance of RCA papers that popularized the transformerless alternatives (ca. 1968). Typical drawbacks of totem pole topology are non-linearity created by the interstage transformer (which is furthermore worsened by the lack of sufficient NFB), cost, requirement of more chassis space, increased amount of wiring length and manual labor in assembling, weight and rarity of suitable transformers (interstage transformers require “customized” winding) plus the fact that short circuit protection schemes that limit the load current are quite difficult to implement. Considering the drawbacks it is easy to see why switching to transformerless designs was such a popular choice. Nevertheless, totem pole has some nice benefits: The bias conditions of VAS and output stage are (often) completely isolated from each other and the indirect coupling prevents the chain destruction of the entire power amplifier in case the output stage fails. It is also possible to push the VAS to saturate the interstage transformer, thus making it appear as a band pass filter – similarly as the output transformers in some tube amplifiers do. This will not necessarily happen in every design though.

Another method of utilizing transformers in output stages is using them to couple the output stage to the load. Today (and even at the beginning of transistor era) this topology was quite rare since it was not necessarily required. In solid-state guitar amplifiers both class-A and push-pull class-AB output transformer coupled topologies are highly rare – nevertheless, they *do* exist. Perhaps the most famous amplifier utilizing both output and interstage transformer is the Pignose 7-100. GMT introduced some output transformer coupled designs in the 70’s as well and their schematics are definitely worth of studying. Due to its inefficiency, this book will not discuss the single-ended class-A output stage. Transformer coupled amplifiers are typically common emitter voltage amplifiers instead of emitter follower current amplifiers. The output transformer reflects the load impedance from the secondary side to the primary, so that the primary can act as a collector load. Due to impedance transfer effect the load impedance seen by the collectors is naturally increased in proportion to transformer’s winding ratio. Basically, the voltage gain of the stage is turned into current gain by the transformer – similarly as in tube amplifiers. Due to circuit’s nature, the voltage gain varies widely in response to load impedance and the damping factor is very low. Linearity and damping can be improved by applying some negative feedback but the amount of it is naturally limited due to phase shifts introduced by the transformer.

Output transformer coupled transistor amplifiers perform like their tube counterparts: Short circuited loads, as well as very small load impedances, can be tolerated without explosions. Too high load impedance – or omitting the load completely – will destroy the amplifier. While this attribute is generally regarded as a tube characteristic it really is just a characteristic of the output transformer coupled circuits.

There are many variations of the topology, two of them shown in the schematic example. The circuit a) uses a split-load amplifier as a phase splitter that feeds the output devices, which are AC coupled from the phase splitter and biased separately. The circuit was very common in the transistor radios of the 50’s. The very same phase splitter circuit variations that are used in tube amplifiers generally work with transistor amplifiers as well and due to this there are numerous embodiments of transformer-coupled circuits. In circuit b) the phase splitter is omitted and replaced with an

interstage transformer that provides the phase inversion. The transformer's primary acts as the load for the preceding stage. The resistor divider network that connects to the center tap of the secondary handles biasing of the output devices. I have also included two other (not so common) tricks to this example circuit: Notice the negative feedback configuration that uses a separate tap of the primary. This is not a very common scenario, yet possible (the usual configuration is to take feedback from the "hot" node of the load). There is also a capacitor connected across the output transformer's primary: This is a high frequency shunt and sometimes used in conjunction with a series resistor.

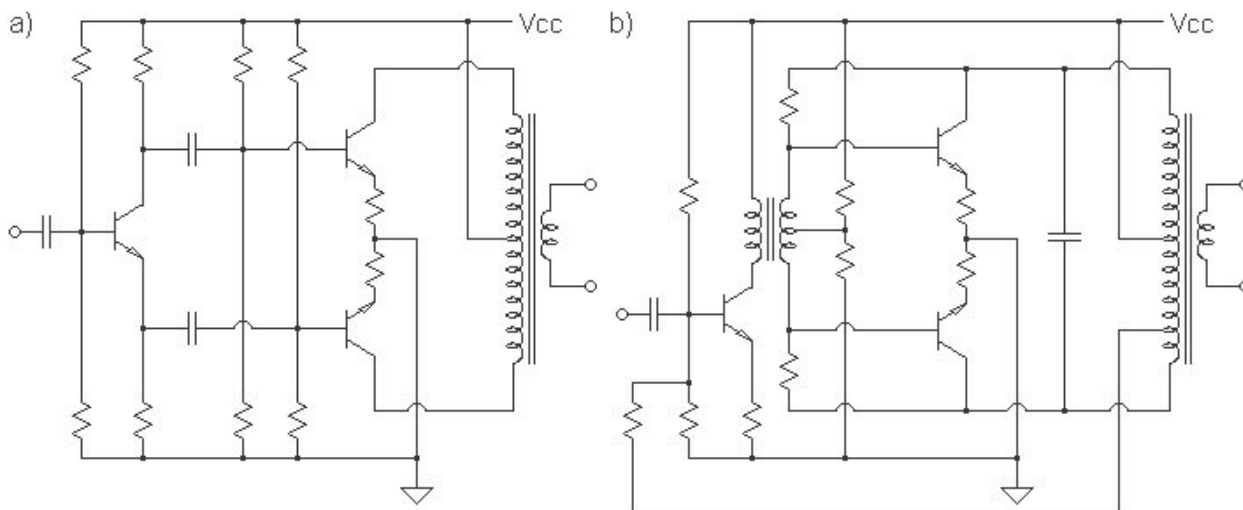


Figure 3.10 Variations of output transformer coupled series feed push-pull circuit.

Basically there are no issues that would prevent the use of high supply rail voltage in a similar manner as in tube amplifiers. High voltage transistors can be difficult to find, they may have lousy specs but nevertheless they do exist. If you are willing to see the trouble of acquiring all the required iron then finding proper transistors is likely not that big deal either. Circuits like this have been built and reputedly they perform and sound exactly like tube amplifiers. In my opinion, these are the only transistor circuits that can be used in any meaningful comparisons between the differences of tubes and transistors *as devices*. Comparing typical tube power amplifiers to typical transistor power amplifiers is like comparing apples and oranges.

3.3 Input stages

3.3.1 Single transistor

A single transistor input stage is essentially a voltage amplifier circuit getting direct negative feedback from the output node. In practice, the common emitter amplifier is biased so that its collector voltage (at idle) without signal is half of the supply voltage. Since theoretically the output node of the circuit has almost same voltage potential as the output node of VAS (which is the collector) the common emitter circuit is *feedback biased*. This also means that the feedback configuration is shunt mixing (feedback mixed with input signal) and voltage sampling - thus the circuit possesses characteristics of a *transresistance* amplifier that converts current into voltage. The equivalent circuits are presented right in Figure 3.11. The resemblance to a basic operational amplifier can be clearly seen from it. The amplifier is inverting and has the characteristics of inverting topology, such as high noise gain, low input impedance (usually defined by the input resistor) and narrow bandwidth.

With a few simple tricks the circuit can yield more open loop gain as well as control over feedback and closed loop gain. First, the emitter resistance of VAS transistor can be AC bypassed with a capacitor in order to effectively increase the gain of AC signals – a trick that is universal to all common emitter circuits. Secondly, the feedback can be controlled more precisely if the feedback path is separated in two distinct resistances for both AC and DC. AC bypassing a variable DC biasing resistor enables controlling the DC offset level without affecting the closed loop AC gain.

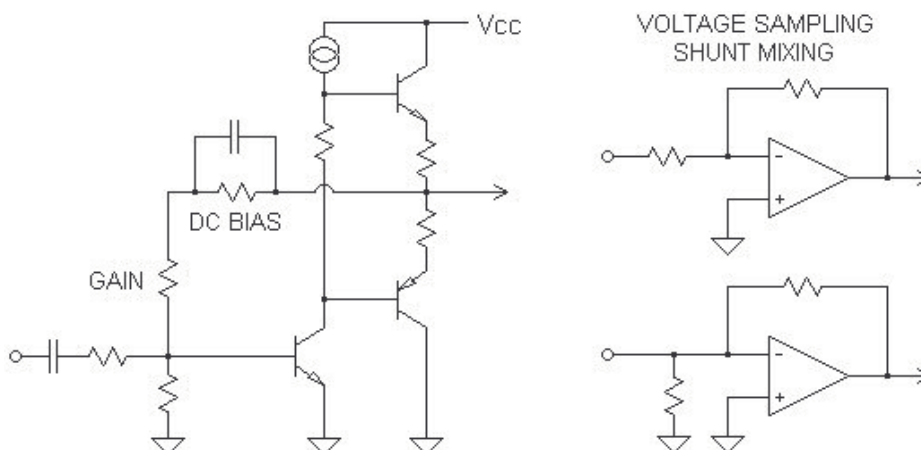


Figure 3.11 Combined VAS and input stages.

The additional resistance from base to common is not necessary in every case but with some feedback loop values the resistor is the only way to bias the stage correctly. It may however cause some problems: *From the input signal's point of view a volume potentiometer in the input, together with the base resistor, forms a voltage divider. A Logarithmic volume potentiometer set to half way has 9/10 of its total resistance between input and wiper and even more when the volume is set to around two or three! Unless the resistance of the potentiometer is notably (i.e. at least 10 times) lower than the base resistance, the volume adjustment will have even a steeper curve than a logarithmic one. To make matters worse, volume control's resistance is a part of the input resistor that defines the gain. Such lack of control is imprecise and useless. Basically this input stage configuration will always require a buffer (like another gain stage) to isolate the volume control potentiometer!* I have seen numerous DIY designs that overlooked this simple detail.

Since only VAS and output transistors exist between the feedback takeoff- and input nodes the phase shift caused by transistors remains very small. An ideal common emitter VAS would introduce a phase shift of 180° (inverted signal), however, a real-life common emitter amplifier will introduce larger phase shifts at higher frequencies. This usually starts happening after 20 kHz or so, depending on slewing rate of the active device. A phase shift above 240° at 1 MHz is a very realistic assumption from any transistor not specifically designed for HF applications - hardly any transistor with moderate power dissipation rating is. A phase shift in feedback will cause instability and in very worst case the negative feedback may turn positive if phase shift exceeds 360° .

As can be noted, the input of the circuit is referenced to the negative supply, which inherently causes major DC offset issues. If the negative reference is common the problem is naturally hidden by the requirement to AC couple the output. In practice, the circuit cannot be used with dual supplies, which means evident degradation of performance unless a high quality and high capacitance output capacitor is used.

Without extensive supply regulation the poor PSRR, common to all single supply amplifiers, will undoubtedly add ripple and noise to the output signal. Reasons mentioned above tend to make the amplifier bulky and expensive – factors that tend to overrule the circuit’s benefit of giving great stability when compared to other topologies.

3.3.2 Combined VAS and input

Effectively the single transistor VAS can be buffered similarly to left side circuit in Figure 3.12. The buffering doesn’t essentially change the operation of the circuit but it does isolate the VAS from the feedback, therefore the buffer stage can be considered as a separate input stage handling both error subtraction and output DC biasing. According to SPICE simulation, similar circuits (one buffered, one using a single VAS transistor) had no major difference in performance except that the emitter follower in the buffered circuit introduced 20 dB of more ripple to the output. The buffered VAS provides a little less phase shift but as this takes place after 300 kHz or so the real-life benefit is practically worthless.

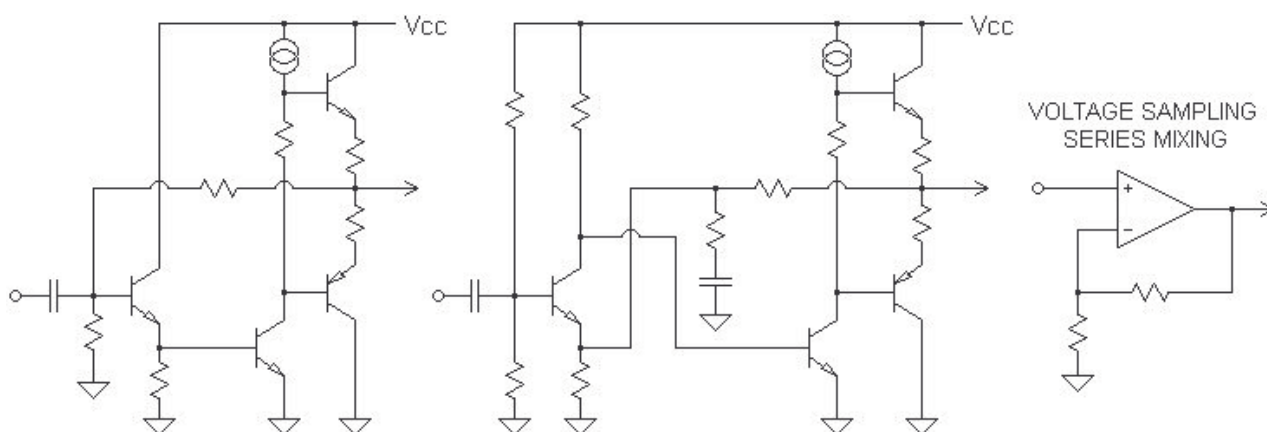


Figure 3.12 Multistage VAS circuits.

When the buffer is replaced with another common emitter circuit and the feedback insertion point is switched from the input transistor’s base to its emitter the operation starts to exhibit more drastic changes: First, the circuit introduces higher gain since it cascades two voltage amplifier stages, secondly, the feedback topology changes from shunt mixing into series mixing (feedback separated from the input signal), which essentially converts the circuit into a plain voltage amplifier. This topology was used in one of the first fully transistorized and OTL power amplifiers (Leak Stereo 30) and since then it has become the basis of modern amplifier topologies – albeit in another form. The topology is non-inverting, thus, when compared to inverting amplifiers, it provides less noise gain, broader bandwidth and higher input impedance.

3.3.3 Current feedback or “singleton”

The successor of the “Leak circuit”, the so-called “Singleton”, is perhaps the most confusing circuit amongst the common input topologies. Not many have a clear view how it actually works and, most importantly, what features it offers. The circuit is also known with another name; that is current feedback (or CFB) -amplifier. Basically, the circuit operates very much like a two-transistor gain stage in series feedback configuration. It is very similar to the one at right in Figure 3.12. However, the first NPN common emitter transistor has been replaced with a PNP one and the typical

biasing arrangement is converted into a circuit that uses the input transistor's emitter as a current sensing inverting input.

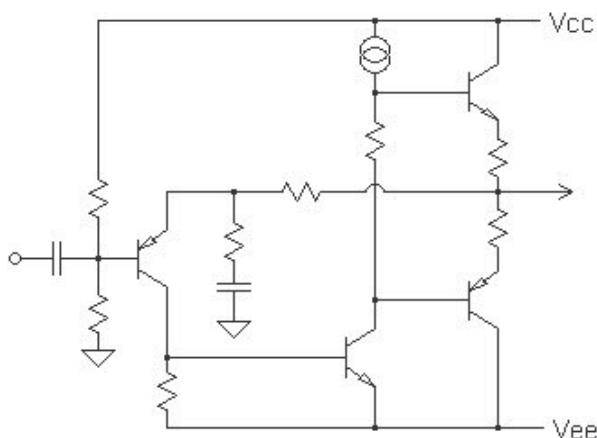


Figure 3.13 A power amplifier with a CFB input stage.

Conversion to this topology introduced some additional benefits: Since the input transistor's emitter is no longer tied to ground the resistor values in feedback path no longer have a limit that the Leak circuit exhibited: For common reason, the emitter resistance was forced to be lower than the collector resistance, thus feedback path resistors were forced to have quite low resistance as well. Consequently, the Leak circuit had an emitter resistance of 220 ohms, which consequently forced to use a resistance of only 10 ohms in shunting the feedback path to ground. For such low impedances the AC bypass capacitor (effectively setting DC gain to unity) had to have a capacitance as high as 1600 μF ! Even few years ago, such capacitors used to be quite large and expensive - during 50's and 60's they simply were gargantuan. For reference, in CFB circuits the corresponding values typically range from few hundred ohms to kilo-ohms for the resistor and from 100 μF to approximately 470 μF for the capacitor.

The CFB circuit can have an extremely broad bandwidth since it introduces very little phase shift at higher frequencies. When designed carefully, the bandwidth can even reach figures higher than 300 MHz! While a broad bandwidth is a totally insignificant feature in guitar amplification the inherent stability of the topology is not. The gain of the circuit is also very high - at least in theory. However, the error subtraction function is quite asymmetric thus the CFB input stage works quite poorly at correcting even harmonic distortion components. When compared to well-balanced long-tailed pair circuit (see 3.3.4), the output signal from CFB will always have a higher amount of second harmonic distortion.

CFB circuits also suffer from DC offset problems and usually require additional compensation circuitry to fight against them. The high DC offset is mostly due to large current affecting over the feedback resistor, which causes a notable voltage potential over it. Since the circuit is not balanced in any way the voltage potential will be amplified and it will wind up to the output. Correcting the DC offset of an amplifier that uses CFB input stage is however quite straightforward: The offset voltage simply has to be balanced by biasing the input transistor's base to a corresponding but *inverted* DC potential. In more demanding systems a separate DC servo circuit is utilized to compensate fluctuations in voltage potentials that cannot be tracked by passive circuitry alone. However, the need to DC bias the input transistor's base causes another shortcoming: The input is now strongly referenced to supply, which will impede the power supply rejection ratio thus making the circuit less

tolerant against power supply ripple. Therefore a large chain of power supply filters at input transistor's base biasing network is often found from old amplifiers utilizing this topology. DC biasing also forces to AC couple the input from preceding circuits.

3.3.4 Long-tailed pair

These days the long-tailed pair (LTP) - or differential circuit - is unavoidably the most common input topology. (Some people also refer to it as "Schmitt"-circuit). The circuit's nature of amplifying the difference of its input signals provides a couple of useful features: It basically cancels all even order distortion, rejects common mode disturbances and makes the amplifier quite immune to fluctuations of supply voltage. In practice, the particular input stage design suffers less from the effects of ripple voltage than other circuits. Another benefit is that the huge amount of DC feedback decreases the output's DC offset in dual supply amplifiers.

The circuit's operation is quite simple: The current at the emitter side of transistors remains nearly constant and is the sum of currents flowing through both transistors. (At least in theory) current divides equally between the two transistors and with equal base currents also the collector currents are equal. Unequal base voltages will cause collector currents to be unequal – the total current, however, does not change. The base of the other transistor gets feedback from the node connected to the load and the amount of it dictates the gain: Zero feedback would theoretically mean an "infinite" gain (large difference between collector currents). Feedback voltage equal to input signal would mean that the current divides equally between the two transistors thus resulting into a gain ratio of only one. The base current naturally controls the voltage amplifier transistor. The huge amount of DC feedback should stabilize the output's DC potential to the same reference potential to which the input is hooked via resistor (or set of them). The reference potential is of course set "in between" the supply rails so with a modern, dual rail supply it is zero. In practice, an ideal zero volt offset is impossible to acquire but it will, however, remain reasonably small.

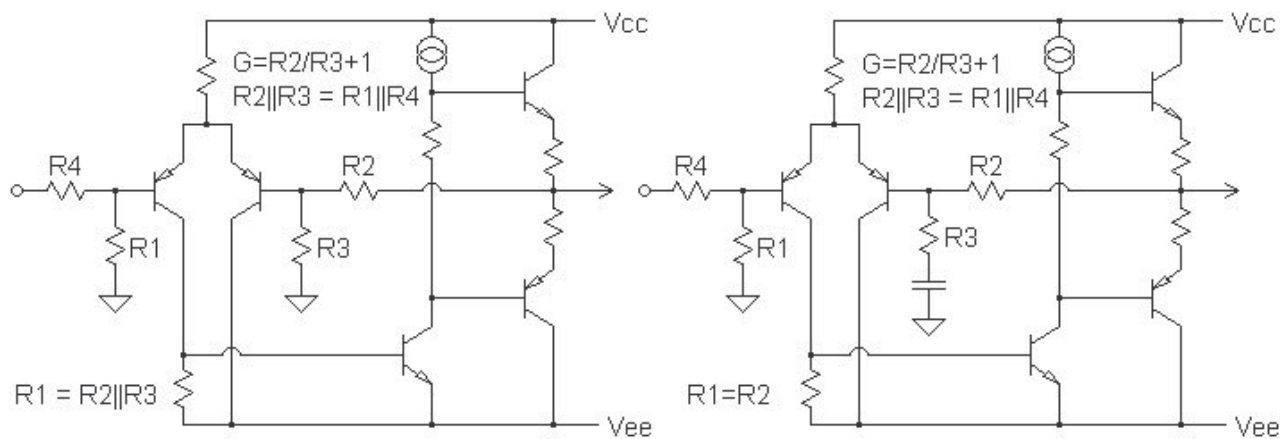


Figure 3.14 Amplifiers using a differential input stage. Left: with DC gain and right: with unity DC gain.

At first glance it may look like the circuit should have an inverted output and the inverting input would not be the negative feedback node. However, since the output is taken from the collector, the differential stage inverts the input signal (as happened in the "Leak circuit" as well) and after being inverted again by the voltage amplifier stage the output signal retains the same phase as the input signal. Since the inverting input (feedback node) is controlled by voltage the circuit is said to be voltage feedback-type (VFB).

Perhaps due to its enormous popularity the simple LTP input stage has become highly overrated. As a matter of fact, an *imbalanced* LTP circuit has nearly all the disadvantages of the CFB circuit, including the relatively high DC offset and limitations of correcting even order distortion. The distortion spectrum is almost identical to any other circuit presented so far and just about low enough to justify the use of an imbalanced LTP. On top of that, the circuit has the narrowest bandwidth and is most prone to oscillation due to large phase shifts at high frequencies. This is not as bad as it sounds: Unless the circuit is very poorly designed the amount of distortion will still stay relatively low – at least low enough not to be detected by most of the people. In guitar amplifiers, the limit of high frequency bandwidth is also quite insignificant concern since the frequency response of most guitar speakers drops heavily after about 5 kHz.

As presented in the schematic of Figure 3.14 the circuit is imbalanced - meaning a different amount of current flows through both transistors. While such thing almost completely ruins the idea of using a long-tailed-pair input stage in the first place, it is interesting to note that most LTP designs used in guitar amplifiers are, in fact, pretty much imbalanced. A balanced circuit becomes more complex but is also much more tolerant against minor component indifferences or temperature changes. For high-power circuits, (100 watts and up), such a tolerance basically becomes a mandatory in order to maintain stability.

As a small but important design detail, *to acquire lowest DC offset both differential inputs should have an equal DC resistance between input and ground*. DC is the keyword here: If feedback path's shunt to ground is capacitively coupled (refer to rightmost circuit in Figure 3.14) the equivalent DC path to ground is through the feedback resistor and load (and the latter is typically small enough to be considered negligible). Thus R2 must equal R1. However, if capacitive coupling of the shunt path is omitted (leftmost circuit) there are two DC paths: Through the shunt resistor R3 and through the feedback resistor R2 (and load). This is naturally interpreted as parallel connection. Thus R1 must equal the value of R2 and R3 in parallel. These are the first rules in minimizing DC offset and therefore all schemes that tamper with the concerned resistance values in real-time should be forgotten and buried. You *do not* want to mess with DC offset levels. The impedances of both inputs should also be kept equal, thus impedance of R4 in parallel with R1 should be equal to impedance of R2 in parallel with R3.

Although most amplifiers tend to utilize the non-inverting configuration of the differential circuit (and there are several practical reasons for this too) it should be noted that it is possible to use the same circuit arrangements as with any ordinary operational amplifier. This means that the differential can also work in *inverting* configuration. Utilizing the alternative design may prove beneficial in bridged circuits since it allows omitting an external phase inverter stage. Although the series resistor and capacitor (that shunt the inverting input to ground) may look like redundant components they – in most circuits – have an important function of keeping the circuit stable. Therefore it is not recommended to omit them.

Noise is always more or less a considerable aspect of design. In essence, any high-value resistor in the signal path will add its share to the noise, thus all impedances should be kept moderately low – especially in the NFB loop. Unfortunately this imposes another design challenge: Since retaining low DC offset requires that the input impedance of the NFB loop is kept equal to input impedance of the non-

inverting input (see Figure 3.14) we quickly run into a limit in the amount of reduction we can apply: In practice, the input impedance must be high enough not to cause degradation of the input signal.

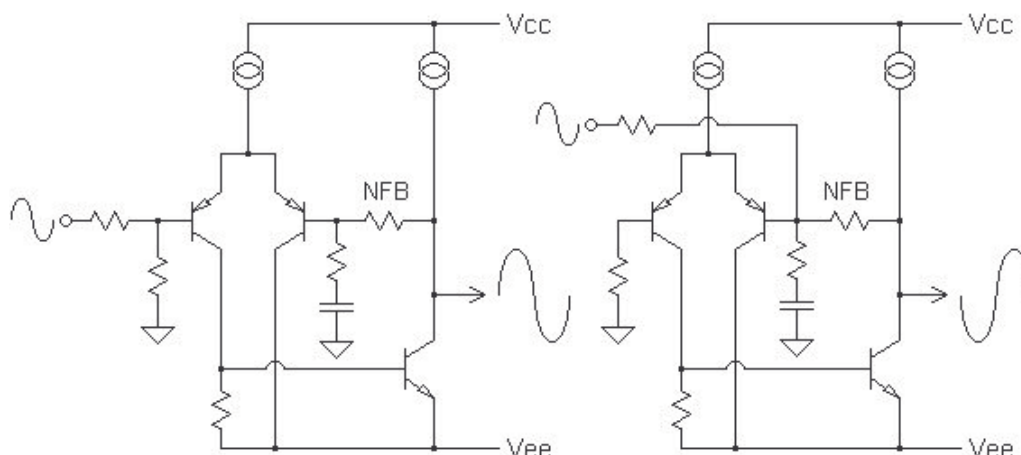


Figure 3.15 Differential amplifier in non-inverting (left) and inverting (right) configuration.

The convenient solution is to *bootstrap* the input. This creates high AC impedance while DC resistance is kept low and therefore DC offset level stays unaffected. In order to effectively bootstrap anything we must have a source that has (non-inverting) unity gain and low source impedance in respect to the point being bootstrapped. A source filling those clauses is naturally the inverting input. Figure 3.16 shows the basic bootstrapping configuration: R_{BOOT} and input resistor together form the resistance $R1$. Its value must naturally be that of the resistor $R2$. The junction of the two resistors is hooked to the inverting input through a capacitor. Resistor R_S isolates the feedback point from stray capacitance and thus enhances stability [3]. In practice, this resistance is kept low value.

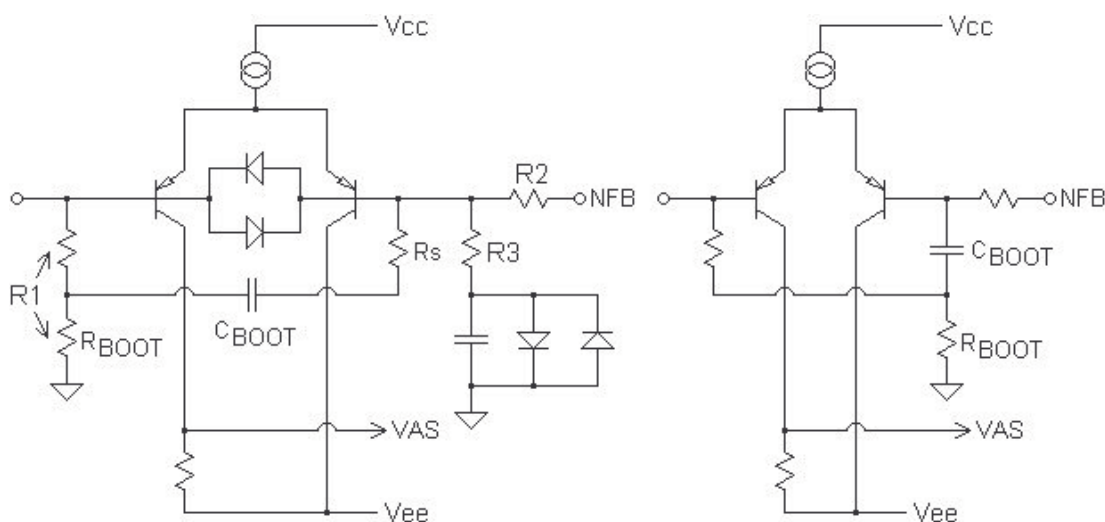


Figure 3.16 Tweaks with LTP: Input bootstrapping plus protections for feedback capacitor and differential. Alternative bootstrapping method is shown at right.

There is a price that one must pay: To compensate the decreased resistance of $R2$ and $R3$ the capacitance of feedback capacitor must be increased – usually to a fairly large value. To keep the capacitor within a reasonable physical size its voltage rating must be quite low. Consequently the voltage affecting over the capacitor must be limited to a safe value, for example, by using a simple parallel diode shunt configuration. This prevents unpleasant explosions should the amplifier fail and saturate to either rail potential.

Shown at right in Figure 3.16 is another bootstrapping method. In practice this is a compromise compared to the configuration described earlier but saves one capacitor and resistor (if R_S is excluded). The inconvenient side effect of the circuit is that it creates an impedance peak instead of fairly flat impedance. The peak can be somewhat “damped” by using high-value feedback capacitor but it is still present. Thus the circuit may react rather inelegantly with sources that have non-linear output impedance or that are capacitive by nature.

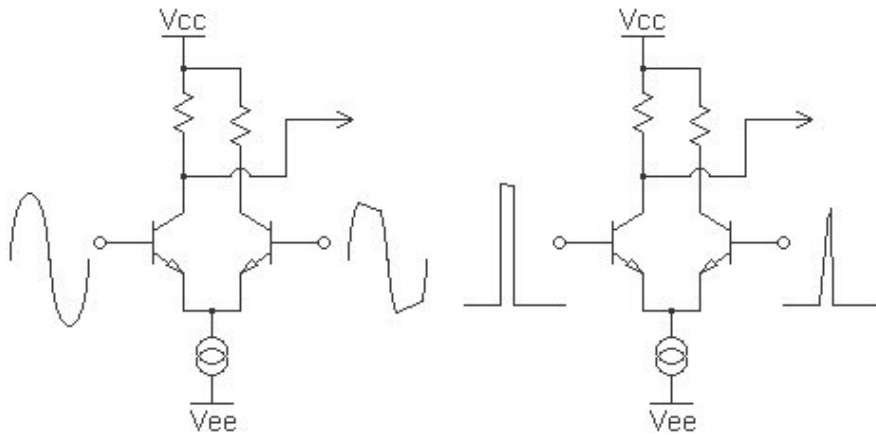


Figure 3.17 Common causes for failure of retaining inputs in balance. Left: Clipping and right: Propagation delay of the amplifier.

The diodes hooking inverting and non-inverting inputs together protect the transistors from reverse-bias. Two common conditions that cause reverse biasing are: 1) too high input amplitude drives the amplifier into severe clipping and 2) the output cannot respond quickly enough to a change that takes place in the input. Under normal operating condition the inverting and non-inverting inputs have nearly equal voltage potential. However, when the amplifier is driven into clipping the feedback can no longer follow the input signal. Consequently, the input signals of the differential are no longer identical and the circuit becomes susceptible to reverse biasing. The amplifier’s inherent propagation delay creates a similar condition although the process is quite different. With the aid of diodes the voltage difference between the inputs is limited to about 0.65 V, which is safely below the base-emitter breakdown voltage of most devices.

3.4 Techniques to improve linearity

3.4.1 Balancing the long-tailed pair

In order to acquire any real benefits from using a LTP input, the currents flowing through the differential transistors must be equalized. Sometimes an attempt to “balance” the stage is made by connecting a resistor from the collector of another transistor to V_{EE} . Although this can certainly ease the PCB layout it may also significantly worsen the current balance.

A lot more effective way is to use a *current mirror*, which is a form of constant current source circuitry. In an *ideal case* it consists of a high- H_{FE} differential dual transistor, meaning the parameters of both transistors are closely matched and the thermal coupling between them is ideal since the transistors share the same package. The current mirror circuit is depicted in Figure 3.18. Note that the transistor at left is essentially just a diode (base-emitter junction). The voltage at the base of the

rightmost transistor must equal rail voltage minus voltage drop over the BE-diode junction. Since the same voltage drops to zero over the “current setting” resistor, the current I_{R1} can be easily calculated using Ohm’s law. Because voltage drop over both BE-junctions is equal, the emitter currents must also be equal, the collector current is thereby mirrored – hence the name of the circuit.

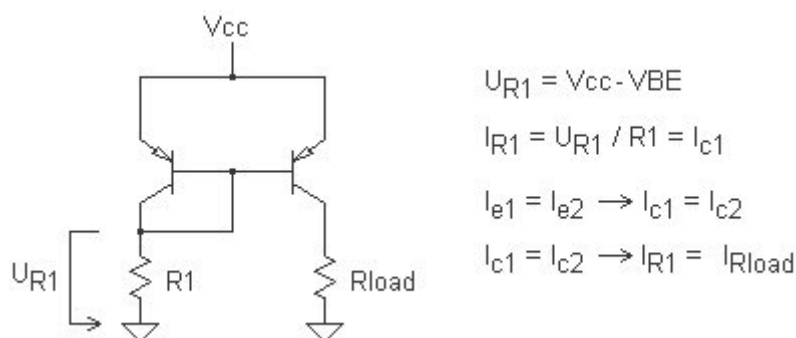


Figure 3.18 A current mirroring constant current source.

The collector load of the LTP input stage in Figure 3.19 uses the same circuit but in reversed polarity – pushing current to ground – therefore when using proper terminology it should be called a current *sink* instead of *source*. The main difference in a practical application is that (instead of a constant resistance reference) the impedance of the input stage’s other half defines the amount of current flow. Note that in the schematic example some of the resistors are replaced by current sources. This is justified since in 99 % of modern solid-state power amplifier circuits at least the VAS is loaded by a CCS that provides an extremely high open loop gain. Since the presented circuits are beginning to approach HiFi-class anyway, learning to consider constant current sources as more efficient and linear loading technique is advisable.

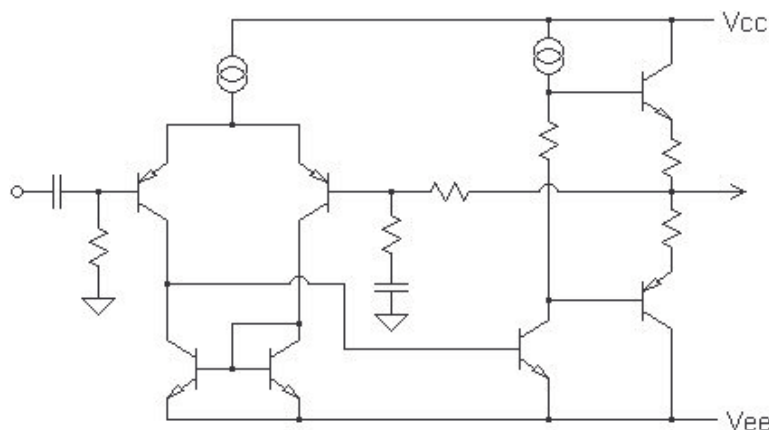


Figure 3.19 Differential input stage balanced with a current mirror.

In real-life, the basic current mirror circuit hardly works as ideally as described and there are a few issues that likely affect the performance: Both transistors require base current and therefore the “mirror” current will actually be adjusted current *minus* the required base current. Buffering the base of the mirror transistor, like in circuit a) of Figure 3.20, will lower the current due to another voltage drop over the base-emitter junction but the current will also be mirrored much more precisely. Another - and bigger - issue is the loading resistance: If it’s too high the other transistor will likely saturate, being unable to provide enough current. This can happen the opposite way as well: Too small loading will draw a higher current than intended. Optimally the load and the resistance setting the current should be equal, which is often impossible to accomplish in practice.

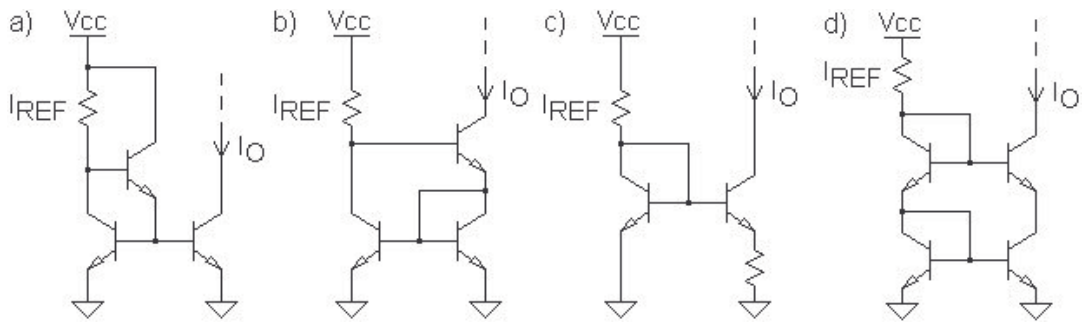


Figure 3.20 Current mirror variations: a) Base current compensation or buffered Widlar's source (or sink), b) Wilson's mirror, c) Widlar's source (or sink) and d) cascode.

The fluctuation of rail voltages during loading causes problems as well since it always affects the current adjustment. In the worst case the collector current may drift as much as 25 %. This phenomenon that is known as “Early-effect” can be cancelled by using emitter resistors that have a resistance that is high enough to create a voltage drop that masks the effects of the voltage drop over base-emitter junction. Usually a voltage drop at least five times V_{BE} is adequate. These drawbacks make current mirror circuit unusable in almost any other application except as LTP circuit's current sink/source.

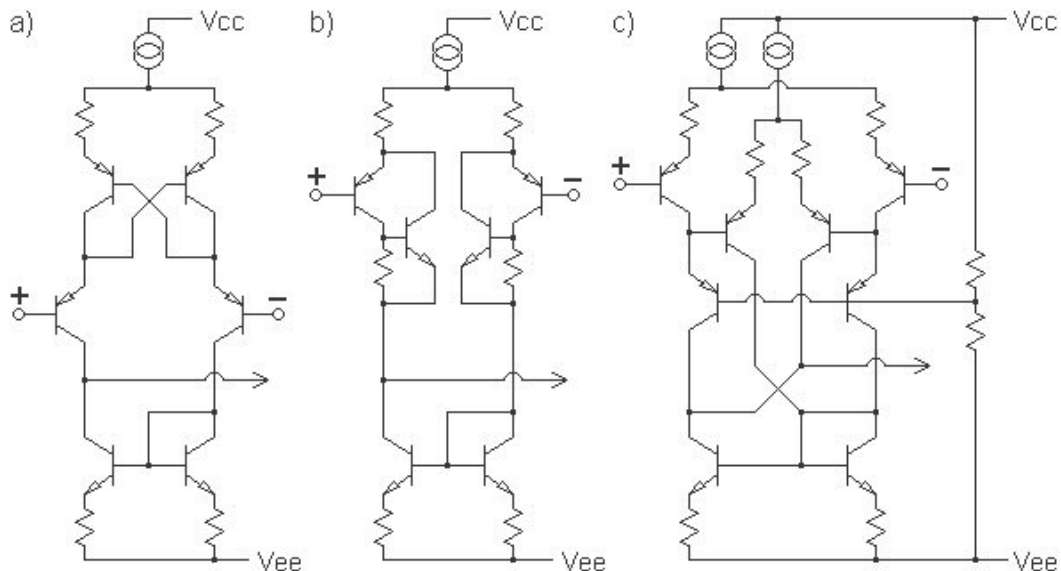


Figure 3.21 Improved input stage balancing methods: a) Cross-quad input, b) CFP input and c) Cascomp input.

Current mirroring is not the only way to improve an input stage's linearity. One technique is, for example, replacing the input transistors with Sziklai pairs (circuit b), which consequently increases the local feedback factor [3]. Other techniques include using a transconductance (“transconductor”) stage in forms such as “Cross-quad” (figure a) or “Cascomp” (figure c). Due to their complexity they have more commonly been used in integrated circuits. The Cross-quad circuit relies on correcting the error caused by non-linear emitter resistance by dividing it with the transistor's beta value. The Cascomp circuit delivers current through cascode transistors that duplicate the input's V_{BE} errors at their emitters. The error is sensed by another differential amplifier and summed to the output signal in reverse phase for correction [3].

An important balancing procedure, which is far too often overlooked, is *thermally coupling the differential transistors*. Ultimately the differential transistors should

share the same package, which would ensure both extraordinary thermal coupling as well as matching of other device parameters. However, such packages are usually quite expensive and limited in selection, therefore differential transistors are often discrete. Being so, they are also prone to operating in different temperatures that (as we all know) results into different collector currents and therefore imbalance. Thermal mismatches become extremely significant when the VAS uses differential topology as well. Common thermal coupling method is to mount transistors so that their flat backs face each other and have some thermal paste in between. Sometimes transistors are bound together with a metal or rubber wrap. This requirement for close proximity of the transistors is an issue one needs to keep in mind when designing board layout. Covering the differential transistors with epoxy further improves the thermal coupling.

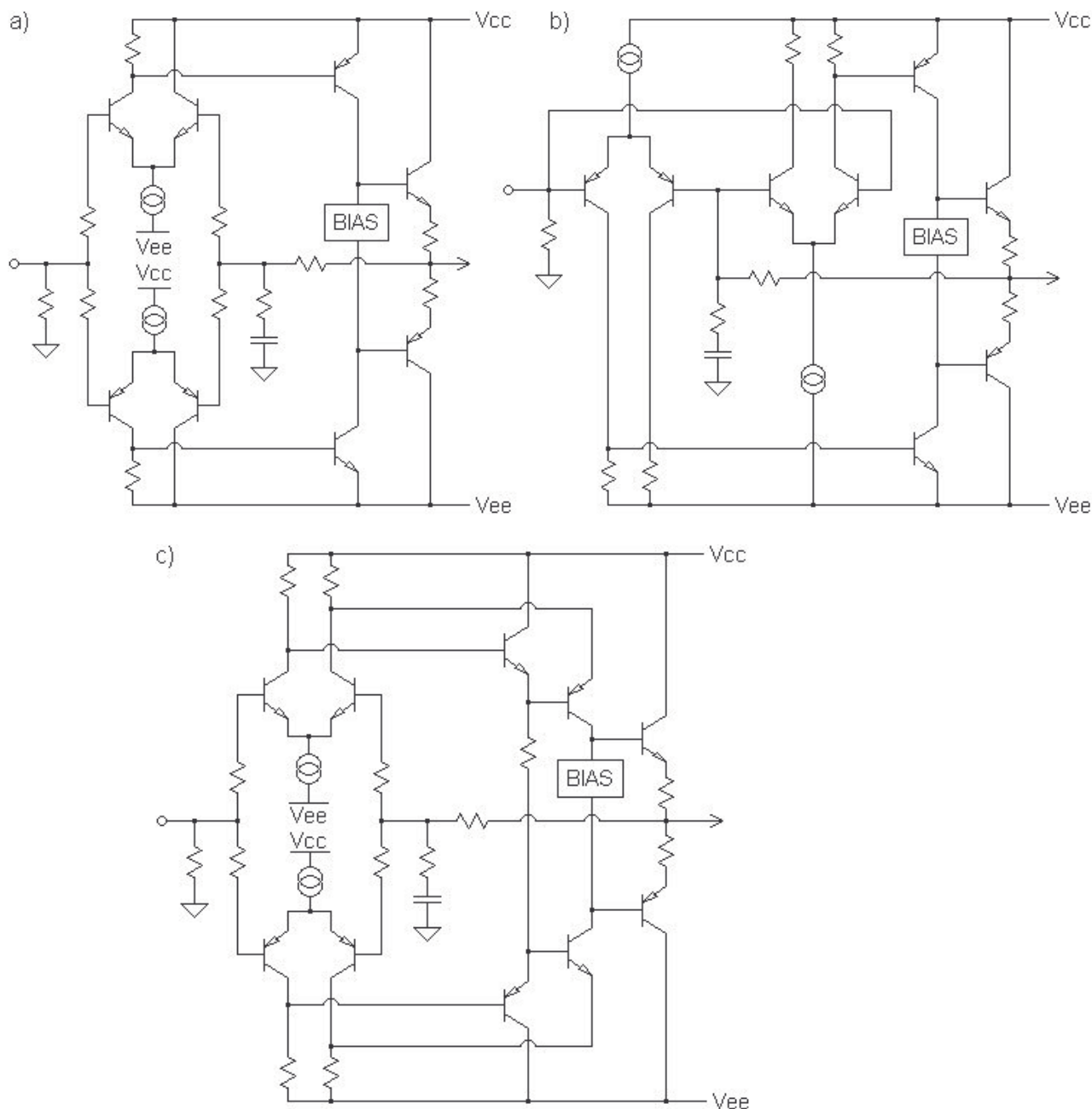
3.4.2 Symmetric designs: Cascoded complementary LTP

The amplifier's symmetry can be improved by cascoding complementary input- and VAS stages. The idea is to use parallel VAS stages that would linearize each other's transfer curves. Effectively, any external VAS loading can be removed as the complementary transistors take turns in loading each other; this allows swinging closer to supply rails and reduces the "rail sticking" during clipping -phenomenon that is somewhat inherent for CCS loaded voltage amplifiers. Simplified schematic examples of cascoded, "mirror image", Lin circuits are shown in Figures 3.22. The circuit in figure c) is an example of "Borbely" topology, named after an Erno Borbely design introduced in Wireless World magazine in 1983. The design utilizes a directly coupled buffer before VAS, which in turn takes "local" negative feedback from the input stage thus improving linearity.

The schematics are very rough examples: Cascoding the LTP *forces* one to use designs that provide excellent matching within the differential circuits (i.e. matched transistors plus use of current mirrors and/or CCS loads in input stage); *otherwise the circuit's performance is even worse than that of an ordinary imbalanced amplifier.* According to SPICE simulation, without current mirrors in long-tailed pairs (or any other way to properly balance the differential currents) the cascoded circuit produced a greatly distorted output and had a lower gain as well as an asymmetric transfer curve. There's also another problem: A single feedback loop is inefficient in compensating the DC offsets of *two* individual input stages. With the increased gain, introduced by current mirroring, this problem becomes even more severe.

By "tailing" the input stages with constant current sources the circuit's performance becomes comparable with an ordinary LTP circuit: A SPICE simulation showed that the cascoded design had a slightly lower DC offset, slightly higher gain and *only* 2 dB less of second harmonic distortion. Due to greater linearity the cascoded circuit starts to introduce more even order distortion when clipping to supply rails. The simulation actually revealed that the only real benefit gained from using a cascoded topology was a better stability that allowed extending the bandwidth. Then again – extension above 100 kHz hardly makes any difference in audio applications and even less when amplifying guitar signals. As is, the differences between ordinary and cascoded LTP are hardly distinguishable by an average listener. Let it be reminded that the simulated circuits also were "ideal" thus they did not suffer from real-life component indifferences that might cause a notable degradation of performance – assumably the effect of this would be *greater in the cascoded circuit* since it heavily relies on component matching. The cascode lacking the current sharing resistors (circuit b) seemed to be worse than a) and exhibited instability in the form of "ringing": Clipping

triggered short-term oscillation bursts if the amount of high frequency NFB was equal to that of circuit a).



Figures 3.22 a, b and c. Cascoded LTP circuits.

Surprisingly, the praised “Borbely” circuit proved to be even worse than the conventional cascoded circuit: The two performed almost identically, except that the Borbely-type circuit had a higher amount of ripple and DC offset voltage in the output - as well as a notably higher quiescent current. It was astounding to find out that this circuit didn’t seem to perform any better than an ordinary LTP circuit either. A deliberate test of *misbalancing* one of the long-tailed pairs by using dissimilar transistors finally revealed some benefits: The typical cascoded circuit began to exhibit quite nasty “ringing” that created a spectrum of unpleasant harmonics but the Borbely design did not suffer from this phenomenon. The asymmetric circuit did not “ring” but it naturally suffered from a *severe* DC offset. The schematics of all simulated circuits are shown in Figure 3.23.

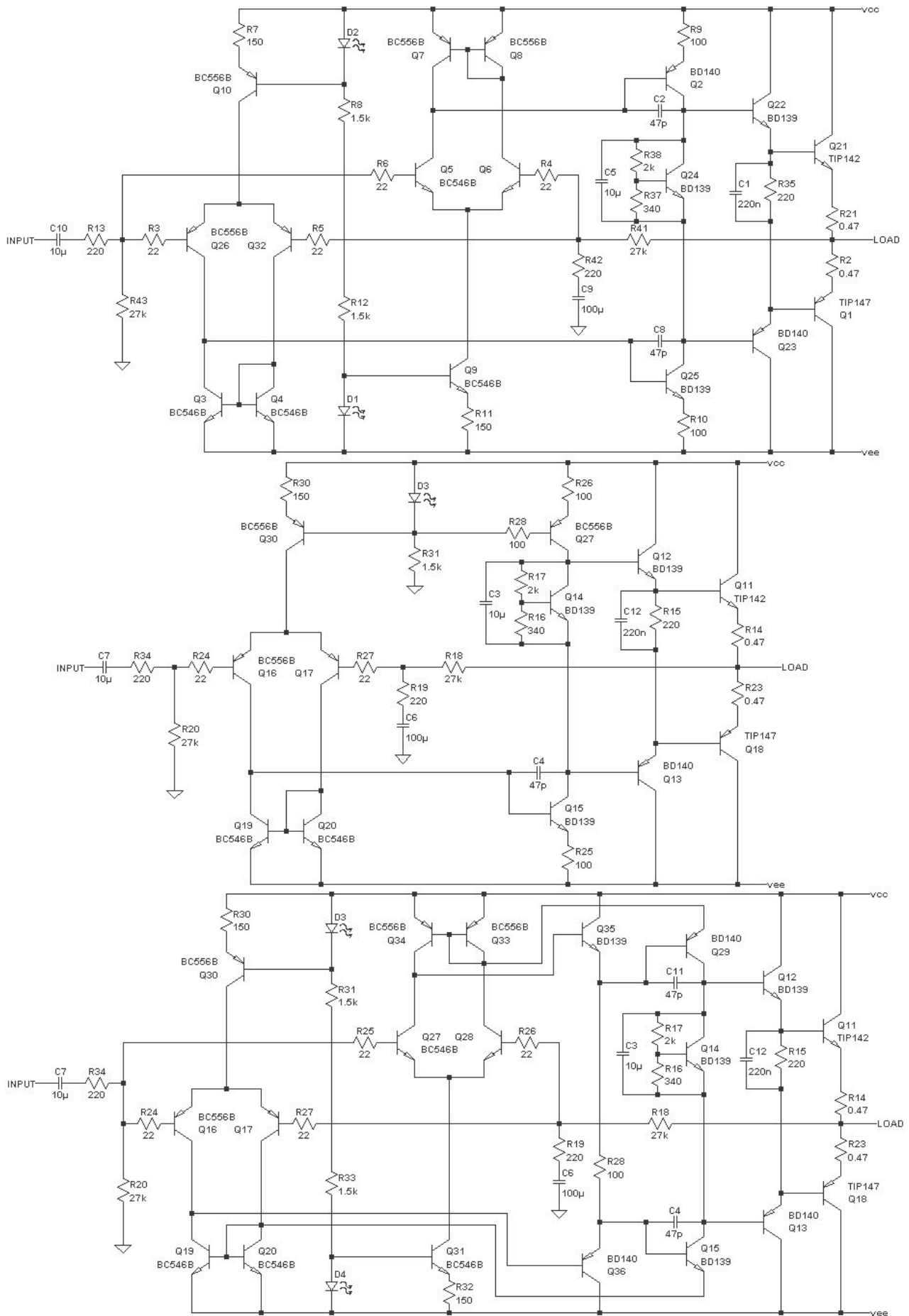


Figure 3.23 Circuits used in SPICE simulation comparing differences between ordinary and cascoded power amplifier circuits.

3.4.3 Symmetric designs: Cascoded complementary CFB and “Semi-symmetry”

Similarly to long-tailed pair, the current feedback input stage can be symmetrically cascoded as well. While the topology has seen much use in operational amplifiers, it is very rarely seen used in modern guitar amplifiers. If the characteristics of the circuit are evaluated this is not a very big surprise: First of all, the need to DC bias the input transistors makes designing a symmetric CFB amplifier a particularly difficult task – especially in dual supply circuits. In comparison to LTP topology, the input cannot be just conveniently referenced to ground but each base has to be biased individually. Furthermore, the designer still must deal with the poor DC offset performance, which is inherent for the topology. Using a single supply can make these troubles somewhat easier to deal with: However, that route often leads to inferior CMRR, PSRR and regulation. The circuit is also noisier than the LTP.

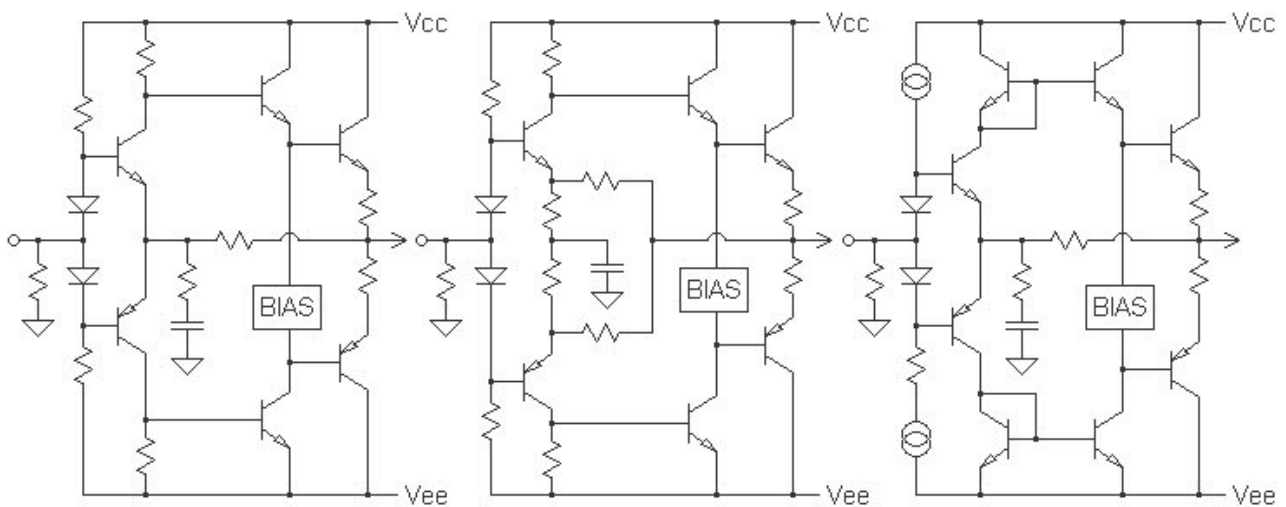


Figure 3.24 Examples of symmetric current feedback cascodes.

The area where the topology shows its benefits is the same as with the single-ended CFB: Due to greater slew rate the bandwidth of the circuit can be increased. Replacing the conventional DC coupling with a current mirroring scheme (as shown right in Figure 3.24) can further increase the linearity and bandwidth. However, while these are perfect features for video or RF applications they are not particularly useful in guitar amplifiers. Practically, it is hard to pinpoint any reasons why using the topology should be even considered.

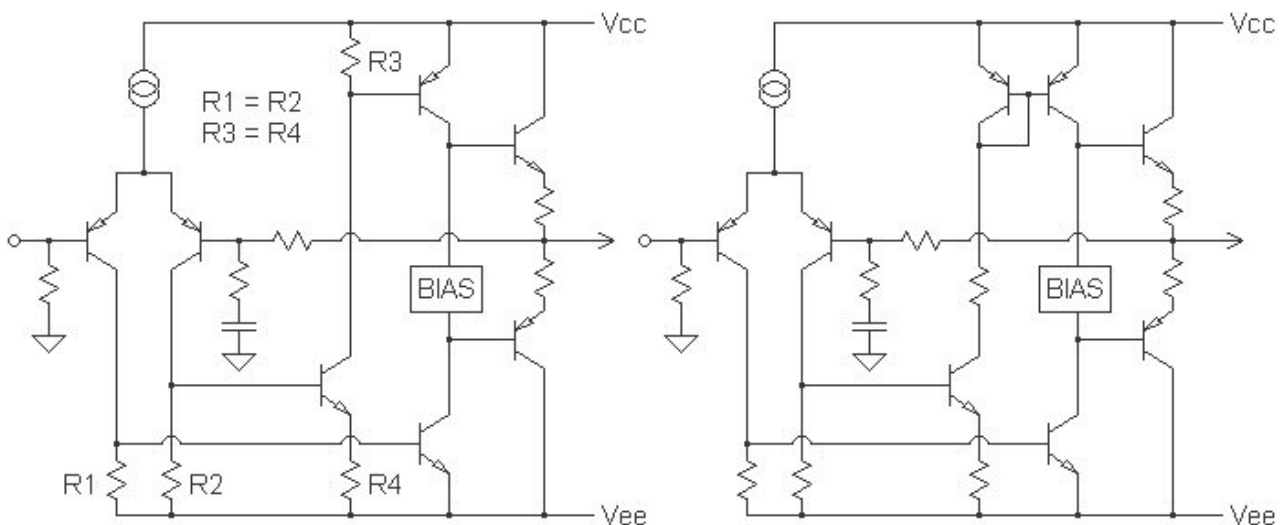


Figure 3.25 Semi-symmetric topology. Left: Direct coupling and right: Current mirror loaded.

The final example of symmetric topologies takes a bit different approach. As one can note, the circuit in Figure 3.25 is not completely symmetric since the cascoded dual input stage arrangement is abandoned and replaced with a phase inverter configuration that provides the suitable drive signal for the “upper” VAS stage. Evidently the merit of the circuit is its simplicity - whether that is a good or a bad thing. If operation of the phase inverter is linear enough and the differential stage properly balanced, the topology will perform in a fairly symmetric manner. In fact, it wouldn't be surprising if a good design outperformed the topology that uses dual cascoded differentials. Additional transistor will undoubtedly introduce phase shifts that impede reliable operation. Nevertheless, this topology has been used in several *very reliable* amplifiers.

3.4.4 Balanced voltage amplifier stages

With minor modifications the basic semi-symmetric topology converts to a semi-symmetric circuit that uses a symmetric differential VAS. National Semiconductor originally introduced this topology as a part of an integrated circuit. The input differential drives a VAS that is configured as another differential amplifier. Since in an ideal differential circuit the common mode errors should be cancelled the circuit should operate with similar linearity as a cascoded complementary circuit does. There is also another approach to achieve a balanced drive, the “Lender” circuit, named after Motorola designer Rudolph Lender. This topology relies on increasing open loop gain by connecting VAS transistor's emitter to the other branch of the differential stage (where the signal is in opposite phase) - a practice that also improves linearity by adding local negative feedback. The transistor in question also serves as the required phase inverter.

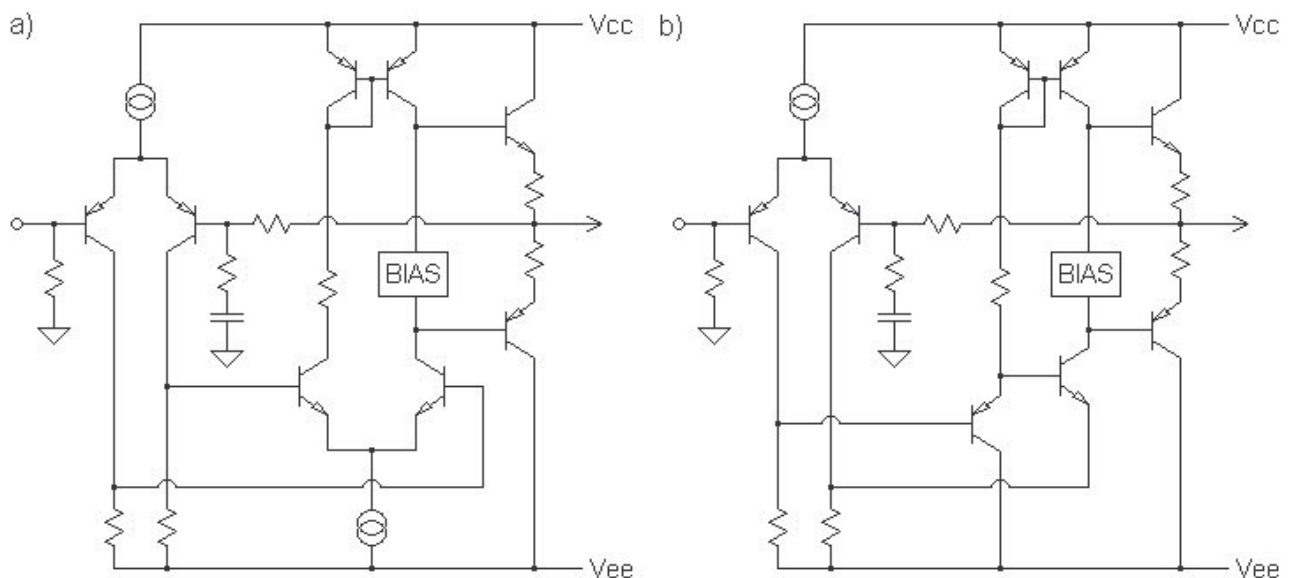


Figure 3.26 Balancing methods for voltage amplifier stages: a) Symmetric and b) “Lender”.

A symmetric VAS has about 10 decibels higher open loop gain than a conventional VAS and to maintain equal stability margins this increment has to be compensated by using a greater amount of high frequency negative feedback [3]. The weaker stability especially shows itself during clipping. However, compared to current source loaded single-ended designs the symmetric VAS arrangement typically provides a steadier clipping behaviour that introduces less high order harmonic content. Both differential and Lender topologies naturally cancel even order harmonics as well. However, this is

a minor benefit since the reduction in second harmonic distortion is only about 4 decibels and the effect is gradually lost during clipping. Reduction of other harmonic components is insignificantly small. The differential VAS circuit also has issues with imbalanced current flow in its branches (bias current is not steady, nor should it be) and with ensuring that both signal paths retain the same bandwidth.

The performance of Lender topology is highly dependant on correctly balancing the input differential and too much imbalance will stick the output very close to potential of the positive rail. The topology essentially presents a substantial design challenge, yet its performance is almost similar with common differential circuit. If you are meticulous about decreasing distortion as much as humanly possible these topologies may come in concern. In practice, however, there are very few reasons to justify the added complexity.

If input differential currents are well balanced the Lender circuit can have an excellent DC offset performance - and so can the differential VAS, which even tolerates slight transistor mismatches rather well. (Although the best practice still is to follow the general rule of device matching and thermal coupling). However, this is a typical characteristic even for the conventional design that uses a single-ended VAS.

3.5 Operational amplifiers as drivers

Since an OpAmp is practically just a voltage amplifier with a differential input one can easily use it as a substitute of discrete input- and voltage amplifier stages. Operational amplifier driven output stages are therefore pretty simple circuits and their operation is quite easy to understand. This is why their introductory was deliberately saved to be the last topic in the theme covering voltage amplifier stages. The most basic configuration of using an operational amplifier as the driver of the OPS is shown left in Figure 3.27. The OpAmp's output is simply hooked to a push-pull buffer with the necessary biasing components. Note that the circuit will not benefit much from bootstrapping or from the use of constant current sources (explained in 3.8 and 3.9) thus it can be left extremely simple. Nevertheless, its sound quality (in low-power applications) is amazingly good. This topology is used in countless headphone amplifiers and it also works very well as a spring reverb driver.

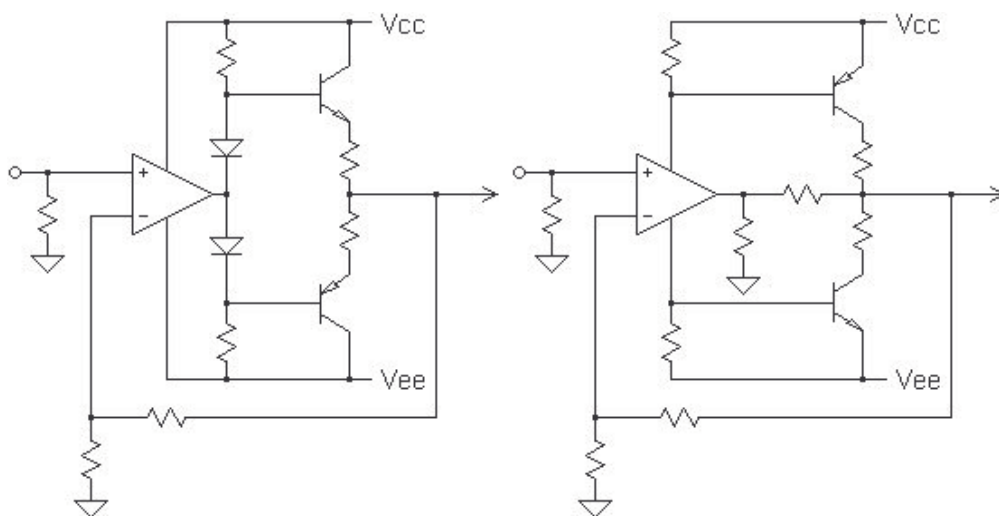


Figure 3.27 Operational amplifier-based voltage amplifier stages, which are limited by rail voltages. Note the configuration of transistors in the circuit on the right.

Another form to drive the output buffer is to sample the current of the supply rails that power the OpAmp. The OpAmp is then usually configured to drive current to the ground. The chips current draw is usually ranging from few millivolts at idle to about 20 – 30 mA. Note the switched order of the output devices. The circuit has worse noise figures and higher distortion than the basic topology - and most importantly, it doesn't have any inherent thermal compensation so its thermal stability is highly questionable. Some designs try to solve the problem by using high-value emitter resistors but a better solution is fitting thermal tracking diodes in series with the rail current sampling resistors. The diodes are configured to track the temperature of the output devices. As is, the topology should be avoided but serves a purpose of showing another way to take push-pull output from an operational amplifier.

The single greatest downside of OpAmp drivers is the restricted output voltage swing they usually have. Basic OpAmp chips generally have a maximum supply rail voltage of about ± 18 V. This effectively limits their realistic voltage swing to about ± 17 V, or even less. To a 4-ohm load this voltage swing equals only around 36 watts of output power (average) - and this is the theoretical maximum! To overcome the limit, several high-voltage OpAmps have been produced. Basically, their typical circuit applications follow the presented topologies but allow using higher rail voltages. For example, LM143 from National Semiconductor can be powered by maximum of ± 40 V rails. Each of these chips is a design of its own so I feel it's unnecessary (and unbeneficial) to go into further details. These chips also have a bad tendency to become obsolete in few years. If you are interested in the topic the best way to learn more is to find a datasheet of such chip and study it.

Using high-voltage ICs is not the only solution to increase the voltage swing of an OpAmp-based driver stage. Let it be mentioned that the basic topologies *can* be bootstrapped to increase the voltage swing towards the rails a bit: In this case *both* resistors hooking the bases to rails should be bootstrapped – After all, OpAmp is a *symmetric* circuit. Principles of bootstrapping are explained in 3.9.

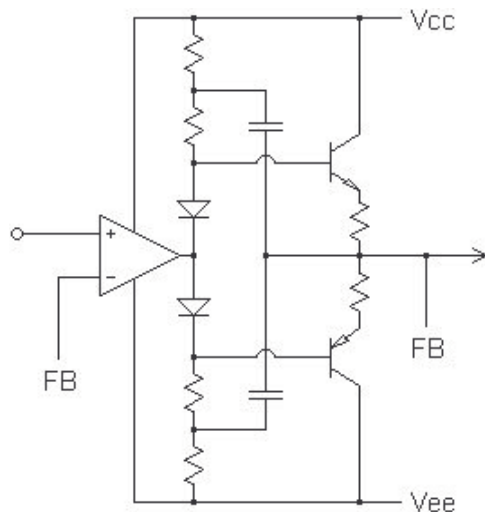


Figure 3.28 Symmetric bootstrapping configuration in an OpAmp-based driver stage.

However, there are alternative (and better) methods as well. The circuit shown left in Figure 3.29 relies on a novel idea: If rail voltage is virtually raised past the chip's "limits", the OpAmp can have a higher output voltage swing. To accomplish this, the OpAmp's supply rails are modulated (with transistors). The basic idea is to have a floating supply voltage (that enables large output signal swings) yet at the same time retain the voltage potential affecting between the chip's positive and negative power input pins constant and fairly low. Typically the pin-to-pin voltage is kept in the order of 30 V, which results into dual ± 15 V. In US this kind of circuit is

described as being "bootstrapped", in UK "wobbled". Further on, connecting parallel zener diodes from transistors' bases to the output node can increase the circuit's reliability. Zener diode value should be selected according to maximum supply voltage of the OpAmp (per rail). This circuit does not form a complete power

amplifier, though. In fact it has a quite poor current feeding capacity. To make a power amplifier out of it the circuit must be buffered, for example, with the circuit shown left in Figure 3.27. The rail voltage of the output buffer must naturally support the higher output signal swing.

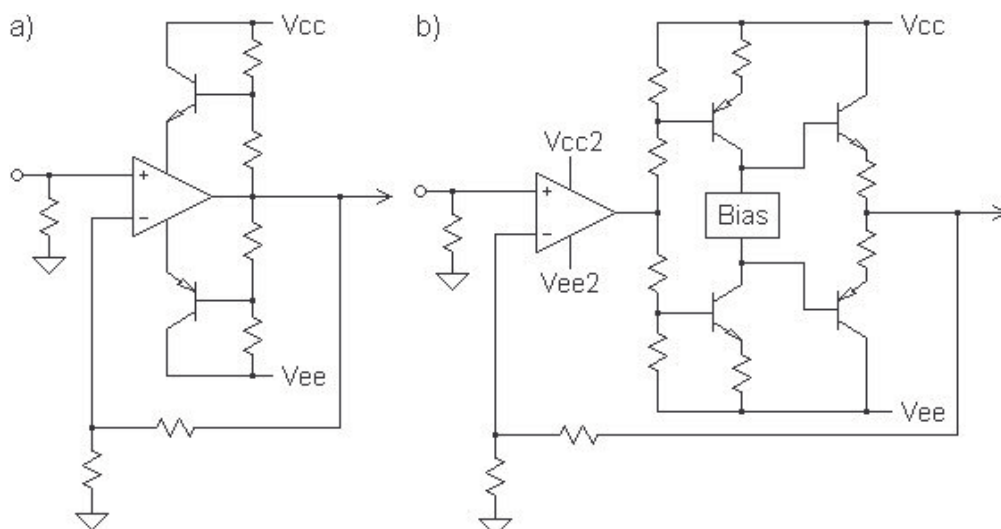


Figure 3.29 OpAmp voltage amplifier stages with increased voltage swing.

Another topology for exceeding supply rail limits is the one depicted at right in Figure 3.29. This one allows powering the operational amplifier with conventional low rail voltage. The OpAmp stage drives a discrete stage that is powered by higher rail voltage. This stage samples the output of the OpAmp utilizing “level shifting resistors”. The combination allows OpAmp output to swing at very low amplitude while the output stage swings at very high. Basically one can set the voltage swing to reach approximately the same levels as limits of higher voltage supply are [8]. The topology has one major drawback, which is the inherent instability: The circuit is very prone to oscillation and also “sticks to rail” if driven into clipping or loaded too hard. Combine these two drawbacks and the outcome is easily a catastrophe. This topology is, nevertheless, used in many excellent and reliable commercial amplifiers. On the other hand, unless you possess outstanding design skills I dearly suggest you to steer away from using this topology.

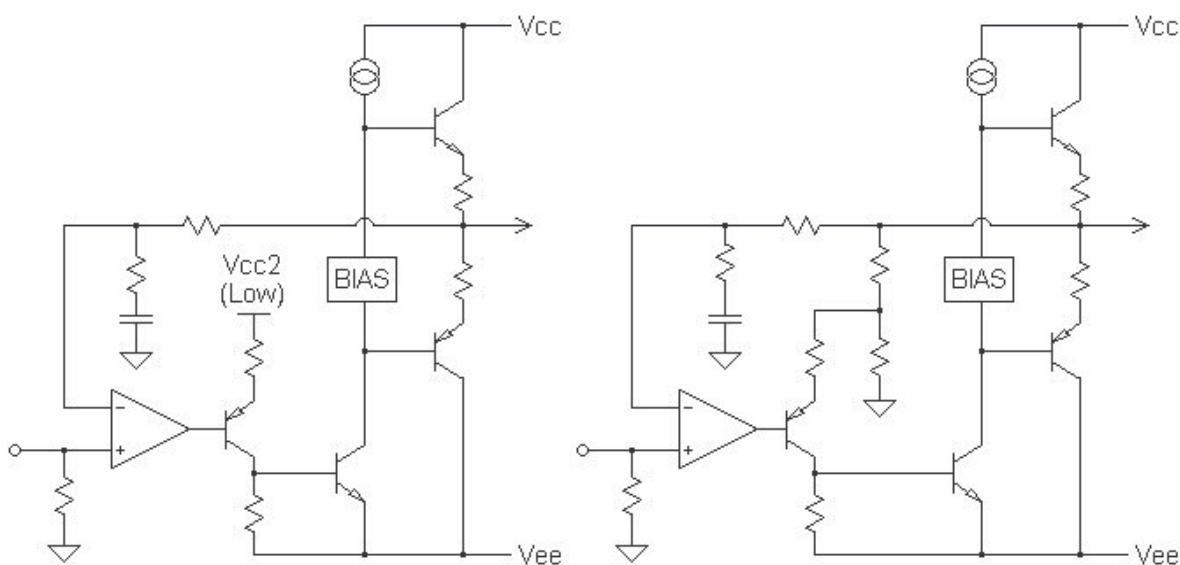


Figure 3.30 Topologies introducing external gain: Two methods of using OpAmp as a differential stage.

Other designs that extend the output voltage swing are depicted in Figure 3.30, Figure 3.31 and Figure 3.32. Using the operational amplifier basically as a mere differential stage is quite tempting since the OpAmp's internal differential stage is extremely linear. (A discrete circuit achieving equal performance would likely be fairly complex so replacing it with a simple and easy-to-use IC is appealing). As shown in the circuit example, the PNP transistor stage buffers the OpAmp, which is then followed by a conventional VAS. The PNP transistor interface also inverts the signal, similarly as a conventional LTP input stage would do. Since the additional voltage amplifier stage provides a fair share of the total open loop gain the OpAmp has to provide only a very small portion of it. Consequently, it has no need to swing to high output amplitudes, which in turn allows using a conventional OpAmp. In essence, the concept is extremely novel; unfortunately, phase shifts introduced by the additional transistors (again) cause the circuit to be inherently very unstable. Another problem is that (usually) the OpAmp is actually introducing *less than unity-gain*, which is a configuration with a highly questionable stability. This is not to say that the topology wouldn't work, on a contrary, it has been used in some highly respected commercial amplifiers. However, all of them were heavily compensated against oscillation.

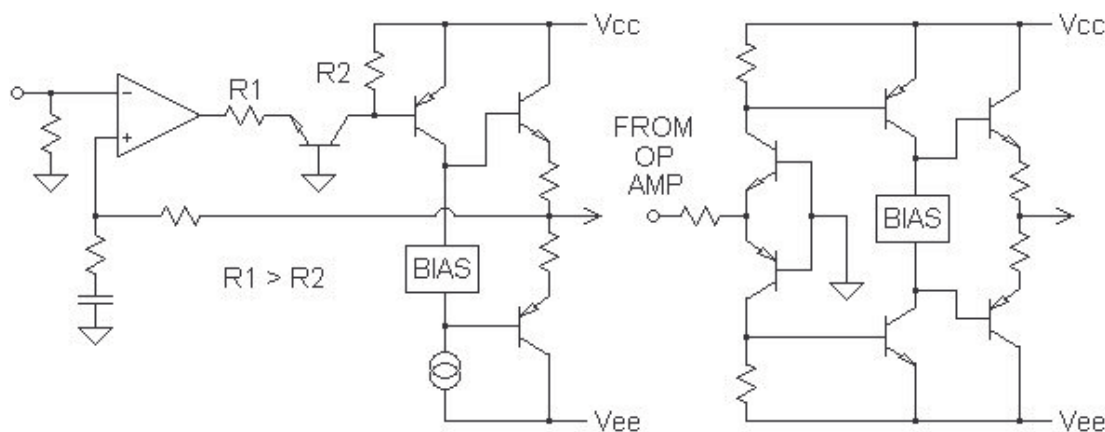


Figure 3.31 Topologies introducing external gain: Common base interface. Complementary symmetric version is shown at right.

An interesting, very similar design is the “common base interface”: This topology is pretty clever and has again seen successful use in some professional-quality PA- and guitar amplifiers. A common base circuit has questionable buffering properties since its current gain is less than unity. However, it has a rather high voltage gain that must be decreased by using an emitter resistor $R1$ that is larger in value than the collector resistor $R2$, which hooks to the positive rail. Quite illogically the common base circuit is also *non-inverting*; this will prove to be an essential detail.

Let's examine the circuit: The operational amplifier controls the emitter voltage, which is accordingly reflected at the collector that directly drives the VAS transistor's base. Consequently, an operational amplifier stage with low voltage swing can be used in driving VAS and OPS that are powered from higher rail voltages. Note that the inputs of the OpAmp are “reversed”, thus inverting input is acting as the non-inverting one. There is a simple reason for this: Due to lack of input stage that would provide phase inversion the VAS and OPS are *inverting*. To retain proper phasing the inputs of the OpAmp must correspondingly be reversed. Similarly to previously discussed topology this one is inherently rather unstable as well. The unfortunate part is that many conventional stabilizing methods tend to worsen the problem.

The last example of designs that introduce external gain is something I recommend to avoid: A driver stage with gain. Although this is quite a novel idea as well (and may be helpful in some MOSFET configurations) there are some very unfortunate shortcomings that seriously compromise the reliability and performance of the circuit. The benefits and drawbacks of the circuit have been previously discussed in 3.2.4 so there is no need for rephrasing them.

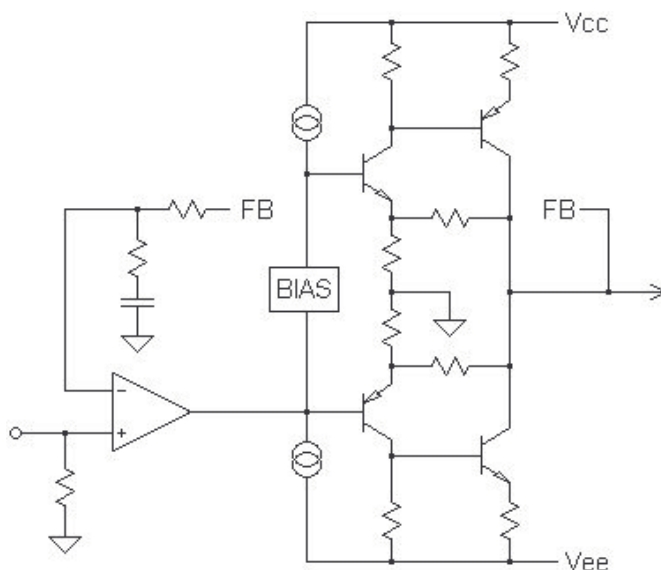


Figure 3.32 OpAmp topologies introducing external gain: Driver stage with voltage gain.

3.6 Negative feedback loops

3.6.1 Basics

Modern solid-state power amplifiers are mostly based on the idea invented by Harold Black in 1927: A very high *open loop* gain is controlled by introducing (usually a large amount of) negative feedback that forms a *closed loop*. On contrary to many beliefs, this mainly offers nothing but benefits: The use of negative feedback increases input impedance and bandwidth. Due to error correction the amplifier remains more stable and the amount of distortion is greatly reduced – this happens at the cost of reduced gain, though. Problems arise only if the amplifier is designed badly to begin with. An ideal amplifier circuit would be *as linear as possible* even without the feedback applied and would have a high open loop gain. Negative feedback should therefore be regarded as a *tool* not as a way to cover faults of the circuitry.

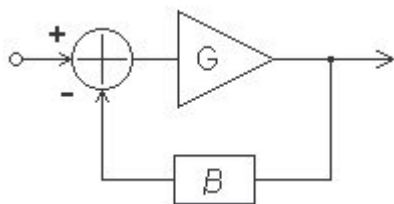


Figure 3.33 Feedback principle.

Feedback loops are categorized being either local or global. Local feedback takes place within the gain stage itself while global feedback extends over several stages. A very usual form of global feedback is the feedback taken from power amplifier's output to the input stage. In Audio Design Handbook Douglas Self even lists the

“Non-linearity resulting from taking the NFB feed from slightly the wrong place near where the power transistor class-B currents sum to form the output” as one of the eight distortions introduced by a power amplifier [3].

In a power amplifier the relation between input and output can be written as

$$V_{OUT} = G \times V_{IN} + \varepsilon \quad \text{(Equation 5)}$$

where G = gain and ε = error. Feedback can reduce this error by a factor of

$$1 + G\beta, \quad \text{(Equation 6)}$$

where $G\beta$ = loop gain. However, stability of the loop gain has to be carefully controlled by reducing its magnitude and phase to zero before its degree becomes 180° . If phase shift introduced by the feedback loop extends beyond 180° the negative feedback turns into positive; this is an incident, which surely triggers oscillation. What this means in practice is that error-correcting properties of feedback are both frequency dependent and less effective at higher frequencies. Note that a typical voltage amplifier inherently introduces a 180° phase shift (inversion) and with the inclusion of another 180° from the feedback loop the signal is completely in phase (360°) with the signal at the feedback injection point, thus triggering oscillation.

Misunderstanding the concept about introduced phase shifts within gain stages and feedback loops sometimes leads some people to believe that a current flow through a physically long feedback loop causes a significant time delay and therefore triggers oscillation. This is not so: Electricity is as fast as light and can travel seven times around the equator in one second! One can safely ignore any “time delay” caused by long wiring. Increased wiring length will, however, increase inductance and capacitance of the feedback path, which in turn introduces more *phase shift*. As you see, this is a completely different phenomenon.

The use of negative feedback also lowers the amplifier’s output impedance, thus increasing the damping factor. This is one of the main reasons why most transformer coupled circuits (including totem pole) behave “like tube amplifiers”; due to phase shifts that are introduced by the transformer, incorporating large amounts of NFB becomes impossible and the output impedance stays a great deal higher. The effects that high output impedance causes to frequency response of an amplifier have already been discussed several times, (refer to Figure 1.5), and such effects - coupled with a fact that an amplifier with a lower amount of NFB most likely also has a higher amount of distortion - may be the key point why people tend to think that amplifiers incorporating large amount of negative feedback sound “sterile”. An older technique, known as *degenerative feedback* is applied in the form of using emitter, source or cathode resistance. This also lowers the gain, raises the input impedance and improves linearity by mixing the input signal with a portion of the output signal. However, degenerative feedback does not have such error correcting capabilities as the negative feedback topology defined by Harold Black has. It also has no effect on bandwidth or output impedance. However, the degenerative feedback is still a form of negative feedback and undisputedly an essential part of electronic circuit design.

The other tasks of the negative feedback - besides correcting the output signal and limiting the gain into a usable (and more predictable) amount - is to handle the DC biasing of the output - and sometimes even input - stage. Apart from the simple benefit of increased linearity, the control of DC level is one of the reasons why most of today’s transistor power amplifiers use global feedback: Their output node’s DC potential is controlled by the input stage, which is feedback biased. There are solutions that allow eliminating the global AC feedback. However, it usually means

that the DC offset level must be controlled with a circuit known as a “DC servo”. Also, if a respectable amount of linearity is to be retained, the circuitry must also employ several and rather complex global error correction schemes.

3.6.2 Current sensing and mixed-mode feedback

A speaker does not present a constant load for the amplifier since its impedance varies with the frequency. Even the enclosure type has a huge effect on speaker’s impedance. The unfortunate fact is that impedance variations are quite large on the whole scale and they may have a tremendous effect on the amplifier’s performance. Figure 3.34 shows a common example of speaker impedance having the graph plotted from a SPICE simulation of an actual speaker. While the nominal impedance remains around 4 ohms it rises as high as 25 ohms at the free air resonance peak at 98 Hz. Due to voice coil inductance the impedance also starts to rise steadily after 1 kHz reaching 25 ohms around 8 kHz. Although the impedance curve is different in each combination of speaker models and enclosures, the example here can be considered as a fairly good presentation of typical speaker impedance.

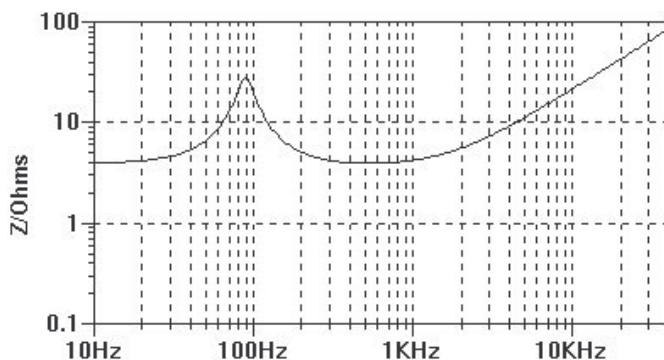


Figure 3.34 An example of speaker’s impedance.

Impedance really comes into play in how the amplifier drives a load. The common (or preferably “stereotypical”) way for transistor amplifiers is voltage driving, also known as *voltage sampling*: Negative feedback that sets the gain is applied with very little reference to load and as a result, the voltage gain remains constant but the current varies with the impedance – and so does the output power. Current driven - or *current sampling* – amplifiers work differently: The amount of negative feedback is referenced to the load impedance and as a result, the current gain remains constant but voltage varies with impedance – and so does the power but this time quite differently. Figure 3.35 shows crude schematics depicting both topologies as well as a schematic of their combination known as mixed-mode feedback amplifier.

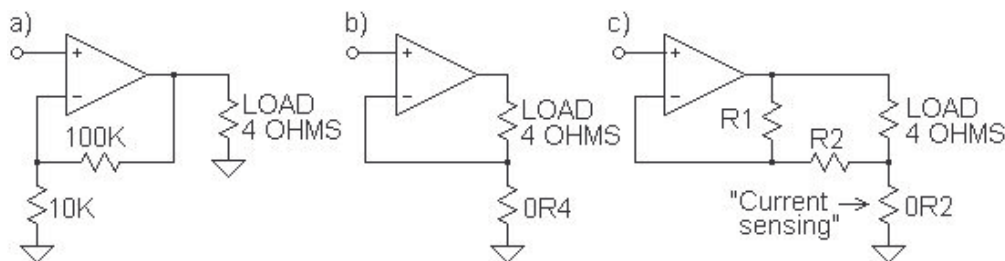


Figure 3.35 Feedback topologies: a) Voltage sampling, b) current sampling and c) mixed-mode feedback.

Typical transistor output stages buffer the constant voltage gain of VAS from the speaker load with an emitter (or source) follower circuit. The buffer has very high

input impedance so it effectively prevents the load impedance from affecting the voltage gain. Since the gain of VAS therefore has little relation to load impedance these amplifiers possess a very flat frequency response. Naturally they will introduce some deliberate high- and low frequency attenuation due to use of coupling capacitors and greater amount of negative feedback at higher frequencies.

However, an amplifier that has either plate-, collector- or drain-loaded output sees the speaker's impedance as the factor that dominates the voltage gain. While the voltage gain varies the load current strives to stay constant. A circuit like this will react very differently throughout the whole bandwidth as can be seen from Figure 1.5. Because circuits that operate this way are usually transformer coupled tube amplifiers it leads many people to believe that the non-linear response is a characteristic of tube amplifiers only. However, it's just a question of loading the amplifier: For example, a transformer coupled common cathode circuit would yield a rather flat frequency response since it is basically a buffer circuit with a constant voltage gain.

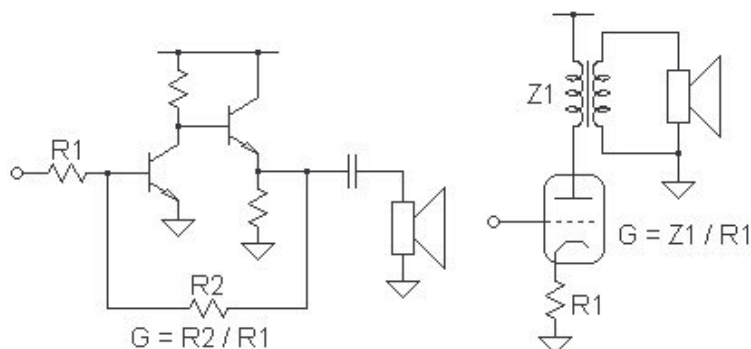


Figure 3.36 Simplified examples of amplifier loading configurations. Left: Buffered high-OLG amplifier. The feedback network defines closed loop gain. Right: Typical transformer-loaded common cathode amplifier without a feedback loop.

Let's examine the circuits shown in Figure 3.35. Gain of the circuit a) is ideally 100 kilo-ohms divided by 10 kilo-ohms plus one, which equals 11. With 1 V input signal the power fed to 4- and 25-ohm loads can be derived using Ohm's law. If the amplifier is an ideal one, it feeds 30.25 watts of power to a 4-ohm load and only 4.84 watts to a 25-ohm load. In the current driven amplifier (circuit b) the efficient gain is only 4 ohms divided by 0.4 ohms, which equals 10. The upper pin of the current sensing resistor acts as a reference point and therefore all the power affecting over the sensing resistor is practically wasted, hence the "+1" is lost from the equation. The output power with a 4-ohm load is now only $U^2 / R = 25$ W. However, current will remain constant throughout the whole impedance range and it can be calculated according to Ohm's law ($I = P / U$) resulting in a figure of 2.5 amperes. With a 25-ohm load the power will then be as high as $I^2 R = 156.25$ W - assuming that the amplifier can provide the required ± 62.5 V swing. Not many can. Let's assume that in another scenario the impedance drops below 4 ohms, say to 2 ohms. In case of the voltage drive the 11 V affecting over 2 ohms will result in 60.5 watts of power, which is logically double fold the amount of power to 4 ohms. The current driven amplifier will provide much weaker results: 2.5 amperes through 2 ohms gives only 12.5 watts of power.

It is clear now that each feedback topology has its advantages and disadvantages. However, it is possible to combine the two by defining variable output impedance between the two extremes. Schematic of a basic mixed mode topology is shown in the figure c). The current sensing resistance is reduced; resistors R1 and R2 define the gain and the circuit's behavior in different loading scenarios. To calculate the voltage

gain under no-load condition the load impedance can be ignored therefore resulting in $U_{UNLOADED} = R_1 / (R_2 + R_{SR}) + 1$, where R_{SR} = the resistance of the sensing resistor. The voltage gain under a load is calculated with a formula

$$U_{LOADED} = 1 / [1 / (R_1 / (R_2 + R_{SR}))] + 1 / [(R_{LOAD} / R_{SR}) + 1] \quad \text{(Equation 7)}$$

Where R_{LOAD} = speaker's resistance during a desired loading condition. The formula simply calculates the parallel gains of two signal paths: Through R_1 , R_2 and R_{SR} and through R_{LOAD} and R_{SR} . The results can be used to calculate the load current and the exact output impedance.

$$I_{LOADED} = U_{LOADED} / R_{LOAD} \quad \text{(Equation 8)}$$

$$Z_0 = (U_{UNLOADED} - U_{LOADED}) / I_{LOADED} \quad \text{(Equation 9)}$$

Finding the “right” R_2 - R_1 -combination, that matches the speaker's impedance properties, is always more or less a compromise. I found the following method quite helpful: I fed the formulas to an Excel chart, which plotted a graph of the output power according to an array of various impedance conditions. This way I was able to come up with some “optimal” values for the following example circuit. These were: $R_1 = 13$ kilo-ohms and $R_2 = 330$ ohms. With these values the voltage gain to nominal 4-ohm load impedance is approximately 13.8 resulting to a load current of 3.45 A. Therefore the output power would be approximately 48 watts. The derived “optimal” values will be used in the following comparison of voltage- and current drive topologies against mixed mode.

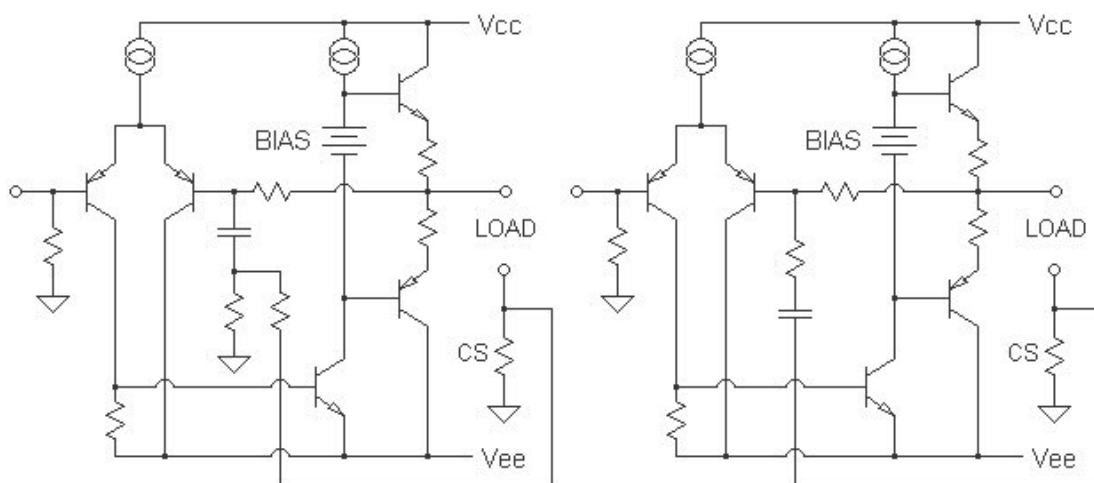


Figure 3.37 Practical methods for applying mixed mode feedback. Left: Conventional method and right: Simple method.

When the impedance rises to 25 ohms the voltage gain is approximately 30.6 and the current 1.22 amperes, resulting into an output power of approximately 37 W. It is a very good figure compared to voltage driven amplifier's output power of only 7.6 W. However, an ideal current drive amplifier would have had an output power of 297.5 watts! In the case of impedance dropping down to 2 ohms the voltage gain of mixed mode amplifier drops to roughly 8.6. Current is 4.3 amperes, resulting into an output power of approximately 37.4 watts. Output power of voltage driven amplifier would increase to 95.2 watts while the output power of the current driven amplifier would decrease to 23.8 watts. It is clear to see that mixed mode amplifiers can't have the

advantages of both topologies but can offer a nice compromise. Further tweaking of values might yield a circuit with even better efficiency. One thing worth noting is that mixed mode amplifiers can be tweaked to mimic the response of “tube amplifiers” or to provide a nearly flat power versus frequency response.

Figure 3.37 depicts some practical methods of applying mixed-mode feedback. The circuit shown on left is likely the most common configuration: Current is sampled and injected to the feedback loop through a resistor. If there is DC in the output (like in the case of single-rail circuits) it is isolated from the speaker due to same capacitor that sets unity DC gain (naturally the output in such case is AC coupled as well). The topology is fairly adjustable for different kinds of responses. You will find this circuit from hundreds of guitar amplifiers. In the circuit shown at right the feedback path is grounded through the current sensing resistor. This is an equally efficient solution but a lot more cost-effective in case of large production runs. While the configuration is not as common as the first one it has been used in numerous amplifiers as well. Unfortunately the circuit does not provide vast adjustment properties and furthermore, it makes the low-current input stage share the common return path of the high speaker load return currents. In case the board and wiring layout is poorly designed this might cause problems.

3.6.3 Resonance, Presence and damping controls

Implementing tone control circuitry to power amplifier’s feedback loop has been a known practice ever since early vacuum tube days. While this technique might seem strange nowadays it actually isn’t that far-fetched - after all, most feedback loops are, by default, configured to introduce basic high- and low-pass filters. There is a very small step from this arrangement to using more elaborate tone controls, such as the “Baxandall” circuit, which for example was used in some Sansui HiFi amplifiers. Merging tone controls and (power amplifier’s) feedback loop has been quite unconventional practice in solid-state amplifiers though. Partially this is so because the local feedback loop of a typical power amplifier doesn’t lend itself to modifications easily: Usually anything but the traditional gain set arrangement (where loop impedance stays constant) causes instability or disturbs the output impedance – or even worse, the DC offset level. For a long time designers were also confined by the “urban myth” that linearity and “transparency” is the preferable configuration for solid-state amplifiers.

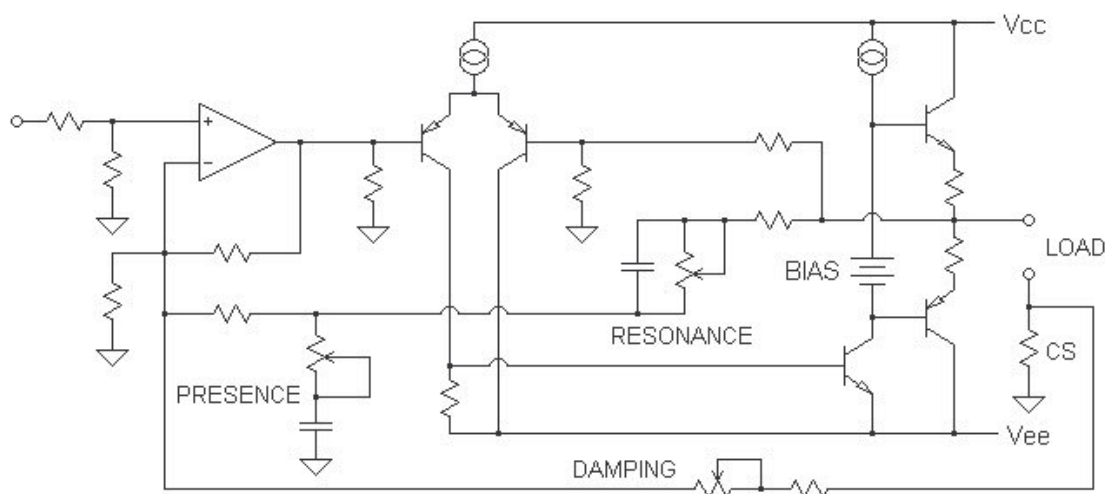


Figure 3.38 Typical tone controls placed within feedback loop.

Since tampering with power amplifier's local feedback is pretty much asking for trouble it is common to create another, global feedback path, which connects to the gain stage feeding the power amplifier. This path can easily house the tone control circuitry without a significant risk of reducing stability or messing up the DC offset level. Traditional tone controls that are implemented to power amplifier's feedback loop are simple "resonance" and "presence" controls. Basically, these are just alternative names given for bass- and treble controls, respectively. The operation of the aforementioned circuits is pretty straightforward so it shall not be discussed in extensive detail. "Resonance" control bypasses a capacitor with a variable resistance. Effectively, this causes the feedback path to present a lower resistance at higher frequencies; this in turn reduces their gain. In systems that have a low damping factor the bass boost of "resonance" control accentuates the effects of speaker's resonance – hence the name. "Presence" control is typically configured to affect the overtone frequencies of human voice range. This can make the instrument sound more "present" or stand out from the mix. The control works by shunting the feedback path to ground, which effectively is a practice that increases gain. Since the shunt path is blocked by a capacitor, higher frequencies see a lower resistance and therefore they are shunted more effectively. Naturally this configuration acts as a boost for mid-range and high frequencies.

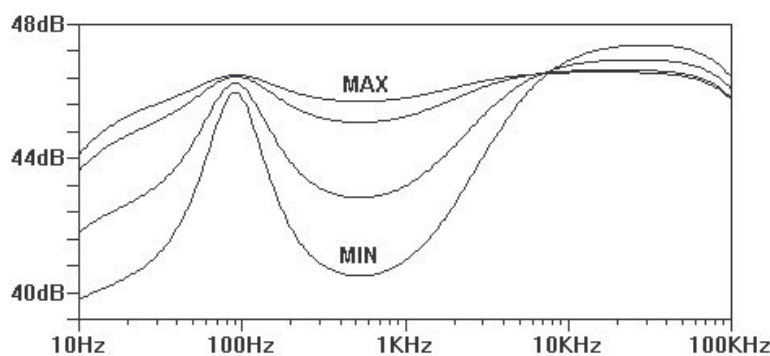


Figure 3.39 Plotted frequency response shows an example of the effect of a "damping" control.

A "damping" control varies the amount of current feedback, which effectively varies the output impedance of the amplifier and in turn affects the damping factor and frequency response. This is essentially a control that accentuates the speaker system's inherent characteristics that typically are perceived as adding quite a bit of "warmth" and "punch" to the signal. In practice, it is quite subjective how people perceive the effects of "loose" or "tight" (damped) system. The damping control is inherently an attenuator since the amount of current feedback is lowest when the damping control is in its maximum position and increase of negative feedback always results into loss of gain. This could be compensated, for example, with a dual gang potentiometer split to control both damping and gain. However, such designs are out of the scope of this section and will not be discussed.

The increased requirement for voltage swing in a mixed-mode feedback topology has been discussed earlier – and undoubtedly the increased voltage swing may result into clipping. Because the current- and voltage feedback paths of the tone controls are now diverted to an ordinary gain stage preceding the power amplifier this stage will respond to changes in (speaker) load impedance by varying its gain. It is rather easy to (smoothly) limit the output signal range of this low current, low voltage gain stage so that it never exceeds the input sensitivity of the power amplifier itself. Numerous designs operating at similar principle have been built. Some of them will be discussed later in the section concerning tube emulation.

3.6.4 Feed-forward

The local compensation of HF feedback in the form of Miller capacitance introduces slewing induced distortion. An improved feedback method is a so-called “feed-forward”. Basically there are many ways how to achieve feed-forward but one thing common for all of them is that two or more signals are added or subtracted to form the output signal. The important point is that the combination of the signals should cancel the distortion. In practical circuits, the summing (or subtraction) can basically happen either in parallel mode using “mixer” resistors or in series mode where the load is connected between the two outputs. In Figure 3.40, G is the main amplifier designed for efficiency and G2 the less effective, error-correcting amplifier designed for linearity. Their input-output relationship can be calculated with an equation

$$V_{OUT} = (G1 + G2 - G1 \times \beta) \times V_{IN} \quad \text{(Equation 10)}$$

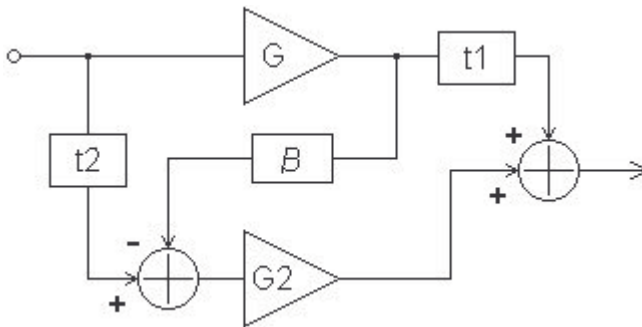


Figure 3.40 Feed-forward principle.

The feed-forward error correction topology is not very common; from the top of my head, I cannot think of any guitar amplifier that would use it. On HiFi systems there are a few examples, like the Quad 405 and its “current dumping” feed-forward scheme. A current dumping amplifier consists of a class-A voltage amplifier that supplies the full voltage swing to the load. The class-A amplifier is aided by a non-linear class-B current amplifier stage that supplies most of the load current. The signals are summed in parallel mode. One of the most astounding features of the topology is that it requires no thermal compensation or bias circuitry to compensate the crossover distortion – thus it is effectively very easy to set up.

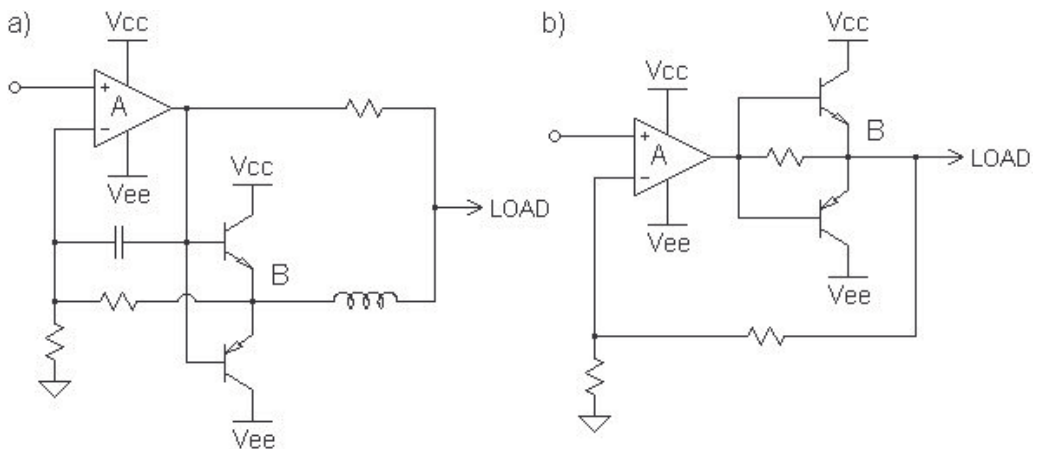


Figure 3.41 “Current dumping” feed-forward error correction topologies: a) Quad and b) a common topology for subwoofer amplifiers that are driven by OpAmp.

Another example of feed-forward is shown in Figure 3.42. According to Marshall Leach [28] “The network is said to be of the feedforward type because it takes the high-frequency signal from a point forward of the output node in the loop-gain transfer function.” At higher frequencies capacitors C1 and C2 begin to couple the signal, which leads to two things: C1 shunts the high frequency negative feedback signal to ground, thus the shunted frequencies effectively have a gain equal to the open loop gain. However, C2 opens a new path for high frequency negative feedback thus limiting the gain to an amount defined by the new closed loop. The resistor in series with C2 limits the amount of feedback in order to prevent oscillation that might occur. (Too much global high frequency feedback tends to make amplifiers unstable). The example schematic does not show the important “Miller” capacitor from VAS collector to base. Again, it cannot be completely omitted but the value of it can be decreased notably while still retaining an equal bandwidth. As the Miller capacitance becomes lower the slew rate of the amplifier improves thus increasing the accuracy of the feedback signal. This in turn leads to an improvement in error correction and decrease in THD.

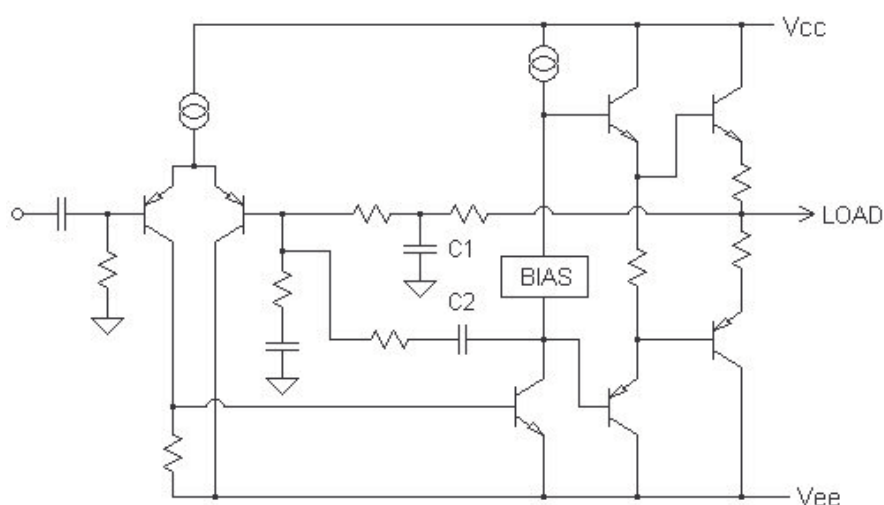


Figure 3.42 High frequency feed-forward feedback.

This topology is used in the famous “Leach” HiFi amplifier with a difference that the feedback takeoff point is located at the emitter of one of the drivers. Both methods work equally well. One may also wonder why not just omit C1 along with one of the series resistors completely and connect the other end of C2 straight to output thus converting the topology into plain negative feedback. This indeed works as well and in fact the method was used in the earlier revisions of “Leach”. However, Marshall Leach, the designer, suggests that his method improves stability by separating the feedback takeoff point from any capacitance of the load. I haven’t verified this but a SPICE simulation shows no difference in operation between alternative methods. Basically, there is a thin line between whether this topology should be considered as ordinary negative feedback or feed-forward.

Figure 3.43 shows a plot that compares total harmonic distortion in amplifiers that use a conventional feedback topology and a feed-forward topology similar to one in Figure 3.42. However, the feedback take-off point was an emitter of the driver. Both circuits behave almost identically to around 20 kHz - after which the detrimental effect of Miller compensation becomes clearly visible: At approximately 33 kHz the THD of the amplifier using conventional feedback topology is nearly 10 % while the THD of feed-forward amplifier is only about 2 %.

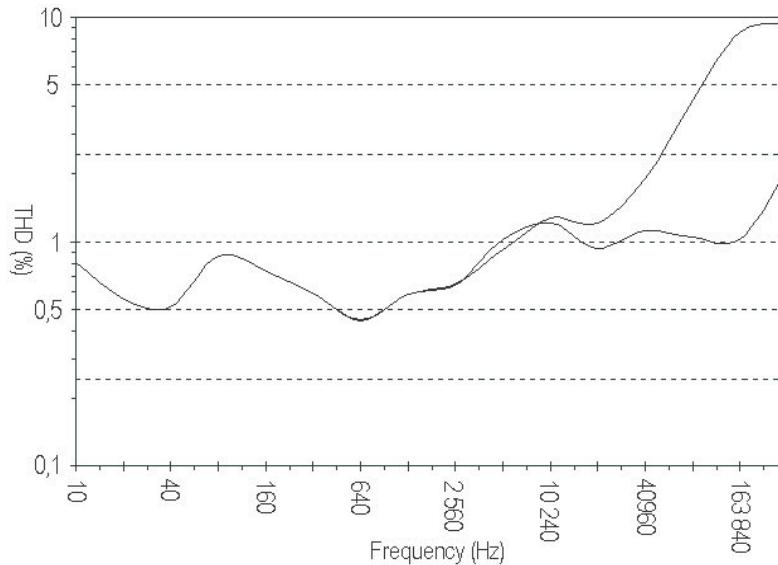


Figure 3.43 Comparison of THD in amplifiers using conventional and feed-forward feedback topologies.

Do not take the following comment as disregarding the efforts made to linearize the output signal and the remarkable results that these techniques can give, but I seriously believe they fall in to the category of “over designing”. I have to admit that I am impressed by the linearity of the output signal but the topology starts to show beneficial improvement in THD only at around 30 kHz. At a moderately high frequency e.g. 200 kHz the output signal is notably linear but the fact is that even bats could not hear a frequency as high as this. Surely improving high frequency reproduction will decrease the amount of intermodulation distortion but this also starts to affect only at frequencies that are *above* 20 kHz, which is basically the limit of hearing range for an average person. Connecting a typical guitar, or even PA, speaker to the output would effectively drop the reproduced bandwidth even more (to around 5 – 10 kHz) and furthermore, basically all frequencies above 10 kHz are considered unpleasant when guitar signals are in question.

3.7 Crossover distortion compensation

Let’s assume a case without a biasing circuit: When there is no input signal present the output voltage of a push-pull emitter follower buffer is exactly mid-way between supply rails. At this point the base-emitter voltages are zero and the transistors do not conduct. As a matter of fact, they do not start conducting until the input signal begins to exceed their forward voltage, which is the voltage over base-emitter junction. It is typically around ± 0.65 V for conventional- and compound transistors and around ± 1.3 V for Darlington pairs. This phenomenon will cause audible “gaps” to appear in the output signal every time the input signal’s amplitude drops below the forward voltage. The effect is naturally most horrid on low amplitude signals. Such a circuit is said to operate in class-B mode and it suffers from huge amounts of *crossover distortion*.

Counteract is to bias the transistors so that their idling voltage never drops below the forward voltage. A specific amount of current, known as *bias current*, is constantly fed to the transistors’ bases in order to ensure that the transistors keep conducting for a desired amount of time. A typical way to feed the bias current is to form a voltage potential between the transistors’ bases. While the forward biasing may not remove all crossover distortion it will certainly reduce it by buying more time for the other

transistor to turn on. When the amount of bias current is increased the amplifier's operation starts to resemble more and more that of a class-A, hence the name class-AB, which is used to describe the compromise between quiescent power dissipation figures of class-A and class-B modes. Due to added power dissipation when idling - caused by the increased *quiescent current* flow - a small amount of the amplifier's efficiency has been sacrificed for a more linear output. Since the transistors do not conduct through a full cycle the circuit still is *really* a class-B amplifier. In some sense AB is really not a valid class and arguably a "better" name would have been "Crossover distortion compensated class-B amplifier". Even this circuit will have some crossover distortion; the amount of it naturally depends on the amount of bias current. If the current is too low the output signal will contain sharp, transient "blips" at the switching points. (Refer to Figure 3.47). Note that blips do not occur at the zero crossing line because the conduction period of the transistors is extended.

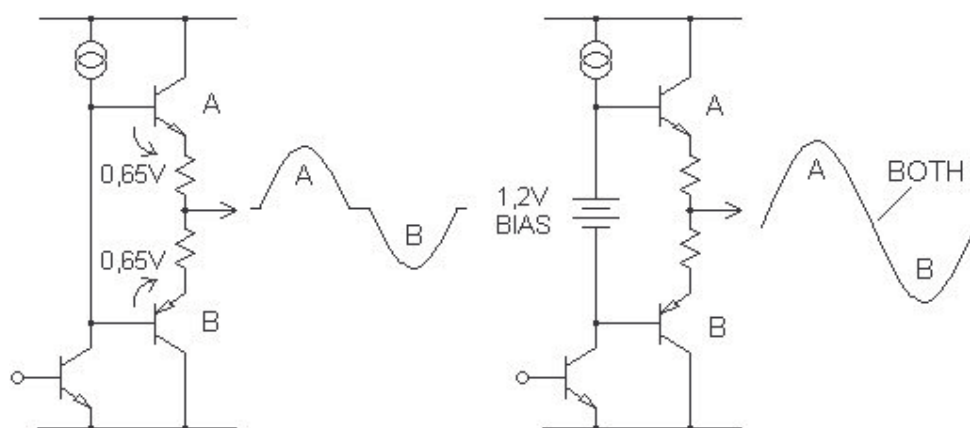


Figure 3.44 Class-B (left) and class-AB (right) push-pull emitter followers.

Creating the required bias current is not as simple as it first seems. The easiest way to cause some current flow between two points is of course to add some resistance; the resistor value and the voltage drop over it will then dictate the amount of current. This, in fact, was the method used by earliest transistor amplifiers. However, transistor's forward current is temperature dependent; it decreases while the temperature increases. The more the output transistors warm up, the less base-emitter voltage they need to turn on. Unless the phenomenon is slowed down somehow the transistors might drift into an uncontrolled operation, known as thermal runaway, which causes them to self-destruct.

What is needed is a biasing circuit that is temperature dependant as well. First solutions to construct one involved using NTC (Negative Temperature Coefficient) – resistors but they were expensive, slow and rather inaccurate so other methods were developed. Using a diode to replace the resistor proved to be a better solution: It offers a voltage drop which is temperature dependant and by matching the thermal coefficient with the transistor the bias current can be kept quite stable. If an accurate thermal tracking is required the diodes are mounted to the same heat sink as power transistors. Since one diode usually wasn't enough, older amplifiers often used integrated series-diode packages that contained several diode junctions. Sometimes a junction of a transistor is used instead of an ordinary diode.

Although the aforementioned methods allowed fairly good thermal tracking that was able to prevent the output stage from drifting into thermal runaway problems with linearity still remained: Distortion increases significantly along with only small changes in the optimum bias setting – whether they happen "up" or "down": *It is not*

that well known but increasing the bias current over the optimum point will also create crossover distortion – this time in the form of increased transconductance. While output devices conduct simultaneously their current gains are simply summed together. This phenomenon is known as “*gm doubling*” and it creates as dreadful effects as conventional crossover distortion. At this point it is worthwhile to mention that some technical literature disagrees with the existence of *gm doubling*, yet comparing the output and input signals can easily reveal that it does indeed exist: Amplifiers having either too low or too high bias settings always have a greater variation in gain around the crossover region than ideally biased amplifiers do. Any decent SPICE simulation, that doesn’t make a mistake of using too simplified circuit models, can expose this behavior. In essence, a “hotter bias” setting will not steer a class-AB amplifier’s operation closer to class-A mode, it just adds more distortion (along with power losses).

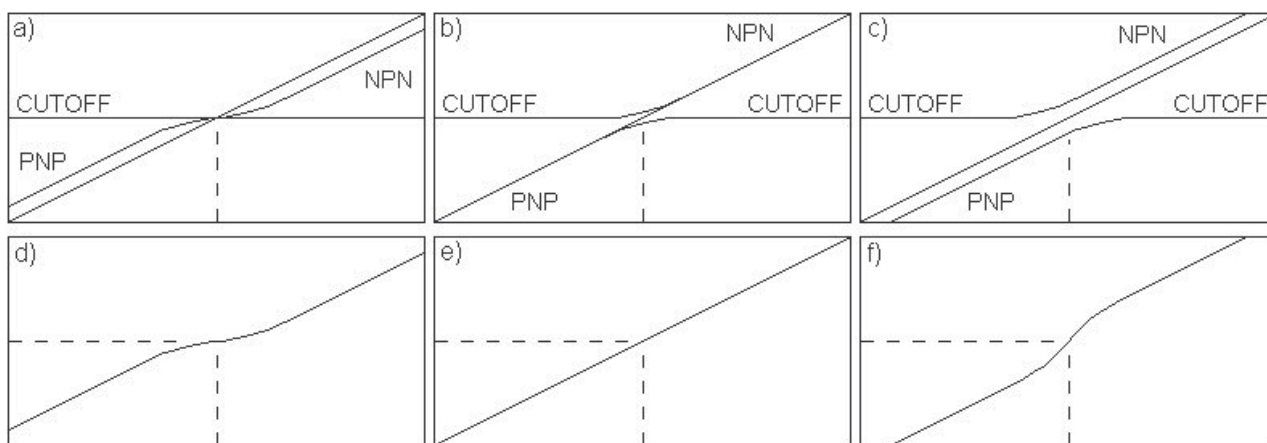


Figure 3.45 Biasing conditions of class-AB circuit: a) Not enough bias current, b) optimal bias and c) too much bias current. The corresponding composite transfer curves are shown in d, e and f.

In a *bias servo*, the diodes are substituted with a transistor that acts as a variable resistor. Effectively, the servo circuit is a constant voltage source - also known as VBE multiplier, amplified diode or, more old-fashionably, a “rubber diode” or “rubber Zener”. The servo transistor is mounted to the same heatsink as the output devices so that it can track their temperature. Temperature changes of the sink affect the gain of the servo transistor; this consequently changes the voltage drop over the servo circuit.

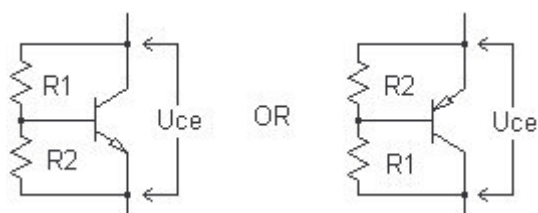


Figure 3.46 The basic VBE multiplier bias servo.

The bias servo has an advantage over ordinary diode strings: In case of Darlington transistors, the temperature coefficient of both transistors is not summed but multiplied, which makes matching up the coefficient with diode strings rather difficult. The rubber diode, however, can amplify its coefficient effect. The circuit is often equipped with a variable resistor that allows fine adjustment of the bias current. Sometimes diodes or temperature coefficient resistors are used in conjunction with the voltage divider circuit. Sometimes they substitute the transistor as the thermal tracking element. Since the bias voltage should remain on a fixed level regardless of output

voltage swing, a common way to improve the servo circuit has been to AC-bypass it with a capacitor, thus making it affect only at DC level.

The required amount of voltage drop over the VBE multiplier is derived from the number of output's base-emitter-junctions: In case of two Darlington transistors the number of these junctions is four and the bias voltage should be approximately four times the forward voltage of a transistor (4 x 0.65 V). Roughly put, the collector-emitter voltage in a basic VBE multiplier circuitry is derived from the ratio of the two resistances using a formula

$$V_{CE} = V_{BE} / R2 \times (R1 + R2) \quad \text{(Equation 11)}$$

The equation naturally consists of two parts: $I \times R$, which results in U . (This is basic Ohm's law). As one can note, the first part of the formula derives current (U/R) from two known constants, which are resistance of $R2$ and forward voltage drop V_{BE} . This is the equation for current through $R2$. If we ignore the (usually low) base current we can assume that the same amount of current must flow through $R1$ as well. This current through the combined resistance of $R1$ and $R2$ (the second part of the formula) causes the required voltage drop ($I \times R$). Note that the procedure ignored the existence of base current, which in fact is a quite important detail. The formula can be simplified as follows:

$$V_{CE} = V_{BE} \times (1 + R1/R2) \quad \text{(Equation 12)}$$

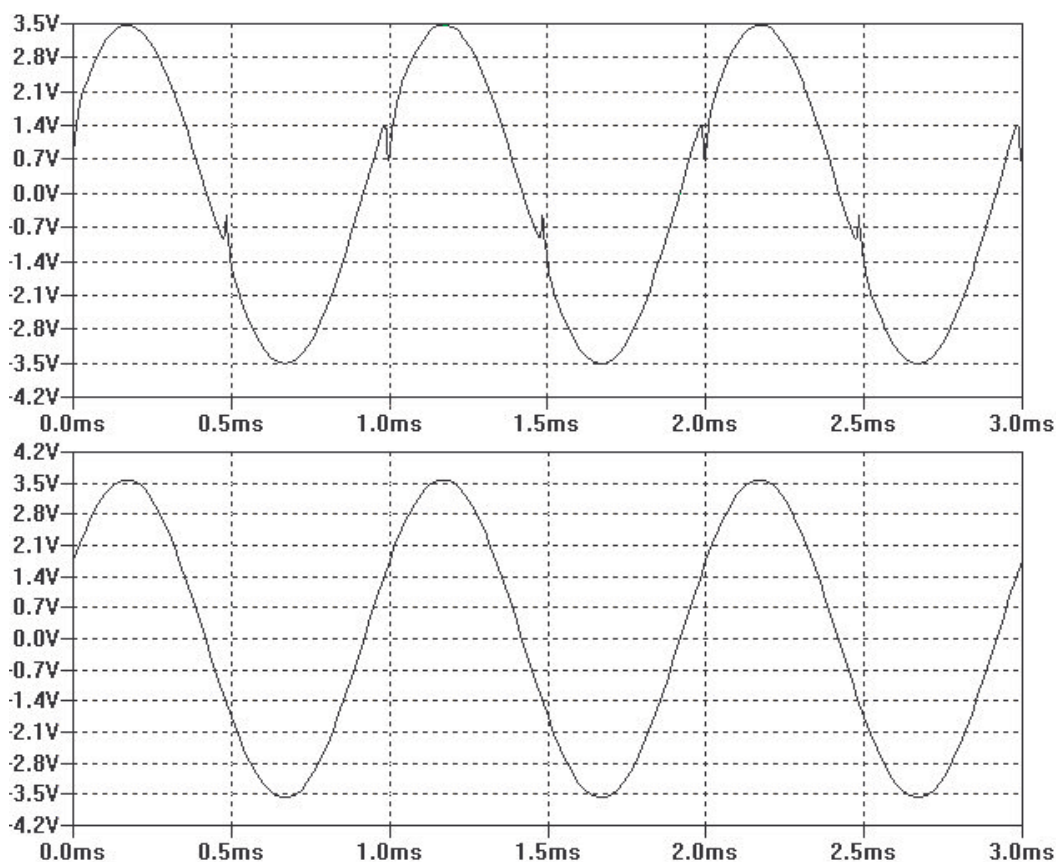


Figure 3.47 Class-AB amplifier's output signals with insufficient (above) and sufficient (below) crossover distortion compensation.

In order to acquire the optimal operating point the current through $R1$ multiplied with 9 should approximately equal the collector current. The equations apply only when the

transistor's current gain (also known as H_{FE} or beta value) is reasonably high to allow ignoring the effect of required base current from the deviation. With low gain transistors this fact can cause very nasty surprises: For example, Figure 3.47 plots the outputs of class-AB amplifier circuits that use two different servo transistors: Upper plot comes from using a low-beta BD139; bottom plot comes from using a high-beta 2N2222.

The basic VBE multiplier can be enhanced with several methods; the most common additions are the inclusion of a current-compensation resistor or increasing the loop gain. Because in reality a voltage amplifier's constant current source load (see 3.8) is an imperfect element - and therefore the VAS current is not so constant - V_{CE} of the basic VBE multiplier circuit will vary according to current fluctuations through the servo circuit. However, in order to keep crossover distortion in minimum the bias voltage should be kept as constant as possible regardless of servo current. The inclusion of current-compensation resistor causes V_{CE} to decrease when the input current increases. The effect is shown in the graph of Figure 3.48. With the choice of $R3$ the steepness of V_{CE} limiting can be controlled. In essence, the value of $R3$ should be optimized so that only slight voltage variations occur around the optimal bias voltage setting. At this point it's perhaps necessary to remind that the enhancements only affect the overall linearity of the VBE multiplier and do not interfere with its basic operation, which is to provide a steady bias voltage that corresponds to sensor device's temperature. In this equation, any variation in input current is an unnecessary variable.

With the addition of another transistor (usually in the form of using Sziklai configuration) the loop gain of the VBE multiplier circuit can be increased. The first transistor serves as the temperature sensor while the second transistor amplifies its effect. The technique can reduce the resistance slope (U versus I) of the multiplier but substantial results are gained only after the inclusion of current-compensation resistor (circuit c). The advantages are a high linearity regardless of input current and a broader tolerance for values of $R3$. Unfortunately, addition of any external, active component is always somewhat an invitation for instability issues. Because reliable operation of bias circuit is vital this is a major concern. Stability can be improved with some Miller compensation within the servo. VBE multipliers that use two transistors are quite rarely seen but have been used at least in many Roland amplifiers.

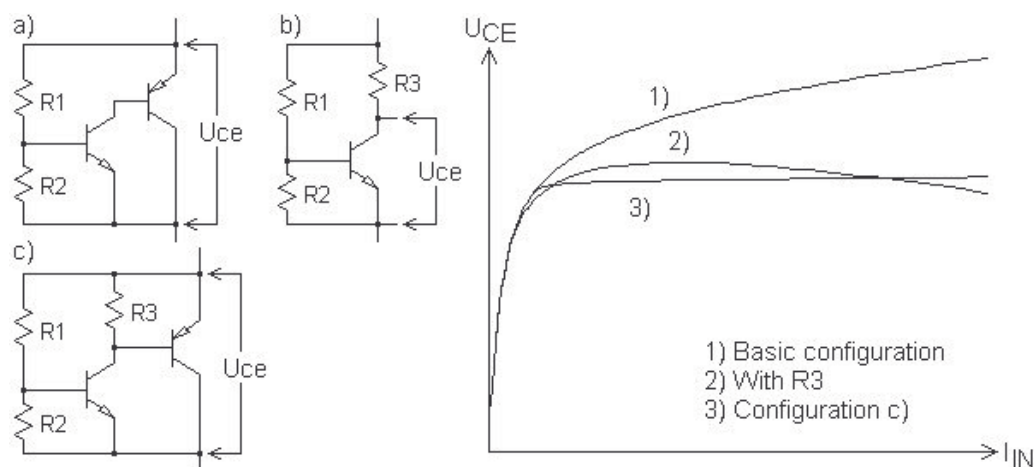


Figure 3.48 Enhanced versions of the VBE-multiplier circuit: a) Increased loop gain, b) “Current-compensation” with resistor $R3$ and c) combination of the aforementioned techniques. Right: The response of different VBE-multiplier variations in input current versus output voltage. Graph 1 shows the response of the conventional arrangement (shown in Figure 3.46).

When designing bias servos there are a couple important things that have to be taken into account. First one is to make sure the servo circuit works as planned with the selected transistor model. Many semi-power transistors have a low beta and therefore the basic design equation will not work with them. Also, *it is very important to ensure that the servo transistor is never allowed to close completely*; otherwise the power transistors will start to conduct simultaneously and overheat. Should R1 become infinite (or just reasonably high) this will happen. Same thing applies when resistance R2 becomes too low. In servo circuits that include trimmer resistors the possibility of a defect in the form of lifted wiper or resistance-introducing corrosion has to be accounted either by using series resistors or parallel diodes. It is recommended to have the trimmer as part of R2 instead of R1 since the first configuration can safely develop a high-resistance-type fault. Trimmer will also require a series resistor that limits the total resistance (of R2) to a certain minimum value when the trimmer is set to zero resistance.

Never replace R1 and R2 with a single trimmer or use the trimmer's wiper as the only contact to transistor's base: This is the worst and the most unreliable design that one could possibly have. Yes, you will likely see it used in many amplifiers but that fact doesn't make it any more reliable. For reference, faults where the trimmer's resistive strip breaks (thus creating infinite resistance) or the wiper lifts off from the track (another infinite resistance) are *very* common. In fact, the wiper can even lift off during adjustments and form a contact again when the tension from the actuator is removed. In a nutshell, the mechanical part of the VBE multiplier circuit should always have a backup because mechanical faults are the ones occurring most often. Naturally, one should use a high-quality, precision trimmer that is preferably multi-turn type.

Sometimes one can see designs where the thermal sensing transistor is mounted in contact with the plastic case of one of the output transistors. My general advice of these kinds of temperature tracking methods is: *Do not do this!* Although the solution may seem compact, the plastic case is (sometimes) just about the worst place to track temperature. The case has a notable thermal resistance and stays substantially cooler than the heat sink. Therefore, regardless of "temperature lag", a heat sink-mounted sensing device usually tracks the temperature of the transistor's die more accurately.

There is nothing wrong in mounting the sense transistor to be in contact with one of the output devices in general – as a matter of fact, it is even recommended: Essentially, *the effectiveness of the bias servo depends more on creating an optimal mechanical layout than an optimal electric circuit*. With metal packages, like TO3, the aforementioned configuration is usually about the best but when it comes to plastic packages do make sure that the configuration works as it's supposed to: That is, minimizing the thermal resistance and temperature delay between the sensing element and power transistors. An alternative way is to mount the sensing device so that it faces the power transistor on the opposite side of the heat sink. With modern, plastic-cased output transistor models this is probably the most practical and effective solution as the thermal lag is reduced to the time it takes for the heat to spread through the bottom plate of the heat sink. Then, of course, there are exceptions: For example, experiments have proved that TO3P and TO225AA packages get hotter from their plastic side [3]. It is also claimed that at least BD139 in TO-126 case tracks temperature better from its plastic side. I have not verified this myself and I dearly

suggest that you stay away from using any odd configuration unless you are hundred percent sure that it works.

Also, it has to be noted that different output topologies require different temperature tracking schemes: In the Darlington complementary circuit the temperature of the output transistors is tracked, in Sziklai complementary circuit the temperature of the driver transistors. This is also a “layout” issue: Darlington complementary circuit benefits from having its drivers thermally coupled to output devices – Sziklai complementary circuit does not. So, with Sziklai compound follower topology it is advisable to use separate heat sinks for the drivers and for the output devices. In Darlington configuration, the drivers and output devices should share the same heatsink. The difference between these two topologies makes the temperature tracking and stabilization of a Quasi-complementary circuit quite difficult since it effectively combines both. In such case, commonly the N-channel is tracked.

3.8 Constant current source loads

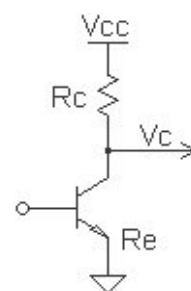
There are a few bottlenecks in a power amplifier circuit’s operation that can be cured by using constant current sources in the right places. Let’s take a quick view of these shortcomings and discuss how they can be avoided.

The gain of a simple common emitter amplifier is approximately equal to collector resistance divided by the transistor’s internal emitter resistance R_E . Basically the voltage at the collector can be calculated from the formula

$$V_C = V_{CC} - I_C R_C, \quad \text{(Equation 13)}$$

which is simply the subtraction of supply voltage and voltage drop over collector resistor. Despite its simplicity the formula shows two very important things: 1. Fluctuations in the supply voltage show up in the collector voltage and 2. Lower voltage drop over a constant resistance means a lower collector current I_C . The latter means that as the transistor’s output voltage increases the collector current must decrease since the voltage difference between the collector and the supply becomes smaller. If the collector current decreases it means that the emitter current has to decrease as well and less emitter current means lower gain. The exact opposite happens when the output voltage decreases and as a result the waveform will become distorted having a flattened top half wave and a stretched bottom half wave.

This bottleneck can be cured by replacing the collector resistor with a constant current source (in short CCS) - as is done in the circuit portrayed in Figure 3.49. To understand how the circuit works let’s examine the operation of a basic, resistor-loaded common emitter amplifier first: In order to produce voltage gain the transistor’s collector current has to be converted into voltage, this happens when the current flows through a collector load. However, like it was concluded earlier, current flow through a resistor is not constant since the affecting supply voltage potential defines its amount. With a constant current source the operation is different: The CCS is seen as a constant high impedance load, through which the collector current remains constant. However, according to Ohm’s law, the base-emitter voltage fluctuation (such as signal input) must show up respectively in the collector voltage.



In theory, the circuit has now become free of limitations caused by current-through-resistor voltage conversion and should be immune to supply voltage changes. Exchanging the collector resistor for a CCS can provide a nice gain boost since the impedance of a constant current source is very high, (in theory it is infinite), and because the collector current does not decrease in relation to difference between supply- and collector voltage. In practice though, constant current sources are far from constant and can have a severe impact on bandwidth.

Figure 3.49 presents some real-life implementations of active constant current sources. The circuit has some added complexity but a closer inspection shows that even now things still remains fairly simple: The emitter loading of the long-tailed pair input stage is achieved by a constant current I_1 through a resistor. Since the voltage V_B at the transistor's base equals V_{CC} minus the voltage drop over the LED, the emitter voltage must equal V_B plus the voltage drop over the base-emitter junction. We now know the current through the resistor; it is V_{CC} minus the emitter voltage divided by the emitter resistor's value. The same current must also flow through the collector load, which in this case is the differential circuit. The purpose of the base-to-ground resistor is to provide the important voltage reference and to limit the LED current. The LED could be replaced with another similar active device as well, such as with a diode string. However, a current flowing through the LED will warm it up, thus its forward voltage will be less dependent on ambient temperature changes.

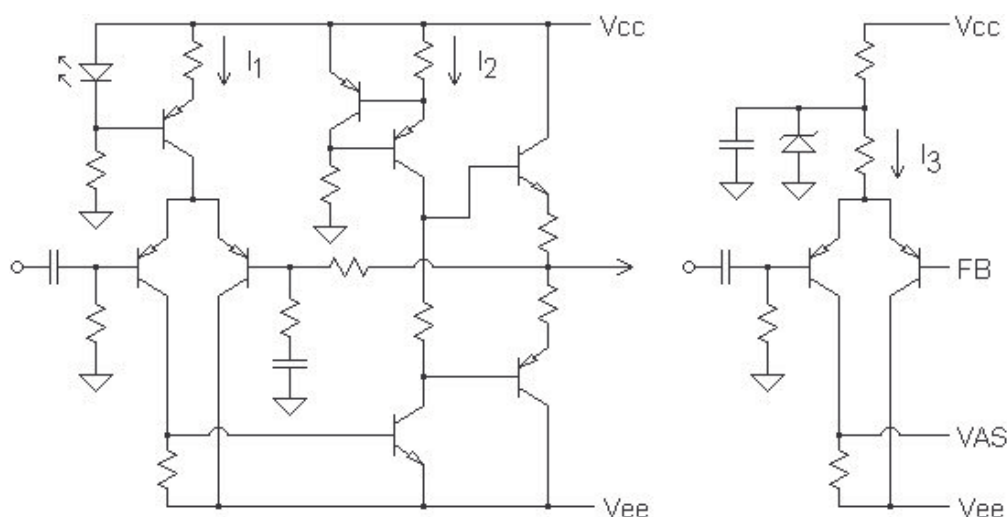


Figure 3.49 Constant current source loads.

Constant current loading of the VAS is another form of the same circuit and has an excellent stability – especially if the two transistors are thermally coupled. The current I_2 will always be the transistor's forward voltage divided by the emitter resistor's value (they are in parallel). The internal emitter resistance is not an issue since the second transistor controls the base current keeping the voltage drop at a steady 0.65 V (or the specific amount of forward voltage for the concerned component).

Shown far right in Figure 3.49 is a circuit that Slone [43] calls a “pseudo-constant” current source. It is essentially a passive circuit that relies on voltage regulation over the zener diode. Since the voltage at the center node between resistors is regulated the current over the bottom resistor should not vary due to rail voltage fluctuations. Keyword here is *should* and compared to circuit solutions presented earlier the circuit is definitely not an ideal one. Yet, it is fairly simple and offers perfectly acceptable performance in most applications. One should note that even the “pseudo-constant”

circuit is way more advanced than most loading solutions found from older guitar amplifiers – even from the ones that are considered excellent.

When relatively high current draw modulates the supply rails of an amplifier it unfortunately affects a realistic CCS circuit as well. Thus, a basic CCS can be improved by regulating its reference. However, this is not as straightforward as it might seem at first: A logical solution would be to filter the reference node conventionally by capacitively shunting AC signals to ground. However, as seen from graph b) in Figure 3.50, this actually makes matters worse. In fact, the circuit would not even work if the capacitor would connect the voltage reference node directly. What is happening here?

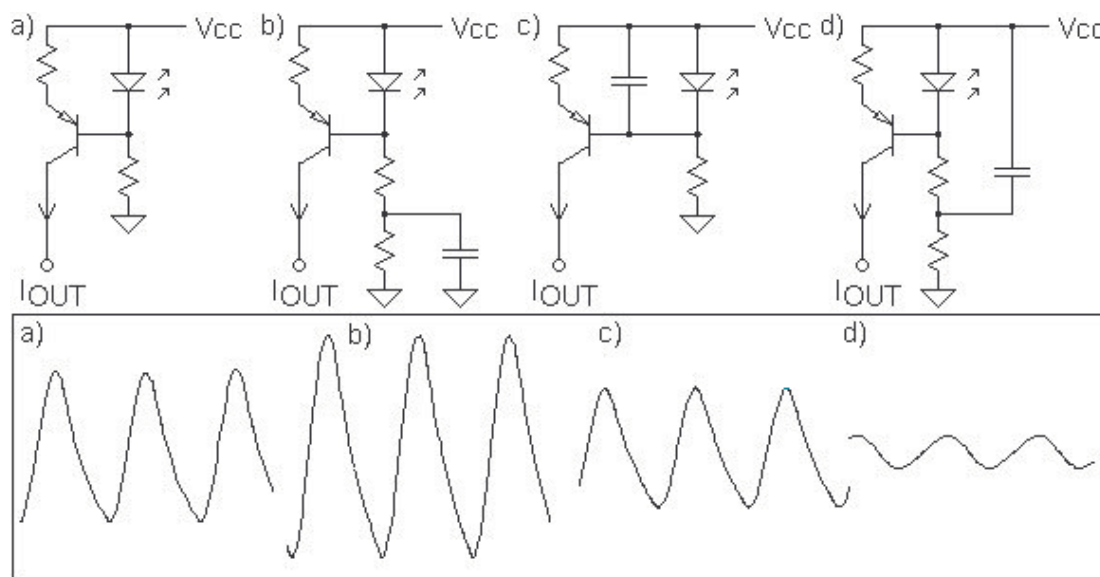


Figure 3.50 Various methods of regulating the CCS reference voltage. Shown below is an “in-scale” plot of I_{OUT} of the different circuits.

The behavior becomes rational when the circuit is inspected from another perspective: The steadiness of current flow through the LED (which determines the amount of “reference” voltage drop) is a key part in forming a constant reference. If the current limiting resistor is *bypassed* with a capacitor it results into lower resistance for AC signals, which consequently *increases AC current flow through the LED*. This means a setback in performance. In order to decrease fluctuation of the LED current the supply rail and the reference potential can be coupled to the same AC potential (circuit c). However, this provides only slightly better performance than the starting point with no regulation at all. Why: Because the transistor’s base is still at the path of AC current flow from supply rail to common. This current flow also modulates the base current. An effective solution is presented in circuit d): The AC source is coupled to the interconnecting node of the current limiting resistors, which effectively minimizes the AC potential across the circuit. Thus, I_{OUT} remains extremely constant.

3.9 Bootstrap constant current sources

Bootstrap was once perhaps the most common form of constant current circuits and it is still found from the voltage amplifier stages of many amplifiers. However, the operating principle of this passive CCS circuit is quite unclear to most people: Two resistors serve as a collector load for the transistor driving a buffer stage. In an ideally operating amplifier the bootstrap capacitor would couple the collector’s AC signal *buffered* back to the joint node of the two resistors resulting in a constant DC voltage

drop across the “lower” resistor. According to Ohm’s law, if the voltage drop over resistor remains constant then the current must remain constant as well. In a realistic amplifier the emitter follower stage will of course introduce a voltage drop over VBE-junction and some other minor non-linearities. The capacitor will also have a different resistance at different frequencies. Figure 3.51 depicts the concepts behind bootstrap loading. The example circuit is a typical common emitter amplifier coupled to a class-A emitter follower buffer that separates the load from the voltage amplifier stage. In a typical power amplifier this simple buffer is of course replaced with a push-pull kind.

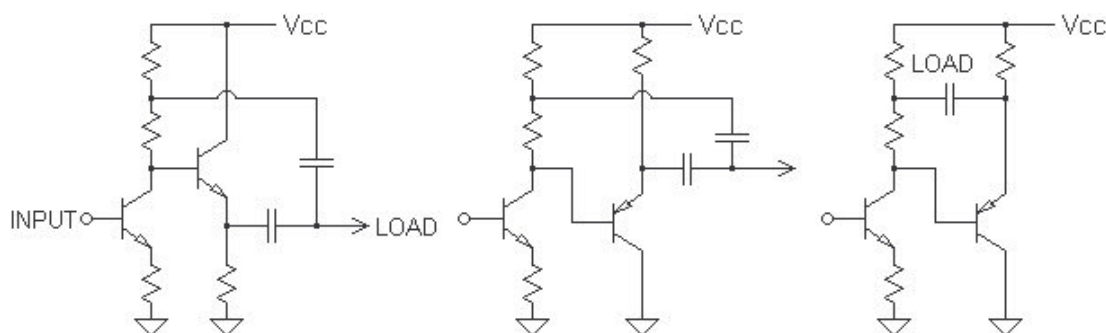


Figure 3.51 “Bootstrap” constant current source.

One reason for misunderstanding the quite simple operation of the bootstrap CCS is the inability to see the capacitor operating as a voltage or current pump: The node between the resistors has a steady DC potential but it is modulated by the AC output signal swing. During a positive half wave in output signal the negative terminal side of the bootstrap capacitor (the output node) swings up in voltage; this potential change is naturally coupled to the positive terminal, which, on the other hand, is connected to the joining node of the two bootstrap resistors. Since the node has a steady DC potential the AC signal can even lift the node’s voltage potential higher than what the supply rail voltage is. Thus a condition where the stage current would be defined by the difference of collector (output) signal and rail potential is avoided because the current pumping effect of the bootstrap capacitor assures that the (collector) current through the lower resistor is “always” kept constant. As expected, in real-life the circuit doesn’t provide a very constant current but it does operate sufficiently enough to be extremely useful.

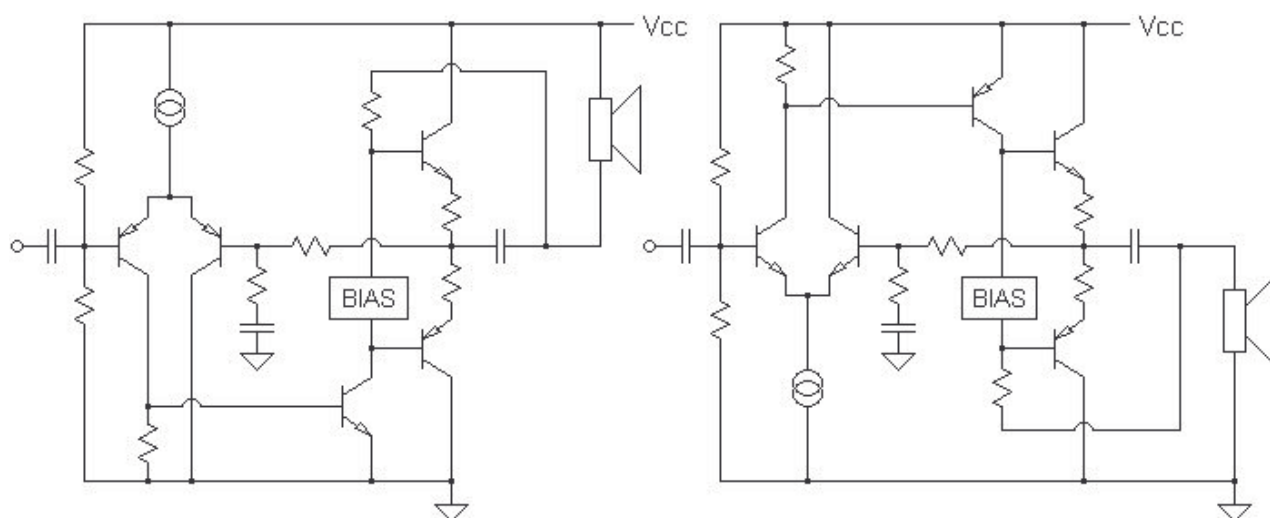


Figure 3.52 Alternative bootstrapping methods.

The circuits in Figure 3.52 present an alternative way to bootstrap the voltage amplifier stage. In this arrangement the load itself acts as the “upper” resistor; the

benefit is “eliminating” the bootstrap capacitor since the output coupling capacitor now acts as one. This arrangement was also shown simplified far right in the Figure 3.51. The speaker can be hooked to supply rail safely since the DC voltage drop across it will be only few millivolts. The topology works also in dual supply circuits - assuming one adds a coupling capacitor to the output, of course – which in fact is not that good idea as the conventional circuit is cheaper, less bulky and likely creates a lower amount of distortion as well.

This rather “unconventional” bootstrapping topology has been used in numerous, cheap portable radios of the era when low-power amplifiers were still made out of discrete components and using single-supply arrangement. *Cheap* is the keyword here since the topology can eliminate the bootstrap capacitor and one resistor. I have never seen this topology used in high or medium power amplifiers – at least in professional applications. There are perfectly reasonable explanations for this: In the configuration shown at left the speaker is referenced to the supply, which *introduces a substantial ripple component to the output* unless the rails are well filtered. The circuit at right doesn’t have this problem but a SPICE simulation revealed another nasty quirk in its performance: The conventional bootstrapping arrangement tended to be more stable if the amplifier was driven into clipping. The alternative arrangement introduced some ringing and the problem persisted even with a purely resistive load.

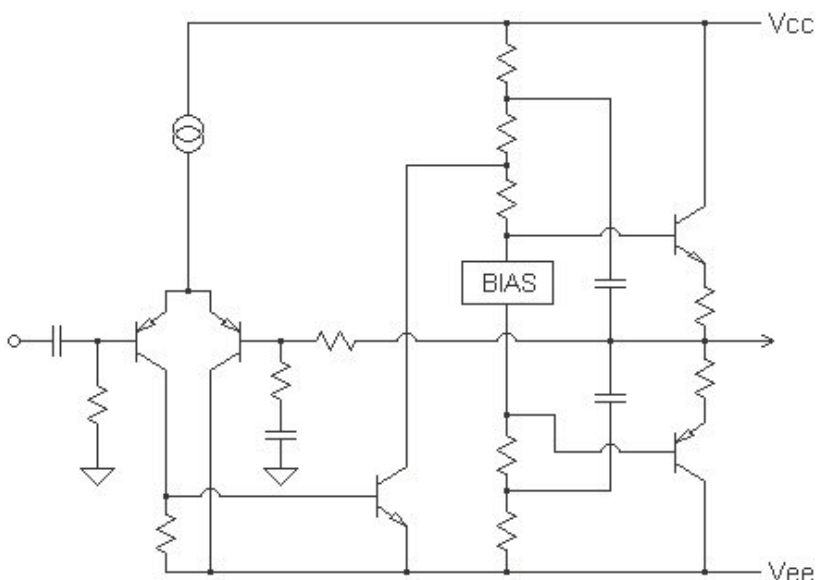


Figure 3.53 Symmetric bootstrapping.

A very rare but novel configuration is the symmetrically bootstrapped circuit shown in Figure 3.53. Here the VAS load is divided to conventional bootstrap CCS in parallel with the bias circuit and another bootstrap CCS for the PNP side. While in conventional, single bootstrap topology nearly all current sourced by the CCS is flowing through the VAS transistor, in dual bootstrap topology the current through the transistor is only the *difference* of the two CCS currents. As a result the VAS uses less current (and therefore dissipates less power) but can still drive the output stage with higher current, which in turn increases slew rate. Also, devices with lower power dissipation capacity are generally faster. A “modern” upgrade would be to replace the bootstraps with active circuits.

The “symmetric” topology also allows the output to swing a bit closer to negative rail. Even though the increased swing is in the range of *about one volt* (with BJTs) it can usually make a huge difference at the concerned rail voltage levels. For example, 24

volts over 4-ohm load gives 144 watts while 25 volts over the same load results to 156.3 W. The extra volt creates a power increase of about 12 watts! The increase in power naturally becomes higher the higher the rail voltages are. Maximizing ability to swing close to rails will become extremely important if the amplifier uses MOSFET output devices, which are generally quite poor in this regard. Unfortunately, the concerned configuration may also introduce ringing or “glitches” when the negative half wave clips. These are result from constant current loading being more susceptible to “rail sticking” phenomenon. Evidently, the design causes this to be a greater concern with the negative half wave proportion of the signal as well. The topology never caught huge success and if symmetry was required the ordinary choice was to resort into using a symmetric cascode circuit instead. The only guitar amplifiers I’ve seen utilizing double bootstraps are some Acoustic Control Corporation models from the 70’s.

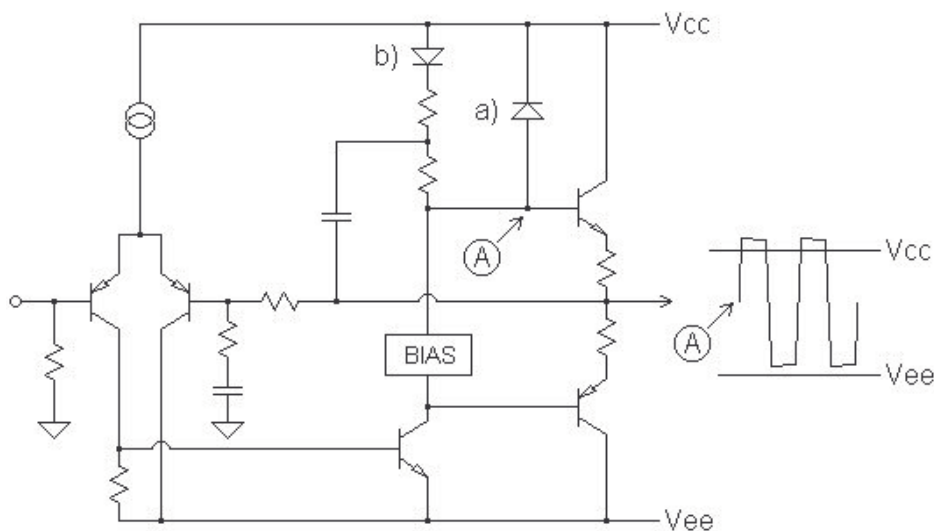


Figure 3.54 Enhancements to bootstrap circuit: a) Diode limiter and b) isolation diode.

The basic bootstrap circuit can be improved with few simple additions: As is known, the bootstrapping effect allows the voltage potential at the interconnecting node of the two “bootstrap resistors” to rise above the rail voltage. However, the very same useful phenomenon also causes a less-ideal side effect: When the positive half wave is driven into clipping the voltage potential at the driver transistor’s base (node A) exceeds that of its collector, which in turn is tied to the supply rail. This “bottoming” causes a *storage effect* that slows down the transistor and leads to cross conduction. The problem can be fought with a diode limiter (a) that limits the maximum base voltage to a level that is one diode forward voltage above the supply rail potential. This simple addition drastically decreases the amount of cross conduction (at least in some circuits) and also improves recovery from clipping. Naturally the diode should have a very low forward voltage (i.e. use a germanium or a Schottky diode).

During positive signal peaks the voltage potential at the interconnecting node of the two “bootstrap” resistors rises above the supply voltage. Accordingly, this phenomenon is what allows the bootstrap circuit to supply constant current in the first place. However, in the typical bootstrapping configuration some of this current is actually bled to the power supply circuit (that has a lower voltage potential than the capacitor plate). The loss of this current leads to non-linear operation of the amplifier when it’s amplifying transient signals. Discharging of the capacitor can be prevented with the utilization of an isolation diode (b). The diode blocks the conduction path

between the power supply and the bootstrap capacitor whenever the potential difference between them exceeds the level defined by the diode.

3.10 Cascoding and rail switching

Increasing the rail voltage is often a practical way to increase the output power capability of a design. However, there is a reasonable limit for this, which is the transistor's maximum collector-emitter voltage. By connecting transistors in *series* they can be made to share the rail voltage. Configuration like this is known as “stacking” or *cas coding*, in the older days people also referred to it as “beanstalk”. K.J. Dean introduced the totem pole cascode circuit in 1964 as a solution to replace expensive high-voltage transistors that were used in driving cathode ray tubes. Naturally the series connection allows the voltage swing across the output to divide (evenly) across each transistor. The topology requires close component matching to retain the (usually even) division of voltage and a failure to do this can create catastrophic results. The capacitors are optional but often required with BJTs to increase the high frequency response. In modern implementations currents sources and Zener diodes may replace the resistor dividers.

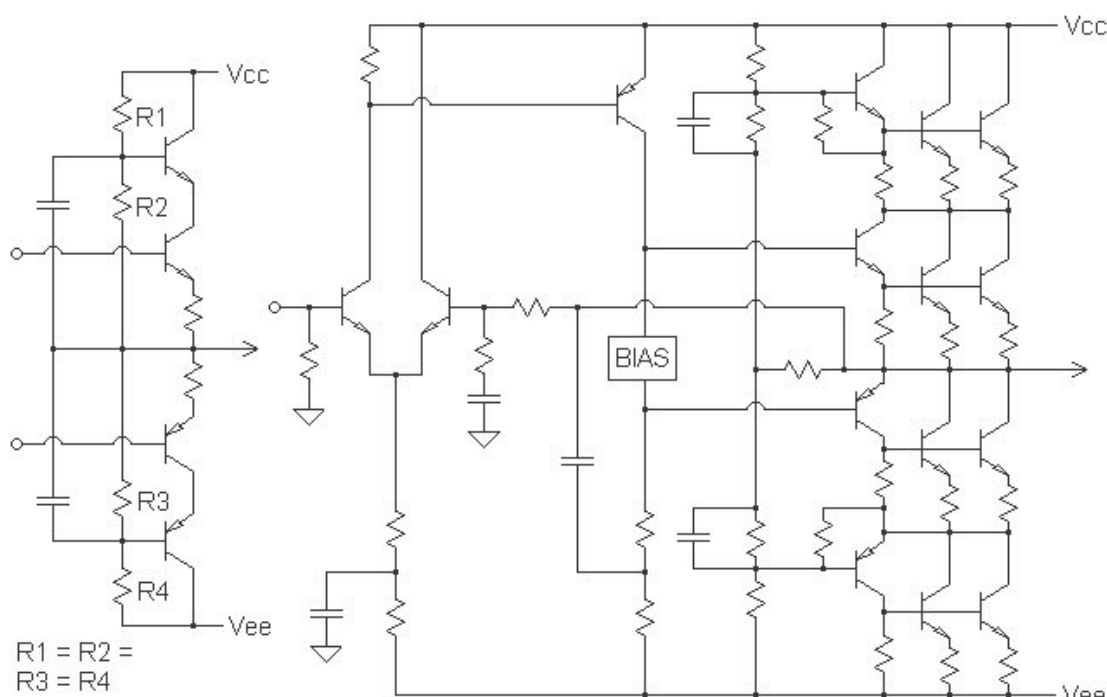


Figure 3.55 Cascoded output stage. Left: Conceptual schematic and right: Typical early 70's GMT-style quasi-complementary cascode circuit.

In a cascode circuit the other power device works as a common base amplifier and the driving device can be operated in any desired configuration, which in the case of output stage is usually an emitter follower. The common base amplifier sees most of the voltage swing, however, the arrangement allows using a lower voltage rated, more linear (and faster) device as the driver transistor. Since the common base amplifier is driven by a typical common emitter circuit the output stage can be protected with an ordinary VI limiter configuration.

Rail switching amplifiers, which were introduced by Hitachi in 1976, are also known as class-G amplifiers. They combine a higher and lower voltage rails making the amplifier draw power from either one according to demand set by the output signal level. The objective is naturally to increase the efficiency by reducing the amount of

waste power that is rather high if quiescent current flows from a high-voltage supply. The higher rail voltage amplifier is initially in standby but becomes active to handle large signal peaks and transients. Since majority of audio signal content is below peak levels the class-G can operate very efficiently. The switching between stages is handled with so called “commutating diodes”.

The voltage ratio between rails does not really matter much. Basically, it’s a compromise between price tag and efficiency/performance. Equal “spacing” between rail voltages allows using same secondaries and capacitors with same voltage rating, which can radically decrease overall cost. Unequal spacing, on the other hand, has higher cost but (usually) allows designs with higher voltage swings and better handling of reactive voltage.

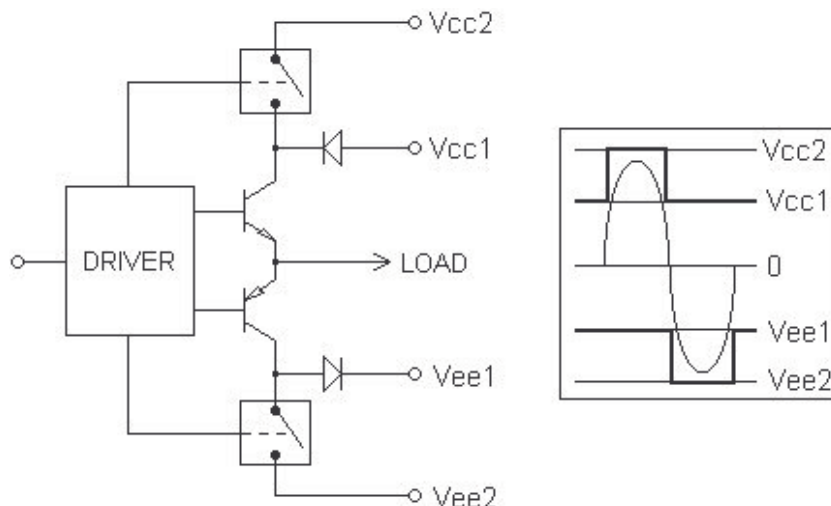


Figure 3.56 Conceptual schematic of class-G rail switching amplifier.

Let’s examine the operation of class-G amplifier on the positive rail: The “inner” device TR4 operates in conventional class-B and TR2 is its usual driver. The circuit also uses the conventional bias circuitry (Bias 1), which is usually the ordinary VBE-multiplier tracking the temperature of the “inner” transistors. Power for the lower supply is drawn through commutating diode D2. When the input signal exceeds the lower rail voltage D1 conducts turning on the “outer” driver TR1 and the class-C transistor TR3. Simultaneously D2 turns off and the whole output current is drawn from the higher voltage supply. The additional bias circuit (Bias 2) is required to turn TR3 on before TR4 runs out of collector voltage. This bias circuit can be, for example, an ordinary zener diode. The negative side works in a similar manner.

The biasing techniques differ from one class G “topology” to another. The original Hitachi circuit put zener diodes in series with the signal path that feeds the inner drivers [3]. Essentially they *subtract* bias from the main bias “generator”, which in turn was biased to a higher voltage than what conventional setups used. However, this design has a very poor thermal stability and linearity, thus it has been recommended by Self [3] to move the additional bias generators to the forward path of outer devices. This allows the bias generator Bias 1 to be a standard type and also prevents the external bias generators from having effect on quiescent conditions [3]. Powering both drivers from the outer rails prevents sudden changes from taking place in the rate in which the driver’s V_{CE} varies. This in turn reduces the effects of “gain steps”, which mean a sudden and non-gradual change of gain occurring at the transition between rails [3]. The gain stepping is one of the inherent non-linearities that troubles class-G amplifiers.

Rail switching circuits never caught huge success: Compared to its benefits the complexity of the topology is found appalling by most and not much information concerning the design aspects has been published. Additionally, some drawbacks make the overall performance inferior to ordinary class-AB amplifiers: The “gain stepping” was already mentioned but there are other quirks as well. For example, switching takes a finite time therefore creating commutating distortion (which is sort of like crossover distortion). Furthermore, the charge storage in commutation diodes creates “glitches” to the signal. Most of the latter can be removed by using fast, Schottky diodes, though. Yet, question remains how to switch cleanly at higher frequencies: A special case is *predictive* rail switching: Mostly this means switching the higher rail on (or off) when the output is within few volts of the limit of the lower rail. However, if the switching arrangement is designed poorly the commutation distortion is still present. High frequency signals together with predictive switching also decrease the overall efficiency. Some compensation methods have been developed to fight this phenomenon. Most common one is boosting up the amplitude of higher frequencies in the signal sent to the voltage sensing mechanism.

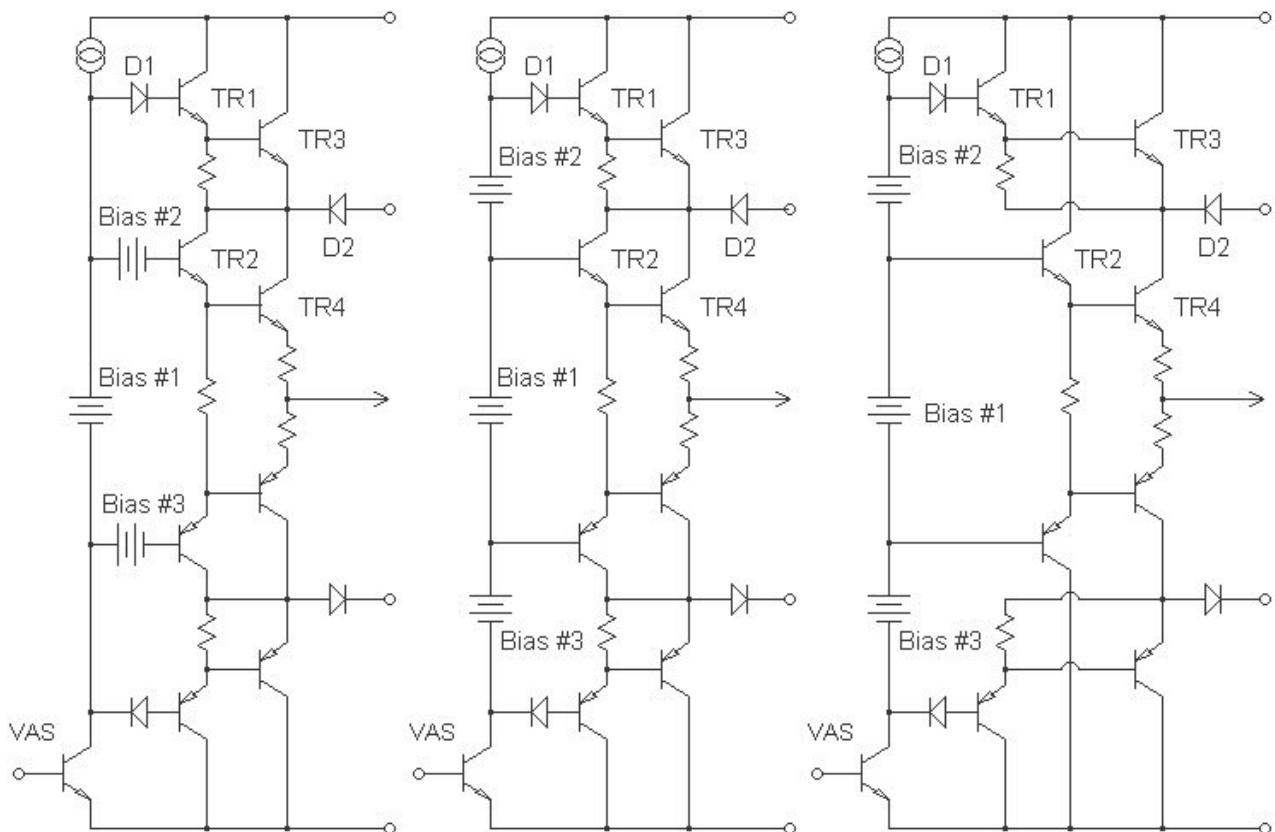


Figure 3.57 Practical rail switching arrangements. Left: Hitachi, middle: Series class-G and right: Drivers powered from outer rails.

Usually utilizing the rail-switching configuration instead of an ordinary class-AB becomes a concern when one wishes to extract output power figures greater than 500 watts from the design. The topology is also fairly popular in bass and subwoofer amplifiers. Class-H takes the rail-switching concept even further creating an “infinite number” of supply rails. Basically, to achieve this the rail voltage is modulated so that it is only few volts higher than the output signal at any given time. *Note: In USA, the meaning of class-G -and H definitions is reversed!*

3.11 Bridged, parallel and grounded output circuits

More output power can be extracted by pairing two similar amplifiers. The most common form of such pairing is a stereo amplifier that utilizes two identical power amplifiers - each amplifying the signal of its corresponding channel individually. Some PA systems can offer even a higher amount of channels than the ordinary two.

Bridged and parallel configurations exploit the power handling capability of multiple amplifier circuits but combine their power feeding capacity to a single channel and a single load. Bridged and parallel circuits see the load differently and this affects their behavior quite drastically: Since in parallel configuration two circuits share the load, each of them sees a *doubled* load impedance. Since each amplifier module also provides half of the total output power the dissipation per module is halved as well. If more parallel amplifiers are added up (yes, this can be done as well) the load impedance will rise equally among them – that assuming the circuits are *identical* of course. To even out circuit indifferences the paralleled circuits must utilize close-tolerance and high-power *current sharing resistors* in the output [9]. The multiplication of load impedance is a practical matter that undoubtedly sets the limit for both number of paralleled circuits and for the load impedance.

A bridged configuration generates a higher voltage potential over the load since it is effectively driven with two opposite-phase signals. This arrangement is also known as “Bridge-Tied-load” (BTL) and while it’s used each amplifier module sees *halved* load impedance. Because the voltage potential over the load is two times higher than what a single amplifier module could produce, and because the load impedance is halved the output power is quadrupled. Note that grounding the speaker from one end will destroy at least one amplifier module (unless it’s protected somehow) so special care is required to make sure this will never happen. Also, mixed mode feedback schemes are harder to implement into bridged amplifier designs. The benefit of the topology is that one can quite easily convert a typical stereo amplifier to a bridged mono amplifier; practically the only requirement is inversion of the input signal in the other channel. Such systems are known as “Tri-Mode” or “Multi-Mode” amplifiers.

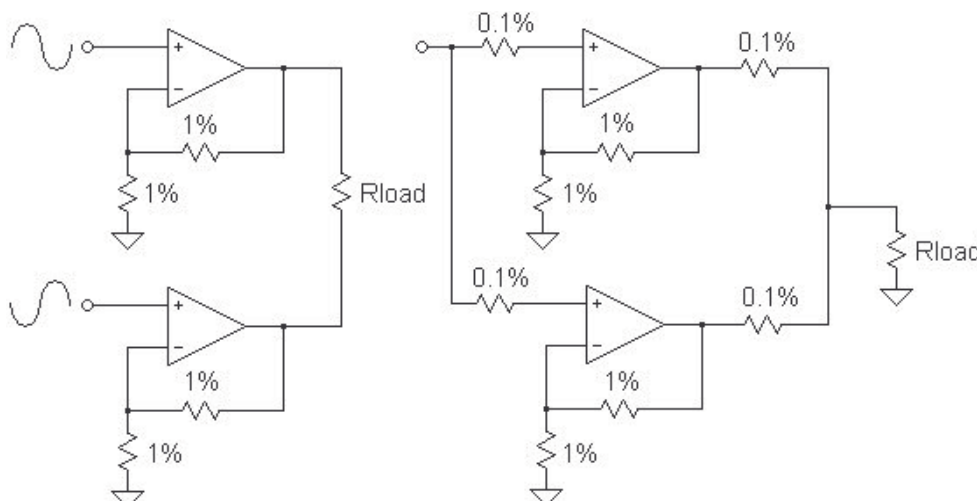


Figure 3.58 Bridged and parallel circuits.

In practice there are no identical circuits and expensive precision resistors have to be used to even out circuit differences. Any divergence in the operation will create heavy loading to one or several of the circuits. The behavior is similar as in unevenly balanced parallel transistors: The transistor with lowest internal resistance will pass

most of the circuit's current. In bridged circuits, the imbalance causes the other amplifier to push power into the other – which effectively has to be dissipated somehow. Because proper circuit balancing with real-life component tolerances is either expensive or difficult, bridge and parallel circuits are notorious for being quite unstable and running hot.

The typical bridging configuration is in more specific terms known as *bridging channels*. Sometimes a term “H-bridge” is used to describe the same thing but this can be a bit misleading since many people who use the term use it incorrectly: H-bridge derives its name from a (schematic) configuration where four switching elements (like BJTs) are located in the “corners” of “H” and the load forms the “horizontal cross bar”. Most known usage for the circuit is controlling direction of a motor. Since such circuit can also be called “full bridge” (which in audio circuitry means a completely different setup) we have a certified semantics problem. Essentially, we better just stick in the “bridged channels” term.

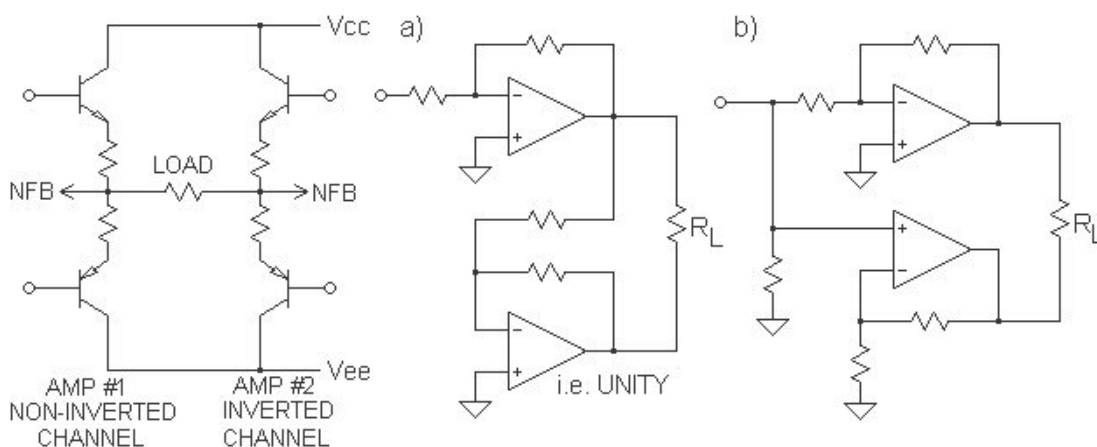


Figure 3.59 Amplifier with bridged channels (or “H-bridge”) and practical ways to realize it (shown simplified): a) tandem and b) non-cascaded configuration.

There are practically two common configurations for creating a bridged channel amplifier: A cascade or “tandem” and a non-cascaded configuration. These are shown in a simplified form in Figure 3.59. Essentially each operational amplifier stage merely represents a complex amplifier circuit. For example, the inverting amplifier may practically consist of an inverting OpAmp stage, which is then followed by an ordinary non-inverting power amplifier. It should be noted that in tandem configuration the first amplifier dominates the overall performance because its output signal (including distortion and noise) is replicated in the second amplifier. The non-cascaded configuration eliminates this drawback and is often more straightforward to realize as well. It is undoubtedly the most commonly used arrangement.

Bridging channels is by no means the only way to create a bridged circuit. Another form is the single-supply bridged amplifier (or “full bridge”) depicted in Figure 3.60. It is pretty similar to push-pull grounded emitter circuit, except that its output stage is an emitter follower circuit. The topology has some interesting features: First of all, it is a bridged amplifier with all the benefits included (voltage swing is “doubled”, power output is quadrupled), secondly, it acts like a dual supply circuit, which allows omitting the output coupling capacitor. Variations of the idea have been used successfully, for example, in Sunn Concert Lead amplifier or in Crown’s patented “Grounded Bridge” topology amplifiers. “Because the power supply is not ground referenced, it is allowed to slide back and forth in a controlled manner” [7]. The “controlled manner” is achieved with another power amplifier circuit, which controls

the potential of supply rails according to drive signal. The output stage can also use an interstage transformer and a totem pole configuration (like the Sunn Concert Lead did), which significantly simplifies the circuit arrangement but also leads to lower damping factor. The downside of the circuit is that it requires another voltage source (secondary winding) to power the rest of the circuitry since any imbalance in power rail loading would lead to serious instability or even destruction of the output stage. It is possible to further bridge two full-bridged circuits together as shown in the lower schematic of Figure 3.60.

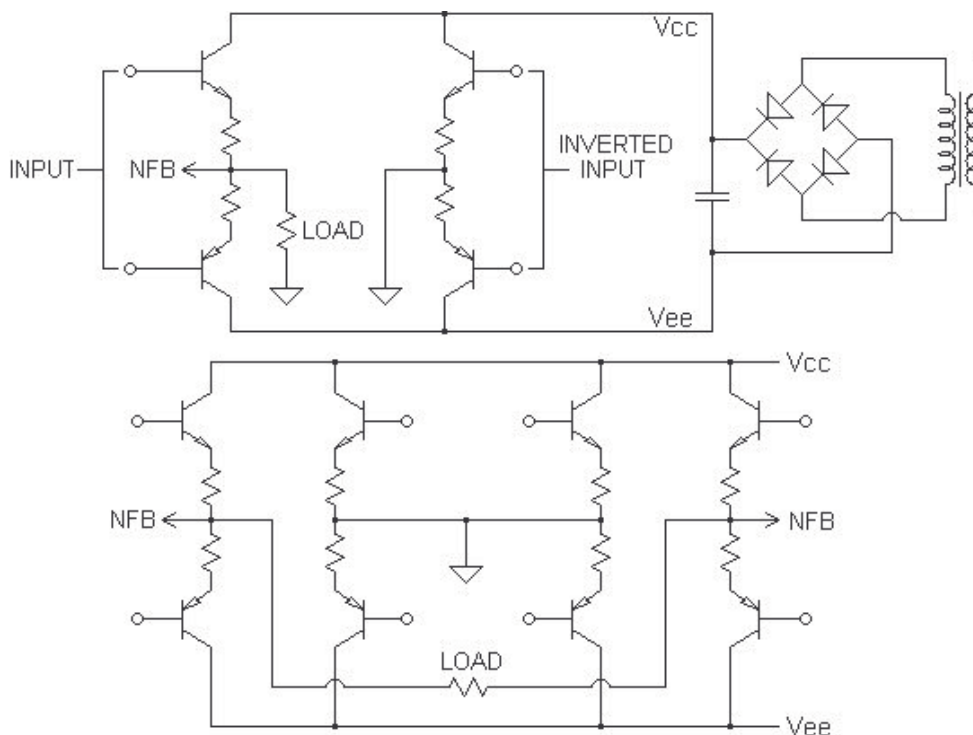


Figure 3.60 Full-bridged amplifier (top) and full-bridged bridge amplifier (bottom).

The circuit shown in Figure 3.61 is a push-pull *grounded emitter* amplifier. Although a different look may deceive at first it is essentially just a basic emitter follower circuit. Controlling the voltage potential of mains transformer's secondary center tap drives the speaker. Since the center tap is no longer directly tied to the ground reference, *the output transistors modulate the rail voltage*. Basically, the audio signal rides atop the supply rails and is coupled to the load via the supply filter capacitors. Because the output is taken from the collector (as opposed to emitter) the signal becomes inverted. Consequently, the input stage must be inverted as well: The inverting input effectively becomes the non-inverting and vice versa. The configuration has been used, for example, in some Fender and Peavey models plus in the Hafler's Transnova series of amplifiers.

In the late 70's QSC Audio Products Inc. (formerly Quilter Sound Company) invented an arrangement that cleverly exploited the grounded emitter circuit: The output stage is configured as compound follower thus emitters of the power transistors are connected to supply rails (via proper ballast resistors). Consequently this allows connecting collectors straight to ground. Since the case of a typical transistor package is (often) connected to the collector, the output devices can be mounted directly to the (grounded) heat sink without any insulation. This provides a major improvement in heat transfer. The circuit was logically named as "grounded collector" - or "AC coupled amplifier". QSC claims that the topology was patented in 1978; however, even a persistent search provided no reference of such patent ever existing.

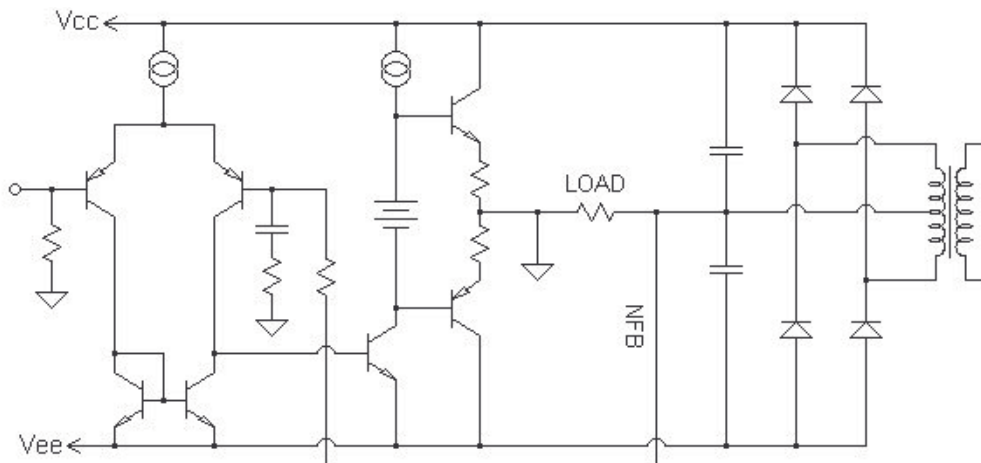


Figure 3.61 A grounded emitter amplifier. Note the “inverted” configuration of the input stage.

3.12 Class-A+ amplifiers

Since class-A amplifiers never turn off their output devices such amplifiers do not suffer from crossover distortion. Unfortunately they suffer from effects of poor efficiency. A push-pull configured class-A amplifier can theoretically have an efficiency of 50 %; nevertheless, it is a too poor rating if high output power is required. In the late seventies a topology known as “Class-A+” was invented to offer one solution to the efficiency issue. The main idea of the topology is to retain a very low voltage drop over the class-A devices, thus the losses are decreased significantly. Although I believe that Class-A+ likely has no potential use in guitar amplifying applications the topology is nonetheless so clever that it definitely must be discussed. If nothing else, it will surely provide some food for thought.

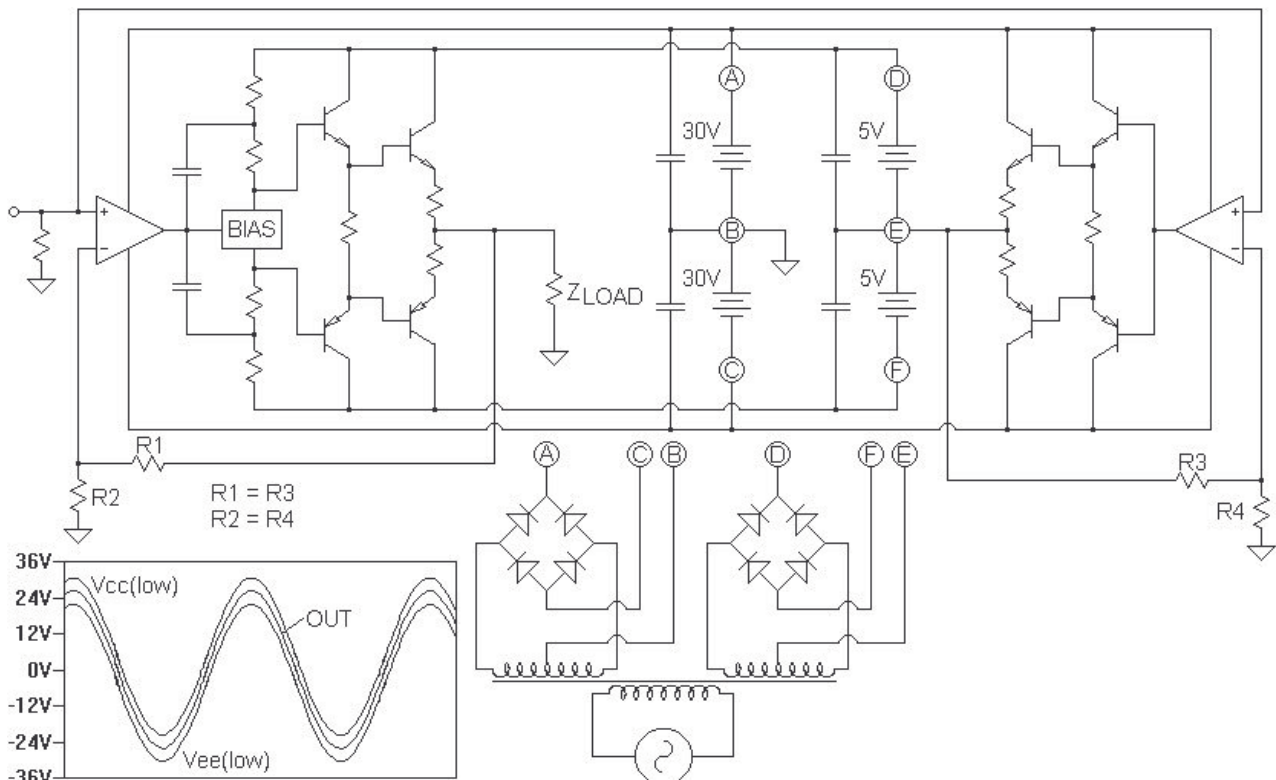


Figure 3.62 Conceptual schematic of a class-A+ amplifier.

An example of a class-A+ amplifier is shown in Figure 3.62. The circuit consists of two individual amplifiers; the leftmost operates in class-A but is powered by a very low rail voltage. The one on the right is biased to class-B operation and uses higher supply voltage. Voltage amplifier stages of both amplifiers are depicted as OpAmps for clarity, they have an equal voltage gain. (Do note that both of them are powered from dual 30-volt rails as well). In some circuit embodiments both class-A and class-B amplifiers may even share the same voltage amplifier stage. The class-A current amplifier runs from only ± 5 V rails. Normally, such a low rail voltage would not be compatible with the high output signal swing of the voltage amplifier stage, which would result into clipping. However, in class-A+ the common reference of the lower voltage supply is modulated by the output signal of the class-B stage, thus *the actual rail voltage is floating* and tracking the output signal. Thus a high signal swing can be amplified without clipping. A catch is that the voltage drop across the class-A stage is still only about ten volts (as oppose to 60 volts of the higher voltage rails). This means a tremendous reduction of losses.

Since current path of the class-A amplifier returns through the class-B stage the topology has been criticized for not being a true class-A. This is incorrect: The speaker is indeed driven by a true class-A amplifier, which just happens to have tracking rails. The crossover irregularities of the class-B stage are indeed coupled to the supply rail of the class-A stage (after being reduced by negative feedback of course) but this does not mean that the class-A amplifier would instantaneously pick them up. If power supply rejection ratio of the amplifier is good enough (what it usually is) only a little portion of these errors will “leak” to the output – just to become corrected by the negative feedback.

Naturally, a configuration as complex as this always has less-ideal properties as well: The class-B output stage must be able to provide rather high current, thus the circuit will have the losses of the class-B stage as well as the (decreased) losses of the class-A stage. In comparison to conventional class-AB this is still rather pitiful performance. Now, I leave it for the reader to judge whether the crossover distortion of class-AB is high enough to justify the need for class-A operation. Replacing the class-B stage with a high-efficiency switching amplifier can naturally decrease the losses but at the same time the complexity will also increase. Also, one must remember that the low voltage supply must be able to deliver an amount of current that *at least* equals the peak load current demand. This is a tough recommendation and means that a bulky, expensive and heavy transformer is required.

3.13 Amplifier stability – Preventing oscillation

All power amplifier circuits are prone to instability. It is mainly caused by both phase shifts introduced by the feedback loop and imperfections of real-life components. The most common result of neglecting these unwanted attributes is high frequency oscillation, which often masks parts of the audio signal (due to blocking distortion) and stresses the output stage with an abnormal loading condition. By introducing some simple, yet effective, circuits one can fight against the causes of instability (at least to some extend).

Circuit in part 1 of the Figure 3.63 is quite commonly preceding any amplifier stage that induces some gain. This simple low-pass RC filter, designed to filter out RF, is a basic building block of input signal conditioning circuits that often include a static voltage surge limiter as well. As is, the circuit is quite crude and works only if the RF

energy is reasonably small. This is really not as great problem as it may seem since the roll-off steepness can be increased easily by connecting more filters in series to form a higher order filter. As a further help, a series resistor following the filter stages would help to get rid of transistor's internal non-linearities that are the main cause for RF detection. Resistor must be placed as close to the transistor's base as possible in order to ensure it has any effect. In most cases the presented circuit, coupled with mounting the entire circuit, into metal chassis is enough. If the input resistor is left out the circuit (as is unfortunately done in some amplifiers) the circuit will not work very effectively. The capacitor may even further help to tune the transistor to catch radio frequencies.

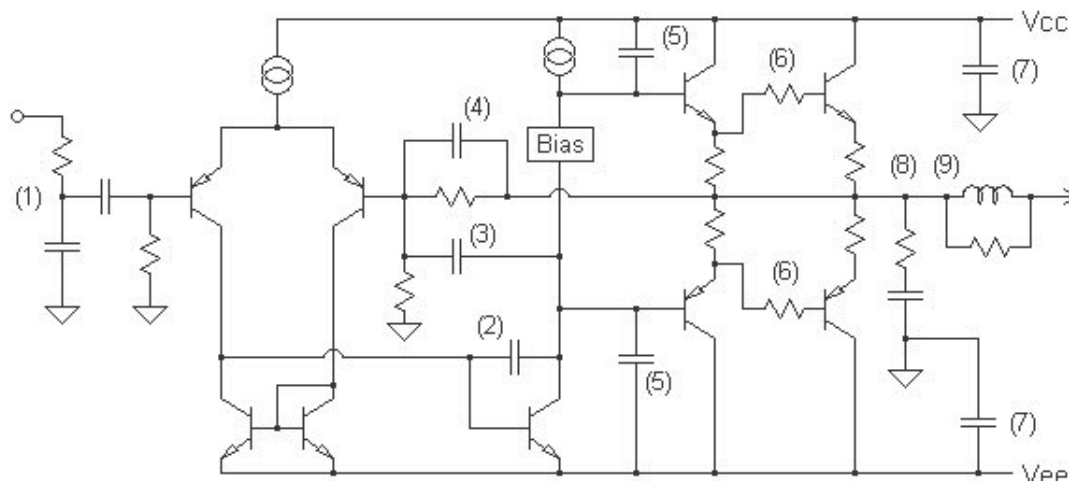


Figure 3.63 Enhancements for stability: 1 RF-filter, 2 Miller compensation, 3 HF “lead compensation” (feedforward), 4 HF compensation in feedback loop, 5 “snubbing” capacitors, 6 base resistors, 7 close proximity decoupling, 8 Zobel network (Boucherot cell) and 9 Output inductor filter (Zobel).

The HF feedback capacitors (2, 3 and 4) are used for decreasing gain at higher frequencies and therefore prevent oscillation. Most designs use only either two of them (2 and 4 or 3 and 4). The global feedback from VAS transistor's collector to base (“Miller capacitor” or “ C_{DOM} ”) is very effective: Basically, its capacitance can be ten times lower than in a case of conventional low-pass RC filter configuration where a capacitor shunts the transistor's base to ground. Also the resistor can be omitted. Especially long-tailed-pair circuits are prone to oscillation and usually never work without the addition of Miller capacitor. Some other VAS topologies, such as the compound feedback pair (CFP), are much more stable and require less high frequency NFB, if any. The drawback of the local compensation is the decrease in slew rate that creates distortion at higher frequencies (see 3.6.4).

If other compensation schemes are used as well the amount of required Miller compensation can be decreased – thus the slew rate is increased. Bypassing the feedback resistor with a capacitor (4) is a conventional method but the capacitor should be selected carefully; values above 47 pF usually cause oscillation instead of preventing it. 22 pF is about correct. The feedforward capacitor of the “lead compensation” (3) essentially works like the capacitor #4 but is free from the phase shifts introduced by the output stage. Very often these capacitors are also equipped with series resistors that stop the gain from dropping too low at higher frequencies, which would again cause instability. The effects of the configuration #3 are also covered in 3.6.4.

Capacitors connected from driver transistors bases to supply rails (5) also attenuate high frequencies, however, these “snubbing” capacitors are not very effective by themselves. More efficient way is to connect low value resistors to driver collectors and exploit the Miller capacitance. Same resistors would also improve clipping recovery. Other similar configuration shunts HF signals from the base to ground with a combination of a capacitor and a series resistor. All of these “snubbing” configurations may help to *dampen* high frequency oscillation but are generally considered as a poor fix since they limit bandwidth, severely decrease slew rate and introduce phase shifts. Needless to say, there are better ways to fix instability issues than any of these methods: It is usually more profitable to concentrate on the real source of problem and improve the layout and PCB design. Amplifiers can be made to work without these components!

Base resistors (6) in output stages *that have parallel transistors* act as “base stoppers” isolating the base capacitances - thus reducing the capacitive spiking (“overshoot”) and possibility of high frequency oscillation [43]. The effectiveness of base stoppers with BJTs (that have a low base capacitance) is highly questionable, however, in output stages with paralleled MOSFETs they are practically mandatory. Some manufacturers include them to BJT stages anyway - just to be safe. I advice that you create slots for these resistors in the PCB but initially jumper them with wire. Practical values range from 10 to 100 ohms. Base stoppers should not be confused with ordinary base resistors (used in non-parallel configurations). These are used merely to increase the input impedance of the output stage – a practice that reduces loading distortion of the VAS [43]. Naturally, the best way to fight instability is not to dampen the high frequency response into oblivion but to re-think the transistor choices and find devices that have a lesser amount of oscillation triggering capacitance.

Close proximity rail capacitance (7) helps in minimizing effects caused by supply rail resistance and inductance. If power is brought to the circuit from an external circuit board, each supply rail of the power amplifier board should have a capacitor of around 10 – 330 μF in close proximity to the DC entry point. The purpose of these capacitors is to avoid DC fluctuations caused by wiring resistance and *especially inductance*: In case of a sudden and fast current demand the increased inductance limits the incoming current, thus creating a noteworthy *local* sag effect. Several smaller capacitors (10 – 100 nF) should be used to locally decouple the rails in very close proximity of the active circuitry (transistors and OpAmps). These capacitors – besides acting as a reservoir for instantaneous current draw – also help blocking RF.

Shown in the schematic, but not given a particular number, is another thing that also enhances stability: Use of constant current sources instead of plain resistor helps each stage to retain a fixed current even when supply voltages vary. This preserves steady open loop gain and operating conditions, which in turn keeps the feedback factor stable as well.

The operation of Zobel network (8), which consists of series connected resistor and capacitor, is once again easier to understand when the load is considered as variable impedance. Typical component values are 10 to 100 nanofarads for capacitance and few ohms for resistance. Since the Zobel is not a compensating circuit for the amplifier but *for the load* (which can be pretty universal) it is common to use a pretty generic circuit from one design to another. The capacitor of the Zobel network is an opposite to voice coil inductance so the two try to cancel each other out. Resistor is mainly acting as a damper. The combination smoothes out the rising curve of the

speaker's impedance. Basically, the main purpose of the Zobel network is either to prevent oscillation or to linearize complex crossover-equipped loads. The main point is that *the purpose of this circuit is not to dampen oscillation but to prevent it from starting in the first place*. While the solutions introduced so far have dealt with improving the *internal* stability of the power amplifier circuit, the Zobel is improving the *external*. Usually the -3 dB point is set somewhere around 100 kHz.

The way to derive component values for the Zobel network is following: The resistor is chosen to “equal” the nominal load resistance, optimally this means 8 ohms for 8-ohm speaker and 4 ohms for 4-ohm speaker, (exact component values have little meaning). Then the required capacitance can be derived using a formula

$$C = Le/R^2 \quad \text{(Equation 14)}$$

where R = the resistor value and Le = inductance of the speaker's voice coil. The value should be found from the datasheet listing the Thiele/Small parameters of the speaker. However, the method is pretty “HiFi” and the result may be even too linear impedance at high frequencies. Zobel networks are often mounted to the circuit board but usually they work more effectively when “hard wired” straight to speaker binding posts. In some occasions Zobel networks mounted on circuit board cause oscillation.

Since RC Zobel also *protects an oscillating amplifier* by providing an alternative shunt for the high frequency signal it may save the output stage from failing. However, in such occasion the Zobel is subjected to rigorous “punishment”. A burned or hot Zobel resistor is always a good indication of HF instability. The resistor should be *flame proof* and able to handle at least 5 watts of power (proper value of course depends on the design and its output power). The capacitor should be self-healing type. For example, a good choice is a high-voltage polypropylene “motor start” capacitor.

In practice, Zobel network is a term for any circuitry, which has a purpose of linearizing the impedance of the speaker load. Sometimes a term “Thiele network” is used to express the same thing. Complex Zobel networks can be used for almost entire linearization of the load impedance but in such cases the component values are highly speaker dependant. You sometimes see such Zobel circuits in speaker crossovers. Most guitar amplifiers tend to avoid excess linearization and use the Zobel network mainly as a stabilizing tool.

The inductor output filter (9) is a trickier form of Zobel and should likely be omitted unless the amplifier is expected to drive highly capacitive loads, such as complex crossovers. As a matter of fact, the inductor filter has hardly any effect on a load that is not very capacitive and hardly any single speaker is. On the other hand, if you are building a bass or acoustic amplifier that uses a configuration of multiple speakers and crossover network you might want to consider using the output inductor. As a downside, the additional circuitry introduces more complexity, expense and a very considerable source for catching RF.

The heart of the circuit is an inductor that both evens out the effects of load capacitance and attenuates higher frequencies; it also isolates the negative feedback chain from the load capacitance. Parallel resistor is used as a “damper” against impedance spikes caused by the inductance of the coil. Typical construction of the circuit consists of 10 to 20 turns of medium gauge (0,8 – 1 mm) coil wire wrapped

around a 2 - 5-watt, 10-ohm resistor, which also serves as the inductor core. The resulting inductance is usually in the range of 0.5 – 15 μ H. Usual values are between 0.5 – 2 μ H. If we want to get excessively technical about this, we realize that an inductor wound around a resistor (or a capacitor) essentially creates a transformer. If the inductor would be wound around a capacitor (this is actually done in some amplifiers) the transformer would have shorted turns. In practice, the coil should have an air core and preferably it should be toroidal to reduce the created magnetic field. It also should be placed away from other magnetic fields.

3.14 Protective circuits

3.14.1 Protecting input stage from static voltage spikes

An input stage may become subjected to transient voltage/current spikes caused by faulty equipment or DC (offset) potential's fluctuation – or even more common: A static discharge. Out of all, due to its high input impedance, the differential input stage is the most vulnerable against such anomalies. Sudden transients may reverse bias a transistor (or FET) but since their power is probably very small they will not necessarily destroy it at once, instead the transistor's junction will deteriorate and get less defined. As a result the device's performance declines and it may start to introduce more hiss and noise. At each occasion of reverse biasing the base-emitter junction (which is significantly weaker than the collector-base junction), the deterioration will advance. Since effects are only slight they are pretty difficult to notice without actual A-B comparison to a similar amplifier in better condition. Eventually the device just ceases to work. Modern transistors will usually reverse-conduct at around 5 to 7 volts. Germanium transistors tend to have a larger base-emitter breakdown voltage; around twenty to thirty volts.

The method of preventing reverse conduction from happening is fairly simple but will involve adding “extra” components to the design. One possible addition is a series resistor, which limits the current to desirable levels. A lot more effective addition is a transient suppressor consisting of two back-to-back diodes connected between the supply rails and input. The diodes will shunt all voltage transients that exceed rail voltage plus diode's forward voltage in amplitude to corresponding supply. To acquire symmetric operation (and highest headroom) a single supply circuit must do the shunting from a node that uses half the supply voltage potential as the signal's “zero” point. Refer to Figure 3.64. For the record, most single-supply powered LTP input stages bias the bases to half of the supply anyway so arranging the shunting configuration should cause no considerable problems.

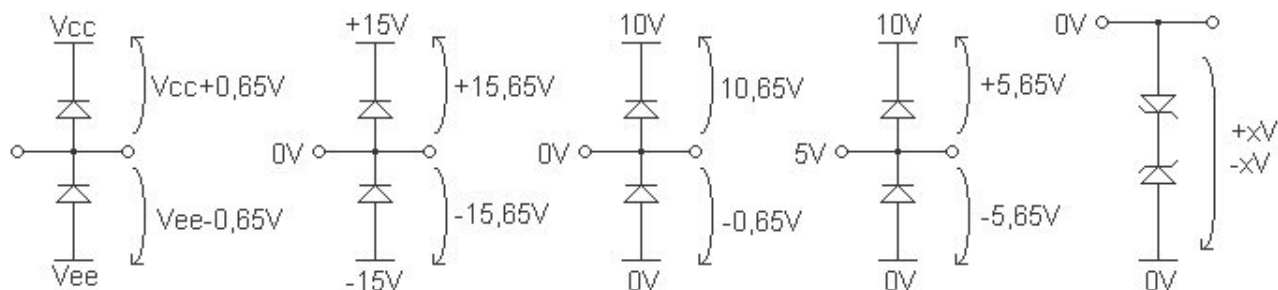


Figure 3.64 Limiter diode thresholds.

An alternative for conventional diodes is a circuit consisting of Zener diodes in series, hooked from signal node to ground and with their anodes connected together. If the node was referenced to ground, it effectively becomes limited to $U_{ZENER} + U_{DIODE}$ in

amplitude. The solution may be little more expensive but works as well providing a more specific voltage limit. Logically, applying more diodes to the chain allows increasing the limiting threshold.

3.14.2 Turn-on delay

When the amplifier is turned on it takes a finite time for the DC voltages to settle into their nominal value. All concerned resistances and capacitances, as well as current draw (loading), define how fast this happens. A steadily rising voltage has an AC component and therefore it passes through signal stages and gets amplified. The transistors do not begin to conduct instantly but will do so when the base bias voltage has risen high enough. All events together create a transient “turn-on thump”, which can be quite harmful to speakers especially in the case of high-power amplifiers. Other than that, the initially low but increasing headroom and other anomalies caused by “under-voltage” of the circuitry can create signals that are not very healthy to either speakers or listeners. It was quickly discovered that similarly to tube amplifiers solid-state amplifiers required a startup delay as well – however, only in connecting the speaker load into the signal chain. Later it was discovered that in most cases switching off the power amplifier’s signal path was an equally efficient solution.

The delay circuitry is often based on RC circuit that sets the startup delay according to its time constant. Basically, a circuit charges a capacitor until a certain voltage potential triggers a switching device. The switching device can be a simple relay or a transistor – or a combination of both. Modern designs tend to favor solid-state switching of either power amplifier’s signal path or driver stage’s supply. This is cheaper and often also more reliable than using ordinary relays. Nowadays, varying the transconductance of the input stage (sort of shutting it off) has become a very popular technique as well. The related theory is explained in the chapter concerning soft limiters.

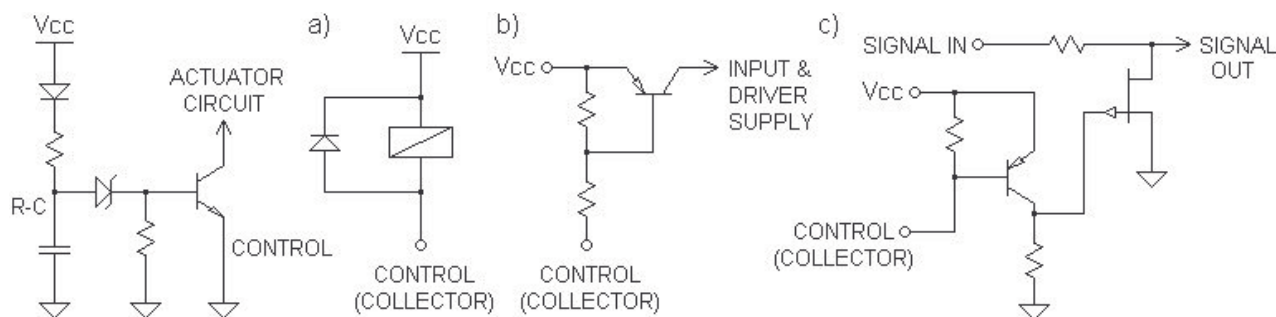


Figure 3.65 Examples of turn-on delay circuits.

Figure 3.65 shows a basic example of a delay circuit (left) and various actuators it may use. When the power is turned on the capacitor starts to charge through the diode and the resistor. When the capacitor has charged itself to its maximum voltage potential - which is the forward voltage of the zener plus voltage over transistor’s base-emitter junction - current starts to flow on transistor’s base thus triggering the actuator. In circuit a) the actuator is a relay that acts as transistor’s collector load. A diode is connected in parallel with the relay to prevent reverse biasing during switching (inductor is a difficult load). In a most minimalist circuit the transistor itself could be replaced with a relay that would trigger after the forward voltage of zener is exceeded.

Actuator b) is a switch: When no collector current is flowing in the NPN control/actuator transistor the PNP transistor does not conduct. When control transistor starts to conduct, the PNP transistor will open. Usually this kind of circuit is used in controlling powering up of the low-current stages of the power amplifier. Note that this circuit can not prevent the startup thump *in single supply circuits* since in them the “half of supply” voltage potential at the output will settle only after switching on the input and VAS. However, in a dual supply amplifier the circuit works perfectly.

The final actuator c) is a voltage-controlled resistor that (together with another resistor) forms a voltage divider to the signal path. This is also known as “muting” circuit. PNP transistor acts as an “inverter” of the control transistor’s signal. In initial state the collector voltage of the PNP transistor is nearly zero and the FET conducts (introducing only a very small resistance) thus attenuating the input signal and muting the turn on thump originating from the preamplifier section. After turn on delay the collector voltage quickly rises towards the supply voltage and the FET closes (and begins to introduce a very high resistance). Effectively this “removes” the voltage divider from the signal path.

3.14.3 Current limiting and flyback protection

The main task of current limiting circuits is to offer protection against short-term short circuits but they also protect the power transistors from large current transients that the supply is capable of feeding for very short periods of time. The most typical current limiter of modern amplifiers is the VI limiter. The operation of the circuit is simple: The current over ballast (or “current sensing”) resistor creates a voltage drop which means that a certain voltage potential must exist at the base of the protection transistor as well. When the voltage potential exceeds the base-emitter voltage (of the protection transistor), the transistor begins to conduct and driver’s base current is diverted to negative feedback loop. This gradually shuts down the power transistor. In common folk terms the VI limiter circuit “samples” the current flowing through ballast resistors and limits the transistor’s base current according to it. It makes no difference for the circuit’s operation whether the power transistor is a single device or a combination of multiple discrete devices, i.e. a Sziklai or a Darlington pair.

With most basic circuitry (all optional resistors omitted) and in standard ambient temperature of 25° C the current that triggers the limiter can approximately be calculated from the formula:

$$I_{LIMIT} = 0.65/R_{BALLAST} \quad \text{(Equation 15)}$$

However, without any external components the safety operating area (SOA) approximation is just a straight line, which is not ideal since SOA is never linear. Aside the typical voltage divider configuration, consisting of a base resistor and a “fold over” resistor connecting the feedback loop, the VI limiter can be equipped with a resistor connected to supply. This resistor generates additional voltage over fold over resistor, which modifies the linear slope of the basic circuit to resemble more the actual SOA curve. Instead of supply rail, the resistor can be connected to common potential as well - in fact, this practice stops supply noise from being injected to the circuit via the resistor. The disadvantage is of course that when the reference to supply potential is lost the variations in supply voltage will create tracking problems. A hefty sum of methods to follow SOA limits more accurately exists but they are left out of

the focus since satisfying operation can be reached with using a two-slope limiter. Sometimes capacitors are connected across base resistors to suppress parasitic oscillations or in parallel with fold over resistors to prevent circuit from operating during transients. If more than one output device per rail is used, each one should have an individual base resistor that connects the emitter of the corresponding output device. The collector diodes shown in the schematic example prevent reverse biasing of the protection transistors.

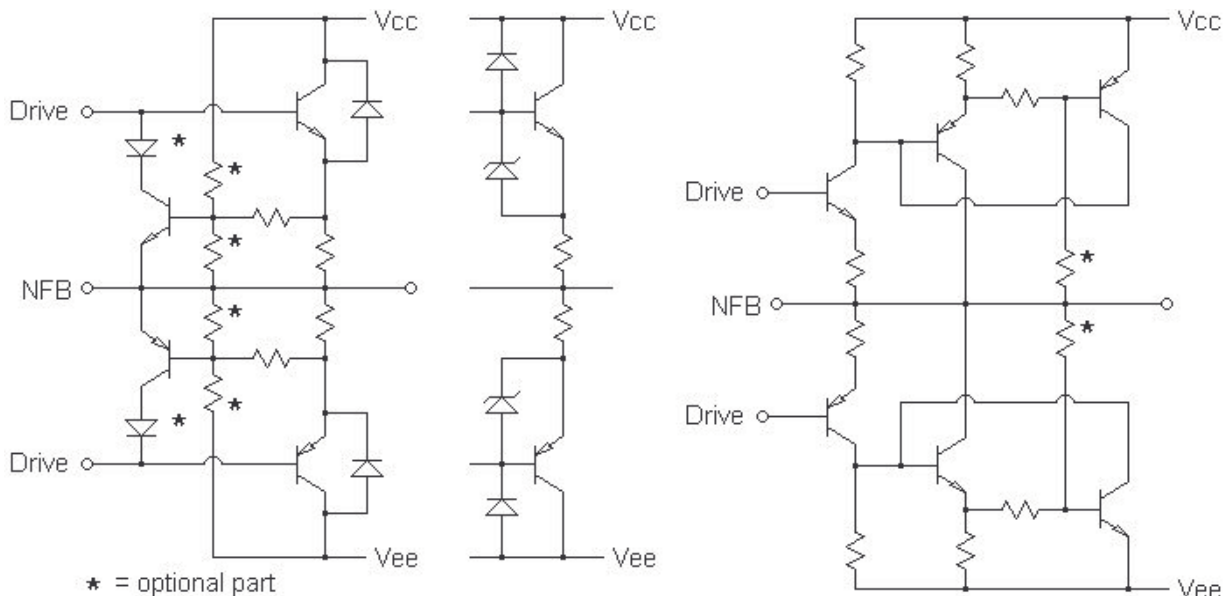


Figure 3.66 VI limiter circuits. Left: The conventional arrangement and right: Alternative arrangement for compound follower output stages.

Safe operation area is a function of temperature and time; however, commonly used VI limiter circuits, depicted in the Figure 3.67, are time and temperature invariant. Design of VI limiter circuits, from simple current limiters up to four-slope limiters, is thoroughly discussed, for example, in United States Patent number 7,061,740B2, which was filed in 2002 by Eric Mendenhall assignee to Gibson Guitar Corporation. Motorola patented the same idea already in 1975 and by no means has neither one of them been the inventor of this circuitry: The oldest *published* application incorporating a VI limiter circuit I've seen so far is the Dan Meyer's original "Tiger" amplifier from Popular Electronics magazine 1969. US Patent application assigned to RCA, which incorporated Jack Sondermeyer's design, was filed in 1967 [50], so was Richard Burwen's design assigned to Analog Designs. Granted is that the Mendenhall's patent presents a much more sophisticated protection method, though. The following equations for deriving component values are summarized from the aforementioned patent and from a whitepaper of Audio Engineering Society convention related to it.

The simple current limiter circuit (a) will not track SOA very accurately as can be noticed. The protection threshold for such a circuit is:

$$(V_{BE}/R_E) \times (R1 + R2) / R2 \quad \text{(Equation 16)}$$

In practice, designing a simple, current limiter-type VI limiter circuit is easier than it might seem – another issue is how useful such a circuit really is. Theoretically, during a short the whole rail voltage is affecting over the load so one can reference datasheet's SOA curve to find out the maximum current at the concerned voltage. At

this state, a voltage of $I_{MAX} \times R_E$ is affecting over the ballast resistor R_E . The current limiting transistor will require a voltage of about 0.65V at its base to turn on so rest of the design is just a matter of calculating the correct values for a *voltage divider* R1 and R2. Note that their resistance should be high enough not to interfere with the emitter resistance value. For example, try 100 ohms for R1.

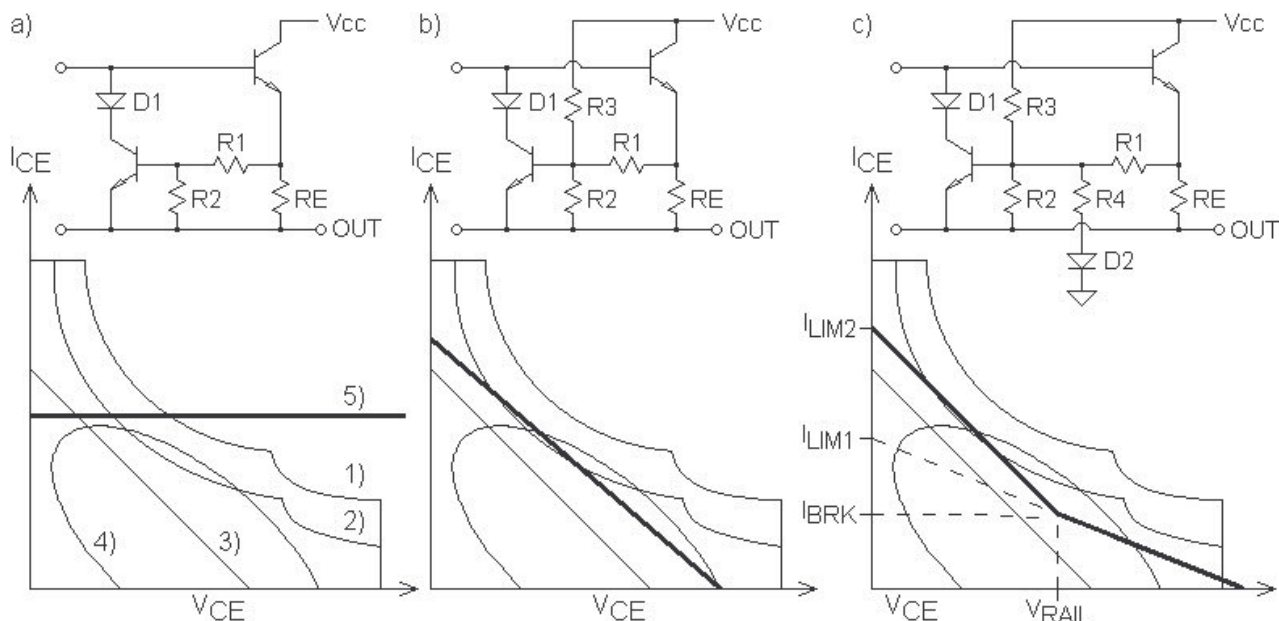


Figure 3.67 VI limiter's accuracy: a) Simple current limiter, b) Single slope limiter and c) two-slope limiter. Graph curves are 1) 25° C steady state SOA, 2) 100° C steady state SOA, 3) resistive load line and 4) inductive load line. 5) Represents the slope of the limiter.

A simple resistor hooked to supply in single slope limiter circuit b) will improve the SOA tracking tremendously and can be used with high-power transistors when rail voltages are reasonably low. There are few basic assumptions concerning the circuit: R3 must be large with respect to R1 and R2 and R_E must be small with respect to R1, R2 and R3. Rail voltage must be high when compared to saturation voltage of the output and voltage across emitter resistor R_E . The protection threshold is:

$$I_{CE} = [V_{BE} - (V_{CE}/R3) \times ((R1 \times R2)/(R1 + R2))] \times [(R1 + R2)/(R2 \times R_E)] \quad \text{(Equation 17)}$$

The value of R2 is chosen to scale all the rest of the resistors, and then the X and Y intercepts for the SOA tracking slope are determined: If $V_{CE} = 0$ then $I_{CE} = I_{LIMIT}$ (Y-axis) and if $I_{CE} = 0$ then $V_{CE} = V_{LIMIT}$. Thus one can calculate the values for R1 and R3 from the following equations:

$$R1 = [R2 \times (I_{LIMIT} \times R_E - V_{BE})] / V_{BE} \quad \text{(Equation 18)}$$

$$R3 = [V_{LIMIT} \times (R1 \times R2)] / [V_{BE} \times (R1 + R2)] \quad \text{(Equation 19)}$$

The two-slope limiter's (circuit b) protection threshold is derived from two slopes set by the operation of the diode D2: First slope - when diode is off and $V_{CE} > V_{RAIL}$ - follows the slope defined in single slope limiter, thus it shares the same basic equation. The second slope, $V_{CE} \leq V_{RAIL}$ slope is derived from

$$\begin{aligned}
I_{CE} = & [(V_{BE} \times (R1 + R2)) / (R2 \times R_E)] \\
& + [((V_{RAIL} - V_{CR}) \times R1) / (R4 \times R_E)] \\
& - [(V_{CE} \times R1) / (R3 / R_E)]
\end{aligned}
\tag{Equation 20}$$

The design process follows the measures of designing the single slope limiter but instead of choosing I_{LIMIT} as the starting point one chooses $I_{BREAKDOWN}$ instead. This is the current where the two slopes cross each other. The equation of the dashed line is:

$$I_{CE} = [(-I_{LIMIT1} / V_{LIMIT1}) \times V_{CE}] + I_{LIMIT1} \tag{Equation 21}$$

Solving I_{LIMIT1} by setting $V_{CE} = V_{RAIL}$ and $I_{CE} = I_{BREAKDOWN}$ leads to

$$I_{LIMIT1} = (I_{BREAKDOWN} \times V_{LIMIT1}) / (V_{LIMIT1} - V_{RAIL}) \tag{Equation 22}$$

This allows deriving R1, R2 and R3 with the method familiar from single slope limiter design. R4 is determined from the Y intercept where $V_{CE} = 0$ and $I_{CE} = I_{LIMIT2}$. Setting V_{CE} as 0 when solving formula 20 yields:

$$R4 = (V_{RAIL} \times R1 \times R2) / [(I_{LIMIT2} \times R_E \times R2) - (V_{BE} \times (R1 + R2))] \tag{Equation 23}$$

By adding similar “shunts” as R4 and D2 to voltage potentials of $V_{RAIL} / 2$ and $V_{RAIL} / 4$ (or wherever the constant power curve is best fit) the VI limiter circuit’s SOA tracking can be made to consist of four slopes [23]. The design process is similar to two-slope design starting from the rightmost portion of the curve when D2, D3 and D4 are off. After this, one proceeds to add breakpoints.

The circuit far right in Figure 3.66 is a basic current limiter that monitors the current flow over emitter resistors. An excessive current will cause the protection transistor to cut off current from the base of the monitored transistor. The concept is almost identical to “conventional” VI limiter, except that the limiting transistors are shunting current to supply rails instead of the feedback path. This basic circuit can be useful in quite many applications; consequently, it suits nicely for protecting power transistors of compound follower output stage. The operation of the circuit is by no means as accurate as of a nicely designed VI limiter but it’s worthwhile to consider that alternative methods for output stage protection exist as well. In his famous amplifier design, Marshall Leach used a combination of this circuit and a conventional VI limiter configuration; however, the concerned circuitry was used to prevent saturation of two VAS transistors, which were preceding a cascoded LTP input stage. Note how the protective device is connected: It cuts off the base current of the *sensed* transistor. This is an essential detail. If one wants to use this circuit, for example in combination with a Sziklai pair, the collector of the protection transistor must connect base of the final transistor, *not* the base of the driver transistor.

Though the basic VI limiter circuit is simple, implementing it is not and may cause unwanted, even catastrophic, results. Usually the limiting circuit has to be designed with the right tolerance of current and voltage in order to prevent trouble. Since speaker load is reactive its requirement for current and voltage can be in a different phase against each other. As a result, the protection circuitry detects a higher amount of base current than what the voltage drop suggests and kicks in. Now, the problems begin: If the amplifier circuit fails to deliver the current that the speaker load requires in normal conditions the voice coil may begin to store the unused energy. As soon as a

chance comes, the stored energy is released in a short burst. This happens twice in each cycle, usually near the zero crossing point. The burst's amplitude is very high - it can easily exceed the supply voltage and *reverse bias* the output transistors thus causing them to destruct. Also, due to its transient-like nature the burst is very hazardous to some speakers.

The “flyback” or “flywheel” diodes, connected back-to-back from emitters to power rails, offer protection against the most severe consequences of problems related to reactive nature of the speaker load. The diodes simply shunt all voltage and current exceeding supply voltage ± 0.65 V to supply, which absorbs them and makes sure that power transistors never get reverse biased. The measure will not, however, protect the speakers from the remainder of bursts. Usually flyback diodes are employed even when a VI limiter circuit is not used since a speaker load - especially when it's coupled with a crossover - can still induce high-voltage transients. Some integrated Darlington transistors may even include this diode as an internal element. Another form of similar protection is to shunt the excessive emitter current to the supply with back-to-back zener diode and diode. When the emitter-base voltage exceeds the forward voltage of the zener the excessive current will be shunted to the supply through a diode.

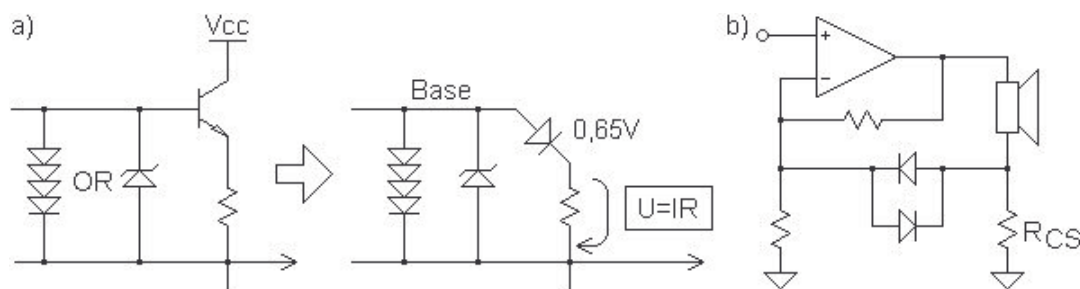


Figure 3.68 a) Principle of crude VI limiting. b) Current sensing gain limit.

Figure 3.68 a) shows the principle of a very crude current limiting method, which was sometimes used in early transistor amplifiers. Like defined by Ohm's law, the ballast resistor introduces a certain voltage drop at a certain current. When referenced to output, the voltage potential at the base of the transistor must be approximately 0.65V higher than what the voltage drop over the emitter resistor is. Thus one can calculate the maximum base voltage according to maximum emitter current. The base voltage can then be limited to a proper value by a zener diode, several diodes connected in series or by a (series) combination of both. The circuit is very simple and crude and naturally can't track SOA accurately - nevertheless, it is often better than nothing. A patent of the circuit was granted to Emerson Electric Corporation in 1971 (US Patent 3,600,695).

A lot more useful idea is the current sensing gain limiter (circuit b). This early RCA design, in slightly altered form and in conjunction with the VI limiter, has been used in many Fender amplifiers. The principle of operation is extremely simple: Whenever current flow through the current sensing resistor R_{CS} is high enough to create a voltage potential that exceeds the forward voltage of the diode the diode begins to conduct and causes gain reduction. The configuration blends in nicely with the existing mixed mode feedback setup and could theoretically even suffice as a soft limiter. Unfortunately there's a good chance that the limiting action will introduce high frequency oscillation. Usually the single diode is replaced with a series connection of a diode and a Zener to increase the threshold of limiting.

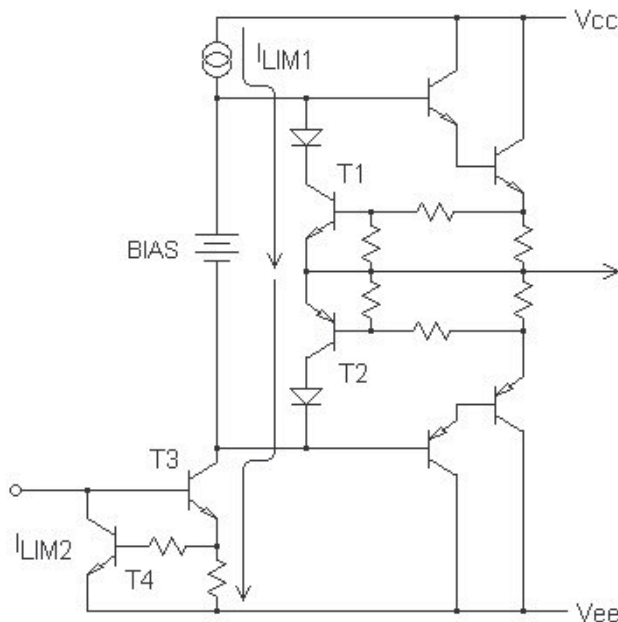


Figure 3.69 Although quite essential, current limiting of VAS is often omitted.

In its basic form the VI limiter circuit is quite far from being a reliable protection, as we will soon learn: During fault condition (which is usually a short circuited load) the maximum current that transistor T1 has to shunt away from (NPN-side) driver's base is limited by the constant current source load of VAS. Unfortunately, there is no such inherent limit for the PNP side and therefore T3 becomes (nearly) directly connected to the (shorted) output through the saturated T2. We now have a high-current path with nothing but moderately low-power devices on it – and one of them is bound to fail sooner or later. Even worse, the output's VI limiter is still incapable to shunt all this excessive current away from the (PNP-side) driver's base, thus the PNP side will conduct more and exhibit a gigantic power loss.

To prevent disasters it is therefore equally essential to limit the VAS current as well - and one suitable circuit for this purpose is the good old VI limiter: T4 now limits the output of T3 by shunting the excessive base current whenever emitter current rises too high. It is unfortunate that the benefits of this circuit are far too unknown: During a short circuit a current limited VAS can easily reduce the power losses over the PNP side as much as 250 watts! It's needless to say that this is not an insignificant figure.

3.14.4 More about short circuit protection

Essentially we need to realize one more thing: *A VI limiter alone is not capable of acting as a reliable short circuit protection!* According to Ohm's law, the higher the output signal amplitude is, the higher the current draw during a short circuit is as well. Furthermore, because resistance of a short-circuited load is so low even extremely low output signal amplitude can create a gigantic current draw. For example, 500 mV over 0.1 ohms creates a current flow of 5 amperes!

In practice this means that the main reason why a short circuit is so devastating is the fact that when the amplitude of driver's input signal is moderately high it, in turn, leads to moderately high output signal, which then causes an extremely high current draw. Without any (base) drive current the output transistors could actually cope happily even with a complete zero-ohm load. This is because they simply would not pass any current while being in *cut off* state. The operation of a VI limiter, in fact, is such that it tries to achieve this condition by attenuating the input signal. However, a

typical VI limiter can do this only at a level that provides a very *short-term* protection. Any input signal will cause current draw and eventually the increased amount of heat dissipation will destroy the output devices. Typically a well operating VI limiter may be able to drop the dissipation from about 400 watts to about 120 watts (per device), which will extend the amplifier’s life from few nanoseconds to few seconds – possibly even to few minutes. This is great but far from perfect.

An effective short circuit protection must therefore cut off the output signal completely – in other words, cut off the input signal from the drivers’ bases. This can be done, for example, by activating the (input stage’s) muting circuitry or by cutting off the voltage supply of the drivers; the latter approach definitely ensures that there cannot be any base drive signal to turn on the devices. One method to do this is to use the thyristor latch circuit presented in 3.14.7. Another very similar method is shown in the schematic example: Transistors T1 and T2 regulate the supply of the input stage, VAS and drivers. Transistors T3 and T4 can bypass the “lower” divider resistors that define the output voltage of the regulators. Effectively this “cuts off” the supplies, thus turning the drivers into cut off state as well. Shunting transistors are controlled by a signal from detector circuit, which is built around transistor T5. This circuit shall be discussed in detail later. Note that this kind of protection should also feature a visual indication of a fault condition so that the user could instantly find out why the sound went off and act accordingly.

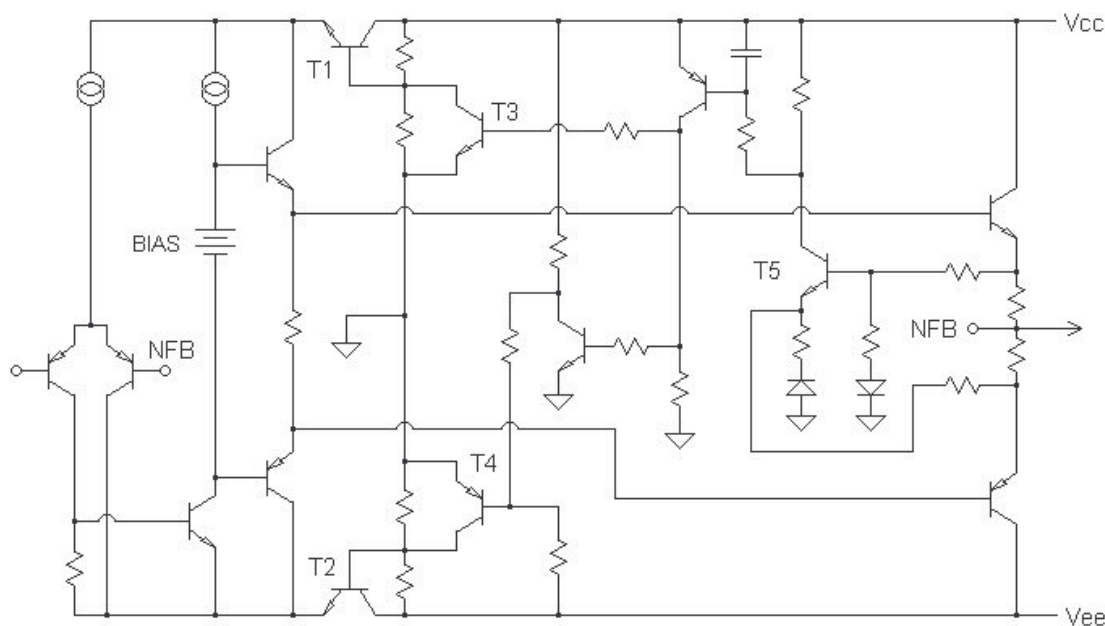


Figure 3.70 Conceptual schematic of a short circuit protection that turns off drivers’ power supply.

Circuitry tampering with rails of the input stage always has some risks: If both rails can’t be (simultaneously) clamped to a potential very near zero volts (or simply cut off) there will be a vast amount of DC offset in the output – and with a short circuited load even millivolts count. Likely relay-based switching of the drivers’ supply rails could prove to be a more reliable solution, yet more expensive and “bulky” as well. The protection is also rather brutal and recovery from protective state takes a finite time and often introduces DC offset since the driver-side rails typically have some capacitance that has to be charged. Due to need to discharge the same capacitance (when the protection trips) there will also be a finite delay, which is fortunate because the protection should not be allowed to trip under any normal conditions.

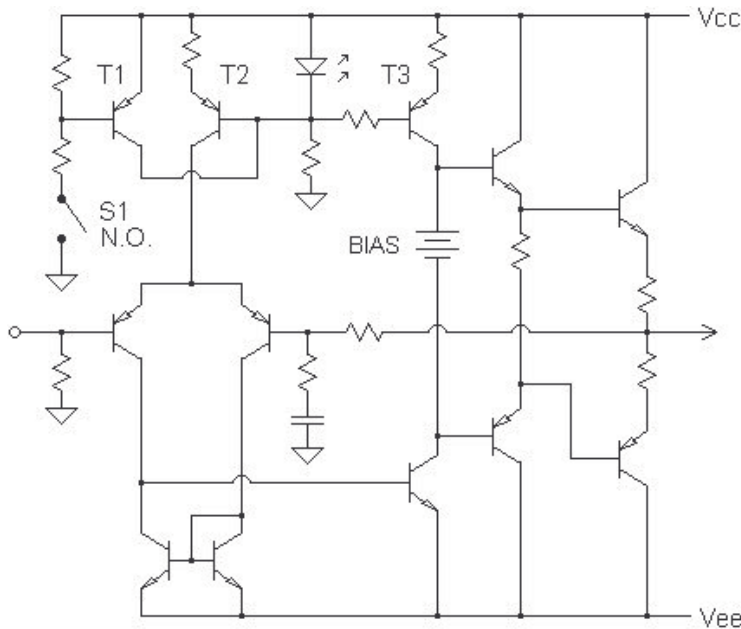


Figure 3.71 Controlling the constant current sources provides a very effective way to realize a “muting” protection circuit.

Fortunately, there is hardly ever need for such complexity in real-life applications. For example, Figure 3.71 shows a very realistic protection circuit: Input differential and VAS are both loaded by active constant current sources (the circuit is discussed in 3.8). Now, if the voltage reference for the current source transistors T2 and T3 is “clamped” to supply rail with a switching transistor T1 the output current (from the sources) becomes so low that it effectively turns off both the input stage and VAS. The switch S1 is normally open type. Basically it can be a simple thermal switch or a semiconductor switch linked to applications such as SOA-limit or startup delay.

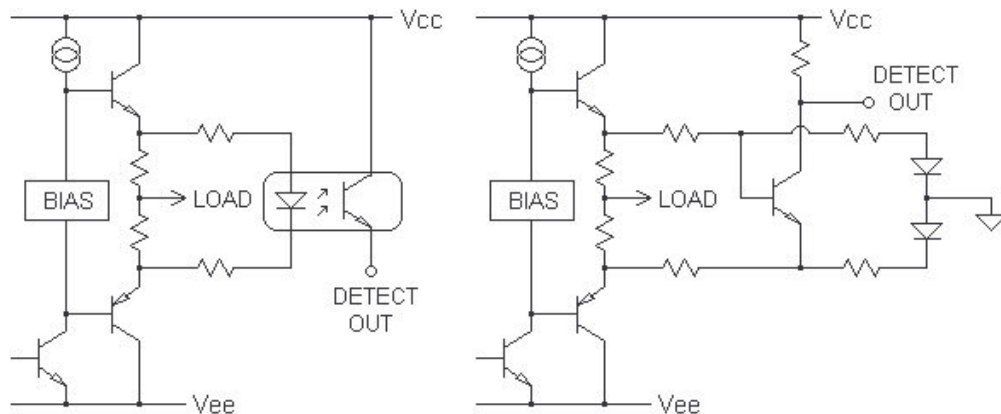


Figure 3.72 Alternative overload detector circuits. Left: Optocoupler drive and right: Pioneer-style “bridge” detector.

The overload detector can be built with numerous ways – and like always, all of them have both advantages and weaknesses. The optoelectronic circuit (shown in Figure 3.72) monitors the voltage potential between output transistors’ emitters. This potential is naturally defined by the amount of current flowing through the emitter resistors (simple Ohm’s law). The LED is triggered to light up when a desired maximum current is exceeded. Often there is a requirement for current limiting resistors that define a proper voltage drop over the LED as well. By today’s standards this is a really crude circuit that is unable to cope with the non-linear SOA of transistors, however, it has a benefit of galvanically isolating the following protection

circuitry. One must also note that none of the presented circuits prevent the use of conventional VI limiter.

The circuit at right in Figure 3.72 is the detector used in the prior example. Pioneer Electronic Corporation initially patented this extremely simple “bridge” detector circuit in 1974 (US Patent 3,814,988); lately it has seen a comeback in new Randall amplifiers like the *Warhead X2*. In normal operating conditions the V_{BE} of the detector transistor is below zero volts and the collector voltage nearly same as V_{CC} . When output current increases enough, a voltage potential exceeding transistor’s forward voltage develops over V_{BE} junction and the transistor begins to conduct. The state of the circuit is determined by monitoring the voltage drop over the collector resistor.

The “Sansui” circuit, named after US Patent 4,216,517 for Sansui Electric (1980), consists of sections for output voltage and load current detection, which are summed together by a discrete AND-gate. To prevent “false” triggering the circuit activates only when the output voltage drops below a predetermined level while a considerable load current is flowing. This is a typical condition when the output has become shorted.

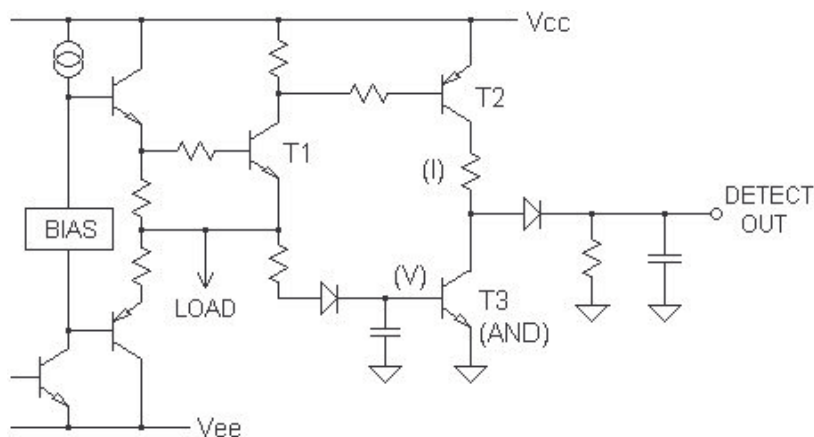


Figure 3.73 Sansui-style detector.

Let’s examine the circuit: The “upper branch” is the detector for current, which traditionally monitors the voltage potential affecting over the ballast emitter resistor of the NPN output transistor. Whenever load current is high enough to create a V_{BE} of about 650 mV (or higher) at the base of T1 the transistor in question begins to conduct. This, consequently, turns on transistor T2, thus opening a path to the positive rail. However, by itself this is not enough to cause a “high” output from the detector because T3 is capable of shunting the collector output of T2 to ground (thus the need for a current limiting resistor between the collectors of T2 and T3). T3 is in cutoff state whenever the output voltage is below a level that creates a sufficient forward voltage to its base.

The circuit shown in Figure 3.74 has been rather common in amplifiers that are German design. Transistors T1 and T2 monitor the voltage potential over the corresponding output transistors. When the voltage potential over the output devices drops, voltage at the midpoint of the resistor dividers drops as well, thus T1 and T2 turn closer to cutoff state causing their collector voltages to drop. The voltage drop at the collector of T2 turns on the transistor T3, which in combination with the collector voltage drop of T1 clamps the voltage at T4’s base to a level where the transistor in question begins to conduct. This creates a low resistance shunt path for the resistor in

parallel with T4 and thus affects the voltage divider connected to negative rail: The voltage at the detector's output drops in reference to negative rail and the indicator LED turns on.

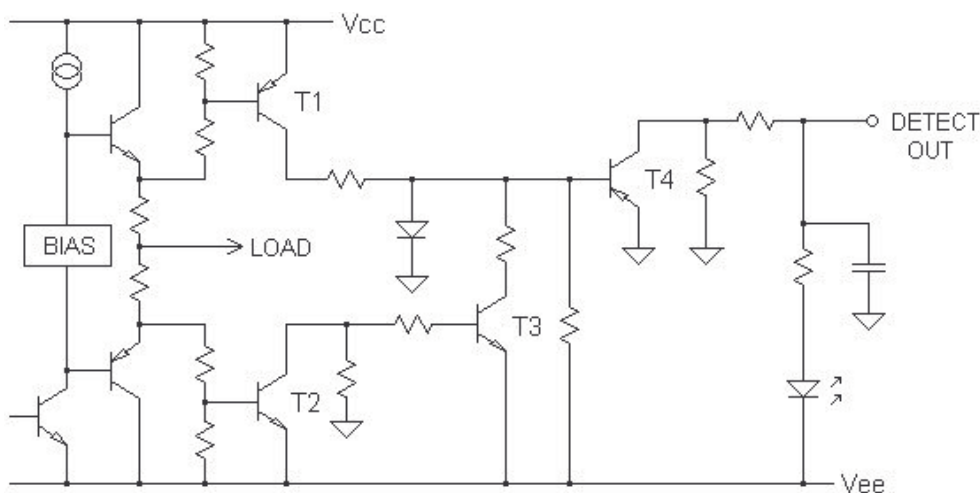


Figure 3.74 Fairly common detector circuit.

Circuit depicted in Figure 3.75 is a really simple one: A short circuit current turns on transistors T1 and T2, which creates current flow from positive to negative rail. The current flow turns on transistor T3, which clamps the detector output to “high” potential.

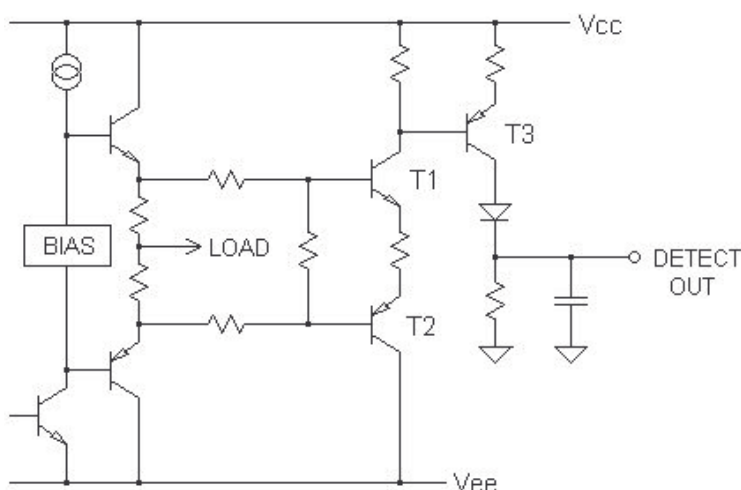


Figure 3.75 A detector circuit that is closely related to the ordinary VI limiter.

3.14.5 Baker Clamps

Although invention of the Baker Clamp circuit is credited to Richard H. Baker (US Patent 3,010,031) it was already common knowledge in 1953 and described in transistor introductory papers that were written by Richard F. Shea. As is known, transistors have a finite release time from saturation state, which can create a multitude of problems and quirks in performance. A Baker Clamp circuit prevents deep saturation by providing a path for excessive base drive current. The circuit is chiefly used in switching applications but it has proven itself useful in audio as well – especially as a part of the VAS. Unfortunately, Baker Clamping is nowadays a rather unknown technique and aside few guitar amplifiers, namely Fenders, used mainly in PA power amplifiers.

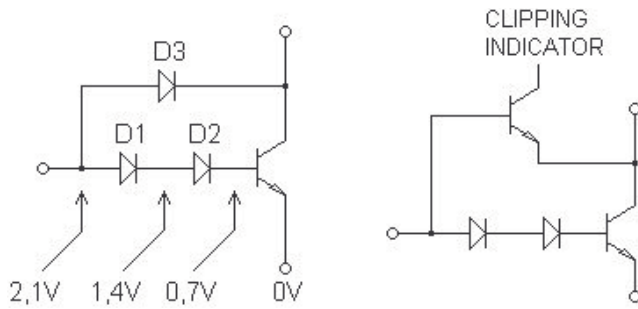


Figure 3.76 Baker Clamp variations.

The basic configuration of a Baker Clamp is shown in Figure 3.76. When transistor goes towards saturation its collector voltage tries to drop below 1.4 V. This creates 0.7 V across diode D3, thus it begins to conduct and shunts excessive base drive current to the collector. A baker clamp diode must have fast recovery time. A novel variation is to replace the diode with a transistor's base-emitter junction, as the collector signal can then serve as a fine clipping detector. The on-state resistance of a Baker Clamped circuit is fairly high thus the method is less suitable for use in the output stage.

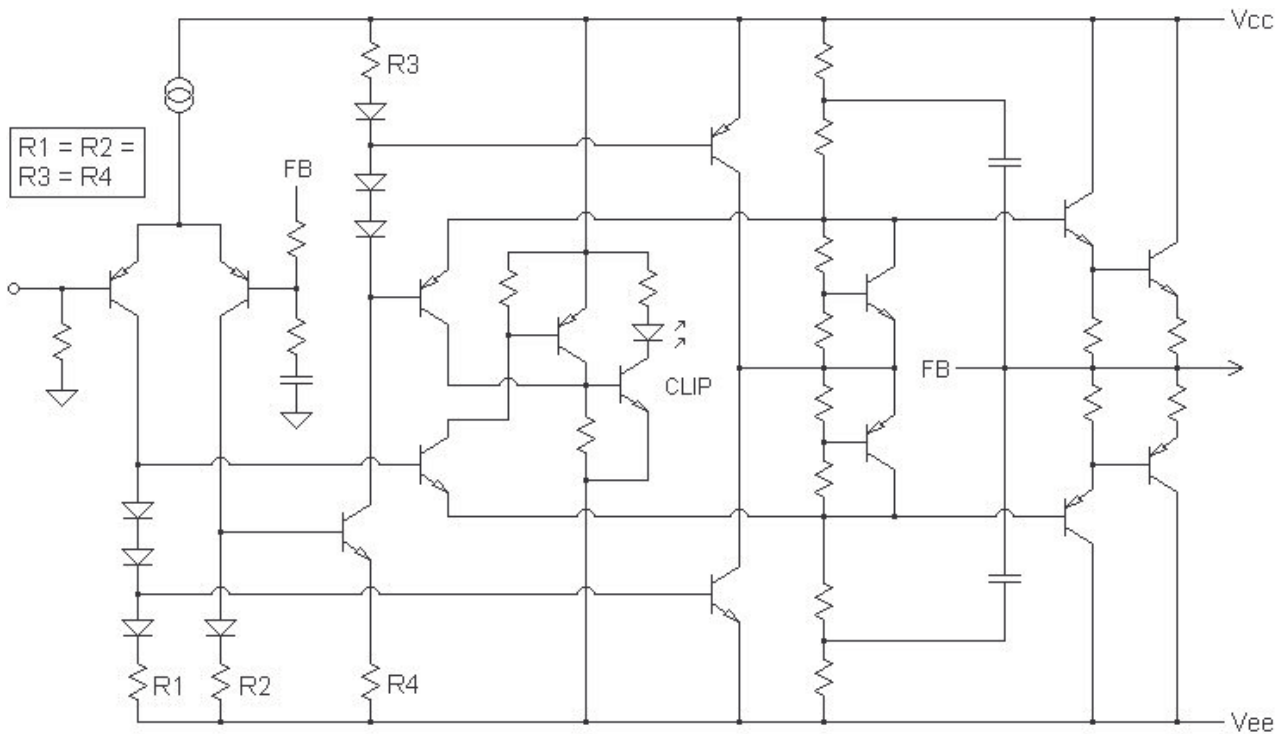


Figure 3.77 Typical implementation of a Baker Clamp circuit in a Fender-style symmetric circuit.

The enclosed circuit diagram, typical to some modern Fender amplifiers, shows one way to implement the Baker Clamp: The voltage amplifier stage is dual symmetric, thus each common emitter amplifier is clamped individually. Note the configuration of diodes in series with the base-emitter junction. Collector-base clamps are realized with a complementary pair of transistors and their collectors feed a clipping indicator circuit. The circuit also shows other ways to enhance symmetry: The usual single VBE-multiplier arrangement is replaced by a dual. Also, the output from the symmetric VAS feeds the interconnecting node of the VBE-multiplier circuits. Now the bias circuit loads both common emitters of VAS equally. This would also be a quite nice arrangement for circuits with an OpAmp-based VAS. Bootstrap loads are naturally symmetric as well. Note that Baker Clamp is usually *only* a protection against VAS saturation. The output devices will still need their effective protection

scheme against short circuits (i.e. VI limiter), which for simplicity's sake is omitted from the schematic.

3.14.6 Thermal switches

Most mechanical thermal switches are bimetal; two metals with different coefficient of expansion welded together. A bimetal piece will bend in one direction when heated and in the other when cooled. Another switch form is a thermal reed switch: When magnetic force is generated in parallel with a reed switch the overlapping reed contacts become opposite magnetic poles, which attract each other. In a thermal reed switch the magnet consists of special ferrite compound that loses its magnetic permeability at its Curie temperature. The third switch form is the “poly-switch”, which is actually a Polymeric Positive Temperature Coefficient (or simply PPTC) resistor: PPTC resistor that has a fast and highly non-linear transition between low and high resistance. They are rated for a specific current and when it is exceeded the device heats up fast and begins to exhibit a high resistance. Therefore they are also known as resettable thermal fuses.

Thermal switches are available in discrete packages designed for mounting on the heat sink or device frame. These are pretty useful since they can be used, for example, to reduce bias voltage, control fans, decrease transconductance of the power amplifier's input stage, or, in a case of emergency, to switch off the mains power. The latter configuration is not generally recommended since it introduces a significant safety risk for repairmen. A typical configuration is to mount the switch to the heat sink and make it activate in the temperature of 70° C.

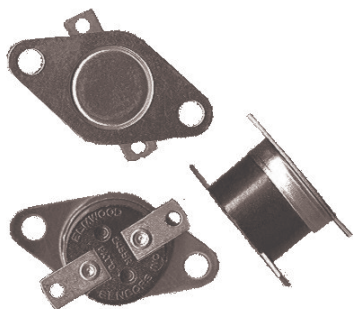


Figure 3.78 Thermal switches come in packages that allow easy mounting.

On contrary to logical belief, one will *not* likely find a conventional thermal switch inside a high-powered transformer as these contain a device called “One-Shot” thermal *fuse*, which is not reset able. Typically the thermal fuse is either bimetal or a spring-loaded switch where the contacts are welded together with a metal that melts in a low temperature. When the metal melts the tension of the spring will cause the contact to break. So, if the winding temperature rises too high the thermal fuse will open and the whole transformer has to be replaced. (It is not practical to open and rewind the transformer just to replace the

fuse). The reason why thermal fused transformers are made disposable is that the fuse rating is selected according to the temperature where winding insulation breaks or melts down, thus replacing the fuse would likely create a possibility for future shorts and faults. In short: *If the thermal fuse has blown, the transformer has likely turned unsafe as well!*

3.14.7 Thyristor latches and crowbars

Thyristor-based protection schemes were used quite often in solid-state amplifiers of the 1970's. A thyristor is basically a diode that can be triggered to conduct from a “gate” node. After triggering the thyristor will conduct just like an ordinary diode until it's anode current falls below a certain “holding current”. Thyristors are also known as silicon controlled rectifiers (SCR).

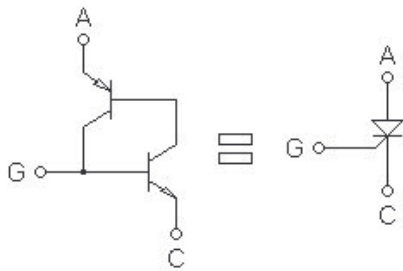


Figure 3.79 Equivalent schematic of a thyristor.

Figure 3.80 shows how a basic thyristor latch works: When the emitter current raises too high (for example during a short circuit) it will create a sufficient voltage potential to the gate of the thyristor, which in turn triggers it. The current is now diverted from the Zener diode and latched to ground through the SCR. This will seriously decrease the output voltage of the regulator. The resulting supply voltage drop in the input stage is usually sufficient for turning off the driver transistors, which in turn prevents excessive current from destroying the output transistors. Similar operation can be incorporated, for example, by using a semiconductor (such as FET) that operates as a voltage controlled switch.

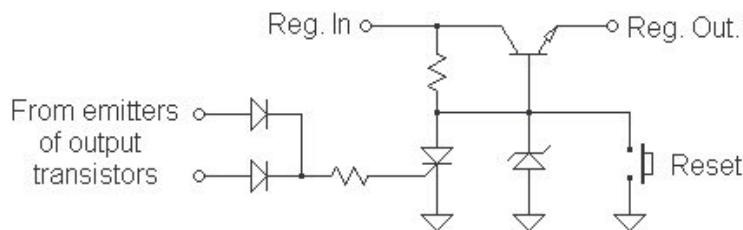


Figure 3.80 Concept of thyristor latch.

The downside of a typical thyristor latch circuit is that it requires sophisticated circuitry to turn the thyristor off automatically. Such circuitry should eliminate anode current when the fault condition has gone away. Usual circuits use a “coarse” method, which is simply expecting that the user will turn off the amplifier for few seconds. This doesn’t sound very nice but is practically a very small price to pay of not blowing up the output devices. The conceptual schematic shows another, quicker alternative for resetting the latch using a momentary switch.

During the days of early totem pole output transistor amplifiers VI limiting was not a well-known concept so protections were often based on thyristor latch circuits. Figure 3.81 shows few typical implementations named after their corresponding patent holders: US Patent 3,930,207 describes the Unicord Incorporated “E.S.P.” (or Electric Short Proof) circuit, used in the Univox “Mobile Ohm” amplifiers. This is a very basic thyristor latch that shunts the driver stage’s supply to ground whenever emitter current is high enough to create a voltage that is sufficient to turn on the thyristor. The driver stage is turned off and the output signal is virtually cut off as well. In turn, the output current becomes so low that a short circuit causes only a little harm.

A patent for the circuit was applied as late as in 1974. At the time totem pole topology was already considered rather “ancient” and in fact, Univox amplifiers housed a power amplifier circuitry that was a lot more “modern”. (Nevertheless, the basic scheme of the protection is the same regardless of power amplifier’s topology). Some may wonder why Unicord did not just incorporate a typical VI limiter. This is explained by the shortcomings of the aforementioned circuit that were discussed earlier: Even when the maximum power dissipation of the output devices is limited to

about half of their rated capacity (this is a rather typical amount) they still dissipate a substantial amount of heat, which in turn leads to early failure. While inconvenient, a properly designed thyristor latch is essentially a more reliable protection.

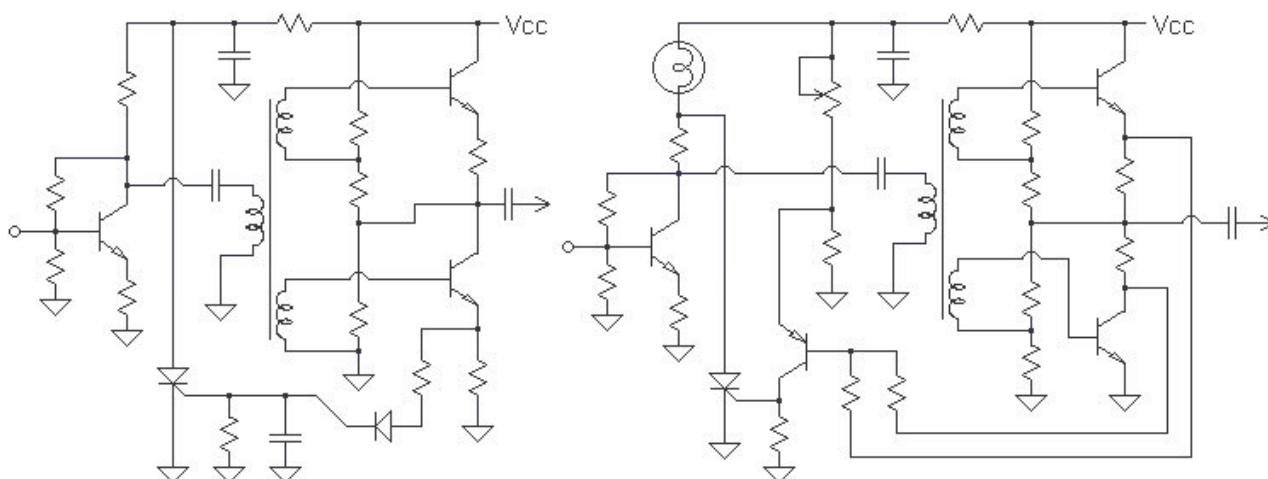


Figure 3.81 Thyristor latch was often used in protecting totem pole circuits. Left: Unicord/Univox and right: Webster Electric.

The Roland O. Rheume's protection circuit, assigned for Webster Electric Company, is a bit older design dating back to 1966. It consists of a shunting thyristor and a detector circuit that allows tuning the protection to trip at a specific level. Once again the thyristor turns off the driver circuit by shunting its supply to ground - and while doing so the increased current flow "turns on" the indicator light bulb. This circuit was in fact the basis for Unicord's protection scheme, which practically just eliminated the transistor-based detector and omitted the light bulb indicator.

One drawback is that the circuit practically requires a complementary output in order to have the detector working symmetrically. The configuration shown in the schematic works but is not an option if more output devices are to be connected in parallel. Although the requirement for complementary output was pretty significant in the sixties it is a minor issue nowadays. The circuit is discussed thoroughly in corresponding US Patent 3,448,394. The patent application also describes a novel way to reset the latch circuit: Instead of a manually operated reset switch that momentarily opens the shunting path, a switch with bimetal contacts reacts to the heat from the lamp and resets the thyristor.

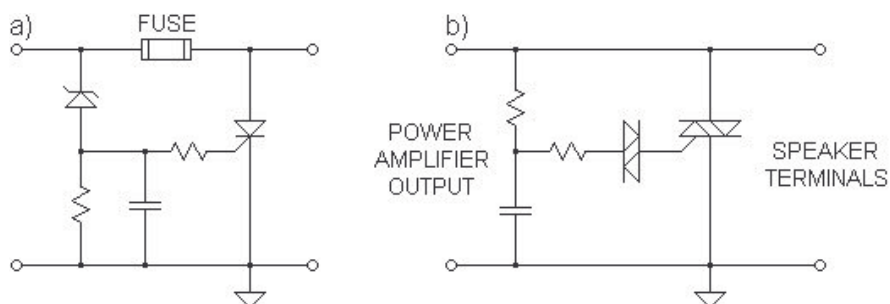


Figure 3.82 Crowbar "protection" circuits: a) Supply rail and b) speaker crowbar.

Since the topic has drifted to old-fashioned protection circuits it might be worthwhile to present a "classic" one: The infamous "Crowbar", depicted in Figure 3.82. The circuit's principle of operation is practically identical to that of a thyristor latch, except that during a fault condition the circuit has a "crowbar" (meaning rough and powerful) method of protection: A short circuit. At first this approach may seem more

than a bit odd but let's consider the operation of circuits a) and b). In the case of *over-voltage* condition the circuit a) will short-circuit the supply rail, causing the high current to blow the protective fuse. This seems like a strange protection method but is extremely useful with digital (and computer) circuits that are more prone to being damaged by over-voltage rather than under-voltage. Practically, the main principle of crowbar protection is to sacrifice the cheap parts in order to save the more expensive ones.

The circuit b) protects the speaker system from DC: The DIAC is triggered to conduct usually at a moderately high voltage, which is further defined by the RC-filter. If there is an excessive DC potential at the output the DIAC will conduct and trigger the TRIAC, which is a two-way thyristor. This, in turn, shorts the amplifier's output. Crowbar is a violent but efficient method to *protect the speaker system* and will likely cause the destruction of output devices. Very often the enormous short circuit current will destroy the TRIAC as well. One must remember that these circuits are products of their time and were once considered pretty sophisticated. In the old days the cost of repairing an amplifier was also a lot smaller than the cost of a new speaker system. Likely it was also assumed that the output stage was faulty already (after all, there was DC in the output). However, this was usually a *wrong* diagnosis! Nowadays we luckily have alternative and better protection methods than using crowbar circuits.

3.14.8 Protecting speakers from DC

The most common methods of protecting speakers from harmful DC offsets are fuses, crowbars and switches. In practice, fuses and crowbars are less-ideal solutions while switches can be categorized into two groups: Those that break the output line and those that shut down the power supply either partially or completely. Typically switching is done with good-old, faithful relays but nowadays a choice to use MOSFETs instead is gaining more and more popularity. A more esoteric solution is the use of PPTC “poly-switches”. Since these are considerably non-linear devices it is unfortunately not a very good practice to hook them in series with the speaker.

Figure 3.83 shows one way of implementing a conventional protection for speaker system. In essence, there are numerous configurations how such protection could be realized, however, most of them tend to follow similar concept of operation as the circuit in the schematic example. How does it work: Capacitance C1 is set large enough to pass all AC signals through but it begins to charge whenever there is a considerable DC potential at the output. The RC filter consisting of R1 and C1 naturally dictates the delay of operation along with other related characteristics. Capacitor *must* be non-polar (e.g. two electrolytic ones in series). Whenever the DC potential increases high enough to exceed the forward voltage of the transistors the base of the relay-driving transistor is clamped to near zero-volt potential. This consequently turns the relay-driver transistor *off*, thus releasing the relay. Effectively, the relay-controlled switch should therefore be normally open type – especially because in this arrangement failing of the protection circuit leads to a safe condition. During regular operation the base voltage is high enough to saturate the transistor, which conducts and energizes the relay.

To prevent “hysteric” operation, it is important to feed the DC detector from a point that is prior to the relay-actuated switch. Because clipping of asymmetrical signals (like that of guitar's) may shift bias from the zero level, thus causing the output signal to have a DC component, it may be wise to design any DC detector to trip only when

there is a moderately high DC potential at the output (i.e. higher than 5 V). However, do notice that this always reduces the level of protection.

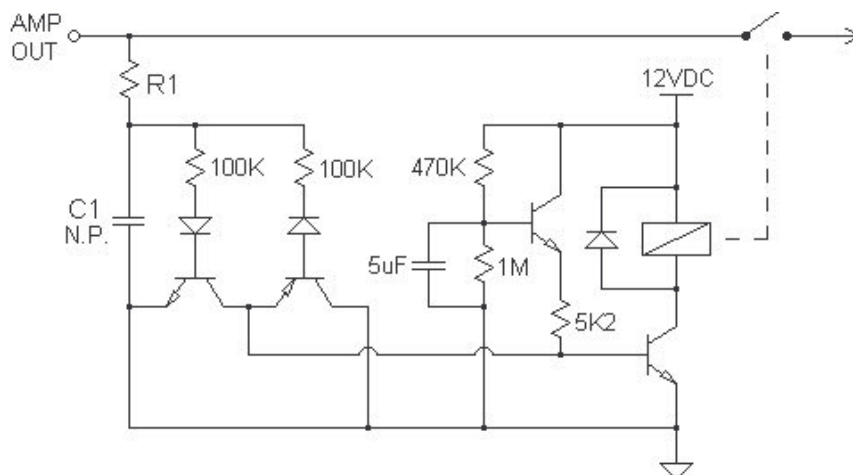


Figure 3.83 DC detection circuit controlling a solid-state switched relay.

Another transistor in conjunction with a second RC circuit provides a delayed speaker load-switching feature (startup muting). When the amplifier is turned on the 5- μ F capacitor begins to charge through the 470-kilo-ohm resistor. This gradually increases the voltage at the base of the relay-driving transistor. The time constant is defined so that the transistor will saturate – and the relay will become energized – only after about three seconds from switching on the power.

Note: In a fault condition the relay must break DC current. This is an important detail because relay's current ratings are lower for DC than for AC. Breaking the inductive speaker load also creates a transient, which is another good reason why the current handling capacity must be generously rated. Ignoring these issues may result into a failure where the relay contacts are welded together due to arcing. Consequently current path to speaker will not be broken and DC destroys the speaker.

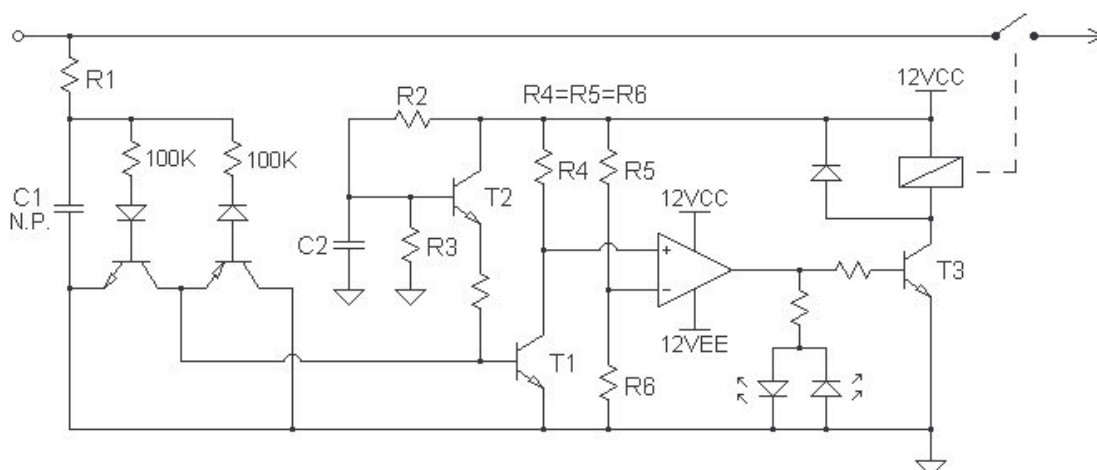


Figure 3.84 DC protection with an integrated startup delay and LED indicator.

Of course there is always room for improvement, like in the form of designs presented, for example, by Sam Electronic Circuits or Elektuur semiconductor guide: A comparator circuit, which controls the relay driver T3 and two anti-parallel indicator LEDs, replaces the original driver transistor configuration. The operation of a comparator is discussed in more detail in 7.3. According to the state of the driver transistor T1, the output of the comparator assumes the potential of either one of the

supply rails. The state of transistor T2 controls the startup delay or “soft start” feature: The delay is naturally set by the charge time of RC filter consisting of R2, R3 and C2. The described circuit is depicted in Figure 3.84.

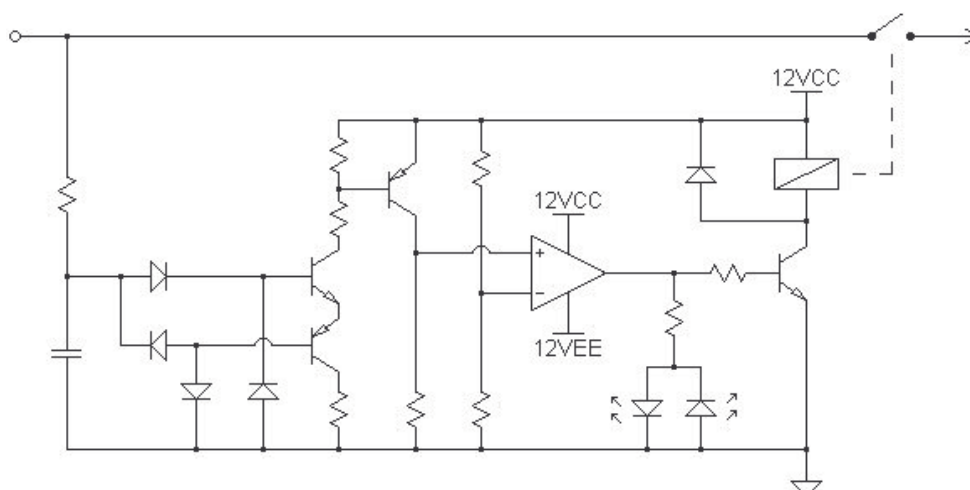


Figure 3.85 DC protection using an alternative detector circuit. Soft start feature is omitted.

Since there are always those who abhor excessive complexity or circuits they are not familiar with – and because of that may even decide to omit the protection - it is maybe worthwhile to mention that such circuits also exist in a convenient IC form. These ICs are extremely reliable and fast in their operation. Naturally they require only a handful of external components. Furthermore, they are typically packed with tons of features such as soft start, relay driver and detectors for DC offset, overload and loss of AC supply. Some examples of such chips are μ PC1237 from NEC Electronics Inc., NTE7100 from NTE Electronics (this is essentially same chip as the μ PC1237) and TA7317 from Toshiba.

3.14.9 About fusing

While fuses are protective devices it is crucial to realize the fact that fuses are mainly used for preventing major hazards such as fires or complete meltdowns caused by overheating, short-circuited components or other alike. A typical fuse (even a fast-acting one) operates too slowly to protect extremely fast solid-state devices during fault conditions; it is even too slow to provide an adequate protection against an electric shock from the device. This should not be taken as an encouragement to omit fuses – on a contrary. However, do note that protection of components (and especially the user of the device) has to rely on faster and more reliable methods than just fusing. Fuses might save a few resistors, circuit board and the transformer from a major catastrophe, and at least prevent the user from forming a high current short for longer time than few milliseconds – but even this small amount of time can be fatal in some occasions. Fuses do work but they are not the best or even the safest solution in every scenario.

Some amplifiers fuse the secondary-side both pre and post the rectifier-filter circuit. Fusing the secondary is a necessity since primary fuse is typically rated in a way that allows the power supply to provide very high secondary current for a prolonged period of time. There are many reasons for such “generosity” – the most important ones being the effects of high inrush- and turbulent charge currents of the filter capacitors. Because the primary fusing must tolerate the transient-like (high) current draw of the secondary it, in turn, becomes incapable of reacting quickly enough to a

sustained high current draw. Turbulent charge- and inrush currents also affect over the pre-rectifier fuses F2 and F3. These are used for protecting the transformer in case one of the filter caps or the rectifier short circuits.

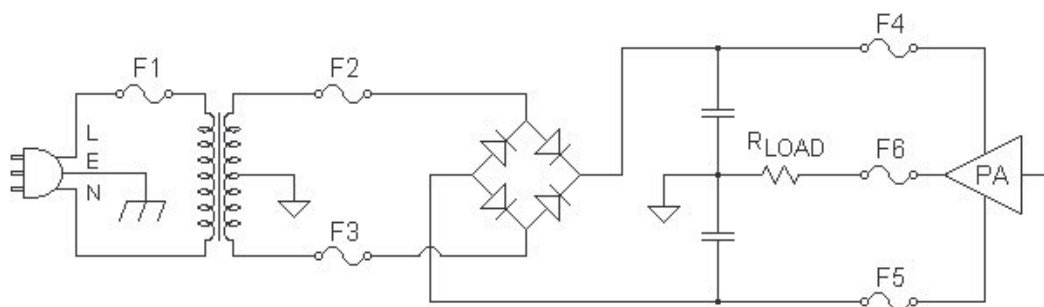


Figure 3.86 Fuse map of a dual supply amplifier circuit.

Rail fusing after the rectifier-filter section is extremely good practice *in single rail amplifiers*. However, in dual supply amplifiers the rail fuses pose a considerable problem: If both fuses do not blow simultaneously the circuit will see a considerable DC unbalance and almost the whole rail voltage will appear across the speaker terminals. In such case the other fuse may blow due to high current demand over the speaker – or it may not! Unless the circuit has a protection against excessive DC offset the speaker is in the risk serious damage. Generally, this is not an excuse to omit rail fusing – rather a drawback that needs to be compensated somehow.

Fusing speaker line is generally a very poor idea: A fuse is essentially an inductor with some resistance (without the latter the fuse wouldn't heat up and burn), thus inserting one in series with the speaker will impede the damping factor and create non-linearities: It has been measured that fuses introduce distortion at low frequencies [3]. Furthermore, even greater problems are created by the inherent fuse behavior: A properly protective fuse is usually fast-acting and has a low current rating, unfortunately such a fuse will likely trip during bass transients. A fused speaker line is a *cheap* protection and overshadowed by alternative protection methods – such as use of relays that trip during faults.

3.15 MOSFETs and IGBTs

MOSFETs are divided into vertical and lateral two types referring to their internal fabrication method. Vertical MOSFETs are optimized for switching applications and therefore they are generally quite unreliable if operated in the active region. Linear amplifiers *can* be made using vertical MOSFETs but usually these are extremely difficult to design and offer no considerable benefits in comparison to conventional BJT designs. Lateral MOSFETs are optimized for linear operation and therefore they can be used in conventional class-A or class-B/AB amplifiers. In essence, most MOSFET amplifiers only use the concerned semiconductor type as output devices and rest of the circuitry is almost without a question realized with ordinary BJTs or – in very rare cases – with FETs. Therefore MOSFET amplifiers basically use the same Darlington, Quasi-complementary or Compound follower configurations we are already familiar with. Due to some imperfect properties of MOSFET transistors the topology where VAS would drive the MOSFETs directly (in plain source follower configuration) is pretty rare – albeit it does work. However, MOSFETs do use slightly different topologies than their bipolar counterparts when it comes to driver stages, thermal stability and short circuit protection.

Before current will start to flow from drain to source MOSFETs require that their gate voltage is a certain amount higher than the source voltage. This voltage is known as *enhancement voltage*. All MOSFETs used in audio applications are enhancement type so they are practically operating in similar manner as BJTs. However, V_{GS} can be in the order of about 4 volts, which creates a serious limit for the output voltage swing unless the driver stage is powered from higher rail voltages than the output devices. To overcome the considerable voltage drop over the MOSFET the supply rail voltage of the input-, voltage amplifier- and driver stages should be approximately $\pm 6 - 10$ V higher than the rail voltage powering the MOSFETs.

Besides the requirement for higher driver stage voltage, MOSFETs have the following drawbacks: They are usually at least twice as expensive as bipolar power transistors. Besides that, they are equally rare as FETs and likely have to be ordered, which further adds to the cost in the form of postal charges. MOSFETs have low gain: This sets a demand for higher gain in the driver circuitry in order to keep the global feedback at acceptable levels. MOSFETs also have a high gate capacitance. This can be in the order of few nanofarads, which essentially creates a huge instability issue and stresses the driver circuitry at higher frequencies. They are also non-linear at lower currents and to reduce crossover distortion MOSFET output stages generally require higher quiescent current. MOSFETs have huge variation in device parameters and require an elegant matching circuitry or preferably have to be handpicked. The layer of metal oxide insulating the gate from the channel is also vulnerable to damage caused by static discharges or excessive voltages. This drawback has to be accounted during the assembling process but also in the design: It is common practice to hook a Zener diode between source and gate to protect the layer from high voltage transients. To prevent the non-linearity of Zener diode's junction capacitance from affecting performance a conventional diode must be fitted in series with the Zener.

The advantages of MOSFETs are speed, rounder transfer curves, easier driving, high current conducting capabilities, no beta-droop at higher currents and, *in case of lateral MOSFETs*, a transition to *negative* temperature coefficient (of current) after drain current exceeds about 200 mA. The latter character makes *lateral* MOSFETs nearly free from the destructive secondary breakdown - however, *only in linear circuits*. In switching circuits the negative temperature coefficient is actually a great problem and may even cause thermal runaway. Vertical MOSFETs (like HEXFETs) are generally worse when it comes to thermal stability because their transition current (also known as "inflection" point) is much higher. Therefore they are quite prone to thermal runaway.

If gate capacitance and requirement for higher driver supply voltage are ignored MOSFETs are generally easier to drive due to extremely low demand of gate current. Their clipping characteristics are also rounder than of BJTs – but not "tube-like", as many salesmen eagerly like to claim. Knee of clipping is the last thing to worry when compared to other clipping-related problems such as rail sag, rail sticking and parasitic oscillation the latter may introduce. In most cases even the term "MOSFET sound" is an overstatement.

The negative temperature coefficient is actually the most ideal characteristic since it solves many issues related to thermal stability. An increased temperature will automatically limit the current, thus preventing thermal runaway. This does not mean that MOSFETs could cope with less heat sinking but it means that since bias voltage no longer has to track the device's temperature many complex issues concerning

accuracy of thermal tracking and layout are eliminated: Lateral MOSFETs do not need to share the same heat sink, neither do they need high-wattage source resistors since the thermal coefficient characteristics automatically correct any imbalance in current sharing [43].

Figure 3.87 shows two most typical MOSFET output configurations; unsurprisingly they are almost identical to their bipolar counterparts. The bipolar drivers in the “Darlington” circuit work as emitter followers and isolate the VAS from the output stage. The buffer stage assures fast charging and discharging of gate capacitances and creates a low source impedance to drive the MOSFETs. As a result the bandwidth increases. The source resistors are not necessarily needed but are in most cases used anyway because they allow hooking up a current sensing scheme of the short circuit protection. In MOSFET designs the resistors are usually quite small - only about 0.2 ohms - and using them in configurations that utilize parallel output devices is recommendable to aid current sharing. Gate resistors are used for isolating the gate capacitance and with MOSFETs they are practically mandatory. Basically the circuit would work as plain source follower without the buffering transistors but due to lower transconductance of MOSFETs that would lead to increase in distortion. Also, the gate capacitance would impose a greater problem since VAS would have to drive the fairly capacitive load directly. Input-, VAS- and driver stage are powered from higher rail voltages than the MOSFETs to ensure sufficient driving.

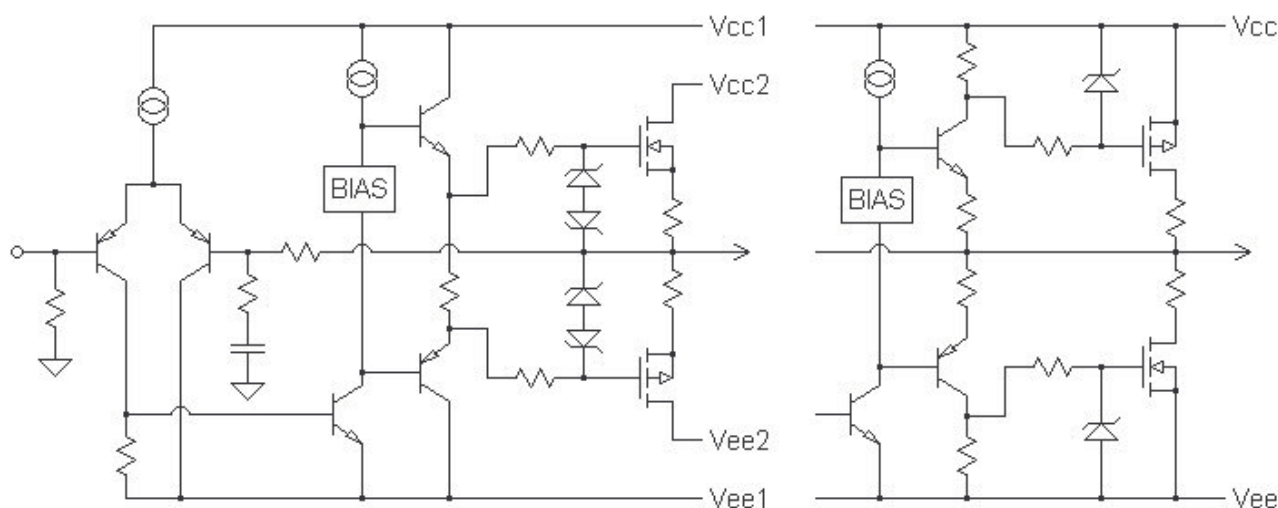


Figure 3.87 Examples of output stages using MOSFETs. Left: Buffered (Darlington) configuration and Right: MOSFET Compound follower or Complementary-Feedback MOSFET OPS.

Note that in schematic symbols of MOSFET the arrows point in opposite directions than in BJT symbols – similarly as they do in symbols of junction FETs. Thus a symbol where arrow points “inside” is a symbol of N-channel MOSFET.

Schematic on the right shows a MOSFET compound follower output stage. Sometimes people refer to it as *hybrid* or complementary-feedback stage. With a similar Sziklai configuration it would be possible to construct a quasi-complementary output stage but in practice these are troubled by high transconductance offset between the output devices [34]. If several MOSFETs are used in parallel the Sziklai arrangement may have to be configured to provide slight voltage gain that compensates the phase-shifted output signal from MOSFETs being fed back to high-gain BJT stage [34].

Assuming the output devices are lateral type, the biasing arrangements for both circuits are not overly demanding. Basically the BJT drivers benefit if their temperature is tracked but in the case of lateral MOSFETs the accuracy is not critical: MOSFETs are not as sensitive to precise bias as BJTs are [34]. The bias circuit might consist of a few series diodes and a trimmer resistor but usually the good old VBE-multiplier tends to outperform all of its alternatives. In case of *vertical* MOSFETs the use of VBE-multiplier and accurate temperature tracking is of course mandatory.

Insulated Gate Bipolar Transistors, also known as IGBTs, are a mix between MOSFETs and bipolar transistors. Their greatest advantage is combining ideal properties of both devices, resulting into a device that has the high power gain of MOSFETs but the low saturation voltage of bipolar transistors. The devices were designed to be *switches* and therefore they are not very good choice for output devices of a linear amplifier. An IGBT is basically a MOSFET driving a high-power bipolar transistor in Sziklai configuration with drain connected to base and source connected to emitter. However, it omits the “speed-up” resistor (which provides a faster response and better linearity) and one cannot be added externally. Also, IGBTs require close matching. This can be problematic because the properties of *both* the bipolar and MOSFET transistor need to be matched. IGBTs are as vulnerable to secondary breakdown as ordinary BJTs, have the capacitive problems of MOSFETs and lack a selection of decent complementary devices. Worst yet, the internal structure creates a “parasitic” transistor that latches-up the output if the collector current of the bipolar transistor exceeds a certain limit. This latch up may cause a catastrophic failure! Truthfully, there are not many features that would encourage or justify the use of IGBTs in linear audio applications. Naturally very nice sounding power amplifiers can be (and have been) built using IGBTs but, as with any esoteric component, the design process will be quite difficult.

3.16 Signal limiting within power amplifier circuit

All solid-state power amplifiers benefit from having their output signals limited to a level that they can handle without clipping. This is a bold statement but very true: The transfer function of a bipolar transistor is very linear up to the point of saturation, which causes transistor circuitry to have sudden, gross and unpleasant clipping characteristics. Some bad designs will exhibit even more awkward symptoms; they may introduce short bursts of oscillation (ringing), have a slow recovery from saturation (“rail sticking”) or they may shift their DC offset point. Some of these symptoms just generate additional high order harmonic content and plainly sound terrible; some are destructive to the amplifier or the speaker load itself. Most PA amplifiers are equipped with clipping indicator and limiter circuitry, most bass amplifiers have a limiting compressor; surprisingly most guitar amplifiers have none of these protection schemes! There is no excuse for that! Even when there is a need for a very “clean” output the fact is that the amplifier is introducing distortion when it is clipping. Most people cannot detect compressive distortion if it has a soft “knee” – a characteristic associated with tube amplifiers – but they will easily detect compressive distortion with a hard knee (characteristic “solid-state” clipping).

Figure 3.88 depicts an old-school signal limiting circuit. An attenuated output signal is rectified and used as a drive signal for LED. The LED controls the light dependant resistor (LDR), which’s resistance drops with increased luminosity. The LDR is a part of voltage divider network in the amplifier’s input and therefore controls the

amplitude of the input signal. The control signal is tied to common reference with a resistor and filtered to achieve a slow decay after signal peaks.

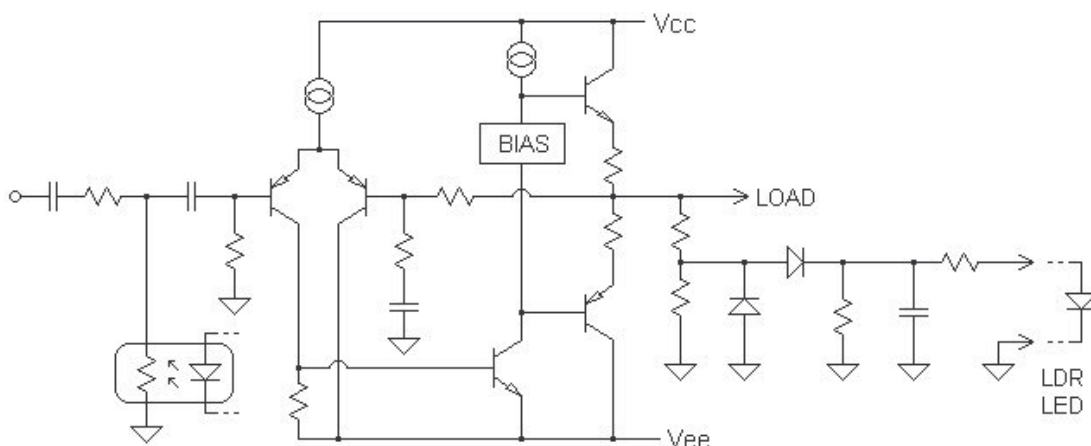


Figure 3.88 A simple limiting arrangement based on optoelectronics.

The presented circuit is crude, yet amazingly effective. On “full gain” the limiter causes a noticeable “crunch” to the signal but this is not as harsh as distortion caused by clipping. As long as the circuit is treated as what it is – a peak limiter – and not abused by maximizing the amplifier’s volume it works very well and is almost transparent. As always, it could be improved by numerous ways: The half wave rectifying “detector” circuit has a poor “attack” since it can detect signal peaks only from the other “side” of the waveform. Also, the circuit would work better if it divided the audio signal to separate frequency bands of bass, midrange and high (perhaps even more) and treated each one separately. The time constants of attack and decay could be tweaked as well. However, one must remember that this circuit is not aiming for HiFi reproduction but to limit the roughest peaks that might cause the amplifier to clip. Substituting the slow LDR with a FET or an OTA, and replacing the detector circuit with something more efficient and accurate, could vastly improve the circuit. The section concerning soft limiters will describe these kinds of limiting methods more thoroughly.

The “light bulb current limiter” trick is also quite handy: The resistance of a light bulb will increase when increased current heats up the filament. In series with the load (either in the output or in the input) this will cause a current limiting effect. However, this effect is even slower than LDR circuits. Since most power amplifiers (often) have a constant voltage gain – regardless of load impedance - the lamp can mainly protect the speaker from excessive current but not from clipping due to exceeding rail voltage limits. Light bulbs are pretty difficult to implement since different lamp brands are quite dissimilar. Also, the lamp has very odd impedance that, together with the speaker, might cause exceedingly peculiar loading conditions. Finding a proper lamp for the application wouldn’t be a problem if manufacturers would publish specs of their products. Unfortunately most consumers are happy with the sole information that the bulb lights up at certain voltage. 12 V “grain of wheat” or “pilot” lamps are the most commonly used devices for this kind of application.

Another very common form of limiting is to control the transconductance of the input differential. The idea is often harnessed in creating a simple “mute” circuit that provides a noiseless start but the very same circuit is applicable to form a circuit that controls the gain of the power amplifier according to certain control voltage. The details of this type of circuitry and more insight to limiters can be found from section 6.12.

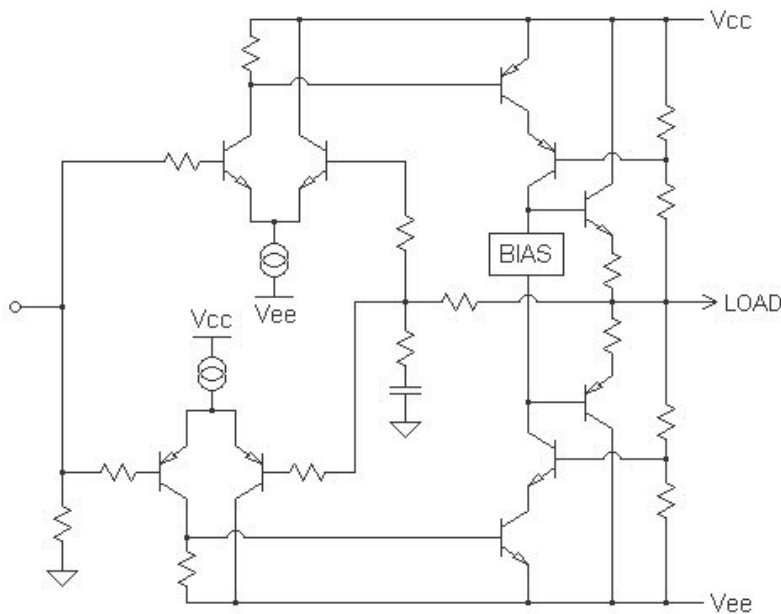


Figure 3.89 "Pseudo-cascoded" VAS exhibits soft clipping characteristics.

A stranger solution to acquire signal "limiting" within power amplifier stage is to use a "pseudo-cascoded" VAS; the solution only works with symmetric cascode circuits, though. The cascoded voltage amplifier stages provide active loading for the push-pull emitter follower and high output impedance (which may be better suited for driving MOSFET outputs). Unlike other cascodes, the circuit lacks a stable voltage reference, hence the name. The reference changes according to voltage potential between output and rails. Since the transconductance of the VAS is that way made variable, the gain decreases when output swings closer to rail. This explains the need for a symmetric voltage amplifier stages: One stage cannot track both rails and the result would be softened clipping only on one of the half waves. Being a "soft" clipping-type limiter, the circuit does not provide as clean output as limiters that work by attenuating the input signal or by controlling the gain.

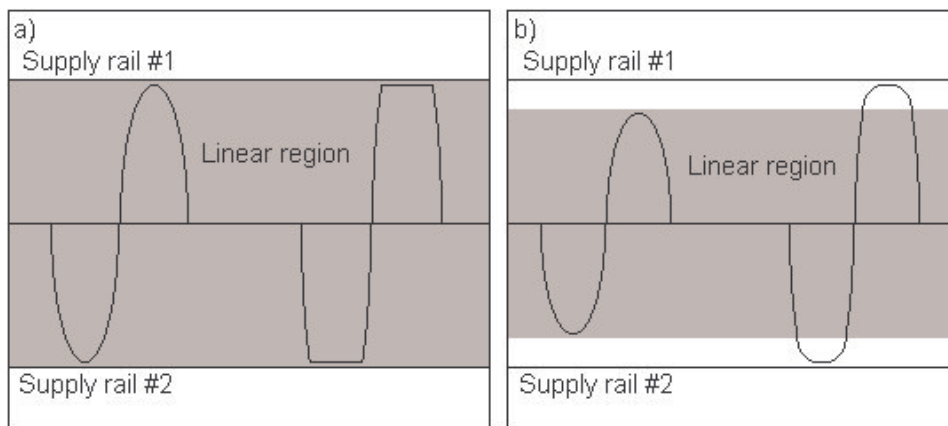


Figure 3.90 a) Greatly linearized amplifier can swing close to its supply rails but has an abrupt onset of clipping. b) Soft clipping narrows down the linear region and therefore decreases the amount of output power at a given percentage of THD. Linear region is shown in gray.

One must note that when powered by an identical supply *all* softly limiting (read: clipping) amplifiers – and this including tube amplifiers as well – are less powerful at their linear region in comparison to highly linearized amplifiers. Negative feedback can extend the linear region to cover nearly the full rail-to-rail voltage swing but this naturally means that the clipping onset becomes more abrupt. An amplifier can't

swing past its supply. Thus to produce an equal amount of output power (at the same level of THD) a soft limiting amplifier needs a more powerful supply. It needs a higher supply voltage than what the linearized amplifier has and more current delivering capacity to keep up with the increased voltage swing. This is one of the reasons behind “tube loudness”: Beyond the linear region of operation the power supplying capacity is just not comparable. Thus an amplifier that produces 60 watts when swinging rail-to-rail cannot use a soft clipping scheme to replicate the behavior of an inherently soft clipping 60-watt amplifier. For that it would need more power reserve.

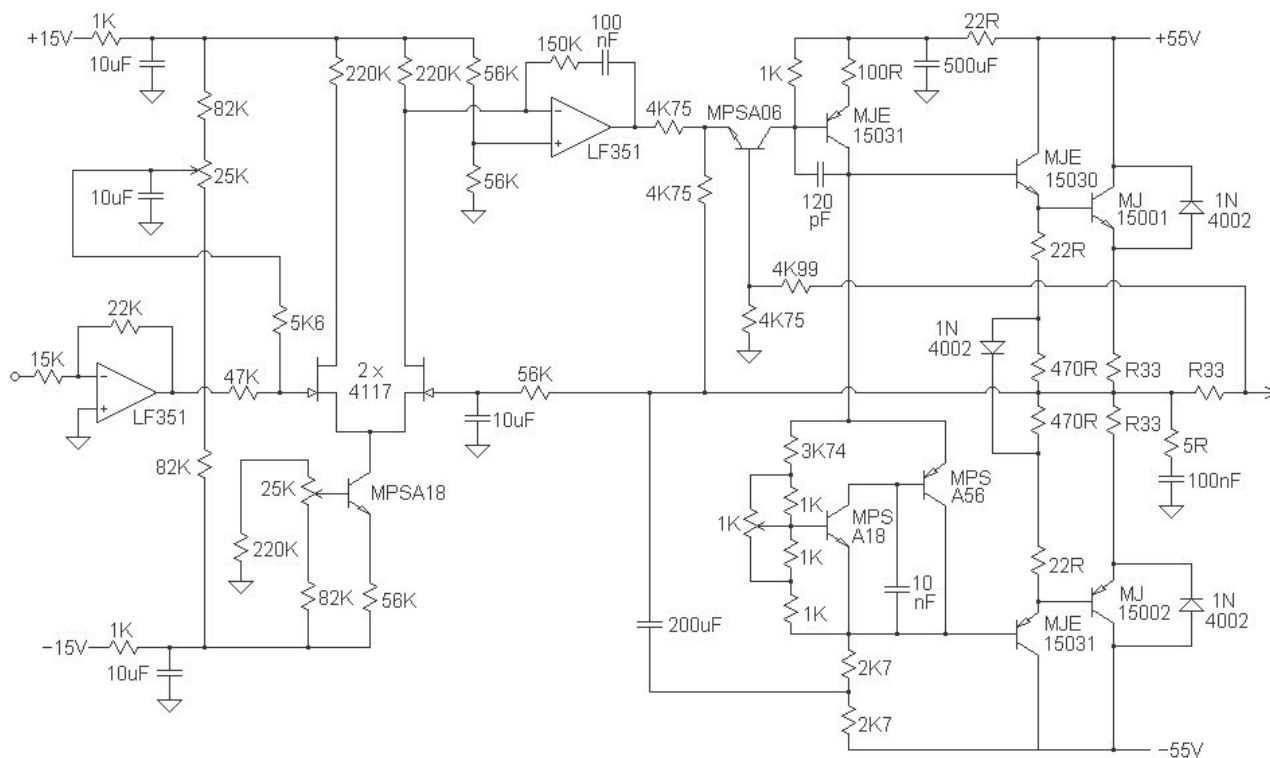


Figure 3.91 Gallien-Krueger 112SC/150GT soft clipping power amplifier. The input circuit is discussed in greater detail in section 6.12. Output devices are doubled in 212SC/300GT.

3.17 DC Servos

DC servo circuits are utilized in keeping the DC offset voltage in the amplifier’s output at extremely low level. Basically, the circuit adds an error correction voltage to the input signal’s common reference, thus a DC servo circuit can provide far lower offsets than the amplifier circuit alone. It also allows eliminating the non-linear, (typically electrolytic) unity DC gain capacitor from the feedback path.

The OpAmp compares the difference between its ground-referenced, non-inverting input and inverting input, which is hooked to the output of the amplifier. Notice that due to C1 the operational amplifier is not configured as comparator but an integrator and the output is not “latching” to rails. A typical error-correcting configuration couldn’t be used as DC servo since it would consider all signals different from ground reference (zero volts) as an error. Because integrator is essentially a low pass filter, which has a very high gain at low frequencies, the error correction is configured to happen below the audio band.

The resistance of R2 and R3 should be about equal to keep the input references in “balance”. Ideally they should also have the same resistance as the input resistor R1,

which in turn should have same value as R4. The purpose of R3 is to compensate the offset current of the OpAmp itself and it can be omitted, although this solution is less ideal. C2 bypasses AC signals to ground and should have a fairly high value, typically few hundred microfarads. C1 is selected so that there is no considerable attenuation in the audio band. Preferably the -3 dB is set at 1 Hz. Its value typically ranges from hundred nanofarads to few microfarads.

Components R5 and C3 are also optional: Most cheap OpAmps have a very poor PSRR at high frequencies, which causes them to pass supply noise to the input – unless the supply is HF filtered well, of course. Splitting the input resistor and bypassing the lower part of it to ground with moderately high capacitance effectively shunts the noises originating from the servo. Unfortunately this happens at the cost of accuracy in DC offset correction. Since a DC servo usually provides nearly perfect zero volt output this would not cast a very big problem though. In a similar manner, DC servos may even have a series second order low-pass filter (for example, Sallen-Key), which further helps in removing power supply originated noises. This filter can be located before or after the integrator (or in both places simultaneously). The additional filtering is usually set to have a -3 dB point that is about ten times higher than that of the integrator's.

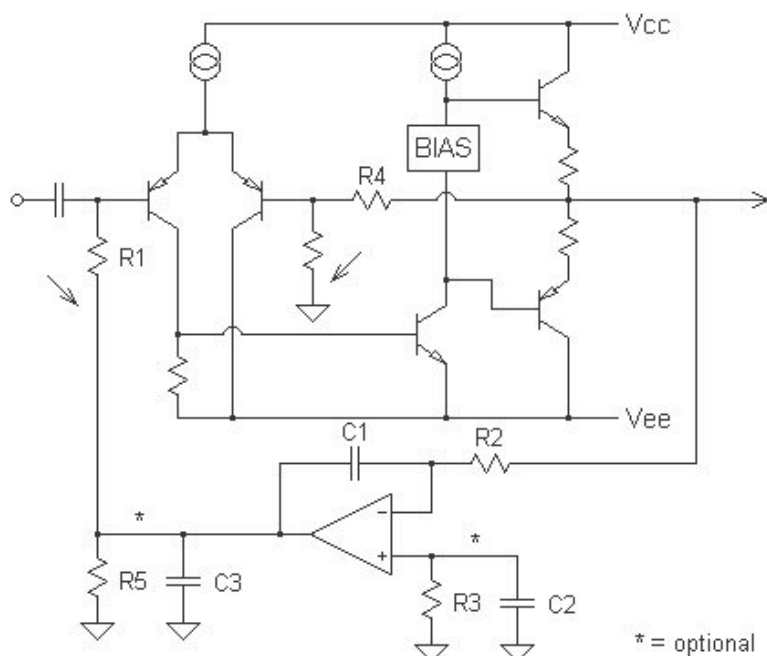


Figure 3.92 Typical DC servo circuit.

If the output of the amplifier can swing higher than the input range of the servo (which very often is the case), the servo's input should be equipped with a voltage divider, which limits the voltage swing to acceptable range. As an additional protection, the input of the OpAmp should be equipped with amplitude limiting diodes.

The basic design of DC servo circuits is actually quite easy and straightforward. Unfortunately the performance of these circuits extensively depends on smaller details, such as component properties. The servo OpAmp needs to be high-impedance and low input offset type. Adequate slew rate is also a fair requirement. Integrator capacitors should be non-polar and preferably high-quality film capacitors with low dielectric absorption (e.g. polypropylene, Teflon, polystyrene). Rest of the capacitors should have a minimum leakage current.

How useful a DC servo is can be argued. Careful balancing of input stage currents usually results into DC offset that is in the level of few millivolts. In theory, a DC servo could drop the offset to micro- or even nanovolts. In practice, such a low level is no more useful than having just a millivolt or two of offset. We are not talking about precise measurement instruments here but guitar amplifiers and as long as the speaker's voice coil doesn't suffer from the DC everything is all right. Some people feel that designs that require a servo in order to work properly should be rejected anyway. Amplifier circuits that do not – or cannot – utilize a DC controlling feedback loop are another issue. In those the servo circuit is likely mandatory.

3.18 Impedance matching transformers and 70-volt systems

A typical solid-state amplifier is designed to provide maximum power to certain load impedance, thus doubling the load impedance approximately halves the output power. This is not an optimal solution since flexibility that allows driving many different load impedances with an equal output power is a virtue. A guitar amplifier that can provide the same output power to 16, 8 and 4-ohm loads can be used with a larger selection of different speaker cabinets than a guitar amplifier that is rated only for a nominal, 4-ohm load and thus provides only about a quarter of its maximum output power to a 16-ohm load.

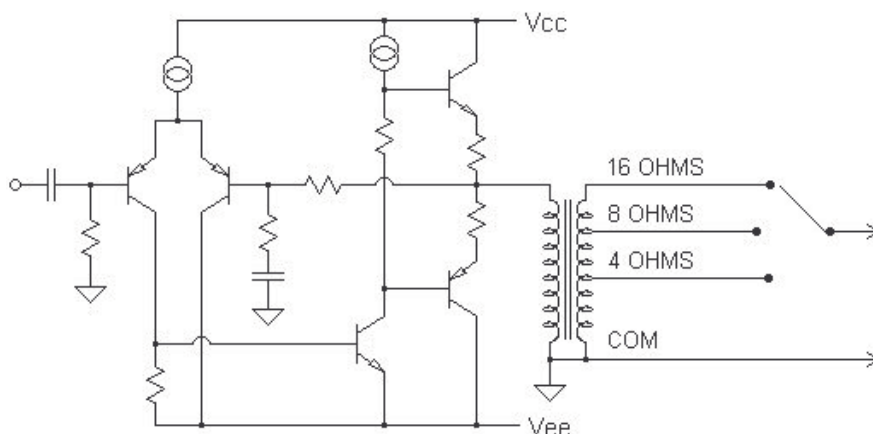


Figure 3.93 Impedance matching with shunt feed output transformer.

Theory behind impedance matching with a transformer is basically really simple: The transformer reflects the load impedance according to its turns-ratio. Thus with a proper ratio, the load impedance seen at the primary side can be set constant regardless of load impedance in the secondary side. The impedance transformation is derived from equation $(N_S/N_P)^2$, where N is the number of turns in winding. The equation happens to be the same for ratio of inductances as well. Thus a transformer with turns-ratio of 1:1 would reflect the secondary loading equally to primary. A transformer with turns-ratio of 1:1.414 would reflect *half* of the secondary load to the primary and transformer with turns-ratio of 1:2 would reflect a *quarter* of the secondary load to the primary.

In practice, the behavior is not so simple, though: Windings of the transformer have a considerable *series resistance* that seriously impedes the amount of current that can be fed through the winding. In short, using the impedance matching transformer always introduces some losses. It also limits the bandwidth and adds more cost and weight.

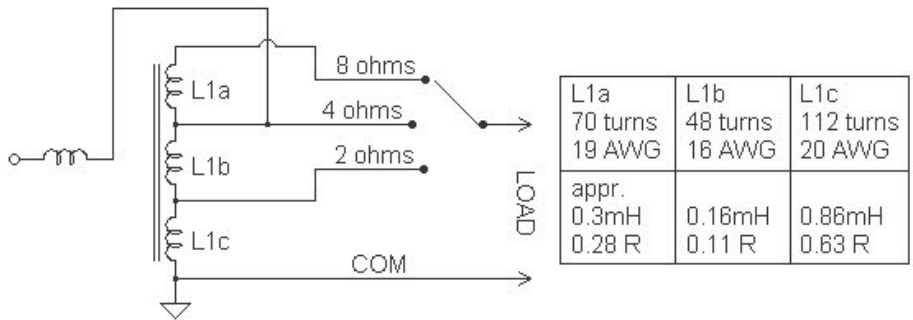


Figure 3.94 Unicord/Univox "Mobile Ohm" technology uses a simple autotransformer for impedance matching.

In 1974 Unicord patented an idea that allowed impedance matching without the use of conventional transformer that consists of a primary and a secondary [62]. The possibility to use an autotransformer, especially in a circuit where there was no DC flow through the windings, meant a considerable decrease in size and weight. Truthfully put, Unicord's patent was not really groundbreaking since the circuit had already been in use in paging systems for a while. Nevertheless, applied to solid-state guitar amplifiers the idea was "new". Like very often happens, the invention seems extremely simple when examined afterwards: A typical amplifier is optimized for specific load impedance. In order to provide matching to load impedances that are higher or lower than the amplifier's rated one, the output of the amplifier is connected to the middlemost "tap" of the autotransformer. Consequently, a voltage potential equal to amplifier's output now affects between this point and the ground reference. This voltage is induced to the autotransformer in proportion to number of turns in its windings. As a result, a higher voltage potential than the one of the amplifier's output now affects between the whole winding of the autotransformer. Correspondingly, the lowest voltage potential can be measured from the tap closest to ground. Ohm's law states that this allows extracting the same output power from different load impedances.

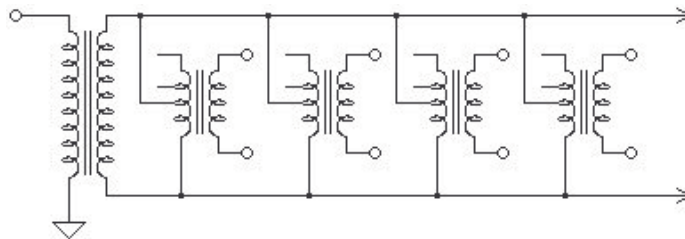


Figure 3.95 Principle of 70-volt sound system.

70-volt distribution systems are generally used in public address applications and in installations where several low volume speakers have to be fed from a single amplifier. Typical applications for distribution lines are paging systems so they are not something you would see in guitar amplifiers. Nevertheless, a small insight will do no harm: A step-up transformer is used for raising the output voltage level higher (simultaneously current decreases) so that the power can be transferred longer distances with fewer losses. The speakers are connected in parallel to the distribution system through step-down transformers. These are usually equipped with many primary side "power taps", which act like a stepped volume control. The transformer significantly increases the impedance, thus a vast amount of speakers can be connected to the distribution line without overloading the power amplifier.

Like impedance matching transformers, the distribution transformers introduce losses as well. As a rule of thumb, about 20 % of the output power will be lost over transformers and speaker cables.

3.19 Examples of hybrid power amplifiers

3.19.1 MosValve and its heirs

Some of you might remember the famous “Chandler Tube Driver” effect pedal or the MosValve series of amplifiers from Tube Works. The name behind these is Brent Butler, a man who has worked with tube and hybrid circuits for decades. At the time of writing he runs Butler Audio Inc., which manufactures high-end amplifiers for cars and households. The current product lineup utilizes puzzling “tube driver” and “thermionic” topologies. The huge success of the original Tube Driver pedal (that was introduced in the late 80’s) likely motivated Butler to employ the familiar name in his future designs as well. While Tube Driver was practically a hybrid, it still utilized a rather conventional overdriven tube gain stage, which was powered from a low voltage supply and enhanced with a clever “pull up resistor biasing” that softened the clipping behavior [64]. Reputedly the pedal’s circuit was later modified to suit car and home audio circuits. However, in this section we’ll be focusing on circuits that present a more distinctive “marriage” of solid-state and vacuum tube devices.

Butler’s view on what is the origin of “tube sound” is obviously the tube’s way to operate by the means of thermionic emission. Whether I agree with him or not, it is hard to argue with this kind of point of view: As long as the signal is, by some means, amplified or “coupled” by the tube the result is essentially *tube sound*. In practice, we can argue if different approaches to attain this “tube sound” are enough to create all nuances and characteristics we (as individuals) file under this extremely indefinite concept.

In the early 90’s Butler introduced a hybrid design that became the basis of MosValve topology and several other tube emulation circuits. Different embodiments of this invention are discussed quite thoroughly in corresponding US patents (refer to [65] and [66]). An operational amplifier is directly coupled to the grid of a triode; signal output and feedback are taken from the cathode. This places the tube inside the feedback loop. The tube and the operational amplifier are both powered from the same dual rail supply, thus the output is automatically biased near zero volts and the circuit can be directly coupled to the following stages. Effectively the circuit inherits the signal transfer properties of the grid-cathode “junction” – including the desirable non-linearity.

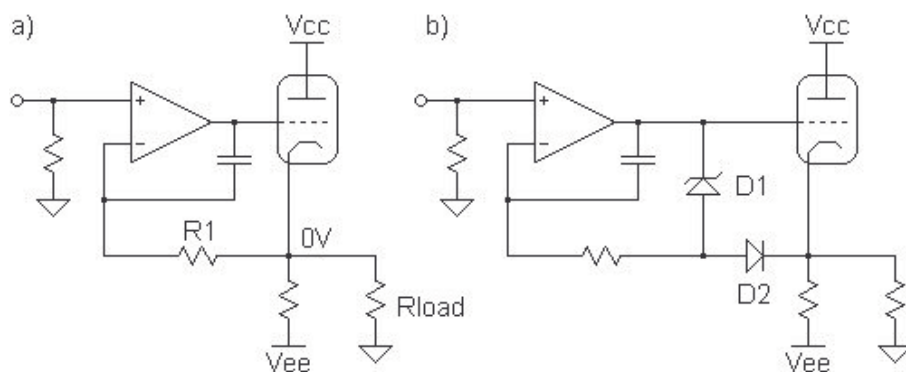


Figure 3.96 Vacuum tube in the feedback loop of an operational amplifier.

In the other embodiment Zener diode D1 is a bypass for the tube and diverts the output of the operational amplifier to the output of the whole hybrid circuit whenever the vacuum tube is warming up or has simply failed. The component is exceptionally vital because without it the DC balance during warm-up period would be extremely upset. Diode D2 is simply isolating the cathode and feedback loop from each other.

To prevent the readers from misinterpreting the following content, I must mention that Butler’s designs and attempts in tube emulation have generally been warmly received in the conservative guitar and audio scene. Yet it goes without saying that (aside driving tone control circuits) cathode follower is not a very typical configuration in ordinary tube amplifiers. Neither is the use of low supply voltage. When overdriven, a cathode follower will clip asymmetrically and due to its integral feedback mechanism it does this rather abruptly. (This is actually preferred by many people). However, feedback via “damping” resistor R1 tends to linearize the circuit. Butler uses feedback to dampen the inherent microphonics of a tube and depending on its amount it either removes a lot or very little of the circuit’s “tube characteristics”.

Essentially the circuit can be treated like a conventional operational amplifier, thus it lends itself to various configurations to drive a current amplifier output stage. One example of such circuit is shown in Figure 3.97. Note that the fundamentals of the following circuits have been discussed already in 3.5. The first embodiment uses “level shifting resistors” to virtually boost up the output signal from the hybrid operational amplifier. This allows the output stage to exceed the low “rail-swing-capability” of the OpAmp itself. The power amplifier is naturally powered from a higher voltage supply. The MOSFETs may (or may not) use a bit lower supply voltage than the BJT driver stage.

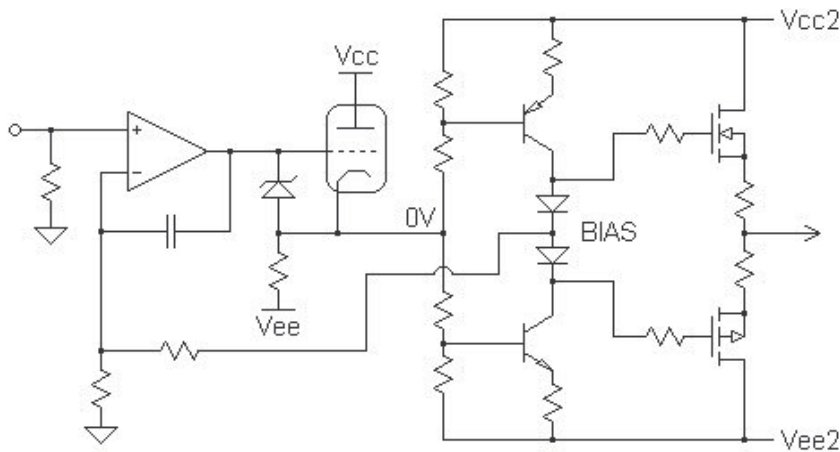


Figure 3.97 Hybrid OpAmp stage driving a MOSFET output stage, which is left outside the feedback loop.

In essence, the output stage is fairly basic having the only difference in the feedback arrangement, which is taken already from the drivers. This was not the Butler’s first attempt to emulate a tube amplifier by using a MOSFET output stage that was left outside the feedback loop [63]. In comparison to the traditional configuration, an open loop output stage has some unique characteristics, however, let’s make it clear since the beginning: Softer “tube-like” clipping is not one of them! Some added softness in clipping is due to using MOSFETs instead of BJTs but the clipping behavior is strikingly similar whether the output stage is inside the feedback loop or not. The most prominent effect from leaving the output stage outside the loop is the increase in

output impedance, which decreases damping and the overall gain. The “tube-like” effects this causes have been discussed earlier.

Another embodiment exploits the hybrid features more successfully: The output signal of the operational amplifier is now amplified by transistor T2, practically leaving the OpAmp to act as a mere differential stage. T1 is a compulsory common base “buffer” that is required for biasing the base of T2. The cathode follower is now directly coupled to the voltage amplifier stage and powered from the same “high voltage” power supply rail (typically 30 – 60 V). This is beneficial because the transfer function becomes more “tube-like” at a higher supply voltage. A conventional MOSFET output stage is directly coupled to the cathode. In some embodiments a BJT driver stage is used in between.

The circuit has two feedback paths: One from the output and one from the output of the VAS. Since the amount of feedback from the output stage is variable, the coloration introduced by the cathode follower and the more or less “open loop” output stage can be controlled. At the other end of potentiometer’s dial the feedback loop from the output stage is completely shunted to ground leaving the output stage open loop and VAS the sole control of negative feedback.

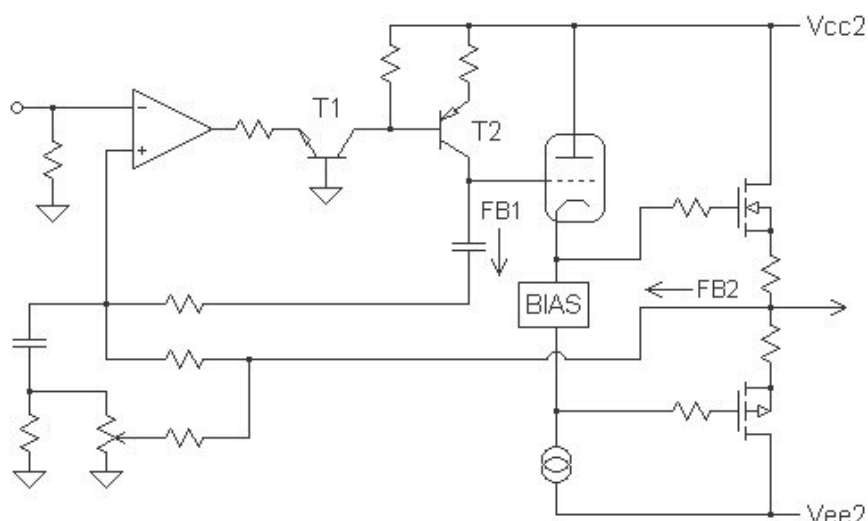


Figure 3.98 A discrete VAS is an alternative method to increase the output voltage swing past the OpAmp’s rail limits. The amount of negative feedback taken from the output stage is adjustable.

Great deal of other embodiments that follow the two basic ideas discussed herein are presented in US Patent 5,705,950. To summarize, it is evident that these circuits do not follow the average solid-state power amplifier topology, which ties the whole output stage into the tight control of negative feedback. However, it is equally apparent that the circuits are far away from an average, output transformer coupled tube topology as well. Undoubtedly the MosValve will have some unique sonic characteristics that render it different in comparison to both.

A patent for the “thermionic” topology, used in Butler Audio designs, was granted in 2003 [67]. This topology is quite different than the ones discussed previously and has evidently been influenced by the Fender’s patented vacuum tube limiter (US Patent 5,727,069). In characteristic manner to Butler’s designs, the output stage is once again left outside the feedback loop. Similarly to typical push-pull designs, the output stage is driven by positively and negatively biased signal sources. Because positive bias voltage and current are present at the grid and plate of V1, the tube in question is constantly near its saturation (or overload) region. At this configuration the tube does

not require a separate power supply. Since its cathode is directly coupled to the base of the driver transistor, the transistor amplifies current from the positively biased signal source. The complementary side, that amplifies the signal of the negatively biased signal source, acts similarly. In order to bias V2 near saturation it is hooked in reversed orientation when compared to V1. Zener diode D1 protects the output stage from tube malfunctions.

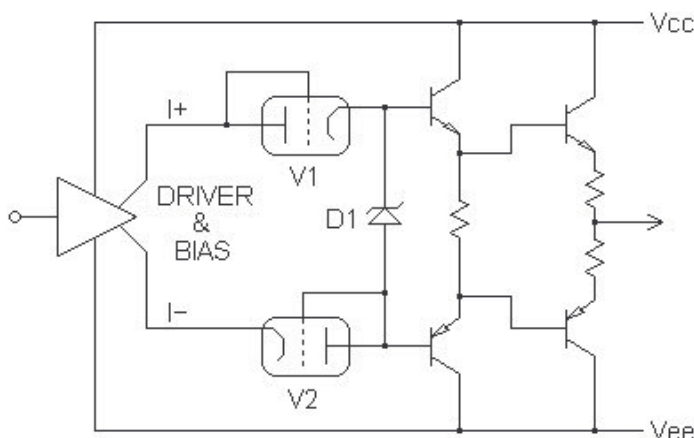


Figure 3.99 Fender’s patented vacuum tube limiter has clearly influenced “thermionic” hybrid topology.

Hooking the vacuum tubes to the bases of the drivers creates some useful features: When the signal from the driver stage is low in amplitude the thermionic emission generates additional bias voltage. Thus the tubes supply additional current during crossover condition, which makes crossover characteristics “softer”. In any properly designed amplifier this effect is rather worthless, though. However, a greater benefit comes from the fact that the tubes automatically try to achieve a proper bias, correspondingly decreasing the amount of bias voltage they introduce when signal amplitude is high. This behavior makes moderately imperfect operation of the bias servo more tolerable.

3.19.2 Peavey

Peavey belongs to the group of few selected manufacturers who chose to craft their early hybrid amplifiers in a fairly unusual manner: Instead of conventional, lighter and cheaper arrangement, where the power amplifier section was made using solid-state devices, Peavey amplifiers used a vacuum tube power amplifier. Well, at least sort of...

An archetypal Peavey hybrid utilizes a BJT long-tailed pair differential circuit as a phase inverter and a gain stage. Aside that fact, the circuit follows the conventional tube push-pull topology that has some major differences when compared to typical solid-state amplifiers: Because there is no complementary output stage there is no separate class-A VAS to drive it either. Instead, each half of the differential stage is essentially a class-A amplifier, feeding its corresponding class-B tube stage. The open loop gain of the differential is set to about 80 and its output begins to clip at approximately 80 volts. The requirement for high output signal amplitude is explained by the fact that a lot of voltage is required to drive the low gain power tubes fully on. The unfortunate detail is that high-voltage transistors are not very ordinary components. Output transformer performs the crucial impedance matching and “mirrors” the speaker load from the secondary side to serve as a properly sized plate load.

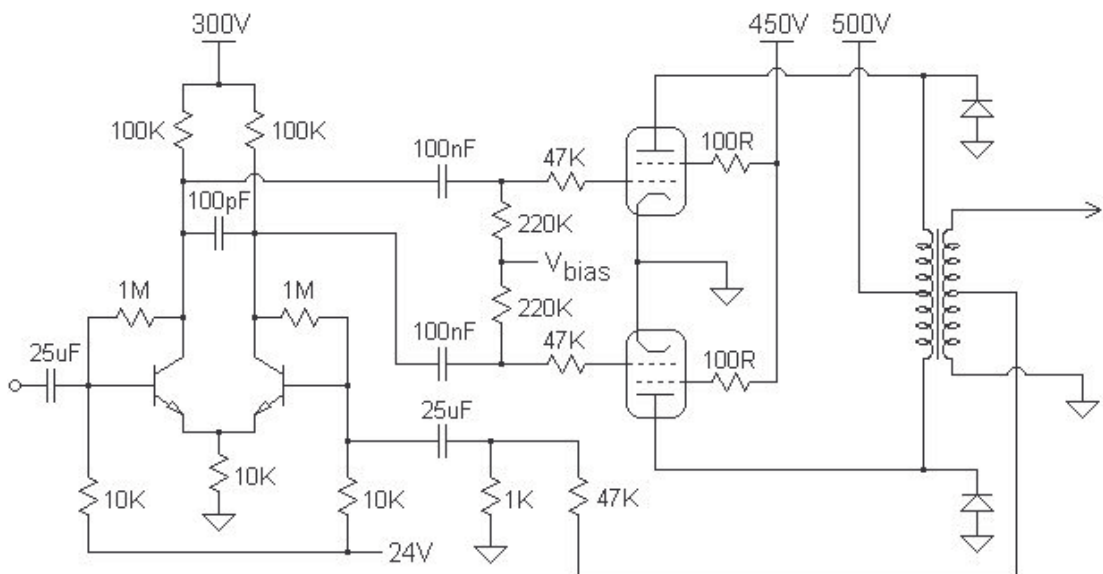


Figure 3.100 Conceptual schematic of a typical hybrid power amplifier from Peavey.

3.19.3 Music Man

From engineering point of view, Music Man amplifiers are likely the most interesting amongst hybrids that employ a tube power amplifier. Like usual tube push-pull circuits the output stage requires two drive signals with opposite phases. Instead of using one of the ordinary phase inverter circuits, like the long-tailed pair, split-load or even transformer, the Music Man hybrids split the signal path in two and reverse the phase with an inverting unity-gain operational amplifier. However, it became evident from Peavey design (refer to 3.19.2) that the power tubes require a rather high input signal to be driven fully on. This criterion cannot be met with an operational amplifier since the maximum drive amplitude it's capable of providing is only about 16 volts. The problem was solved with an arrangement to which many people refer as "cathode drive". Essentially it's just an ordinary *cascode*.

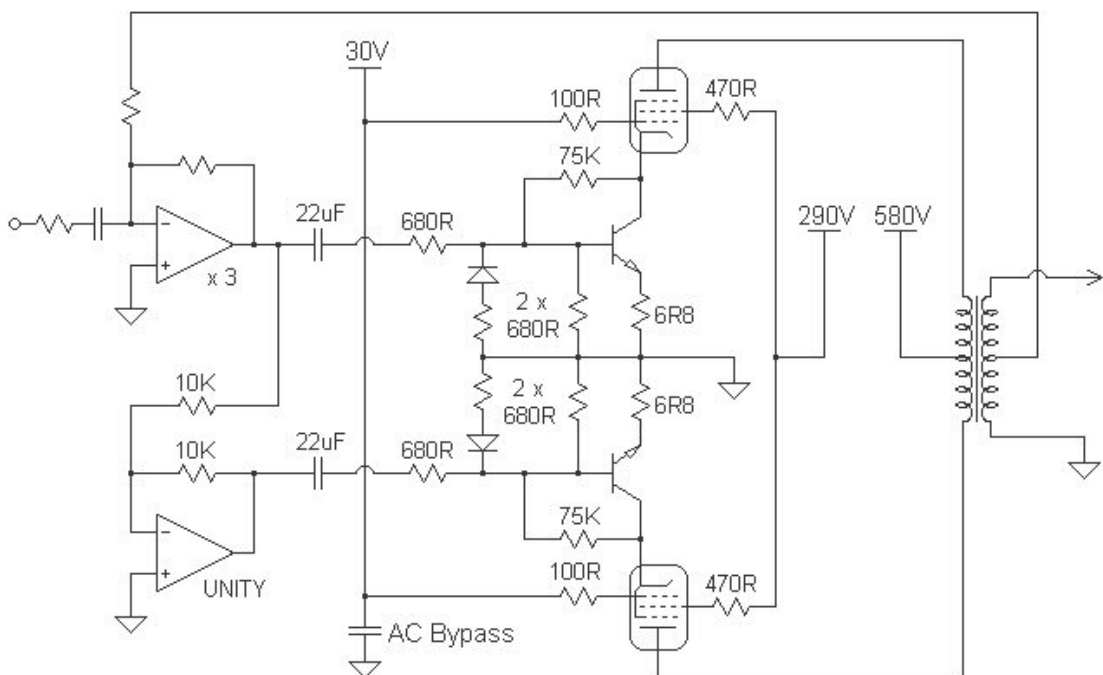


Figure 3.101 A schematic of a Music Man hybrid power amplifier.

The cascode in question consists of a BJT common emitter amplifier driving a power pentode common grid amplifier. In such configuration the pentode acts as a collector load for the BJT. Effectively, the transistor sees only a fraction of the total voltage swing and consequently high-voltage devices are not required (typical V_{CEO} of the transistors used here is about 70V). The common grid circuit is known better by the name of “grounded grid” and while the grid looks like to be tied to 30 VDC the supply’s smoothing capacitor ensures that the grid is, in fact, grounded when AC signals are considered. Since a grounded grid amplifier effectively cannot suffer from the effect of “Miller” capacitance the cascode has a very broad bandwidth. Gain of a cascode circuit is also very high (being the product of individual gains of common emitter and common grid stages), thus the low output signal amplitude of OpAmps is acceptable as an input signal.

3.19.4 ADA

During the period between the late 80’s and the early 90’s ADA Signal Processors, Inc. was one of the most prominent manufacturers introducing hybrid technology. The ADA MP-1 preamplifier is a classic in its own right, yet ADA also introduced few hybrid power amplifiers and an amplifier simulator known as the “Ampulator”.

The Ampulator follows the typical push-pull tube power amplifier topology although it works at much higher impedance levels and therefore cannot provide any considerable power amplification. The first gain stage and the unity-gain phase inverter amplifier following it are reminiscent of the Music Man hybrid, yet they include some interesting differences. Instead of being AC coupled to the following circuit the input stage plays an important part in forming a fixed-bias arrangement: The bias voltage is applied to the reference potential of the operational amplifiers, thus the signal rides a top this DC component. Due to DC coupling the grids are referenced to the bias voltage as well. A simple resistive divider allows varying the bias – and consequently the mode of operation from class-A to class-B. A comparator circuit is harnessed to compare the V_{BIAS} with the cathode voltage and to provide a real-time visual indication of the current class of operation.

The output stage works in push-pull and is the same transformer-based configuration used by countless of other tube amplifiers. The secondary side drives a high-impedance mockup of a speaker load (omitted from the schematic example) and also serves as the negative feedback takeoff point. Feedback loop houses a classic presence control circuit. The final stage is somewhat unconventional: The active devices form a tube-BJT compound circuit in which the transistor effectively acts as a unity-gain buffer for the tube. This increases current gain and allows driving lower load impedances. The B+ voltage of the Ampulator can be varied in order to acquire a variable amount of overdrive from the stage. This regulator-based control is not shown in the example schematic. The amplifier also includes a comparator circuit that provides visual indication of the signal level. Its operating principle is essentially the following: A high-impedance circuit samples, rectifies and filters the AC signal riding atop the DC “plate” voltage. The output signal of the rectifier circuit is proportional to signal amplitude at the plates. It is compared to a set of steady DC references and if these are exceeded the comparator circuit triggers indicator LEDs accordingly.

Apart from the clever biasing scheme the Ampulator’s “power amp emulator” contains no exceptionally novel ideas concerning hybrid power amplifier circuitry. It also requires a high B+ voltage and an output transformer, which is a big minus to

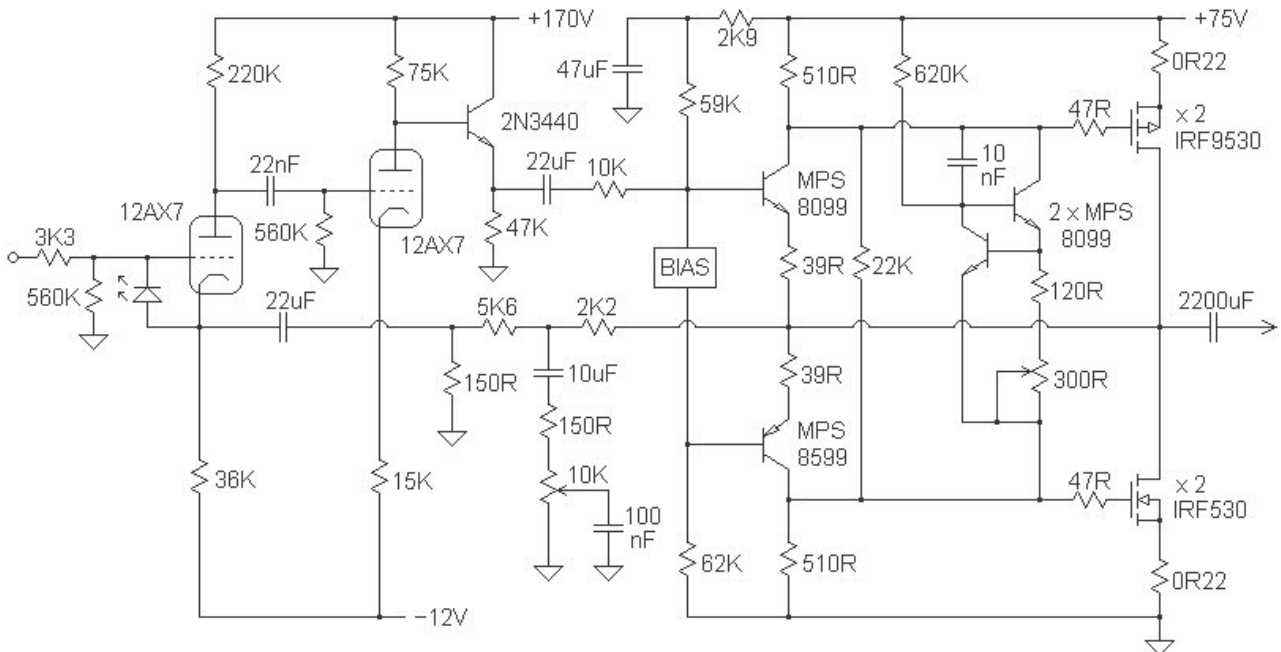


Figure 3.103 Simplified schematic showing the ADA hybrid power amplifier topology. (Microtube 200).

The feedback path is quite “old school” and very similar to that of tube amplifiers, as it is AC coupled and connected to the cathode of a single-ended input stage. What is more unusual is the fact that the employed feedback is *positive* and has a gain-increasing effect. The circuit still retains a sufficient stability. The presence control is, as usual, a shelving high frequency control that has a flat response when fully in counter-clockwise position. When in clockwise position it introduces a 9 dB boost with a corner frequency around 3 kHz.

When it comes to hybrid circuits, most of us are interested in knowing how well they exploit the tube characteristics. Unfortunately this is the field where the design fails to live up to its promises: For example, the output stage of the Microtube 200 clips at about 28 volts, yet the input stage has clean headroom up to approximately ± 48 volts, at which point the BJT emitter follower begins to clip. As is, the input stage is fully capable of pushing the output stage into that harsh and sudden clipping that is characteristic to solid-state amplifiers. This feature would remain even if the headroom of the MOSFET output stage would be increased. On a contrary to all marketing hype, this design provides no tube overdrive whatsoever.

ADA’s “Quad Tube” hybrid amplifiers follow the same circuit topology with few minor exceptions. The presence control is omitted from the feedback loop and the additional biasing circuitry and output protection are also realized a bit differently. These modifications make no significant difference topology-wise. The input stage is similar but powered from higher B+ voltage (220 V). Some component values are different so the circuit provides earlier breakup in the first gain stage. Following along the lines, the headroom of the output stage is raised even further by increasing the supply voltage to 87 V. Mostly the clipping distortion originating from the circuit is still that coming from the overdriven output stage. Anyway, for fairness sake it must be mentioned that it seems like ADA tried to extract as much as clean power and headroom from the topology as possible. The approach evidently caused the circuit to have a sudden clipping onset and a hard knee in the transfer curve. There is no reason why the topology could not be tweaked to provide only tube clipping but in that case the amount of headroom would naturally be decreased.

the zero volt references, which in turn modulates the constant current sources. Thus the output retains a sufficient DC balance.

The maximum output signal swing of the tube input stage at highest B+ voltage is about 100 volts (peak); this is more than enough to drive the MOSFET buffer, powered by ± 65 V rails, into clipping. At this level of overdrive it will clip very abruptly and “flat top” like an ordinary solid-state amplifier. At lower input signal levels the clipping is less harsh since the output stage is not wrapped inside a global feedback loop like in typical solid-state topologies. At the lowest B+ settings the input stage clips before the signal amplitude exceeds the level that is capable of driving the output stage to clipping. Consequently all overdrive is strictly tube overdrive originating from the input stage. Naturally this setting also decreases headroom and output power.

All in all, the circuit is pretty clever, relatively simple and most importantly, *very stable*. Since there practically are no feedback loop -induced phase shifts the only compensation required to fight against oscillation are the gate resistors of the MOSFETs. Undoubtedly the circuit could also be scaled down to use lower rail voltages in order to suit less gargantuan output power demands (the SVT hybrid is a bass amplifier). This would also mean that B+ would be a lot lower, less lethal and likely attainable with “conventional” mains transformers. Does the circuit perform like an ordinary tube power amplifier? Most certainly not but it is, nevertheless, a nice hybrid circuit that possesses many characteristics of a tube amplifier.

4 Heat sinking

4.1 Basics of heat generation in a semiconductor

Heat sinking is an extremely important issue in semiconductor amplifier design. Transistors are considered to have an infinite lifespan but this is assuming that their junction temperature never exceeds the rated limits. For silicon the maximum rating of junction temperature is between 150° C and 200° C, for germanium between 70° C and 100° C. Note that the value is indeed the *maximum*; exceed it and the device will surely be destroyed. If you run the semiconductors little below their maximum value they might take it – but only for a short time. Reducing the junction temperature by 10° C will approximately double the semiconductor's expected lifespan and halving the temperature will quadruple it. Also, the operation of the transistor is highly affected by the temperature: In typical silicon transistor the base-emitter voltage will decrease about 2.1 mV per an increase of one degree Celsius. Lower forward voltage would be ideal but unfortunately this means that the current gain will also increase – often in an uncontrollable manner. Unless this phenomenon is slowed down, compensated or prevented completely the transistor will automatically self-destruct. This incident is known as thermal runaway. Germanium transistors are especially notorious of this.

The transistors will always generate waste heat to which people commonly refer as “thermal power”. The power is, as usual, the function of current and voltage drop over the transistor. The base current can be left out of the equation since it's usually low enough not to generate a notable portion of waste heat. Thus the thermal power can be calculated from formula:

$$P_D = I_C \times V_{CE} \quad \text{(Equation 24)}$$

The formula also shows why heat will not be a major problem if transistor is used in a switching application: In saturation state (fully conducting) the collector-emitter voltage (V_{CE}) is very small and in cutoff state the collector current (I_C) is nearly zero.

However, in conventional, linear amplifier circuits transistors operate in the active region, which means they pass a lot of current and simultaneously control a substantial voltage; thus the thermal power will rise considerably high. A typical class-AB amplifier has the poorest efficiency when its output signal is zero. At this state of operation the output transistors still have to dissipate their idle power defined by quiescent current. Usually this is rather small, though. The efficiency increases together with output signal's amplitude – being greatest at the clipping region. The thermal power (amount of waste power) will increase as well but only up to a point where the amplitude of output signal *at the power transistor's emitter* is approximately half of the rail voltage per each half wave. At this state the power amplifier is under most stress. Due to correlation of current and voltage drop the thermal power will actually decrease after this point as shown in Figure 4.1.

When driving a combination of capacitive and inductive loads, such as a loudspeaker, the load impedance defines the phase difference of current and voltage, thus it is possible that the transistors will have to momentarily deal with fairly high thermal

power levels (refer to Figure 4.1). The *phase angle* is also the reason why it is a good practice to use transistors with exaggeratedly high power rating. The average power of the two waveforms (top and bottom) is still nearly equal.

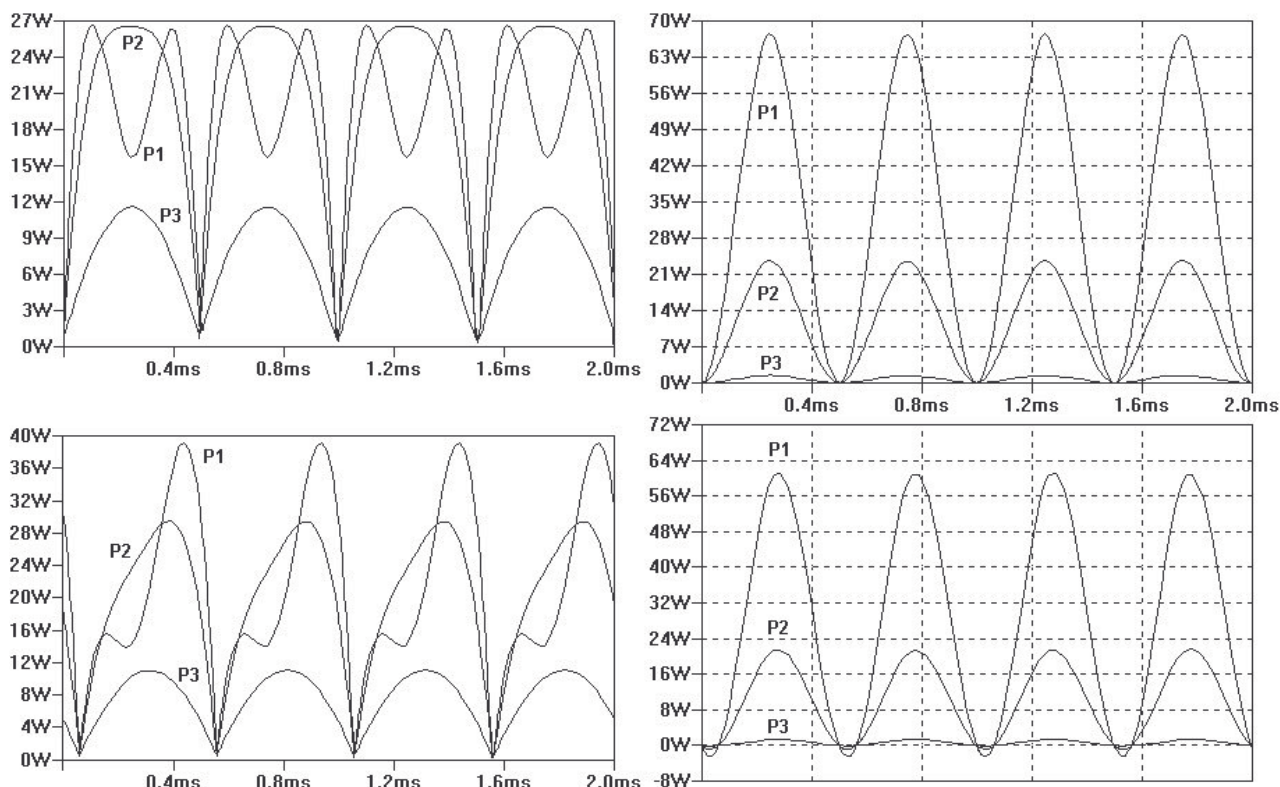


Figure 4.1 Left: Thermal power of output transistors in a class-AB amplifier driving a purely resistive 8-ohm load (top) and a speaker with nominal impedance of 8 ohms (bottom). Right: The output power of the aforementioned circuits in equal conditions.

Especially at low output signal levels the efficiency of the plotted amplifier may seem horrid (≈ 10 watts of thermal power for ≈ 100 mW output) but one has to remember that class-A amplifier's idling point is exactly the same where the efficiency of a class-B amplifier begins to increase. Thus a similar class-A amplifier would constantly dissipate the concerned 27 watts that represent the most stressful point for the class-B amplifier. The idling power of the plotted circuit was only few milliwatts and it was deliberately biased fairly hot. With colder biasing and larger amount of crossover distortion the efficiency at lower power levels could be doubled.

4.2 Heat sink math

The heat sink's *disability* to dissipate the heat applied to it is known as thermal resistance. It is stated in a form of T/P, for example 2K/W. This means that the heat sink will get 2 Kelvins hotter than the ambient temperature per every watt increase in wasted power. 60 watts of waste power would mean an increase of 120 Kelvins. Note that Kelvin and Celsius share the same scale so you don't have to make any conversion between units. Assuming that the suitable junction temperature for silicon transistor is 100°C , the proper thermal resistance can be derived from the formula:

$$100^\circ\text{C} = P_D \times R_{SA} + T_A \quad \text{(Equation 25)}$$

In the equation T_A = ambient temperature (usual assumption is 25° C), P_D = thermal power and R_{SA} = thermal resistance. However, it is more practical to consider the junction temperature as variable T_J . Thus the formula converts to:

$$R_{SA} = (T_J - T_A) / P_D \quad \text{(Equation 26)}$$

The thermal resistance of selected heat sink should always be *smaller* than the derived thermal resistance. A lower thermal resistance means a *better* heat sink. If the thermal resistance is negative (it can be) it means the sink should operate as a heat pump, also known as *cooler*. An example of where to find one is a refrigerator.

The basic formulas of calculating heat sink's maximum thermal resistance apply only if transistor's junction and heat sink have exactly the same temperature – which is virtually impossible. Usually the thermal resistance is much larger due to junction-to-case resistance $R_{\theta JC}$ and case-to-heat sink resistance, better known as interface thermal resistance $R_{\theta I}$. The first variable should be found from the datasheet and the second variable is explained later. You can simply add all thermal resistances together and then derive the junction temperature from the formula:

$$T_J = P_D \times (R_{\theta JC} + R_{\theta I} + R_{SA}) + T_A \quad \text{(Equation 27)}$$

The real thermal resistance requirement for heat sink is then derived from the formula:

$$R_{SA} = (T_J - T_A) / P_D - (R_{\theta JC} + R_{\theta I}) \quad \text{(Equation 28)}$$

To improve reliability and lifespan of the output devices it is advisable to assume that the ambient temperature is higher than the standard 25° C (onstage it very likely is) and use lower than maximum junction temperature value.

Also, transistor's datasheet will usually show a figure of power derating above 25° C. The figure is in the form of P/T - watts per increase of temperature. Increase in temperature will decrease the maximum power dissipation ratio and this variable usually drops heavily after 25° C. Figure 4.2 shows an example of power derating curve; it's an excerpt from the datasheet of MJE3055, which was a common choice for a NPN power transistor for decades. The maximum ratings for junction temperature and power dissipation of the transistor are 150° C and 75 W. The thermal resistance from junction to case is 1.66. The graph shows that the power dissipation rating of 75 W only applies when the case temperature is kept below 25° C. On a case temperature of 75° C the dissipation rating has dropped to only 45 W and the junction temperature is 75° C x 1.66 = 124.5° C.

Generally, reducing the heat sink's thermal resistance with various methods is cheaper than using a bigger heat sink. This is not a very good practice and just because it can be done doesn't mean it should be. Mounting a fan to the heat sink can reduce the thermal resistance by 1 to 2° C/W but to trust the fan as the only cooling solution is not reliable practice. If the fan develops a fault there must be an alternative way to prevent transistors from overheating. More *effective and reliable way to reduce the thermal resistance is to double the number of output devices*: The thermal resistance for each device remains the same but each dissipates only half of the power. This effectively halves the temperature rise.

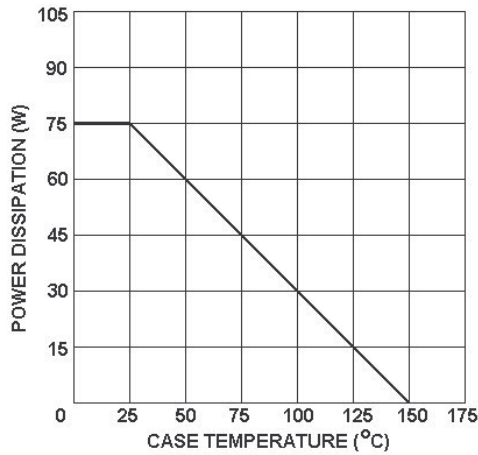


Figure 4.2 An example of power derating curve.

As an example, let's assume that we have a transistor that dissipates 50 watts of heat. We want to run it in an ambient temperature of 30° C and never let the case temperature grow higher than 105° C. The thermal resistance from case to heat sink ($R_{\theta I}$) is 1° C/W. Thus the heat sink needs to have a thermal resistance rating that is less than

$$(105^{\circ}\text{C} - 30^{\circ}\text{C})/50\text{W} - 1^{\circ}\text{C}/\text{W} = 0.5^{\circ}\text{C}/\text{W} .$$

If two devices are used in parallel each of them dissipates only 25 watts. One device would then require a heat sink with a thermal resistance less than

$$(105^{\circ}\text{C} - 30^{\circ}\text{C})/25\text{W} - 1^{\circ}\text{C}/\text{W} = 2^{\circ}\text{C}/\text{W} .$$

Effectively, the parallel pair will require a heat sink that has a lower thermal resistance than 2° C/W divided by two, which equals only 1° C/W or 1 K/W. We see that the rating is two times higher than for one device alone. Basically, doubling the amount of transistors usually reduces the amount of introduced heat even more effectively than adding a fan to the heat sink. However, inclusion of a third device would mean a slighter reduction: The maximum thermal resistance would be around 1,17 K/W. (Parallel connection of three thermal resistance requirements of 3.5 K/W; the 50 W is divided by three).

Note: The above rule is one of the main caveats of “chip amplifiers”: They cannot effectively parallel output devices to reduce the dissipation and they simply lack enough surface area that is in contact with the heat sink. Some older chip amplifiers, that came in nearly palm-sized cases were slightly better because they usually were “hybrid” designs and actually contained two to four internal, discrete power transistors that were spread equally along the rear metal surface of the case.

The above math can also be done easier since thermal resistances can be placed in series, parallel or in other various combinations. They are perfectly analogous to electrical resistance. Thus we can use a formula

$$R_{SA} = (T_J - T_A)/P_D - [(R_{\theta JC} + R_{\theta I})/n] \quad \text{(Equation 29)}$$

where n = the number of devices in parallel, to calculate the required thermal resistance for heat sink with multiple devices – assuming devices are identical and use

a similar interface, of course. For the aforementioned example the math for three parallel devices would be

$$R_{SA} = (T_{CASE} - T_A) / P_D - (R_{\theta I} / n)$$

$$R_{SA} = (105^\circ C - 30^\circ C) / 50W - (1^\circ C / W) / 3 = 1.1667^\circ C / W$$

Note that in this case, there was a requirement for *case* temperature of 105° C, thus the $R_{\theta JC}$ has already been taken into consideration and the thermal resistance consists of only the resistance of interface. To perform the following math in relation to junction temperature use equation 29.

4.3 Interface properties

The thermal resistance of the interface between heat sink and the output device is one of the most important variables since it's basically the only parameter (besides heat sink selection) that the designer can improve and tweak with correct choices. The interface between the device's case and heat sink can be either metal-to-metal or some kind of an insulation. Both types are defined as either "dry" (direct contact) or "lubed" (thermal paste is used to smooth out surface roughness). In most cases the use of thermal compound *halves* the thermal resistance of the interface because air (which has very high thermal resistance) is excluded from the mating surfaces. However, the thermal paste has to be applied correctly since it is an insulator of heat as well. There is only need for a little amount of paste that evens out the microscopic surface roughness. Applying thick gobs of paste (like you unfortunately see done in some commercial equipment) basically does more harm than good.

The thermal resistance of the interface will also vary according to how the device is mounted to the heat sink. For example, some devices are fastened with screws and some are clamped with spring clips. Clamping is generally a better method since it spreads out the pressure more evenly. For the same reason *there is a specific torque momentum for the screws that hold the device*. The screws should not be fastened too loose – neither too tight. Spring washers should also be used in order to stop the screws from loosening due to thermal effects.

Table 4.1 Properties of thermal interface materials

Material	Thermal resistance (°C/W)
Aluminum oxide	≈ 0.4
Sil-Pads	≈ 1 to 1.5 (cannot be lubed)
Mica	≈ 0.3 to 1.6 lubed, ≈ 0.8 to 4.3 dry
Kapton	≈ 0.9 to 1.5
None	≈ 0.1 to 1.3 lubed, ≈ 0.15 to 2 dry

The aluminum oxide is the usual surface material of anodized heat sinks - it is also an insulator. However, the anodized surface is very fragile and scratches in it can easily cause the device's case to short to heat sink's potential. Sometimes a metal-to-metal connection of the device and heat sink (together with some added thermal greasing) is a good solution: If heat sinks are insulated (and protected from accidental touching) they can serve as very low resistance *supply rails*. In such cases it is essential to use thermal paste that does not have isolative properties.

4.4 Heat sinks and cooling solutions

Heat sinks come in many shapes and sizes. One common rule that applies to all of them is: The bigger, the better. One approximation of the size requirement is

$$V = 0.8 \times R_{SA}^{1.47}, \quad \text{(Equation 30)}$$

where V = volume in liters. For instance, the volume of a 1 K/W heat sink would be approximately 0.8 liters. However, this equation is useful only if the heat sink has a reasonable amount of fins filling its volume. For example, 0.8-liter solid block of metal is not as efficient as a finned design since *the surface area of heat sink will account to heat transfer ratio*. Large surface area can transfer more heat than small, thus finned heat sinks are more efficient than simple stamped designs. Sometimes the fin surface is “ruptured” and wavy, which further increases the surface area. The smaller the fin area is, the better the efficiency is as well, thus a heat sink with needle-like fins will be more efficient than a heat sink with plate-like fins - assuming the surface area and other properties of both sinks are equal.

If the heat sink’s mass is too small it cannot draw heat from the source effectively enough. In other words, the sink draws more heat than it dissipates. This is because the amount of introduced heat, Q (measured in Joules), is directly proportional to change in temperature ΔT and mass m , thus

$$Q = c \times m \times \Delta T \quad \text{(Equation 31)}$$

where c = specific heat capacitance, which is a property of a material and doesn’t depend on volume. Its unit is J / (kg x K). The values of c are typically very small with metals, when compared to - for example - water. (On aluminum 0.9 kJ / kg x K).

Basically there are three elements that define how efficiently the heat is transferred away from the sink: Convection, conduction and, to a lesser extent, radiation.

The convection can be further divided into two methods; free (natural) and forced. In free convection, warm fluid (heated air) rises up and is replaced by cooler air. The higher the velocity and the flow rate of air are, the more effective is the cooling. By increasing turbulence this process can be made even more effective since the *laminar* or “skin” effect (flow velocity is lower near a surface) is reduced. To increase the velocity the air can be channeled with fins or, even better, chimney-like tubular structures that narrow the flow area. (Remember that flow velocity increases when cross-sectional area decreases). For greatest efficiency the fins should always be positioned vertically so that the heated fluid has a free access to rise upwards.

For a given heat sink volume, there exists optimal fin spacing. When the trends of fin spacing and heat transfer coefficient are balanced, the heat sink has the optimal heat transfer properties: If fins are too close together the fin channels just fill up with warm air and heat transfer coefficient reduces. This phenomenon is also known as “mixing of the boundary layers”. Naturally spacing fins closer together increases surface area per given volume and compensates the reduction of the heat transfer coefficient. Note that boundary layers begin to mix significantly as well if the heat sink is too tall, thus a heat sink taller than 15 mm typically begins to lose its efficiency.

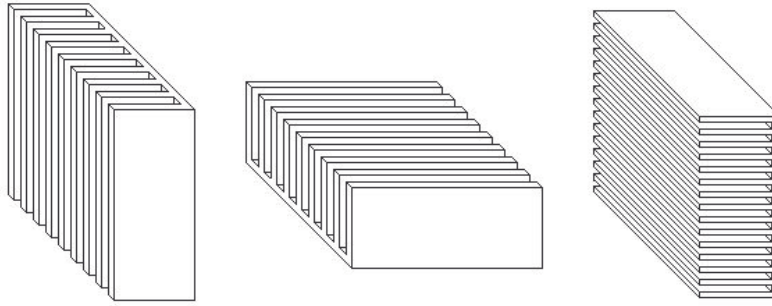


Figure 4.3 Heat sink configurations for natural cooling. From left to right: Vertical baseplate and fin channels – best configuration for natural cooling. Horizontal baseplate fins facing up – nearly as good as the vertical/vertical configuration. Vertical baseplate with horizontal fin channels is only slightly better than a simple metal plate.

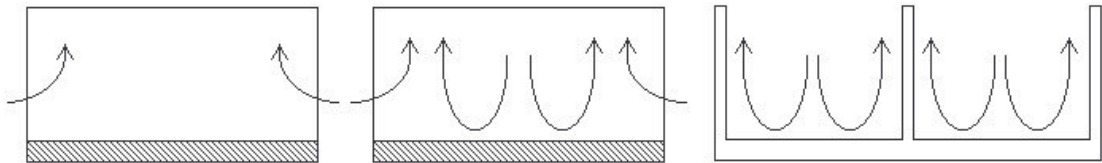


Figure 4.4 Flow patterns in horizontal/vertical heat sink. Left: Inflow from ends, middle: longitudinal thermosiphon and right: Thermosiphon between fins.

What may be a bit surprising is that the horizontal/vertical heat sink configuration is nearly as efficient as the conventional vertical/vertical configuration. One reason narrowing the gap in their performance is that short vertical surfaces have higher heat transfer coefficients than the taller ones. If fin spacing is large a horizontal/vertical configuration actually becomes a more efficient solution than the vertical/vertical. Basic flow pattern to horizontal heat sinks is from the ends to the middle but the effects of thermosiphon may become considerable as well, especially if gaps between fins are large and fins are short. Any significant thermosiphon effect does not appear in vertical/vertical configuration.

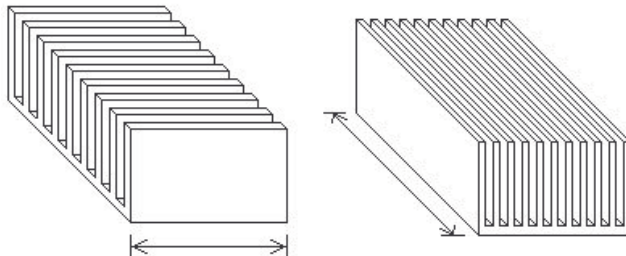


Figure 4.5 Fin channel orientation in horizontal/vertical heat sinks having a same volume. Left: Short channel depth (efficient) and right: Long channel depth (inefficient).

Forced convection imposes the effect of natural convection: If fans are used, they should usually *blow* air through the fins – this is more efficient than sucking air with a fan that is directly facing the fins. Usually the turbulent airflow created by the fan alone is sufficient enough to have the cooling limited only by the conduction along the sink’s baseplate. In such cases, a thick-finned sink that is designed for use with a fan may prove to be more efficient than using a sink with thin fins designed for natural convection. Note that while turbulence is beneficial for cooling it should be avoided at any cost when the exhaust air is removed. Circulating air in the “output compartment” will reduce the effectiveness of heat removing. In most cases it is worthwhile to channel the flow of warm exhaust air to a duct that allows a straight and non-turbulent exit route.

A very common forced cooling arrangement known as “tunnel cooling” involves using a “heat tunnel”, which is a tunnel formed by two or more heat sinks into where a fan blows air. In some cases another fan is used in the other end of the tunnel to suck air away. “Center-driven” heat tunnels have a fan in the middle that tries to perform both jobs (blowing and sucking).

The heat tunnel is an effective solution but suffers from a problem of equalizing the heat sink temperature along the tunnel: The sink temperature at the end of the tunnel (from where the warm air exits) is higher than at the beginning of the tunnel from where cool air is blown. This causes the devices at the furthest end to be warmer, which is not an ideal situation since equalizing the temperature of the output devices is rather important for reliability. Also, any thermal protection must be rated according to warmest devices. Poor thermal coupling between devices easily leads to a destructive phenomenon known as *current hogging*: The warmest device has the highest current gain so it draws major part of the current that was originally distributed evenly across all devices. This both substantially increases power loss over the device and introduces a very high risk for the device to drift into thermal runaway. A usual solution to the problem is increasing the fin length towards the end of the tunnel. In the year 2003 Peavey patented another solution, which is a very smart tunnel arrangement called “Turbo-V”: The heat sinks are angled to form a V-shape that makes the tunnel narrower towards the end. As the walls get closer, the speed of airflow increases due to decrease in cross-sectional area. Also, flowing air collides more with the walls making the walls extract more heat as well. At the end of the tunnel the heat extraction is at its maximum and the temperature along the heat sinks is equalized.

Another inherent problem of all heat tunnel constructions is the vast amount of dust they tend to collect. Fitting in a dust filter is more than recommendable. In practice, the intake port should also be larger than the exhaust port in order to keep the case pressure larger than the pressure outside – otherwise the system acts like a vacuum cleaner. The positive case pressure also compensates losses caused by the fan grilles and dust filters. And speaking of them, forget using those stupid stamped structures like holes, slots and spirals as a fan grille, they may block as much as 50 to 70 percent of the “opening” and usually just create turbulence and noise.

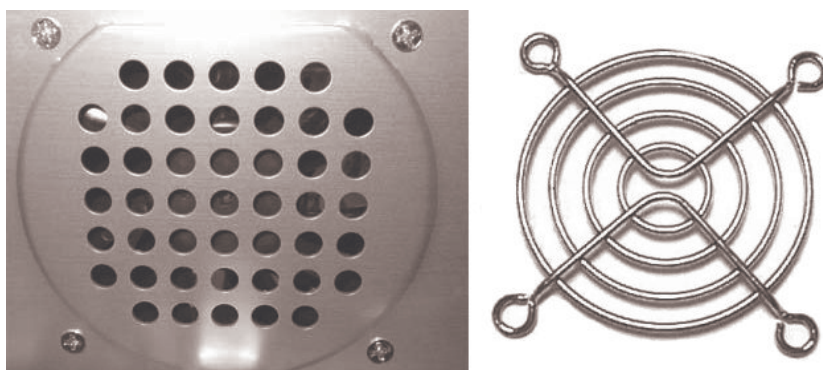


Figure 4.6 Compared to standard “fan guard” grille, the stamped grilles (left) are usually very poor.

Heat conduction means transmission of heat across matter and is always directed from higher temperature to lower temperature. Conduction spreads the introduced heat on a larger area, which is essential to improve cooling done by convection and radiation. Process is more effective in objects that have a large volume since thermal current \dot{Q} is directly proportional to area of the conductor A and temperature gradient $\Delta T / \Delta x$.

Temperature gradient is a constant, which means that temperature will change in direct proportion to distance. Thus we come to Fourier's law

$$\theta = \lambda \times A \times (\Delta T / \Delta x) \quad \text{(Equation 32)}$$

Where material constant λ is thermal conductivity measured in W / (m x K). Metals have a high thermal conductivity: With pure silver its 418.7, with pure copper 372.1 and with pure aluminum 209.3. Duralumin (the usual "aluminum" heat sink material) has a thermal conductivity of 129.1, brass (depending on alloy) around 100 and steel (depending on alloy) approximately 50. For the sole reason of increasing conductivity high-quality designs often use copper "heat spreaders" that distribute the heat around a large duralumin heat sink. The devices can also be "sandwiched" between two spreaders using the top one as a clamp that holds the devices in place. When constructed correctly, this is a very efficient solution and a novel way to get rid of all individual (and often troublesome) mounting screws.

A very typical construction is to mount the output devices on a separate, fairly thick metal bracket that then mounts to the heat sink. It has been observed that in the typical configuration the heat spreads quite evenly along the bracket; which is essentially a side effect we do not require. Equipping the bracket with a "thermal break" slots between the devices forces heat to spread into the right direction, which is of course towards the main heat sink. Compared to conventional configuration, utilizing thermal breaks may improve the cooling by lowering the case temperature as much as 20 %. Very often the mounting bracket has an "L"-profile. This is not an optimal solution since bends (especially sharp ones) deform the metal's crystal structure, which consequently increases thermal resistance.

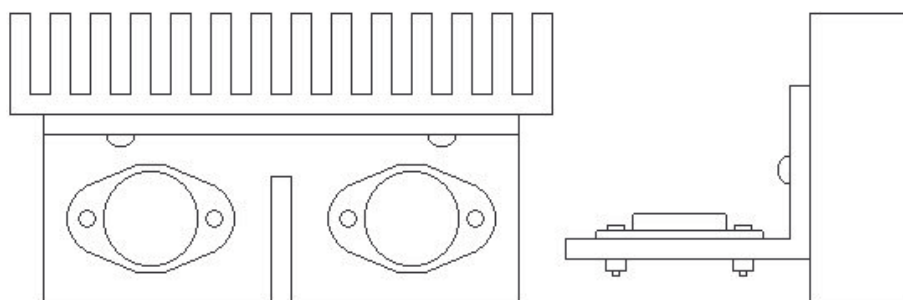


Figure 4.7 Thermal break in mounting bracket.

The type of metal and its thermal conductivity do not make a big difference when the conduction has to pass an interface of air, which very effectively reduces heat transfer; 1 mm of air will conduct heat as poorly as a 20 meter long plate of copper. Most of the heat dissipation will therefore happen by convection. However, *the greatest variable affecting the heat transfer is the temperature difference between the heat sink and ambient air*; this variable can actually even decrease the sink's thermal resistance. Logically, heat transfer is greater to cooler air.

Some properties of different heat sink dimensions have been presented in the following list. It will provide a rough guide about the effects that various dimensions cause. The third column from the left will explain the effects of increasing the related dimension and the fourth column will explain the effects of decreasing the related dimension. What one can conclude from the list is that increasing the volume of the baseplate is beneficial but almost every parameter considering fin dimensions is a compromise between other parameters.

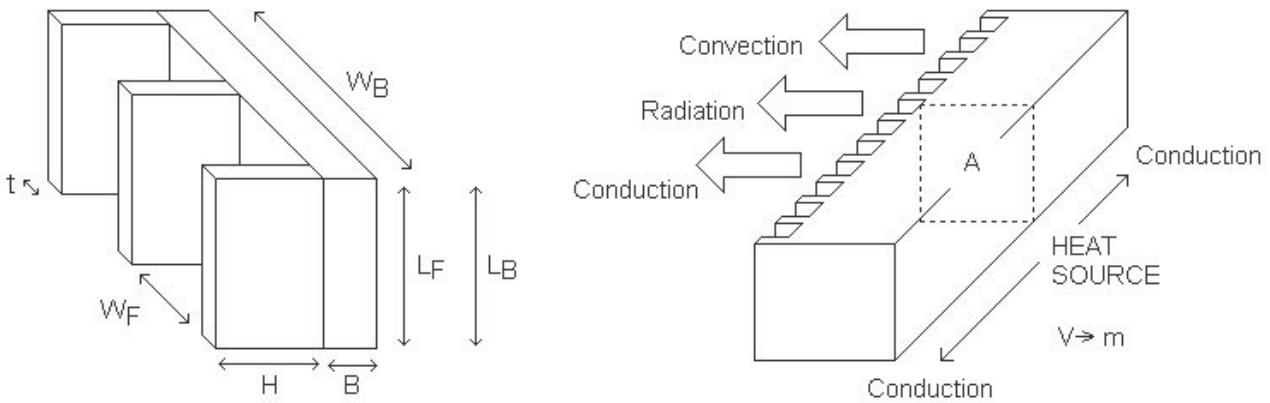


Figure 4.8 An extruded heat sink.

B	Baseplate thickness	Mass increases; sink can suck more heat from the source. Conductivity increases. Thermal inertia increases.	Conductivity decreases. Thermal inertia decreases. Can suck less heat from the source.
L _B	Baseplate height	Doubling the height = increase of heat dissipation capability approximately by 1.4x.	Halving the height = decrease of heat dissipation capability approximately by 0.741x.
W _B	Baseplate width	Doubling the width = Increase of heat dissipation capability approximately by 2x.	Halving the width = Decrease of heat dissipation capability approximately by 0.5x.
L _F	Fin height	Can channel airflow to achieve greater flow velocity. Loses efficiency by radiating more heat to adjacent fins. Increase in overall surface area.	Ability to channel the airflow weakens. Radiates less heat to adjacent fins. Decrease in overall surface area.
H	Fin width	Increase in overall surface area. Loses efficiency by radiating more heat to adjacent fins.	Decrease in overall surface area. Radiates less heat to adjacent fins.
W _F	Fin spacing / density	Up to a point an increase in heat dissipation capability. Restricts airflow. Fins radiate more heat to adjacent fins.	Decrease in heat dissipation capability. Allows a greater airflow between fins.
t	Fin thickness	Mass increases; sink can suck more heat from the source. Thermal inertia increases. Heat dissipation from fin decreases.	Increase of heat dissipation capability. Decrease of thermal inertia and conductivity.
n	Number of fins	Increases surface area.	Decreases surface area.

When the amount of convective heat transfer is normal *radiation* can count as much as 25 % of the sink's efficiency in heat dissipation. One simple rule applies to the emissivity of surface: The higher the reflection of the object, the lower the radiation.

Thus a mat black surface material will evaporate heat more effectively than a shining, metallic surface material. The sexy, polished copper heat sinks that you may see in some applications are ridiculously ineffective. However, *never paint a bright heat sink black!* In fact, never paint a heat sink at all. The interface of paint and the sink's metal surface - as well as the interface of paint and air - will create an enormous thermal resistance.

The last important parameter is the conductor's thermal inertia, which is actually just a modulus of conductivity, heat capacity and density of the material. Materials with high thermal inertia warm up and cool down slower. A higher thermal inertia allows storing high thermal peaks and dissipating them later when the need for thermal power dissipation is smaller. As a downside, the heat sink may for a small period of time seem more effective than it actually is. Aluminum has lower thermal inertia than copper.

A heat pipe is an extremely effective way to boost up heat transfer from the heat source; unfortunately it is expensive as well. The pipe is usually a sealed, hollow thermoconductive metal tube filled with a small quantity of coolant, which is usually water, mercury or ethanol. The remainder of the tube is filled with a vapor phase of the fluid. When heated, the fluid will vaporize and travel to lower temperature, which is usually a finned area - another heat sink - around the pipe. When cooled, the vapor condenses back to fluid and returns to heat source for another cycle.

5 Power supplies

5.1 Transformers

Due to modern safety requirements the transformer has become the heart of an amplifier. In the old days (1920's – 1940's) it was quite common to power up circuits without the expensive mains transformer. However, back then, even having electricity wasn't as self-evident as it is today: Most households were not connected to electric distribution network - neither was there a “modern” distribution network. Some households had an AC distribution, some DC. The ways to handle electricity were totally different than what they are today and probably that's why the safety aspect of galvanic coupling via mains transformer was ignored as well. It has to be noted that a circuit, which works without a mains transformer, is potentially very lethal because it has a direct connection to the mains distribution network. The modern safety measure is to at least galvanically couple such a circuit with a 1:1 transformer, also known as unity or isolation transformer.

In theory, every transformer can also be flipped around in which case the primary becomes the secondary and the secondary becomes the primary. This is an important point because some mains transformers may also qualify as “acceptable” output or interstage transformers - and vice versa. In some vacuum tube circuits chaining up transformers (for example, two 230 VAC to 12 VAC) to provide both filament and B+ supply is also common. However, this “reversing” of transformers is generally not recommended because their construction is *optimized for specific voltages and currents*.

The most notable distinctions in construction of secondary and primary windings are due to compensating differences in current and voltage. Using different winding gauges and perhaps a better insulation does this. Now, if the transformer is reversed it might be that its inner construction becomes insufficient to handle the new, reversed operating conditions. The same thing applies when the transformer is used with other “input” voltages than what it was designed for. At best, the transformer will likely only experience a slight loss of efficiency; at worst there is a risk of severe damage. A good rule of thumb is to use transformers only in applications for which they were originally specified and designed. Experiments may work – they do so even more likely if they are carefully planned and researched beforehand – but a good practice in such cases is to always expect a less reliable performance.

The efficiency of the transformer is specified by a VA rating, which is the product of maximum current and voltage the transformer can safely provide, therefore *the choice of the mains transformer pretty much dominates the maximum obtainable output power of the amplifier*. When the power demand grows higher than the VA rating the transformer will less likely be able to fulfill it, instead it just overheats and introduces a supply voltage drop. To prevent meltdown of insulation some transformers come equipped with thermal fuses or switches that are non-resettable, thus overheating may permanently destroy the transformer. The winding's series resistance mostly determines the VA rating, which is also one of the key factors in determining how much the supply voltages will sag when loaded. Also, under maximum load conditions the power taken from the mains distribution is always higher than the VA rating due to other transformer losses.

Transformers with higher VA ratings generally have better load regulation (ability to maintain constant output current during loading) than the smaller ones. The amount of regulation is given in percentage of output variation and for a 500 VA transformer it is about 5 % while for a 5 VA transformer it is only about 30 %. A typical 250 VA transformer (which is somewhat ideal for 60 – 100 W amplifiers) has a rating of about 7 %. The figure may not sound much but considering that under loading the supply voltage may, for example, drop from 60 to 55 V (or even lower) the effect of insufficient regulation starts to seem rather important. Even a 5 V drop in supply voltage can count as much as approximately 36 watts of lost power if the load impedance is 4 ohms. Besides, sag induces earlier clipping, which is never an ideal thing in a solid-state amplifier. Regulation is also nearly always specified for fully resistive load and assuming the current has a sinusoidal form, thus the rating is quite useless in practice since realistic loading is never like this. A rule of thumb is to assume that regulation during a realistic loading condition is about two or three times higher than the quoted figure.

Many types of transformers exist but audio circuits usually utilize only few of them. Autotransformers (single winding tapped at some point), polyphase transformers (for three-phase power), spark plug transformers and transformers used in microwave ovens (MOT), televisions (flyback), x-rays, laser, radar and broadcasting/receiving technologies are out of this topic.

The traditional core material for a typical mains transformer is silicon steel, which develops more thermal losses when frequency rises. For frequencies above 1 kHz the core material is usually ferrite. This material is often used in flyback transformers. Switching power supply transformers typically operate between 50 - 1000 kHz and they may even have an air core. Military and aviation transformers are designed to operate at 400 Hz. *This is important because all transformers are not interchangeable! Using a wrong supply frequency might induce more losses than what is considered economical.* For example, aviation transformers do not work as mains transformers.

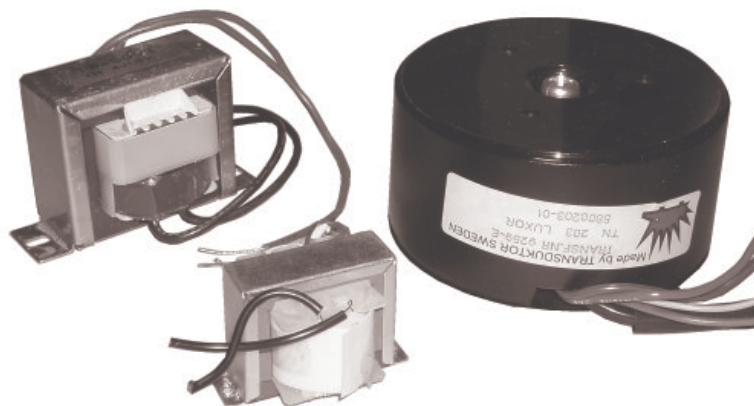


Figure 5.1 Mains transformers.

Typical mains transformer has either a laminated silicon steel EI-core or a toroidal core. The laminated structure decreases the flux of eddy currents that heat the core, thus the efficiency is improved. Toroidal transformers have a ring-shaped solid core. Sometimes the transformer may have bell-end covers to protect the windings, or it can be fitted into a housing. The latter is a good way to reduce electromagnetic interference but that only if the housing is material that is capable of doing that. High-

power transformers may be housed inside a box containing non-conductive transformer oil, which acts both as an additional insulation and coolant. Some transformer housings contain tar-like substance instead of oil.

Toroidal and EI-core transformers have some differences: Typical toroidal transformers are more efficient than cheap EI-cored types. They are also smaller (with equal power), have better regulation, less leakage inductance, mechanical hum and a lower exterior magnetic field. While some attributes might seem better than of EI-cored transformers, the toroids do suffer from *higher inrush current*. They are usually more expensive as well. Generally, EI-cored transformers have shallower frequency response (resulting into better mains noise filtering qualities), less problem with core saturation, lower capacitance between windings and better heat transfer characteristics. If more than one EI-core transformer is used the transformers should always be placed so that their laminations are at right angles to each other. If transformers poles face the same direction they will strengthen each other's magnetic fields or cause flux coupling. If many transformers are used place them in odd angles. Always place transformers and other elements that handle magnetic currents (like speakers, relays and inductors) as far away from each other as possible.

Transformers that convert voltage up are called step-up transformers and, vice versa, transformers that convert voltage down are called step-down transformers. Typical transistor amplifiers utilize only step-down mains transformers; tube amplifiers usually have a mains transformer that mixes both.

Aside winding configurations that offer multiple secondary voltages, transformers have three regular secondary winding configurations. They are known as single, center tapped and symmetrical windings. Single winding, as presented in Figure 5.2, is the least versatile one of the three since without any external circuitry it can only provide an unipolar voltage supply. The configuration is commonly known as “single supply”. Center tapping the secondary winding introduces a whole lot more of versatility for rectifying and supply polarity configurations. There will be more about this later. In short, windings on both sides of the center tap are symmetrical, which allows using a very simple rectifier circuit for single supply operation or (these days more commonly) allows the utilization of dual supply rails. The usual configuration is to use the center tap as the common reference point for zero-volt potential.

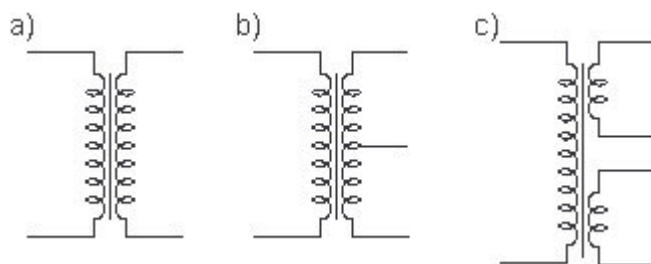


Figure 5.2 Typical transformer winding configurations: a) Single, b) center-tapped and c) symmetric dual.

Even more versatility can be gained by completely separating the symmetrical windings from each other. When secondaries are not connected together, a voltage potential proportional to turns ratio of each secondary, affects over the windings. Since the VA rating determines the obtainable current and voltage of *the whole transformer* it means that in the case of two symmetric secondaries the power available for each one must be half of the total VA rating. (For the sake of simplicity we assume that these are the only secondary windings in the concerned transformer).

Depending on how the current from the secondary windings gets rectified, it is possible to achieve a combination of either two positive or two negative supplies - or a combination of one negative and positive supply.

In the case of secondaries, one can follow electrical laws concerning the connections of power supplies. If supplies are connected in parallel the current feed capacity is that of all supplies together (current feed capacity is doubled in the case of two symmetrical secondaries) but the voltage potential is that of a single supply. Connecting the supplies in series has an opposite effect: The voltage potential is the sum of all supply potentials but the current feed capacity is lower (half of total in the case of two symmetrical secondaries). However, if the supplies are connected in series but *in opposite phases* the voltages are subtracted. The voltage and current available from each winding can always be derived from the VA rating, since it is a product of total obtainable current and voltage. *To avoid unnecessary loading, parallel configurations should only be used with symmetrical windings.* Series anti-phase wiring will not generally harm the transformer but parallel anti-phase will destroy it - unless the fuse blows fast enough, of course. Common referenced center tap is practically a series connection - after all, an equal voltage potential affects between the output nodes. Thus power for each supply side is half of the VA rating. Refer to Figure 5.3 for math.

The configurations work as well with primary side: A usual technique is to add a node that bypasses a little bit of the winding in order to select between 230V and 240V mains voltage. The primary may as well be symmetric, which allows parallel configuration and selection between 120V (primaries in parallel) and 240 V (primaries in series). This is basic circuit theory, yet one of the most frequently asked questions in various electronics forums.

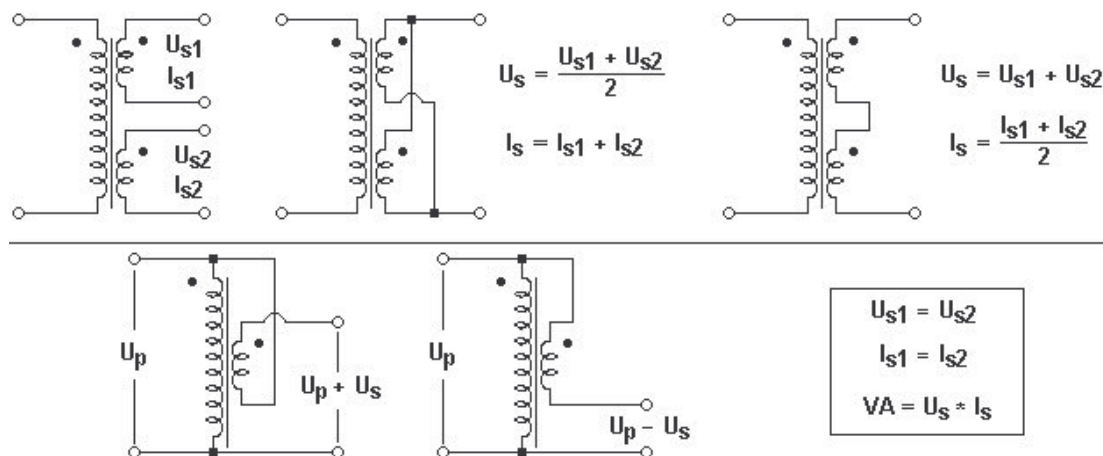


Figure 5.3 Configurations for a symmetric dual winding transformer and transformer configured as “booster” or “bucker”. Note the polarity indicator dots.

I don't recommend using the topologies presented in the bottom half of the Figure 5.3 in mains supply circuits since they fail to isolate the supply from the distribution network. They can prove worthy for some applications but coupling with an isolation transformer is generally advisable. The “booster” configuration connects the voltage source of the secondary in series with the voltage source of the primary, thus the voltage potential affecting between the output nodes is U_{PRIMARY} plus $U_{\text{SECONDARY}}$. In some circuits this might be an effective way to boost up the secondary voltage higher. Directly heated tube rectifiers utilize a very similar idea by the way. The “bucker” circuit affects in an opposite manner: It creates a series connection of the windings as

well but this time in reversed polarity, thus the voltage potential between the output nodes is effectively U_{PRIMARY} minus $U_{\text{SECONDARY}}$. As a side effect the common mode disturbances are reduced. As the name already implies, the circuit's principle idea has become much more exploited in "hum bucking" pickups.

Some equations become useful when dealing with transformers. They are:

$$N_p/N_s = I_s/I_p = U_p/U_s \quad \text{(Equation 33)}$$

$$N_p/N_s = \sqrt{R_p/R_s} = \sqrt{Z_p/Z_s} \quad \text{(Equation 34)}$$

$$N_p/N_s = \sqrt{C_s/C_p} = \sqrt{L_p/L_s} \quad \text{(Equation 35)}$$

N stands for number of turns in p = primary and s = secondary. As can be noted, all electrical variables have a distinct relation to turns ratio.

5.2 Rectifiers and supply polarity

A rectifier is a circuit or device that converts AC into DC by allowing current pass only into one direction. While common people nowadays might associate the term "rectifier" with Mesa Boogie amplifiers, the rectifying process is by no means anything unusual and has been known almost ever since the first vacuum tubes were invented (first vacuum tube was a diode). Rectifiers and Mesa Boogie amplifiers are related as much as engines and cars. Practically every household item utilizing DC has a rectifier circuit inside it.

The simplest way to control the direction of current is to use a diode. In fact, the very first rectifier circuits were vacuum tube diodes. I feel it's unnecessary to explain in technical detail how a diode does what it does - which is passing current only into one direction. More essential is the fact that diodes are imperfect devices and always introduce a certain amount of voltage drop. In vacuum tube devices this voltage drop is fairly high due to their high internal resistance. In semiconductor devices the voltage drop, caused by forward current requirement of the PN-junction, is usually very small. The voltage drop means a loss in efficiency and the diode has to dissipate the wasted power in the form of heat. Also, diodes can block only a certain amount of current; this limit is known as reverse breakdown voltage.

First semiconductor rectifiers were made out of copper oxide. They offered a good conduction but had a reverse voltage rating of only about 6 V. Selenium rectifiers, invented in 1933, quickly surmounted the copper oxide devices. Although selenium's forward current resistance was higher the reverse voltage rating was higher as well, this time about 20 to 45 volts. Like copper oxide the selenium was a "metal"-type semiconductor and did not require a tricky point contact technology that had to be used with "crystal"-type semiconductor materials, such as germanium and silicon. Basically the internal structure of Selenium rectifier consists of metal plates stacked over the semiconductor material. Selenium was not very reliable, though: Like all metallic rectifiers it is prone to aging and humidity; in time it develops a very high internal series resistance, which usually proves to be fatal for the device.

Overcoming the problems of point contact semiconductor manufacturing allowed the use of crystal materials, germanium and silicon. The crystal semiconductor diodes

possess a very high reverse voltage rating and good conductivity. First diodes that saw the mass markets were made out of germanium. In comparison to silicon, germanium has a lower forward voltage (around 0.3 V). In low-voltage supplies or in rectifying applications that require high precision this might rise to be an important factor. As a downside, germanium is very intolerant against heat and is destroyed in much lower temperatures than silicon. This variable essentially limits the maximum power obtainable through the diode. Today, modern Schottky diodes possess a low forward voltage almost equal to germanium.

Solid-state diodes always introduce some switching noise known as “hash”: The switching creates EMI in form of transient peaks that modulate the output signal by the same principles as ripple does. Radio receivers in a close vicinity of the device can even pick up the switching noise. Typical silicon diodes do not instantly turn off when the current through them drops to zero and reverses. Instead they conduct for a period of time called *reverse recovery time*. During reverse recovery the leakage inductance of the power transformer stores up energy that is finally released in a quick burst as the diode finally stops conducting. The flyback burst can have amplitude of several hundreds of volts and usually lasts a few microseconds. At the expense of slowing down the diode’s operation, the magnitude of spikes can be decreased by connecting a capacitor in parallel with each rectifier diode. Anything from 10 to 100 nanofarads seems to be a “standard” value for this purpose. This practice also lowers the ringing frequency thus decreasing EMI. Toroidal transformers also reduce diode hash because they have lower leakage inductance. Another, although not as effective, way to remove the diode hash is a *snubber* circuit, which consists of a resistor and a capacitor in series (“Zobel network”) shunting the supply to ground. The idea of the circuit is to create a very small resistance for high frequencies. The snubber can also dampen the LC *tank circuit* resonance, which is caused by parallel connection of capacitance and series inductance of large filter capacitors.

Note: The notation used on marking the supply polarity is originated from transistor technology: The abbreviation of positive rail node (V_{CC}) comes from the fact that it is the (usual) rail of the power supply that is connected to the collector terminal of a transistor. Similarly abbreviation of negative node (V_{EE}) refers to voltage (often) connected to emitter. Thus, in some occasions, V_{EE} rail is considered the same thing as the “common” rail. In some occasions you can see corresponding V_{DD} and V_{SS} abbreviations (FET circuits). Another common notation is V_{PP} (positive) and V_{NN} (negative). The first one is, however, easy to mix up with “peak-to-peak” so I don’t recommend using it. In vacuum tube circuits the supply abbreviations are based on battery markings of old radios: A = high current, low voltage filament battery, B = high voltage, low current plate supply battery (HT in Britain) and C = low voltage, low current bias battery. Usually these are also followed by “+”, which marks the use of positive battery terminal whereas the negative terminal typically is the common.

Common’s voltage potential (in reference to Earth) can be chosen quite erratically without affecting circuit’s performance. Usually the selection of common point is, nonetheless, based on logical reasons. Circuits a) and b) in Figure 5.4 show a quite unorthodox method of achieving a single supply (half wave rectifying). The diode will pass through only either positive or negative half wave (depending on the orientation) of the alternating secondary current. Since only the other half of the wave gets rectified it is clearly an efficiency issue. Note that without an “artificial” common point reference the other end of the winding cannot have an AC potential due to canceling effect of the load return current. Thus obtaining dual supply from a single

ended secondary winding always requires additional circuitry. This topic will be discussed later.

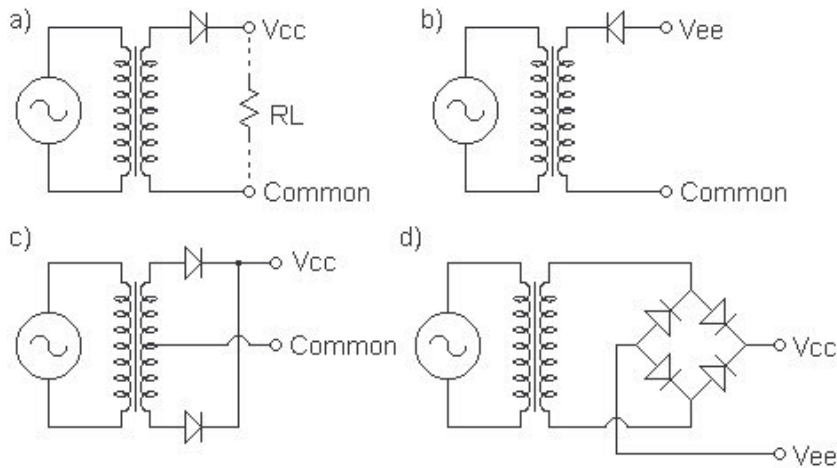


Figure 5.4 Rectifier configurations for achieving single supply polarity.

Circuits c) and d) are full wave rectifiers. The circuit c) relies on taking the common point reference from the center tap, which in turn causes ends of the windings to have a 180 degree phase shift in relation to each other. Two diodes can now be used to rectify both half waves of the alternating supply current. Depending on the orientation of the diodes the output's polarity can be either positive or negative. The bridge rectifier in circuit d) is configured to single supply configuration where it uses either one of its output rails as the common reference point. The secondary winding is not directly referenced to common but will always be in a very close potential through a forward biased diode. Notice how alternate pairs of diodes conduct at both negative and positive half waves thus forcing the current to flow only into one direction across the load that is connected between V_{CC} and V_{EE} rails. Depending on which rail is picked as the common reference the output polarity can be either positive or negative.

While both circuits - c) and d) - accomplish essentially the same thing they do have some differences: Unlike the bridge rectifier, the full wave rectifier in circuit c) cannot utilize the full winding during each half cycle of the AC wave. Consequently, to acquire the same voltage output than from a circuit with a bridge rectifier, each winding must have the same amount of turns that a single winding in a bridge rectifier-equipped circuit would have. This, in turn, means at least doubled winding resistance and therefore decrease of efficiency.

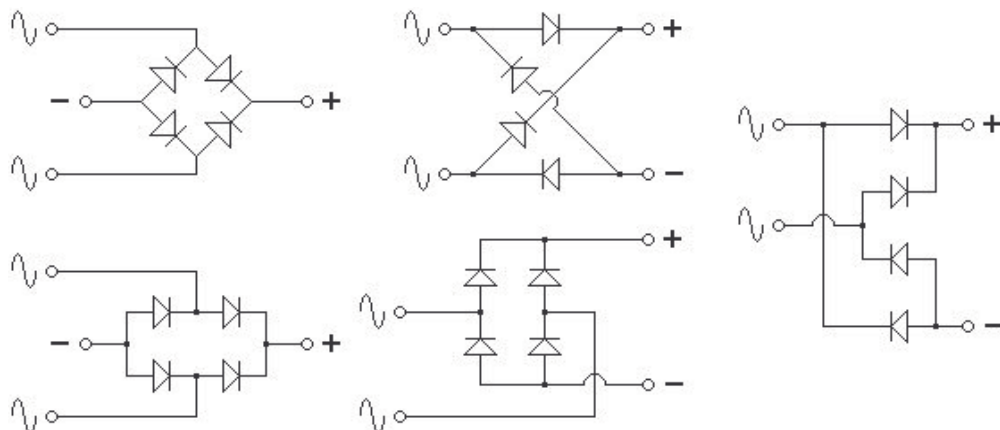


Figure 5.5 Alternative ways to draw a bridge rectifier.

As a side note, there are few alternative ways to draw a bridge rectifier. Some of these are shown in Figure 5.5. It is advisable to memorize the uncommon “forms” since it will speed up the process of examining schematics. Sometimes the bridge rectifier is depicted as a box that has two inputs, and two outputs and a symbol of a diode inside it.

Figure 5.6 depicts typical rectifier configurations for achieving dual rail supplies. Circuit a) is half wave rectifier and should require no further examination. The supply voltage for each rail is half of the total voltage potential across the whole secondary winding minus diode’s forward voltage drop. Circuit b) is a center tapped full wave rectifier. Since the common reference is taken from the center tap each output rail of the rectifier can be utilized and this allows dual supply polarity in reference to common. The configuration uses half of the total winding, which means that the voltage of each rail is only half of the total measured across the whole winding.

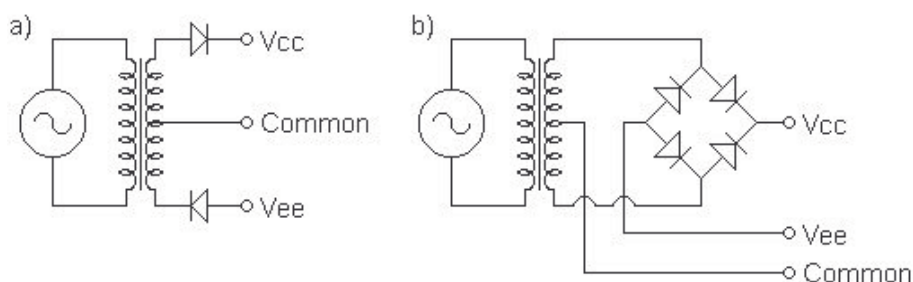


Figure 5.6 Rectifier configurations for achieving dual supply polarities.

Circuits depicted in Figure 5.7 are typically used in more powerful and refined systems. The circuit a) utilizes the dual winding and rectifiers to create two separate but single polarity supplies. The circuit can be configured to provide two positive, two negative or one positive and one negative supply – depending on which rails are tied to common. In stereo amplifiers using individual supplies will reduce the crosstalk between channels since the voltage fluctuations (sag) of the supply do not affect each other as much. Also, requirements for rectifier’s current handling and power dissipation are reduced when comparing to using only a single rectifier.

The circuit b) forms a dual supply from the parallel connection of two rectifiers. Think of the configuration as series connection of batteries: If the common reference is taken from the mid-point of them, one gets both positive and negative supplies. This configuration will again reduce current handling and power dissipation requirements of the rectifiers. Due to small (but significant) differences between rectifier bridges, the circuit *should not* be operated in a single supply configuration, which is equal to series connection of batteries *without* the mid-point being common. Circuit c) is simply a parallel connection of the rectifiers, which forms two dual secondary supplies. Both share the same common node, which is the center tap of the secondary. If the transformer has symmetrical dual windings the center tap is formed as usual.

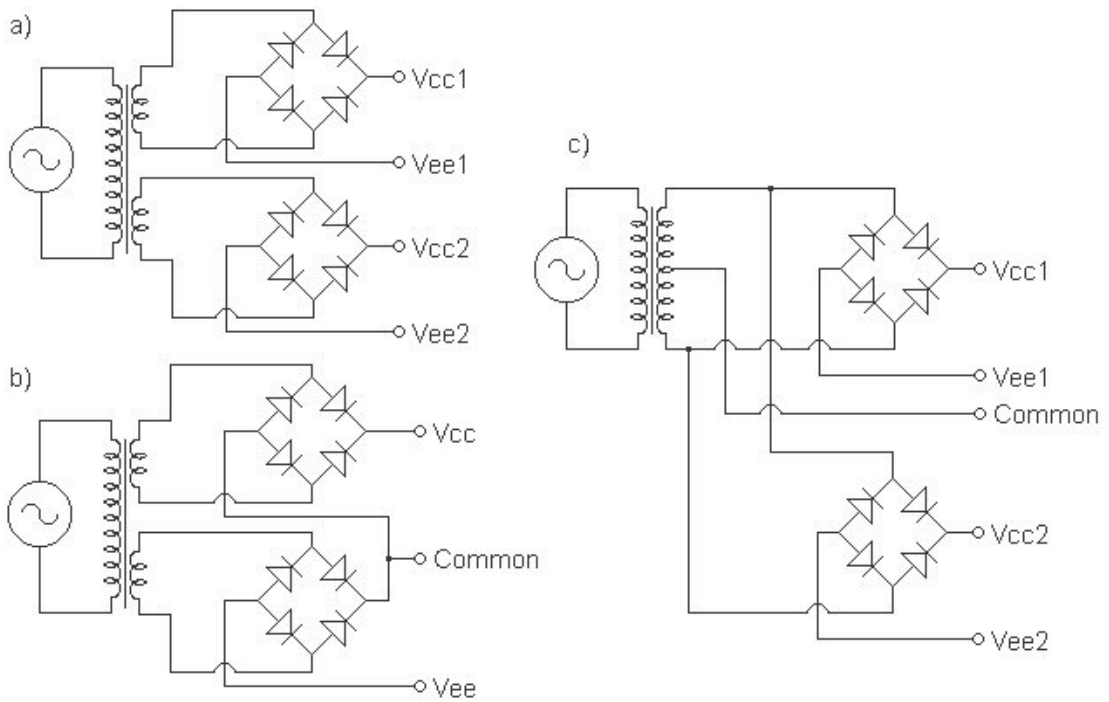


Figure 5.7 Rectifying methods utilizing symmetrical dual windings and dual rectifiers.

Figure 5.8 shows a less common variation of rectifier configuration: This is basically a very typical configuration – except that instead of connecting the center tap to the common reference, one end of the rectifier is connected to it instead. Effectively this forms a supply that is capable of providing voltage potentials, which correspond to full winding (secondary potentials summed) or half of it. Figure 5.9 shows another example of uncommon but rather clever wiring configuration. The transformers used in the circuit must be identical.

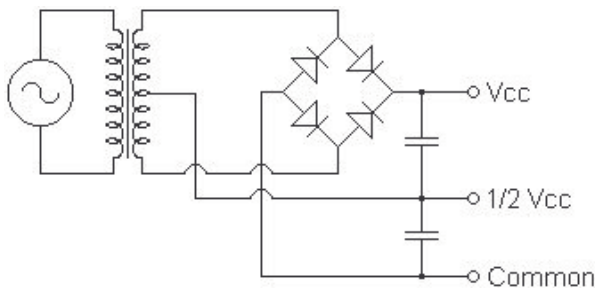


Figure 5.8 Single-ended supply with half-voltage tap.

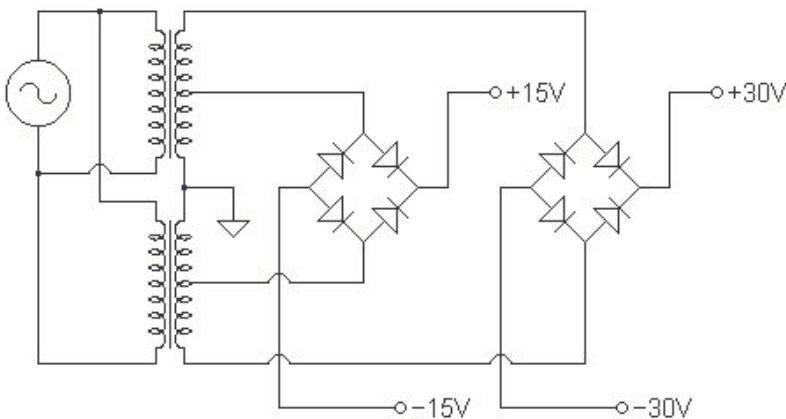


Figure 5.9 Another example of creative thinking.

The virtual earth (also known as phantom ground) circuit a) in Figure 5.10 is included here mainly because of its educative qualities: It takes a common reference from a potential that is supposedly half of the potential affecting over the whole winding. Trick is that the secondary does not need a center tap because a corresponding reference is formed “artificially” by using a resistor network. Thus in reference to common point the bridge rectifier output rails are both positive and negative.

The reason why this example circuit does not work very well in real-life is the circuit’s tendency to become unbalanced from unequal rail loading. Any offset voltage in the output will cause a DC current flow over the load. This current must originate from the rail-splitter; which in turn means that it must have a shift in ideal voltage potential. From DC point of view, the load sees the rail-splitter as two resistors in parallel and since the resistors are identical the math is easy. According to Ohm’s laws, the voltage shift must be $I_{\text{OFFSET}} * R_{\text{RAILSPLIT}}$, where I_{OFFSET} is the current flow through the load caused by the offset voltage in the output. This offset is usually significant with low-impedance loads (typical loud speakers) and will increase as the impedance decreases. Decreasing the value of the splitter resistors helps but in turn consumes more idle current. Too low values will also form a substantial voltage divider in together with the rectifier diodes.

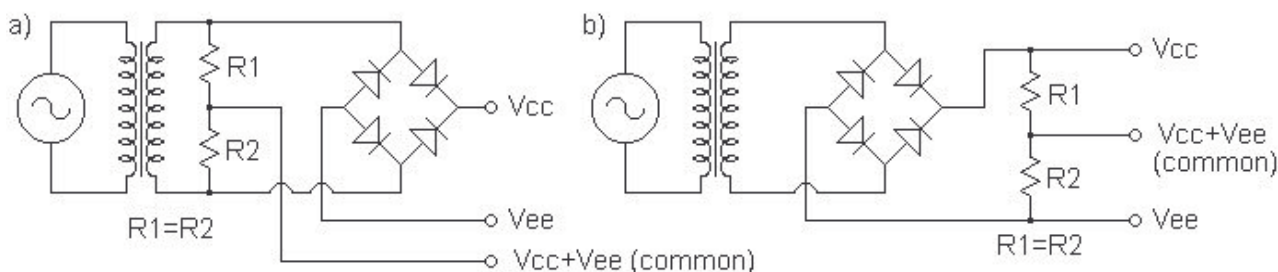


Figure 5.10 Dual supplies utilizing virtual earth.

For reference, an average DC offset of an amplifier with a long-tailed pair input stage is around 10 mV. If the nominal load is 8 ohms and the rail-splitter resistors are 10 kilo-ohms each, the common point offset with the particular DC offset in the output (10 mV) would be 12.5 V! For some circuits this means a serious limit in the ability to swing close to the supply rails.

Despite the obvious drawbacks the circuit topology is still fairly commonly exploited, especially in the form of circuit b). The circuit forms the reference potential between positive and negative rail after the rectifier. One can see this configuration used in low current applications where there is a requirement to power up operational amplifier stages with negative and positive supply polarities from a single ended supply – such as a battery or wall wart. Another form is the similar half-supply “bias” arrangement used in single supply OpAmp circuits: The common point is zero volts but the bias point is balanced between the supply and common rail.

Each approach benefits from having a buffer amplifier circuit that controls the current at the “rail split” node. The input of the buffer is the half voltage reference and the buffer’s output acts as the virtual earth point. With an ordinary OpAmp used as the buffering device the unequal currents can be compensated up to about 15 mA. Higher currents can be handled by either connecting more OpAmps in parallel or by building the buffer from discrete circuitry. Buffering circuits will usually even provide some output DC protection, which is a result of the circuit’s current delivering limit. One example of such a discrete circuit is shown in Figure 5.11. The particular circuit was

used in some Quad power amplifiers and when equipped with medium power transistors it is able to handle the current demand of at least 50 watts to 8 ohms power amplifiers. Circuits like this come very handy when the amplifier calls for a dual supply but you only happen to have a transformer with a single secondary. The drawback compared to OpAmp-based circuits is that no feedback is used to regulate the voltage at the “center” node. One can always use a similar buffer in the output of an OpAmp circuit though. Note that the *capacitors used in the circuit must have a voltage rating that is equal to total supply voltage - not half of it*, which at first would seem more logical. If there’s a requirement for an indicator LED or a pilot light they (and their current limiting resistors) should always be connected *between the rails* to achieve equal balancing.

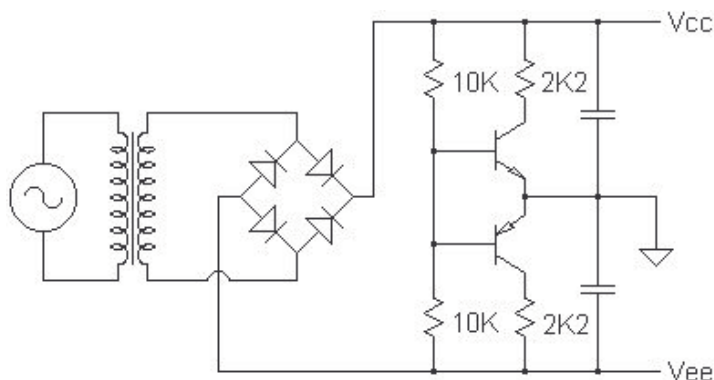


Figure 5.11 Buffered virtual earth circuit (Quad power amplifiers).

5.3 Supply filtering

The pulsating DC voltage output from the rectifier is not suitable for powering amplifier circuits and it has to be smoothed out in order to create a steady supply. The smoothing circuit usually consists of a capacitor in parallel with the load. The capacitor charges when supplied current exceeds the amount of drawn current - and discharges vice versa. In reality the filtered voltage is never completely stable and a small AC component, known as ripple, rides on top of the DC signal.

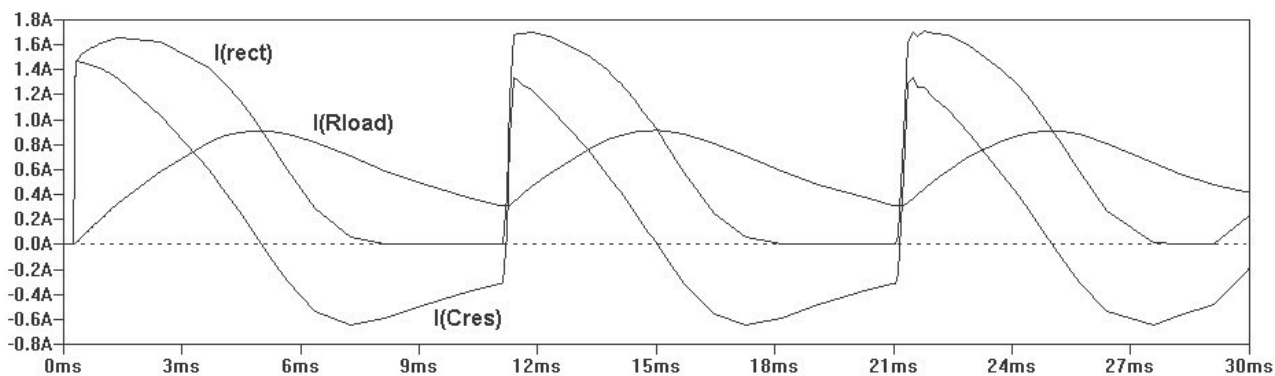


Figure 5.12 Currents in low impedance -loaded circuit.

Figure 5.12 makes clearer what goes on in a circuit formed by AC supply, full wave rectifier, reservoir capacitor and a load. During each half-wave, pulse current from the rectifier (I_{RECT}) is divided between the load (I_{RLOAD}) and the reservoir capacitor (I_{CRES}). Thus the supply current is always the sum of I_{RLOAD} and I_{CRES} . The load current is lowest at the zero crossing point of the sinusoidal supply waveform (and zero when powered up). This means that as soon as the voltage has risen enough to exceed the rectifier’s forward voltage the current pulse will quickly charge the

capacitor “full”. As the voltage of the sinusoidal supply signal raises the load current increases being highest at the peak of the sine wave. Up to this point the load current is completely fed by the supply. The residue of the supply current charges the capacitor.

After peak point the current of sinusoidal supply signal begins to decrease. At the point where the load current draw exceeds the supplied current the capacitor starts feeding current to the load as well – at the same time its current turns “negative”. After the current from the rectifier has dropped down to zero the load current is entirely supplied by the capacitor. As expected, the capacitor drains faster the higher the current draw of the load is.

If the load impedance is very high the amount of current drawn by the load is very low. This means that the capacitor doesn’t drain nearly at all. Since the supply current is the sum of load current and capacitor current the only charging current pulse from the rectifier will exist at the point where there is the highest current draw. This is the only event when the capacitor exhibits a little bit of discharging. During these conditions the charging pulse from the supply can be even ten times higher than the (tiny) current drawn by the load. This phenomenon is known as “current turbulence”. Since load current is formed from the division of supply and capacitor current the charging transients do not affect it.

One simple rule of thumb can be applied to determining the proper supply capacitance: The higher the power demand, the higher the requirement for reservoir capacitance. Usually the capacitance is best to rate as high as possible. Since hardly any power supply in a solid-state amplifier uses vacuum tube rectifiers a high filter capacitance will unlikely stress the rectifier too much – that is of course assuming that the rectifier is properly rated in the first place and that the rail capacitance is not insanely high. One could calculate the required capacitance but usually that is not necessary since there are rules of thumb that can be quickly assimilated: 1000 μF per rail is helplessly a too low capacitance by modern standards and 4700 or 6800 μF (per rail) have become common “standards” for a suitable value. Lower power amplifiers (i.e. less than 30 watts average) can cope with about 2200 μF (per rail). Capacitance exceeding 10 000 - 20 000 μF per rail is probably overkill and for such high power reserve the charging currents would remain very high for extensive period of time. Since the supply essentially sees a short circuit during the charging of filters this can indeed cause a notable stress for both transformer and rectifier.

Under loading the voltage potential of the supply drops. This is a phenomenon known as “sagging”. The reason for sagging is not an inadequate supply filtering but the internal resistance of the supply. The secondary winding resistance of a typical transformer can usually be anything from 0.5 to 5 ohms depending on the VA rating - enough to cause a notable voltage drop when the load current is high enough. Usually when the supply begins to “sag” the filter capacitors are charged with higher current than what the supply can provide. The filter capacitors thus try to compensate the voltage drop by discharging. This creates an increase in ripple since the typical discharge rate is increased. If the loading is constant, the filtered supply voltage gradually drops to the level of charging voltage as a function of capacitance and current draw. If there is a lot of capacitance the process is slower – however, it is *inevitable*. Fortunately, constant and severe loading of an amplifier is fairly uncommon and majority of loading takes place as fast transients. Logically, a large

supply capacitance is beneficial when the amplified signal has a lot of high-amplitude, low frequency content.

Although tube amplifiers (supposedly) benefit from “sag” solid-state amplifiers absolutely *do not*. Since sagging and ripple go hand to hand the preamplifier stages usually will amplify all supply anomalies. The ripple modulates the signal creating “ghost notes”, which are very often out of tune. (This also happens in tube amplifiers). The supply voltage drop may affect the operation of constant current sources and it definitely lowers the headroom by causing earlier clipping. Even in tube amplifiers a more ideal way to accomplish the benefits of sagging would be to use a proper limiter circuit in front of the power amplifier stage.

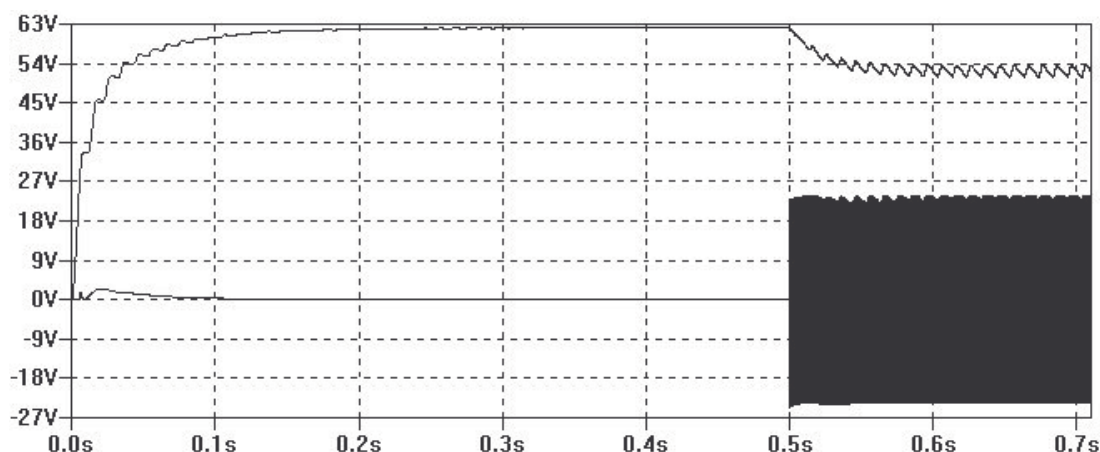


Figure 5.13 Amplifier's rail voltage behavior at startup and during loading.

Figure 5.13 shows a plot of typical amplifier performance during startup and considerable loading. The plot shows both the output and rail voltage of a single supply amplifier. Once the power is turned on (without input signal), the rail voltage rapidly climbs to its maximum level and stays fairly “clean”. The AC component from the climbing rail, however, is passed through the circuit and it causes a “glitch” at the output. This is the infamous “startup thump” and, depending on the circuit topology and component properties, its amplitude can be a great deal higher. If such glitch is passed on to the signal path at a very early stage of the preamplifier circuit it usually is amplified drastically. After 500 ms the input signal is applied. One can note that this immediately begins to load the supply causing considerable sag (about 10 V) and eventually an increase in ripple. The resolution of the picture is a bit poor but one can notice that the ripple also gets injected to the output signal, which is on the verge of clipping (and in fact clips each time the supply rail voltage hits the “bottom” of the ripple waveform). The effect of loading is most prominent on positive half wave of the signal – which is an inherent property of most amplifiers. Note also that at the portion where rail voltage is just falling the signal does not clip and exhibit ripple. The behavior during loading is the key why only continuous, average ratings can tell anything concrete about amplifier's output power.

Figure 5.14 Shows some typical filter configurations. The choke input filter in a) is more typically used in tube amplifiers since inductors are rather ineffective in high current, low impedance solid-state circuits. Therefore the simple capacitor input circuit (b) has become somewhat a standard setup. When chokes are used in the power supply circuit of a transistor amplifier they are usually exploited in the Pi-filter configuration (c): The circuit is essentially an AC bypassing filter followed by a low-pass filter formed by the inductor and another capacitor. Ideally, the AC resistance of

the inductor is very high while its DC resistance is very low, thus it is effectively possesses better filtering characteristics than a plain resistor used in same configuration. While in tube circuits that have low current demand the choke values are typically in order of few Henries, in high current solid-state circuits the required inductance is a lot less: Only few milli- or micro henries. Each filter has its qualities and selecting the “best one” is a task related to the desired purpose of the power amplifier circuit. The single shunt capacitor filter has the highest ripple as well as the highest output voltage. The choke input filter has medium ripple and medium output voltage. Pi-filter is a compromise between the two: Output voltage and ripple are medium.

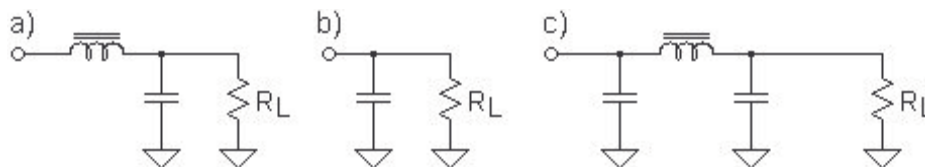


Figure 5.14 Power supply filtering circuits. a) Choke input, b) capacitor input and c) capacitor input “Pi-filter”.

Because solid-state amplifiers benefit from having a massive rail capacitance this creates a problem of slow startup and shutdown. In some amplifiers it takes several seconds before the supply voltage has risen to its maximum potential and all voltages in the circuit have settled. The rising voltage will be interpreted as AC by the circuitry and thus passed through to load – this will create a “startup thump”. Correspondingly a similar phenomenon takes place when power is shut down. The low draining of supply voltage will cause all kinds of anomalies to the output signal; few of them are “healthy” for the speakers. Prevention of this behavior is discussed elsewhere in this book.

The high rail capacitance is also a potential safety risk for servicemen: In modest and high-power amplifiers (aside practice and recording there is no sense to use anything else) the voltage potential between rails is quite high. In a 60-watt amplifier it is usually something around 60 volts and in a 150-watt amplifier around 100 volts. Even though it’s only DC such potentials begin to be hazardous. On top of that, the instantaneous current charges from the capacitive element are enormous. Even modest power amplifiers are capable of quickly discharging of about 5 amperes. Typical figures are closer to 10 amperes! The resulting arc from such current will have a serious heating effect and the shock is potentially lethal. To improve safety and speed up the shutdown process the reservoir capacitors should always have a parallel “bleeder” resistor that drains them when the power is turned off. A proper time constant for draining is few seconds. (Some designers have a rule of thumb where the time constant is equal to time it takes to open up the chassis). Ultimately the bleeder resistance should be switched on only when the amplifier is shut down; otherwise it will cause additional loading. If the resistance is “fixed” one has to make a compromise between suitable time constant and loading. A power indicator LED, together with its series resistance, is usually a pretty effective bleeder. In dual supply rail circuits such bleeder can be simply hooked up between the rails – there is no rule saying that capacitance must discharge to common potential only.

Ultimately there is another solution for bleeding, which is incorporated in a topology that is still fairly commonly used in some specific occasions. If the supply voltages are very high it might become hard to get a good deal of “cheap” capacitors with high enough capacitance. The physical size of high-voltage capacitors is also quite

massive. The problem can be solved with a simple way: Two (or more) lower voltage capacitors are connected in series. This effectively halves their capacitance but also doubles their voltage rating since each capacitor sees only half of the voltage affecting over the series connection. To maintain voltage over each capacitor in an exact potential two series resistors are connected in parallel with the capacitors and the “center nodes” of both are hooked together. The same resistors also effectively bleed the capacitors when power is disconnected.

One might think that addition of three more parts to replace one is bad design practice but this is not necessarily so: The cost of lower voltage capacitors might be a lot smaller even though they have two times more capacitance – usually it’s the voltage rating that adds up more size and cost. The cost of resistors is also very small. For example, let’s consider a rail voltage of 90 V: The difference in physical size and cost between one 100 V, 2200 μF capacitor and two 50 V, 4700 μF capacitors is quite remarkable. Naturally, any smart commercial manufacturer chooses the deal, which is cheapest in mass production.

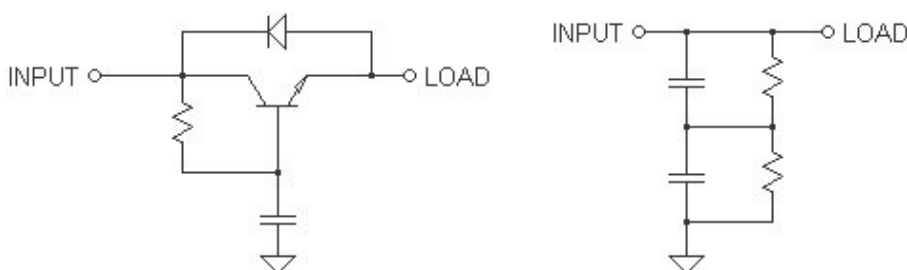


Figure 5.15 Filter circuit “oddities”: Capacitance multiplier (left) and series connection of capacitors (right).

More unusual filter circuit is the capacitance multiplier depicted in Figure 5.15. It consists of a transistor that is configured to multiply the capacitance at its base by the value of its current gain. As a result, the load will see a multiplied capacitance. The load current imposed on the capacitor (at the transistor’s base) is reduced by the factor of transistor’s current gain as well. The diode is used for preventing reverse biasing and the resistor and the capacitor together form a low-pass filter that smoothens out ripple. Larger values of R and C reduce the ripple by increasing time constant, unfortunately, too large values have a side effect of causing a lot of voltage drop over the circuit and also make the voltage rise very slowly towards the required value. The circuit benefits from any additional capacitance connected to the input in parallel with the load, especially if the circuit is used in class-AB amplifiers that have a massive variation in current draw. Though the circuit may look like a voltage regulator (see 5.4) it is not one: Although the circuit introduces some voltage drop like regulators, the output voltage varies according to input voltage – which is against the basic principle of voltage regulation.

5.4 Regulated supplies

A voltage regulator, by definition, is a device or circuit that maintains the terminal voltage of generator, or other voltage source, within required limits *despite variations in supply or source*. A typical voltage regulator is pretty helpless against voltage variations that drop the input voltage below the regulator’s control voltage but unlike capacitance multiplier it can keep the voltage at a steady, limited level during input voltage variations that exceed the control voltage.

Figure 5.16 shows three common methods of creating a voltage regulator. For simplicity, reservoir capacitors that are typically employed in the supply are omitted from the schematics; preferable there should be a high rail capacitance before the regulator and a lower one after it.) The leftmost approach, the *shunt regulator*, is very simple and based on a single Zener diode (two diodes are used in dual supplies). Logically, the output voltage of the regulator can never be higher than the zener voltage. However, input voltage should always exceed the zener voltage; this is actually the main operating principle for all regulator circuits: The regulator must always introduce some voltage drop in order to work. The higher the voltage drop is, the better the regulator circuit operates – and the higher the *losses* are over the regulator circuitry. The minimum amount of voltage drop the rectifier can work with is known as *dropout voltage*.

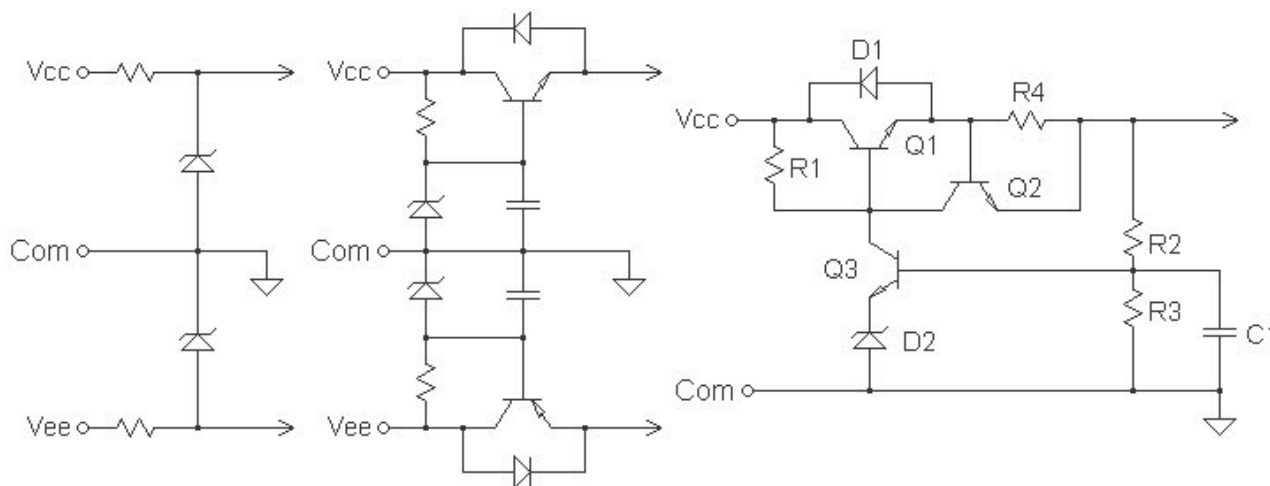


Figure 5.16 Discrete voltage regulator circuits.

In shunt regulator a resistor separates the regulator’s input and output voltage (including ripple) from each other and prevents the Zener from directly loading (and shunting) the supply; in normal operating conditions it also determines the amount of current supplied to the loading circuitry connected tot regulator’s output (the load including the Zener diode). Shunt regulator is always quite inefficient: In circuits, where low supply voltages are derived from the high voltage rails, there is commonly at least a 15 V drop over the series resistor. If such regulator would be configured to support a (realistic) current draw of 100 mA there would be a constant power loss of 1.5 W over the series resistor.

The implementation in the center is the most common approach in realizing a voltage regulated supply; it also mends some drawbacks of the shunt regulator. The Zener diode is placed at the base of a current amplifier transistor, or “pass device”, which effectively ties the output voltage to a potential of voltage drop over the Zener minus voltage drop over the pass device’s BE-junction. Additionally, the base circuit can have some additional capacitance to improve the steadiness of the voltage in case of sudden increase in load current. A resistor from rail to base is measured to feed enough of base current for the transistor, which in turn amplifies load current changes and keeps the voltage steady. As a downside, drop in input voltage will also decrease the base current, which in turn decreases the current gain as well. The circuit is partially temperature compensated: If temperature rises along with the zener voltage, the base-emitter voltage drops correspondingly. To increase current gain of the regulator, Darlington or Sziklai configurations can be utilized and to raise the power

handling capacity more transistors can be connected in parallel – in the latter case one should not forget the appropriate emitter resistors.

The greatest benefit of an active voltage regulator circuit is that when the load current draw is low the power losses are low as well. In the passive shunt regulator there is a constant current flow either to the load itself or to the Zener reference. Because the circuit also requires quite small series resistance to minimize sagging the power losses may become quite considerable. The active circuit practically “isolates” the Zener reference thus it doesn’t load the supply in a considerable way. The transistor then modulates the current draw according to the need, which helps to keep the losses at very minimal level.

The rightmost regulator utilizes negative feedback. The base current is once again fed through a resistor on the input side but the controlling device Q3 gets its base current from the output side. The voltage over R3 is kept at steady potential of zener voltage plus voltage over BE-junction. The current over the resistor defines how much current from the base of Q1 gets shunted off. As a result, the output voltage is regulated. The most significant trait of the circuit is that resistor R2 or R3 can be substituted with a potentiometer to make the output voltage adjustable. As an additional feature the schematic shows a current limiter arrangement: When the voltage drop over R4 exceeds V_{BE} of Q2 the transistor turns on and shunts off Q1. This feature will cause a further voltage drop, size of V_{BE} , in the output voltage. The circuitry shown in the schematic (obviously) includes only the positive rail side.

In both active circuits, substituting the NPN pass device for PNP at positive rail side, (and conversely PNP to NPN on negative rail side), allows creating a regulator that has a very low dropout voltage. This is beneficial since low voltage drop means also lower losses and decreased requirement for power dissipation. As a downside, the low-dropout regulator has the highest ground pin current requirement. This current is effectively wasted (it does not power the load) but still must be supplied by the source [25]. The minimum voltage drop for the standard (NPN for positive rail) regulator is

$$VD_{MIN} = V_{BE} + V_{CE} \quad \text{(Equation 36)}$$

With low-dropout regulator it is only:

$$VD_{MIN} = V_{CE} \quad \text{(Equation 37)}$$

All linear regulators require a certain amount of time to adjust the output voltage to correct level after a change in load current demand. This delay is a characteristic that defines the transient response of the regulator, which is a measure of how fast the regulator returns to steady-state condition after a load change [25]. The speed of the regulator should exceed the speed of the fastest gain stage. If this clause cannot be fulfilled the regulator is not fast enough to react to sudden current changes occurring in the gain stage and bandwidth will be limited.

Due to large fluctuations in load current, supplies of class-AB amplifiers should be regulated to acquire optimal performance. This is theoretically beneficial but the practical downside is that power loss and power-handling capacity of the regulator would have to be enormous – such is not good engineering. Fortunately, the effects of sag and ripple can be quite safely tolerated in power amplifier circuitry; also the transient response of an unregulated supply is faster. An adequately good practice is to

regulate only the low-voltage (and especially the low-current) circuitry, such as preamplifier and possibly the input stage of the power amplifier.

Note: Sometimes the switching devices in a switched mode power supply are called regulators or regulator switches. At the same the whole switched mode supply can be called as a regulated supply. However, the two circuit concepts – SMPS and linear regulator - work in an entirely different way and should not be mixed up with each other. Even though the output voltage or current of a typical switched mode supply stays constant under varying load and therefore can be considered as “regulated”, the supply should be called what it is; a switched mode power supply and the active devices that do the switching should be called switching devices.

5.5 Inrush current and how to limit it

When the power is switched on to a typical electrical device two things will happen: 1. The mains transformer will draw a very high current until its flux has stabilized and 2. The capacitors show up as “short circuits” until they have been charged. Both startup conditions have a tendency of raising the initial current draw (known as inrush current) notably higher than what even the maximum loading current is – this is also the reason why most of the faults occur when the device is turned on. The inrush current introduces very nasty side effects: It can affect the operation of other devices connected to the same AC line, weld switch contacts and more importantly, it stresses components and dictates the size of the mains side fuse since one cannot apply the “turns ratio is proportional to current” –rule. Mains fuses should be “slow-blow” type (US type “T” or euro-type 2 or B, C and D). It is highly recommendable to use an inrush current limiter with every transformer that has a higher VA rating than 300.

Limiting inrush current to less destructive levels can be done with multiple ways. Probably the easiest solution is to use a resistor in series with the primary. The resistance limits the initial current to “safe” level, which is usually around 200 % to 500 % of the maximum load current. A relay-controlled switch will short the resistor after a time delay equal to few mains cycles. The resistor has to withstand a lot of power: With a limit of 200 % a proper resistor value - and its dissipation - are:

$$R_{LIMIT} = U_{MAINS} / [(VA/U_{MAINS}) \times (200/100)] \quad \text{(Equation 38)}$$

$$P = U^2 / R_{LIMIT} \quad \text{(Equation 39)}$$

With a 250 VA amplifier and 230 VAC mains distribution this results to 105.8 ohms with a power rating of 500 W! The power dissipation requirement seems rather high but it has to be remembered that the resistor is in the circuit only for a few cycles – which equals around 100 to 200 ms – thus one can get away with values that are even a decade smaller. Resistance values that are too low do not limit current enough and therefore cause higher power dissipation. Too high resistance may either prevent the amplifier from reaching a stable point or interfere with delay circuits that are based on time constants of charging capacitors. Neither event is wanted. Naturally the relay operation must be as reliable as possible or else the current limiter will disintegrate with nasty smoke.

A NTC resistor can act as another type of inrush current limiter. Usually the negative temperature coefficient resistors that are indented for use as inrush current limiters

(also known as surge limiters) have a relatively high resistance at turn on (when the resistor is cold). The resistance typically ranges from 220 to 0.25 depending on the amount of protection desired. When current is applied the resistor heats up and its resistance drops *logarithmically* to very low value, which can be for example 0.01 ohms. This happens in few milliseconds so the NTC resistor does not have time to heat up excessively and burn.

Some amplifiers use an inrush current limiter that is realized with a triac, which is a two-way thyristor. A triac does not have a polarity but terminals marked A1 and A2 (or MT1 and MT2). During positive half wave A1 has a negative polarity in relation to A2 and during the negative half wave A2 has a negative polarity in relation to A1. The gate of the thyristors is shared. Essentially a triac requires a different type of control for each half wave. For positive half wave the gate must have a positive potential in relation to A1 and for negative positive in relation to A2. The triacs in inrush current limiters are usually “Logic Level” –type. It means that they can be triggered with a low gate voltage provided by outputs of typical logic circuits outputs. Usually the controlling circuits are fairly complex. An inrush current limiter realized with a triac always requires switching suppressor filters.

5.6 Surge limiting with MOVs

MOV, or Metal Oxide Varistor, is a voltage dependant resistor (VDR) that exhibits a low resistance in high-voltage, thus it can be used as a surge protector. MOVs are usually wired in parallel with the protected circuit. This configuration is known as *shunt protection*. When the voltage rises, the resistance of the VDR decreases causing voltage sag. This in turn dampens the surge. Voltage dependant resistors are manufactured for voltage ranges of 1 to 1000 V and for current ranges of 1 microampere to 100 amperes. The protection they provide is pretty effective against typical distribution network transients.

What is less well known is that MOVs wear out in use. The more surges are diverted, the shorter the life of a MOV gets and *when it finally fails it usually overheats to high temperatures*. Failure mainly occurs in three ways: Conductors inside MOV weld together (short circuit), conductors inside MOV open circuit (no protection), or conductors form a high resistance heater (most common failure mode). *These components have started many fires!* To make matters worse, MOVs usually give no indication of becoming too faulty. A certified surge protector needs an indicator that tells whether the MOV is in a working order or not. It also needs a circuit that prevents powering up if the MOV has become faulty. At the time of writing no device using MOVs is permitted to carry UL certification marking unless MOVs are coupled with thermal fusing.

5.7 Power supply overview

Figure 5.17 shows a schematic of the innards of a top quality linear power supply circuit – excluding possible use of regulators. The concerned schematic essentially assembles things taught by this chapter. It also shows some little details that prove to be quite essential when a power supply design is realized. These things are very logical when one sacrifices some time to think of them but usually seem less important during the design process. Most of these things also fall into the category of “beginner’s mistakes” that a knowledgeable designer knows to avoid already.

Therefore a less experienced one can learn a lot from the schematic – and especially from the description of it.

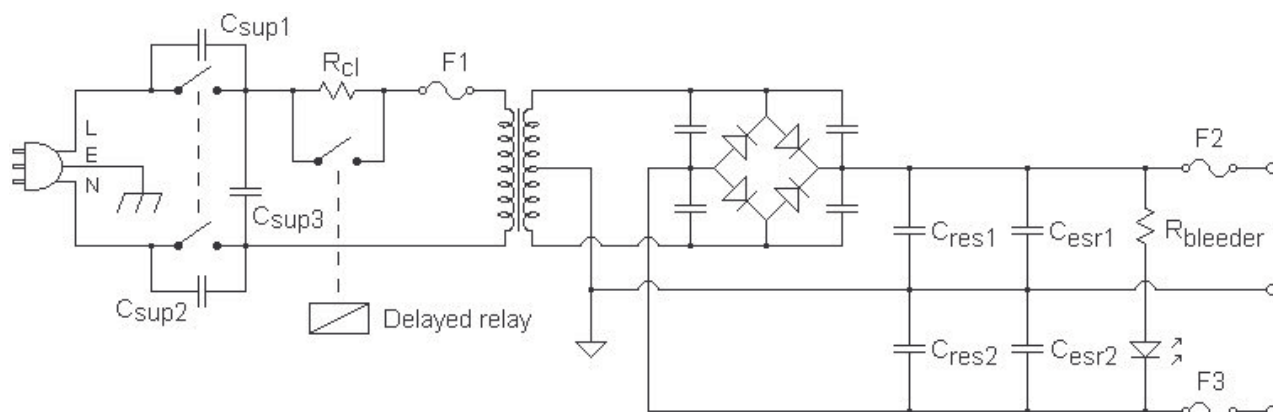


Figure 5.17 A top quality power supply.

The mains switch actuates *both L and N* (live and neutral) in order to assure that no mains voltage will exist in the circuit if the power is switched off. The problem in switching off only one of the mains inputs is to assure the switched one is L. In most electrical systems there is no way to make sure of this; even when using an IEC connector, that eliminates the possibility to plug the mains cord in with different orientation, the mains could have an alternate arrangement in the house wiring. If neutral (N) is switched off, the “neutral” potential at the device side of the switch will actually become same as mains potential. Logically, no current will flow until the point when a careless serviceman touches the “neutral” of the device side – or another potential earth. As a side note, service work should *never* be performed while the amplifier is plugged into the mains.

Even more important point is that the *switch is on the primary side*. I have seen countless devices that switch the power from the secondary side and while this works it will continuously waste power over the primary winding and all primary-side circuitry. So far I have not come up with any other benefits of this arrangement than the possibility to skimp on the switch’s ratings and save a few bucks. The absolutely worst way to realize this arrangement is to place the switch after the rectifier and filter capacitors and wonder why the circuit tends to die very often.

Capacitors marked C_{SUP} are surge suppressor filters that attenuate mains and switching transients. The suppressor capacitor hooked across the mains is located *after the mains switch*; this is done to eliminate the effect of leakage current over the capacitor. Common practice is to solder this capacitor to switch terminals. The suppressor capacitors have *very strict requirements* concerning voltage rating and power dissipation. Unless you know how to select them do not use them! The least requirements are that the capacitor is *rated for AC* (instead of DC) and at least 400 VAC. Ordinary DC capacitors will fail in this task sooner or later.

The resistor R_{CL} , which is switched off after a proper delay, limits the inrush current. If the transformer has a low VA rating and the filter capacitance is not insanely high this feature can be omitted. Other possibilities to limit the inrush current have been discussed earlier.

The mains fuse F1 is located *after* the mains switch to make sure it will not trip due to any switching transients. Usually fuses are slow and quite tolerant for momentary

current peaks but a surge can still trip them every now and then. Even if the inrush current tends to kill mains fuses periodically do not omit them! In such case you just have to select a higher fuse rating. Also, make sure the fuse is “slow-blow” type. Omitting the mains fuse is reasonable – yet highly questionable - only when it needs to be equally rated as the fuse in the breaker box. If so, there is either a too small fuse in the breaker box or the design is simply *terrible* and, say, needs an inrush current limiter. I don’t recommend tampering with fuse ratings of the breaker box since (if they are proper ones) they are based on house wiring’s capacity to carry current.

On the secondary side, the rectifier diodes have parallel capacitors to remove switching noise (“diode hash”). With modern, super fast diodes these have become quite optional though. The main reservoir capacitor banks (marked C_{RES}) have lower value parallel capacitors C_{ESR} that both dampen the tank circuit resonance of the supply and, more importantly, lower the equivalent series resistance of the capacitor banks. A bleeder resistor drains the capacitance and drops the supply voltage to safe level when power is switched off. In this circuit a current-limiting resistor and a power indicator LED form the “bleeder”. To get *balanced* loading they are hooked from rail to rail. Finally, the secondary side rail fuses F2 and F3 are located *after* rectifier and capacitance bank so that they will not trip due to high current “turbulence” (caused by rectifying and charging of the filter capacitors). Optional fuses can also be fitted to the supply lines before the rectifier. These must be able to handle the high current transients. Bleeder circuit is located *before* the rail fuses to ensure that filter capacitors are drained even in the case of a failure. The concerned fusing configuration is excellent in single-rail amplifiers; however, subsection 3.14.8 discusses the unfortunate side effects that trouble dual supply circuits. The same section also discusses fusing in detail.

5.8 Quick introductory to switched mode power supplies

A switched mode power supply (SMPS) is a circuit that uses active devices, switched on and off at a high frequency, to stabilize the output voltage or current. There are four types of switched mode power supplies: Buck converter, boost converter, buck-boost converter and flyback converter. Three first ones are unisolated from the mains, which is a bit risky from the safety point of view. The most common SMPS type – the flyback converter – offers isolation through transformer, although the switching circuit (primary side) still remains unisolated. A switch mode supply with isolation is often being referred as “Off-Line” circuit.

All SMPS circuits must first convert their input AC voltage to DC; this is essential for their operation since the input voltage that is switched must be constant. The rectifier stage is followed by an inverter stage, a “chopper”, that switches the rectified input voltage on and off at very high frequency. The frequency is kept well above 20 kHz to make switching inaudible; the fast devices of today allow using switching frequency as high as 800 kHz! At high frequencies the volume and copper losses of inductors and transformers decrease - so does the requirement for capacitance. This effectively decreases component size, weight and cost - although the capacitors have to be high grade, low ESR and low inductance type.

The switched pulses must have as steep corners as possible since any operation in active state creates an excessive amount of waste power and heats up the switching device. Also the diodes used in the rectifiers *must* be fast recovery –type – preferably Schottky – because all diodes remain in conducting state for a brief period when they

are suddenly reverse-biased. In a high speed switching mode supply a slow, standard diode would fail in seconds as the reverse current causes a huge power loss.

Usually the oscillator circuit controlling the switches utilizes feedback to compensate variation in input voltage. If the supply does not use feedback it is said to be *open-loop* type. In an Off-Line SMPS the feedback is typically galvanically coupled with an optoelectronic device in order to isolate the primary side from the mains circuit.

The chopper stage is either followed by or in conjunction with a voltage converter stage. The buck, boost and buck-boost circuits are based on the inductor's ability to release the current charged to it. The buck converter reduces the input voltage in direct relation of switch "on" time and total switching period. The output voltage of boost converter is always greater than the input voltage. Buck-boost converter has an inverted output but it can be greater, lower or equal to input voltage. The flyback converter uses a transformer that operates on a very high frequency.

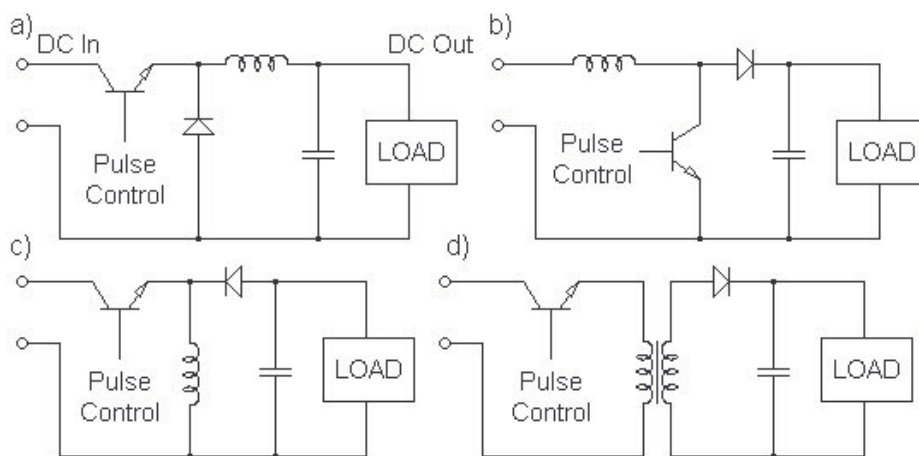


Figure 5.18 Different voltage converter topologies: a) Buck converter, b) boost converter, c) buck-boost converter and d) flyback converter.

The operation of switched mode supplies can be divided in two circuit configurations: On- and off-state. The operation of the buck converter is rather simple: It either connects the inductor to source to store energy or disconnects the source causing the inductor to discharge its energy into the load. The booster converter operates a bit differently: When the switch is closed the inductor is the only loading element for the supply thus the current through it grows very high. When the switch is opened the inductor transfers the stored high energy to the load. Buck-boost converter supplies current to the inductor while the switch is on. When the switch is off, the inductor supplies current to the load through the diode.

A flyback converter uses a high-frequency transformer for voltage conversion. It basically operates similarly as the buck-boost converter except that the inductor element is replaced with a transformer. The output of the chopper appears to the transformer as high frequency square wave signal. In on state the energy from the source is stored to the primary inductor and in off state the energy is released to the load from the secondary inductor. The secondary side signal is then rectified and filtered as usual. The configuration can be used in a push-pull switching configuration as well.

Figure 5.19 shows a practical realization of a flyback SMPS. The source signal is pre regulated with a triac to ensure the primary side reservoir capacitor will not be

exposed to transient surges in the mains. The active MOSFET switches are controlled by a pulse control that can be a logic circuit designed for such a task. The control board is powered from unisolated mains. A coupled inductor transfers the switched source signal to the secondary side where it is rectified and filtered from RF. A common mode choke is used as a filter in addition to high frequency filter capacitors. The current flows through the choke's windings but in opposite directions so the common mode signal is cancelled.

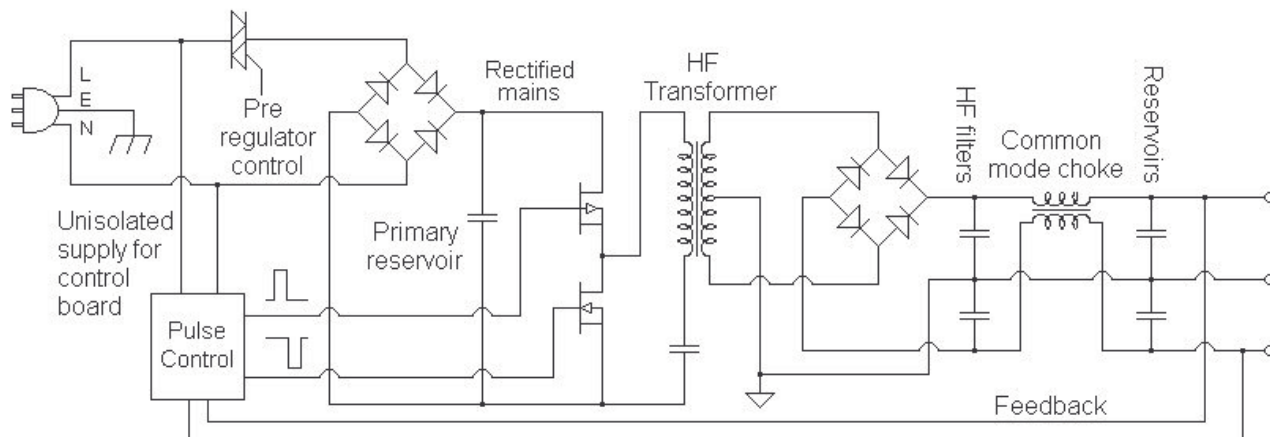


Figure 5.19 Practical realization of a flyback SMPS. (Half bridge type).

A resonant switching supply uses a “tank circuit” (an inductor in parallel with a capacitor) to replace the capacitor in series with the high-frequency transformer. This arrangement improves efficiency a little and more importantly the cleanliness of the supply since typical SMPS tend to induce a lot of switching noise to the supply. The resonant tank circuit effectively converts the switched square wave voltage into a quasi-sinusoidal form that is switched at the point of zero current – or zero voltage. On a contrary to pulse width modulated and push-pull type SMPS the resonant SMPS has a fixed “on” time but the period of the switching varies [8].

PWM Power (Pulse Width Modulation power) is not actually a switched mode power supply in the original meaning of the term but uses some elements of it. In the topology, the mains transformer’s duty cycle is controlled by a switching triac, which in turn is controlled by pulses that have a constant frequency but a variable width that changes according to loading. The width of the pulse naturally dictates how long the triac is energized. Under standby condition the pulse width is kept small but it increases in relation to power demand.

5.9 Safety precautions

When dealing with electrical circuits – especially with those that are in contact to mains circuit – following some safety precautions is essential and although this book has a main focus on *design* – not on actual building process – I feel an introductory to these precautions is truly on place:

Rule number one: Never work on a “live” circuit unless you have to. There are some occasions when dealing with live circuit becomes mandatory - for example, because of troubleshooting, finding a lead dress arrangement with least hum etc. Bear in mind that these are exceptional occasions and under normal circumstances there is practically no need to *work* on a live circuit. A device connected to mains is essentially live – even when the power is switched off. Since you likely have to probe

around anyway, wrap some electrical tape around your probes leaving only their tips exposed. This will prevent a lot of accidental shorts.

Rule number two: Work with only one hand and keep the other hand in your pocket. This way you avoid the most fatal form of electrocution where current flows from one hand to another – directly through your heart! If you don't have a probe's reference end that is equipped with an alligator clip then it's your first priority to get one.

Before you start to work on a circuit, appropriate measures have to be taken to discharge all reservoir capacitors that may contain a harmful voltage potential. A good rule of thumb of what to consider as a harmful potential is anything above *few volts* – for reasons explained soon. *Never drain capacitors by shorting them to ground with a screwdriver!* Some capacitors may partially regain their charge after a sudden (and short term) discharge. This phenomenon is known as dielectric absorption or “soakage” and it used to plague older (especially oil-filled paper) capacitors. Let it be mentioned that the effect of soakage is more prominent the faster the discharge time is and by short-circuiting capacitor terminals you practically beg for it. With modern components soakage is quite insignificant concern and the real problem in “screwdriver method” is that *the short circuit current may be high enough to do excessive damage for the capacitors*. Note that if you short the capacitor terminals the only current limiting factor is the very low equivalent series resistance of the capacitor! *Even more likely the short circuit current will damage delicate semiconductors in the circuit*. To summarize, even when your life is not on the line (like in the case of high-voltage circuits) you might still cause damage by following inappropriate procedures: Starting to work on a low-voltage circuit without bleeding it first may cause a destructive short. Use a proper “bleeder resistor” and *always* monitor with voltmeter to ensure that capacitors have indeed discharged.

Remember this: It's not the voltage that kills; it's the *current!* Human body has resistive properties that depend on point of contact and additional parameters like humidity (i.e. sweat), skin thickness etc. If the resistance of skin is left out of equation, the resistance between ears is about 100 ohms and from hand to foot about 500 ohms. Skin resistance can vary from 500 kilo-ohms (very dry skin) to a kilo-ohm (wet skin). If we consider that a current of about 10 mA produces a painful shock and currents above 100 mA kill one can *estimate* how high the voltage has to be in order to induce injuries or death. Usually at low voltages the high resistance of skin will prevent high currents from flowing but one should never count on this. Let it also be mentioned that victims of high-voltage shock usually respond to artificial respiration better than victims of low-voltage shock.

Do not work near potential grounds you may accidentally contact and *never* work barefooted or on a wet surface. Best practice is to wear rubber-soled shoes and stand on a rubber carpet. If you have long hair or a habit of wearing necklaces make sure they cannot touch the circuit while you work. Do not wear rings or other jewelry in your hands. Most low-power transistor amplifiers use low secondary voltages but even they can cause a serious injury if the current is high enough: If you short a low-voltage, high-current supply with high resistance of your skin, chances are that nothing happens - but a short circuit through a ring or bracelet can *melt* metal giving you nasty, third-degree burns. It will also increase possibility of a severe shock: If your skin *burns away* the total resistance between contacts will be extremely low!

Also, wear safety goggles (or eyeglasses): Capacitors may explode if they are subjected to excessive voltage or incorrect polarity, transistors may explode when they are subjected to high current flow, cut leads fly around with high velocities and soldering is all the time a little bit messy. Due to latter reason, wear appropriate clothes while soldering. Always keep a fire extinguisher near you. Even if you only work with 9-volt battery-powered circuits (i.e. effect pedals) it may become handy. For example, the fine metal wool often used in preparing PC boards can burst into flames when you subject it to short circuit currents by forgetting to disconnect the battery. Trust me – I know. No matter how skilled you are, accidents do happen; be prepared for them.

Know the equipment that you are working on! If you are unsure of what you are doing then do not fool around with circuitry that is capable of delivering high currents!

Some high-power public address amplifiers can use rail voltages that are as high as 100 V! Although their DC rail voltage is not considered fatal the AC voltage at speaker terminals can be deadly. Naturally, the voltage across the speaker terminals is even higher if the amplifier is bridged. Although a bit out of topic, let it be mentioned that you also cannot sense a shock from a high frequency source until you have already burned pretty badly.

People who have worked with computers know the procedure of wearing a “ground bracelet”, “grounding strap” or “anti-static strap” (whatever you want to call it) that helps to prevent the built up of static electricity in a human body. Modern amplifiers are full of circuits, which are practically computers. We don’t even have to talk about DSP or circuits alike: A basic CMOS circuitry, which is often used in switching circuits that have some “logic”, is an example of one of the technologies that are very prone to being damaged by static discharges. Since you really are not working with a computer the grounding strap should have a high-resistance path, preferably over 100 kilo-ohms. *Never* use metallic straps – at this point this should be self-explanatory.

There are guides considering unit construction that you should follow as well. Study various chassis safety classifications (for example, European IP standards) to find out the construction requirements. For example, the IP standard consists of two-number coding. The first number tells the amount of protection against foreign objects (including dust) and the second number the amount of protection against water and humidity. You are most likely planning to build a device with an IP-class 20 casing and earthed chassis. For such a device you can apply rules bulleted below:

- Parts in mains voltage potential should have at least a 6 mm aerial distance and 3 mm insulated distance to conductive parts that the user of the appliance can touch.
- All mains wiring should be either soldered or otherwise secured tightly. Cover all exposed wiring with heat-shrink tubing.
- The mains switch should not have a metal lever or any other metal parts that can be touched by the user. (Those are illegal in some countries.)
- Mains side should always be fused. (Mandatory in some countries).
- Unless the device is appropriately *double insulated*, earthing the frame/chassis is mandatory! The frame should include transformer’s chassis as well. The connection between earth and frame must be *solid!* For example, do not fuse it!

- All conductive parts that can be touched by the user have to be earthed! Note that in the case of most guitar amplifiers this includes the secondary side common as well since the guitar's common (strings, etc.) is hooked to it.
- DO NOT – under any circumstance - connect the Neutral and Earth together inside the device! There are exceptions to this but they are better left unmentioned in order not to stir up any confusion.

If you do not understand every single one of the bulleted safety precautions do not mess with power supply circuits; you are not experienced enough to deal with them! I have dedicated a whole chapter for discussing matters that concern grounding, earthing and the terminology it uses. Read it before designing power supply circuits.

I feel that it is appropriate to include some details about the bulleted matters: The mains switch should have a proper current and voltage rating. Do not use a switch with barely adequate ratings. In most cases the amplifier's inrush current is far higher than the maximum load current. Such current can weld switch contacts and even melt flimsy switches. The possibility of arcing is always present as well. If the switch has a parallel "transient suppressor" capacitor it should be rated properly for the task. The switch should preferably break (or establish) the connection of both L and N simultaneously. Switching only the neutral off will leave many parts of the primary circuit floating in mains potential. In the case of an open circuit the voltage potential at the "break" point is the same as that which feeds the circuit, thus "opening" the circuit by breaking only the Neutral wire will leave the whole primary circuit to mains potential.

The power cord should have an adequate strain reliever. A knot in the cord (as seen in some vintage or "hacked" equipment) is definitely not a proper strain reliever! The earth wire of the cord should connect the chassis close to the point where mains is brought in. To accompany the rule, leave the earth wire longer than L and N wires: it should be the last one to break loose under tension. Better yet, use an IEC socket instead: Best ones are equipped with a built-in fuse holder, on/off switch and mains filter! By using something like this the reduction of wire clutter is astounding.

When earthing the chassis, do consider the fact that the chassis may likely corrode during its lifetime. Always use a teathed washer that penetrates the anodized or corroded layer and secures a solid electrical connection with least resistance as possible. Thermal changes may affect the tightness of screws: Use locktite to ensure a sustained mounting. Also, make sure that heat sinks are properly earthed if necessary. Some heat sinks have an anodized surface that acts as an insulator and the earthing connector must penetrate it.

It is recommended that primary side circuits have a different circuit board than secondary side circuits. In most cases it is recommendable to simply point-to-point construct the primary side: Separate wiring usually takes less space than a PC board that has to have traces that are thick enough to withstand the high currents. The layout of filter capacitor bank is also much easier to arrange if the capacitors are chassis mounted since capacitors always tend to come in various shapes and sizes. Also, integrated bridge rectifiers will require chassis mounting or some form of heat sinking to meet their stated power dissipation figures.

Since most chassis contain ventilation holes make sure that the user cannot touch any parts of the circuit through them by accident. An IP-class 10 will not allow objects

larger than 50 mm to enter the chassis. The figure is rather large but becomes smaller when the safety class advances. At IP-class 40 the size is reduced to one millimeter. This covers objects like screwdrivers or nails. While the possibility of someone sticking a long screwdriver through a chassis ventilation hole is somewhat extreme, a perfect example of a very common accident is sticking a metal-framed phone plug through one. This can happen unbelievably easy: Commonly people are lazy enough to “poke around” the places that are not directly under their eyesight; “The jack is somewhere over there...” Once again, to prevent such mishaps from causing excessive damage or injuries, I stress the importance of using heat-shrink tubing.

6 Preamplifier circuitry

6.1 What is a preamplifier?

This is surprisingly a very frequently asked question. By definition, preamplifier is an electronic amplifier, which precedes another amplifier and prepares signal for further amplification or processing. The definition is basically excellent but naturally as imprecise as the function of a preamplifier itself: In other words, preamplifier is just a circuit that does some “preparing” for the signal. It may be a (less than) unity-gain buffer that merely “prepares” a proper impedance match between the guitar and power amplifier - or a complex circuit with dozens of controls and features that “prepare” volume levels, tone and special effects – and very often also handles switching of alternative signal paths for a wide selection of different “preparations”. Since the definition leaves so much room for variety it should be realized that preamplifiers could just as well be found inside amplifiers, guitars, rack units or foot pedals. It is not mandatory for a preamplifier to contain active circuitry, nor does it have to introduce any voltage gain. (Note that gain figure, which is *less than unity* is still considered as “gain”). In fact, the signal “preparation” is the one and only distinct feature defining what is a preamplifier. Unfortunately, for a novice the terminology can cause a lot of confusion, which then manifests itself as questions like “is my distortion box a preamp?” Everyone having an insight of this understands that the answer is *yes*; however, in most cases the important point is not the *semantics* but the fact does the “preamp” have the certain (and required) features required to *interface* the guitar with the power amplifier. This is what counts.

6.2 About transfer curves and clipping characteristics

Clipping characteristics of a gain stage are not a result of one single variable (such as active device type) but many of them. Generally they are:

- Transfer characteristics (gain/transfer curve) of the device
- Gain setting of the circuit
- Amount of feedback applied
- DC bias point
- Rail voltage
- Post -and pre filtering.

Transfer curve defines how much current will flow through the amplifier at a given input current or voltage. The ratio between input current/voltage and output current does not change linearly when the operation is getting closer to saturation or cutoff regions of the device. Basically the transfer curve defines the gain margin and the requirement for operating conditions that (supposedly) should keep the operation as linear as possible. A perfect linear device would have an ultimately steep and long linear area in the curve, which would lead to a high gain without distortion. Unfortunately there are no perfect devices: Transistors can offer a steep linear part but it will typically end with a “hard knee” at cutoff and saturation regions. This causes a sudden and nasty clipping. Tubes generally have a softer knee, thus softer clipping as well, but the linear part of the curve is less steep, which makes these devices poor for high gain applications. Generally, even the linear area is never completely linear,

which causes some distortion as well. It is commonly approved that vacuum tube triodes possess the most linear portion of the curve when compared to any other active device. Transfer curves of field effect transistors are generally something between tubes and transistors. They have a softer knee at saturation and cutoff than transistors and steeper linear area than tubes. Nevertheless, these characteristics are merely generalizations.

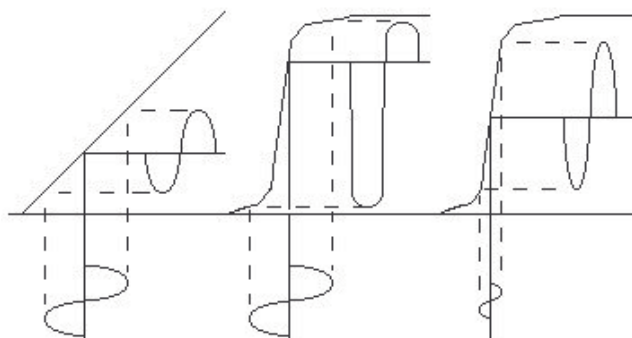


Figure 6.1 Transfer curves and their interaction with DC bias point: Left: Linear with a gain of unity. Middle: Linear with high gain, incorrect DC bias point and high amplitude input signal. Right: Linear with proper DC bias point and low amplitude input signal.

Gain of the actual device defines how steep the linear area of the transfer curve is since the curve is a function of base current to collector current or gate-source voltage to drain current. Higher gain devices have a steeper linear area and vice versa. Field effect transistors and tubes have a huge variation in gain figures even among identical device models. Transistors are usually more similar but even among them variation does exist.

The gain of the circuit is only loosely dependent on the gain of the actual device: The function of the surrounding circuitry is to define proper DC conditions for the device so that it will operate in a controlled manner. For example, the voltage gain of a transistor is only a function of emitter current, load resistance and temperature, not H_{FE} . The latter variable *will*, however, make a difference. This will be explained later. Usually the maximum gain is limited by applying a certain amount of negative feedback, which also improves the stability of the circuit. Generally, this technique is used in every circuit that provides some gain regardless of the active device type it uses. For example, the voltage gain of a transistor stage is often defined as a function of collector resistance divided by emitter resistance, where any emitter resistance actually means using feedback. Bypassing some of this feedback (with, for example, capacitor from emitter to common) will increase AC gain.

Also, by applying negative or degenerative feedback, non-linearities in the transfer curve can be smoothed out – but only up to a certain level. Usually the effect of extensive negative feedback is an extended and extremely linear area that ends up with a very hard knee. In many cases having less distortion and an extended linear operation area is beneficial, unfortunately the transfer curve’s knee gets harder and harder the more feedback is applied. This character is, again, common to all circuits that use negative feedback, regardless of the active device type. Operational amplifiers that use extremely high amount of feedback are infamous because of their abrupt clipping.

DC bias point sets the “zero” crossing point’s location on the transfer curve. It should be usually in the center of the linear area to ensure maximum headroom. An incorrect

DC bias point will lead the amplifier circuit to reach either saturation or cutoff region earlier on either half wave thus causing asymmetric clipping. *A claim that only vacuum tube circuits exhibit asymmetric clipping is blatantly inaccurate.* Every active device that is biased “imperfectly” (whether intentionally or unintentionally) will clip asymmetrically; this includes operational amplifiers as well. Supply rail voltage defines the headroom of the stage (since it extends cutoff and saturation regions). This variable usually has only a mediocre effect on the amount of gain – and in some systems it doesn’t affect the gain at all.

Then there is the effect of filtering: Low-pass before the gain stage attenuates higher frequencies. If the circuit is driven into clipping all attenuated frequencies will naturally stay clean longer. High-pass filtering before the gain stage has an equal effect – this time on lower frequencies – and it can tremendously reduce the phenomenon known as “blocking distortion” that is caused by low frequency saturation (or cutoff) that clips the signal allowing no low-amplitude, high-frequency signal content to exist. Low-pass filtering after the gain stage attenuates higher frequencies rounding up the hard knees of clipped signal portions. It can therefore remove a lot of unpleasant, high order harmonic signal content. High-pass filtering after the gain stage attenuates lower frequencies bringing “detail” to sound by removing accentuated bass content. The filtering substitutes the flat “DC” portion of a clipped signal with a smooth “descending” portion that, again, removes unpleasant, high order harmonic distortion content. The high pass filtering also affects the DC bias point in the preceding circuitry.

6.3 Basic gain stages

6.3.1 Common emitter

The series feedback circuit (schematic a in Figure 6.2) is the basic “schoolbook” example of a common emitter amplifier. The voltage gain of the stage is approximately R_{LOAD} / R_E . With no “external” load connected R_{LOAD} equals R_L but with a typical load consisting of another series feedback common emitter circuit the R_{LOAD} is a parallel connection of R_L , R_1 , R_2 and the input base resistance of the loading stage. If the stage is loaded by a volume potentiometer, the parallel connection of R_L and the resistance of the potentiometer form R_{LOAD} . Any coupling capacitors will naturally present a load that varies according to frequency. The output impedance is practically defined by R_L since the internal impedance of collector is several megaohms.

If the base voltage rises the emitter voltage rises as well. Since the current flows to common (or V_{EE}) through the collector resistor, transistor and emitter resistor, the collector resistor will cause a voltage drop that lowers the collector voltage. Similarly the collector voltage increases as the base voltage decreases. Thus the topology is inverting. The collector voltage is:

$$V_C = V_{CC} - I_C \times R_C \quad \text{(Equation 40)}$$

To ensure symmetric operation and maximum headroom the collector should be biased to voltage potential that is halfway between the positive rail’s potential and emitter voltage. The base current I_B should be approximately ten times lower than the collector current to ensure that the calculatory base voltage is correct. Emitter current I_E equals I_C plus I_B . The voltage gain of a transistor is:

$$A_V = (R_{LOAD} \times I_E) / T_A \quad \text{(Equation 41)}$$

T_A = ambient temperature. R1 and R2 define the base bias point while R_E provides some degenerative negative feedback that stabilizes emitter current and ensures thermal stability. In theory, the circuit would work with just R1 and R_L but the additional components make sure the operation is repeatable with different devices and in varying temperatures.

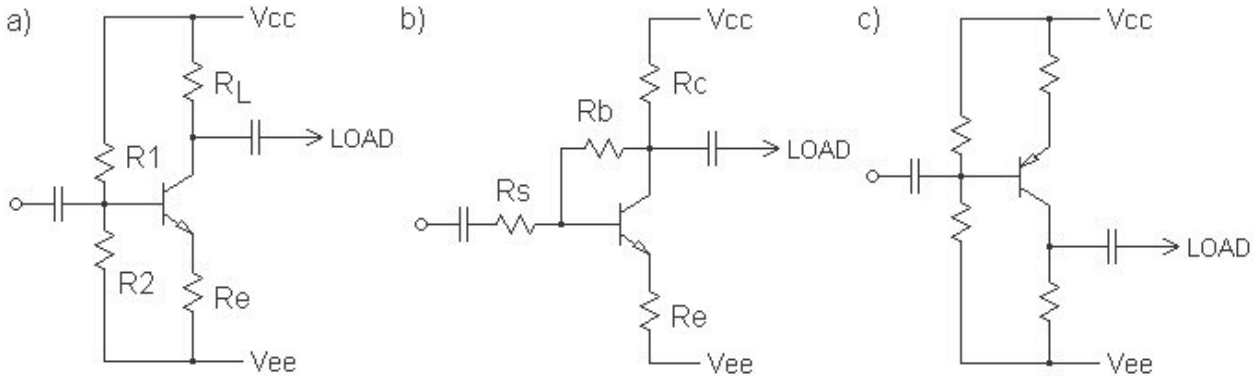


Figure 6.2 Common emitter configurations: a) Series feedback, b) shunt feedback and c) series feedback using PNP transistor.

Calculating component values for the basic circuit is quite easy: The first thing to do is to choose the amount of emitter current. Usually a current of 1 mA is suitable but if the noise level is an issue the current should be decreased. However, note that amplifying higher frequencies essentially requires higher current. The next decision is selecting the emitter voltage. The higher the voltage is, the more stable the emitter current will be during variations in temperature and H_{FE} . A value between 1 to 2 volts should be adequate but if supply voltages are very high one may wish to use higher emitter voltage as well. While doing calculations don't stress yourself with the rail polarity; one can consider a dual supply as a single supply and the following equations will work out just as right. All you need to do is to calculate that

$$V_C = [(V_{CC} - V_E) / 2] + V_E \quad \text{(Equation 42)}$$

is approximately half of the rail-to-rail voltage potential. For example, treat dual supply of ± 15 V as a single supply of 30 V. In this case the V_C should be around 15 V. As the two variables are known the value for emitter resistor works out to be

$$R_E = V_E / I_E \quad \text{(Equation 43)}$$

Since base current is about collector (or emitter) current divided by H_{FE} and should be around ten times lower than I_C , the resistance values for R1 and R2 can be calculated from the following formulas:

$$R1 = H_{fe} \times (V_{CC} - V_B) / (10 \times I_E) \quad \text{(Equation 44)}$$

$$R2 = (H_{fe} \times I_B) / (9 \times I_E) \quad \text{(Equation 45)}$$

Remember that the base voltage is V_{BE} higher than the emitter voltage. The R_L is then calculated from:

$$R_L = (V_{CC} - V_E) / (2 \times I_E) \quad \text{(Equation 46)}$$

Also, note that the impedance of the following stage has to be considerably higher than R_L so that it will not affect the gain. The input impedance of the following transistor stage is essentially R_1 in parallel with R_2 and transistor's base resistance. This is where high H_{FE} figures actually begin to matter since the base resistance of a transistor is equal to $H_{FE} \times R_E$. If there is some capacitance at the emitter (i.e. bypass capacitor for emitter resistor) it must be taken into consideration when doing calculations.

One way to increase input impedance is to connect the base to a separate biasing network. This effectively changes the input impedance to combination of $H_{FE} \times R_E$ in parallel with the "input" resistor. Also, since the base is not directly connected between rails but to a point that is (assumably) well filtered the stage is less vulnerable to power supply induced signals.

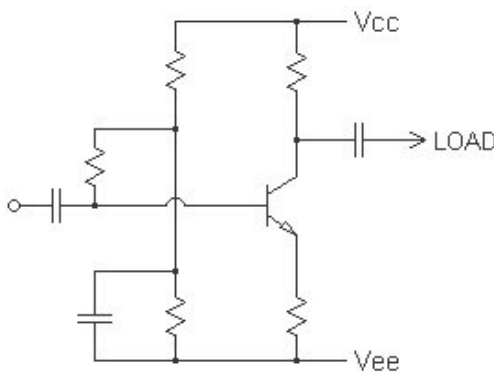


Figure 6.3 Biasing the base with an external bias network.

Since gain of a series feedback circuit is essentially defined by the ratio of collector and emitter load it can be increased by AC bypassing the emitter load either completely or partially. A good rule of thumb is to use a capacitor, which at the lowest frequency of operation has reactance less than R_E divided by 10. It is often useful to design a stage with a fixed gain and define different emitter resistances for both AC and DC. In the example configuration, the emitter is connected to common via two resistors, R_{E1} and R_{E2} , which are in series. R_{E2} is then bypassed with a capacitor C_E . The gain ratio A_f becomes $(R_{E1} + R_{LOAD}) / R_{E1}$, thus one can derive resistor values from

$$R_{E1} = R_{LOAD} / (A_f - 1) \quad \text{(Equation 47)}$$

$$R_{E2} = R_E - R_{E1} \quad \text{(Equation 48)}$$

where R_E is the calculated value that sets the desired emitter current. This method can yield a stable voltage gain up to a figure of 10.

Shunt feedback stage (circuit b in Figure 6.2) utilizes negative current feedback that is realized by connecting a resistor from collector to base. Since the resistor has collector current flowing through it, it sets DC conditions for the base, thus the stage is *feedback biased*. The value of R_B should be selected so that the collector voltage is half of the supply voltage. The approximate voltage gain of the stage is once again collector resistance divided by the emitter resistance. The effect of the feedback

resistor has to be taken into consideration, of course. The input and output impedances of the stage are low since the voltage feedback decreases them. AC bypassing the emitter resistance can yield more gain similarly to series feedback configuration.

A little side note: Most bipolar transistors are inherently reversible, meaning one can swap the place of emitter and collector and the device will still work like a transistor. However, since the junctions are very asymmetric, operating parameters will naturally change quite dramatically, usually resulting into increase in noise and reduction in both speed and beta (which often drops to 10 or so). V_{CE} max also converts to V_{BE} max and vice versa. “Reversing” the transistor will not convert it to a different polarity!

6.3.2 Stages using multiple transistors

The simplest way to construct a multiple transistor stage is to DC couple two or more common emitter stages. Basically, there is a thin line between whether directly coupled transistors should be considered as a single gain stage or many separate ones. I like to consider them as one due to very close interaction with each other: In such circuit the DC bias point of a transistor is the collector voltage of the preceding transistor, thus the stages do not operate independently. The interaction is also the inherent problem of the topology since the direct coupling can lead to a serious *domino effect*: Should the DC condition in one “stage” become flawed enough, all transistors within the directly coupled stages can be destroyed – or at least their operation will be significantly disturbed. Naturally AC coupled stages that have separate DC conditions are much more tolerant against failures such as this. They are also easier to design and troubleshoot.

Simple, directly coupled stages do not employ much negative feedback, which helps them to pertain “smoother” clipping characteristics. Naturally, the drawback of this approach is that the amount of distortion in the output signal is notably greater than with designs that use feedback. However, omitting the negative feedback will have more serious consequences: The thermal stability of a stage consisting of multiple directly coupled transistors is very poor. An increase in junction temperature will cause an increase in the collector current. Since each transistor, except the first one, biases itself from the collector of the previous stage, the increased collector current will alter the bias point. In directly coupled circuits containing several transistors the alteration will multiply itself and grow very large.

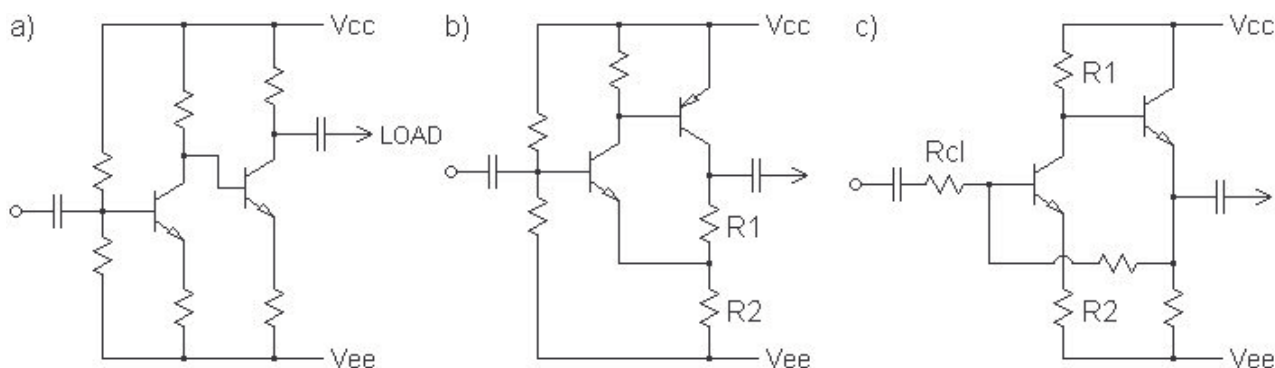


Figure 6.4 Gain stages using multiple transistors: a) Directly coupled NPN common emitter stages, b) series feedback and c) shunt feedback.

If the base or emitter voltage of the first gain stage is controlled by a signal fed back from the last stage the temperature alterations become compensated. Since this is

means applying negative feedback the amount of distortion will decrease as well, along with the amount of gain. The clipping also becomes more sudden since the linear area of the transfer curve (of the whole combination of stages) is “artificially” extended.

Dual-transistor series feedback stage (b) is formed out of NPN-PNP pair. There are two inverting stages in the circuit so the signal is fed back to the emitter of the first gain stage in order to retain the correct phase. Since the output is separated from the input the closed loop gain is defined by the divider ratio of R1 and R2 as $G = (R1 / R2) + 1$. The circuit can be modified by adding an emitter resistor to the second transistor (degenerative feedback) and by shunting the feedback node to ground through a capacitor (higher AC gain) but the additions make defining the gain harder. The THD of the circuit is approximately 0.003 % – 0.08 %. Acquiring a good figure of course requires a very clean supply, a clean common reference and proper tweaking to obtain optimal component values. THD is strongly dependent on the collector resistance of the first transistor.

Two-transistor shunt feedback stage (c) consists of a common emitter gain stage and an emitter follower buffer stage. Because the circuit is inverting the output signal is fed back to the base, which provides DC biasing and temperature stability. A resistor marked R_{CL} provides some current limiting. The voltage gain of the circuit is approximately $R1 / R2$, although the feedback (along with the internal resistance of the transistor) will decrease the amount of gain a bit. The THD of the circuit is approximately 0.05 % – 0.2 %. However, this type of feedback arrangement causes the stage to have very low input impedance, which in turn loads the previous stage.

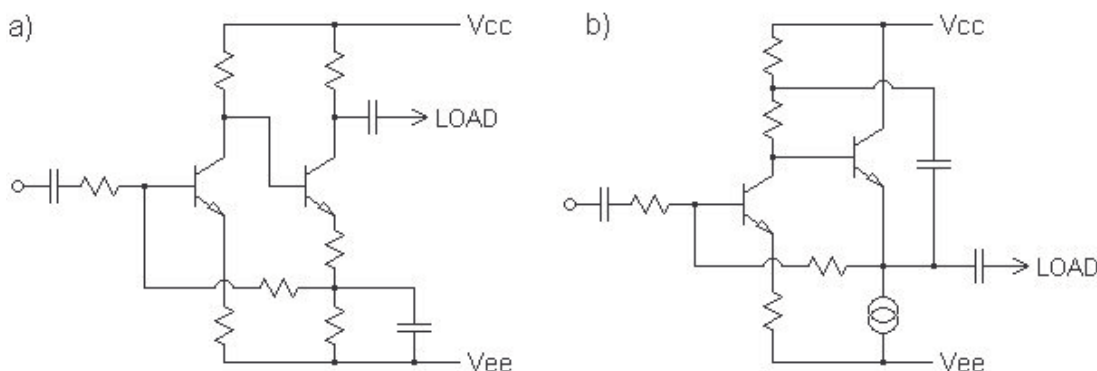


Figure 6.5 Modifications: a) Shunt feedback amplifier with two common emitter stages and b) bootstrapped shunt feedback amplifier with constant current emitter load.

Since the output of shunt feedback stage is basically a buffer, replacing the collector resistor of the first transistor with a network of two series resistors and connecting a bootstrap capacitor from their center point to the output can make the stage constant current loaded as in figure b. The configuration will allow the circuit to provide an enormous gain (if required) but at the same time it increases THD, which rises closer to 1 %. Obviously the bootstrap could be substituted with an active CCS to improve the matter. Also, the emitter resistor of the buffer stage could be replaced with a CCS as well. However, I feel it’s unnecessary to focus more on these modifications. Even the basic circuit should have plenty of gain.

Another shunt feedback circuit directly couples two common emitter circuits and utilizes feedback to adjust bias and thermal stability. This circuit can have a very high gain, especially if the emitter of the second stage is bypassed with a capacitor. However, the THD figure is related to gain: At a gain of 10 it is around 0.02 % but has

increased to only 0.9 % when the gain is 100. Also, the amount of gain is hard to define since it's not directly a function of collector and emitter resistances. Basically the gain of the first stage decreases if the gain of the second stage increases. If gain of the second stage is 1, the gain of the first stage is about collector resistance divided by the emitter resistance. However this value decreases to about 2 % of the initial value when the gain of the second stage is set to 100.

As someone may have already noticed, the circuit is identical to that used in the famous *Fuzz Face*. I find it quite funny that so many people persistently criticize "transistor sound" and the use of negative feedback but still praise the tone of this circuit since it essentially combines both in the "worst" possible way. This circuit clips asymmetrically like almost any other combination of common emitter stages that has an indefinite bias point.

6.3.3 Common source

Field effect transistors operate very similarly to bipolar junction transistors: In a BJT the base current controls the collector current while in a FET the gate voltage controls the source current. Where FETs differ from BJTs is that the gate current is nearly non-existent, while the base current of transistors is always higher than zero. Due to this, the input impedance (or gate impedance) of a typical FET is very high, usually several megaohms. In practice, the gate resistor that is in parallel with the gate impedance defines the stage's input impedance. Similarly to transistors - and especially to vacuum tubes - the gate has some capacitance in relation to source (known as C_{SG} or input capacitance); this capacitance is usually proportional to the type of the FET's casing as follows: TO-72: 2 - 30 pF, TO-205D and To-226A: 30 - 500 pF, TO-220: 500 pF - 3 nF and TO-218: 3 - 5 nF. Beside input capacitance, FETs also have some drain-to-gate capacitance (known as C_{DG} or "reverse transfer capacitance"). The amount of it is smaller than C_{SG} but it is being "amplified" by Miller effect - similarly as in BJTs. Since impedances of stages using FETs are generally high, these capacitances become very notable factor even when one's not dealing with high frequencies. One effective way to reduce the input capacitance is cascoding. Also, the V_{GS} is notably higher than the BJT equivalent V_{BE} ; with FETs this variable is about 1 - 4 volts depending on the model. However, *the PN-junction of a FET is reverse biased so the source is always more positive than the gate.*

The characteristic curves show the similarity to bipolar junction transistors and vacuum tubes: When there is no drain current, the drain voltage is zero and the FET is in a cut-off state. This condition is met when voltage difference between gate and source is lower than the gate-source pinch-off voltage, $V_{GS(OFF)}$. When the drain-source voltage rises high enough ($V_{DS} > V_{SAT}$) the FET enters the saturation state where an increase in V_{DS} does not lead to any further increase in I_D . Between these states the FET operates in resistive region where an increase in V_{DS} results to a nearly linear increase in I_D . The voltage that defines the line between saturation and resistive regions is called the "pinch-off" voltage. The saturation state is also known as "pentode" region and the resistive state as "triode" region [29]. Sometimes saturation state is called pinch-off state.

Considering their characteristics, *FETs are much closer to pentodes than triodes*: Basically, when V_{DS} is greater than V_{SAT} , the drain current is controlled solely by V_{GS} . *The saturation state of a FET therefore does not behave similarly as the saturation state of a BJT!* Actually most JFET AC amplifier circuits are operated in the

saturation region: Since the V_{GS} curve is straight it provides a more linear ratio of control over I_{DS} . P-type JFETs operate similarly but the direction of the current and polarity are reversed.

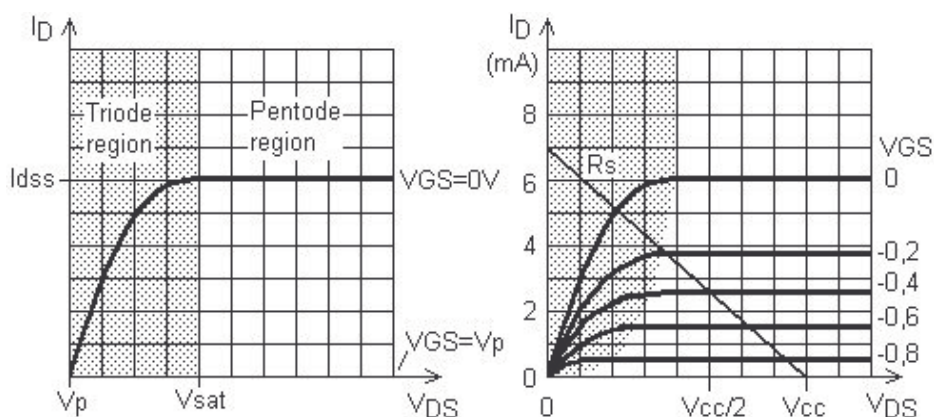


Figure 6.6 Basics of FET characteristic curves. Left: Output characteristics shows relation between I_D and V_D on a given V_{GS} , right: A load line defines operating point (Q-point).

The current flow through the FET is controlled with voltage: N-type FETs are open when the gate voltage is zero. When V_{GS} is made more negative, the current flow decreases. A load line can be used to define the Q-point: It is in the intersection of load line and a suitable V_{GS} curve. The output characteristic curve in the datasheet of a concerned FET will define how much a change in V_{GS} will change the drain current. For example, with a J309 JFET a change in V_{GS} from -2.7 V to 0 V will cause the drain current to increase from 0 to 23 mA. With a drain resistor of 1 kilo-ohm this corresponds to a voltage change of 23 V. This is equal to voltage gain of 8.5 ($2.7 V_{GS} \rightarrow 23 V_D$). *Note that the datasheets list the maximum drain current, which should never be exceeded. A safe design principle is to limit the maximum drain current to about 75 % of the value listed in the datasheet.*

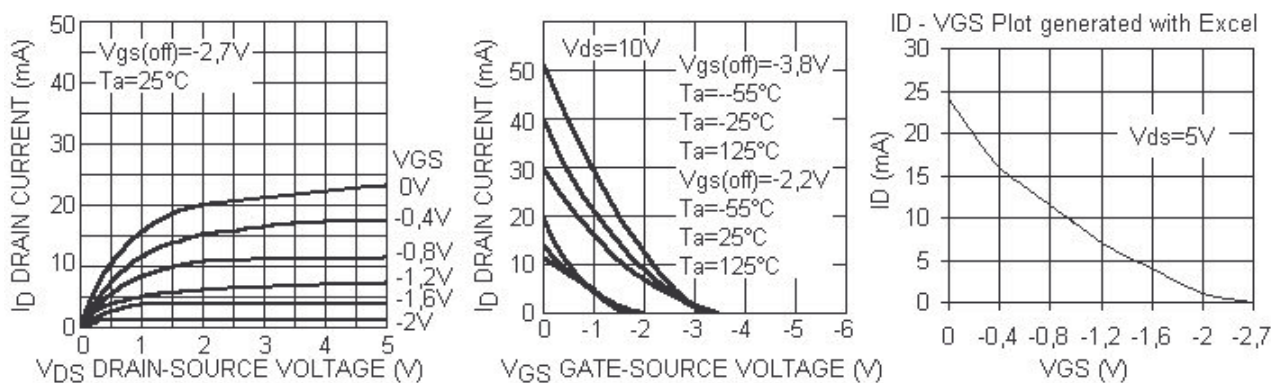


Figure 6.7 Examples of characteristic curves of J309 N-type JFET.

The drain is usually biased to voltage potential between positive and negative rail to ensure maximum headroom and output signal swing. I_D that corresponds the voltage potential is defined by the output characteristic graph. The middlemost curve does not necessarily present the “zero” point of the input signal where the V_{GS} would be about half of the whole range of input voltage swing. The important parameters to define this are $V_{GS(OFF)}$ and curve for V_{GS} of 0 V. The required drain current is usually referenced according to them. For example, to bias the drain of J309 to half of 20 V supply the graph indicates that I_D of about 7 mA is required ($V_{GS} = -1.2$ V). (The graph does not show more than 5 V V_{DS} but the curves assumably continue very linearly). The drain resistor’s value is worked out using Ohm’s law: 10 V divided by 7

mA yields 1.43 kilo-ohms. We choose the nearest standard value. The circuit still needs a source resistor to stabilize the current. Since the “zero” point of the input signal is the V_{GS} curve of -1.2 V we want to bias the source to a steady voltage potential that is 1.2 V higher than the gate voltage when no input signal is present. According to Ohm’s law, the source resistor must then have a value of 1.2 V divided by 7 mA; this yields 171.4 ohms.

With FETs, terms voltage or current gain are not really used. Similarly to vacuum tubes, the datasheets list a variable called *transconductance*, which is a function of drain current and source voltage when V_{DS} is constant. Since with FETs the relationship between input and output quantities is nonlinear a graphical approach to define operation conditions may prove faster, although it is less accurate than a mathematical method. Some generic mathematical definitions can still be applied to all FET circuits. The most important rule is

$$I_G \approx 0A \rightarrow I_D = I_S = I_{DS}, \quad \text{(Equation 49)}$$

which shows that since gate current can be completely ignored the currents at both drain and source must be equal. For JFETs and depletion type MOSFETs the drain current as a function of V_{GS} is derived from the Shockley’s equation:

$$I_D = I_{DSS} \times [1 - (V_{GS}/V_P)]^2, \quad \text{(Equation 50)}$$

where I_{DSS} = saturation drain current, that is dependant on V_{GS} , and V_P = gate pinch-off voltage, which is the gate-source voltage that reduces I_D to a very low value. Basically $V_{GS(OFF)}$ listed in the datasheet is a very good estimation of V_P . The equation is true while V_G stays within a range of $V_P \leq V_G \leq +0.5$ V. The requirement shows that the Shockley’s equation is therefore valid only for operation at saturation region. The formula also allows plotting the characteristics of a FET when all the variables are known.

FET circuits can be biased mainly by three ways: Using constant-voltage bias, using constant-current bias or using source bias, which is also known as automatic or *self-bias*. The latter technique is the simplest one and suits well for AC amplifiers. The self-bias establishes the operating point by applying the voltage dropped over source resistor to the gate. To keep the gate-source junction reverse biased (gate is negative in reference to source), the source voltage must be higher than the gate voltage. If the gate is tied to common potential (0 V) with a resistor then the voltage potential at the source resistor is effectively higher than the base voltage, thus circuit is biased. The important equations are:

$$R_D = (V_{CC}/2)/I_{DS} \quad \text{(Equation 51)}$$

$$R_S = V_{GS}/I_{DS} \quad \text{(Equation 52)}$$

V_{GS} and I_{DS} are derived graphically from the datasheet. Note that deriving the V_{GS} from another curve than the middlemost is a valid technique as well. With higher V_{GS} figures the drain current is lower and consequently the related resistances are higher. For instance, the example J309 circuit would perform well with $R_D = 6.8$ kilo-ohms and $R_S = 3.3$ kilo-ohms if the supply would be about 12 V. At these values the V_{GS} would be approximately -2 V and $I_{DS} \approx 0.9$ mA. This corresponds to the lowest V_{GS}

curve shown in the graph. However, the circuit has a different gain ratio, which goes hand-to-hand with the current consumption figure: A higher drain current will naturally yield a higher gain since the voltage drop over drain resistor gets higher as well. *In a properly biased circuit the variations in V_{CC} do not affect the gain but the collector's voltage potential*, unless the circuit drifts too much from the bias point and clips to rails of course. However, increasing the supply voltage and *re-biasing* the stage can yield more gain.

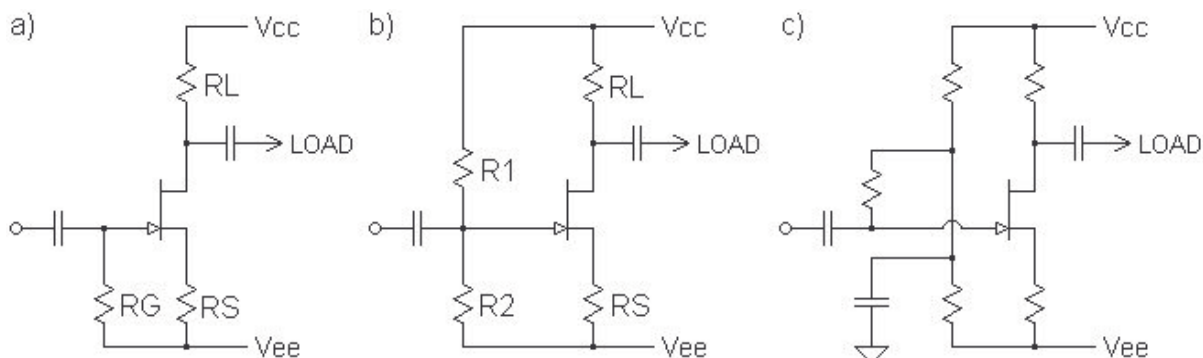


Figure 6.8 Biasing configurations for JFET: a) Self-biasing, b) voltage divider biasing and c) offset gate biasing.

The self-biasing method can be substituted with a voltage divider bias configuration, similar to what is used in bipolar junction transistors. Such configuration is more commonly used with MOSFETs but works just as well with ordinary JFETs. Basically, the biasing procedure follows principles and laws of self-biasing method; the most important is that the source must be positive in relation to gate. Gate current can be ignored but since R_1 and R_2 define the gate's voltage potential they affect the biasing process. Based on self-biasing principles the required gate-source difference should already be known, this also indicates the value of I_{DS} . Also, known is that the V_{DS} should be half of V_{CC} for maximum output signal headroom. *Note: Similarly as when calculating BJT circuit values these equations work for dual supplies. Make all calculations treating the dual supply as single supply: For example, $\pm 10\text{ V}$ dual converts to 20 V single.*

The only way to determine V_G is from the relation between R_D and R_S , which also affects the gain. Essentially this works out to be a compromise between gain and input impedance since the lowest resistance in the input is dominant. The higher the ratio between R_1 and R_2 is, the lower the value of R_S is in proportion to R_D . Ratio should always be at least higher than 2 to ensure that the circuit operates correctly. Basically the exact gate resistor values have little of importance in delivering gate voltage from

$$V_G = V_{CC} \times [R_2 / (R_1 + R_2)] \quad \text{(Equation 53)}$$

After finding out a proper ratio for gate resistors the source potential is derived from

$$V_S = V_G - V_{GS} \quad \text{(Equation 54)}$$

Note that since V_{GS} is a negative value V_S will be higher than V_G . After solving these, both R_S and R_D are calculated as usual. This circuit topology tends to have very low input impedance due to practical reasons: A modest ratio of 10 between R_1 and R_2 would mean that with R_1 of 1 Megaohm R_2 would be only 100 kilo-ohms. Input impedance would be approximately 90 kilo-ohms. Again, the way to make preceding stages see a higher input impedance is the "offset biasing" presented in figure c).

In his article “Triode Emulator” (AES Convention 116, May 2004, preprint 6008) Dimitri Danyuk proposes an alternative biasing method that strives to follow the three-halves power law inherent for vacuum tube triode operation. I have not read the article in question but the technical terms used in presenting the topology have stirred some confusion and I feel it’s therefore best to make a subtle review of the concerned biasing technique. Basically, the three-halves power law is known as Child-Langmuir law and it “governs space-charge limited flow of electron current between two plane parallel electrodes in vacuum when the emission velocities of the electrons can be neglected”. The law actually becomes more useful when it’s applied to an equation that estimates cathode current of an electron tube:

$$I_A = \alpha \times (\mu \times V_G + V_A)^{3/2} \quad \text{(Equation 55)}$$

I’m quite sure that Danyuk has based his theory around this principle since the original Child-Langmuir equation is quite irrelevant for semiconductor technology. The most important point is the similarity to Shockley’s equation. The catch is that the Shockley’s equation has an exponent function of 2 while the three-halves power law has an exponent function of 1.5 (“three halved”). Different exponentials in the transfer curves cause difference in the harmonic content of the result signal. Some of the benefits – and drawbacks - of transfer functions based on square law and three-halves power laws can be realized when examining plots of exponent functions. The square law transfer clearly provides much more gain than the comparison but it is inherently very nonlinear as well. Three-halves power law transfer function does not have as much gain but it tends to provide a more linear output.

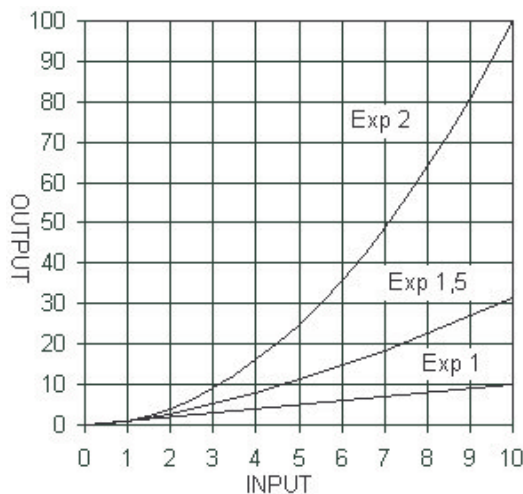


Figure 6.9 Plots of exponential gain functions: Square law (exp 2), three-halves power law (exp 1.5) and linear or unity-gain (exp 1).

Danyuk suggests that in order to acquire transfer characteristics following three-halves power law the source resistor value should be

$$R_S = (0.83 \times V_P) / I_{DSS} \quad \text{(Equation 56)}$$

The equation is basically an assumption of a “correct” V_{GS} value based on values of pinch-off voltage and saturation current. Instead of half-supply the drain’s voltage potential is set to:

$$V_D = 0.66 \times V_{CC} + 0.7 \times V_P \quad \text{(Equation 57)}$$

Therefore the required drain resistance is,

$$R_D = 0.9 \times (V_{CC} - 2 \times V_P) / I_{DSS} \quad \text{(Equation 58)}$$

Where this all leads is acquiring a load line that differs from the load line of the common biasing arrangement by being steeper – thus remaining more at the linear region. To do this, headroom, V_{GS} , is sacrificed causing the circuit to clip earlier.

6.3.4 SRPP

SRPP topology is known with many other names including SEPP (Single-Ended Push-Pull), Micro-amplifier (“Mu”), Totem Pole, Mu Follower, Cascoded source follower and the original name *Series-Balanced amplifier* (US patent 2,310,342 February 1943). It is controversial what the acronym SRPP means but likely it stands for Single-Ended Regulated Push-Pull, Single-Ended Reflexive Push-Pull or Series Regulated Push-Pull. The circuit is based on vacuum tube designs that were used in driving modestly low output impedances in a push-pull configuration and without an output transformer. The topology has a *very close* resemblance to BJT totem pole power amplifier arrangement presented earlier and even closer resemblance to push-pull grounded source, emitter or cathode amplifiers. Basically the topologies are identical, except for the signal coupling and phase inversion configurations.

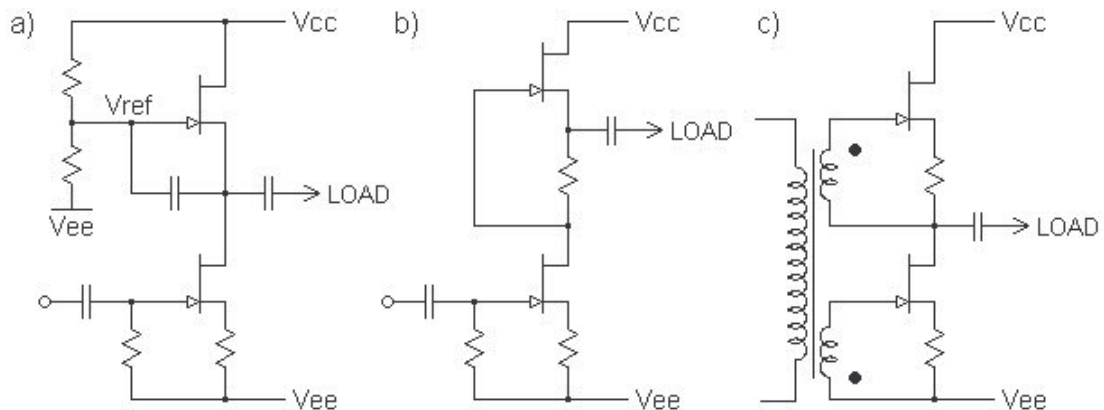


Figure 6.10 SRPP JFET topologies: a) Fixed bias and b) symmetrical. Circuit c) is an interstage transformer coupled push-pull grounded source amplifier.

Although often believed otherwise, the SRPP circuit is really not a single-ended circuit but a true push-pull amplifier that just happens to comprise its own phase inverter. However, class-A operation is mandatory for the functioning of the circuit. Current flow through the lower FET varies its drain voltage, which is then fed to the gate of the upper FET either directly or through a capacitor. The upper FET then amplifies the changes of its gate voltage. Even though the FETs are in series they will operate in a push-pull configuration since the currents on source and drain have an opposite phase.

In capacitor coupled circuit (a) the gate of the upper FET has to be biased (usually) to potential between positive and negative rail to ensure proper operation. The depicted method can be improved by connecting the gate to a biasing voltage divider through a separate resistor, similarly to operational amplifier circuits driven from a single-ended supply. (See 6.3.6 for more details). If divider network resistors are kept low value

and the bias point is filtered the noise of the circuit can be dramatically decreased. With this method a single biasing circuit can also be used for multiple SRPP stages.

Compared to a typical common source amplifier the SRPP provides a higher gain and softer clipping. The gain of the capacitor-coupled circuit also tends to be lower than the gain of the directly coupled one. Basically, the lower the source resistance is, the lower the gain is as well. More gain can be acquired by bypassing the lower source resistance with a capacitor. On a contrary to typical push-pull circuits, no device or drive voltage matching is required. This is a great merit when it comes to FETs. The limitation of the circuit is that it can only be optimized for one specific load impedance: Since the drive voltage of the upper FET is defined according to load current a wrong impedance will lead to an unbalanced operation. This usually does not impose a problem in preamplifier circuits where load impedances stay rather constant anyway.

6.3.5 Discrete buffers

Buffers are used in providing impedance transformation from one circuit to another. Typically a buffer is a circuit that prevents a low input impedance circuit from loading a circuit that has high output impedance. This happens because a buffer is essentially a current amplifier with high input impedance and low output impedance. Having (nearly) unity voltage gain *is not* an inherent property of buffers, although it's very common one. Buffer circuits with unity voltage gain are often called "voltage followers".

A BJT emitter follower has a voltage gain little less than unity but the current gain is fairly high, thus the circuit works well as a buffer. *On contrary to common generalization, the input impedance of an emitter follower is not necessarily high:* It is H_{FE} times R_E in parallel with any accompanying biasing resistors, similarly to common emitter configuration. In practice, the internal input impedance is (usually) high enough to cause the biasing resistors to become the sole defining factor of the actual input impedance. Calculating base biasing resistor values of an emitter follower buffer circuit follows the methods described in calculating the resistor values for common emitter amplifier. However, the base voltage is made equal to voltage potential between positive and negative rail plus voltage over BE-junction. Essentially this means that *the emitter voltage is half of the supply voltage*, which is a logical configuration since it provides the greatest headroom. Once again, a dual supply should be considered as a single supply while doing the calculations. R_E is calculated from the equation:

$$R_E = (V_{CC}/2)/I_E \quad \text{(Equation 59)}$$

Even though FETs have a lower output current delivering capacity and a higher output impedance than BJTs, the high gate impedance makes them more ideal active devices for buffering circuits than bipolar junction transistors are. This is because high input impedance loads the source less. The input impedance is defined by the gate resistance in parallel with the gate impedance, similarly as in common source configuration. Although there are quite many ways to load and power the source (meaning source of the FET), the follower configurations can be divided into two main categories: Ones with, and ones without gate feedback. Input impedance of a circuit without gate feedback is approximately equal to R_G . With gate feedback the input impedance can be increased past this since connecting gate to source essentially acts like a bootstrap.

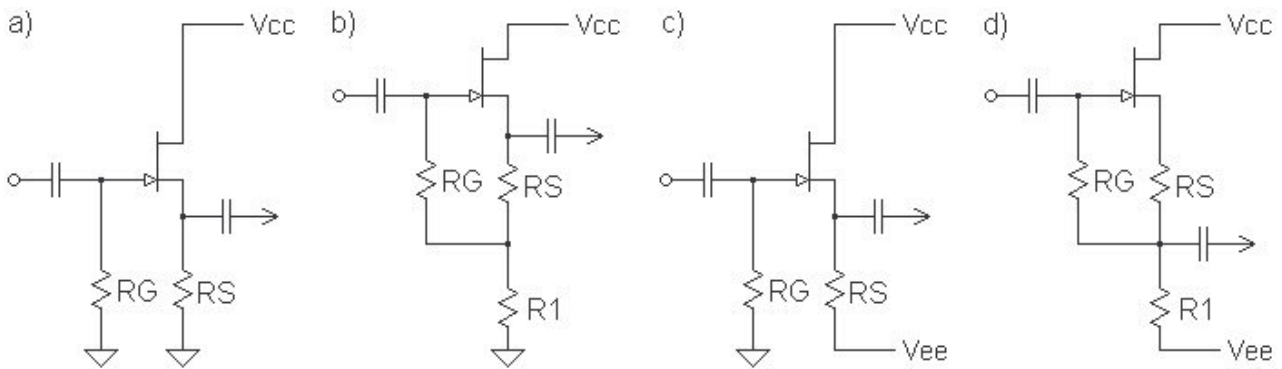


Figure 6.11 Source follower configurations: Circuits a) and c) no gate feedback, circuits b) and d) gate feedback.

In practice, the source resistance can be anything from one kilo-ohm - or even less - to about 10 kilo-ohms. Ideally the loading impedance should be considerably higher not to interfere with the buffer's operation. For AC signals the source load is (naturally) always R_S in parallel with loading impedance.

If one wants to get meticulous and derive the component values (for self-biasing configuration) the basic assumptions withstand: The gate current does not count and $I_D = I_S$. There is no "correct" way to figure out a proper value for I_{DS} . Basically aiming for high V_{GS} (e.g. half of the supply voltage as in BJT circuits) will result in very low I_{DS} . That in turn means using a very high R_S , which is far from ideal. The same goes in opposite direction as well: Small signal stages dealing with low currents hardly require enormously low R_S values - save them for high current stages. Basically, it is worthwhile to set the I_{DS} to a maximum of 1 mA, which will usually mean a V_{GS} of about -1 to -3 V - and this is what the datasheets will indicate as well. The source resistance is then derived from:

$$R_S = V_{GS} / I_{DS} \quad \text{(Equation 60)}$$

Note: One should avoid using JFETs with low I_{DS} - especially in buffering circuits. For example, J201 is a commonly used FET - probably because it was once used in the famous "Till" guitar preamplifier and "FET Preamp Cable", both designed by Donald Tillman. However, the gate cutoff voltage $V_{GS(OFF)}$ of a J201 is about the lowest amongst all depletion mode JFETs and with moderate source resistor values this device can't even handle input signals that are greater than few hundred millivolts peak-to-peak. This FET is a horrible choice for buffers and basically for common source circuits as well: In an equal circuit, a higher current FET, like J309, can handle input voltages higher than 1 V_{PP} and even offer slightly greater gain. Although popular, J201 is really not that marvelous device. Note that Donald Tillman originally substituted a higher current model with a J201 solely because of improved noise performance. Let this be a lesson to you: Always base your component selection principles on circuit theory - not to a fact that a particular component was used in some famous circuit! It might have worked there - likely it will not work as well in another application.

The circuits employing gate feedback are biased similarly as corresponding circuits without feedback. However, R_G is returned to a point that provides almost unity feedback [32]. R_1 is usually higher in value than R_S but not *much* higher, as some sources like to claim: With some FETs, five times higher value than R_S is about the highest one can safely use and the factor decreases while the R_S increases. *Too much*

resistance between source and ground will prevent any FET circuit from working correctly and this rule is equivalent to common source amplifiers as well. Basically, the lower the current the FET uses, the higher the factor can be between R_1 and R_S . For example, J201 can tolerate even 22 times higher values. Main point is that one cannot automatically assume that values, that are *much* higher than R_S would work.

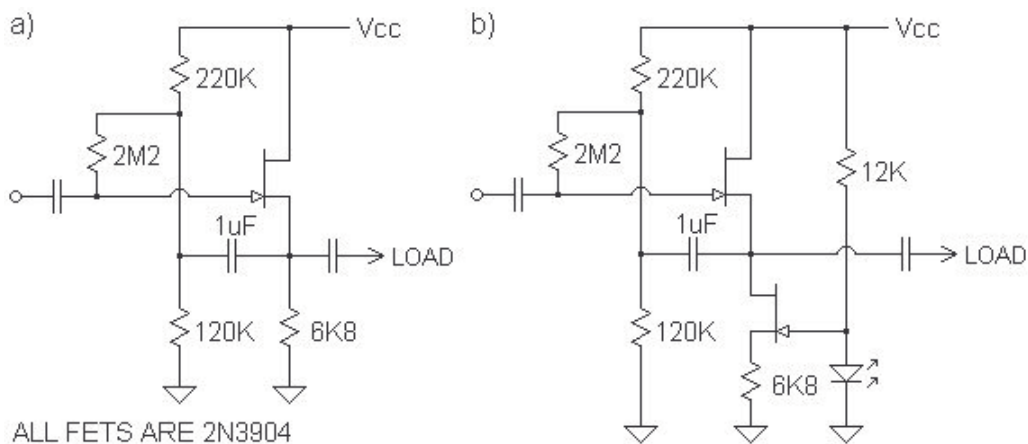


Figure 6.12 Increasing input impedance: a) Bootstrapped configuration and b) bootstrapped configuration with constant current source load.

Another method to increase the input impedance is to substitute the resistor-type gate feedback configuration with a bootstrap capacitor. With properly chosen components the simple configuration can lift the input impedance very high. The circuit in figure a) has an input impedance of 44 Megaohms and the similar circuit using a constant current source load has an input impedance of 500 Megaohms! However, for the later configuration to maintain the high input impedance, the load must be high impedance as well – preferably another source follower [33]. High input impedances of both circuits are acquired by using typical component values, which is a great merit since sourcing for resistors that are higher in value than 10 Megaohms usually proves to be very difficult. Also, noise performance is much better than when using high value resistors. However, I believe that benefits of having higher input impedance than, say, 10 Megaohms are questionable. Many guitar amplifiers with even lower input impedance than 500 kilo-ohms have been made and they sound very good too. *Note: Generally, too high impedance can even become detrimental since it requires an extremely good layout and board design to prevent the circuit from catching interference, becoming unstable or just being extremely sensitive for stray capacitances.*

Since the biasing configuration differs from self-bias the bootstrapping trick also works very well with BJTs. Ultimately, using the topology in conjunction with an ordinary transistor might induce better performance than FETs could ever provide since transistors generally offer higher output current and lower output impedance – both ideal properties for a buffer circuit.

6.3.6 Operational amplifiers

An operational amplifier is essentially a *differential* amplifier, which means it amplifies the difference between its inputs. Contradictory to conventional differential circuit an operational amplifier has only one output, though. The open loop gain of an operational amplifier is very high but generally they are never used in open loop mode, instead the gain is controlled by the amount of negative feedback. *An operational amplifier tries to keep its inputs in the same voltage potential and if it*

fails to do so the output will assume the polarity of the most positive input. This simple rule explains the basis of every operational amplifier circuit.

The circuit a) is a non-inverting amplifier. It has very high input impedance so it can tolerate large series impedances. Still, in most of the cases the input impedance is defined by R3. The resistors R2 and R1 form a voltage divider that defines the gain. Because the operational amplifier tries to maintain voltage potentials at its inputs equal, connecting the center node of the voltage divider at one of the inputs means that the OpAmp must adjust its output to a higher voltage. However, since signals at the inputs are in phase, the gain defined by the divider increases by 1. The gain is therefore

$$A = 1 + (R2/R1) \quad \text{(Equation 61)}$$

The resistor R3 essentially ties the input to reference voltage V_{REF} , which is half of the rail voltages feeding the OpAmp. (Zero volts in dual supply circuits). R3 is extremely important in cases when the feedback loop's connection to reference (or common) through R1 is also capacitor coupled. R3 (along with other parallel impedances) also sets the input impedance. To prevent the bias currents of the input from interfering with the offset voltage the value of R3 should ideally be equal to parallel connection of R2 and R1.

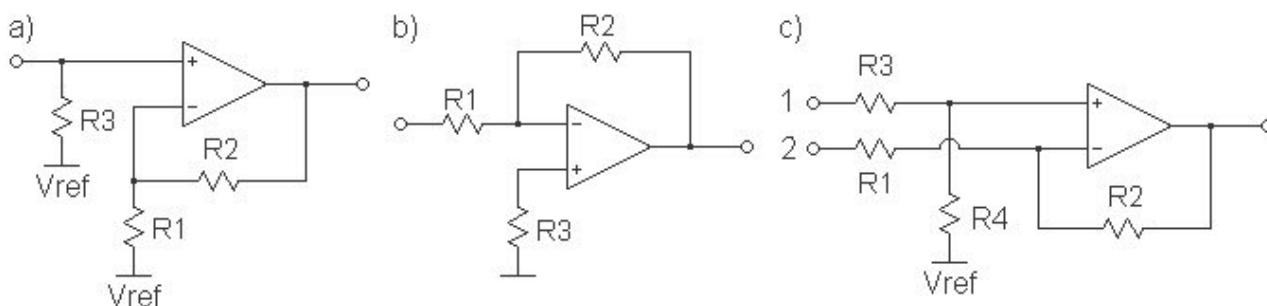


Figure 6.13 Operational amplifier configurations: a) Non-inverting, b) inverting and c) differential.

The circuit b) is an inverting amplifier and its gain is set similarly by the voltage divider consisting of R2 and R1. Because inverting input is used, the signal from output has an opposite phase with the input signal, which in turn causes them to cancel each other out at the input node, thus forming a “virtual ground” with potential equal to reference voltage. This both causes the voltage potential to be equal with the non-inverting input (which is tied to reference potential) and sets the input impedance to R1. The concept is exploited in a summing amplifier configuration where several sources are connected to the non-inverting input through summing (or mixing) resistors. The gain of each source is defined by the ratio of the summing resistor and R2. It is important to note that the source impedance is *in series* with R1, thus it must be fairly low or otherwise the gain will decrease. For the same reason devices having output impedances that shift on a large scale (like dynamic microphones) are not ideal signal sources for this configuration. The gain of the inverting circuit is calculated using a formula:

$$A = - R2/R1 \quad \text{(Equation 62)}$$

The purpose of the resistor R3, connected from non-inverting input to V_{REF} , is again to prevent the bias currents of the input from interfering with the offset voltage. The

resistor's value should be equal to that of parallel connection of R2 and R1. The resistor will not affect the gain.

Circuit c) is a differential amplifier like the first two. However, this time instead of shunting one of the inputs to V_{REF} both inputs have a signal source. The gain is equal to

$$V_{OUT} = [(R1 + R2)/(R3 + R4)] \times (R4/R1) \times V_{IN2} - (R2/R1) \times V_{IN1} \quad \text{(Equation 63)}$$

If R1 is equal to R3 and R2 is equal to R4 the output voltage can simply be expressed with equation

$$V_{OUT} = (R2/R1) \times (V_{IN2} - V_{IN1}) \quad \text{(Equation 64)}$$

A differential amplifier is especially handy in amplifying balanced signals since the configuration removes all common mode errors and distortion that have been induced to inputs. Due to configuration the impedances of the inputs are different. A practical circuit would require either a way to match the impedances with component values or with buffered inputs. Also the components used in the circuit have to have a close matching.

An important point to notice is that an operational amplifier does not necessarily reference its inputs to common/ground: It references them to voltage potential of V_{REF} , which is preferably $(V_{CC} + V_{EE}) / 2$ – whether that's zero volts or not. The bias reference is often set to be the common but it is false to assume that this would always have to be the case. Essentially this means that the inputs and output will assume V_{REF} as their “zero” potential and swing around it. Thus incorrectly biasing the operational amplifier will create DC offset and cause asymmetric clipping.

Operational amplifiers are not created equal. Generally they can be divided into two main categories: Ones with FET input and ones with BJT input. FET-input OpAmps have very high input impedance while BJT-input OpAmps have a lower one. Therefore most BJT input operational amplifiers can be replaced with FET input models but the other way around might cause a quirk in performance due to impedance mismatch. While slew rate, noise figure, CMRR and PSRR of different OpAmp models may vary the effects of that are usually not drastic enough to totally impede operation of a conventional audio circuit. However, there are more important differences to consider: Some operational amplifiers have a different range for supply voltages and in some rare cases the pinout of different OpAmps may also diverge from the standard. These are things you should always check out. In fact, always consult the datasheet of any IC before using it. Note that different OpAmp models have different ranges for input voltage and output voltage swing. Some operational amplifiers also have an internal compensation capacitor that increases NFB at higher frequencies. If an OpAmp with internal compensation is replaced with a one that does not include it there is a chance for oscillation. Some operational amplifiers, especially high-speed ones, are also very sensitive to proper supply rail decoupling in close vicinity to supply pins. There is a good chance that replacing a less-sensitive OpAmp with a more sensitive one also starts oscillation if the rail decoupling is realized poorly.

Few other less well known but important facts concerning operational amplifiers are:

- On a contrary to typical transistor circuits, the noise of an operational amplifier is higher with lower current consumption.
- Most OpAmps have an internally built short-circuit protection that creates a limit for maximum output current. It is safe to assume that 600-ohms is about the lowest impedance an OpAmp is capable of driving. Some OpAmps can't even drive that one properly.
- An operational amplifier cannot drive a purely capacitive load. For capacitive loads the operational amplifier requires a series load resistor.
- The unused OpAmp sections (in dual and quad ICs) have to be terminated by connecting the inverting input to output and hooking the non-inverting input to potential between positive and negative rail. Inappropriately terminated pins may catch interference and obstruct the performance of the other stages within the IC package.
- Most operational amplifiers can work as comparators but *only in low frequency applications*. This is due to internal high frequency compensation that most of them have. *Note that OpAmps are not designed for open loop applications, such as being used as comparators. Generally, manufacturers cannot, and will not, assure a predictable operation in such cases. A comparator does not work as an operational amplifier.*
- There are operational amplifiers with an internal channel switching function: These typically consist of two differential circuits A and B that share a common output stage. Controlling which one of the input differentials is used at a time is done by a logic level signal. See 7.4.5 for more details.

Operational amplifiers located in the input stages should be protected from static discharges that may reverse bias the differential stage transistors. A simple and effective protection method is hooking up two parallel diodes in reversed orientation (“anti-parallel”) across the inputs. This will clamp the maximum voltage difference between the inputs to a safe level. (Remember that an operational amplifier tries to keep its inputs in the same potential so normally there is a difference of only a few microvolts between input potentials). Some devices (for example the 553x series) already apply this protection method internally.

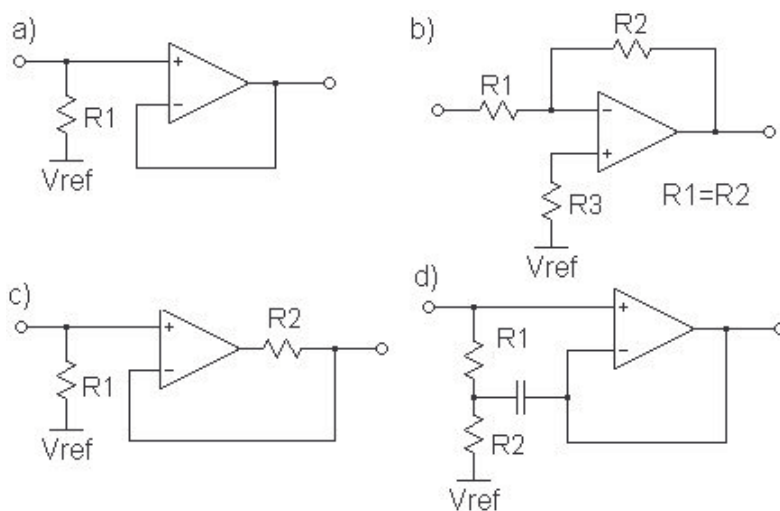


Figure 6.14 Operational amplifier buffers: a) non-inverting, b) inverting unity-gain stage, c) non-inverting buffer with compensation against capacitive loading and d) bootstrapped high input impedance buffer stage.

A conventional operational amplifier buffer circuit has unity voltage gain. There are mainly two ways to realize one: Form a loop with 100 % negative feedback or use the inverting configuration in unity gain. The first one is the most common buffer version because it is non-inverting (no change in signal information) and requires no external components. Circuit b) is used in occasions when there is a need to get both inverted and buffered output. For example, the circuit would be an ideal choice when making an inverter stage for a bridged stereo amplifier configuration. The circuit c) is a modification of the circuit a) with higher output impedance. However, it is more tolerant against capacitive loading due to series resistor in the output.

It is possible to apply some active filtering within the operational amplifier stage. Most common configurations are first order low- and high-pass filters. High-pass filtering, as applied in circuit a), is a very versatile technique because it also sets DC gain to unity, thus any offset voltage present in the input will not be amplified. The low-pass filter b), on the other hand, rolls off high frequency gain improving the amplifier's stability and reducing the risk of oscillation. Some operational amplifiers are internally compensated with similar filter – some are not. The filter gets more crucial the more gain the stage introduces. The figure shows both filter types in separate circuits but naturally they can, and should be, combined into a single stage. The roll-off points and gains for filter circuits a) and b) can be calculated from equations:

$$f_{-3dB} = 1/(2\pi \times R2 \times C1) \quad \text{(Equation 65)}$$

$$Avf = 1 + [(R2 + X_{C1})/R1] \quad \text{(Equation 66)}$$

$$Avf = (R2 \parallel X_{C1})/R1 \quad \text{(Equation 67)}$$

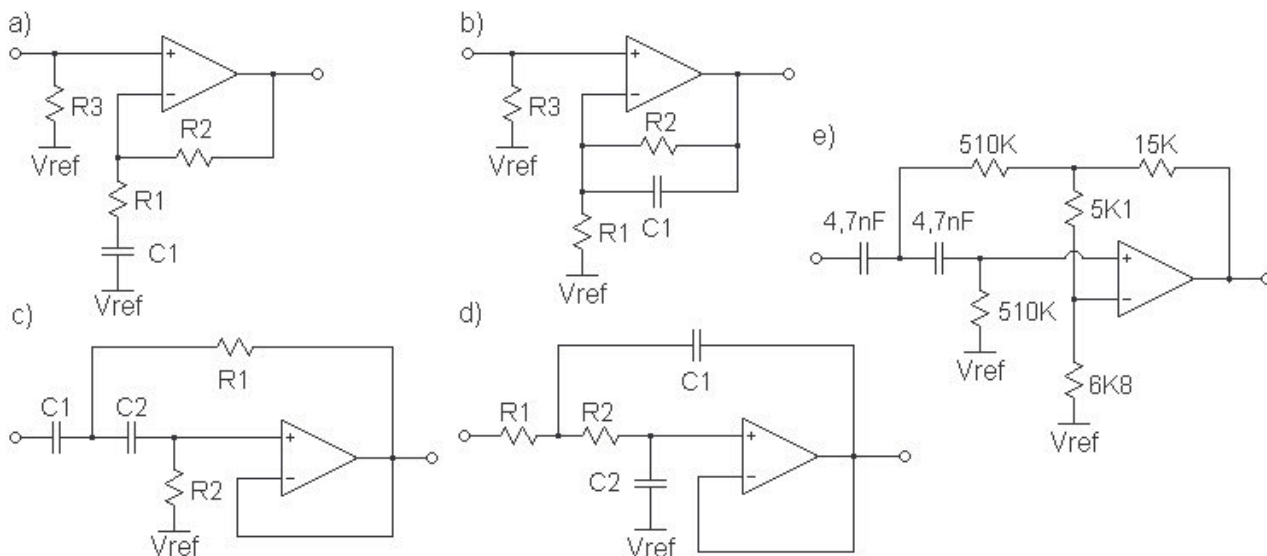


Figure 6.15 Filter configurations: a) First order high-pass, b) first order low-pass, c) Sallen-Key second order high-pass, c) Sallen-Key second order low-pass and e) input stage of Fender Deluxe 90.

Second order filters act like two first order filters in series; that is, the -6 dB decrease per octave of the first order filter converts to -12 dB. The schematic shows a unity-gain buffer stage in conjunction with the filter - this is solely to simplify the calculations for frequency response and component values. The stage can be configured as a “conventional” amplifier with gain as well but it will make component value derivation less straightforward because the filter tuning and style change

according to gain. Fundamentally, the high-pass filter is ideal to be used in conjunction with the input gain stage (see circuit e) as it can quickly roll-off the unnecessary bass content that “muddies” up the sound and causes blocking distortion when stage is overdriven. The roll-off point for a (Butterworth-style) second order filter is:

$$f_{-3dB} = 1/2\pi \times (R1 \times R2 \times C1 \times C2)^{0.5} \quad \text{(Equation 68)}$$

Sometimes a need arises to drive an operational amplifier from a single supply. This can be done by utilizing principles of “virtual earthing”, explained in 5.2. The operational amplifier’s input is biased to half of the supply rail voltage using rail-splitter resistors. The absolutely wrong way to do this is shown in figure a). The circuit has relatively low input impedance since it is a combination of two resistors in parallel. Any change in supply voltage will also change the bias voltage set by the divider network, which means that any common mode signal in the rails will be amplified, along with the input signal. If the operational amplifier must supply large currents it may also couple the signal voltage to the supply line – unless the supply is well filtered or regulated - and if the input is directly referenced to supply lines this can cause serious “motor boating” issues [16]. Basically the circuit a) is only good for battery powered applications.

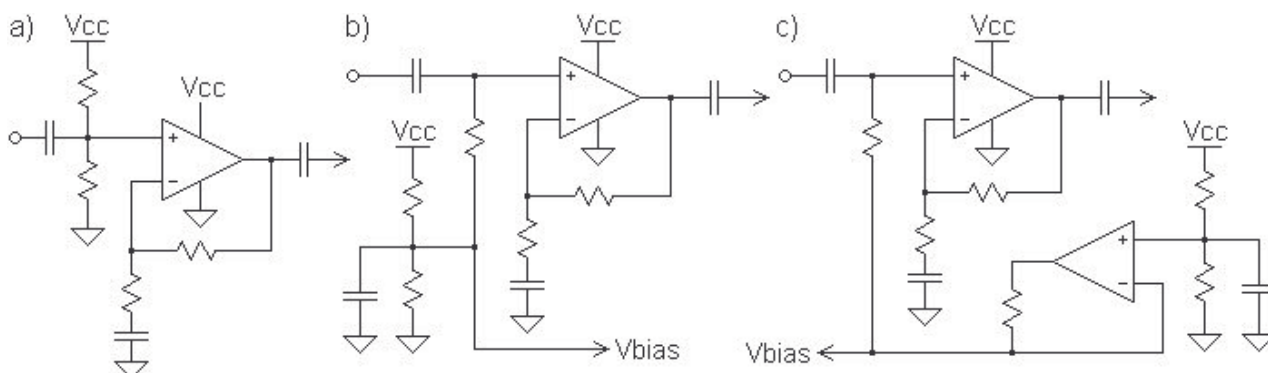


Figure 6.16 Reference biasing methods: a) Wrong way, b) better way and c) buffer driven potential (“best” way).

The circuit in figure b) is slightly better and perhaps the most commonly used. Expect to find this circuit from most guitar pedals that are powered by a 9 V battery and from older OpAmp-based preamplifiers that run from a 30 V single supply. The “half supply” reference, often called V_{BIAS} or V_{REF} , is a separate circuit that is filtered with a capacitor that bypasses AC signals. The resistance of the divider network should be as low as feasible but lower values will also load the supply harder, which means that the designer faces a compromise. The resistor that connects the bias point to the input practically defines the input impedance. Since the reference voltage is formed in a separate circuit it can be used for supplying a multiple amount of operational amplifier stages. As is, the reference is still affected by supply voltage variation, however, much less since it is filtered. The circuit can be further improved by replacing the bottom resistor with a Zener diode that ties the virtual earth potential to a constant value.

The most sophisticated technique is to buffer the divider network in order to form a low impedance reference. Since the divider network of the voltage reference is separated from the load by buffer it can use higher resistor values and therefore it can also be filtered with lower capacitance. The buffer has to be able to feed enough current for the load though; otherwise the circuit’s operation is disrupted. If the load current is high this might become a too tough job for a single operational amplifier

stage. In some applications the OpAmp stage is accompanied with an emitter follower buffer. Another very nice method is to use a low-power LM386 amplifier chip as a reference circuit with both inputs tied to ground as the chip internally biases its output to a potential that is between positive and negative rail (common). Downside is the restriction of maximum supply voltage to 12 V. This is no problem with applications using a 9 V battery, though. The performance of both circuits, b and c, will deteriorate on lower frequencies where the operation of the filter capacitor weakens.

The AC coupling of both input and output is mandatory since the signal's zero point has a DC voltage potential equal to potential between positive and negative rails. However, AC coupling is not needed when two stages that use the same virtual ground reference are connected together.

Operational amplifiers work in class-AB but they are usually biased very near class-A operation since the low currents involved allow this without considerable losses. Some operational amplifiers can be forced to work in class-A by inserting a current source inside the feedback loop at the amplifier's output. The stage has to be buffered with another operational amplifier that actually forms the loop. Simplest way to create the current source is to hook up a resistor from V_{EE} rail to output, more sophisticated method involves building an active current source to do the same thing. The required current depends on the device but it should be higher than the maximum current the device will provide in the concerned circuit. However, *the trick only works with operational amplifiers that have a complementary output stage*. Quite many devices use a quasi-complementary output instead, which doesn't lend itself to the aforementioned configuration.

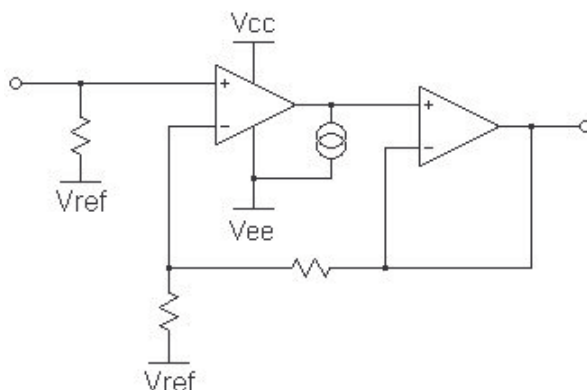


Figure 6.17 Forcing operational amplifier into class-A.

I feel it's unnecessary and unbeneficial to go to into fine details and try to present every possible OpAmp circuit variation here, as there are quite many of them. For that purpose there are greater design references, size of an entire book. For example, I highly recommend:

- Texas Instruments Design Guide SLOD006B, "Op Amps For Everyone – Design Reference, Revision B" by Ron Mancini
- Texas Instruments Application Report SBOA092A, "Handbook of Operational Amplifier Applications" by Bruce Carter and Thomas R. Brown.
- Op Amp Applications Handbook by Walter Jung, Analog Devices, Inc. Newnes, 2005.

The bulleted items are very comprehensive reference books – even for a skilled electronics engineer. On top of that, at the time of writing, the first two are available

in portable document format (PDF) for free downloading. The first one in the list is probably all that one will ever need.

6.4 Magnetic pickups and input stage

In short, a magnetic pickup is an inductor wrapped around a magnet or several of them. A vibration of ferromagnetic material near the pickup causes a flux in the magnetic field created by the pickup's coil. The flux, that is basically alternating current, is then processed by the guitar's internal electronics and sent to the amplifier. As a rule of thumb, the nominal output amplitude of most pickups is between 60 to 200 mV (single coils) and 200 to 600 mV (humbuckers and other "hot" pickups). However, with "hot" pickups the picking transient can be as high as 2 V_{PEAK}. Extremely hard picking (or strumming) of electric basses that have active pickups has reputedly produced peaks as high as 6 V. Pickups are resonant, which means that at the correct frequency the amplitude may be even significantly higher than the "rule of thumb" level. As you see, there is a lot of variation. Pickup manufacturers are usually pretty skimp on publishing accurate specs on their products (at least the kind that would satisfy a person knowledgeable of electronics). While telling what magnet material a pickup uses and how many wires comes out of it is fascinating, it is also quite meaningless, as these "specs" tell nothing about the pickup's electrical properties.

The frequency response of a guitar is far from being flat but due to simplicity of the pickup (and the surrounding electronics) itself it is not as complex product of different variables as, for example, the frequency response of a loudspeaker or dynamic microphone. One reason for this is that a pickup does not have any mechanical, moving parts. As known, a pickup is basically an inductor and a real-life inductor is never an ideal one; it has some series winding resistance and parallel winding capacitance. Some manufacturers state the winding resistance of their pickups but this is pretty meaningless value on its own. The winding capacitance is usually around 10 to 200 pF. The winding resistance for single coil pickups is normally about 6 to 8 kilo-ohms and with humbuckers around 8 to 22 kilo-ohms. The inductance is usually around 2 to 10 henries.

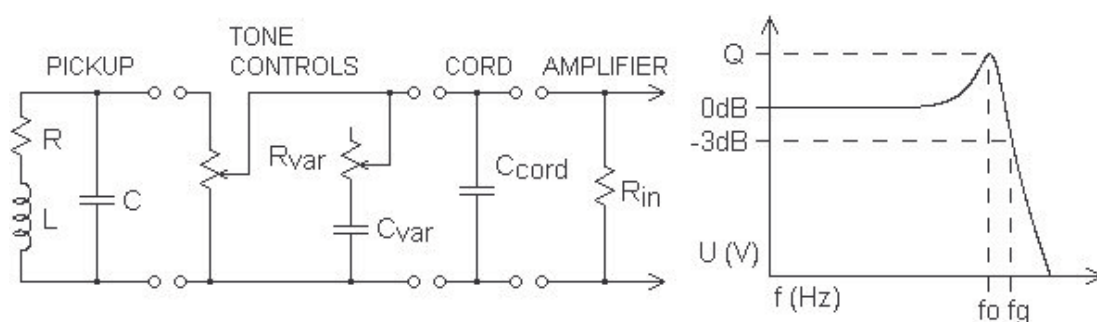


Figure 6.18 A typical pickup-input stage-circuit.

The inductance and the capacitance form a *parallel resonant circuit*, which's resonant frequency can be altered by the value of the capacitance. The winding capacitance C defines the natural resonant frequency of the pickup; it is typically at a kilohertz range. Lower capacitance values shift the resonant frequency f_0 higher while higher capacitance values shift the f_0 lower. Where the magic happens is when the pickup gets paired with a – usually higher value - parallel capacitor. Above f_G the resonance starts to roll off -12 dB per octave.

Any series resistance in the resonant circuit effectively dampens the amplitude of the resonance peak - this is why humbuckers can sound quite dark; their high winding resistance provides a lot of damping. Even as often the resonance is damped by other electronics: The resistance of volume potentiometer, the resistance of the tone control potentiometer (this is actually intentional) and, last but not least, a too low input impedance of the amplifier.

The resonant frequency is a very dominant characteristic amongst those that define pickup's tone so to retain it the amplifier should have very high input impedance. Bypassing the volume potentiometer's "input" lug to wiper with a "treble bleeder" capacitor is a common guitar modification that decreases the effect of high frequency damping caused by the series resistance over the potentiometer. Connecting the pickup straight into a high input Z active buffer that is followed by tone and volume controls has gained a lot of popularity as well. In almost every case when a guitar does not employ fancy, complex wiring or active electronics it has a basic low-pass filter as a tone control. Making the parallel capacitance value switchable between different capacitors can make the tone control far more versatile. Effectively this slides the resonant frequency along the audio band. The main topic of this book is not guitar electronics, though. However, discussing these basics helps one to understand what amplifiers should do in conjunction with the guitar. *This is why an amplifier is considered as an equally important part of the instrument!*

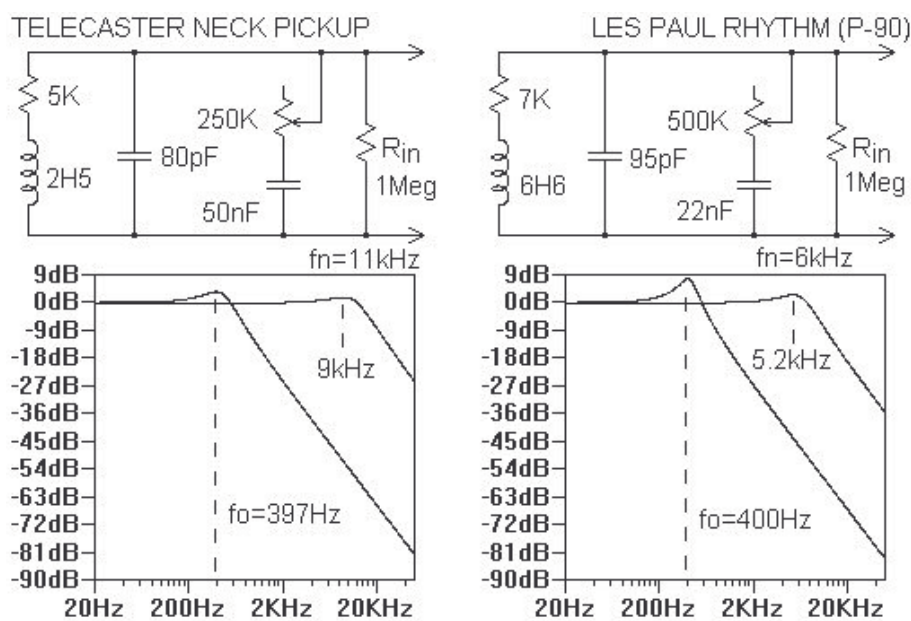


Figure 6.19 Frequency responses of Fender Telecaster's neck pickup and Gibson Les Paul's "rhythm" mode (P-90 pickup). Volume potentiometer is set to "10" and the frequency response is shown with tone potentiometer in maximum and minimum positions.

Effectively, one does not (usually) want to add any more capacitance in parallel with the pickup than the one in the tone control since the additional capacitance sums up to the existing, and shifts the resonant frequency lower causing attenuation of the high frequencies. This can make the guitar's tone very "dark" and "undefined". The guitar cord may have a capacitance between 100 pF to 1 nF depending on the type. The cord capacitance effectively causes no mentionable difference when it's in parallel with the tone control's capacitance of few nanofarads (tone control potentiometer set to zero). However, the situation is different when the tone control's capacitor has a high series resistance (tone control set to "10" or "bright") causing the pickup's winding capacitance to become dominant. Since the winding capacitance is in the order of

picofarads any parallel cord capacitance will cause a significant shift of the resonant frequency. Sometimes this is used as a trick to shape the tone: For example, Brian May and B.B King deliberately used long, coiled cords (very capacitive) to cut off high frequencies.

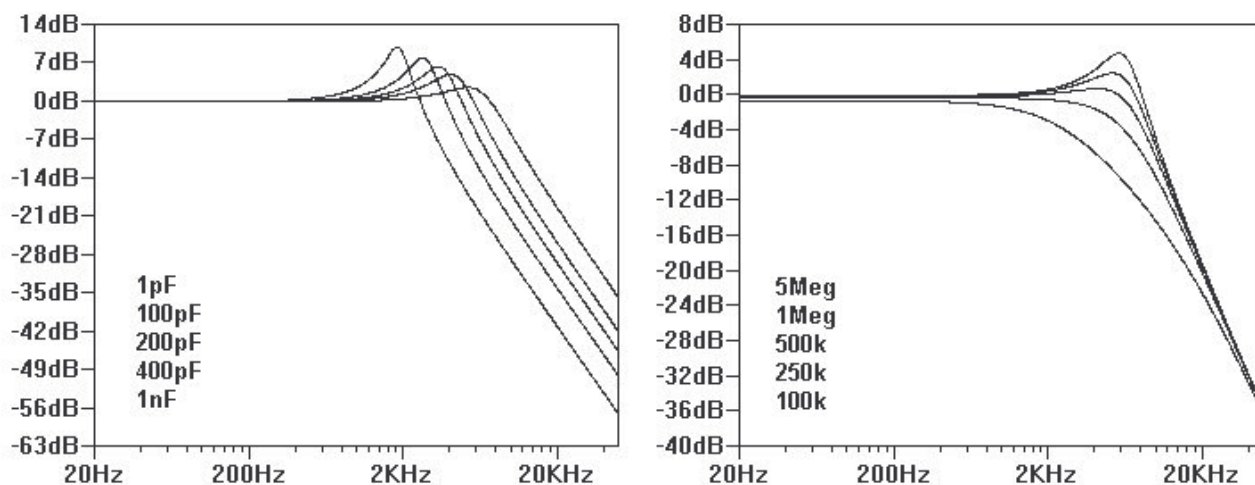


Figure 6.20 Left: Guitar cord and input stage capacitance shifts the resonant frequency lower and increases the amplitude of the resonance peak. Right: Lower input impedance attenuates high frequencies and the resonance peak.

Another factor, which may become detrimental, is the input impedance of the amplifier. Low input impedance will attenuate higher frequencies and dampen the resonance. The aforementioned facts are main reasons why people wish to plug their guitar into a high impedance buffer stage. Basically, if there is a chain of effects (or even just one) between the guitar and the amplifier the latter cannot be responsible of retaining the tone of the guitar. However, since shunt capacitance and low input impedance are even generally detrimental introducing them should be avoided.

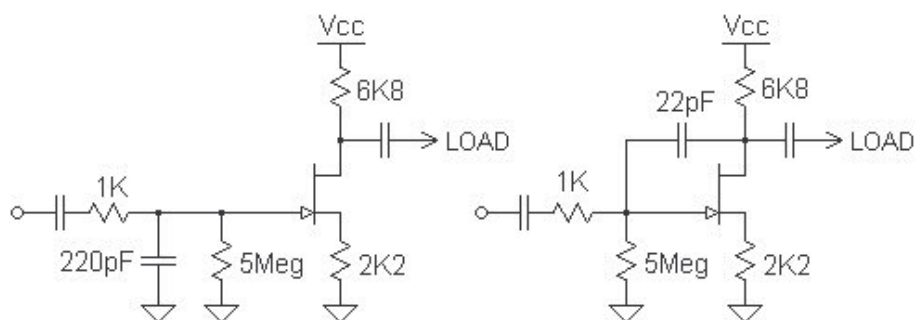


Figure 6.21 Two ways to construct an input stage with low pass filter.

Let's consider a practical design example: The circuits in the above schematic have an equal frequency response when amplifying a signal from a purely resistive source. The circuit at left is quite commonly used, however, in conjunction with *the impedance* of a guitar's pickup it shifts the resonant frequency 1 kHz lower thus resulting into a different tone than what you get from the circuit at right. No one can tell which one is better "tone-wise" because that is a highly subjective issue but the important point is that little details like this can in some cases become very important. While people are nowadays quite educated on the effects of insufficient input impedance, it is surprising how few designers actually consider what kind of an effect the input *capacitance* can create.

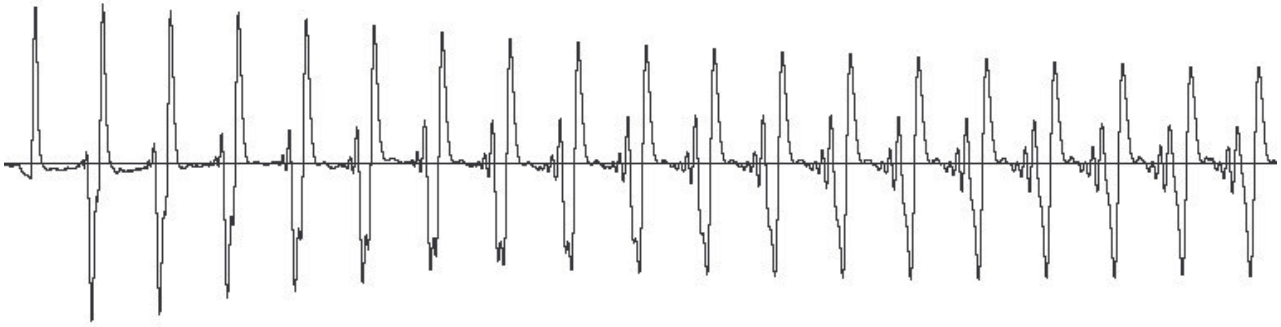


Figure 6.22 Time domain plot of electric guitar's output.

In conclusion, it is worthwhile to analyze the guitar signal in time domain. The Figure 6.22 shows a plot of open low-E string plucked gently with a medium pick. Before proceeding with the analysis, it must be thoroughly understood that different guitars have different “tones”, which are even further tailored by factors such as string gauge, pickup type, internal guitar electronics, pick type as well as picking strength and technique. Therefore the plot can only be taken as a plain generic example of guitar's signal. However, few universal conclusions can be drawn from the plot: The most obvious one is that guitar signal is far from sine wave and its structure is more that of a sound signal of a *horn*. In fact, if guitar signal did not have its characteristic envelope it would sound a lot like a tuba. The very same reason explains why some “fuzz” type distortion effects tend to make the guitar sound very much like a horn instrument: The Envelope is “smoothened” while the amount of “edgy” harmonics gets increased.

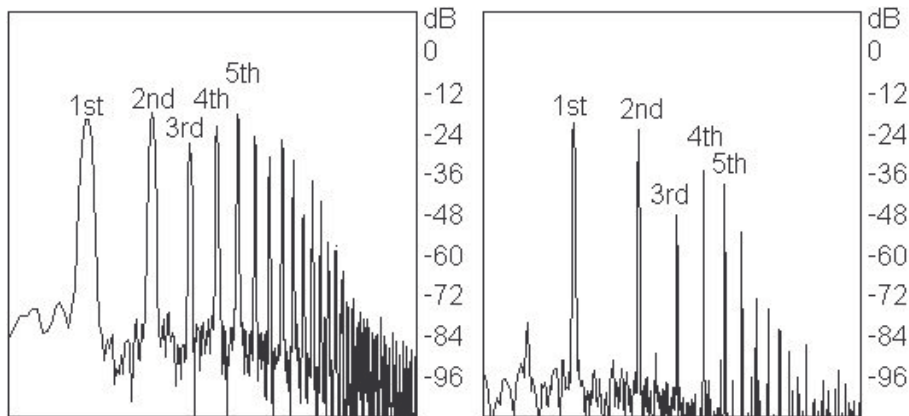


Figure 6.23 Frequency domain analysis of guitar signal. Left: Immediately after plucking the string and right: Decaying vibration of the string.

The majority of signal information exists in (“dull-edged saw tooth”-like) transients, while the “color” or “tone” of the instrument is defined in higher frequency harmonics that are significantly lower in amplitude. When string is plucked these harmonics are really strong but when the note begins to decay the harmonics decrease in strength and amplitude. Due to dominance of high amplitude signal content it is very expectable that gain stages *will be* overdriven. It is also easy to see why compressors and limiters are useful tools in preventing hard clipping and enhancing the guitar's characteristic tone. In many cases the signal from a guitar is also highly asymmetric.

The harmonic structure of the signal is quite surprising as well. When string is plucked and begins to vibrate the most dominant frequency *is not only the fundamental but several of its harmonics as well*. Even the 9th harmonic is only few decibels weaker in amplitude than the fundamental. When the vibration of the string decreases in amplitude the harmonic structure gradually changes: The strength of the

fundamental increases, yet second harmonic remains equally strong. Higher order harmonics significantly decrease in amplitude but the third, fourth and fifth still stay within few decibels of the fundamental.

6.5 Piezoelectric pickups

Piezoelectric pickups are based on crystal's ability to convert applied mechanical stress into voltage, thus a piezoelectric element has an ability to pickup vibration, which is commonly the vibration induced to the guitar's saddle. Since a piezoelectric element operates electromechanically it can pick up sound of non-metallic strings as well. Piezos also tend to work better with acoustic instruments since huge part of the "acoustic" sound is vibration of the body that forms the acoustic chamber. This kind of pickup is also called *contact microphone*.

The piezoelectric element is usually a thin disc size of a coin. The diameter of the element determines its resonant peak; using smaller element raises it. Also the mounting location tremendously alters the response. Piezoelectric elements commonly have very high output impedance and capacitive character caused by the metallic sensor plates of the element, therefore mismatch of input and output impedances usually causes low frequency attenuation. Usually piezoelectric pickups require input impedance as high as 10 Megaohms to even work ideally.

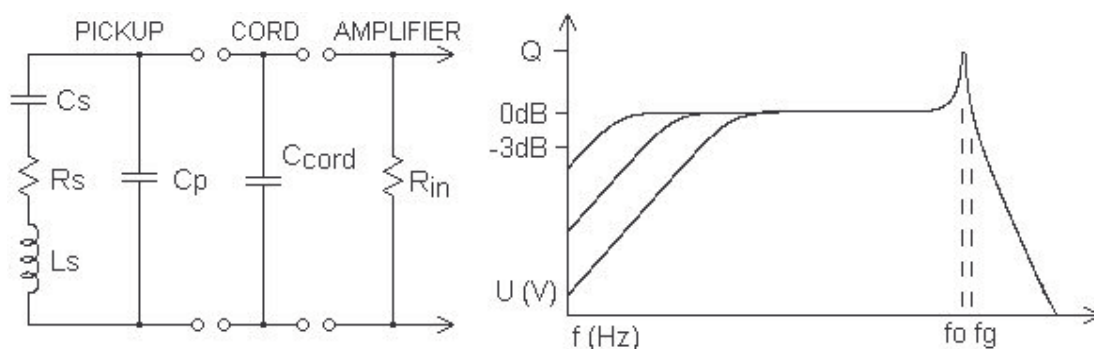


Figure 6.24 Electromechanical properties of piezoelectric element create a resonance peak. Lower input impedance attenuates lower frequencies.

The response of a piezoelectric pickup is also far from linear, however, the resonance peak it exhibits is typically located far above the audible region and therefore attenuated by the following electronic circuitry. The characteristics of a piezoelectric element can be modeled with an electromechanical model where mechanical mass, stiffness and damping of the crystal are represented by an inductor, capacitor and resistor. The model is similar to that of a conventional crystal. C_P typically defines the impedance, unless the circuit is in resonance. Any increase in parallel capacitance to C_P (like, for example, that introduced by the guitar cord) will attenuate the whole signal and shift the resonant frequency lower. Lowering the value of R_{IN} causes earlier low-frequency roll-off (we knew this already) but does not affect the resonance peak notably. As a conclusion, on an audio band a piezoelectric pickup can have a very linear frequency response - assuming it is paired up with a high input impedance buffer. A piezoelectric pickup plugged straight into a guitar amplifier with a typical ~ 1 Megaohm input impedance would probably sound very "thin".

6.6 The effects of frequency response

Unlike HiFi and PA amplifiers, guitar amplifiers do not *often* aim for a flat frequency response. Even the very first guitar amplifiers with crude tone control circuits – if any – had a very odd response due to high output impedance. The frequency response of an amplifier is often measured in dB scale of voltage instead of dB scale of power. The two scales are interchangeable, though: A 6 dB increase in voltage equals a 6 dB increase in power. However, in electrical units a 6 dB increase means a voltage gain of 1.995 and power gain of 3.981. The compatibility of dB scales is essential since it gives a direct estimation how much power will be fed to the speaker system. The speaker system converts the power into acoustic energy (sound pressure) usually with a very poor efficiency so the *output power figure does not actually tell us anything about the actual loudness* (see chapter about speakers). The dB scale will, however, tell something about relative loudness levels between peaks and bottoms of the frequency response graph.

Table 6.1 Decibels and perceived loudness

dB	x Loudness	x U	x P	dB	x Loudness	x U
33.2	10	45.71	2100	-33.2	0.1	0.0219
20	4	10	100	-20	0.25	0.1
10	2	3.162	10	-10	0.5	0.316
6	1.5	1.995	3.98	-6	0.66	0.501
1	1.07	1.122	1.26	-1	0.93	0.891

Attached table shows few measures of perceived loudness in relation to dB scale. It also shows the power and voltage gain that is required to produce the gain stated in decibels. The table is just a mere approximation and follows a linear behavior, which human hearing is not. A more accurate presentation of a natural, unlinear perception of loudness is the Fletcher-Munson curve – or its improved version the Robinson-Dadson curve (refer to Figure 1.10). It relates the sensed loudness of different frequencies to the intensity of the sound. Also, let's emphasize the fact that output power *is not* a direct measure of amplifier's loudness: Loudness depends on the speaker system's efficiency of turning electric power into acoustic energy. The difference between attaining a twofold loudness can be either increasing power tenfold or using a speaker that is 10 dB more efficient. However, the table is useful when one ponders the effects of the tone control circuit's frequency response.

Similarly to nearly every instrument, a guitar has a unique harmonic timbre. When a string is plucked it vibrates on a certain fundamental frequency and with a multitude of its harmonics. These give the guitar its characteristic tone. Further on, the pickup colorizes the tone by attenuating frequencies that are higher than its resonant frequency, enhancing frequencies in the range of its resonance peak and keeping the frequencies below the resonant frequency unaltered. The timbre of guitar string's vibration picked up by a magnetic pickup is very different from the timbre of guitar string's vibration, which is amplified by an acoustic chamber along with vibration of the soundboard. Although the overall timbre of the tone can infallibly be identifiable as belonging to a guitar instrument the magnetically picked vibration has a strong middle frequency emphasis.

Since the 1960's guitar amplifiers have mostly tried to compensate the aforementioned emphasis by scooping the middle frequencies, sometimes pretty heavily. A Fender advertisement from 1972 states: "*In Fender amps the low and high*

ends are exaggerated and the mid-range falls down a little bit in accent. ... We do it to compensate for the irregularities of a typical guitar signal. The result is perfectly rounded response curve that lets your guitar sound the way it's supposed to sound". For example, a Fender Twin Reverb-Amp scoops the frequencies around 440 Hz about -17 dB (when all controls are in the "middle" position). The difference in perceived loudness between peak and bottom points is approximately threefold.

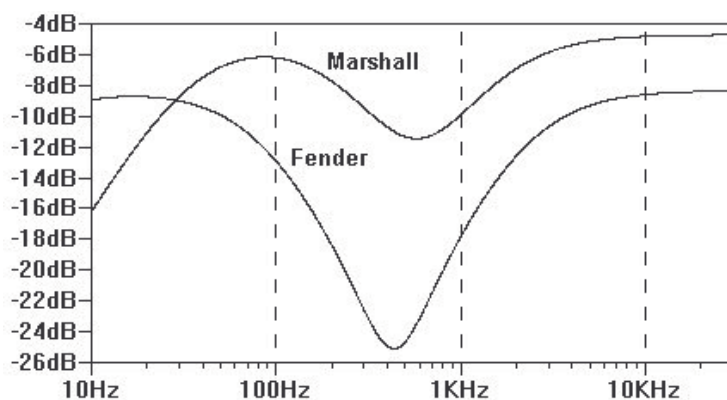


Figure 6.25 Frequency responses of tone controls in "Black Face" Fender Twin Reverb-Amp AB763 and Marshall 1959 Super Lead amplifiers. All potentiometers are set on half position.

Note: Most guitar amplifiers that are manufactured after the early 70's have a fairly linear frequency response when the "middle" control is turned to fully clockwise position and "bass" and "treble" controls are turned to fully counter-clockwise positions.

Since guitar amplifiers usually alter the frequency response so that a signal from a magnetic pickup will have a pleasant tone it usually causes that linear signal sources (such as piezoelectric pickups or conventionally miked acoustic guitars) may sound strange. Having a more linear frequency response is not "bad" though: In fact, the frequency response of a Marshall amplifier is a lot more linear than that of Fender's, although it also exhibits a similar scoop (about -6 dB at 580 Hz). Some guitar amplifiers have a totally flat frequency response (aside normal high and low pass filtering). Most amplifiers for acoustic instruments amplifiers have linear response but also employ a steep filter that can be used in attenuating feedback. The differences in frequency responses create the tones we familiarize amplifiers with.

Perhaps the most important purpose of tone controls is to shape the guitar tone so that it blends in nicely with other instruments in the band. Sometimes this requires using a considerably "thinner" tone. Basically, the fewer instruments in the band, the fuller the guitar tone can be. If the band's instruments already consist of a synthesizer, bass, drums and two guitars (plus a vocalist), the guitar tone of each guitar should probably be limited to a very narrow bandwidth so that they won't compete with each other and with the other instruments, as this easily drowns out the vocals or just plainly "muffles" the overall sound. You hear this happening often when bands play club gigs without a professional soundman. Talented musicians know that considering bandwidths where individual instruments should operate already starts from the song arrangements.

Another important function of tone controls is to set up the amplifier to operate correctly with varying room acoustics and loudness levels. Anyone familiar with Fletcher-Munson curve knows that at higher loudness levels the low and high frequencies need less emphasize. Even more: Amplifying bass frequencies consumes

about 65 % of the total power and therefore attenuating them (even a little) can greatly reduce distortion, which brings clearness to the tone. The room acoustics can create a huge difference in tone as well: Basically the room creates sound reflections that resonate on frequencies that are proportional to wavelength and room's dimensions. Even equal sized rooms will sound completely different if the walls of the other room are damped with material that absorbs sound and artificially extends the room volume. On top of that is the factor of speaker cabinet placement: Is it in a corner, on the floor or on a stand - and is it perhaps tilted. The reflections created by amplifier placement can either enhance or attenuate certain frequencies – this is a well-known trick as well. An amplifier should be equipped with tone controls that can help the musician to cope with these issues. This is especially important if the amp does double-duty as a monitor (which often is the case). To add to the confusion, due to all created reflections and resonances, the tone of a guitar – and especially bass - amplifier system is completely different few meters away from the speaker cabinet. Now you have to think whether you want to dial a good tone for yourself or for the audience. Luckily these days the guitar cabinets are often close miked and their sound is reproduced through a much more linear PA system so you can, in most cases, somewhat ignore these effects.

It should be realized that the design process of a tone control circuit has to be based on the function of the amplifier: Is it intended for low-power bedroom use on minimal loudness levels (typically small spaces with passable acoustic properties), practice sessions and occasional gigging with a band (both small and medium sized rooms with widely varying and usually horrible acoustic properties – also tends to require at least 40 to 60 watts of power), studio use (too many aspects to comment in short – mostly depends on facilities and on how the musician likes to work in a studio) or for use as a high-power workhorse for stadium concerts (very large halls and maybe even outdoors but probably miked through PA anyway). The higher the requirements, the less likely an amplifier with only crude tone controls can fulfill them.

6.7 Tone controls

6.7.1 Quick introduction to tone control terminology

A shelving equalizer is a tone control that affects all frequencies above or below the control point by either attenuating or amplifying. Basically controllable high- and low pass filters are shelving equalizers. The basic forms of Baxandall and Bingham tone controls consist of two shelving equalizers (one for highs and one for lows) that use an equal control point, which is typically located somewhere in the midrange area.

A graphic equalizer splits the frequency into multiple bands that can be controlled separately. Usually the control for one of the bands interacts heavily with the frequencies of the adjacent bands making the circuit a bit difficult to adjust. Basic FMV circuits are graphic equalizers with a very minimal amount of bands. Even a two-knob FMV circuit can be considered as a graphic equalizer since it just has a control for one of the bands in a fixed position. An advanced form of Baxandall circuit also turns into a graphic equalizer. If the rotating potentiometers would be replaced with sliders the “graphic” character of the circuit would become clearer. The sliders also offer a more visual indication of the frequency response.

A parametric equalizer uses three controls to set frequency, bandwidth of the control and the amount of gain/attenuation. The most sophisticated systems combine a graphic

equalizer with a parametric equalizer. This kind of circuit provides an ultimate control over frequency response but tends to be very complex and requires extensive tweaking. A quite common form of the idea, which is more suitable for guitar amplifiers as well, is a circuit that uses three-band graphic equalizer with a selectable middle frequency point and no bandwidth control.

6.7.2 Active tone controls

How an active tone control differs from a passive one is that it usually introduces some gain. While passive tone control circuits are based on the idea of enhancing certain frequencies by *attenuating* the others, active tone controls enhance frequencies by *increasing their gain*. Since there is less attenuation (also known as “insertion losses”) involved in the process, the signal to noise ratio is notably improved. The idea of using negative feedback to control frequency response was, if not invented at least popularized, by Peter Baxandall and presented in Wireless World article “Negative Feedback Tone Control, Independent Variation of Bass and Treble Without Switches” in 1952. Simple pass filters within the feedback loop can be considered as active filters but the circuits can be made even more complex than that. Basically, it is possible to convert almost any passive tone control stage to an active one.

The conversion process is fairly straightforward: The common node is connected to the output of an inverting amplifier and the output of the tone control circuit feeds the input. The only clause is retaining a solid DC feedback path between input and output. In practice, the output impedance of the previous stage has to be low. In most cases this means that the stage feeding the tone control is a buffer. Without resistors R1 and R2 the active stage has no gain since it is technically a basic buffer. When R1 and R2 are added up they form a voltage divider that forces the output to assume a higher voltage potential: The circuit now has gain but retains the frequency response. It is also inverting.

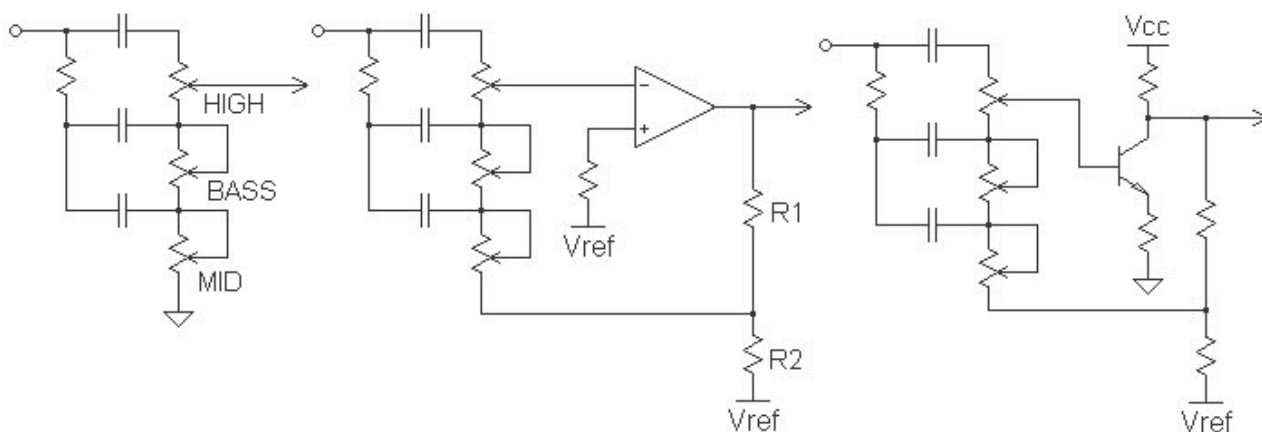


Figure 6.26 Converting passive “FMV” tone control circuit to active.

6.7.3 Scaling impedance and frequency

Frequency scaling means shifting control points up or down in the audio spectrum. Most tone control topology variations usually show only small modifications in component values and the overall ratio between different components remains about the same. However, often there also arises a need to scale impedances up or down, for example, to change 1-Megaohm potentiometers to 100 kilo-ohm ones, yet still retaining an equal frequency response.

Both scaling processes start from calculating the scaling factor as follows:

$$FSF(\text{frequency}) = f_{NEW} / f_{OLD} \quad \text{(Equation 69)}$$

$$ZSF(\text{impedance}) = Z_{NEW} / Z_{OLD} \quad \text{(Equation 70)}$$

Once the scaling factor is known, the new component values for a frequency scaled circuit can be calculated using the following equations:

$$R_{NEW} = R_{OLD} \quad \text{(Equation 71)}$$

$$L_{NEW} = L_{OLD} / FSF \quad \text{(Equation 72)}$$

$$C_{NEW} = C_{OLD} / FSF \quad \text{(Equation 73)}$$

For impedance scaling the equations are:

$$R_{OLD} = ZSF \times R_{NEW} \quad \text{(Equation 74)}$$

$$L_{NEW} = L_{OLD} \times ZSF \quad \text{(Equation 75)}$$

$$C_{NEW} = C_{OLD} / ZSF \quad \text{(Equation 76)}$$

Scaling functions become very useful when you wish to use a certain tone control circuit with different source and load impedances.

6.7.4 "Pass"-filters and other crude tone controls

Considering requirements for "compensating" room acoustics, pickup linearity, and loudness versus frequency, simple high- and low-pass filters just aren't ample enough to act as tone controls in any task that is more demanding than low-power bedroom practice amplifier. The controls might work great in half-watt applications aimed for private jamming and practice (so does lack of them) - or perhaps in a preamplifier stage of an effects loop, reverb or another special effect circuit. A crude high pass filter might also suit the input stage when acting as a circuit that selects a proper -3 dB point for 6 and 7-string guitars. For a serious need of compensating room acoustics or shaping personal tones they are practically worthless.

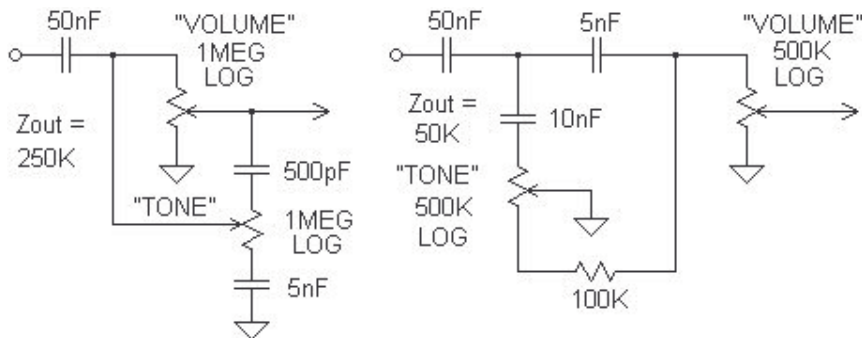


Figure 6.27 Left: "Tweed tone controls" and right: Tone controls from 18 W Marshall.

In guitar amplifier world there exists a concept known as "tweed tone controls". This is actually somewhat a poor name: Before the invention of plastic covering materials,

such as tolex, tweed was a very common cloth used in covering items like suitcases, shoeboxes and guitar amplifiers. So, practically all guitar amplifiers of the concerned era were “tweed”. Due to such variation it is difficult to pinpoint a literal example of a “tweed tone control” circuit. Essentially the term is used when referring to tone controls used in some of the earliest Fender guitar amplifiers and from this variation it is quite easy to pick an adequate example: Aside slight variations in the source impedance, *exactly* the same tone control circuit that is depicted left in Figure 6.27 was used in Fender Princeton 5C2 and 5F2-A, Deluxe 5E3 and 5D3, Super 5D4, Pro-Amp 5C5, 5D5 and 5E5, Tremolux 5E9-A and 5G9, Harward 5F10 and Vibrolux models 5E11 and 5F11. Similar tone stack with slightly different component values or with some series resistance and capacitance before the volume control was used in the rest of the “tweed” Fender models. The most advanced circuits separated the tone- and volume controls from each other with a gain stage: This was good practice, as we’ll soon learn.

Hooking the tone control in conjunction with the volume control was a dreadful mistake since the two tremendously interact with each other. While in guitar world nearly anything including words “tweed” or “vintage” seems to enjoy somewhat a glorified status the “tweed” tone controls do not: Most people abhor the inherent interactivity of the circuit and dialing settings for good tone in alternating environments can be torture - especially because the single tone control is quite limiting. The later “tweed era” tone control of Fender amplifiers was almost similar to “James/Volkoff” circuit (discussed later) and was eventually replaced by the “FMV” circuit of the “Black Face” amplifiers. Personally, I feel that a single tone control is in most cases inadequate but if I had to use one I would prefer the less-interactive “18 W Marshall” -type circuit. *When potentiometers are set to center position the frequency response of both tonestacks is rather flat throughout the fundamental bandwidth of guitar. As opposed to “scooped” tones of modern amplifiers this creates an “older, vintage vibe” – perfect for rockabilly and alike.*

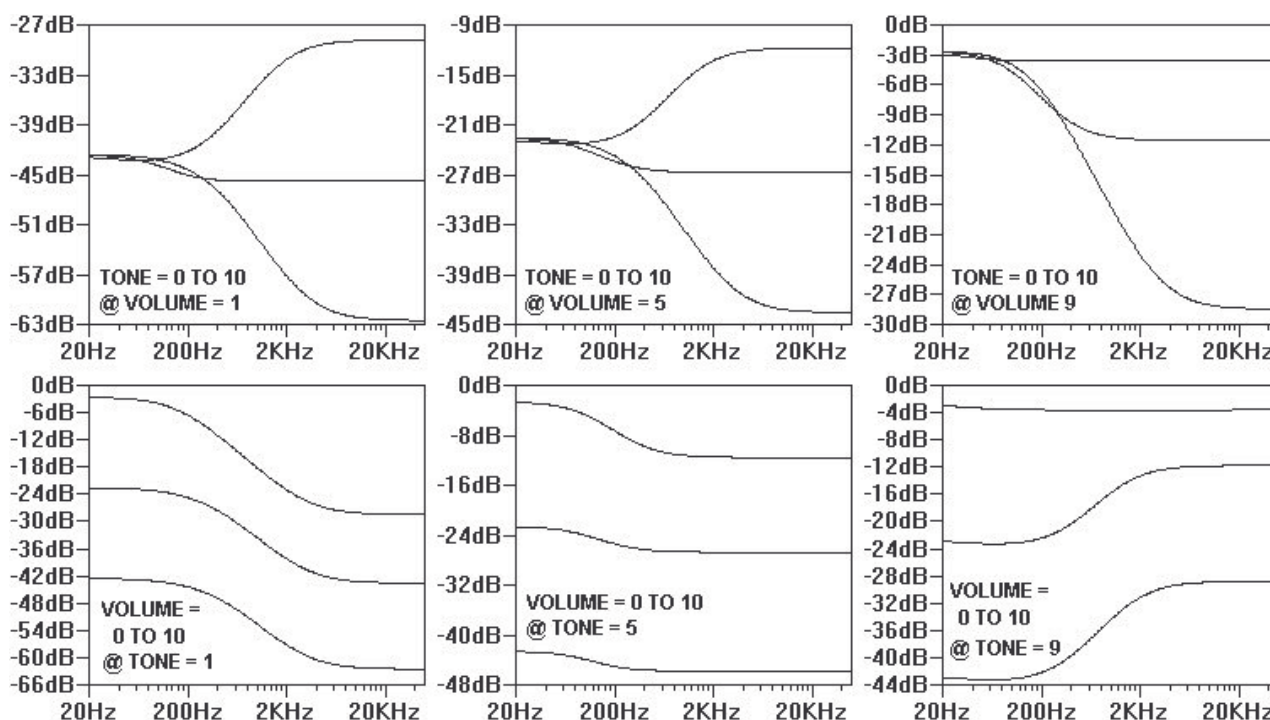


Figure 6.28 The interaction of volume and tone controls in a “tweed” tone control circuit.

The plots in Figure 6.28 show the interaction of controls in the depicted tweed circuit. First notice the inherent insertion losses of the passive circuitry: Compared to many other passive controls this is a positive feature since the losses at maximum volume and tone control set to middle are only few decibels. The rest is not so positive: At volume levels between “1” and “5” the tone control works quite well providing either quite linear response or some bass or treble “boost”. Notice that the bass “boost” is created at the expense of cutting highs quite extensively. When the volume level increases the circuit begins to cut high frequencies more and more and initially the “treble boosting” quality is completely lost. This effect is most prominently shown in plots that show alteration in volume control setting while tone control has a fixed position. When the tone is set “high” the overall alteration of volume is less linear than in any other case. The inherent loss of “highs” (or bass boosting) at higher volume settings is prominent even when the tone control is set to middle position.

6.7.5 Bingham

“Tilting” tone control circuit, named after its inventor John Bingham, is mainly intended for use in HiFi amplifiers. It combines high and low pass filters that operate around a center frequency, which is typically located at middle band. The circuit is good for passable compensation of room acoustics but cannot sufficiently compensate pickup’s natural mid-range oriented tone. Frankly put, tilting control is not very versatile circuit if one wishes to tweak out different tones. The circuit might work with low-power practice amplifiers or with acoustic instrument amplification.

The real benefit of a Bingham circuit is its simplicity: The circuit allows balancing between high and low frequencies with just one potentiometer – for low-power private jamming amplifiers (or headphone/line level out circuits) something like this might actually be all that is needed. Turning the potentiometer either boosts up highs while attenuating lows or vice versa. With the potentiometer in center position the circuit has no effect to frequency response. The effects of the control are very easy to visualize from the classic “tilt” arrangement used in Quad preamplifiers.

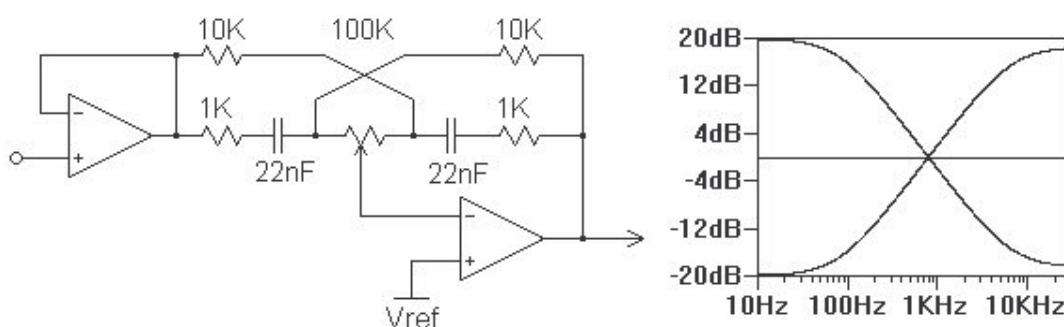


Figure 6.29 Bingham’s tone control circuit provides a balancing control for high and low frequencies that works with a single potentiometer.

As a historical side note, it maybe should be mentioned that the original circuit from John Bingham was quite different than the one used in Quad amplifiers (refer to Figure 6.31). While the operating principle and frequency response of the actual Bingham circuit are alike, the circuit itself is not: Basically it consists of parallel active high- and low-pass filters that are “mixed” together with a potentiometer, thus achieving the desired amount of “tilt”. The upper low-pass filter circuit is pretty straightforward; mixing its (inverted) output with the opposite phase input signal of the lower inverting amplifier cunningly forms a hi-pass filter. Consequently the “crossing” frequency is fully adjustable with a simple change of capacitor’s value.

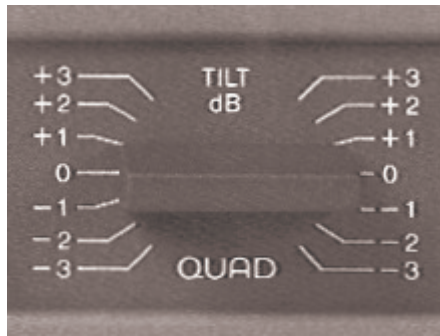


Figure 6.30 "Tilt" control, as used in Quad 34 HiFi preamplifier

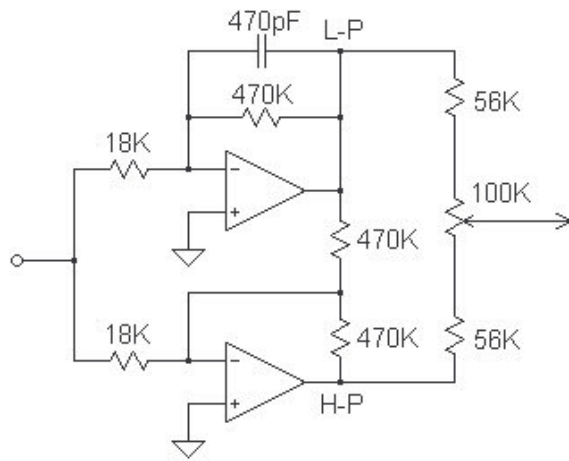


Figure 6.31 The original circuit from John Bingham

6.7.6 Baxandall circuit and its variations

Duncan Munro's Tonestack Calculator software, which among guitar amplifier modifiers is quite famous, shows a tone control circuit, which goes by the name of "James". The depicted, passive circuit is often mistakenly called "Baxandall" tone control. One should note that the actual and *true* Baxandall circuit is *active* while the James circuit is *passive*. A little peek in history might be in order here: As far as I know, Michael Volkoff designed the original, passive circuit (which is actually a combination of two generic filter circuits) and it dates back to 1935. Later in 1949 E. J. James introduced the circuit in a *Wireless World* article. In 1952 Peter Baxandall's article, that described how to convert the circuit into an *active* form, was published in *Wireless World* magazine. The latter invention was quite revolutionary because it utilized an already well-known tone control circuit in a way that was never done before. Although the configuration (in both passive and active forms) became a "standard" in HiFi amplifiers the invention was left almost uncredited for decades. During the time the circuit has seen numerous modifications.

The "passive Baxandall" – or preferably James/Volkoff - circuit generally has *more insertion losses than a standard "FMV" circuit* but less interaction between the controls. Increasing the value of the resistor that separates the bass and treble controls can furthermore minimize the amount of interaction. The sweep plot of controls is similar to one of Bingham's - except that high- or low frequency response remains nearly flat while adjusting the opposite control. The circuit doesn't "tilt". The basic form lacks a mid-range control but one can be included pretty easily. The most "crude" way to do this is to make the resistor marked "middle" variable. This is indeed a *very crude* way since the concerned implementation basically acts pretty much like a simple volume control. Yet, it has been used in various amplifiers. There is an alternative and better way to implement a midrange control into the circuit. It is more complex and introduced later. One quite common choice is just to use an external mid-range control circuit like, for example, the one presented in Figure 6.38. *On contrary to false belief, the basic James/Baxandall circuit does not have a very practical way to boost up middle frequencies, other than setting bass and treble controls to zero.* This will create a frequency response graph resembling a hill but the "boost" will also affect frequencies on a pretty wide frequency band.

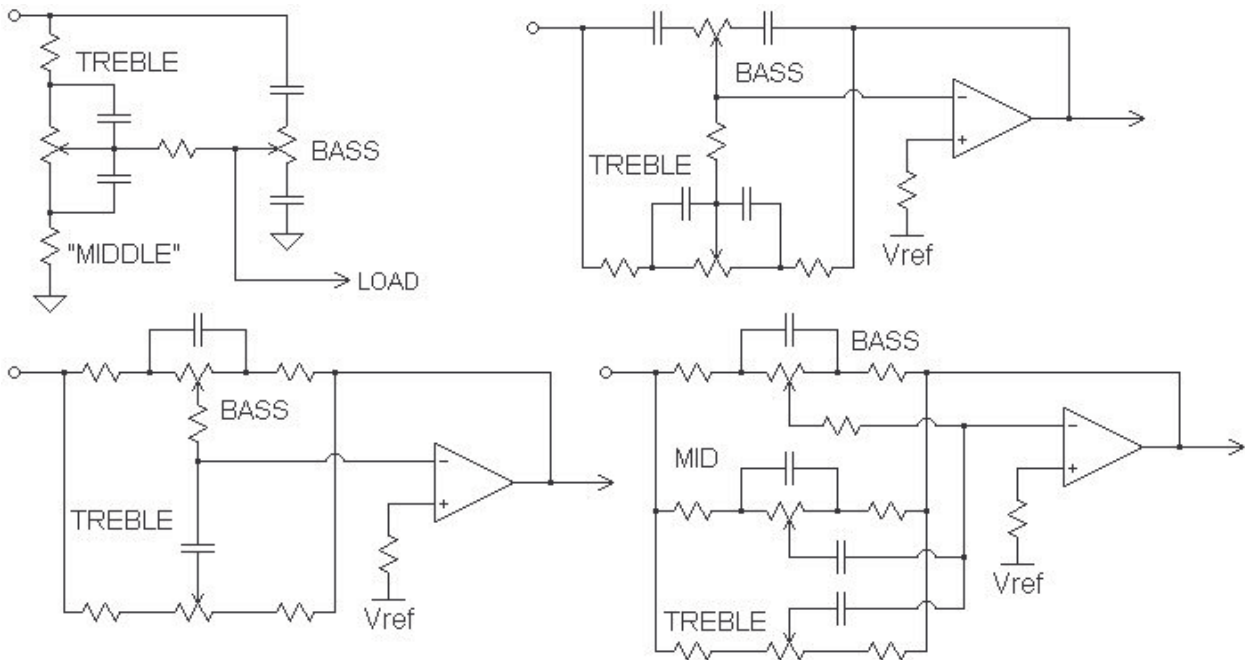


Figure 6.32 The "Baxandall" tone control has many forms; top right is closest to the original Baxandall circuit, which was active. Top left is closest to the passive circuit introduced by E.J. James.

The James/Baxandall circuit is quite common although not as "mainstream" as the FMV control. It has been successfully used at least in some Ampeg, Fender, Gibson, Orange, Magnatone and Stadel amplifiers. Accompanied with a mid-range control the circuit was used in active configuration in the famous Lab Series amplifier models L5, L7, L9 and L11. As a generalization, the common opinion tends to be that a Baxandall control works less ideally in "high gain" circuits. Generally tone controls of such circuits should introduce some mid-range scoop, like the "FMV" circuit does. This scoop can be dialed in but while in "scoop mode" the controls usually have only very little range to operate. If the circuit is designed to provide a scoop with controls in center position the effect of high frequency control is usually weakened. The basic circuit was designed to have a *flat* frequency when all controls are in middle position and it is not very adaptable to other configurations.

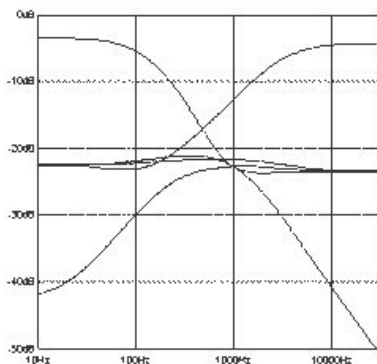


Figure 6.33 Control sweep plot of a simple, two-control Baxandall circuit.

Peavey amplifiers use an embodiment of the circuit where component order is rearranged. If you ever consider using the Baxandall circuit then at least one midrange control should be added to it - otherwise the circuit is more suitable for amplifiers that are intended to be used by jazz musicians playing acoustic or semi-acoustic instruments. Even in that purpose having one additional tone control does no harm.

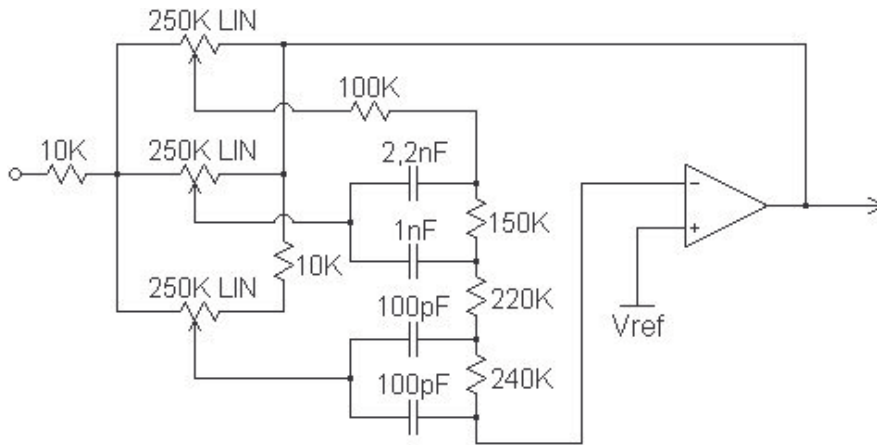


Figure 6.34 Variation of the three-band Baxandall circuit (Peavey Amplifiers).

6.7.7 “FMV” – tonestack: Fender, Marshall and Vox

FMV is a common abbreviation for the tone control circuit topology used in most Fender, Marshall, and Vox amplifiers – plus countless of others. Basically the name is very misleading and bad (as usual) since all mentioned companies, even including Marshall, have used plenty of other tone control circuit types as well. The most famous Fender amplifiers *did not* use this circuit and Vox copied the tonestack to its “Top Boost” circuit from Gibson. Nonetheless, nowadays the topology is best known as Fender/Marshall/Vox (or just “FMV”) circuit.

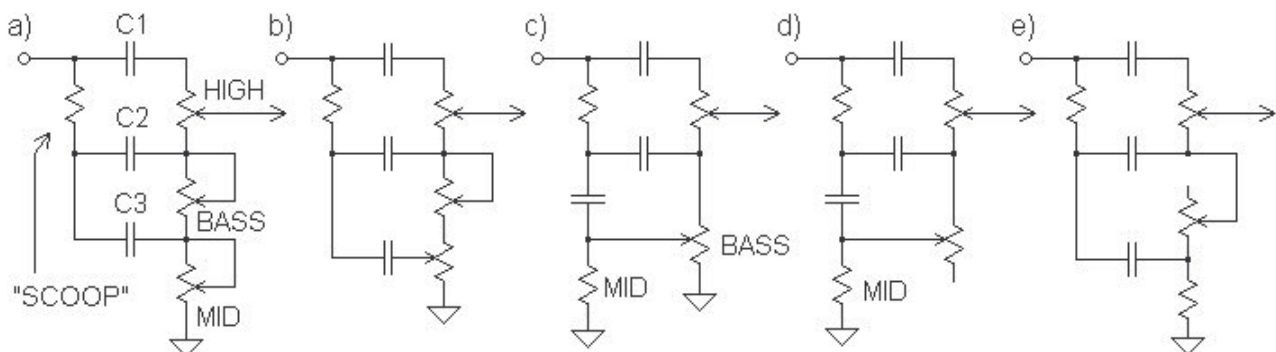


Figure 6.35 Variants of “FMV” tone control circuit: a) Fender, b) Marshall, c) Vox, d) Gibson and e) Magnatone. The variations in using fixed resistance values, switchable amount of capacitance or resistance and different component values are endless.

Calculating impedance of the FMV circuit is very complex process so it is advisable to either breadboard the design or to use some computer software (e.g. SPICE, Duncan Munro’s Tone Stack Calculator or Dave Cigna’s Cigna) to do the math and simulate the design. At the time of writing this, Duncan Munro’s otherwise excellent software obviously contains some errors in the algorithm deriving frequency response of Vox-type stacks (even aside the case of incorrect source and output impedances) and I’m afraid I have to say that on that part it should not be trusted as a design tool. The following list should provide some help in defining some basic rules that guide the design process.

The following list shows typical component values and describes the effect of shifting them up or down:

C1	100pF - 1nF	Higher values enhance high frequencies and shift f_{SCOOP} lower	Lower values attenuate high frequencies and increase scoop's bandwidth
C2	10 - 200nF	Enhances bass, narrows the scoop	Attenuates low frequencies
C3	10 - 100nF	Shifts f_{SCOOP} lower, attenuates low frequencies, increases scoop's bandwidth	Enhances middle frequencies, too high values remove the scoop
R_{SCOOP}	10 - 500k Ω	Attenuates low frequencies, shifts f_{SCOOP} down	Attenuates high frequencies, shifts f_{SCOOP} up
R_{HIGH}	10k Ω - 1M Ω	Attenuates high frequencies, shifts f_{SCOOP} up, increases scoop's bandwidth	Shifts f_{SCOOP} up, weakens the scoop effect
R_{BASS}	10k Ω - 1M Ω	Enhances low frequencies	Attenuates low frequencies
R_{MID}	10k - 1M Ω	Decreases insertion losses, weakens the scoop effect	Increases insertion losses, up to a certain value enhances the scoop effect
R_{SOURCE}	-	Increases insertion losses, attenuates high frequencies	Less insertion losses
R_{LOAD}	-	Less insertion losses, attenuates high frequencies	More insertion losses, enhances high frequencies

The controls in FMV circuit are very interactive: In the Fender-type circuit the bass control works pretty well but the mid-range control also affects low and high frequencies. The high frequency control is the greatest caveat: Adjustment alters the notch or “scoop” frequency f_{SCOOP} and also interacts with low frequencies. If the middle control is turned to “zero” the stack can very effectively cut frequencies above or below the notch point. Unfortunately the effect of f_{SCOOP} shifting is the most severe in this configuration. The special “feature” of the stack is that when all controls are on “zero”, no signal is passing through.

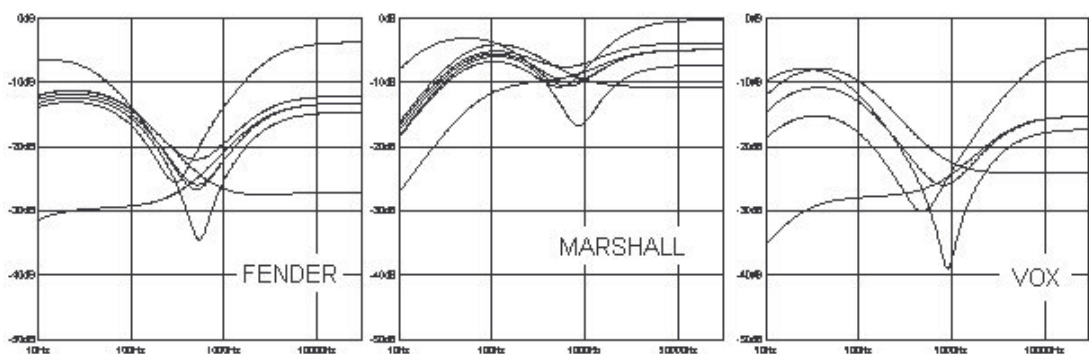


Figure 6.36 Control sweeps of three basic configurations.

Because the mid-range potentiometer in the Marshall-type circuit is configured differently the circuit won't cut out the signal with all controls on zero, instead it offers some mid-range emphasis – similarly to basic James/Baxandall circuit. With the exception of the aforementioned detail the Marshall circuit operates as interactively as the Fender one.

The basic Vox circuit lacks a mid-range control, however, it can be added by substituting the fixed value “MID” resistor with a potentiometer. Adjusting the bass causes a huge dip in the mid-range frequencies at the end of the potentiometer range. If coupled with a mid-range control the circuit probably stands out from the bunch as the one being most difficult to adjust.

The greatest drawback of the FMV circuit is the incapability for any practical mid-range boosting. When bass and treble controls of Fender and Vox circuits are at “zero” and middle is at “ten” the frequency response is nearly *flat*. The Marshall circuit can slightly “boost” mid-range frequencies if all tone controls are set to “zero” but the insertion losses are high and the bandwidth of the enhanced frequencies is very large. Frankly put, a typical “FMV” tonestack tone control circuit is pretty bad for its job – however, most guitarists have already gotten heavily accustomed to its simplicity, tone and interaction of controls. For them anything else can be quite impractical or strange sounding.

The FMV tonestack introduces a considerable insertion loss. This “feature” was (and still is) exploited in a novel boost configuration that simply controls whether the tonestack circuit is ground-referenced or not. If the ground reference is removed then only the “treble” control is left working. However, insertion losses are reduced tremendously. In practice, implementation of this control is very simple to any other FM-stack type than the one used in Vox amplifiers; simply make or break the ground path from the “middle” resistor/potentiometer.

6.7.8 “Big Muff” tone control and other ideas

In the long run sticking with the basic “FMV” tonestack is pretty boring from the engineering point of view. The circuit has been used in 99 % of most modern amplifiers with just little modification to component values. Sometimes it’s better to think outside the box and consider other solutions – even when you would choose to abandon them at the end and stick with a conventional solution. This is one of the reasons why I have always admired Gibson corporation: While many other companies were building nice but *technically boring* circuits the company was experimenting with “strange” tubes, unconventional tone control topologies and circuits, addition of negative feedback and even with the use of primitive vacuum tube voltage regulators. For example, Vox owns a lot to Gibson: The tone control circuit of the famous AC-30/6 was copied from the *Gibson GA-77 “Vanguard”* & *GA-70 “Country and Western”* models - all the way down to the schematic’s drawing error that showed one end of the “bass” control potentiometer grounded; it should have been left floating. The error made little difference for the functionality of the circuit and changed the tone only slightly– mainly by widening the range of controls.

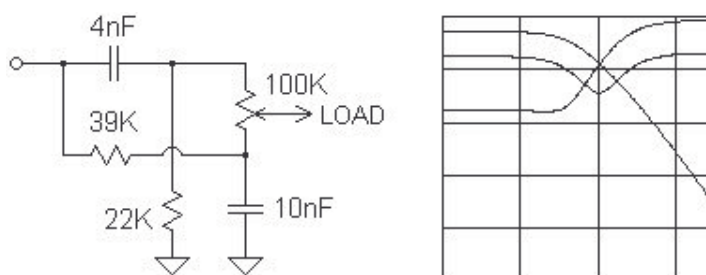


Figure 6.37 A “Big Muff tone control”.

A very simple and effective tone control has become famous from its application in Electro Harmonix' Big Muff distortion pedal. Essentially the circuit is just a combination of high- and low-pass filters that are mixed together by a single potentiometer. The cutoff points are (usually) designed so that their interweaving effect introduces a middle frequency scoop when the mixer control is set to middle position. The circuit is shown with its "typical" values that result into scoop frequency of 1 kHz. The potentiometer is linear type. Effective variation is making the cutoff frequencies of the RC circuits variable (replace resistors with potentiometers) and fitting series resistors to mixing potentiometer (to prevent the filters stages from loading each other).

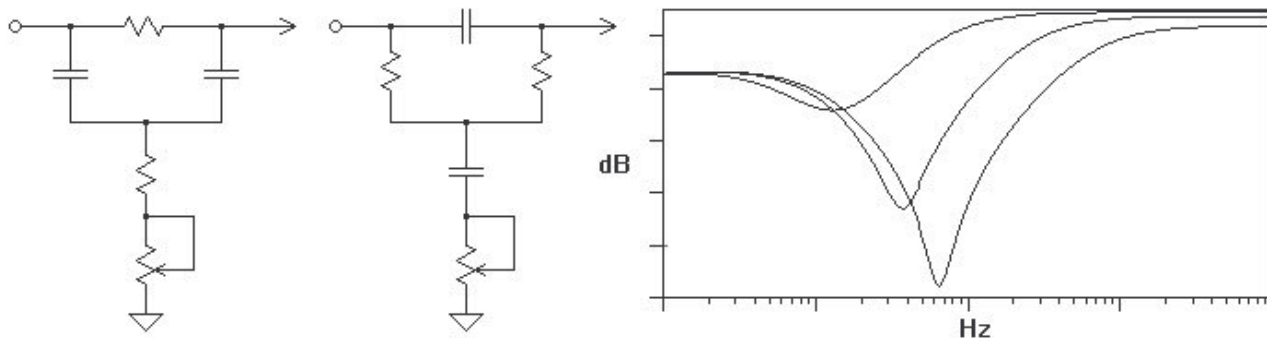


Figure 6.38 An addition of a potentiometer can turn a basic notch filter into a handy middle frequency control.

A very common technique is to use the standard Baxandall-type tone control for bass and treble and then introduce an external circuit for middle frequency control. One approach in doing this is to utilize a basic notch filter circuit: These can either have a fixed value or a potentiometer for a variable notch control. However, like the plot shows, the control is not very linear and the scoop frequency interacts with the potentiometer's resistance. A stranger mid-range control is SWR's "Aural Enhancer" that was introduced in their product lineups in the early 80's. It is practically a much more "linear" band-stop filter that ranges from high band boost to a steep narrow band cut. Both circuits are passive and cannot provide a true mid-range boost. Fortunately, there are alternative solutions to all aforementioned problems.

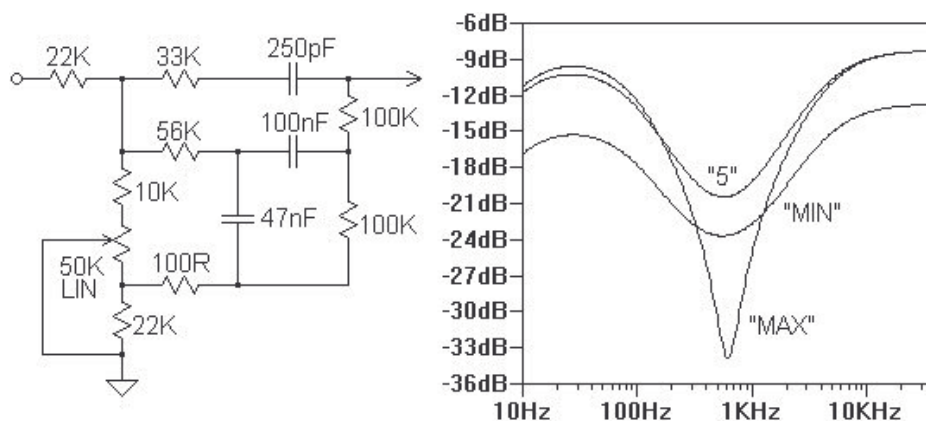


Figure 6.39 SWR "Aural Enhancer": Rather linear band-stop filter that goes from high-bandwidth boost to steep cut.

While Baxandall circuit with added midrange controls can bring a lot of variety there are even more ways to do things differently. Switchable capacitance values are simple and common tricks to add more control but they do not cure the inherent feature of a typical FMV stack: The controls are too interactive with each other. One considerable tonestack variation is presented in Figure 6.40; it combines tone stack designs of Steve Bench and Acoustic Control Corporation. The circuit is almost identical to one used in

Acoustic Control model 370, also known as “Steve Bench” circuit – however, this version uses an *active* configuration *and allows substituting the inductors with gyrators*. The resemblance to tone control circuits used in Acoustic models 230 and 320 is also very prominent; after all, designing is often just refining ideas and topologies that work. There is no need to reinvent the wheel.

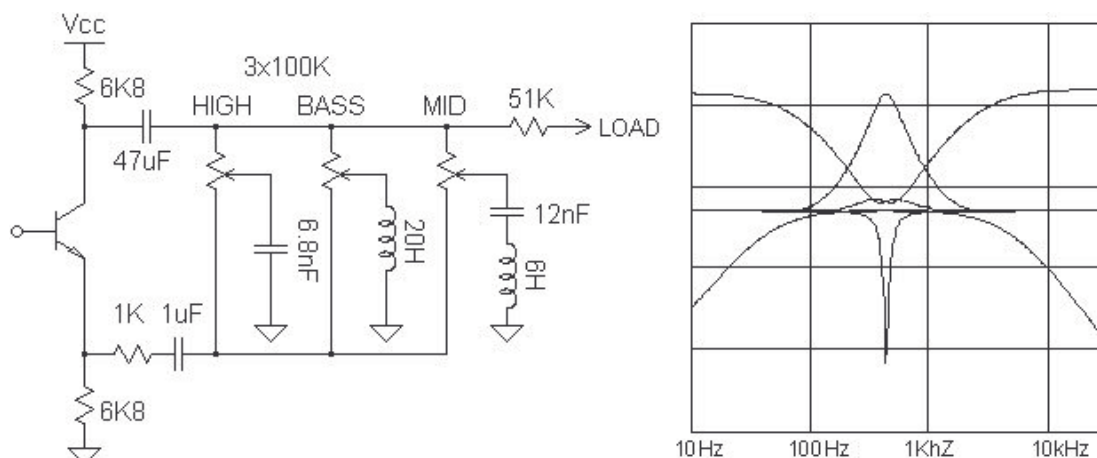


Figure 6.40 Tone control circuit utilizing resonance.

The filters introduce *resonance* that allows a steep control to affect at a very defined frequency band. As can be observed, the sweep plot of the controls is much more precise than in a FMV circuit: This means that the interaction of controls is eliminated and one has a good estimate of what frequencies the controls really are affecting. The circuit is also capable of providing a *real* midrange boost. The design consists of a parallel combination of low-, high-pass and notch filters. Whether this topology is any better than FMV is subjective and depends on the taste of the musician. However, the undeniable fact is that *the circuit is more versatile and intuitive*. In 6.7.9, a similar, yet even more flexible, design is introduced. It will also teach how to get rid of the bulky inductors.

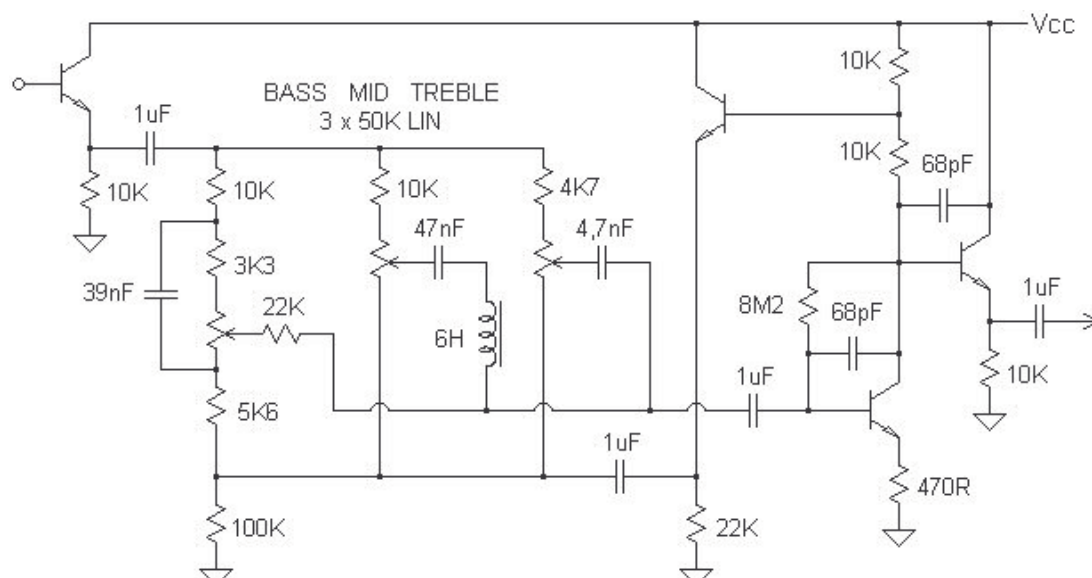


Figure 6.41 Tone control circuit in Acoustic Control Corporation’s Model 370 is pretty close to what is known as “Bench” stack. Note the buffered feedback path.

Everyone who is serious about learning how to design solid-state guitar amplifiers should study schematics of the old Acoustic Control Corporation amplifiers closely. The company produced some very respected amplifiers and instead of just sticking

into using ordinary solutions they really explored advantages of the transistor technology and various negative feedback configurations. Even today these amplifiers stand out with their quality. The Pulteq-inspired “VariAmp” circuit in Figure 6.42 is a nice example of using a basic series resonant circuit, which in the world of guitar amplifiers may seem unorthodox to some. The circuit can introduce either a notch or a boost (effect is controllable with the potentiometer) at a selected frequency. The circuit exploits the same idea as the tone control circuit introduced earlier.

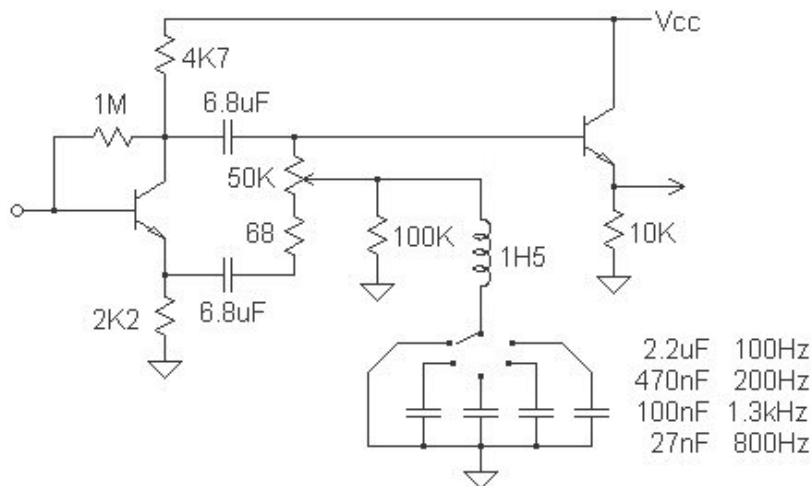


Figure 6.42 A “VariAmp” circuit – or simply just a series resonant circuit in the feedback loop?

6.7.9 Utilizing gyrators

A gyrator is a circuit that inverts impedance. Concept of the circuit was invented around 1948 and its most common application is simulating inductors since those tend to be expensive components and sometimes rather large in physical size. The aforementioned so called *active inductor* circuit works by inverting the impedance of a capacitor. Because characteristics of a real-life capacitor are often closer to ideal than characteristics of a real-life inductor the circuit using gyrators may actually outperform a circuit using real inductors. The inductance of the operational amplifier and the discrete gyrators shown in the example schematics is roughly:

$$L = C1 \times R1 \times R2 \quad \text{(Equation 77)}$$

Resistor R2 effects to frequency and R1 to Q-factor, which determines the height and width of the resonance peak. By making these resistors variable the versatility of the circuit will leap to a next level. The effect of resistor R3 in the discrete circuit is quite insignificant but can make a difference if it’s not chosen accordingly. In practice the calculated inductance will vary a little because transistor’s gain is not unity and component values have some variation. Also, the presented series resistor value equation is merely a rough approximation. However, it is decently realistic for applications that do not require an outstanding precision.

Gyrator circuits have some limitations, though: There is a practical restraint for the minimum value of R1, which means the Q of resonant circuits utilizing gyrators is limited to about 10. An active inductor also cannot store energy similarly to a real inductor and the “flyback” pulse of gyrator is in practice limited by the supply voltage. There is one even greater drawback, which is that *the simulated inductor must have its other end tied to ground* (or appropriate reference voltage). Effectively this means that a signal cannot pass through a gyrator circuit unless it’s being shunted to a

fitting reference potential. There are gyrator circuits that get around this limitation but they tend to be fairly complex. Luckily inductors in audio circuits are often used in shunting configuration so the limit does not create a very considerable problem. The “VariAmp” circuit shown in the previous chapter can work with gyrator as well – assuming that the place of capacitors and inductor is switched. This can be done since they are connected in series.

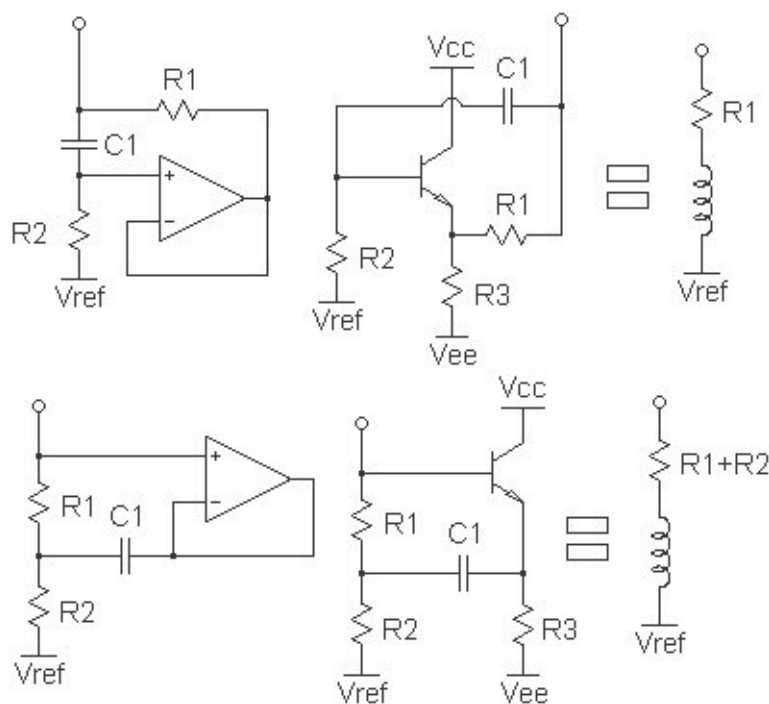


Figure 6.43 Few examples of gyrator circuits that simulate an inductor.

Adding an inductive element to tone control circuitry is usually very beneficial. Due to introduced inductance the circuits may exhibit resonance on a much narrower area. This makes them more ideal for use in graphical equalizers that have multiple bands. Further on, varying the resonant frequency, thus converting the circuit into a parametric equalizer, is very simple with a gyrator: In conventional circuit one would need either a variable capacitor or a variable inductor. Both components are generally rare, expensive and bulky. However, gyrator turns into a variable inductor simply with an inclusion of a potentiometer.

The operation of a graphic equalizer circuit is basically very simple: The inductor (which is usually a gyrator), resistor and a capacitor form a series resonant circuit. If this is located in the input of an operational amplifier (potentiometer’s wiper at input) the resonance does not have a voltage peak but a notch since the input is shunted to ground. That is, the circuit is attenuating the signal. When the circuit is shifted to shunt the feedback loop the effect becomes inverted: The notch becomes a peak. The potentiometer hooked between inputs controls the circuit’s operation between these modes. When the wiper is directly in the center position the resonant circuit “drops out” leaving frequency response flat. The resonant and center frequencies plus the Q value of the filter can roughly be calculated from:

$$f_r = 1 / (2\pi \times \sqrt{L \times C}) \quad \text{(Equation 78)}$$

$$f_o = 1 / (2\pi \times L \times C) \quad \text{(Equation 79)}$$

$$Q = X_L/R = fr/(f2 - f1) \quad \text{(Equation 80)}$$

If many series resonant circuits are connected in parallel in such a circuit one has effectively constructed a graphic equalizer. Basically the only limit to amount of bands is the Q value of the resonant circuit since it defines the bandwidth. Knowing the Q value is important because it indicates the bandwidth of the filter defined by the two -3 dB points, f1 and f2. Since a graphic equalizer typically has fixed frequency and Q it should be set up so that the total bandwidth exhibits no peaks or notches when all controls are set to middle position. In practice this means that corners frequencies of the bands have to interweave each other slightly to compensate the -3 dB loss.

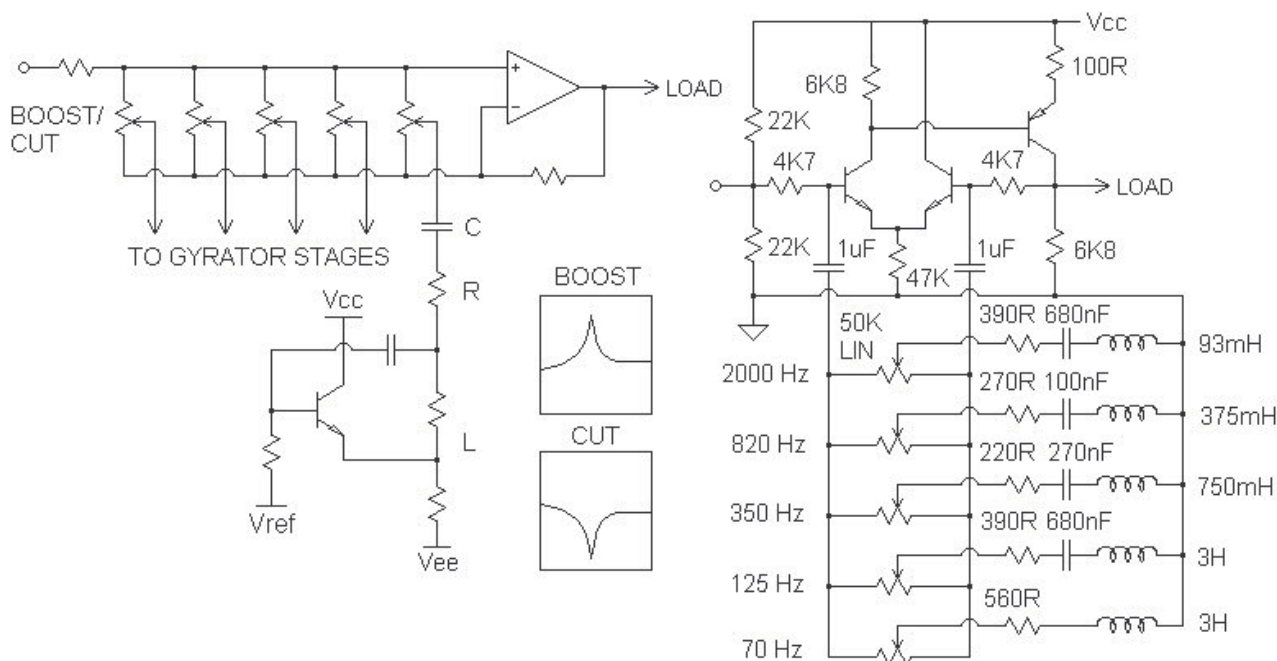


Figure 6.44 Graphic equalizer circuits. Left: “Modern” circuit using operational amplifier and discrete gyrator inductors. Right: “Vintage” circuit with discrete differential amplifier and real inductors.

Since vintage “purists” probably wish to know how things worked back in the good old days when circuits were built from discrete components, a schematic showing an example of a “vintage” design is also included: It depicts the equalizer circuit from the Acoustic Control Corporation amplifier model 230. Similarly to “modern” version, the resonant circuits are placed inside the feedback loop of a *differential amplifier* (a discrete OpAmp). In theory, the two circuits are completely interchangeable (the circuit is not aimed to clip or overdrive so the choice of active device makes no difference); one just has to calculate the proper component values for the gyrators. Only real differences are the poorer PSRR and CMRR of the discrete version - along with higher amount of distortion. A discrete version can naturally provide more headroom than the OpAmp circuit - assuming rail voltages are increased. However, the original circuit ran from 30V single supply, which corresponds to modern configuration of using ± 15 V rails.

Marshall Amplification has patented another interesting application for gyrator circuits [48]: When a series resonance circuit is placed into the feedback loop of a typical non-inverting gain stage its frequency response effectively begins to resemble that of an amplifier with high output impedance driving a speaker load. The series resonant circuit causes the gain peak and the series resistor and capacitor, which form

a high-pass filter, are tweaked to provide a substantially high cutoff frequency. The troublesome inductor (which would be quite high in value) is once again replaced with a conventional gyrator.

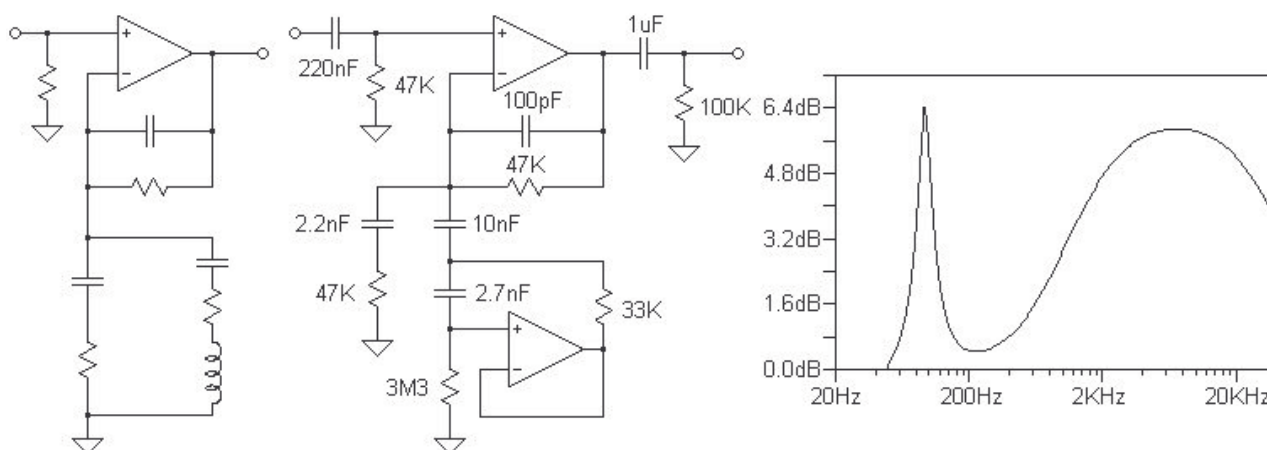


Figure 6.45 Gain stage simulating the frequency response of an amplifier with high output impedance driving a speaker load.

6.7.10 T-Filters

Another common technique to obtain a typical frequency response with scooped middle frequencies is to use a notch-T filter. This is a band-stop filter and making the resistors variable can alter its operating frequency. Easiest way to do this is to use a dual-ganged potentiometer. With the configuration shown in the schematic the notch frequency is:

$$f = 1/(2\pi \times R \times C) \quad \text{(Equation 81)}$$

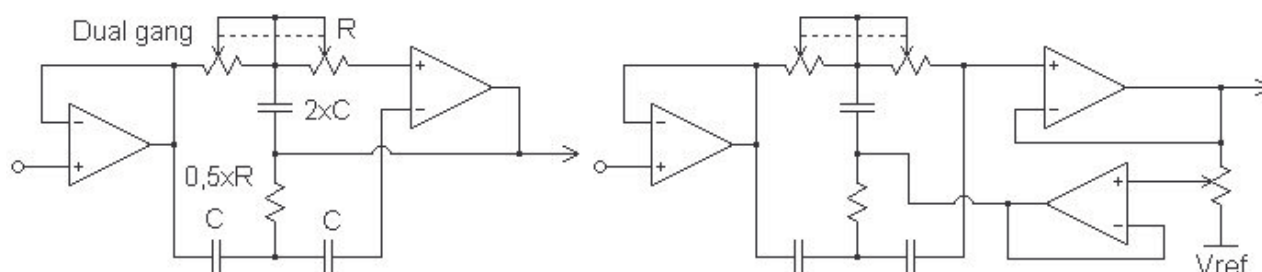


Figure 6.46 Notch-T filters. Left: The standard configuration with adjustable frequency and right: Buffered circuit for portable adjusting of Q.

If the “ground” point is buffered and the buffer’s input signal amplitude made variable the Q of the filter can be improved. With high-quality and close-tolerance parts such a circuit can be used as a precise band-stop filter to remove mains frequency hum or acoustic feedback signal occurring on a certain frequency.

6.7.11 About placement of the tone controls

Such a simple thing as placement of the tone control circuit can make a drastic difference, especially if the amplifier uses effects loops or utilizes circuits that add distortion. Logically, all effects that are fed post tone control sound different than effects fed before it. The traditional “British” high-gain approach inserts EQ after distortion stage(s) while the “American” Mesa Boogie approach inserts EQ before the distortion stage(s). The difference is drastic. Some people have described the tone of

the first configuration more aggressive, thus better for metal music styles, while the latter configuration is more suitable for blues and alike. In high gain circuits there is also a serious danger that pre-distortion tone controls will be found lacking in their adjustment range, as the dynamic range of signal is severely limited due to high gain and clipping.

Another issue is the insertion point's effect to SNR: If tone control circuits are passive they usually introduce quite high insertion losses. If the circuit is fed with a low amplitude signal this means a huge decrease in overall SNR. Usually active circuits tend to give better results in solid-state amplifiers – at least they introduce less noise and improve headroom of the preamplifier stages since high gains are not needed to compensate insertion losses of several decibels.

Some amplifier circuits using more unorthodox concepts can spread the tone controls all over the circuit. This was very typical in old Gibson amplifiers. Another great example of similar technique is placing “presence” and “resonance” controls into the negative feedback loop of the power amplifier. Some amplifiers even went further and incorporated a whole active tone control circuitry as a part of the feedback loop.

6.7.12 Automatic frequency-compensated volume control

Due to non-linear hearing curve it is necessary to introduce some bass and treble boost to the signal on low listening volume levels (refer to Figure 1.10). Incorporating a volume control that would do this automatically to a guitar amplifier is not very common practice – not that it would be even in HiFi amplifiers. Nevertheless, configuration is possible so short introduction is beneficial. Because perfect reproduction of the fairly intricate human hearing curve would require a very complex circuit, most frequency-compensating volume controls are quite imperfect.

There are basically three methods to realize a frequency-compensated volume control circuit: 1. Stepped controls, 2. Using tapped potentiometers and 3. Incorporating “inverse volume expansion with multi-channel amplifier” [47]. The last method basically combines the curve for specific volume level from the signals of three individual channels; one set flat, one set to provide a required bass boost and one set to provide the treble boost. Circuits incorporating the third method are generally too complex to have any practical use so they are left out of scope of this book.

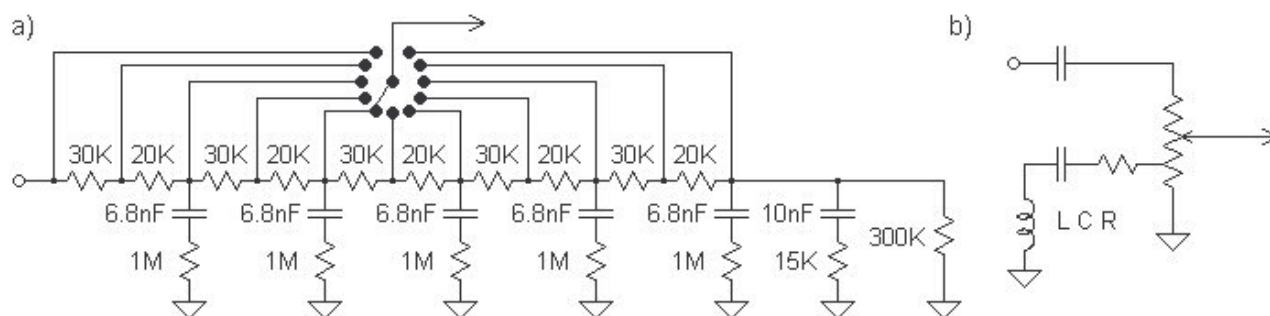


Figure 6.47 Examples of frequency-compensated volume control circuits: a) Stepped volume control and b) tapped potentiometer method.

A typical method used to be to use a variation of circuit b). L and C form a series resonant circuit, which is (typically) tuned to about 1 kHz. R is then used as limiting resistance for bypassing middle frequencies. Tapping point is usually 1/5 of the total resistance (14 dB below maximum level). If treble boosting is not desired L can be

omitted. More ideal performance can be extracted by using potentiometers with multiple tapping points [47]. Problem today is that tapped potentiometers have become very rare and expensive.

Circuits incorporating method 1 are usually more complex but require no special components. Stepped controls do not provide a smooth control but they are quite precise and generally better quality than bargain bin potentiometers. They don't get scratchy either. However, stepped control neglects the treble boosting (which is fortunately considered less-important) [47]. The design in circuit a), which was introduced in RCA manual from 1953, uses five 6 dB filter sections, which are divided into two parts. The circuit works with stages that have low source resistance and bass reproduction of hearing curve is fairly accurate.

6.8 Volume- and gain controls

Even as simple parts of the circuitry as gain- and volume controls have some features that should be discussed. However, before advancing any further it is important to make a distinction between volume- and gain controls: A volume control is, by principle, a control that affects the apparent loudness level. A gain control is very similar (and in large extent the two are interchangeable); however, the function of a gain control is strictly literal: It controls gain. The important distinction is that increasing or decreasing gain (of a certain stage) does not necessarily have to affect the perceived loudness level. This is particularly apparent in cases where a gain stage is driven into clipping. Since a gain control may as well serve as a volume control (and it often does) it is a lot more worthwhile to make a distinction between a *gain* control and an *attenuator*. Like the name implies, the latter is typically a passive control and quite *opposite* to a gain control.

Since loudness perception is logarithmic by nature attenuating volume controls are (typically) logarithmic as well. An "audio taper" potentiometer attenuates the signal 9/10 when volume is set to "5" and at lower volume levels the attenuation becomes more and more substantial. A cheap trick of some manufacturers is to use a linear potentiometer to give a false impression of the amplifier's power since such an amplifier is extremely loud on low volume settings. In reality it's no louder than it would be if it had a logarithmic potentiometer – it only has a limited (usable) range of control: Past a very low "volume" setting (usually 2 – 4) the only thing increasing is distortion – not perceived loudness.

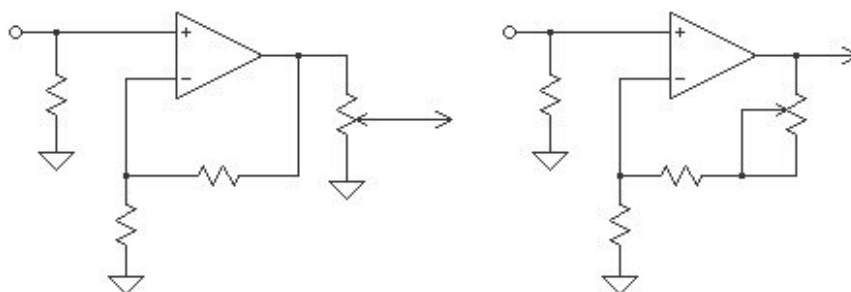


Figure 6.48 Left: Attenuating control (passive) and right: Gain control (active).

Two possible locations to place an attenuator are before and after a gain stage; both configurations have advantages and drawbacks. The inherent problem of a post attenuator is that there is no way of limiting the input signal of the preceding stage. With hot enough input signal the gain stage is overdriven and virtually nothing can be

done to prevent this. If the attenuator is placed in front of the gain stage the problem in question can be avoided; unfortunately, this brings in another drawback: The noise level introduced by the gain stage remains constant, however, the input signal amplitude may be significantly attenuated; this means an obvious decrease in SNR.

The combination of the controls was an early attempt to compensate the drawbacks: The ganged input and output attenuators are linear and together they have a “taper” function, which is “*slightly logarithmic*”. Practically the taper curve is in between those of linear and logarithmic potentiometers. Due to use of linear potentiometers the gain stage is cleverly fed with higher input signal: For example, with volume set to “1” the input signal amplitude is 10 % of the original instead of 0,0007 %, which would be the case if the input attenuator was logarithmic. This results into a significant improvement in SNR. Naturally this also leads to “hotter” driving of the gain stage but bear in mind that the potentiometer still introduces fairly considerable amount attenuation (9/10). However, since the logarithmic curve of the combined tapers is not as steep as in a certified type A “audio taper” potentiometer there will be some issues with the sensitivity and usable range of control.

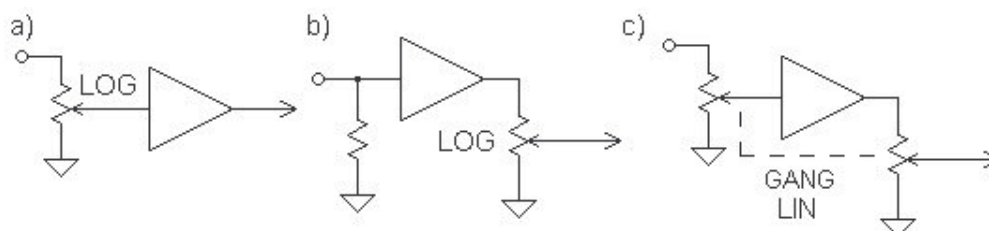


Figure 6.49 Configurations for attenuating control: a) Preceding gain stage, b) post gain stage and c) combined.

A gain control is practically free of the problems associated with attenuators since it allows adjusting gain to a level where overdrive will not occur while SNR still remains reasonable. It is not a completely trouble-free arrangement though: Typical OpAmps may turn unstable in less than unity gain (read: attenuating) configurations, which makes substituting a traditional volume control quite hard. One must also note that conventional potentiometers are rather vulnerable to DC current and in a position where they are exposed to constant DC potential they become noisy or “scratchy” very fast. This is a problem that pesters many single-supply powered circuits. The rule of using logarithmic attenuators does not apply to gain controls, which’s selection is very case-specific. Depending on the application, the taper function that is required to achieve a logarithmically increasing (or decreasing) gain may be linear, logarithmic or (annoyingly) even “antilog”.

A typical potentiometer has substantial stray capacitance in reference to its casing (consider this as a capacitor shunt from wiper to ground). A typical amount of shunt capacitance is about 30 pF and in many circuit realizations this figure is often increased by stray capacitance to hood of a shielded wire. If all stray capacitances are taken into consideration the worst-case estimation is in the order of about two hundred picofarads! In high impedance circuitry this causes a considerable treble loss. Note that the wiper sees the highest impedance when the control is set to half resistance, thus losses are highest at this setting. The treble loss can be compensated with the familiar treble bleeder circuit that offers a shunt path for higher frequencies.

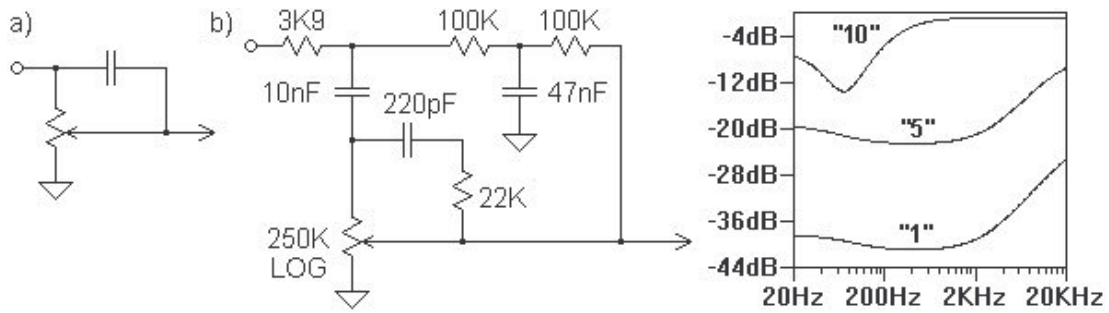


Figure 6.50 Variations of attenuating control: a) Treble bleeder and b) “Gain” control from Hughes & Kettner Vortex amplifier.

Excluding the occasional use of “treble bleeders”, gain and loudness controls have traditionally been rather linear in respect to frequency response. Apart from inherent simplicity of the linear configuration there are no noteworthy reasons to back up the use of such design. On the contrary, there are more reasons why the controls should not be linear:

- **Loudness:** As shown by Fletcher and Munson, human hearing does not respond linearly to changes in loudness. At lower volume levels our ability to sense bass frequencies decreases, which creates a need to tweak the tone control settings in relation to loudness level. This is rather annoying and various attempts to compensate the problem have been made throughout decades – all of them seem to require fairly complex analog circuitry. (Perhaps DSP will effectively solve this issue in the future). To make matters worse, “guitar amp snobs” have traditionally been extremely conservative and resentful towards new ideas rejecting any unconventional design as poor or “tone sucking”. Regrettably, tonal effects of different coupling capacitor brands in a 5E3 clone tend to spawn more discussion and interest than eccentric new circuits.
- **Gain:** If gain control adjusts the amount of overdrive (meaning signal clipping in a certain stage) it (often) benefits from an arrangement that attenuates the bass frequencies when gain increases. This is done to prevent blocking distortion but also to compensate our non-linear hearing. There may also be a need to intensify higher frequencies when the amount of gain is decreased. Such designs have been employed with good results, for example by Hughes & Kettner. Peavey and Fender have also introduced gain control arrangements that limit the bandwidth in proportion to increased gain.

6.9 Spring reverbs

Spring reverbs are electromechanical devices that emulate the natural reverberant effect with the aid of transducers and vibrating springs. Transducer is a device that converts electric energy into kinetic energy or vice versa. An electric current from a “driver” circuit creates a magnetic field that is proportional to the drive signal. This magnetic field controls a solenoid (transducer), which excites a set of springs. Thus the electric energy is converted into kinetic energy. Another transducer, located at the other end of the springs, picks up the movement of the springs. The kinetic energy is then converted into current and fed to a “recovery” circuitry that amplifies the weak signal. Perhaps the most famous manufacturer of these electromechanical devices is Accutronics®, which started out as an unit of Hammond Organ Co. but at the time of writing is a division of Sound Enhancement Products, Inc. Another noteworthy manufacturer is Belton Engineering Co., Ltd.

Spring reverb technology is rather old dating back to the late 1930's. Bell Telephone Laboratories had developed the earliest spring reverb, which was devised to simulate the delay of long distance calls. Laurens Hammond (the founder of Hammond Organ Company) saw the opportunity of using the invention as a way to add artificial reverberation sound to the organs he manufactured. In 1939 he filed a patent describing a mechanical reverberation system (US Patent 2,230,836). The unit described within was about four feet tall and had to be enclosed inside a speaker cabinet.

The requirement for artificial reverberation effect ensured further development: In 1944 John D. Stack, assigned by 20th Century-Fox, filed a patent describing a less colossal device (US Patent 2,437,445). This was followed by Hammond Organ Company's "necklace reverb" named after the way the springs hung in the air. Patent for it was filed in 1958 (US Patent 2,982,819). The "necklace" was a huge success due to its light weight and compact size but the "unconventional" spring suspension proved to be a failure: There was nothing to prevent springs from banging against each other or the metal frame of the device. In 1961, Alan Young – a designer for Hammond Organ Company and a member of "necklace reverb" design team – introduced a new design, which was initially known as Hammond Type 4. This was the first "conventional" spring reverb unit and was happily adopted by various organ and amplifier manufacturers – Including Leo Fender who fitted the unit into the Twin Reverb amplifier. Patent was granted in 1963 (US Patent 3,106,610). US Patent 3,159,713 (W.C. Laube, Jr) – which was also assigned for Hammond Organ Company - followed in 1964.

Patents for a spring reverb unit were also granted to such companies and individuals as D.H. Baldwin Co. in 1965 (US Patent 3,199,053), H.E. Meinema in 1968 (US Patent 3,363,202) and Wurlitzer Co. in 1968 (US Patent 3,391,250). The Baldwin spring reverb is worth of extra mention since it utilized *only one spring*, which was, however, "tapered" in order to create the same "full" reverb tones that units using multiple springs had. O.C. Electronics manufactured Herbert Meinema's patented "Folded Line" reverberation devices. These were used at least in Sunn amplifiers. The tanks are famous for their sticker that stated "Manufactured by beautiful girls in Milton Wis. under controlled atmosphere conditions."

Similarly as in loudspeakers, the driving force of spring reverb's solenoid is *current* instead of voltage. Since solenoid is also an inductor the spring reverb circuits have to deal with variations in drive force caused by the effect of impedance: If the driver amplifier has a constant voltage output, the reverb produces less output with higher frequencies. This forces the recovery amplifier to use higher gain that in turn leads to decrease in SNR. Effectively, spring reverbs should be current driven – at least partially. See 3.6.2 for more details about the technology. The current drive and its effects are once again an important factor in explaining why some people tend to prefer reverb circuitry that uses vacuum tubes.

To reduce hum and vibration noise (of mechanical reverb units) it is desirable to have as high drive signal of the reverb unit as possible [12]. Accutronics technical literature goes as far as suggesting that the reverb coil should be driven to near saturation (yet actual saturation should be avoided). This may at first seem a bit illogical since controlling the drive signal's amplitude would seem like the most natural way to control the amount of reverb. However, we want to reduce the gain of the recovery

stage as much as possible since this will also reduce the possibility of feedback and microphonics. Typical practice is to mix the “wet” signal from recovery stage to the clean “dry” signal using a constant-value resistor network. Other option is to use a potentiometer to control “blend” between wet and dry signal sources.

Spring reverb transducers cannot exploit low frequencies. Some of those can even cause the reverb tank to go into resonance and emit low frequency “howl”. The driver amplifier should therefore include a high pass filter set at approximately between 150 - 300 Hz. The low frequency response is a matter of taste but to remove the “rumble effect” the high pass filter’s cutoff frequency should, at least, be something between 50 - 100 Hz. One should experiment with values since different tanks behave differently.

When determining the drive level look for a tank’s variable called maximum amperes per turns. This tells how much current the inductor can take before going into saturation. If the value is listed in RMS convert it to peak value. You also have to know the amount of turns in the inductor. The relation between turns and impedance at a certain frequency ($Z @ f$) gives an approximation of the required drive voltage on a drive frequency f_D :

$$U = (amp - turns_{MAX} / turns_{TOTAL}) \times Z \times (f_D / f) \quad \text{(Equation 82)}$$

If the values are unobtainable then there is no other choice but to base the design on the assumed drive level. This procedure should be self-explanatory. Again, let it be stressed out that *the driving force is current*, thus the driver amplifier needs to have a sufficient current feed capacity. Input impedances of typical reverb tanks (at 1 kHz) can be close to 8 ohms or as high as several kilo-ohms. It is clear that some tanks need a rather powerful amplifier to push enough current to the transducer load. Typically tanks that have a higher load impedance require less drive current but a higher drive voltage and vice versa. The following data table of Accutronics reverb tanks can be used as guidance as long as it’s not interpreted as a gospel. As always, the best practice is to consult the data presented by the tank’s manufacturer.

Table 6.2 Accutronics reverb tank properties

Impedance @ 1kHz (±10 %)			DC Resistance in Ohms (±10 %)	# of turns	Nominal drive current (mA)	Max I (mA)	Drive voltage (V _{RMS})
Input Z (Ohms)							
	Type 9 or 8	Type 4 or 1					
A	10	8	0.81	124	28	242	3
B	190	150	26	538	6.5	56	10.5
C	240	200	27	600	5.8	50	12
D	310	250	36	688	5.0	44	13.5
E	800	600	58	1100	3.1	27	21.5
F	1925	1475	200	1700	2.0	18	35
Output Z (Ohms)							
A	600	500	42				
B	2575	2250	200				
C	12000	10000	800				

The current feed capacity of the driver is actually more important than the dimensions of the reverb tank: I have heard 5-inch tanks that sounded pretty good, although they

did not have the long reverberation time of larger tanks. I have also heard a similar 5-inch tank that sounded horrible because the driver stage could not do its job correctly and push enough current to move the springs: The reverb had a short decay time and a distorted sound originating from the clipping driver stage.

Good and simple choices for driver devices are, for example, NE5532 or LM386 chips. Other solutions exist as well: Many amplifiers use an OpAmp-based driver that is buffered with semi-power transistors. Using several OpAmps in parallel connection is a very common configuration as well. Another factor contributing to the reverb tank's tone is the number of springs having different tensions. Having more springs creates a fuller sound.

The output impedance of a reverb tank at 1 kHz can be anything between 100 ohms and 10 kilo-ohms and some – depending on the model. The input impedance of the recovery stage should be chosen in respect and it should be at least high enough not to load the source and attenuate higher frequencies. Note that the source is inductive, thus the output impedance will rise at higher frequencies. The output voltage is typically in the range of 1 – 10 millivolts. The gain of the signal recovery amplifier must naturally be high enough to amplify the reverb signal to match the mean signal level at the insertion point. *I have seen recovery stages with gains ranging from 1000 to zero.* Since each circuit is always an individual case it is hard to state anything definite about this subject. You (likely) *do not* want to overdrive the recovery stage, though. For the same reason it is advisable that the reverb is the “last” effect in line and only followed by dynamic processors or other “soft” and transparent effects.

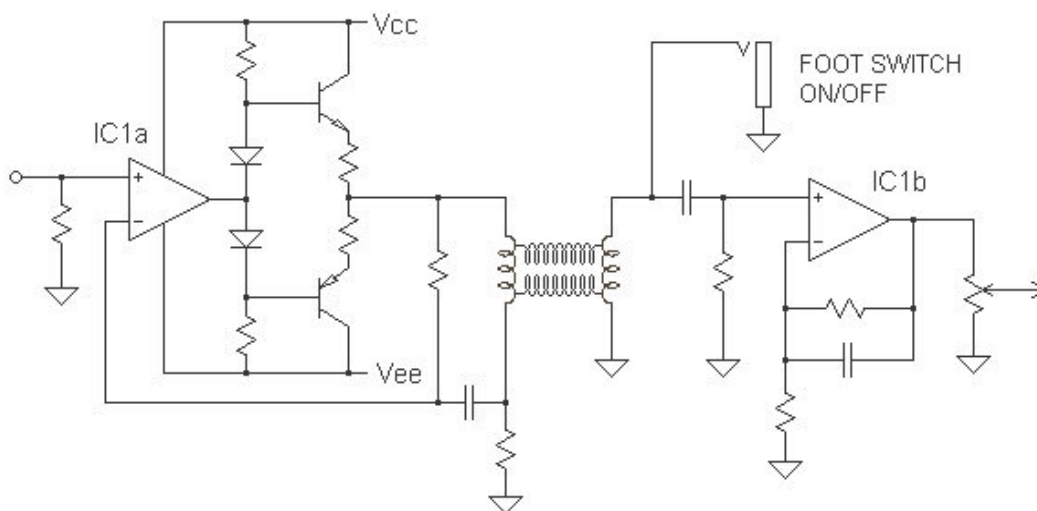


Figure 6.51 A spring reverb driver/recovery circuit with discrete buffer and footswitch toggle.

Example of a pretty versatile driver/recovery circuit is shown in Figure 6.51. The transducer is driven by an operational amplifier stage that is buffered with a discrete emitter follower in order to increase current feeding capacity. Suitable output devices are for example semi-power BD139/BD179 or BD140/BD180. If requirement for drive current is not very high and the operational amplifier is a model that can cope with higher currents the discrete buffer stage can naturally be omitted. To mute the circuit, the output of the spring reverb can be shunted to ground with a footswitch that simply connects tip to sleeve.

Tube technology purists will of course rant about the tube sound and the superiority of tube-based reverb units. To be fair, tubes do little good in reverb circuits: The impedance matching provided by the transformer is the only considerable sound

enhancer and causes similar effects as conventional current drive topology (that is often the recommended scheme anyway). Soft clipping of tubes does no good either since overdriven reverb signals - in both driver and recovery stage - generally just deteriorate the sound integrity. Drive signal can be easily limited with any soft clipping arrangement and the simplest limiting arrangement is the trick used in some Hammond organs: Insert a light bulb (or few of them) in series with the tank input.

Mounting the spring reverb tank is not so straightforward as one probably thinks either: Actually the worst way to mount most of them is the usual way of placing them horizontally on the bottom (or top) of the cabinet. Horizontal or even vertical placement to one *side* of the cabinet is a more ideal solution considering magnet alignment. This does not apply to every tank but to most of them anyway. Naturally, the tank should be located away from magnetic sources and should never be mounted on a resonating section of the chassis or the cabinet. Use of rubber grommets is highly recommendable. Consult the tank datasheet for further details.

6.10 Effect loops and lineout circuits

Effect loop -and lineout circuits are simple and addition of them can increase the versatility of the amplifier greatly. In practice, both circuits work passably in their simplest form but increased versatility and professional quality comes from adding some extra features. For example, simple effects loop is nice but if one can toggle it by footswitch it is even nicer – not to mention a feature of footswitch controlled selection between *two* effects loops - each one equipped with tone controls and “blend” potentiometer that allows mixing dry signal and signal return from the loop. Circuits are pretty basic and designing them practically involves use of plain logic instead of vast knowledge about circuit theory. Question is, how much complexity is one willing to add or how much of it is really even needed.

Since most of the circuit ingredients used in lineout- and effect loop circuits have been presented already I shall not discuss them too extensively. If one wishes to add up a tone control to an effect loop he surely should have some good ideas of how to do it after reading the section that discusses them. If I have felt that there is a need for a detailed schematic or description it is included.

Effect loops are basically best to be driven by buffer circuits that provide both sufficient current and reasonably low output impedance for optimal voltage transfer. The greatest design dilemma is the input sensitivity of most effects: 99 % of pedal effects are designed for input signal amplitudes that a *guitar* provides; this means a nominal level of maximum 100 mV or so. Most rack mount effects, however, work better with “line-level” signals. Note that in many amplifiers the purpose of the effect loop is specifically feeding rack equipment. Sometimes driving typical “floor pedals“ with high amplitude input signal is beneficial - sometimes it is not. My personal opinion is that the nominal amplitude of the signal send to conventional pedals should never exceed 1 V. There are lots of effects out there that clip their input stage even when the amplitude is as low as 200 mV. The best solution is of course to make the amplitude of the send signal variable.

FX loops can be categorized to series and parallel types according to their placement in the signal path. Typical configuration is the series loop. Parallel FX loops typically have some very nice features like “blending” but also some quite significant problems: First of all, *signal phase* may become an issue and in the worst case that turns the

amplifier into an oscillator. Usually phasing errors just cause weird attenuation of the signal. Secondly, parallel loops are quite a poor choice if the looped effect tampers with the frequency response. Blending between dry and equalized wet signal is not very efficient – and again the phasing issue may result into strange attenuation or enhancement of certain frequencies - as well as oscillation. As a rule of thumb, graphical equalizers, and all similar effects, work better in series loops.

Shown in the following example are two versions of an effects loop circuit. The circuit a) shows a conventional and simple application: The return jack (on the right) has an internal switch mechanism that establishes a connection to the tip when no plug is inserted. Inserting a plug into the return jack breaks the signal line and it can only be reformed by creating an effects loop between send and return jacks. (In some cases this may be nothing else but a single cord). The signal levels are adjusted with resistance dividers formed by R_{ATT} and effect's input impedance or R_{ATT} and Z_{IN} , the latter in the case when loop is not in use. As one may note, the loop jacks can also serve a double duty as individual “preamp out” and “power amp in” jacks. This is one more reason why signal levels at the loop point are subject to careful consideration. The circuit is simple but since it is *passive* it may increase SNR.

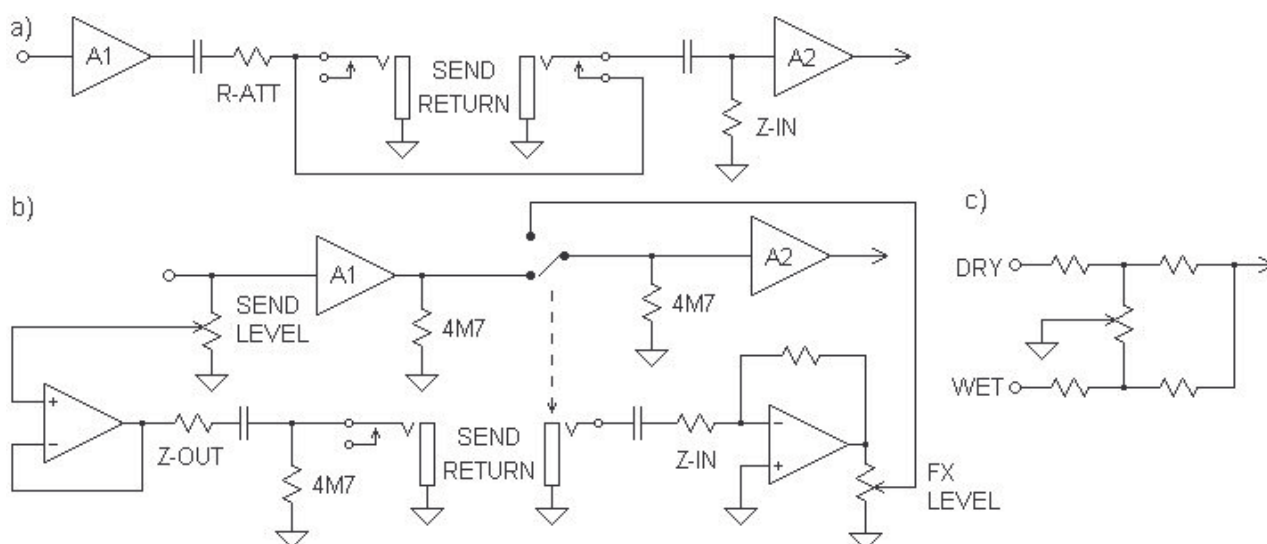


Figure 6.52 Effect loop circuits: a) Conventional series loop and b) versatile. Circuit c) is a simple “blending” mixer for parallel loop.

One alternative implementation feeds the effect loop through a buffer. The send level is adjusted with a potentiometer that also defines the input impedance of the gain stage in parallel with the loop. Output impedance of the buffer OpAmp is extremely low so it's practical to define the real output impedance with a low-value resistor Z_{OUT} . 4.7 Megaohm resistors act as “click suppressors” by keeping DC voltage potential at steady zero at switching nodes. The return jack has an internal DPDT-switch that is actuated by inserting a plug. One of the switches is left vacant so it can be used, for example, to control an indicator LED. Return signal is amplified and the output amplitude can be adjusted to a preferable level. The circuit introduces no insertion losses – other than those of the potentiometers. A “blend” feature (and conversion to parallel loop) is very easy to implement; just remove the switch and replace it with a circuit depicted in figure c).

Lineout circuits are used for plugging the amplifier straight into PA and studio equipment so they can be considered as direct injection “DI” circuits (at least sort of). There are standards for the *nominal* line-level amplitude, which specify it being either

-10 dBV (consumer equipment) or +4 dBu (professional equipment). The dBV unit is referenced to 1 volt so the decibel math indicates that the line-level signal should have RMS voltage of 0.3162V. The dBu unit has a reference of 0.775 V so the +4 dB change correlates to RMS voltage of about 1.228 V. Using higher line-level decreases SNR. As always, some instances “break” the standards by creating more of them; for example, German studio standard is +6 dBu (1.55 V_{RMS}).

Standards for input and output impedances are somewhat “lacking”. The 600-ohm input and output impedance-standard - suitable for *optimal power transfer* – never worked well with audio and has been long gone. As a general rule, the input impedance of the receiver should be high and the output impedance of the source should be low to guarantee optimal *voltage transfer*. What low impedance here means is a range from 100 to 150 ohms - or less (although some systems go as high as 1 kilo-ohm). It is advisable to maintain the source impedance at least 10 times lower than the receiver’s input impedance. Typical input impedances for mixers vary a lot but usually they range from 10 kilo-ohms to 1 megaohm. It is safe to base the design on an assumption that the input impedance is a little bit less than 10 kilo-ohms; this should assure good signal transfer even to badly designed equipment.

Professional systems use *balanced* signal lines to remove common mode noise induced by the cable. The transfer line uses three conductors, one for normal signal, one for inverted signal and one for common. The receiver side has a unity-gain amplifier that amplifies the difference between signals thus removing common mode error. Although being unity-gain, the amplifier has output amplitude two times higher than the input amplitude. Sometimes a passive “balun” element (abbreviation for balanced-unbalanced) is used instead. This is usually a center-tapped, ferrite-cored inductor that acts as a mixing element. The component is small enough to be fitted inside a standard phone or XLR jack.

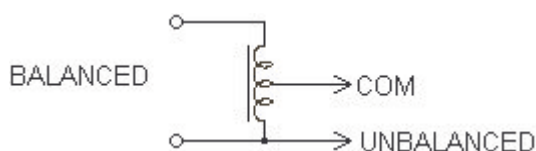


Figure 6.53 Balun.

Table 6.3 Pin orders for balanced line driving.

XLR		TRS phone connector	
Pin	Signal	Pin	Signal
1	Signal common	Sleeve	Common/shield
2	“Hot” Normal	Tip	“Hot” Normal
3	“Cold” Inverted	Ring	“Cold” Inverted
Case	Shield/Earth		

The lineout feeds use three kinds of connectors: Three-pin XLR, RCA and 6.3 mm (1/4”) phone connectors. The phone connectors are divided into two types: TS (Tip-Sleeve) - or simply just “Mono” - and TRS (Tip-Ring-Sleeve) or “Stereo”. Typically use of RCA connectors is reserved only for “consumer” line-level output (-10 dB), while XLR and phone connectors are used with both line-levels. The latter can also provide both balanced and unbalanced signal sends – assuming that the phone connector is TRS-type.

Note that great amount of equipment (again) mess up the standard by having pins 2 and 3 (in XLR) reversed. If balanced signal is fed to an unbalanced system the

negative signal is simply shunted to ground; this can load the feeding stage hardly if the output impedance is very low.

Table 6.4 Usual pin orders for “insert”.

XLR		TRS phone connector	
Pin	Signal	Pin	Signal
1	Signal common	Sleeve	Common/shield
2	Send	Tip	Return
3	Return	Ring	Send
Case	Shield/Earth		

Another way to use the same connectors is to construct an effects loop known as *insert*, which is a bit different than basic effects loop concept of guitar amplifiers. These are quite common in professional mixers, though. The loop utilizes only single, stereo cord. Insert circuits have been used at least in some Crate amplifiers. In patch bays the tip connector of a TRS phone jack serves as the output and the ring connector serves as the input, thus in the *feeding* system (the amplifier) the corresponding order is: Tip: Input (“Return”) and ring: Output (“Send”).

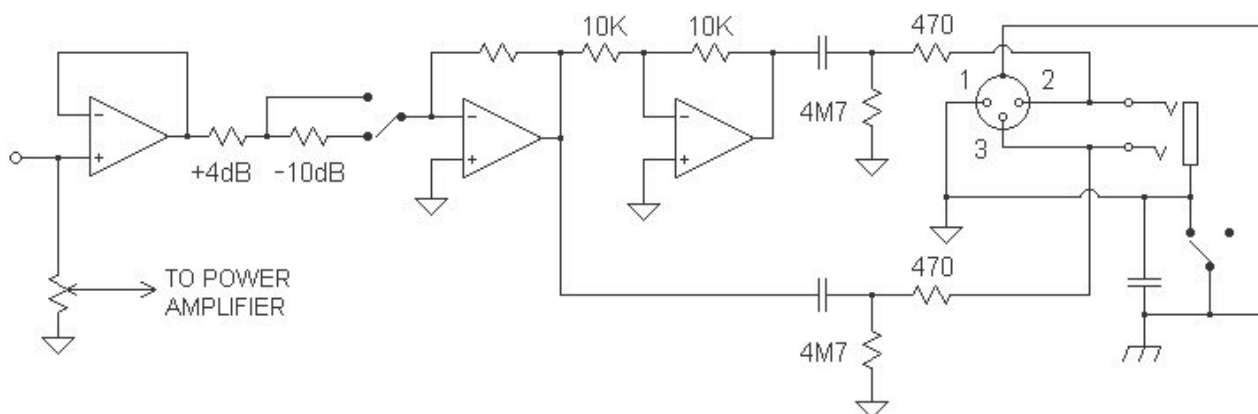


Figure 6.54 An example of a typical balanced lineout circuit.

Shown is an example of a typical lineout circuit. A signal from a suitable point in the circuit (which in the example is the point feeding power amplifier and before the master volume control) is fed to a buffer that isolates the following circuitry. An inverting amplifier, with variable gain for +4 dB and –10 dB line-levels, precedes the buffer. After the inverting amplifier stage the signal is split to a unity-gain, inverting amplifier that corrects the phase. The outputs of both stages are fed to XLR and TRS jacks through 470-ohm resistors that set the effective output impedance. 4.7 megaohm resistors are used to “bias” the coupling capacitors, which will probably be electrolytic and therefore need a correct polarity. The operational amplifiers will likely have a small *negative* offset voltage. The signal common can be lifted away from the chassis (Earth) with a switch.

A more versatile solution would also introduce a “cabinet simulator” after the buffer. The circuit would shape the frequency response to mimic the effects created by the guitar cabinet and power amplifier’s interaction with speaker impedance. These are important factors since people usually dislike the “clean” tone of the amplifier itself. A crude “cabinet simulator” is a high order low pass filter that has a –3 dB point at around 5 kHz.

6.11 Clipping circuits

Clipping circuits are used for introducing either compressive or “enhancing”, amplitude distortion by clipping off the signal portions that have either very high amplitude (compressive) or very low amplitude (enhancing). The latter configuration is very handy in *noise gate* circuits but tends to introduce an unmusical “crossover” effect. The compressive limiting of dynamic range is what most people seek from a typical “distortion” effect anyway, thus it has become nearly synonymous with the term “clipping”.

The most basic signal clipping circuit is an overdriven gain stage and these can exist in so many configurations that it is impossible to introduce them in a comprehensible and short form. As a generalization, the clipping characteristics of a gain stage mostly depend on the amount of negative feedback applied – the more the feedback is used in linearizing the circuit the harsher the clipping becomes. Field effect transistor circuits often use only small amounts of negative feedback (and that mostly degenerative type) and their transfer characteristics are also “rounder” than of BJTs. This has made them preferable devices for some people. For example, many Randall amplifiers create their heavy distortion tones by overdriving a cascade of FET gain stages. Similar circuits are also employed in numerous distortion pedal designs like “Dr. Boogey”. Some people, in fond of different tone, tend to prefer “fuzzier” BJT circuits that introduce a rather high amount of negative feedback (e.g. Fuzz Face).

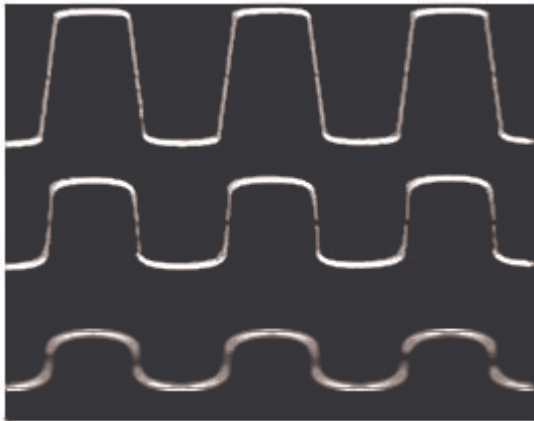


Figure 6.55 Oscilloscope capture showing clipping of sine wave when using LEDs (top) silicon diodes (middle) and germanium diodes (bottom).

Possibilities are virtually endless and some people like to get lyrical about “diode picking”. In practice, most notable effects result only from large shifts in threshold levels and transfer curves, i.e. from replacing silicon diodes ($V_F \approx 0.65$ V) with germanium or Schottky types ($V_F \approx 0.3$ V) or LEDs ($V_F \approx 1.5$ V). Introducing asymmetric clipping (different V_F for each half wave) also has a quite notable effect.

There are also other solutions to replace the ordinary diode element: Collector-base and emitter-base junctions of bipolar transistors are essentially diodes so using them as the nonlinear element literally falls into the category of diode clipping. Besides using BJTs, drain-gate and source-gate junctions of FETs work equally well. Fender Roc-Pro amplifiers utilized the same concept but used a dual triode with plate and cathode voltages of about ± 2 V. Circuit provides “tube clipping” without expensive transformers or lethal voltages. Will it sound like crunch of an all-tube amplifier – not really. The clipping characteristics are in fact quite “harsh”.

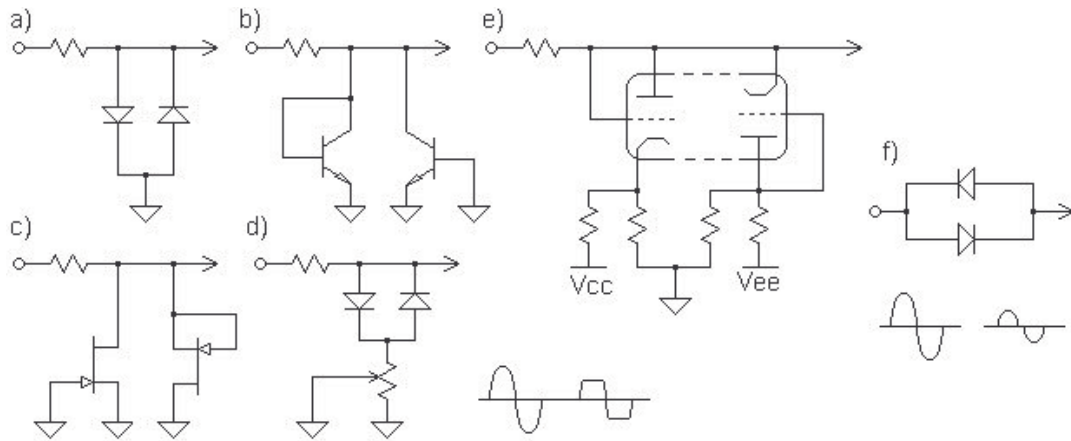


Figure 6.56 Passive diode clipping concepts: a) The basic configuration, b) conventional diodes replaced with diode junctions of BJTs, c) diodes replaced with junctions of FETs, d) increased series resistance, e) vacuum tube diodes made out of dual triode and f) crossover clipping.

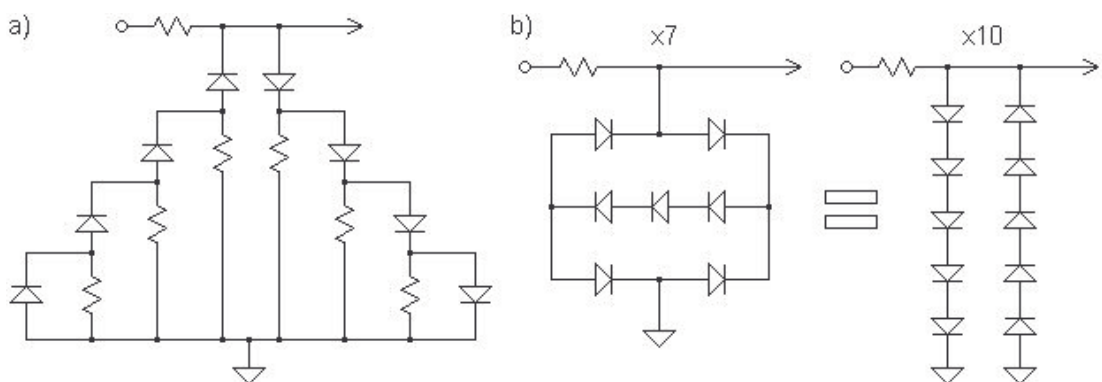


Figure 6.57 a) Resistor-diode ladder circuits can be used to replicate almost any imaginable transfer curve and b) Using “Marshall JCM900” –type diode bridge arrangement can save up few components.

Introducing some series resistance to the nonlinear element can increase the clipping threshold resulting into softer clipping. In creating a “drive” control this offers an alternative for controlling the gain of the previous stage. Utilizing the effect of series resistance in resistor-diode ladder circuits allows replicating very complex transfer curves. If the diodes are paired with series capacitors the clipping can be made frequency dependant. This is also the only way to acquire symmetric clipping if the signal is referenced to other than zero volts. Marshall JCM900 amplifiers use a diode bridge arrangement that essentially replaces a shunt configuration requiring six diodes with a similar configuration that requires only five. The concerned arrangement is even more useful if one wishes to go crazy with the amount of series diodes – especially if the diode string in the center is replaced with Zener diodes. This practice reduces board space and if amplifiers are manufactured in mass quantities also the saved costs from omitting few components become quite noteworthy.

The other example portrays some typical configurations of implementing a clipping element into the feedback loop. The active device does not necessarily have to be an operational amplifier; FETs, BJTs and even tubes work as well, however, since they usually have a notable DC voltage potential difference between collector and base, (or drain-gate, plate-grid whatever) the diodes should be AC coupled. The circuit’s operation is fairly simple: When the voltage at the output exceeds the forward voltage of the diodes they begin to conduct and limit the initial gain of the stage. In practice, the forward voltage of a device hooked to feedback drops a bit lower (i.e. 650 mV of

silicon diode converts to about 400 or 500 mV) and the transition to conduction is “softer” than in shunting configurations.

The clipping threshold of the active stage can be increased with an inclusion of a series resistor. The higher the value of this resistor is, the less signal passes through the diodes, which means that in practice there is an optimal value that the series resistor should not exceed in order to get any clipping at all. This is not the most favorable solution for all applications: The stage usually has a very high gain, which is reduced rather roughly when the forward voltage is exceeded. As a result the stage exhibits a very low overall gain and introduces crossover distortion in the form of enhancing the lower amplitudes. The higher the amplitude of the driving signal is, the less the clipping actually effects. This phenomenon is soon to be described in detail. *In practice, the circuit works perfectly with low series resistance values but in those cases it can never exhibit a very high output signal swing.*

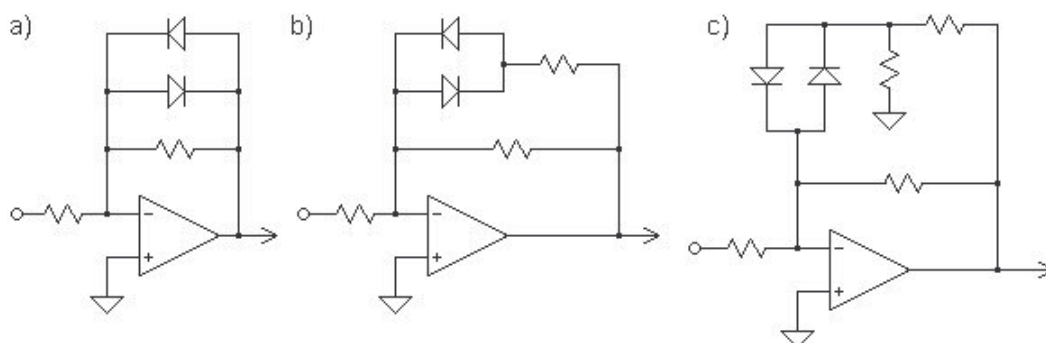


Figure 6.58 Signal clipping element in feedback loop: a) Clipping threshold near forward voltage, b) series resistor increases clipping threshold but adds crossover distortion, c) using voltage divider to increase output swing and clipping threshold and to prevent crossover distortion.

Circuit c) overcomes these limitations: The node connecting diodes to series resistor is turned into a voltage divider; as a result the output signal needs to exhibit a higher voltage swing in order to exceed the forward voltage of the diodes. At amplitudes lower than V_F the resistance of the diode network is nearly infinite and the parallel resistor defines the gain. For signals exceeding the forward voltage, the value of the series resistance essentially defines the gain. If its value is equal to input resistor of the stage the gain will be about unity. As a result, the closer the rails the output signal swings, the lower the gain will be, thus the signal clips softly. Vox used the idea in creating a softly clipping power amplifier for the Busker model. The power amplifier section in the concerned circuit was a TDA2030 power OpAmp. Reputedly the same circuit is used in one of Trace Elliott bass amplifier models as well.

Note that if the circuits are used in non-inverting configuration the wave is not clipped but “extended” from the middle section; thus *soft clipping applied in non-inverting configuration creates crossover distortion*. This crossover distortion is, however, dissimilar from the conventional form where all signal content lower than V_F is clipped. How come this happens? The answer relies in the differences between operational amplifier configurations: As was mentioned, when using inverting configuration, the potential at the inverting input acts as a “virtual ground”, where the other end of the diodes is basically shunted. However, signal passed through the diodes also decreases the closed loop gain thus causing the clipping to remain softer than in a case where the output node would shunt the signal to common in passive configuration.

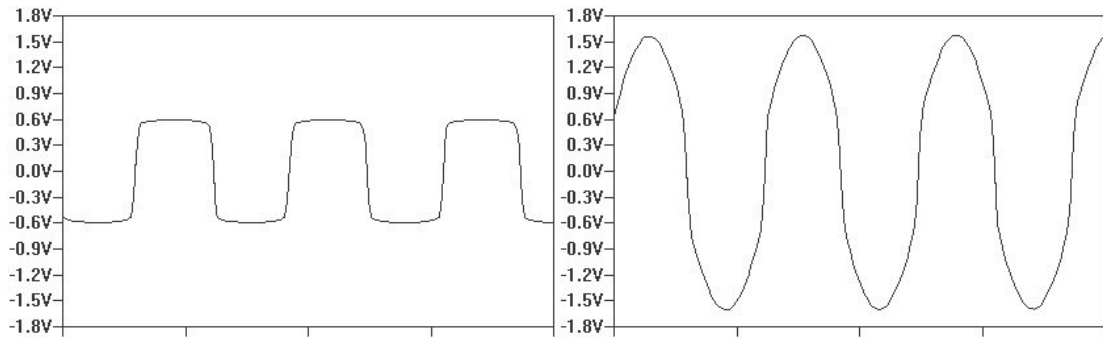


Figure 6.59 Plotted wave of diode introduced clipping in conveniently clipped signal or “normal configuration” (left) and enhanced gain of low amplitude signals causing crossover distortion (right).

In non-inverting configuration the input is not a virtual ground so the diodes are practically in series with the (feedback) signal path, however, they do not conduct until V_F is exceeded. Therefore, the alternative feedback signal path, which is (typically) the feedback resistor in parallel with the diodes, dictates the gain. When V_F is exceeded the circuit begins to introduce a lower gain. As a result, the lower amplitude portions of the signal get overly emphasized. Removing the parallel resistor would result into a nearly infinite gain with amplitudes lower than V_F .

With low amplitude input signals the clipping of both circuit configurations sounds nearly the same. If plotted, the harmonic spectrum from both circuits is actually nearly identical with the exception of stronger emphasis of the fundamental harmonic in non-inverting configuration. When the input signal’s amplitude is increased the signal from the inverting configuration gets more and more “compressed” while the signal from the non-inverting gets “cleaner”. However, it will exhibit an annoying “buzz” – similar to one heard when a radio station is slightly off tune. After serious “overdriving” the signal will finally hard clip to rails. However, the rail clip threshold is met only when input amplitude + V_F equals the OpAmp’s maximum swing voltage towards rails.

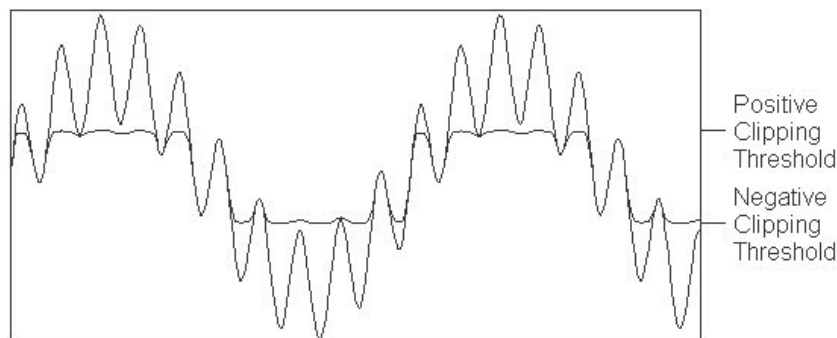


Figure 6.60 Blocking distortion “masks” the high frequency signal content.

With all clipping circuits the pre/post tone shaping is essential and in practice, it has a major effect on tone. This is a point that cannot be emphasized enough. Design of good-sounding clipping circuits is mainly *filter design* and the way in which the signal is clipped usually has only a secondary importance. The “voicing”, as this tone shaping procedure is called, is a very subjective process and therefore lacks major guidelines. However, it *does* make a significant difference. A usual configuration is to attenuate low frequencies before clipping so that the annoying amount of blocking distortion is minimized (clipping low frequency content “masks” the high frequency content). For the same reason it is fairly common to boost up the top end to preserve the high frequency harmonics that give definition and character to guitar’s tone (think

treble boosters). In “classic hi-gain” distortion circuits that consist of multiple overdriven stages the same effect is acquired by gradually increasing the amount of bass frequency coupling towards the output stage. Sometimes complete “bandwidth reduction” (meaning drastic cutting of low- and high frequencies) before clipping is preferred instead. A very common processing method is to attenuate high frequencies post clipping in order to remove the sharp edges of the clipped signal. This removes the “fizzy” “cricket” or “razor” –like frequencies, which are considered unpleasant by most people. These are typically frequencies above 3 kHz and therefore *the speaker plays an important part in smoothing the clipping tones*. Some prefer boosted bass post clipping. In addition, creating notches and peaks at certain frequencies (post and pre) is a certified way to make the distortion tone sound more “professional” and “unique” rather than annoying and sterile fuzz box buzz of the sixties. It is, for example, quite typical to boost frequencies around 1.5 – 4.5 kHz, just before the steep low-pass filter starts to have effect. Many speakers inherently introduce a complex comb filter at right about these frequencies, which has a huge effect to distorted signals.

Since it’s virtually impossible to describe tones of various filtering configurations on paper, I highly recommend experimenting with two graphic equalizers that have a signal-clipping element inserted in between them. You will note that even when the clipping element is as simple as a two-diode shunt limiter you will be able to dial in hundreds of different tones. You will also find out that you can get a lot more variation this way than by doing the “conventional” modification of tweaking only the clipping devices. When you have dialed in the desired frequency responses replicate them with the required circuitry. However, don’t be amazed if the result is very complex: This is just the nature of the game. Since variations in lower- and upper mid-range response generally tend to have a greater effect after the clipping stage it is also quite easy to see why some amplifiers are more “pedal friendly” than others.

A slight peek in history is also in order here: During the 60’s using distorted tones became popular but up to this point the amplifier design had mainly focused on how to *minimize or remove* the distortion completely. When musicians began to demand that their amplifiers actually needed to distort, the (mostly) conservative designers were pretty much amazed and clueless. To serve the demand they harnessed the new solid-state technology to produce distorted and buzzing tones that were thought to be what musicians wanted. Mostly they were not. Naturally these simple circuits could not replicate the complex processes that created the distorted tones of tube amplifiers either. Most of these processes were even unknown up to the early 90’s! For defense, I must say that many fuzz box effects did, however, create unique tones that had not been heard before.

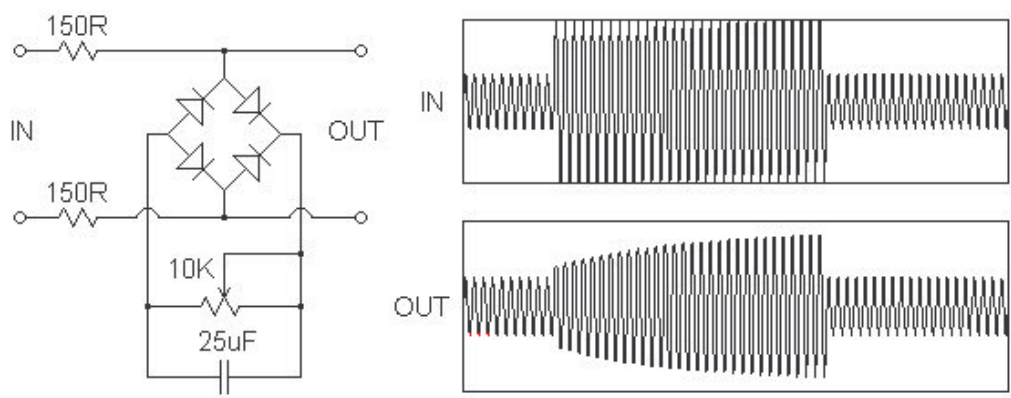


Figure 6.61 Crowhurst’s “Fuzz”.

Finally, let's take a small look at some "oddities": In "Electronic Musical Instruments" (published in 1962) Norman H. Crowhurst criticizes the "new" craze of achieving apparent loudness by distorting the signal calling such a fad "temporary insanity". Crowhurst presents a quite unusual circuit to produce the desired "fuzz" tones: The audio signal is shunted with a full wave rectifier that charges a capacitor. A resistor in parallel to the capacitor is used to allow a rapid discharging; the resistance also controls the amount of clipping. The surprising feature of the circuit is the "attack": The amount of clipping is gradually decreased during sustained overdrive. In practice, the amount of attack (or the rate of clipping) is controlled by the capacitor's charge, thus the circuit will react quite interactively to playing dynamics: Acquiring the heaviest clipping requires tranquil passages that allow the capacitor to discharge. Unfortunately the circuit is not conveniently single-ended and requires special drive arrangements like use of audio transformers or balanced circuitry. Crowhurst's method to produce "fuzz" was pretty much buried in silence.

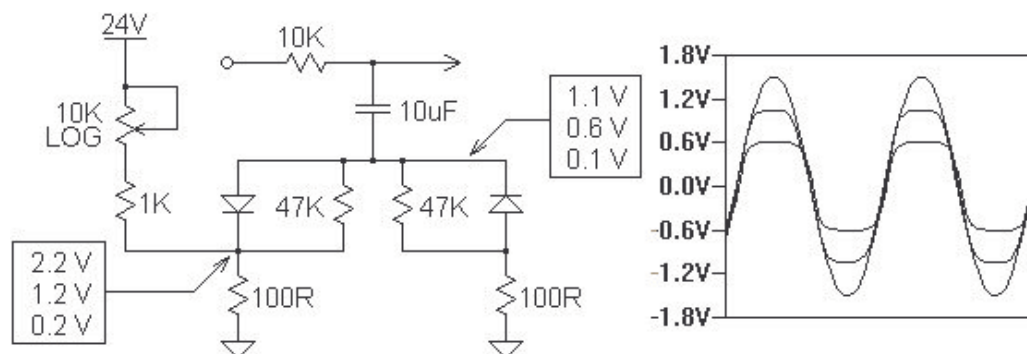


Figure 6.62 The "Peak limiter" circuit of Thomas Organ Vox amplifiers.

The "peak limiter" circuit, employed in many Thomas Organ Vox amplifiers, had a much better success. While the circuit is quite ingenious and even managed to attain a quite "legendary" status it still failed to gain a greater popularity. The operation of the circuit is very basic: A potentiometer that determines the amount of limiting varies the DC voltage levels of the shunt circuit. Because this variation is essential part of the circuit's operation the shunt path must be AC coupled with a 10-μF capacitor.

At its lowest resistance the potentiometer and the following kilo-ohm resistor create the highest-possible voltage potential (about 2.2 V) at the node of the resistive divider that is formed in conjunction with the 100-ohm resistor. A proportional voltage potential is coupled to the shunt circuit's input through the voltage divider formed by the two 47-kilo-ohm resistors (the 100-ohms at the end is low enough to be ignored). The AC signal rides atop this DC potential. Consequently, the threshold of signal clipping for both half waves is approximately 1.7 V. When the potentiometer is turned to its maximum resistance the voltage potentials defined by the resistive dividers change: They are now only about 0.2 V and 0.1 V, respectively. The change naturally reflects to clipping threshold, which is decreased to about ± 600 mV for both half waves. Potentiometer settings between these two extremes vary the clipping level accordingly. Logarithmic potentiometer is required to obtain a linearly changing threshold adjustment.

We must acknowledge that different applications require different clipping behavior. Typical instrument amplifiers try to achieve "musicality", which often translates into requirement for soft and "transparent" clipping by the means of having a smooth transition to operating area with limited dynamic range. In contrast, many PA

amplifiers try to achieve linearity and in them the dynamic range is often limited for *protective* purposes only – for example, to prevent too high signal amplitudes from destroying the input stage. Thus “musical” tone has secondary importance when compared to reliable operation of the amplifier.

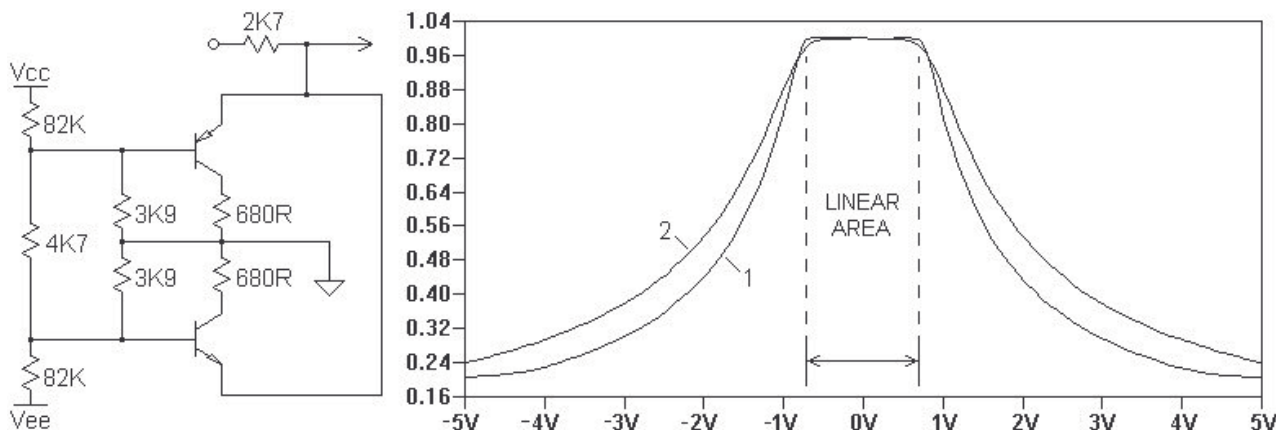


Figure 6.63 Otis PA amplifier’s limiter and its difference to diode clipping. The plotted curve at right shows gain versus input signal amplitude for 1) Otis PA limiter and 2) two series diodes connected anti-parallel.

Figure 6.63 shows an input signal range limiter circuit from Otis PowerStation public address amplifier. One of the circuit’s merits is the great tweak-ability. Let’s compare this circuit to a rather typical clipping arrangement of guitar effects and amplifiers, which consists of two series shunting diodes connected anti-parallel: The PA amplifier limiter circuit has less smooth transition to clipping (harder knee) to keep the signal longer within the linear region. When limiting takes place it has a rather steep curve, which ensures proper attenuation of signals exceeding the input range. No doubt this will create a harsher clipping, which is often considered less “musical”.

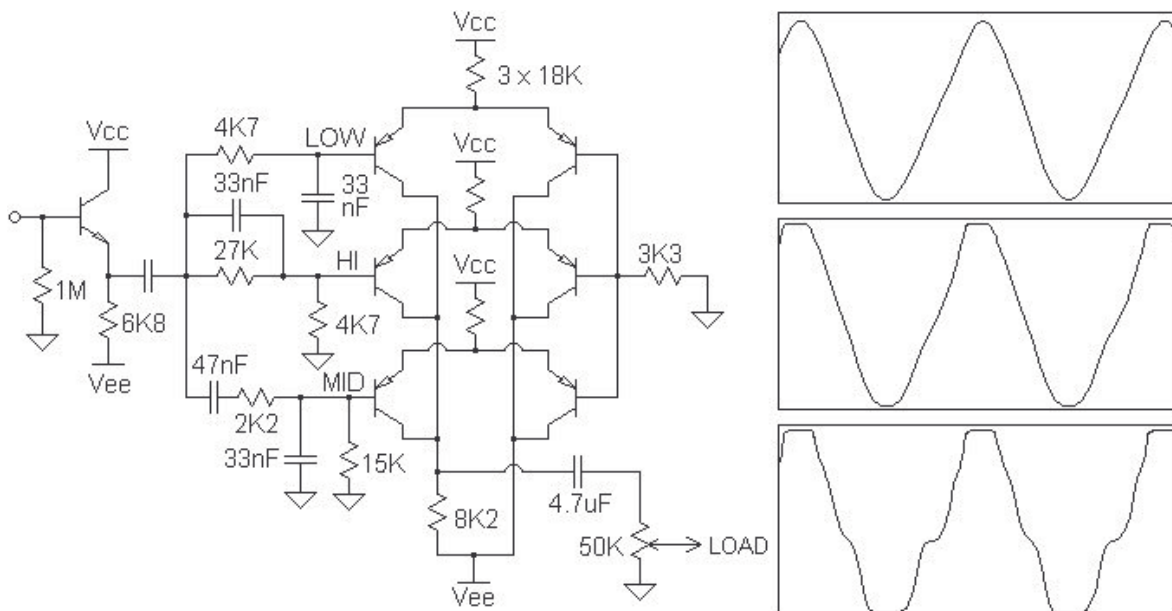


Figure 6.64 Traynor-style clipping circuit introduces tube-like “soft” crossover distortion that increases along the level of overdrive.

Another peculiar circuit can be found from old Traynor amplifiers. This circuit, based on three differentials, emulates the soft crossover distortion of overdriven class-B/AB tube amplifiers. The circuit was introduced in the late 70’s. Peavey introduced (and patented) the same concept only few years later but their circuit was based on an actual class-AB gain stage. A refined version of the circuit (which was also patented)

is utilized in the new TransTube amplifiers. Crossover distortion often sounds downright horrid but when it occurs only during signal overdrive (and in a soft manner) the tone is – subjectively – rather ear pleasing and the crossover components quite hard to perceive (as opposed to crossover distortion introduced by a badly biased amplifier). Thus attempts to replicate these less-ideal circuit characteristics are somewhat interesting. I suppose they add their share of “reality” to the emulation. Another similar example is replicating the signal modulation caused by higher amount of supply rail ripple. For example, in Pritchard amplifiers this characteristic (and the ghost notes that it introduces) is called “fat”.

6.12 Soft limiters

Limiter circuits can be roughly divided into two categories: Ones that are controlled by an amplitude detector circuit and ones that are not. The element driven by the detector circuit, which is usually a LDR, FET or OTA (Operational Transconductance Amplifier), can either be a part of an attenuating or gain inducing circuit. Some limiters are based on integrated circuits that often can work either as dynamic range enhancers or limiters (compressors). These chips are usually versatile and perform well but have a tendency of becoming obsolete as new models are introduced. This book will not discuss these ICs and if you need more information about them the best practice is to look for it from the corresponding datasheets. The solutions presented here are conventional and mostly work with ordinary, discrete components.

A word about semantics: Voltage-controlled amplifiers (VCA) have a gain relative to a control voltage. A sub topology of them, automatic gain control amplifiers (AGC), use either feedforward or negative feedback to automatically adjust the gain to keep the signal amplitude within certain limits. Practically all limiters that are controlled by detector circuits are therefore AGC circuits and the detector is their feedback circuit. This also means that a very similar circuit – with detector replaced - can be used in, for example, building a digital volume control, a mute circuit or other similar circuits.

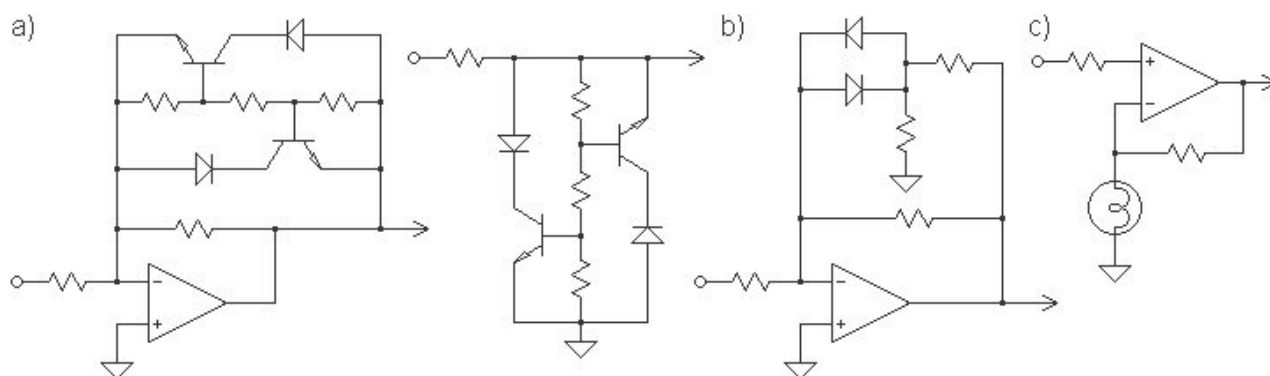


Figure 6.65 Soft limiter circuits without detector: a) Active clipping control in inverting active and non-inverting passive configurations, b) adjusting the clipping region with voltage divider and c) lamp limiter.

Limiters without detector circuit are mainly based on diode clipping principles. Practically there is a thin line defining, which circuits exhibit clipping and which ones soft limiting, as both terms may essentially mean the same thing. The Figure 6.65 displays some examples of circuits commonly considered as “limiters”. The circuit a) has been used as a “compressor” in some Fender and Music Man amplifier models. Some Crate amplifiers utilize it as a distortion circuit. The circuit varies the voltage potential across collector-emitter junction according to drive signal at transistor’s base. Basically, the circuit is nothing else but our well known “rubber diode” (VBE-

multiplier) configured to control two anti-parallel “transistor diodes”. Conventional diodes are also fitted in series with the collectors to prevent reverse conduction. The circuit is usually seen in conjunction with a unity-gain inverting amplifier – however, it naturally works in the passive shunting form as well. When passive, the circuit is non-inverting – just like an ordinary diode clipper. Properly tweaked, the circuit can accept a huge range of input amplitudes without going into hard clipping.

The circuit b) is already familiar to us from previous examples. I have included it to the list because it practically falls into the category of soft limiters. In comparison to circuit a) this one usually exhibits higher gain (therefore working with lower input amplitudes) and also swings closer to rails. The range for input amplitude is not as high as in circuit a) and if overdriven too much the circuit will clip hard to rails.

Circuit c) exploits a very old idea: When the filament wire heats up due to current flowing through it, its resistance increases limiting the current. Series in signal path the light bulb will act as a current dependant resistor. For example, in some Leslie organs two light bulbs were used to limit the drive signal of the reverb tank. Lamps can also be used in creating an active stage with a variable gain and in this form they have been commonly used to tame down oscillator circuits. The ordinary problems with light bulb limiters are lifetime, cost (typical lamp is usually more expensive than few transistors and resistors) and, even more importantly, the variety of different lamps: The commonly used 6 V or 12 V lamps have pretty inconsistent resistance versus current characteristics so not all lamps work equally in a specific circuit. The most significant factor defining these characteristics is of course the power rating of the lamp. Light bulb limiters also have very slow attack and decay since these parameters are essentially defined by thermal inertia. Most common lamp type used in these circuits is a so-called “grain of wheat” lamp that has ordinary leads and therefore requires no lamp holder socket.

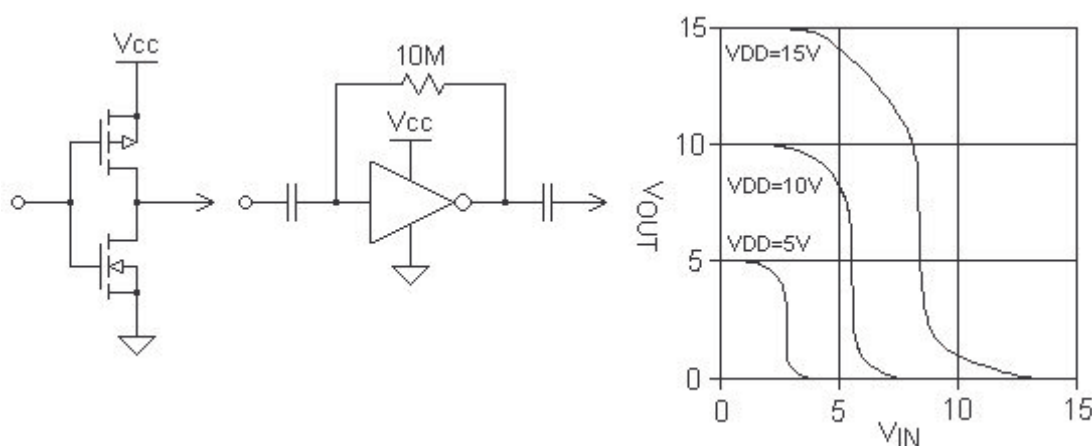


Figure 6.66 CMOS Inverter configured as amplifier and transfer characteristics of a 4007UB CMOS inverter buffer.

Another quite famous soft limiting technique is using a CMOS inverter as an amplifier. If a high value resistor is wired between the input and output the circuit will automatically bias itself to center of the supply. Counter intuitively the circuit exhibits a higher gain with lower supply voltage. For example, V_{CC} of 5 V provides a gain of 40 dB (x 100 V) on a certain chip. With V_{CC} of 10 V the gain is about 31 dB (x 35.5 V) and with V_{CC} of 15 V about 29 dB (x 28.2 V). The clipping characteristics of most logic circuits are very soft and practically softest with supply voltage of 15 V. Using a suitable input resistor in front of the circuit allows tailoring the “drive” characteristics. It is controversial who is the inventor of this circuit: It was successfully used at least

in Sunn “Beta” lineup (circa 1977) and Electro-Harmonix Hot Tubes (the original, solid-state version from 1978). However, the undisputable fact is that the idea was published and popularized by Craig Anderton around the same time that first commercial products became available.

While examining old designs we may pick up an interesting one from the very late 70’s: Figure 6.67 depicts the “Gate Induced Valve Effect” (G.I.V.E.) circuit from Gallien-Krueger amplifiers. The circuit uses a JFET differential to drive an OpAmp-based voltage amplifier stage. Introducing some local negative feedback in the OpAmp stage can reduce the open loop gain of the circuit. This subjects the differential itself to clipping. If designed with care, the circuit can provide very soft clipping behavior. While this is just one of the many soft clipping schemes that are out there, the novel part was that distortion channels of Gallien-Krueger amplifiers were constructed from *cascades* of these circuits, thus they were able to replicate the performance of “hi-gain” tube amplifiers that also employed cascaded gain stages and careful voicing within. On top of that, the circuit was even used as the “OpAmp” driver of the power amplifier - that, of course, with the aid of a proper output voltage-boosting scheme.

Compared to limiters operating on clipping principle the detector-based circuits have a finite speed of reacting to amplitudes exceeding their threshold. This variable is known as “attack”. The slowness of either the control circuit or the limiting device itself usually limits the attack time to about 10 – 30 ms (100 Hz – 33.3 Hz). A typical LED-detector circuit combination has a response time of about 15 ms so it’s a lot faster than a light bulb. Although not extraordinary, this is usually fast enough for guitar audio applications. Limiters for fast speed applications like AD converters would naturally require faster attack times. As a generalization, FET based limiters are “sterile” and operate faster while LDR-based limiters are slower, “smoother” and “color the sound”. In practically every case a compressor/limiter adds very distinctive colorization to the signal – especially when chained up with a distortion effect. Often the tone of this dynamic processor is more distinctive characteristic than the tone of the distortion unit. One can definitely “hear” compressors/limiters.

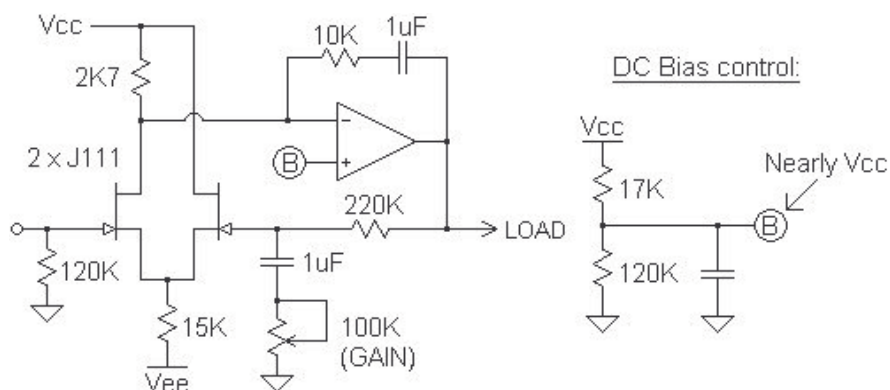


Figure 6.67 Early Gallien-Krueger-style soft limiting circuit.

Note: People with trained ears can even distinct some radio stations from each other due to their characteristic compressor tones. Compressors these stations use undoubtedly professional quality and operate on multiple frequency bands in order to be as transparent as possible. They are also set-up professionally. This example should bring home a point that simple compressor circuits that are used in guitar amplification devices are typically very audible.

Detectors for LDR limiters mainly work by two ways: They either introduce high enough amplitude to drive the LED directly or control a transistor that uses the LED as collector or emitter load. The latter is more advisable configuration. A component combining both LED and LDR is known as a “Vactrol” and using one might be easier than wrapping a LDR and a LED inside a black heat shrink. The important parameters for the design are LDR element’s maximum-, typical- and dark resistances. It is important to note that in dark the element slowly *saturates* and after about five seconds the dark resistance has usually grown around three times higher. The speed and severity of this characteristic varies and this greatly effects the decay time of the circuit causing different Vactrols and LDRs to have different tones. The dark resistance values for typical LDRs range from 5 Megaohms to 180 kilo-ohms when saturated (value of few Megaohms being more common) and from one Megaohm to 50 kilo-ohms after being only about 1 second in the dark (few hundred kilo-ohms being the more common value). The minimum resistance (in light) is usually from 100 kilo-ohms to few kilo-ohms.

It is claimed that LDR elements have the best response with yellow LEDs and this is in most cases quite true. However, do not take it as an absolute truth. The highest sensitivity of a LDR element is reached at its *spectral peak*, which is the wavelength of light to which the LED reacts the most sensitive. This is often around 500 to 600 nm. Visible light covers wavelengths from about 400 to 700 nm and the wavelength of yellow light is about 570 nm. Green light would have a wavelength of about 510 nm and an orange light a wavelength of about 590 nm. Yellow falls in between so it’s often a good match to LDRs. The wavelengths for violet, indigo, blue and red lights are 400, 445, 475 and 650 nm respectively. Note that there are many shades of colors so the wavelength is indicated only for a *pure* color. In practice, if you want the best match you have to find out the spectral peak of the LDR and the wavelength of the concerned LED. Some LEDs and LDR elements work at infrared region, which is a reasonable demand for applications that must work while exposed to light.

Table 6.5 Properties of standard LEDs.

Color	I _F max.	V _F typ.	V _F max.	Intensity	Wavelength
Red	30 mA	1.7 V	2.1 V	5 mcd@10mA	650 nm
Bright red	30 mA	2.0 V	2.5 V	80 mcd@10mA	625 nm
Yellow	30 mA	2.1 V	2.5 V	32 mcd@10mA	590 nm
Green	25 mA	2.2 V	2.5V	32 mcd@10mA	565 nm

Methods that replace LEDs have been used with success as well: Before LEDs became as common and cheap as they are now light bulbs were used instead of them. A light bulb has a slower decay time than a LED, which usually “turns off” quite suddenly after driving voltage has decreased past V_F. However, filament of a light bulb is basically too slow to react to transients. Another alternative is an electro-luminescent panel. Some claim that these give a better compression tone than LEDs. Perhaps so, but the fact is that electro-luminescent panels are quite expensive, operate on tricky voltages and have a finite life: That is, they deteriorate slowly in use and loose some of their luminescence. It is true that the older, “classic” compressors used them but it might have been for the sole reason that they were the only option at a time when LEDs were still considered hi-tech.

In FET-based limiter circuits the field effect transistors acts as voltage-controlled resistors. In this mode the FET operates in triode region and it has to be “linearized” by adding half of V_{DS} to V_{GS}, this is done with resistors R1 and R2. The procedure

ensures that the FET's resistance R_{DS} is not dependent on V_{DS} . $R1$ and $R2$ also decrease the control sensitivity and their resistance should be kept high since they effectively shunt R_{DS} . Source impedance (not meaning FET's source) must be kept low or value of $R1$ has to be tweaked to compensate it. Since the FET's source practically needs a steady potential reference (i.e. common) it is tied to virtual ground at the inverting amplifier's input. $R3$ is a parallel resistor that can be used in defining the gain. Since R_{DS} is usually quite high the resistor $R4$ usually has value of only few kilo-ohms. It is advisable to drive the circuit with a buffer amplifier.

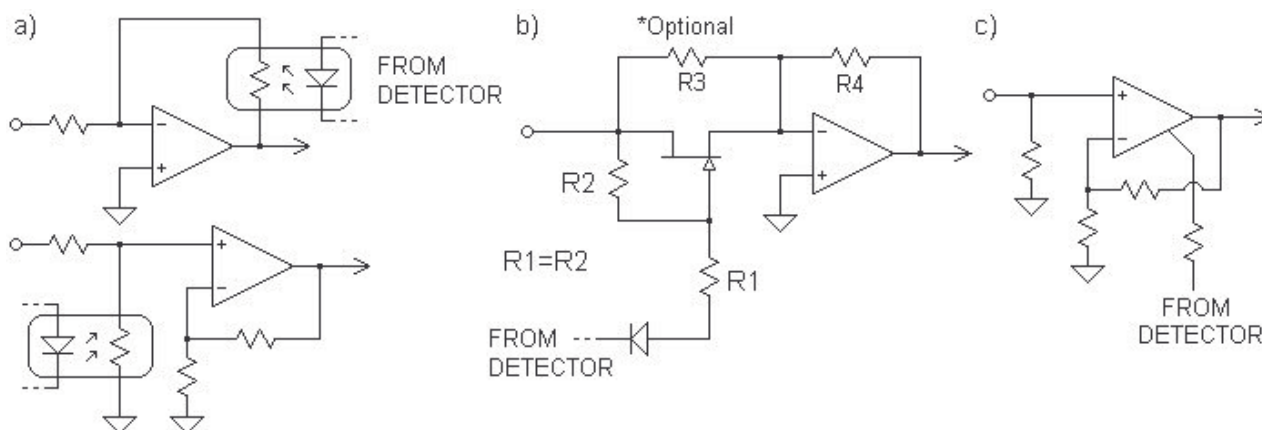


Figure 6.68 Limiter circuits requiring a detector: a) optical LDR-LED or “Vactrol” limiter (active and passive), b) FET limiter using N-channel junction FET and c) operational transconductance amplifier.

Note that both amplifier and detector circuits have to be tweaked according to different FET characteristics: Some FETs have higher R_{DS} and exhibit more decrease in it with lower signal amplitudes from the detector than others. Also, remember that this is just a mere example of how a FET behaves as a voltage controlled resistor and other, more sophisticated, circuit topologies will likely provide better performance.

Basically the circuit would work in a passive form as well but the greatest limitation is the interaction with signal from the detector: The FET's resistance is higher the more negative the signal fed by the detector circuit is. Since the FET's source would have to have a common connection the only effective resistance divider configuration would cause *less* attenuation when detector introduces higher control signal amplitudes. Compensating this behavior would introduce a lot of unnecessary complexity to the detector circuitry. Anyway, the active configuration has better SNR and likely even operates better so why bother.

Note that the control voltage of the FET, fed by the detector circuitry, is negative. The diode shown in the schematic is usually part of the detector and shown merely as a reminder of this. In practice, acquiring either negative or positive control signal is just a matter of diode orientation arrangement.

Implementing voltage controlled gain around operational transconductance amplifiers, or OTAs, is also a rather common approach. An OTA is a voltage controlled current source (this is sometimes indicated in the schematic symbol by drawing a controllable CCS into the “tip” of the OpAmp), while a conventional operational amplifier is a voltage controlled voltage source. The input voltage controls the output current by means of device transconductance, or g_m , and what makes OTAs so versatile is that the transconductance can be varied with external current known as the amplifier bias current I_{ABC} . The output current of OTA is the return of g_m multiplied with voltage difference between the inputs. Since the OTA is a current source it, on a contrary to

conventional OpAmp, has high output impedance and can be used in a configuration that employs no negative feedback. To compensate the high output impedance the chips usually contain an “on-chip” discrete buffer.

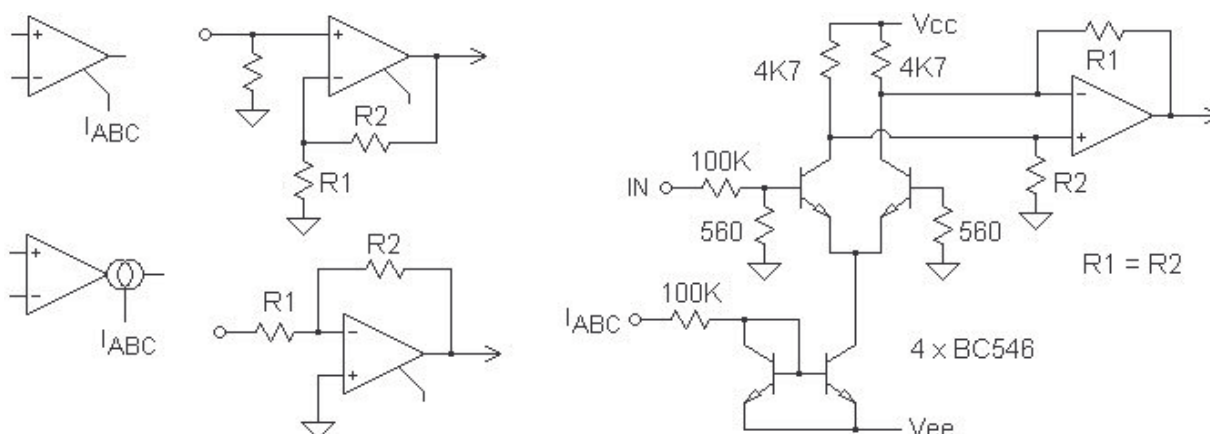


Figure 6.69 Typical OTA configurations and schematic symbols. Discrete implementation to a conventional OpAmp is shown at right.

The usual VCA configurations employ OTA in one of the typical OpAmp configurations where negative feedback is used to set the gain and decrease output impedance. It should be noted that OTAs operate on a very small input signal range and in order to avoid serious distortion the input amplitude must usually be kept way under $1 V_{PEAK}$. The voltage gain of OTA in inverting configuration is:

$$A = (1 - g_m \times R2) / (1 + g_m \times R1) \quad \text{(Equation 83)}$$

Common OTA models are NE5517 from Philips, CA3080 from RCA, Intersil and Harris (an astounding amount of compressors, limiters and other VCAs is based on this chip), SSM2024 from Analog Devices and LM3080 or LM13600 from National Semiconductors. Do not confuse OTAs with “programmable OpAmps” that have their *operating current* set by an external resistor. These are completely different species of OpAmps and the operating current is a parameter that does not affect gain.

There is a lot that can be learned from old analog, synthesizer technology, as it has practically been the “playground” for high-tech analog audio signal processing. Most of these synthesizers sport circuits that average amplifier designers do not even know to exist. To get back on topic, old analog synthesizers used VCAs that were made out of discrete transconductance input stage. This is basically a differential circuit that has its constant current source load controlled by the detector signal. It has also been the basis for *all* IC operational transconductance amplifiers so far. The input of the differential was usually fed to conventional OpAmp but it can be preceded with a discrete circuitry that forms an OpAmp as well. Even nowadays the circuit sees plenty of use as a “mute” control implemented straight into power amplifier’s input stage. Sometimes the very same “mute” circuitry is harnessed to gain controlling. All OTAs shown in the examples work by lowering the transconductance the more negative the control signal (V_{ABC}) gets.

The *detector* circuit is the element controlling the gain control circuit. In most limiter configurations the detector monitors the signal amplitude and varies its output signal according to it. In monitoring, detectors utilize either feedforward (signal sampled before VCA) or feedback (signal sampled after VCA). In practice the detector circuit is usually a combination of a rectifier and a filter (capacitor) that produces a higher

DC voltage the higher the signal amplitude is. To work correctly the feedback path should not load the signal path and it should work fast enough to respond to amplitude peaks. The response speed is defined by “attack”, which is the delay before limiting starts to have effect. Since the signal is usually filtered, the capacitor’s charging time constant mostly defines the attack. It takes a finite time for DC voltage to decrease as well: The speed of this change is known as “release” and it is a great part in defining sustain of the limiter. The time constants of both characteristics are always somewhat a compromise. Controlling the detector’s output amplitude will vary the rate of limiting operation.

Figure 6.70 shows two examples of detector circuit implementations. Circuit a) is active and uses an OpAmp-based half-wave rectifier. These circuits are also known as precision rectifiers because they work on very low forward voltage. In concerned application the greatest merits of these circuits are that they introduce gain and act as a buffer between the VCA controls and sampling point. The output of the rectifier charges the capacitor C_A through resistor R_A . The RC time constant determines how fast the circuit reacts to transients. Due to finite attack time, a limiter will typically introduce some clipping distortion when a transient occurs after a quiet passage. During lower amplitude output from the rectifier the capacitor discharges through R_D and limiting circuit in series with R_{CL} . The time constant defines the “release” characteristics. This is another drawback of typical limiters: Quiet passages after loud ones are attenuated because the capacitor has a finite discharging time. The capacitor can be removed to shorten both attack and release times but it has been observed that circuits like this do not work very well in practice. Since R_A and R_D form a voltage divider it is beneficial that R_A is considerably smaller than R_D . Resistor R_{CL} is a current limiter that is needed in some cases.

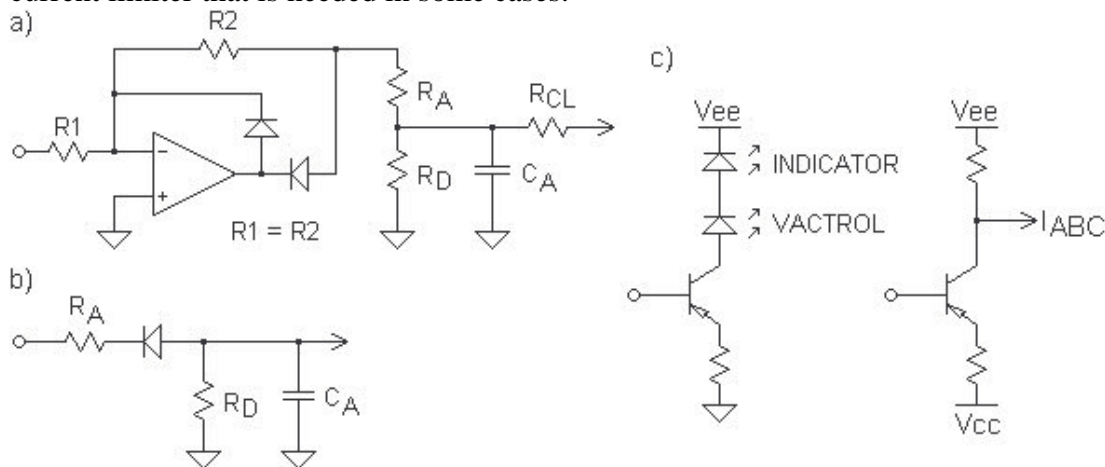


Figure 6.70 Detector circuits: a) Detector with active half-wave rectification, b) detector with passive half-wave rectification and c) transistor interfaces for driving LEDs or OTAs.

Circuit b) is one of the simplest passive detector circuits: With low amplitude sampling point it will require a gain stage to boost up the control signal amplitude and to cover the losses introduced by the diodes and resistors. However, if the sampling point is, for example, the output of a power amplifier there is likely no need for an amplifying rectifier circuit. The circuit is quite identical to the active one with few exceptions: Resistor R_A is also the current limiting resistor and it also defines the loading impedance along with R_D . In high impedance circuits the current is in most cases very small so they likely require an interface made out of a transistor - or similar device – to drive LEDs.

Circuits in c) are suitable interfaces for LEDs and OTAs. They buffer the controls and amplify current of high impedance detectors. FETs are driven by voltage and most OTAs can work with a control current of few microamperes, which means that they likely will not require an interface – other than for buffering purposes. LEDs are another deal since they require a current of few milliamperes. Depending on the LED’s color the forward voltage will vary – so does the reaction of the LDR element.

Naturally, resistors R_A and R_D can be replaced with multiple variations of potentiometers and their series resistors. Controlling the gain of the feedback path, or the ratio of the voltage dividers will of course alter the limiting threshold. The speed and precision of both circuits could be increased by substituting the half-wave rectifiers with full-wave ones and by buffering the control outputs. I leave it up to the designer to figure out how to implement these features - or to ponder whether those are even necessary in the concerned application.

Since majority of the VCAs work with negative control voltage the examples of the detector circuitry are shown using it. At zero volt output there should be no limiting introduced. Naturally, at a certain point either the detector circuit or the VCA will saturate. I have not included component values to the circuits since one should always derive them individually for the concerned application. The values are dependant on required limiting threshold, sought attack- and release times, the maximum voltage and current limit of the VCA, and signal amplitude at the “sampling” point. Once you understand how these circuits work deriving the values should be fairly easy.

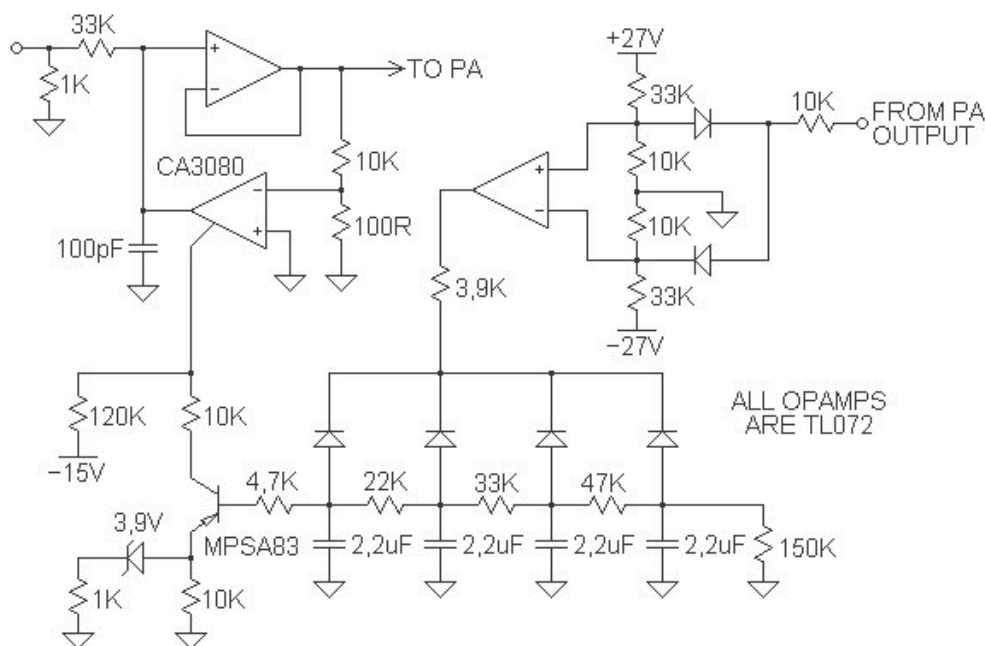


Figure 6.71 Real-life OTA-based limiter realization: “DeltaComp Limiter” circuit from Fender Frontman 25B.

6.13 Tremolo circuits

Tremolo as an *analog* onboard effect is becoming very rare these days. However, it was fairly popular up to the mid 70’s. For starters, let’s make a distinction between tremolo and vibrato: Tremolo is an effect that modulates the signal amplitude while a vibrato modulates the signal’s frequency (“pitch”). The outcome is essentially the same but in practice a varying pitch effect is harder to realize with electronic circuits. Therefore, when you see an amplifier with a model name hinting to vibrato function

you can be 99 % sure that it actually is a tremolo instead. Some amplifiers *do* have an actual vibrato effect but this is pretty rare.

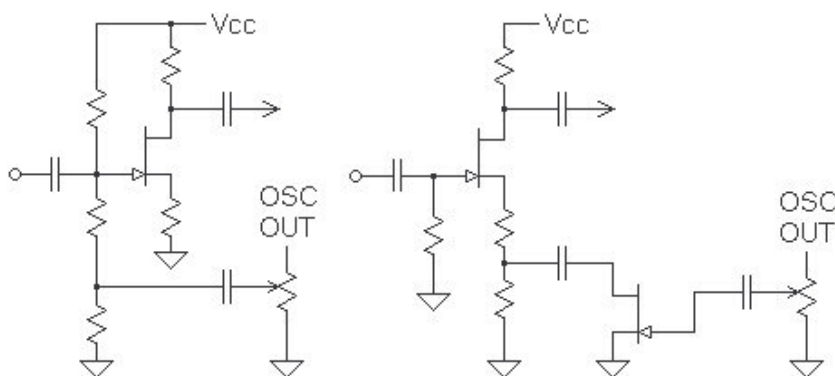


Figure 6.72 Two common ways to apply tremolo. Left: Varying gain by varying bias level and right: Varying gain by varying resistance of AC feedback.

Typical tremolo has a frequency that is variable between 2 to 15 Hz. Generally the heart of every tremolo circuit is a low frequency oscillator that controls a gain-controlling element. Early vacuum tube and some FET circuits employed a solution that controlled the gain by varying DC bias level of the grid or gate. When the output signal from the tremolo circuit is fed to the bottom of the grid/gate circuit through isolating resistors the existing bias voltage is affected by addition and subtraction of the signal from the tremolo. This will result into variation of gain. A common problem from this approach is that a very low frequency AC component is introduced to the signal. The component is inaudible but the audible signal “rides on top of it” causing anomalies like earlier clipping or modulation of noise and hum that can be heard as annoying “wobble” when no guitar signal is present.

Old tube amplifiers battled the problem by mixing the tremolo signal in the phase inverter stage, this created two opposite phase signals that cancelled each other out when no guitar signal was present. However, there are more sophisticated solutions: Most oscillator circuit outputs are perfectly capable of driving LEDs (LDR as gain-controlling element), FETs and - with some modifications - OTAs.

Sine wave oscillators, like the one depicted in Figure 6.73, are the heart of most tremolo circuits. The presented circuit is likely the most simple one can use while still having the most important features like control for oscillation frequency (“speed”) and amplitude (“depth”) - along with a toggle control that in most basic form just shunts the feedback loop to ground causing the oscillation to seize. The general operating principle of most tremolo oscillators is to delay the feedback signal with adequate phase shift. Outside the guitar amplifier electronics world these circuits are therefore known as *phase shift oscillators*.

A phase shift oscillator circuit will begin to operate when its gain is higher than unity and total phase shift exceeds 360° . A typical circuit will provide the 180° phase shift (inversion) inherently and achieve the rest with a combination of RC *phase lead* or *phase lag* networks. The shown schematic uses the first configuration; swapping the positions of capacitors and resistors would create the latter. Because theoretical phase shift from one RC network is only 90° a minimum of three networks is required. When examining the example schematic one notices that the phase lead circuit is basically just a high-pass filter and the phase lag circuit correspondingly a low-pass filter.

The -3dB point of the filters corresponds to a phase shift of 45 degrees. The corner frequency is same for both filters and derived using the basic equation $f_{3\text{dB}}=1/(2\pi RC)$. Derived from appendix B of [60], the phase shift of three RC networks is 180° at the frequency of $f_{\text{osc}}=1/(2\pi 6RC)$. The equation assumes that capacitor and resistor values in all networks are identical. Feedback factor for such a circuit is $1/29$ so to start oscillating the amplifier must have a gain of 29 or higher.

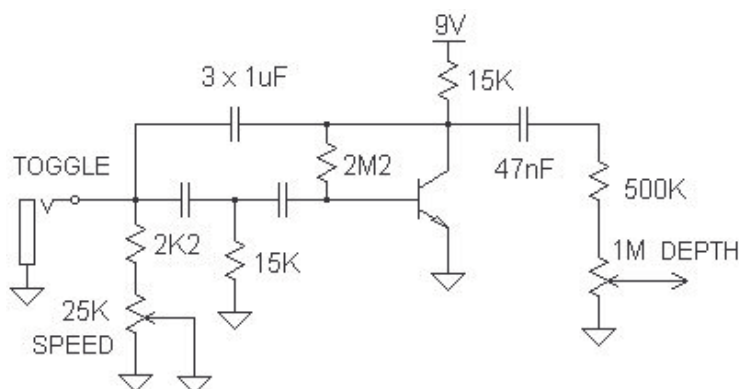


Figure 6.73 Typical oscillator circuit.

The effect of RC networks is attenuating and has to be considered when setting up the closed loop gain. To minimize distortion the oscillator amplifier should operate at exactly unity gain. In reality this is hardly possible due to inherent variations in device gains and even more: Adjusting the frequency of the oscillation may also change the RC network's level of attenuation. Therefore the gain is usually set higher and the output of the circuit filtered to remove distortion harmonics. If the amount of RC networks is increased the total attenuation of the loop will decrease.

For wider range of control the adjustment for oscillation frequency should be located in the last RC filter section rather than the first one. The potentiometer should have no larger range than about five times the nominal resistance value – otherwise at the extremes of the rotation the oscillation may seize. This limits the frequency variation to about 3:1. Loop should naturally be designed to provide oscillation at a frequency that is halfway between the requirement of maximum and minimum tremolo frequencies. Note that in this condition the resistance presented by the frequency control potentiometer is naturally half of the total resistance.

One important point to realize is that the oscillator amplifier cannot oscillate unless it has something to amplify: If the oscillator is to be switched on and off (like in the shown configuration) it will need a proper “kick-start” signal that starts the oscillation again after the loop shunt is removed. The higher this signal is, the faster the oscillation starts and increases to its nominal level. Usually the transient that occurs when the bias voltage shifts from a short to its natural level is a sufficient kick-start but in some designs there may be elements that slow down or limit the magnitude of this transient, thus rendering it less effective. Such elements are, for example, emitter bypass capacitors. Note that the oscillation must start as quickly as possible since the performer will not likely appreciate the fact that the tremolo effect kicks in several seconds after toggling it. Phase shift oscillator circuits are delicate and have to be designed carefully based on the properties of the active device, rail voltages and frequency range needed. It may be hard to believe but it is actually quite difficult to get a circuit to oscillate *in a controlled manner*.

6.14 Cabinet simulators

The effect that the combination of speakers and guitar cabinet has on the frequency response is quite radical therefore it is useful, often even necessary, to emulate this effect in systems that do not employ a guitar cabinet (i.e. headphone and DI out circuits etc.). It is pointless to delve into technical aspects of designing cabinet simulator circuits since that is virtually only *filter design*. A typical circuit usually consists of a notch T-filter, a second order high pass filter and a fourth order low-pass filter. The latter is practically two second-order filters in series. Most designs are active in order to minimize insertion losses. One should read the chapters concerning speakers and cabinet design and accumulate all necessary information from them. Next step is to acquire datasheets of different guitar speakers and study their frequency response and impedance plots carefully. This should provide a good start in understanding what is going on. The basic approach is just to construct a collection of series filter circuits that create a sufficient emulation of the speaker system's frequency response. Some eccentric circuits may go as far as using some dynamic processing (e.g. clipping) to simulate the effects of speaker distortion at the extremes of cone travel.

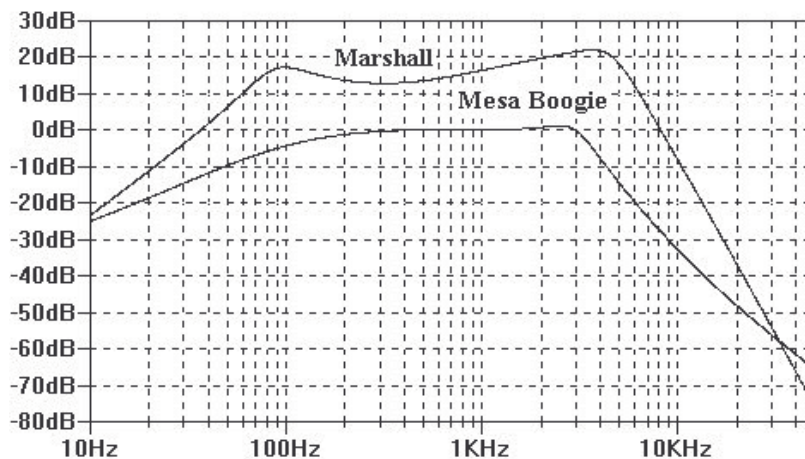


Figure 6.74 Frequency responses of Mesa Boogie and Marshall cabinet simulators.

The plots above are acquired from speaker simulator circuits used in Marshall and Mesa Boogie V-Twin amplifiers. For both, there are some common nominators and some individual features: The frequency response of both circuits falls heavily after about 5 kHz – Mesa Boogie circuit starts to attenuate at around 3 kHz, Marshall at 5 kHz. This is pretty typical for most guitar and PA speakers as well: Most of them roll off heavily after about 5 kHz. Compared to Mesa Boogie, Marshall introduces some emphasize at the frequencies of about 90 Hz and 4 kHz. The latter corresponds to typical upper mid highlight of Jensen speakers and the first to interaction of speaker's resonance and amplifier's low output impedance. In simulation, this emphasize is created with a notch T-filter. The effects are several decibels so they have a radical effect to the tone of the amplifier.

Real speaker cabinets introduce very similar effects, although their frequency response is even more irregular and contains delicate nuances. The lack of these details can make the simplest cabinet simulations sound unexciting, unnatural and sterile. Other important aspects that most cabinet simulators fail to successfully emulate are room reverberation and acoustic feedback. However, in most cases using even the simplest cabinet simulator circuit provides better results than omitting it completely. As always, faking real-life behavior requires complex algorithms and

circuits, thus it is my strong belief that a DSP approach in cabinet simulation would be a superior solution when compared to using analog circuits.

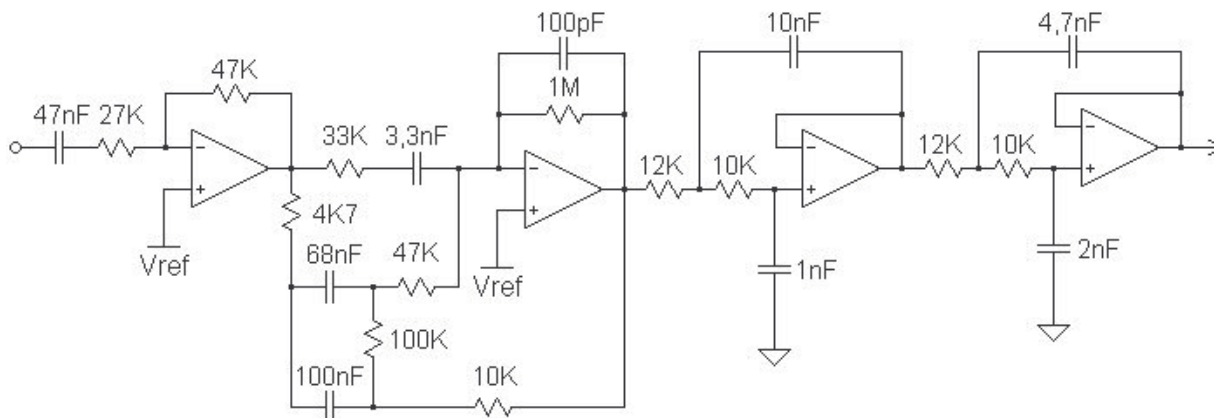


Figure 6.75 Example of a cabinet simulator circuit. (Marshall).

6.15 Phase splitters

Sometimes phase splitter circuits are required for special effects that utilize phase shifting. Even more often signals simply need to be inverted.

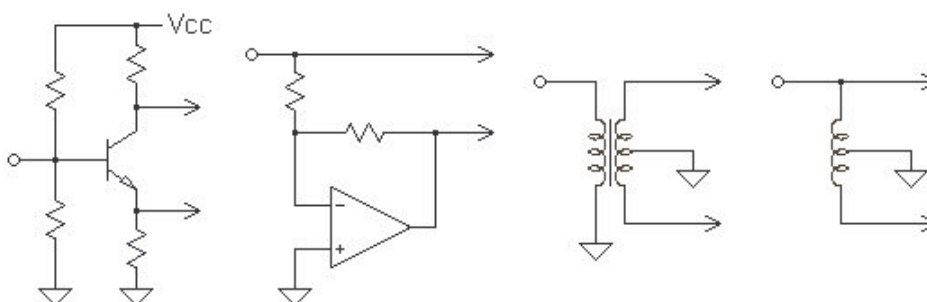


Figure 6.76 Phase splitter circuits. From left to right: Split-load amplifier, unity-gain inverting amplifier, center-tapped interstage transformer and center-tapped inductor.

Circuit concepts like split-load and unity-gain inverting amplifier are pretty basic and likely need no further explanation. Properly done the passive solutions based on transformers or center-tapped inductors are very nice solutions – but also the most complex and expensive ones. The center-tapped inductor or “center-tapped choke” is basically identical to transformer circuit – except for the coupling properties. Since the polarity of windings is different the signal appears in opposite polarities at the ends of the winding. The center tap is either tied to common or used to feed the bias voltage. The inductor is usually ferrite cored so practically a primary or secondary side of an ordinary transformer might qualify as a suitable component. Actually, converting the transformer topology to a center-tapped inductor was once a quite common way to “fix” broken phase inverter stages of vacuum tube amplifiers. In this case, the inductor has to be capacitively coupled of course. The topology would be eligible as a replacement of the transformer section in “totem pole”-type amplifiers.

6.16 Examples of tube emulation circuits

6.16.1 TransTube

Peavey’s TransTube is not a specific circuit or topology, merely a name for a concept of emulating characteristics of a tube amplifier. Thus there are actually quite many

circuits that belong into the TransTube category. “Hartley Peavey Whitepapers Chapter 3 – TransTube” acknowledges five characteristics, which are typical for tube amplifiers. These are output transformer’s saturation during overdrive (which causes bandwidth to change dynamically), bias drift during overdrive, asymmetric clipping in multiple gain stages, power supply sag and its effects as well as high output impedance and its effects. These characteristics – or only some of them - are simulated with various methods (and results) in the huge variety of TransTube series of amplifiers. Because the scope of this chapter is not to present each circuit in extreme detail, I kindly suggest that anyone who is interested in the subject will check out the references 51 – 58, which are US Patents and open for public view. The most interesting ones are likely the distortion circuit that uses a ganged gain- and bandwidth limiting “saturation” control [51], circuit that introduces asymmetric clipping and bias shift in multiple gain stages [56 and 57], push-pull distortion stage that introduces bias drift-related crossover distortion [54] and T-Dynamics circuit [58].

First circuits following principles of TransTube concept appeared around the early 80’s. These were simple circuits that incorporated an OpAmp gain stage with a bandwidth limiting gain control ganged to a potentiometer controlling the clipping threshold. The idea was to simulate the output transformer’s saturation that takes place during high gain overdrive and is excited by the DC component of the clipped wave. The schematic presented in this book is actually not an exact copy of the circuit that was used by Peavey but it incorporates the same principles and a similar gain stage. Note that the intention is not to present a schematic of a Peavey amplifier that you could copy but to present the basics of the tube simulation methods they use. If you are anxious to see the exact circuit and its component values refer to the Peavey schematics from the concerned era (e.g. Bandit 65, Renown 400 etc.) or the US Patent 4,405,832. Peavey also made circuit revisions that included addition of an independent gain control, brightness switch and an additional diode clipper within the gain stage.

Let’s examine the schematic shown in Figure 6.77: It can be seen that in one of its extremes the gain controlling potentiometer shorts the 6.8 nF capacitor, thus effectively shifting the 2-kilo-ohm resistor and the 470 nF capacitor from the feedback path to the output and converting the circuit into a plain non-inverting OpAmp stage. In the other extreme, the feedback path is interrupted with a capacitive shunt between the two resistances and the stage converts into an active band-pass filter that has its Q limited by the 2-kilo-ohm series resistor and the 6.8 nF capacitor bypassing the potentiometer’s resistance at higher frequencies. Effectively the circuit has a high gain on a considerably narrow bandwidth. The potentiometer blends between these modes.

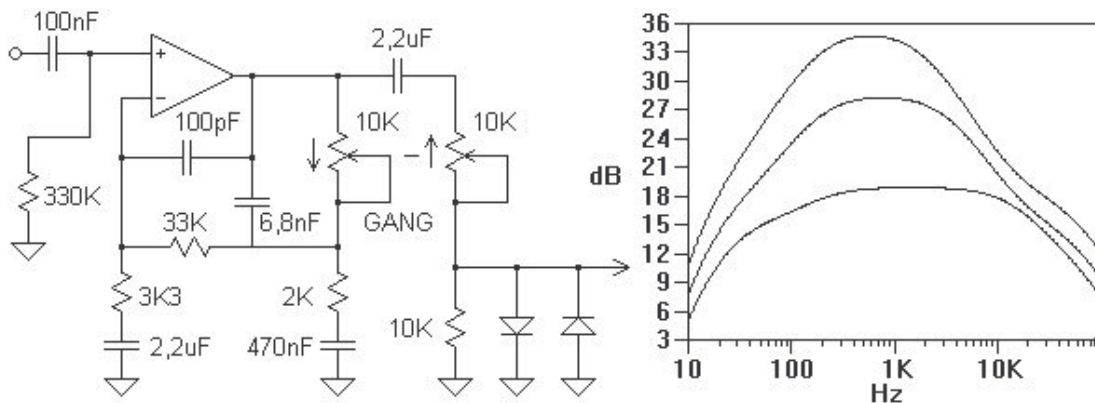


Figure 6.77 Bandwidth limiting gain control.

The final stage is a simple diode clipper with variable voltage divider that controls the signal amplitude at the “input terminals” of the clipping diodes, and consequently the amount of clipping as well. Note that the circuit *does not* introduce a “dynamic” bandwidth limiting effect that would vary in real-time according to signal amplitude. Since the bandwidth limit is only related to stage gain – not to amplitude of the DC component, which is generally higher during higher amplitude signals - the circuit essentially fails to simulate the dynamic saturation effect of a real output transformer. Nevertheless, this was a very good attempt and a step towards the right direction. Undoubtedly the potentiometer controlling the bandwidth could be replaced with a voltage dependent resistive element (like PTC resistor or FET) in order to capture some of the dynamics. Why Peavey never did it will remain a mystery – I am pretty sure they have at least thought of it. What happened instead was that Peavey abandoned the idea of using ganged potentiometers and a variable clipping threshold control. At the end of the 80’s the topology had evolved into cascaded gain stages of “Super Distortion” circuit [53]. These had a fixed clipping level but still used bandwidth limiting gain stages.

In the early 80’s Peavey Electronics Corporation also did some experiments with push-pull grounded emitter amplifiers that would couple the signal using a real output transformer. Up to my knowledge none of these designs went into a production although a patent was issued. While some Peavey amplifiers did use transformers in the output stage they were either used for impedance matching (autotransformers) or interstage coupling.

In the mid 90’s another topology to simulate the clipping characteristics of push-pull tube stages was devised [54]. The circuit consists of two class-AB OpAmp buffers that have a floating offset that is controlled by the signal amplitude. The offset biases the output diodes into slight forward conduction, which in turn clips off a major portion from the corresponding half wave. When the two output signals are summed together with mixing resistors, they create a complete full wave. The circuit is effectively a push-pull stage. The summing process also introduces some inherent error correction and softens the clipping. When input stage is driven hard enough also the input diodes begin to clip; this over-biases the swing at the output, which essentially causes a shift in the zero crossing point. When the output diode clips its corresponding half wave, it will now also clip the over-biased portion of the other half wave. In this case the significant asymmetry of summed signals gives birth to crossover distortion. Similar crossover distortion also occurs in clipping push-pull tube amplifiers.

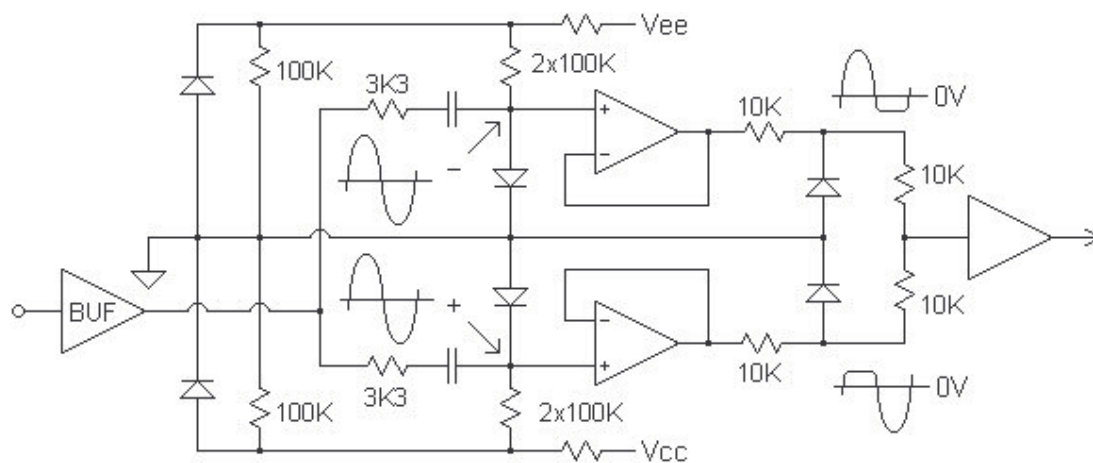


Figure 6.78 Asymmetric clipping in parallel stages.

The late 90's circuit, that employs cascaded FET or Darlington transistor gain stages, is practically the Peavey's simplest tube emulation circuit. Around the time it was introduced Peavey also started to categorize tube emulation methods as "TransTube". The purpose of this circuit is obviously simulating the preamplifier circuitry of conventional tube amplifiers. In US Patent 5,619,578 circuit consisting of four gain stages is presented: The gain stages are typical AC coupled common source or common emitter amplifiers with attenuating resistor divider inputs and some AC bypass at emitter/source. Only significant difference in comparison to usual (discrete and simple) multi-stage circuit topology is the diode shunt that limits positive signal swing at the gate/base to approximately 500 mV thus creating an asymmetric clipping effect in each overdriven stage. The diode is also the novel addition of the invention since it doesn't only clip but also shifts the bias of the stage whenever the circuit clips. The coupling capacitor, along with series- and shunt resistors, defines the attack and release time constants for the bias shifts.

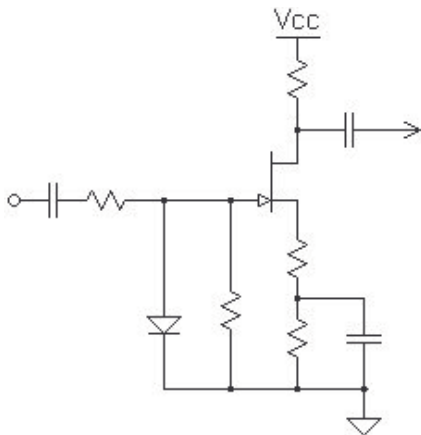


Figure 6.79 Common source amplifier with limited gate voltage swing.

Another embodiment of the same idea, presented in the patent, uses discrete Darlington transistors due to their higher input impedance and gain (when compared to single transistor), consistency (when compared to FETs) not to mention lower cost and possibility to tweak the turn-off characteristics by altering the emitter-shunt resistance of the first transistor in the pair (when compared to integrated Darlington). This is also the version that went into production. In practice, the previously introduced circuit experiences some minor changes: Gain stages are now biased from a separate bias voltage source, which is formed with a simple resistor divider network.

(Note that schematic omits the filtering capacitors for simplicity). The input stage differs from the rest by having a separate and higher bias voltage supply, as well as a separate reference for the diode shunt. Emitter resistors are also bypassed with parallel signal path to gain some high frequency boost. Some common component values are shown but they are subject to change according to application. Schematic also omits the Miller capacitor of the first stage.

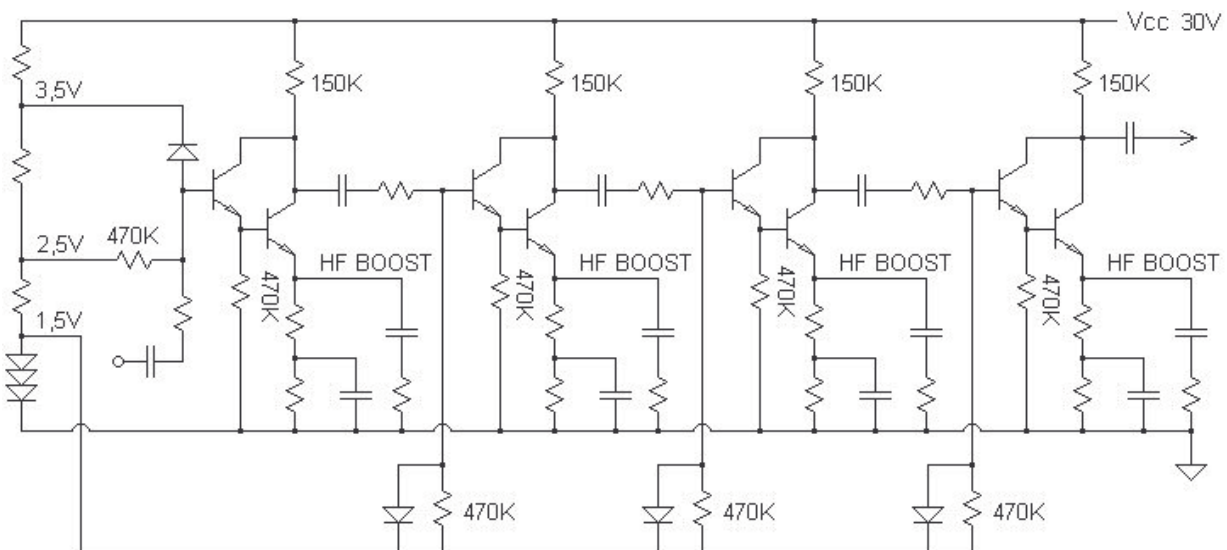


Figure 6.80 "TransTube" circuit utilizing Darlington transistors.

Due to inherent minimalism the performance of both circuits is - as expected - pretty much dominated by the “artistic” choice of component values that define coupling, gain and bandwidth characteristics. Naturally some of the resistor dividers can be replaced with potentiometers to create variable gain characteristics.

Addition of “Tube Dynamics” or “T-Dynamics” took place in the late 90’s as well. The circuit is essentially a power attenuator that simultaneously varies the compression characteristics of the amplifier. It may also introduce resonance and presence controls that adjust the damping factor at high- and low frequencies. The circuit can be divided into few major parts, which are gain stages 1 and 2 (latter being part of the power amplifier), signal limiter stage, feedback circuits of the gain stages and the “main” T-Dynamics circuit that controls both feedback paths. Modern Peavey amplifiers tend to use grounded output topology (refer to Figure 3.61) and therefore such circuit is also depicted in their patent [58]. Note that differential stage of grounded output amplifier uses inverted pinout: For example, *non-inverting amplifier uses the inverting input as the signal input* and non-inverting for the feedback. Consequently, schematics in the T-Dynamics patent are drawn incorrectly. T-Dynamics works just as well with a conventional circuit topology and this is used in the conceptual schematic example since it is easier to follow.

Both gain stages are essentially non-inverting amplifiers. They are separated from each other with the limiter circuit in series with the signal path. The second gain stage may look complex at first glance but when it is broken down into sections one can note that it is in fact nothing but an ordinary non-inverting amplifier. The concerned stage is also buffered so it essentially is a power amplifier. To overcome the limit of OpAmp’s output voltage swing Peavey mainly utilizes “level shifting resistors” (see Figure 3.29). However, this is just a minor detail and will not affect the operation of the T-Dynamics topology; in essence, the stage could be discrete just as well. The (voltage) feedback loop from the output splits into two branches (FB1 and FB2) that feed both the first gain stage and the driver OpAmp of the power amplifier. The gain of the power amplifier stage is controlled with a potentiometer and limited to proper range with the two resistors in series and in parallel with it.

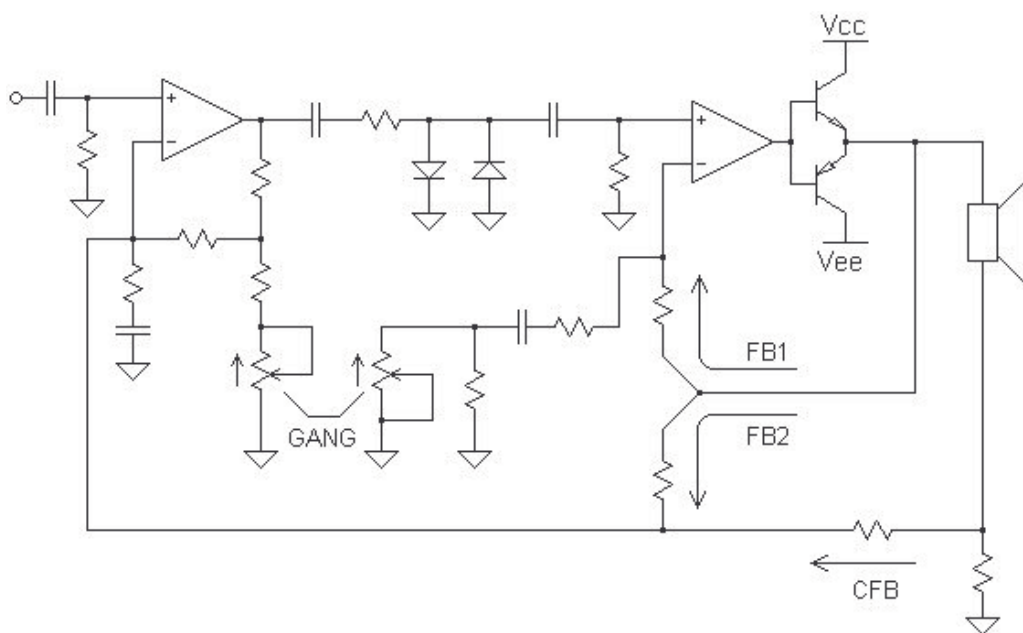


Figure 6.81 Conceptual schematic of T-Dynamics circuit.

You may have noticed that the circuit also utilizes current feedback. Mixed-mode feedback, by the way, has also been a key part of Peavey's tube emulation circuits at least since the early 80's. However, this time the current feedback is mainly directed to the first gain stage, thus causing its frequency response to resemble that of an amplifier with high output impedance driving a speaker. This is an essential detail because the tube-emulating limiter circuit will now react to dynamic gain changes caused by the varying speaker impedance. In its simplest form the T-dynamics circuit would utilize a simple diode clipper stage as the signal-limiting element. In practice the topology rather utilizes the more refined clipping circuit depicted in Figure 6.78.

The other feedback branch, which is mix of voltage and current in nature, is hooked to the summing (inverting) input of the first gain stage. In one embodiment of the circuit part of the resistance that is setting the amount of voltage feedback is made variable, while simultaneously bypassing it with a capacitor, thus creating a control for the low frequency damping ("Resonance"). "Presence" control is created by splitting part of resistance and by shunting the joining node of the two resistors to ground through a potentiometer that has its wiper connected to capacitor. The other end of the capacitor is connected to ground, thus the circuit essentially forms a resonant circuit that controls the high frequency damping. These circuits have been discussed earlier in this book.

The first gain stage, as already mentioned, is non-inverting. However, it differs from the conventional topology since it places *a voltage divider* into the feedback path: When the ganged-potentiometer of the first gain stage is in the fully clockwise position we can presume that the resistance of the potentiometer and its series resistor are high enough to render the voltage divider "invisible". Thus the two "conventional" feedback resistors in series define the gain of the stage. In the other extreme of the potentiometer's setting the feedback path is "shunted" by the attenuating voltage divider. At most severity this would mean a complete disconnection of the feedback path that would result into infinite gain, however, the series resistor is there to prevent this. Thus it can be concluded that varying the amount of attenuation caused by the voltage divider varies the stage's gain. The ganged potentiometers are connected so that the gain of the second stage increases while the gain of the first stage decreases - and vice versa. To obtain supreme accuracy the potentiometers use opposite audio tapers. As a result, the limiter has more effect when power amplifier's gain is set low - and vice versa. Below clipping levels of the limiter circuit or the power amplifier itself the gain characteristics of the amplifier are identical regardless of T-Dynamic control's setting.

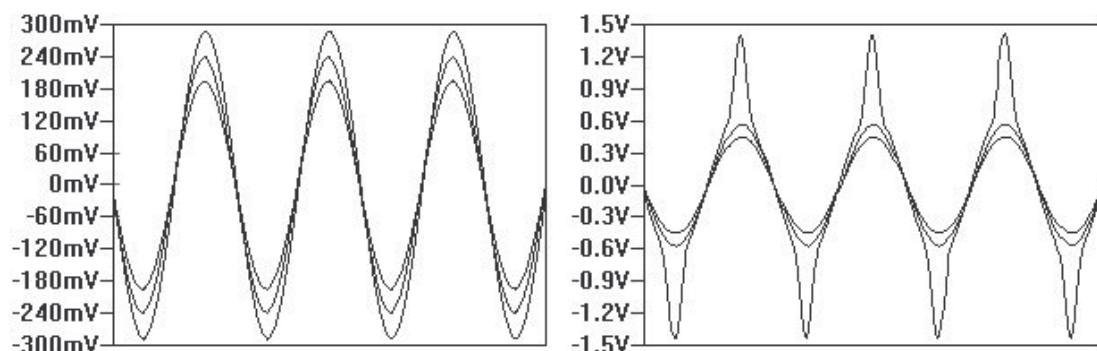


Figure 6.82 Output signal from first gain stage with T-dynamics set at "10", "5" and "zero". Left: Low input signal amplitude and right: Higher input signal amplitude.

The elegance of T-Dynamics lies in the interaction of gain stages: One must remember that the effect of feedback is defeated whenever the amplifier clips, thus the feedback path from the power amplifier's output to the first gain stage is defeated as well (including the effect of presence and resonance controls). As a result, the feedback no more limits the gain and the gain of the first stage increases. This in turn exaggerates the signal peaks and consequently drives the signal limiter harder, thus preventing overdriving of the power amplifier. This “self-correcting” interaction virtually makes the power amplifier immune to rail clipping. A sufficient input signal *can* drive the first OpAmp to clip into rails, though. Note that engaging the signal limiter also decreases the amount of feedback, thus affecting the presence and resonance controls as well. Overall, the T-Dynamic is, as its name implies, a very dynamic topology.

6.16.2 Carvin SX series' Classic Tube Emulation

In 1983 John Murphy invented a diode clipping topology that performs “duty-cycle modulation” by dynamically shifting the bias points in order to cause asymmetric clipping. The invention was first used in Carvin's SX series of amplifiers that were introduced in 1987. For some reason patent was never applied for and the circuit became public domain. Peavey introduced their alternative method for duty-cycle modulation a few years later. Note that both topologies mainly focus on clipping the signal with similar asymmetry as gain stages in tube amplifiers do and the clipping method alone does not guarantee a good tone. Like many others, Murphy has concluded that pre- and post clipping equalizations are “absolutely critical adjustments” as well.

The SX series tube simulator circuit uses two “single-sided” diode-clipping stages that are buffered with inverting operational amplifier gain stages. The stages are AC coupled and diodes are tied to common with resistance reference in order to make asymmetric clipping to cause DC bias shift. To make clipping harder the feedback loops of OpAmps also employ diodes. The circuit at “bottom” varies the DC offset levels of the gain stages according to the amplitude of the input signal it samples. The presented circuit is merely a rough sketch and can be modified extensively by increasing HF feedback within gain stages, tweaking resistor values and forward voltages of the diodes. Possibilities are endless.

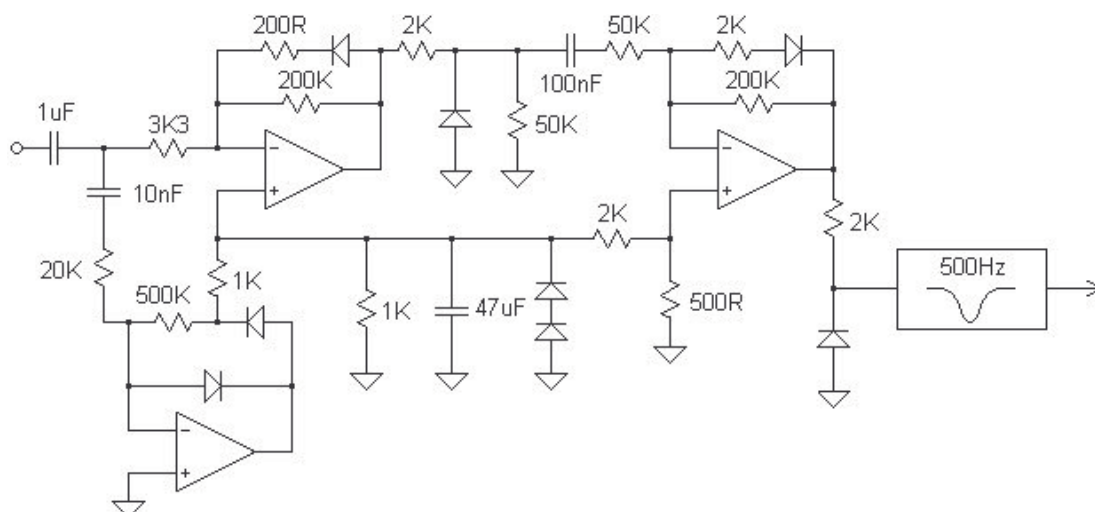


Figure 6.83 Conceptual schematic of SX series “tube simulator” stage.

Filters circuits that control frequency response are omitted from the schematic but generally the circuit contains a treble-preserving attenuator in the input (bass is reduced at higher volumes). The output of the clipping circuit feeds a band-stop filter stage that creates a typical “gull wing” response (similar to basic response of a FMV tonestack) with a -5 dB notch at 700 Hz. This is reasonable since Carvin SX Series amplifiers use an active, three-band Baxandall tone control that has a *flat* initial response. Tone controls are located after the distortion circuit. The band-stop (notch) filter is followed by a bass band-boost stage that has center frequency at 50 Hz; boost is about 5 dB.

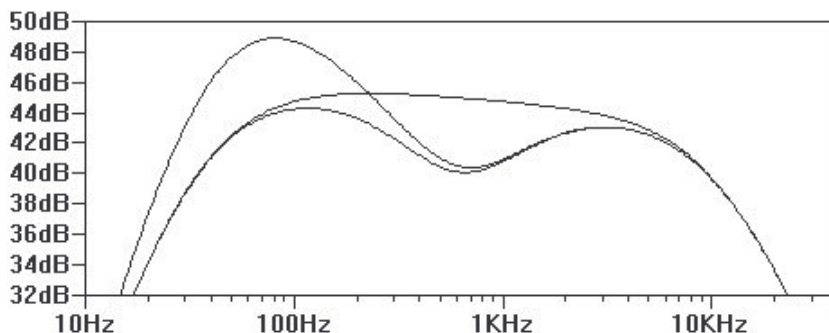


Figure 6.84 Frequency response shaping after duty-cycle modulated clipping in Carvin SX200.

6.16.3 FlexWave

FlexWave topology, used in Crate amplifiers, was patented in the early 90's [59] and therefore a thorough description of it is open for public. Principally amplifiers utilizing FlexWave technology employ mixed-mode feedback output stage as well, which seems to be the common “standard” of guitar amplifiers nowadays. The aforementioned FlexWave circuit is usually located in the separate overdrive channel, which consists of a cascade of clipping and tone control stages. What comes to origin of the topology's name, it is explained by Crate that with the aid of the circuit the waveform “flexes” in response to picking touch. The circuit discussed hereafter is the distortion circuit patented by Saint Louis Music, Inc.

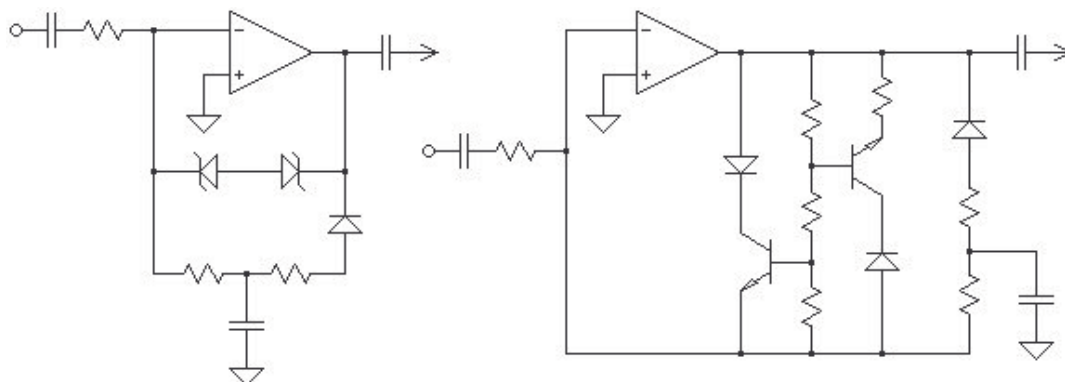


Figure 6.85 Inverting signal limiter stages with dynamically shifting DC offset.

The FlexWave circuit uses a diode to half wave rectify the output signal. The resulting DC signal is used in charging the capacitor between the two resistors with correlated time constant. Furthermore, the DC signal is also shunted to the virtual ground of the inverting input resulting into a variable DC offset in the output. In practice, the output signal's DC offset shifts positively when the amplitude of the input signal increases. In a conventional circuit the clipping level is determined by two zener diodes connected back-to-back. These typically have different zener voltages so the clipping

is asymmetrical. However, the shifting offset also reduces the asymmetry, which is at its summit when the signal is applied and gradually decreases. As a result, there is also a shift in harmonic content that is claimed to be similar to that introduced by an overdriven tube amplifier. Another variation of the idea is to replace the zener diodes with actively controlled “rubber diode” configuration depicted at right. Typically two clipping stages are used in series. The stages have to be AC coupled to prevent bias shifts from transferring from one stage to another. A very generic FMV tonestack and a heavy notch filter tuned to 1 kHz usually follow the distortion circuit.

6.16.4 Valve Reactor

Korg’s patented Valve Reactor and “Virtual Power” (variable power or VariAmp) topologies, used in Valvetronix series of amplifiers, are covered in United States Patent 6,350,943 B1 (Suruga et al.). The patent describes an electric guitar amplifier that “emulates audio characteristics of a traditional vacuum-tube type amplifier”. All hype that has been written about the Valvetronix amplifiers has also created a lot of misconceptions, mainly because some of the written material is outrageously inaccurate. Let’s take a look what Valve Reactor topology is and what it is not.

Note: The “REMS” DSP-based preamplifier is a great contributor in creating the tones of the Valvetronix amplifiers but firmware code of digital circuits is not the topic of this book.

Valve Reactor and VariAmp circuits can be divided into five major parts, which are the differential amplifier (and its feedback circuit), push-pull tube gain stage, the “simulated output transformer”, combined variable gain and feedback stage, and the power amplifier. These parts are interacting with each other more or less. The power amplifier circuit is deliberately left out of this focus since it’s mainly a generic non-inverting IC power OpAmp stage (for example, LM3886 in AD30VT) and nothing exceptionally interesting.

The long-tailed pair differential and phase splitter amplifier is a very basic circuit and based around two bipolar transistors. Its outputs drive a push-pull 12AX7 tube gain stage that is powered from starved plate voltages. Cathodes are biased from the negative -15 V supply rail. Resistances R1, R2 and R3 are variable to achieve either more drive (R1 and R2) or to vary bias between class-A and class-AB mode (R3), the latter of course only in the tube gain stage. In real-life realization of the circuit the aforementioned resistances are controlled by digital logic switches, which in turn are controlled by either CPU or footswitch. The variable feedback filter circuit is basically just a feedback path controlled by digital logic switches that bypass some of the resistors and capacitors. In the basic form it essentially adds a switchable high-pass filtering to the gain stage. The feedback path can also be disabled. Since all features are CPU controlled the preamplifier’s tone patches essentially affect the analog Valve Reactor circuit as well.

A virtual output transformer follows the tube gain stage. This circuit consists of two inverting unity-gain OpAmps that handle phase inversion and error correction by summing the output signals together through mixing resistors R4 and R5. Note that the output signals of the push-pull stage have opposite phases, thus the tube stage on top can be error corrected by feeding back the inverted signal from the tube stage at bottom and vice versa. The circuit does not simulate saturation and inductive effects of a real output transformer.

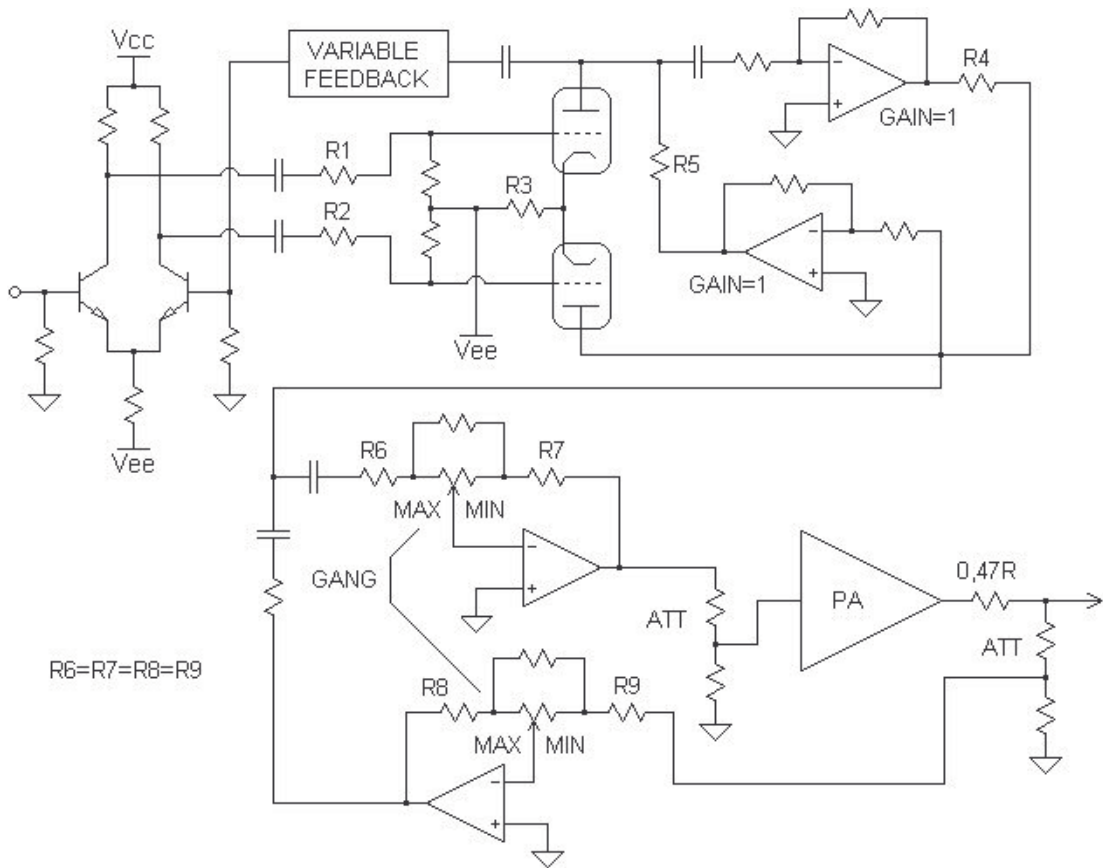


Figure 6.86 Conceptual schematic of Korg’s Valve Reactor and Virtual Power topologies.

Combined variable gain and feedback stage is wrapped around the power amplifier circuit. One of the functions of this stage is to provide a variable level of mixed-mode feedback, which reflects as “amplified load” to the virtual transformer circuit. Thus the gain of the push-pull stage is partially controlled by the (speaker) load impedance – like in most tube-, and mixed-mode feedback amplifiers. Essentially *this makes the tube stage function like it would be driving a real speaker*. The 0.47-ohm resistor is used to deliberately increase the output impedance to enhance the effect. The most important point in the operation of the circuit is that the amount of feedback has to be amplified (or attenuated) in inverse proportion to gain/attenuation of the input signal - otherwise the damping characteristics and gain of the whole circuit will vary. This happens because the resistor attenuators render the voltage gain of the power amplifier “invisible”, thus essentially converting it into a power buffer for the push-pull tube stage. Fender has also patented an astonishingly similar concept, which they call a transimpedance amplifier [61].

6.16.5 Valvestate

Aside employing a conventional mixed-mode feedback configuration that provides a “tube-like” frequency response, Valvestate amplifiers contain no circuitry that would specifically emulate characteristics of a tube amplifier. However, excluding the low-power models of the first two Valvestate lineups the amplifiers do employ an actual dual triode ECC83 in the preamplifier circuit. Yet that fact just makes them simple *hybrids* - and not even very novel as such. It can be argued whether mixed-mode feedback should be considered as tube emulation technique or not but based on the fact that some Valvestate models include a switch between “Valvestate” (current feedback path toggled) and “Linear” (current feedback path open) it is clear that

Marshall has formed their opinion about the issue. Yet, current feedback is not even exclusive to Valvestate lineup: For example, it was already in use in the MG series where the FDD (Frequency Dependant Damping) feature simply disabled a low cut filter that attenuated the resonance peak.

It should be noted that, excluding the tube stage, circuit designs of various Valvestate amplifiers are very different from each other. Basically, the more advanced (and expensive) the model is, the more sophisticated circuit it contains. Considering the triode circuit, first and second Valvestate lineups were virtually identical. In the AVT series (Advanced Valvestate Technology) the section is revised. A typical Valvestate preamplifier has nearly separate signal paths for clean- and overdrive channels. Common nominator is sharing the tube stage. Solid-state stages of the clean channel are basically configured to be virtually transparent within the realistic range of input signals, thus the channel will provide *only* tube overdrive. Overdrive channel, however, utilizes diode clipping and provides both solid-state and tube overdrive. As a rough generalization, the first and second lineup used a shunting diode clipping configuration while the circuits of AVT series utilize diodes in the OpAmp feedback loop, thus being less prone for driving OpAmps to clip into rails. The limiting stage is located before the tube stage but the placement of gain-, and in some cases tone, controls varies widely in different models. Typically tone controls follow the tube stage forming the cathode follower buffering arrangement that is similar to many Marshall circuits.

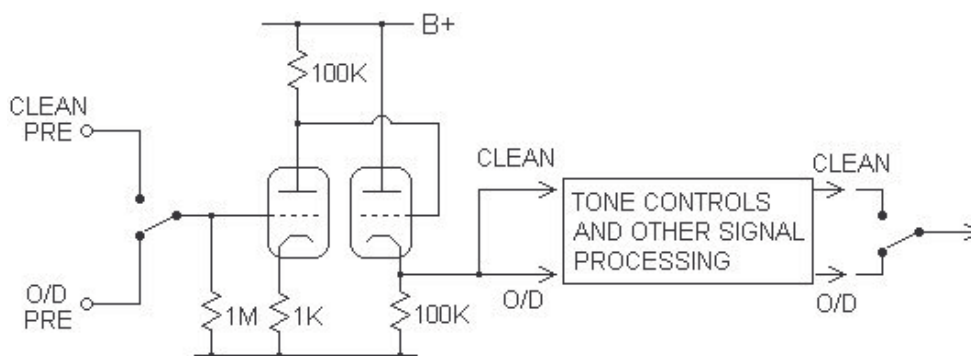


Figure 6.87 Tube section of the Valvestate series 1 and 2.

The dual tube stage of series 1 and 2 Valvestate amplifiers consists of a common cathode stage, which is directly coupled to a cathode follower. This has been a typical circuit in Marshall amplifiers ever since JTM45 and it allows driving the tone stack from fairly low impedance source. The B+ voltage of Valvestate amplifiers is approximately 109 V so it is somewhat lower than the B+ of typical tube amplifiers. The moderately low value is pretty typical for all similar hybrid circuits. Although the voltage gain of a cathode follower is less than unity the whole tube section, counter intuitively, has *fairly symmetric clipping characteristics*: According to magnitude of drive signal the clipping turns from asymmetric (cathode follower is overdriven) into “semi-symmetric” (the common cathode stage clips as well). Both configurations also have different clipping characteristics: Cathode follower generally clips “harshly” while the common cathode has “softer” clipping. The input impedance of the following stage also has huge effect on the behavior: Purely resistive load creates very harsh clipping while the irregular loading impedance of a typical FMV tonestack circuit can actually soften it a bit. Aside from having lower gain the performance of the circuit is identical to similar circuit in JTM45, Fender Bassman and various other “classic” tube amplifiers.

In the AVT series the tube stage is configured to provide more gain by wiring the second tube to common cathode configuration as well. The typical cathode follower concept is still maintained but the tube is replaced with a MOSFET that shares the high rail voltages. Not shown in the schematic is the zener diode protection that connects between the source and the gate and consists of two back-to-back 15 V zener diodes. It is claimed that a MOSFET is a very “transparent” replacement of a tube in a follower configuration. While being partially correct the claim is also a bit misleading: In the concerned circuit the cathode (or source) resistance is very high, which *biases the follower so that it is driven into clipping*. (This is exactly the same thing that happened in the earlier Valvestate circuit as well). If one would decrease the cathode resistance from the typical value of about 47 – 100 kilo-ohms to, say, 3.3 kilo-ohms the follower would operate in the linear region and would not clip the signal. However, knowing this doesn’t change the existing design to anything else than what it already is. In this case the follower *will* clip – and this is an essential ingredient of the circuit’s tone. However, I haven’t yet seen a MOSFET that in the source follower configuration would have similar clipping characteristics as a triode in cathode follower configuration. This is not saying that the circuit couldn’t clip “softly” or “tube-like” – just that a MOSFET does not behave exactly like a triode. MOSFET choice is an extremely important aspect of the design and the device choice dominates the clipping characteristics of the preceding tube stage as well.

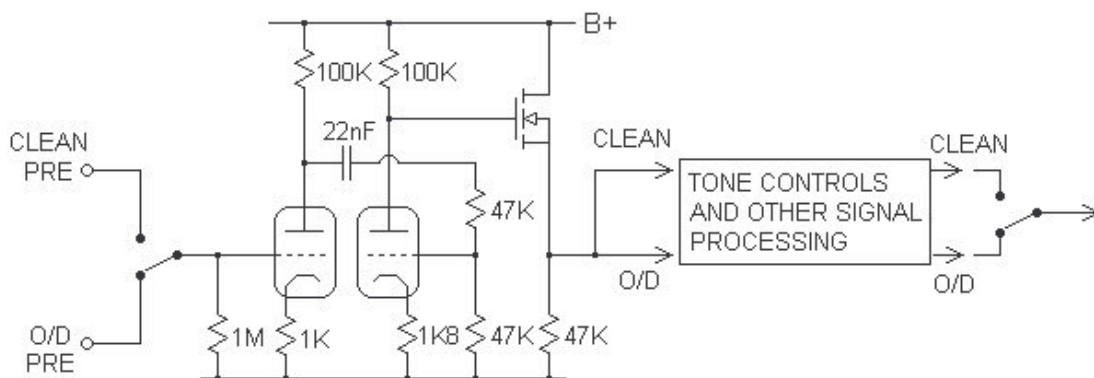


Figure 6.88 Tube section of the AVT series.

6.16.6 SansAmp

Tech21’s SansAmp product range is known from its amplifier emulation technology, which was initially introduced by Andrew Barta in 1989. It would be unfair to claim that SansAmp doesn’t include a tube emulation circuit (in fact it emulates a complete signal path from amplifier’s input to the microphone picking the sound from the speakers). However, quite fair is to state that SansAmp technology utilizes *nothing but conventional circuitry and prior art*. Because SansAmp contains no “special” or novel circuitry I was a bit skeptical about discussing it in this book. What helped me to make up my mind is the fact that SansAmp, in all its “simplicity”, is an extremely perfect example of what can be done with *carefully designed filtering*. Take for example its distortion tones: SansAmp products rely strictly on such “heresy” as *overdriving OpAmps*. If with proper filtering this can be made to sound like an overdriven tube amplifier then pretty much anything can – the vast amount of FET-based “tube emulator” circuits essentially prove the same point. Note that my statements should not be considered as belittling this type of “amplify-and-filter”-approach: For example, of all the tube emulators presented here SansAmp is perhaps the most respected and praised.

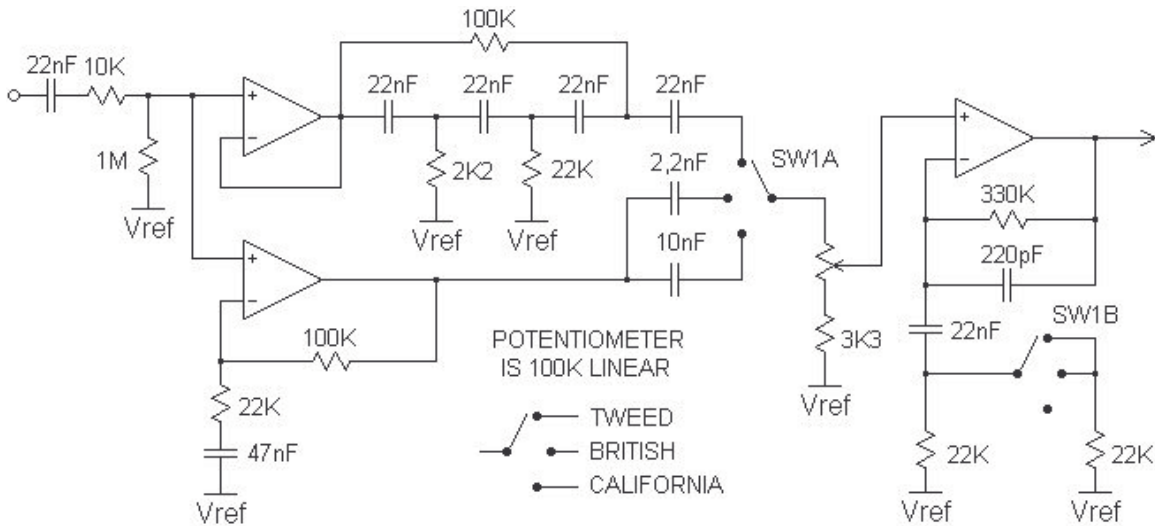


Figure 6.89 SansAmp GT2: First stage (Amplifier selection).

There is no universal “SansAmp circuit” and thus the insight is limited to SansAmp GT2. However, what are universal are the approaches and methods of tone shaping used in the concerned circuit – and these can be harnessed to create an infinite amount of different tonal characteristics. For example, it should be noted that circuitry in SansAmp TRI-A.C., TRI-O.D. and Bass Driver DI is not much different. When we follow the signal path of GT2 starting from the input, we see it splits in two branches: The non-inverting stage at the bottom is a conventional non-inverting amplifier and configured for some bass-cut - similarly to all signal coupling in the first stage. The “upper” path, which is likely much more interesting, is a form of passive bridged-T filter. This can have even higher Q than the twin-T variant. Here the filter is configured to provide the 1-kHz notch for the “Tweed” amplifier model. The “amplifier selection” switch is configured to select proper corner frequencies for the hi-pass filters, as well as to set gain of the second gain stage. It can be concluded that the main function of this circuit section is to do some basic, pre-distortion “voicing”. With highest gain setting (“British”) the second OpAmp may also provide some signal clipping.

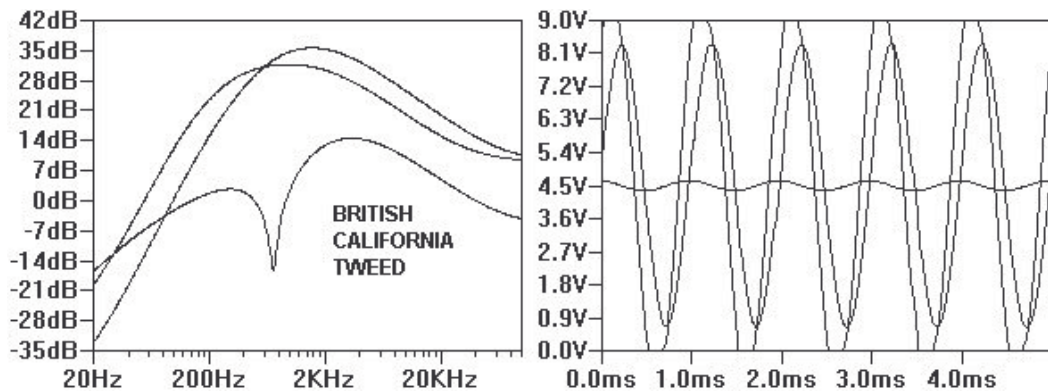


Figure 6.90 SansAmp GT2: Plotted frequency responses and output signals of “amplifier selection” stage with 71 mV 1 kHz sine wave input signal. From top to bottom: British, California and Tweed.

The second circuit section starts with a fairly high gain OpAmp stage. Once again the gain is controlled by switched resistance value, which also affects the corner frequency of the internal hi-pass filter. This is actually a very important arrangement because the decreased bandwidth at higher gain prevents blocking distortion, which is the main cause of “muddiness” or “farting”. Two second-order Sallen-Key low-pass

filters follow the gain stage. First one of them is also equipped with series RC shunt that creates more roll-off at the mid-range frequencies (where simple Sallen-Key filter would stay flat), thus creating the pronounced bass-boost that can be seen in the frequency response plots of the concerned circuit section. Note that the plotted responses show *only* the effect of the section depicted in Figure 6.91. This is due to providing a clearer demonstration of various filters and their effects. The same principle is followed hereafter.

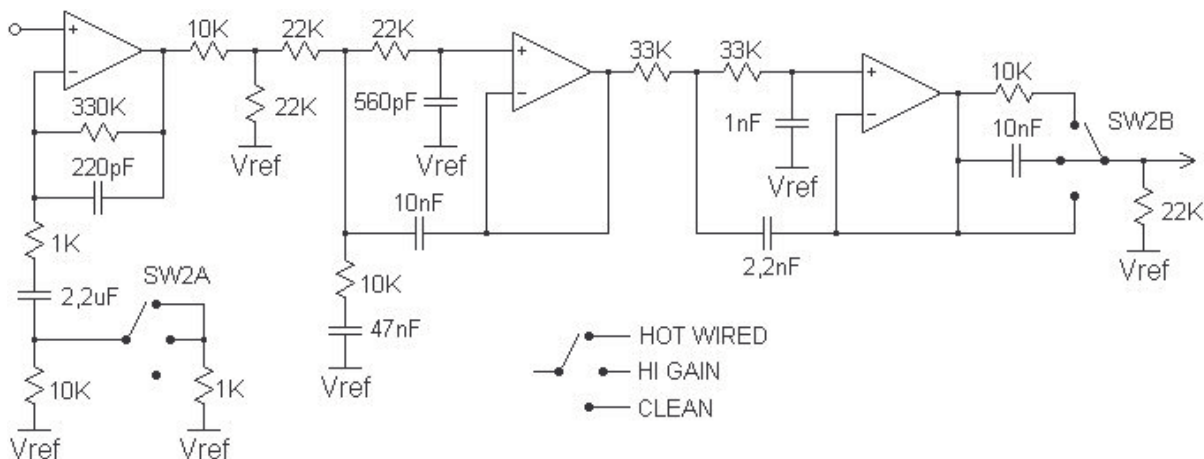


Figure 6.91 SansAmp GT2: Second stage (“Mod” selection)

Finally, the second part of the “Mod” switch is used to select between attenuating, hi-pass or unaltered signal path. Again, the circuit section is a better example of artistic voicing than mentionable novelty. As can be noted, the circuit utilizes no diode clipping, asymmetry or DC shifting - or any other imaginable configurations to make it more “tube-like”. Yet the waveform plots from the output are far from the “flat tops with hard knee” -stereotype associated with solid-state clipping. What we can see, however, is the pronunciation of different harmonics in “British” and “Hi Gain” configurations. This is how clipping gets its characteristics tones.

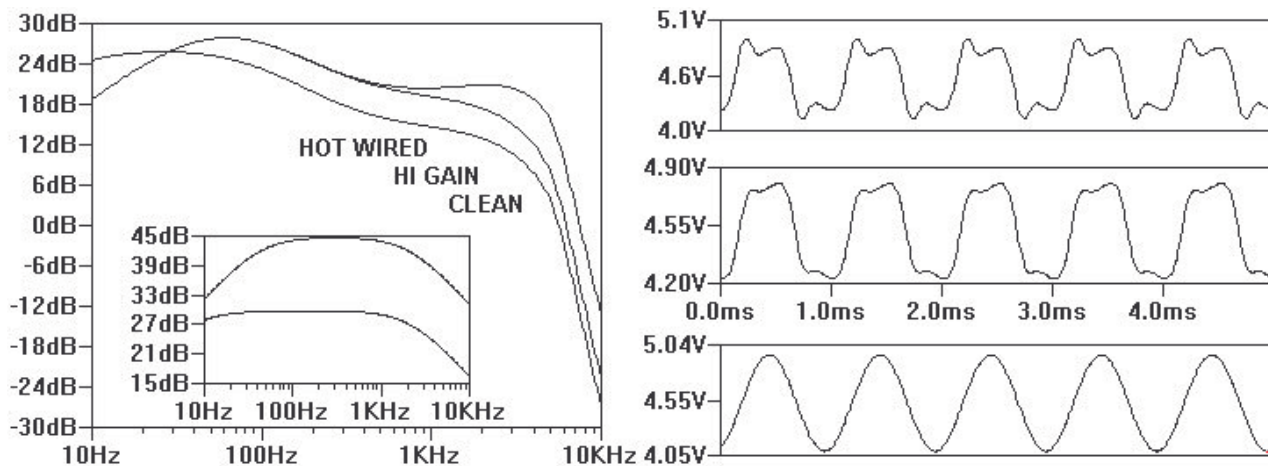


Figure 6.92 SansAmp GT2: Plotted frequency responses and output signals of “mod selection” stage with 57 mV 1 kHz input signal. The small plot depicts the bandwidth of the first gain stage with clean and distortion setting.

Following the “Amplifier” and “Mod” selection circuits is the emulation of microphone position. When this circuit is broken into sections according to switch positions we can see that the signal path for “Classic” configuration is pretty much an amplitude halving voltage divider (the series connection of 4,7 nF capacitor and 6.8 kilo-ohm resistor is pretty much insignificant). Both center and off-axis configurations form a Bridge-T filter with a variable notch frequency.

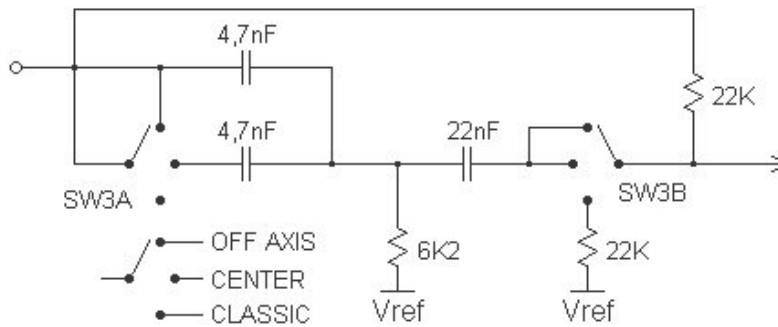


Figure 6.93 SansAmp GT2: Third stage (“Microphone selection”).

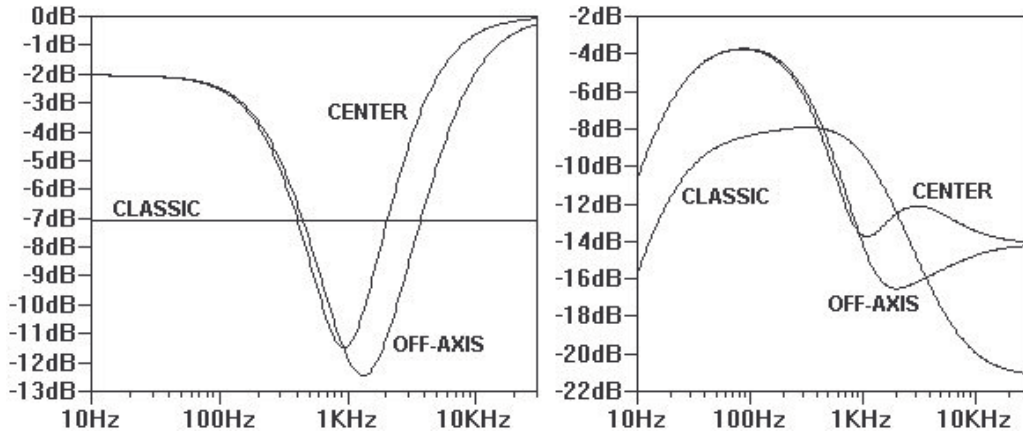
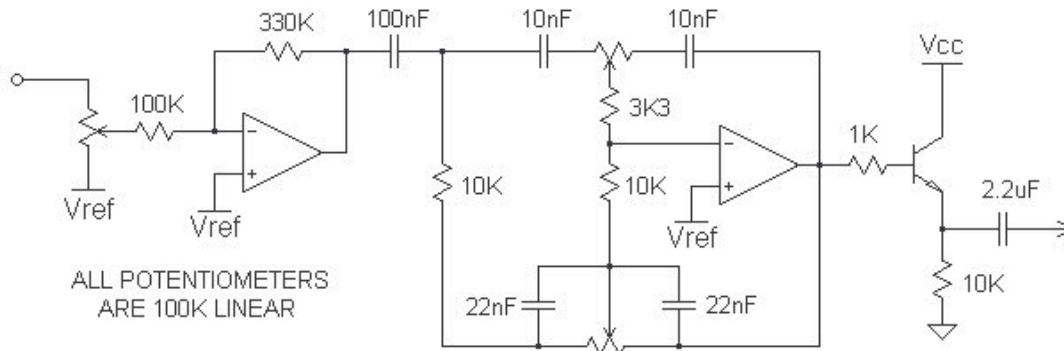


Figure 6.94 SansAmp GT2: Left: Frequency responses of the “microphone simulation” settings. The combination with low-pass filter of the final stage is shown at right.



ALL POTENTIOMETERS ARE 100K LINEAR

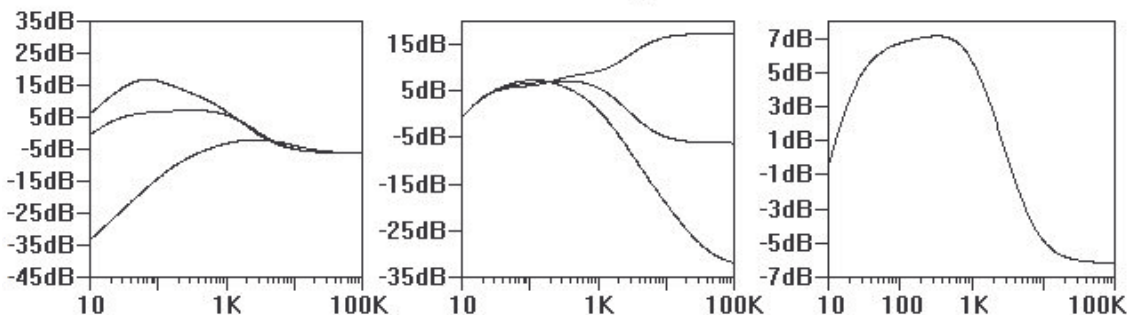


Figure 6.95 SansAmp GT2 final stages: Volume and tone controls, and signal buffering. Depicted below are frequency responses of the controls and far left the response with both tone controls in middle position.

The final stage of the SansAmp GT2 is an extremely basic configuration consisting of an attenuating volume control, low gain, inverting amplifier and an active Baxandall tone control circuit followed by a BJT buffer. The catch of the circuit is the odd voicing of the tonestack, which forms a noteworthy band-pass filter when the controls are set to middle position. The corner frequency of the high-cut filter is around 5 kHz so the circuit additionally imitates the cutoff introduced by a typical speaker. The

controls are designed to provide a very large range from considerable cut to some boost.

6.17 What is an “Acoustic Amplifier”?

An acoustic amplifier is an amplifier that is specially designed for amplifying vocals and acoustic instruments (such as acoustic guitars). By experience I know that many musicians are not particularly educated about amplifiers and their features, thus they often see no difference in a guitar amplifier and an *acoustic guitar* amplifier. In reality, the requirements of acoustic instruments (and other acoustic signal sources) create quite considerable distinctions to circuitry when compared to “ordinary” guitar amplifiers. I feel that a small introductory is more than necessary:

- An acoustic amplifier (by default) *must have a linear output*.
- Speaker system of an acoustic amplifier should be able to reproduce a *broad bandwidth without extensive coloration or distortion*.
- Besides having an interface for “instrument” inputs, an acoustic amplifier should also provide *proper interfaces for line level sources and microphones*.
- Preamplifier’s channel handling, EQ and other features should serve the *unique requirements of acoustic instruments and vocals*.

Let’s discuss the aforementioned points a bit more thoroughly: Quite often an acoustic amplifier is used in a configuration where a guitarist plugs in both his guitar and vocal microphone, thus using the amplifier to reproduce both signal sources. In such case there is nothing but harm from such typical features of conventional guitar amplifiers as smooth breakup to distortion, mid-range notch introduced by the tonestack or low-pass filter introduced by a single mid-band speaker. In essence, the output of an acoustic instrument is typically rather “complete” already and usually only needs to be reproduced a bit more loudly. Some other special requirements will be discussed in greater detail later. Anyway, it is rather safe to say that when designing acoustic amplifiers the typical design rules concerning conventional guitar amplifiers are pretty much ignored.

The requirements for broader bandwidth and more linear output naturally dictate speaker choices, thus employing piezoelectric elements, horns or tweeters, dual-cone drivers, bass reflex systems and crossovers is quite common. Guitar speakers having a “smooth breakup” or “signal coloring” non-linearity in frequency response are typically discarded in favor of more “neutral” PA drivers. Lately, designers have also begun to employ various cabinet construction techniques that enhance an “omnidirectional” dispersion of sound. High efficiency (SPL) of the loudspeaker system is still, of course, an important issue. Naturally the extended low frequency response and requirement of “clean” reproduction also increases the requirement of output power. Considering the aforementioned facts it may not be a bad idea to use some sort of “bi-amp” configuration. Many acoustic amplifiers of today have followed the path of bass amplifiers and switched to using lighter and more efficient switching topologies in both power amplifier and power supply circuits. Partial reason for this is that companies that manufacture good acoustic amplifiers tend to be the same ones that started out by manufacturing good bass amplifiers. Going “class-D” is a nice choice but unfortunately out of the scope of this book.

An acoustic instrument is typically either close miked with a conventional microphone or contains an internal – usually piezoelectric – microphone. Both require a different

interface than a magnetic pickup. This book has already discussed properties and requirements of piezoelectric pickups (the most important being the need for extremely high input impedance) so I feel there is no need for rephrasing the same content. Quite often an acoustic guitar also has an internal preamplifier with line level output, and – for the record – quite often those come equipped with highly versatile tone controls (such as a graphic EQ) as well. Due to existence of such active guitar electronics, as well as various outboard preamplifiers for microphones, acoustic instruments usually have fewer problems with proper interfacing than conventional electric guitars.

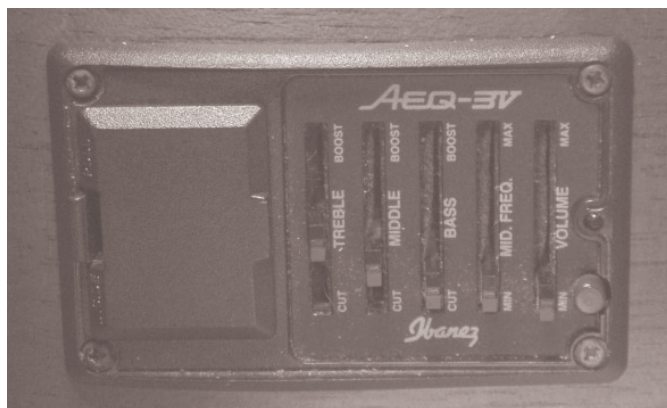


Figure 6.96 Built-in preamplifiers of acoustic guitars are often quite versatile. For example, Ibanez AEQ-3V is equipped with shelving graphic EQ for bass and treble plus a parametric control for mid-ranges. Tone-shaping capabilities of this preamplifier are far better than what the majority of guitar amplifiers can offer.

Special case is the interface for microphones, which I feel is quite necessary to present in some detail: Microphones are often “low-Z” devices, which means they have very low source impedance. By definition, anything between 150 to 600 ohms is considered as low-Z and anything between 600 ohms to 2 kilo-ohms as medium-Z. All microphones with 3-pin XLR connector are, by definition, low-Z. Connecting a device with low source impedance to a device with high input impedance (“impedance bridging”) causes no signal deterioration but the opposite configuration should naturally be avoided. Impedance bridging of microphones is usually done in a ratio of 1:10. Impedance matching usually only results to decrease in SNR and gain and is therefore avoided.

The nominal output amplitude of a typical microphone (sometimes inaccurately called the “MIC level”) is also extremely low: Normally it is in the order of few millivolts RMS. General consensus is that a microphone signal has to be amplified 40 dB (100x) to reach consumer audio line level (–10 dBV) and 50 dB (316x) to reach professional audio line level (+4 dBu). The exact values naturally depend on the microphone and those are as unique as guitar pickups.

It is common for some microphones to require *phantom power* in order to work. Phantom power refers to sending a polarizing DC voltage for the (condenser) microphone’s capsule. The same power supply is also used by the internal active electronics that convert the capsule’s extremely high source impedance to a proper low-Z output. The electronics are usually based on a FET (while tube microphones generally require a battery or an external device for power supply). Another condenser microphone type, an electret, requires no polarizing voltage, however, it still needs phantom power for the electronics that handle the impedance conversion. Dynamic

and ribbon microphones principally do not require any power to function but some of them may need it for powering the internal active electronics.

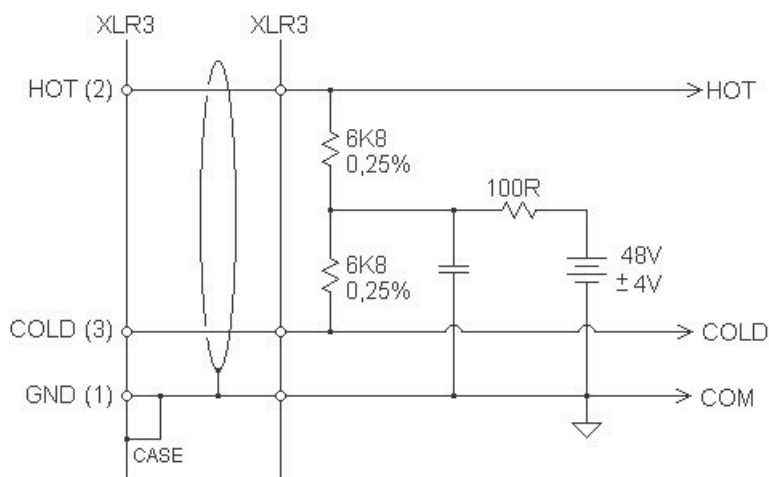


Figure 6.97 Concept of phantom powering. Note that the hot and cold leads carry 48 VDC and therefore the input must be AC coupled with either capacitors or an audio transformer. Typical input stage to follow is a differential amplifier.

There is little help from delving into design details of microphones but let it be mentioned that the details of phantom power concept are defined in DIN 45596 standard (which is at the time of writing referred to as IEC 61938 or DIN EN 61938): The 48 VDC phantom power is fed to signal lines through two 6.81 kilo-ohm resistors. The absolute value of these resistors is not critical but the standard requires that the resistor values should not differ from each other no more than 0.4 %. For the record, the typical tolerance choice of 1 % allows a maximum mismatch of 2 %. Fortunately “resistor sharing”, as shown in the schematic example (note the 100-ohm resistor), can improve the matching. Logically, good matching is vital for properly balanced operation and mismatches will increase the level of interference.

According to standard, a typical phantom-powered microphone should draw a maximum current of 10 mA and the power supply should provide a maximum of 14 mA, the latter during a short circuit of the supply line. For reference, a typical microphone has a current draw of about 2 – 4 mA. However, the maximum rating of current draw has been raised periodically and some digital microphones may need as much as 20 mA to work properly. In all conditions the supply should maintain its tolerance of ± 4 V. Another common, and standard, configuration (which is gaining more and more popularity) is to use a phantom power supply voltage of 12 VDC. In this case the standard value for the supply resistors is 680 ohms. Configuration using 24 VDC and 1.2 kilo-ohm resistors never caught success and likely it will be removed from further editions of the standard.

The single conductor 1.5 – 9 VDC “bias” or “plug in power” supply, used in some consumer-grade devices, is not the same thing as phantom power, although it is often inaccurately referred to being one. There are many versions of these and they have various names (usually each manufacturer also prefers to use their own setup). My advice is that the best practice is simply to ignore (or even boycott) all of these: Trying to provide an universal bias supply for a bunch of cheap microphones is not worth the effort and even more, microphones using such non-standard configurations are subject to being damaged by *true* phantom power systems.

The final distinctions of acoustic amplifiers are in the channel handling and tone control features. Voicing is, of course, a matter of taste and what is presented here is just a general guide. It was already mentioned that an acoustic amplifier should (by default) have a rather linear frequency response, which allows instrument's own tone to be reproduced without any annoying coloration. In my opinion, the versatility of the tone controls should not be overlooked but this is an issue that has already been discussed before. Let's just mention that any bi-amp crossovers may integrate nicely with the tone control circuitry.

Since acoustic instruments usually have a soundboard (e.g. guitar's "body") they are extremely prone for catching feedback. Consequently, an acoustic amplifier should include a "feedback eliminator" circuit, which is typically a high-Q, narrow-band notch filter. Naturally, the filter should have an adjustable notch frequency - otherwise the filter is pretty much useless. Do note that acoustic instruments (and microphones) typically have two feedback frequencies: Low frequency "howl" and mid-frequency "whine". Therefore actually two eliminator circuits are needed. Recently commercial amplifiers have started to employ circuits that perform automatic feedback detection and elimination.

Acoustic instruments (including vocals) have a remarkable dynamic range and therefore including features for compression and peak limiting (to prevent clipping) are more than gently recommended. Let it be mentioned that without the addition of slight compression many acoustic instruments (and especially vocals) tend to sound quite unnatural when reproduced through a loudspeaker system. Generally there is no need for the usual guitar's special effect - which is distortion - as it tends to sound quite pitiable with acoustic guitars. Special effects like doubling, delay, chorus, tremolo, reverb and echo may be far more useful additions. Modern acoustic amplifiers tend to employ DSP in creating these. Features such as FX loop and line level (DI) or tuner outputs are also extremely useful.

Requirements for channel selection and handling are also quite different when compared to conventional guitar amplifiers. In a typical guitar amplifier the musician may usually choose between two or three channels, which offer alternative voicing. For example, the channels are often configured to provide a clean tone and two distorted tones for lead and rhythm. In this case, a single guitar is also the sole instrument plugged into the amplifier. Acoustic amplifiers, on the other hand, are used in a similar manner as early guitar amplifiers: There is often more than one instrument (or signal source) plugged in. Typical configuration is an acoustic guitar and a microphone for vocals, both amplified simultaneously. Effectively, there may not be much need for channel switching (as all channels should be on at the same time). More useful arrangements could be switches for effects or specific voicing configurations.

6.18 Practical design example

The current trend seems to be that hobbyists, who are taking their first steps in guitar electronics, start by building an undemanding IC-based power amplifier kit, which is typically a variation of the popular "Gainclone" circuit. After completing the project they find themselves in a need of a simple preamplifier design that would match up their skill level. Such designs do not really exist so the easiest solution is to design a preamplifier. That way it will also lend itself to individual requirements for features and tones. Once you are familiar with basic circuit building blocks, such as OpAmp-

or transistor-based gain stages, or few of the most typical passive filters, designing basic preamplifiers becomes very easy.

The following section will present a simple dual-channel design that will get you started and give you a basic idea of the whole design process. The circuit is limited in its features and may have to be tweaked and modified to suit your individual needs. This is a task where you unfortunately must do your own footwork. If you have some preferences over certain features (like distortion tones) or voicing configurations the best practice is to hunt down various schematics of devices possessing these and study how they are accomplished circuit-wise. This book has so far given you a wealth of information that should help you in understanding the circuitry involved. The rest is up to you.

Anyway, let's first take a look at the very basic requirements and variables of the design starting from the power amplifier: A typical "Gainclone" is designed for HiFi applications and therefore has an input sensitivity of about $1 V_{PEAK}$. Input sensitivity means the input signal voltage that drives the amplifier into its maximum power output. If we do not know the input sensitivity we can determine it from the amplifier's gain and maximum output voltage swing. For example, a typical Gainclone amplifier has a gain of 22 and its output swings about $22 V_{PEAK}$. When we divide the maximum voltage swing by gain we get the input sensitivity (in V_{PEAK}). We now know the most important variables of the power amplifier.

Let's say we want to have a volume control potentiometer in front of the power amplifier and that we want to get the full output power when the potentiometer is dialed to "5". The potentiometer is logarithmic in order to compensate our natural hearing; this also means that at the halfway of the dial setting the output voltage of the potentiometer is one tenth of its input voltage. Effectively, the potentiometer must be fed with a $10 V_{PEAK}$ signal to meet the power amplifier's input sensitivity.

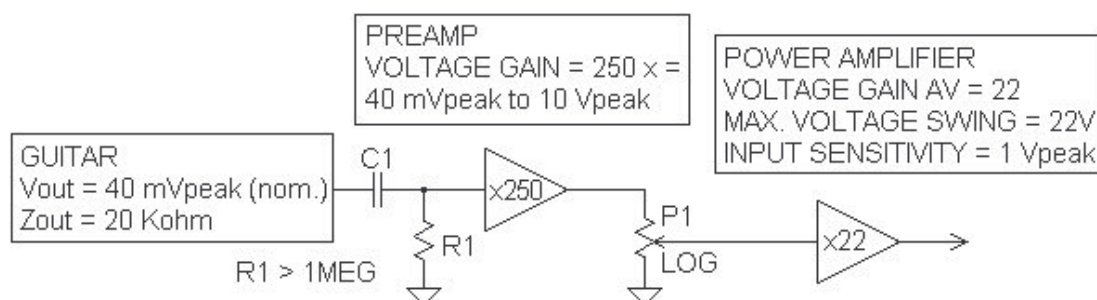


Figure 6.98 The design process starts from finding out the requirements and variables of the system.

At this point we also need to examine the source of the signal input, the guitar: It is impossible to say anything specific about guitars but we may assume that the nominal level of a guitar signal is about 40 mV_{PEAK} . This is far from the required $10 V_{PEAK}$ and therefore we need an interface that will introduce a voltage gain of 250. (The need for such high preamplifier gain is one of the drawbacks of Gainclone design but that's another story). The source impedance of a guitar is typically very high and therefore proper impedance bridging is required to avoid signal degradation and "tone loss". A convenient rule of thumb is to interface the guitar to input impedance that is at least 1 Megaohm. A capacitor separates the guitar from any DC potential of the preamp and protects the pickups in case of a preamplifier failure. The capacitor also protects the preamplifier from signal sources that have a high DC offset. Aside the input RC

circuit and the volume control potentiometer our preamp is still pretty much a “black box”.

The next design process is the decision of features. This typically presents a compromise between simplicity and versatility. We wish to have a moderately versatile but an extremely simple preamplifier, which means that the design can only have few very basic features. Also, the amount of gain stages and special effects must be limited to minimum. If we want features like reverb, cabinet simulated output, multi-band tone controls, protections against everyday abuse, “soft” dynamic processing, FX loop or foot switching capability the expense is always sacrificed simplicity.

In most cases, the classic feature expectations are:

- Two channels: One for clean and one for overdrive
- Sufficient tone and gain controls.

This is about the minimum requirement. As the intention is to keep the design simple some serious compromises must be made: 1) Aside few simple low- and hi-pass filters the clipping circuitry cannot contain any serious voicing circuits. Band pass or band stop filters, that would make the distortion sound really nice and “individual”, would introduce too much complexity and therefore must be omitted. Unfortunately the requirement for simplicity means that we must live with less-ideal distortion tones. 2) Both channels will share the same tone control stage. Individual controls for both channels are much more versatile but tend to bloat the design. 3) Similarly, both channels will also share many gain stages. This decreases component count by a great deal and also simplifies the design a lot. However, it also weakens the possibilities to voice each channel individually. 4) Tone stack will be a basic, passive FMV design. Active, multi-band designs are nice and many of them are much more versatile than FMV, unfortunately they will require more components and gain stages. Drawbacks caused by this simplification are, in my opinion, appalling but it should be noticed that the goal is to design a *simple* preamplifier, not a super versatile one.

Features the design absolutely must have are: 1) Gain control for the overdrive channel. A combined volume and gain control simply will not suffice. 2) Individual volume controls for each channel. These allow balancing the loudness levels so that you can sovereignly switch from channel to channel without having to tweak the volume levels in the process. 3) A master volume control. One must note that these design choices only reflect the preferences of the author.

The aforementioned are likely ideas the designer has already pondered for quite a while. The next step is to sketch a function block diagram of the preamplifier. This should preferably include all individual gain stages and circuits, although the way they are exactly realized is yet of lesser importance. Do not overlook this step: *This is basically the most critical part of the whole design process.* Thorough planning in this stage will tremendously ease the final stage of the design, which is designing the actual circuit. The function block diagram will show you how to put things together, and what design challenges will come up when you do so.

The first principle we deduced was that our preamplifier must have a gain of 500. When gain levels in each stage are kept within reasonable limits this will require about three or four gain stages. The various potentiometers, diode clipping circuits and tone

controls will introduce some signal losses along the way. They must be acknowledged and the lost gain must be recovered somehow. Fortunately the function block diagram tells us where each of these parts is located so planning the stage gains becomes easy.

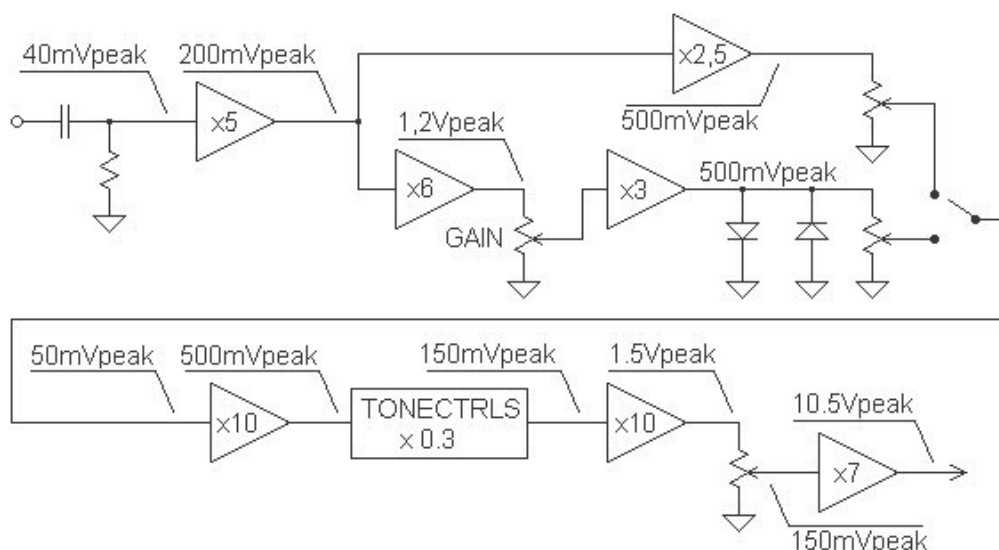


Figure 6.99 Function block diagram plots the features and clarifies the gain requirements per each stage.

In this design, both channels share the first gain stage. Why: The signal at this point has a very low amplitude and pretty much the only processing that can be done here is to amplify a little to improve the SNR. Only after the first gain stage it becomes reasonable to split the signal path to feed the individual channels. The diode clipping circuit of the overdrive channel introduces another limit: The amplitude at this stage is limited to about 500 – 700 mV_{PEAK} regardless of preceding gain. If we’re going to “mix” the signals of the two channels without any additional “gain recovery” stages it makes no sense to amplify the signal of the clean channel any higher than this level. The channel-specific volume controls must naturally be located before the “mixing” stage. These are logarithmic potentiometers and we assume that they normally attenuate the signal level to one tenth. This corresponds to dial setting of “5”.

At this point it is obvious that the signal amplitude of only 50 mV_{PEAK} is still quite far from the goal of 10 V_{PEAK}. We also need to insert a passive tone control somewhere and it will have a nominal “gain” of about 0.3, which means a significant attenuation. Since signal amplitude is fairly low it is best to increase the signal level with a gain stage before the tone controls in order to maintain a satisfying S/N ratio. Another “gain recovery” stage is put to follow the tone control circuit. After these stages the nominal signal level is still only 1.5 V_{PEAK}.

Now, at the following stages we could amplify the signal to the required level and have the master volume potentiometer as the final circuit element. However, there is a better solution: The master volume control is put to follow the recovery stage of the tonestack and the final gain stage is configured to follow the master volume control. This decreases the amount of series gain stages that lack a volume control. Thus we can use the master volume control to prevent overdriving of the final gain stage and to dial in some *headroom*. Secondly, a gain stage as a final link effectively buffers the preamplifier and isolates the volume control from the input of the power amplifier. This is beneficial in many systems. We can also design the final gain stage to contain a soft limiter setup that never allows exceeding the input sensitivity of the power

amplifier. Since this is intended to be a simple design the latter feature is omitted, though.

Now that we have a clear vision of the stage layout the rest is just basic designing using the circuit blocks we are convenient with. Since this design aims for simplicity the natural choice is to use OpAmp-based gain stages. Due to their awesome linearity they are an excellent choice for the clean channel in virtually every design. For the overdrive channel they are a sufficient solution. Cascaded FET stages, tubes or complex soft clipping schemes are simply out of the question due to their usual complexity and higher component count. The complete circuit is shown in Figure 6.100. All stages are basic inverting or non-inverting amplifiers. Shunt clipping diode configuration shown in the function block diagram is replaced with a configuration where diodes are in the feedback loop of an inverting operational amplifier. Channel switching is manual but the circuit shown in Figure 7.17 could directly replace the switch SW1 without any further alterations to the circuit. Gain levels are rolled down a bit from the original plan.

The next process involves deriving the component values. This is the “artistic” part of the design and basically one could use the same schematic and still create dozens of circuits that would all sound different – just by varying the component values. This voicing process has few circuit-dependant principles that will be discussed soon. Anyway, a pleasant way to start is to ignore the RC filter circuits for a while and just derive the basic resistor values from required gain figures. There are few important issues in this procedure, though. All of them relate to impedance bridging. 1) The signal source (guitar) must be impedance bridged to a load that is preferably at least 1 Megaohm. 2) All “passive” potentiometers (in this case the volume controls) should be impedance bridged to a load that is about ten times higher than the potentiometer’s resistance. 3) The response of the passive tone control circuit is affected by both source- and load impedance. To minimize losses, the source impedance should preferably be very low while the load impedance should be very high. 4) OpAmps have high input impedance and low output impedance. It means that an OpAmp gain stage can pretty effectively buffer all typical circuits and that we must only focus on the input impedance defined by the input resistor. 5) Using too high impedances is begging for trouble since it accentuates the capacitive effects. This may turn the circuit instable or cause notable high frequency losses due to shunt capacitances. Using too low impedances makes no sense either because it forces to use very large capacitance values. Examine the example values of the design and see how they fulfill these clauses.

Once the resistor values are worked out we come to filter design, which is basically the stage where we “voice” the circuit to our liking. In this circuit, C1 and R1 form a hi-pass RC filter. This is the first one out of the many that will introduce gradual bass frequency roll off. Other hi-pass filters are formed by capacitors C2, C3, C4, C6, C8, C13, C16 and C17 along with the resistors involved. Together these form a high-order filter that should limit the low frequency response very effectively. Special cases are filters formed in conjunction with C3 and C4. These are the overdrive channel’s hi-pass filters that precede the overdrive stage and therefore have a serious effect to clipping tone. Preferably they should introduce enough low cut to avoid harsh “farting” during clipping. Too much low cut, on the other hand, will make the clipping sound “thin” and weak. The “right” values are highly subjective.

Capacitors C5, C7, C12 and C15 form low-pass filters. All others than the C5 are mainly stabilizing the gain stages by limiting the bandwidth outside hearing range. Too high values (that will make the filters cut notably at audible frequencies) will make the circuit sound “muddy”. C5, on the other hand, is a special case since it is used for attenuating the harsh high order harmonics resulting from clipping. Again, too high values will make the sound “muddy”. Too low values cause the clipping to have a “harsh” and piercing, obnoxious tone that hurts ears. In extreme, a too low value may introduce oscillation.

Now, it should be noted that once the schematic is completed the real design process *merely just begins*: First, we need to think about the issue of powering the circuit. A typical and good solution is to, with the aid of two regulators, derive the ± 15 V rail voltages from the supply rails that power the power amplifier stage. A satisfactory rule of thumb to calculate the current draw is to multiply the number of OpAmps with 5 mA. The supply must be well filtered and power pins of each operational amplifier should be bypassed with 100 nF capacitors.

The next step is to plan out a component layout and sketch up the circuit board. The section 9.8 discusses issues related to this stage more comprehensively. Once the board is assembled and all issues concerning mechanical design (for example, how to mount big components securely, how to mount the board to the chassis, how to route the wiring etc.) are worked out the circuit should preferably spend its few first weeks at the clumsy hands of those non-technical persons who abuse their equipment with wacky experiments one after another. (For example, plug power amplifier’s speaker output to the preamplifier input, drive the input with obscure signal sources, plug devices in and out without turning down the volume level or the amplifier itself etc.) Preferably these “beta testers” should also gig frequently and treat their equipment as roughly as possible while doing so. After the device has reliably lived through this testing period you can proudly announce that the design process has ended. When problems *do* occur, they will teach you invaluable lessons of how to build and design reliable equipment. As is, the simple preamplifier pretty much lacks all protections against the described, severe abuse.

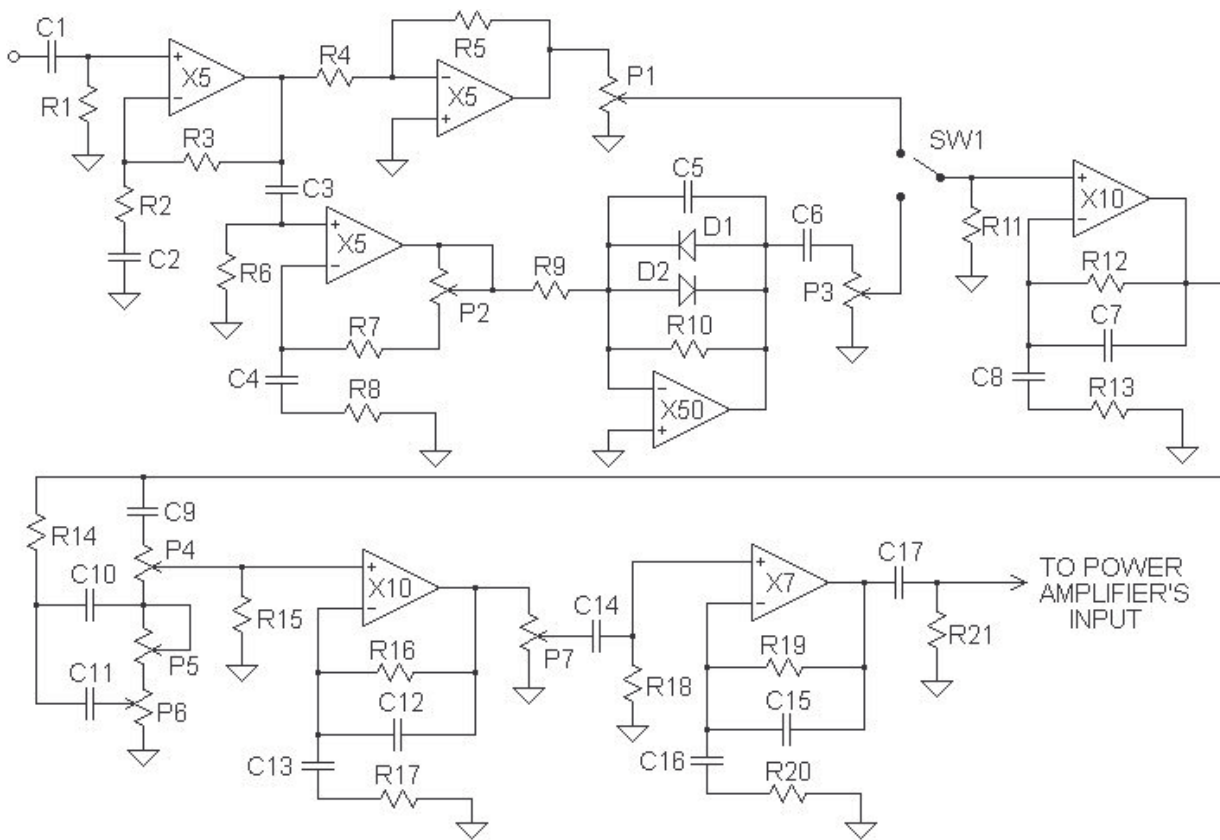


Figure 6.100 Preamplifier design example.

R1	1 MEG	C1	56 nF	P1	10 K Log
R2	12 K	C2	520 nF	P2	50 K Log
R3	47 K	C3	22 nF	P3	10 K Log
R4	10 K	C4	56 nF	P4	250 K Linear
R5	25 K	C5	470 pF	P5	1 MEG Log
R6	470 K	C6	1 μ F	P6	25 K Linear
R7	10 K	C7	47 pF	P7	10 K Log
R8	12 K	C8	520 nF		
R9	33 K	C9	470 pF	SW1	Any SPDT
R10	100 K	C10	22 nF		
R11	100 K	C11	22 nF		
R12	47 K	C12	47 pF		
R13	8,2 K	C13	520 nF		
R14	33K	C14	1 μ F		
R15	1 MEG	C15	47 pF		
R16	22 K	C16	520 nF		
R17	8,2 K	C17	22 μ F		
R18	100 K				
R19	47 K	D1	1N4148		
R20	10 K	D2	1N4148		
R21	470 K				

OpAmps can be any generic dual and FET input ones. The supply rails are bypassed with 100 nF capacitors near the power pins of each operational amplifier. Rail voltage is preferably ± 15 V.

7 Miscellaneous circuitry

7.1 Light bulb current limiter

Light Bulb current limiter is a simple and useful little tool for troubleshooting and testing designs that cannot be powered directly from mains network. This device is also probably the best use for a spare mains receptacle and light bulb socket one can imagine and will gain its price back in short time by saving precious components from blowing up during power on. Amazingly the device has caught way too little amount of attention and far too many people misunderstand how it should be built and used.

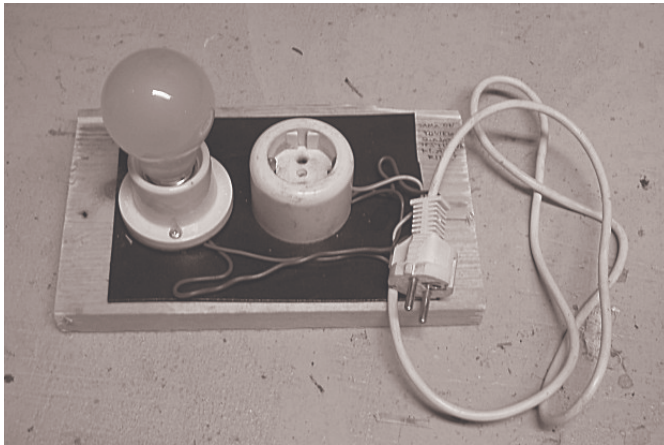


Figure 7.1 A crude, yet excellent light bulb current limiter can be built out of scrapped electric parts.

First of all, the light bulb limiter should never replace a conventional fuse: A fuse is a safety device that is supposed to trip under excessive current and open the circuit. However, even under a “hard” short this will take a finite amount of time and usually the short circuit current is capable of doing some real damage before the fuse blows. On contrary to a too common belief, *fuses do not protect the device; they protect the user from a faulty device!* A light bulb current limiter, on the other hand, is a protection for the device - not for the user - and *fitting one permanently inside a circuit is hazardous to health* since the bulb limits current, thus preventing fuses from tripping under fault conditions. The current, although limited by the bulb, may still be high enough to kill you!

How does it work? Current limiting light bulb acts like a current dependant resistor: When the filament current increases the filament heats up and its resistance grows higher. The added resistance is usually high enough to turn a “hard” short into a “softer” one (i.e. increases the short circuit resistance) – which is often more tolerable by delicate circuitry. Also, the light of the bulb acts as a handy indicator of excessive current demand. In normal operation the light bulb should glow brightly about a second after turning the device on and then dim to slight orange glow or no light at all. (There is practically a short circuit during capacitor charging, after this the circuit starts to show a high resistance). It is also perfectly normal for the light to brighten up if the tested circuit is loaded harder. In some cases a “pulsating” behavior reveals oscillation that otherwise might just blow the amplifier.

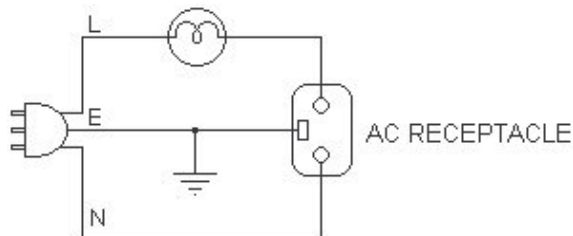


Figure 7.2 Light bulb current limiter.

The basic light bulb current limiter is a very simple circuit so it's amazing how so many people can wire it up incorrectly. The light bulb is connected (and only works as a limiter) *in series* with the load introduced by the tested device and that is basically all one needs to know. The power rating of the lamp is proportional to resistance of the filament. The basic Ohm's law $R = U^2 / W$ should tell this already. In short, the resistance introduced by the bulb is lower the higher the power rating of the lamp is. Safety ground continuity is naturally maintained throughout the circuit and no fancy ground lift circuits should be used. If switches are required there can practically be two of them: One breaking the continuity of *both L and N* and one bypassing the light bulb with a short. More complex switching could be used to create a sort of a "Variac" by adding more light bulbs in series. Basically, I don't recommend intricate switching configurations as they introduce (and increase) the possibility of human error. In the basic arrangement - without any switches - one definitely knows when the limiter is on and when the tested device is connected to mains.

7.2 Fans

Fans offer a very practical solution to keep heat sinks or other parts of the circuitry cool. If treated right, the typical 12 V fans are very reliable and can outlive many other components. Most fans are considerably noisy, which might become an issue when the amplifier is put to service in studio or in quiet practicing sessions. In live situations the stage loudness and even the ambient noise is sufficient to cover the noise made by the fan. If one wants to get meticulous about this, the usual options are either to run the fan with lower voltage (not recommended), replace the fan with a quieter model, improve the suspension of the fan or, more commonly, activate the fan only when a certain temperature limit is reached. The latter approach is better from the efficiency point of view as well.

The most common fan sizes are 60, 80, 92 and 120 mm. Larger fans generally generate more airflow and operate at slower speeds and due to this they generate less noise. The noise also has a lower pitch. However, larger fans generate less air velocity and pressure than smaller ones. This may become a problem since both attributes significantly decrease efficiency if airflow becomes obstructed. Stamped grill structures may also contribute to noise by blocking as much as 50 to 70 percent of the intake opening. The type of bearing usually affects the fan's performance and especially the amount of noise the fan emits. Fans with sleeve bearings are generally the cheapest ones; they are less durable than the others, intolerant against high temperatures and perform poorly in any other mounting orientation than the horizontal one. Ball bearing fans are substantially better: They do not suffer from orientation limitations; they tolerate more heat, are more durable and produce less noise. Fluid bearing fans are even sturdier and less noisy. Unfortunately they are generally quite expensive. Magnetic levitation (MagLev) fans have a levitating rotor, which usually makes them rather quiet.

Fans typically draw about 200 mA to 1 A of current, which means their typical winding resistance is about 12 to 60 ohms. It should be noted that fans draw more current when they are operated with lower voltages, on account of having less torque. A 12 V fan that would normally draw about 600 mA may draw nearly an ampere of current when running on 6 V. Since the currents are quite high it is not a practical solution to power fans from high rail voltages using voltage-dropping resistors; the power loss over them gets really high. Because fan is a motor it will create all kinds of noise to the supply rail, which unfortunately is very often the “sensitive” preamplifier rail. The best practice is to isolate the fan circuitry completely and use a separate secondary winding or a transformer for powering it up.

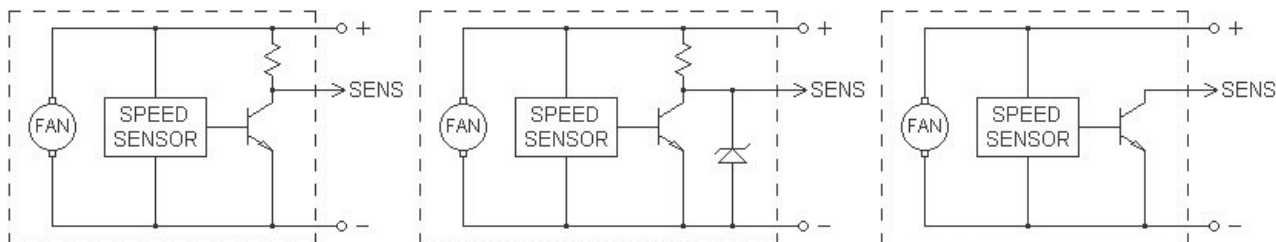


Figure 7.3 Fan sensor outputs. Left: Pull-up resistor is TTL incompatible, middle: Zener diode ensures TTL compatibility and right: Basic open-collector has maximum flexibility but requires external resistors.

Modern fans usually contain a safety circuit that prevents them from turning on if the supply voltage rises too slowly – even if it would eventually reach 12 V! A best measure is to immediately feed the fan with more than 8 volts to make sure it will really turn on. In practice, all fans are better to be driven using solid-state switching that actuates when certain conditions are met. Relays are another option but they tend to be more bulky and expensive. Some fans intended for use in computer cooling have a third wire that monitors the rotation of the fan. This monitoring can basically be categorized into alarm and speed sensors. Alarm sensors typically output digital signal pulses when the fan speed has dropped a certain amount; unfortunately the type of these alarm signals usually depends on what a certain manufacturer has considered as a good idea. Speed indicators output series of digital pulses at a frequency proportional to fan’s RPM. Standard is to give two pulses per revolution. Since amplifiers using fans are also dependable on them in most cases, it is advisable to exploit the rotation detection and devise a circuit that shuts down the amplifier if a fan malfunction is detected.

Table 7.1 Typical Connector pin assignments

Typical fan	
Black	Common
Red	Supply
Yellow	Speed sensor (RPM)
PWM fan	
Black	Common
Red	Supply
Yellow	Speed sensor (RPM)
Green	PWM wire

Lowering the fan voltage is not necessarily as good idea as it first seems: Most importantly, the airflow and cooling effect is proportional to revolution of the fan, which (in case of DC fans) is dependable on the supply voltage. Lowering the supply voltage will reduce both the noise and *efficiency* of the fan. Note that the internal speed sensing circuitry is powered from the same rails as the fan, so any tampering

with supply voltages will also mess up its operation. Greatest causes for fan noise are obstacles in airflow and improper suspension that creates resonance with the case – or simply just using a fan that has poor-quality bearings. Focusing on these details, or acquiring a less noisy fan, is usually a better solution to decrease fan noise than lowering the revolution rate.

7.3 Clipping indicators

Clipping indicator is a little misleading name for circuits that for the most part are actually *clipping detectors*. The *indicator* part of them is just a lamp or LED - and possibly a driver for it, such as a transistor. Because of the separation, most clipping detector circuits are perfectly compatible to control VCAs and can therefore work as limiters that kick in when amplifier starts to clip. (For reference see Figure 6.71). I have made a division here since a usual detector circuit for a limiter works based on amplitude while a clipping detector circuit works based on occurrence of clipping.

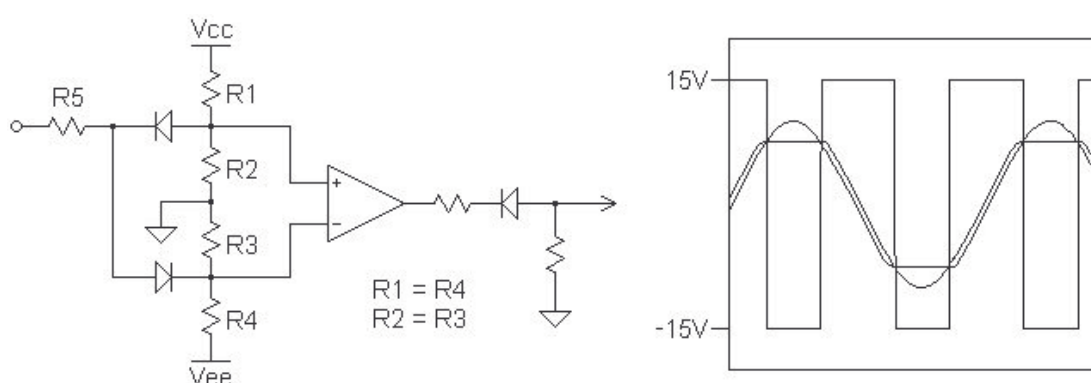


Figure 7.4 Clipping detector based on a comparator.

A common approach in realizing a clipping detector circuit is to use some kind of a comparator. An operational amplifier without feedback basically acts as one, meaning that the output assumes the polarity of the most positive input. Let's examine the behavior of the circuit in Figure 7.4: When the voltage potential of the non-inverting input is higher than the voltage potential of the inverting input the output of the comparator assumes the potential of the positive rail. This is the operating condition without clipping. Resistor networks R1, R2 and R3, R4 control the voltage potential affecting over the diode, thus they, in conjunction with resistor R5, determine when the diodes clip the input signal. If this happens it will cause the inverting-input to have higher potential, which in turn shifts the comparator's output to polarity of the negative rail. The diode in the output passes only the negative signal through, which is theoretically all that is needed to drive a VCA, LED, transistor or a FET. In practice, an attenuating network and some capacitance to add up some hysteresis is probably needed as well. What should be noted is that this circuit does not, and cannot, actually detect whether the amplifier is clipping, it just *estimates when this might happen*. In most cases the estimation is also correct so we're ok with that. The rails feeding the resistor network are the rails of the power amplifier or the gain stage being "monitored".

Another plausible circuit for clipping "detection" is a window comparator. Like the previous circuit this one does not actually detect clipping either but estimates it based on input amplitude (A) and clipping threshold voltage (B and C). Each time A exceeds either B or C the comparator triggers. R3, R4 and R5, R6 are simple voltage dividers that define B and C. Voltage divider R1-R2 defines the "zero point" reference of A.

Usually it is between rails equaling zero volts. Note that resistor R7 forms a voltage divider together with resistors R1 and R2, which means that voltages B and C have to be scaled according to that. Comparator output shown in the plot (the square wave) is not in any particular scale with other amplitudes, rather just an example of the output.

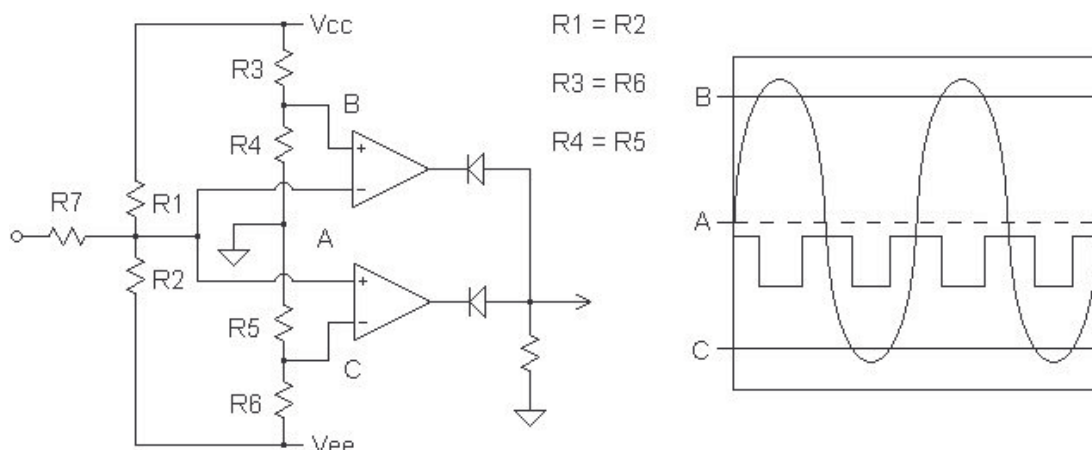


Figure 7.5 Window comparator.

Question rises why not just compare attenuated output signal with the input signal in order to *really* detect clipping. This works in theory but in reality the idea is not that plausible: The comparator will trigger abruptly from any difference between input and output signals. In addition to clipping, this also includes effects of bandwidth limiting and phase shifts caused by power amplifier circuitry and speaker load. Basically the sampled input signal should be treated with an “emulation” of the power amplifier stage, which is virtually impossible.

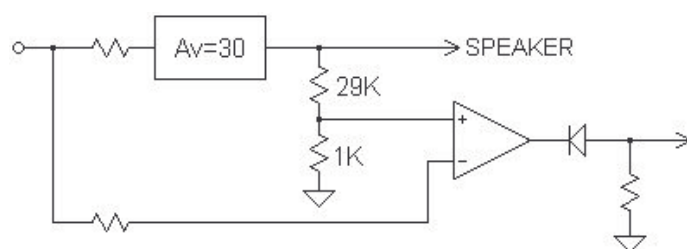


Figure 7.6 Comparing output and input does not qualify as plausible clipping indicator.

There is another circuit that utilizes the aforementioned idea but does it in a slightly different manner: The external resistor divider is just replaced with the one that is already in the feedback loop and the non-gradually operating comparator is omitted. In normal operating conditions the input signals of a long-tailed pair differential amplifier are almost identical having perhaps only few millivolts of difference. However, this condition changes when the circuit clips: The feedback signal is now clipped and the differential amplifier loses control over signal correction. If another differential amplifier (such as an OpAmp) is hooked up to the inputs of the long-tailed pair any difference between the inputs can be amplified. The output of this circuit can then be treated accordingly. Note that this method is, again, more a complete “distortion detector”, as it increases detector’s output signal level at higher frequencies where phase shift and slewing distortion increase. Therefore the circuit cannot work accurately as a limiter when hooked directly to circuit controlling signal amplitude (i.e. VCA). Peavey amplifiers with DDT (Distortion Detection Technique), for example, solve this problem by hooking up the output to a circuit that is similar to one shown in Figure 7.4. The comparator’s output is then used for controlling an OTA, which in turn controls the gain of the stage preceding the power amplifier. Some

amplifiers just rectify the output signal and drive a clipping/distortion indicator LED. Possibilities are endless.

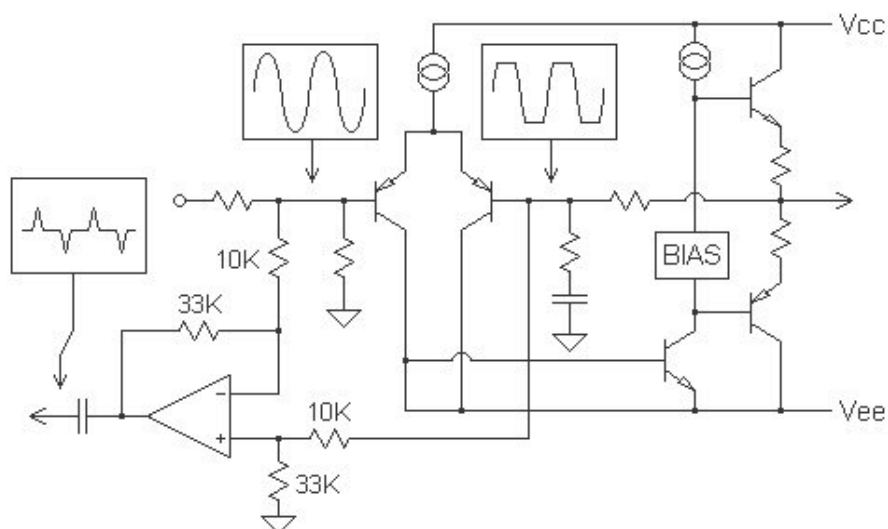


Figure 7.7 Amplifying the difference of differential inputs.

7.4 Switching circuits

7.4.1 Jack switches

Jack switching schemes are the most simple ones since they are purely mechanic. The common nominator for all jack-switching designs is that inserting a plug to the jack (or removing it) actuates the switch. The most basic switch was created by accident and it can be used in creating a connection to common (thus creating a closed circuit) when a plug is inserted into jack. This switching mechanism uses TS (mono) plug and a stereo jack. When plugged into a stereo jack, a TS plug effectively shorts the sleeve and ring together. Since in most applications the sleeve is hooked to common potential any node connected to jack's ring terminal will be "grounded". The trick is very common in guitar pedals and one can likely find applications for it from amplifiers as well.

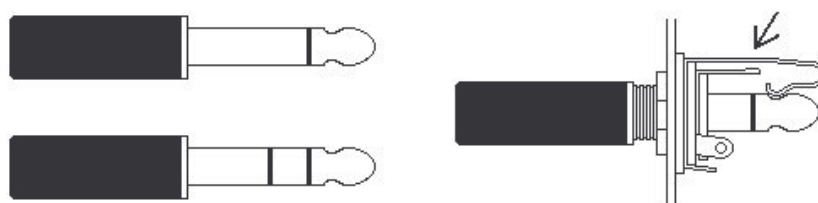


Figure 7.8 Left: Mono (TS) plug shorts the connection of ring and sleeve. Right: Basic switch jack has an internal spring mechanism that breaks connection to tip when a plug is inserted.

Another typical switching jack comes with a normally closed (NC) contact that connects the tip when no plug is in the jack. Stereo jacks commonly provide two such contacts. This feature was originally used (and still is) to shunt all unused inputs to common - a procedure that may drastically cut the amount of noise in some amplifier, as well as keep them stable. (Some circuits oscillate with no input connected). Effectively the arrangement has proven to be most useful in creating effects loop circuits and alike.

Figure 7.9 shows some basic jack switches. Black arrows indicate electrical contact and white, larger "arrows" mechanic. Thinking that the tip and ring contactor will

bend when plug is inserted will help to examine the schematics more clearly. The depicted schematics show the most usual arrangements one will encounter but, for example, at the time of writing *Switchcraft lists 37 different switching combinations in their catalog!* Bottom line is: When acquiring jacks with switches, *always* test that their arrangement is similar to that required. In addition to switch arrangement the pinout may be different as well. The variety of different configurations and pinouts is one of the reasons why jacks should never be mounted on a circuit board. For example, one can convert jack with isolated SPST switches (c) into a jack with normally closed connection to tip (a) but this is usually plausible only when the jacks are not mounted on a circuit board.

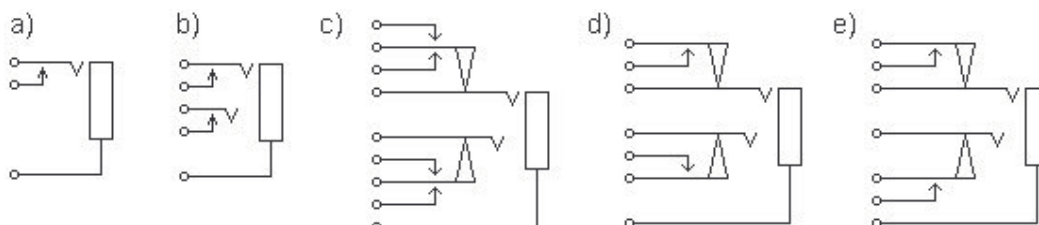


Figure 7.9 Common jack switch arrangements: a) TS jack with connection to sleeve when no plug is inserted, b) TRS jack with connections to tip and ring when no plug is inserted, c) TRS jack with two isolated SPDT switches, d) TRS jack with two isolated SPST switches – both normally closed and e) TRS jack with two isolated SPST switches – one normally closed, one normally open.

7.4.2 Switches in remote boxes

Controlling signal continuity, shunting signal to ground or controlling whether an electronic circuit is closed or open are likely the most used switching methods. All this switching can be done remotely (i.e. using switches located inside a footswitch box) but in most cases this will cause trouble: Leads to remote switch unit are usually very long, which means that they introduce resistance along with parallel inductance and capacitance, they may catch interference or cause ground loops and in worst cases trigger oscillation. It is preferable to keep signal path lengths as short as possible by keeping switches inside the amplifier chassis and using remote units only to control the switches. Practically, simple circuits where the switch is located in a remote box outside the amplifier work well with circuits that shunt part of the signal to common. “Proven” application is, for example, shunting the spring reverb tank’s output to ground (reverb toggle). Even this circuit will benefit from bringing the switch inside the amplifier. If you for some reason wish to (or must) use a remote signal line switch then at least buffer the stages feeding the cable.

7.4.3 LDR switches

LDRs qualify as “switches” since some of them have an enormous resistance in dark and moderately low resistance in light. LDRs are not as ideal as ordinary switches but since they require no mechanical action their controlling arrangements can have a lot more versatility. For the purpose, one will likely prefer LDR models with highest “off-state” resistance and lowest “on-state” resistance as possible.

The operation of depicted LDR switching circuit should be pretty self explanatory to anyone but let’s examine it anyway: Let’s assume the LDR has an on-state resistance of 4.7 kilo-ohms and off-state resistance of 4.7 Megaohms – maybe not an optimal component but a pretty realistic estimation of a LDR anyway. A circuit controlling the switching will take care that while the other LED is on the other one is off. Effectively

there are two signal paths to output but due to high off-state resistance of the LDR element the other one of them will always be highly attenuated. In order not to attenuate the “effective” signal path too much the preceding stage should have low output impedance since that, in series with the on-state resistance of the LDR, forms a voltage divider with the load impedance. The example circuit selects between different “low cut” configurations. The same principle works as well in channel switching.

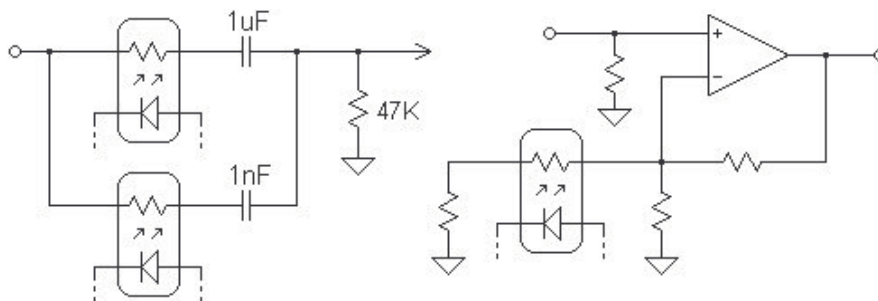


Figure 7.10 Configurations to select between two alternate signal paths or boost up gain.

To show few examples, the circuit configuration on right can be used to increase gain or to form “treble boost” circuits that “bleed” high frequencies across attenuating elements like volume controls. In off state the high resistance of the LDR element causes the impedance of the alternative signal path to be high enough to be ignored. In on state the resistance of the LDR is moderately low so the alternative signal path will have effect. In the example circuit this effect is increasing gain of the operational amplifier stage.

LDR switches can as well be used to toggle alternative signal sources like reverb amplifiers. The configuration to do this is not as straightforward as one might first think: In the scheme shown left in Figure 7.11, switching on the LDR provides the signal another considerable path to common (through the alternative signal source R_S) and this path, in parallel with the 47 kilo-ohm resistor, increases attenuation of the signal. As a result, signal’s amplitude will be lower when the external input is switched on. There is really no effective way to battle this: The source impedance of the external input can be raised but this will attenuate its signal and therefore is not a very good solution either.

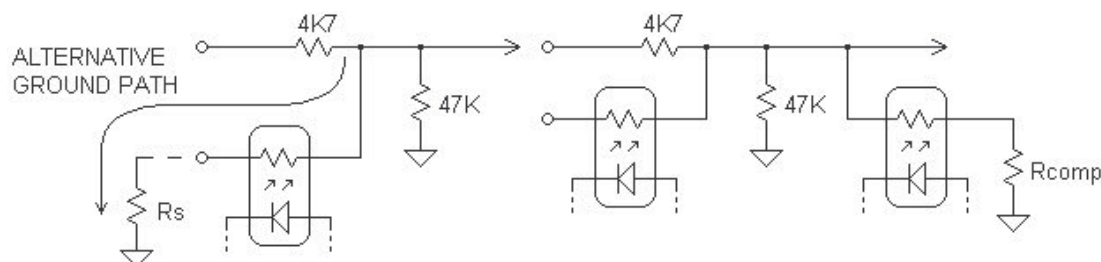


Figure 7.11 Configurations for mixing another signal source. Left: The wrong method and right: The sophisticated method.

A compensating configuration is a bit more complex and uses two switches. When the external signal source is “connected” the LDR at right is closed and presents a high resistance for the signal. The circuit still has the attenuating effect but one just has to live with that: When the external signal source is “disconnected” with high resistance of the LDR on the left, the LDR on the right “opens” creating a “compensating” path to common that attenuates the signal. The resistor values should be matched so that the signal has equal amplitude in both states.

The greatest merit of LDR switches is that they require no bias voltage and can therefore control even quite high signal amplitudes. Typical voltage limits for LDR switches range from 100 to 200 volts. Solid-state switches generally are incapable of dealing with such high amplitudes. LDR switches also do not introduce a low pass filter like the solid-state ones do. The drawbacks are the typically high on state resistance and only moderate off state resistance (which in some amplifiers leads to disturbing “signal bleeding” between channels). However, some LDR elements can offer an outstanding performance with characteristics in the order of $R_{OFF} = 150$ ohms and $R_{ON} = 50$ Megaohms.

7.4.4 Bipolar transistor and FET-switches

Semiconductor switches controlling analog signals are known as *signal or analog switches*. Optocoupler (“opto-isolator”) transistor switches are likely the simplest ones of them. Instead of base current, a phototransistor is controlled by LED’s luminance, which means that the transistor can be driven into saturation when the LED has its highest luminosity. Conventional BJT switching of analog signals is pretty self-explanatory and follows the general principles of all BJT switching. The AC input signal is converted into DC (with AC component) in order to make the switch to function. To do this, the signal at the input (which is the collector node) is biased to suitable DC potential, after this the conduction of the transistor is controlled by the means of varying base voltage.

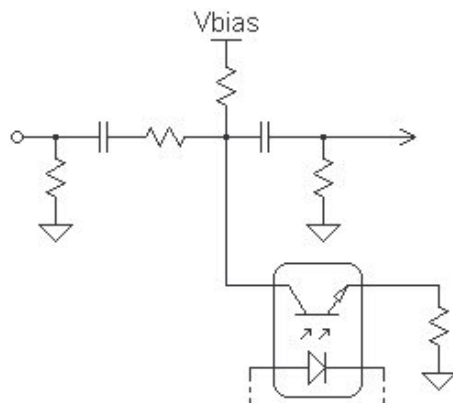


Figure 7.12 Phototransistor as analog shunt switch.

A simple application is a shunt switch that usually shunts part of the signal to ground – or alternatively splits the signal into two paths. In the schematic the arrangement is depicted as a simple shunt that can be used to control gain – or more accurately, the amount of signal attenuation. When the phototransistor switch is “open” (LED is off, transistor does not conduct) the signal is attenuated by the voltage divider formed by the input resistor, bias resistor and load resistance. When turning the LED on closes the phototransistor switch an alternative resistance, formed by the transistor’s internal impedance and resistance at its emitter, adds up. As a result the signal is attenuated more. Note that at this state an exact copy of the collector signal is present at the transistor’s emitter; if the transistor won’t conduct the emitter signal is near zero. This means that another type of switch, toggling the signal on or off, can also be realized.

Figure 7.13 shows some switching arrangements for conventional BJTs. These circuits work but they can easily be surpassed with FETs or LDRs. The circuit a) is series in the signal path and controls the collector-emitter resistance. When the transistor’s base is tied to ground with toggle switch no base current will flow, thus transistor does not conduct and is presenting a high resistance for the signal. When base is instead tied to switch supply potential, the base current will turn the transistor on and it starts to present only a moderate resistance for the signal. The arrangement is more usefully exploited in circuit b) where two parallel switches are used in selecting a signal source. An external control circuitry makes sure that switches are not on

simultaneously. These circuits are dealt later. Circuit c) is a shunt switch adapted for use with conventional BJT.

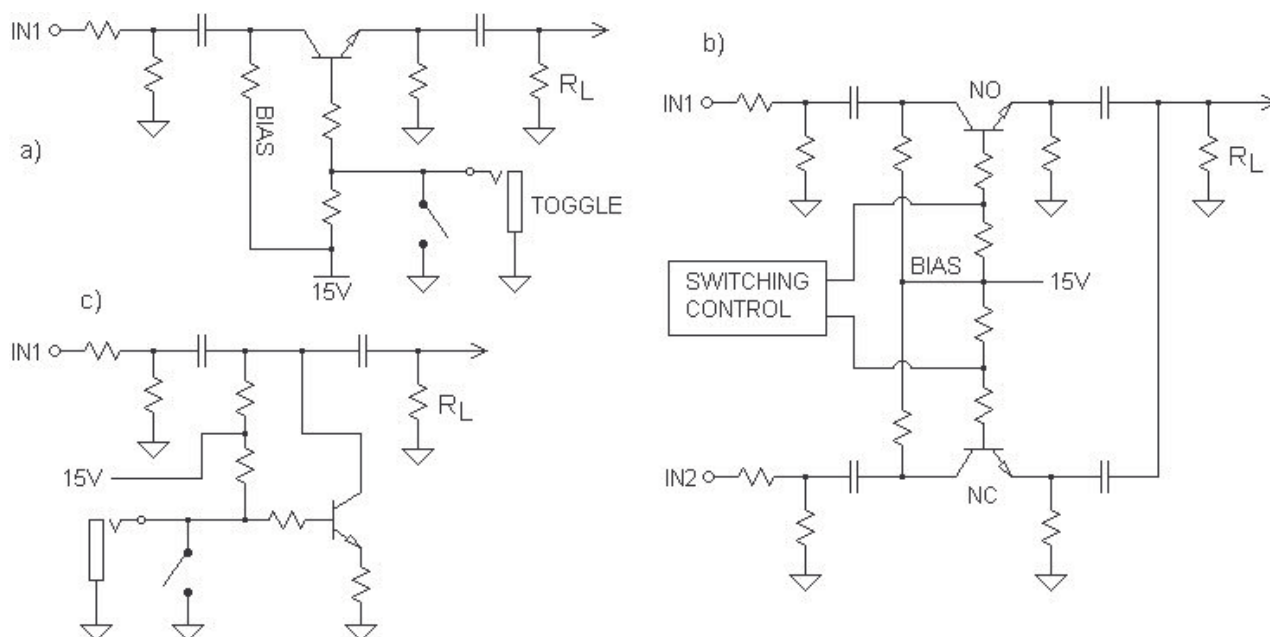


Figure 7.13 Bipolar analog signal switch configuration. a) Series in signal path, b) channel selection and c) shunting signal path.

With all BJT circuits it's important to know the current being switched. After the load current is known it is easy to calculate the value for the current limiting base resistor, it is:

$$I_B = I_C / h_{fe} \rightarrow R_B = V_B / I_B \quad \text{(Equation 84)}$$

Use the worst-case (minimum) value of H_{FE} listed in the datasheet for the calculation.

BJT switches cannot reliably control signals that have amplitude levels exceeding less than half of the controlling voltage. Past a certain point a saturated transistor will just clip off the top half wave of the signal. There are other drawbacks too: To increase the input voltage range the switch can handle, the circuits usually require a quite low bias resistor value. Since the bias resistor forms a voltage divider together with the source resistance the signal losses over the switch circuit can become quite high. The on state resistance of a transistor is quite high so it cannot attenuate the signal completely when in shunt mode either. Due to DC potential the switches also need coupling. Best practice is just to avoid using BJTs in analog switching applications completely since their performance can easily be surpassed with FETs. In load and "logic" DC switching applications BJTs work fine.

Similarly to BJT, a FET switch operates as a voltage controlled resistor that has a high off state resistance and a low on state resistance. In switching application the drain and source functions of N-channel JFETs are interchangeable, which means the "drain" will either be at the input or output end of the FET depending on whether the gate is forward or reverse biased. This allows the output voltage track the input voltage. Essential part of operation is controlling the biasing of the diode: If the diode is reverse biased (gate control voltage is at least V_F higher than most positive input signal amplitude) the JFET switch is "closed" and signal passes through. If the diode is forward biased (gate control voltage is at least V_F more negative than the pinch-off

voltage of the FET) the switch is “open” and the FET presents a high resistance. V_F is the forward voltage of the diode.

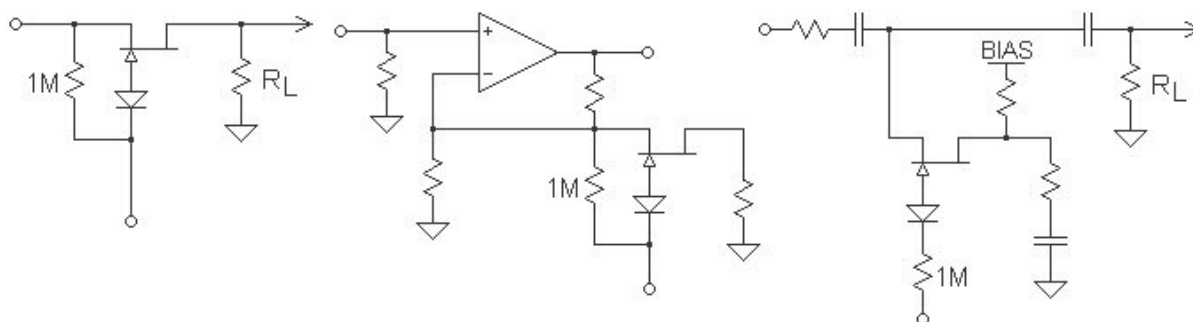


Figure 7.14 FET as analog signal switch. Left: Switching signal path, middle: Switching gain and right: Shunt switching.

In practice, the control voltage (again) sets the limit for the amplitude of the input signal: The maximum amplitude of the signal passing through the switch without distortion can be only about half of the gate control voltage. This means that with control voltages ranging from zero (switch closed) to -5 V (switch open), the maximum signal amplitude is $2.5\text{ V}_{\text{PEAK}}$. *After exceeding this range the switch will clip the signal in closed state or begins to pass signal peaks through in open state.* Likely one will wish to set the ratio of control voltage and signal amplitude so that such overshoots will *never* occur.

The FET switch can effectively be used as a part of the feedback loop if switched gain configurations are needed. In this case it is equally essential not to exceed the voltage limits. Remember that the potential of inverting and non-inverting inputs is equal in non-inverting operational amplifiers so the input signal range for the gain stage is the same as it would be outside the feedback loop.

If controlling FETs using negative and zero control voltages is not a compact solution the way out is to bias the signal potential at the switching point to have a positive voltage offset; this allows to control FETs with positive and zero control voltages. The process involves building a simple voltage divider bias circuit and omitting resistors that connected gate to source. Instead, they are replaced with hookups to “switch bias” potential. Since practically all guitar amplifier circuits use at least 15 V positive rail voltage it is used for switching. These practices ensure that at least $\pm 10\text{ V}_{\text{PEAK}}$ signals can still be switched. Because the signal has DC offset it has to be AC coupled, the demand is also relevant for switching circuitry that shunts parts of the signal path to common (e.g. attenuating voltage divider switches).

Shown are couple examples of switching applications utilizing JFETs: The first circuit provides selection between two inputs (“channel switching”) by controlling the state of the switches simultaneously. When the controlling node is grounded the bottom JFET is “open” and presents a high resistance. At the same time, the control transistor does not conduct. This means that the top JFET has a gate voltage near supply potential, thus the switch is “closed” and the JFET presents a low resistance. If hookup to ground is removed the situation becomes opposite: The transistor begins to conduct dropping the gate voltage of the top JFET to few millivolts thus “opening” it. The gate of the bottom JFET is no longer hooked to zero volts but close to supply potential causing the switch to be “closed”.

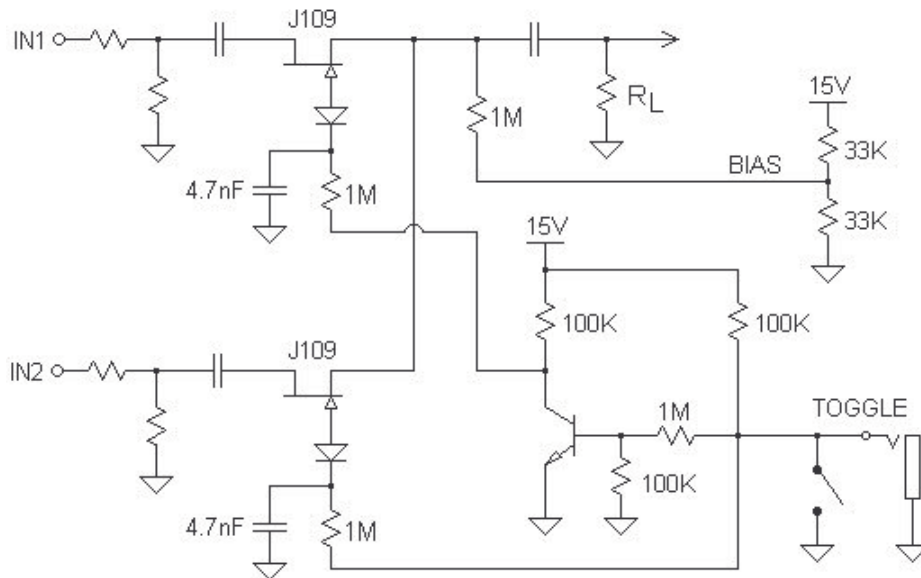


Figure 7.15 JFET Channel switching.

The signal's DC offset at the switches is biased to half of "switch supply" potential using a simple voltage divider and therefore the switching stages are naturally AC coupled. 1-Megohm resistors protect the gates from high current peaks and 4.7 nanofarad capacitors are used to dampen the switching transients.

The other application is very similar – after all, both are *only* switches: This time three switches are used in controlling whether the signal goes through a diode clipping stage or whether the stage is bypassed. In bypass mode, the signal path to diode clipper stage is blocked with the high resistance of the "open" JFET preventing the stage from loading the signal.

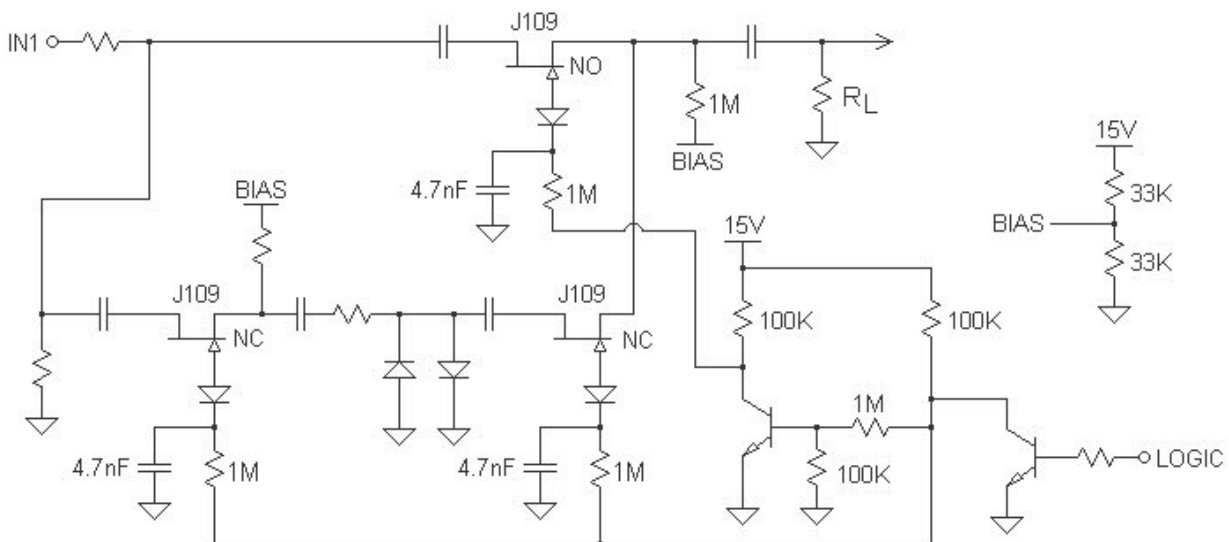


Figure 7.16 Logic-controlled FET switch arrangement for a diode clipper circuit.

The biasing and control arrangements for the switches are very same as in the other circuit, except for the additional transistor that replaces mechanical toggle switches. This is for the purpose of showing an example how the switching can be controlled by a logic circuit operating from 5 V. When logic output is "high" (5 V) the transistor conducts and shunts the collector to ground: The JFETs at bottom are driven "open".

When the logic output is “low” (zero) neither transistor conducts, thus top JFET is “open” and bottom ones are “closed”.

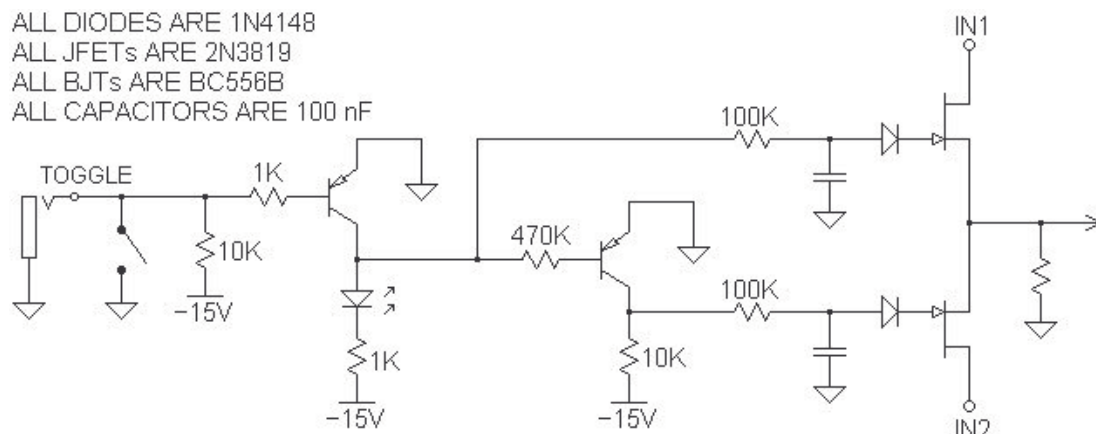


Figure 7.17 Real-life example of a FET-based channel switching circuit (Laney amplifiers).

FET-based switching circuits can even be realized without a bias voltage. In this case the switching devices are actually biased by conveniently referencing them to zero-volt potential. Since the logic is now powered from a negative supply rail the transistor are likewise chosen to be PNP type. Operation of the circuit is similar to that of the previous examples with the addition of an indicator LED.

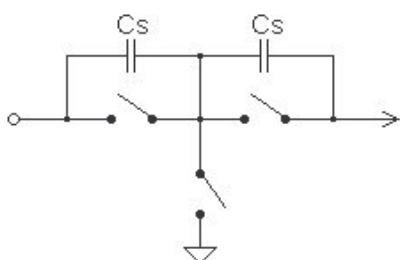


Figure 7.18 T-switch configuration.

Hooking up switches to “T” configuration provides better isolation resistance than using a single switch configuration. It also lowers the RF induced by the stray capacitance that affects between FETs gate and source. The T-switch consists of two switches in series and of one ground-shunting switch connected to their adjoining node.

The greatest drawbacks in using solid-state switching are the substantial amount of complexity the switching circuitry tends to add and the limited input voltage range. Even though the basic FET switch practically consists of only three components (plus the control circuitry) a practical application requires a lot more parts and still has a very limited input voltage range when compared to LDR switches.

7.4.5 Switching OpAmps

One handy way to realize simple switching applications is to use switching operational amplifiers. These typically come in a standard 8-pin DIP package that contains two selectable input differential circuits that share a common output stage. The switching between input differentials A and B is actuated by a logic level signal. Apart from having a switchable output, the features of these ICs are pretty similar to any conventional OpAmp: For example, in the Mitsubishi M5201 the gain of operational amplifier inputs A and B can be set individually. The stages will, in fact, behave exactly like individual OpAmps. The IC also works with both unipolar and dual power supplies.

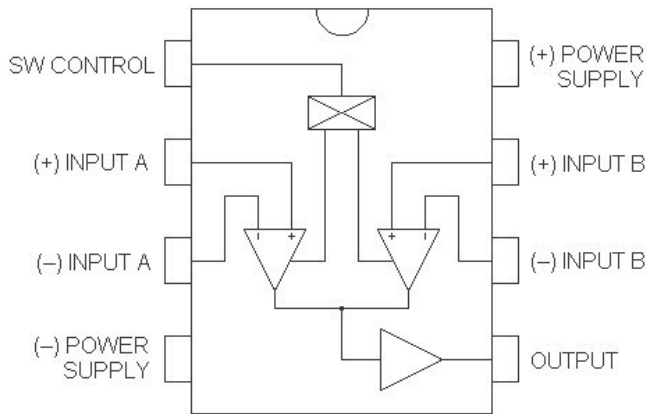


Figure 7.19 Pin configuration for M5201L switching operational amplifier.

7.4.6 CMOS-switches, logic circuits and microcontrollers

CMOS switches are integrated MOSFET switches that are actuated with CMOS logic level signals. Since CMOS-switches are practically solid-state switches packed inside an integrated circuit they are troubled by the same drawbacks; introduction of low pass filter formed by the active device and more importantly *limited input voltage range*. In some cases the latter is less than 5 V (which is an operating voltage of most CMOS logics). The on state resistance of CMOS switches typically varies in proportion to logic supply voltage – resistance being higher at lower supply voltage levels. For instance, a pretty commonly used CMOS switch circuit 4066 has an on state resistance of about 80 ohms with V_{DD} of 15 V but the value increases to 270 ohms with V_{DD} of 5 V. Respectively the input voltage range is 6.75 V and 2.25 V.

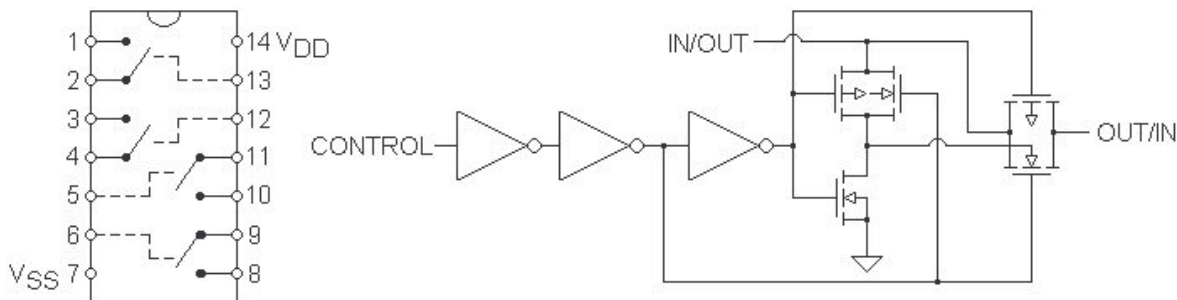


Figure 7.20 Connection and schematic diagrams of Quad Bilateral CMOS Switch CD4066BC.

These days a great selection of analog switches packed inside an easily operated and compact IC circuit is available. Some of these switches can even control moderately high currents and signal amplitudes. Trying to make a small insight to vast amount of applications offered by manufacturers like Texas Instruments, Fairchild, Maxim, Analog Devices or ST Microelectronics would require another chapter and is practically futile since the information would not likely stay up-to-date very long. Most of the switching that guitar amplifiers require can be done without extraordinary complexity or high fidelity so it is up to the designer to ponder whether IC circuitry-based switching is even needed and to acquire the necessary datasheets. If it seems like there will be a requirement for vast amount of switches then using CMOS ones might be the best solution.

Lot of the switching can be aided by logic circuits. These are divided in “families”; two most important being TTL (Transistor-Transistor Logic) based on conventional bipolar transistors and MOSFET based CMOS (Complementary Metal Oxide Semiconductor). Generally digital logic circuits have two basic states: On (“high” or

“1”) and off (“low” or “0”). The real-life voltages are not that exact in order to allow some “hysteresis” to exist: TTL logic family has a low range from 0 to 0.8 V and high range from 2 to 5 V. The operating voltage for TTL chips is 5 V. Circuits of this family are fast but consume a moderate amount of current. CMOS circuits use lower currents and tolerate a wider range of supply voltages. With 5 V supply the acceptable input ranges require a bit tighter fit towards 0 and 5 volts than with TTL. The CMOS chips are extremely sensitive for static electricity.

Logic circuits are conventional way to implement some “smart” functionality to switching - with or without assistance of microprocessors. A “truth table” is used to determine the output of digital circuit with certain values of input. The amount of inputs may vary but some circuit functions are always identical regardless of this; for example, the output of OR function is always high if any of the inputs is high. Output of AND circuit is always low if any of the inputs is low. A circle symbol denotes inversion. Very versatile systems can be – and *are* - constructed from basic logic blocks but this is out of the topic of this book. *Unused logic input should always be terminated with a proper way*; this means either hooking them to “low” or “high” depending on the logic application.

AND			OR			NOT / INVERSION		NAND			NOR			XOR			XOR		
I1	I2	Q	I1	I2	Q	I1	Q	I1	I2	Q	I1	I2	Q	I1	I2	Q	I1	I2	Q
0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
0	1	0	0	1	1	1	0	0	1	1	0	1	0	0	1	1	0	1	0
1	0	0	1	0	1			1	0	1	1	0	0	1	0	1	1	0	0
1	1	1	1	1	1			1	1	0	1	1	0	1	1	0	1	1	1

Figure 7.21 Basic logic functions.

The current available from logic output may or may not be high enough to drive relays or even LEDs and often a transistor interface is required. These interfaces can be made to exhibit a certain behavior as well: For example, it is possible to drive two LEDs simultaneously so that one of them is off with high input and on with low input. Since most switching circuits are based on “dual action” of two switches such circuits are pretty useful. The transistor circuits in the example are used for driving LEDs but it should be obvious that with little modifications they can be configured to drive virtually anything. *Note that if several LEDs are driven in parallel configuration each one of them will require an individual series resistor.* LEDs behave identically to any other solid-state device and the one with least internal resistance will pass most of the current and burn; thus need for current balancing series resistors.

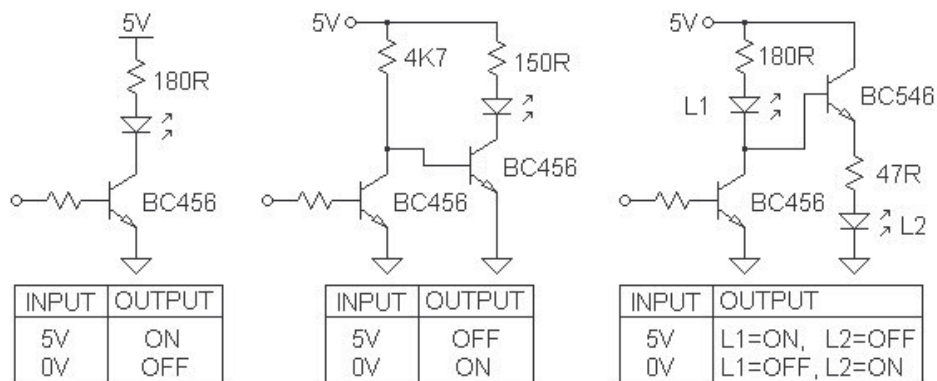


Figure 7.22 Useful configurations for using transistors to switch LEDs.

Transistors can also be configured to respond differently to a logic state: They can either activate a switch with high input signal or activate a switch with low input signal.

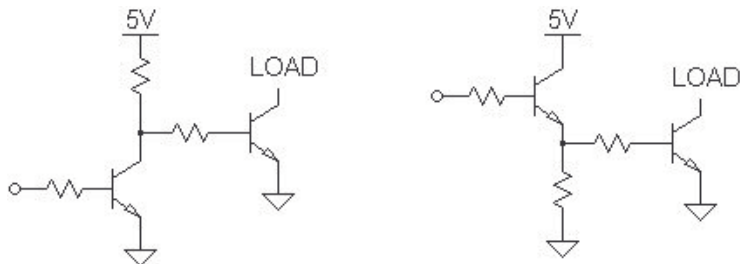


Figure 7.23 Switching arrangements for logic outputs. Left: Pull-down (transistor switch activates with zero input signal) and right: Pull-up (transistor switch activates with 5V input signal).

Other useful circuits include flip-flops; depicted is one of them called SR (Set-Reset) flip-flop: When the circuit detects a high pulse at “Set” input it changes its output into high state and holds it there until a high pulse is detected in the “Reset” input. When reset, the output will stay low until set again. A flip-flop circuit does not require latching switch action to control inputs so momentary switches can be used in controlling it. A simple circuit also provides an inverting output.

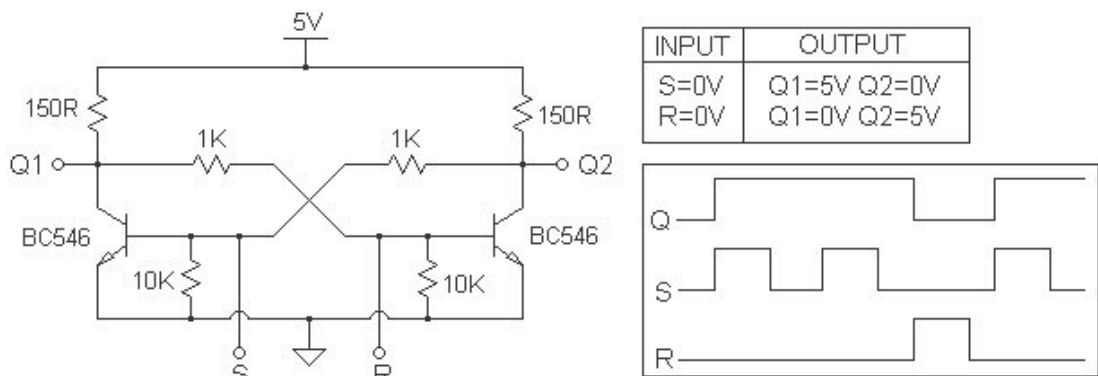


Figure 7.24 A discrete SR flip-flop.

Another useful trick involves controlling the state of two LEDs: When the input is controlled with a digital signal the other LED is always off while the other is on. Using a special bi-color LED, that encapsulates two LEDs with common anode-cathode connection in one casing, allows the same LED to display two different colors in various states.

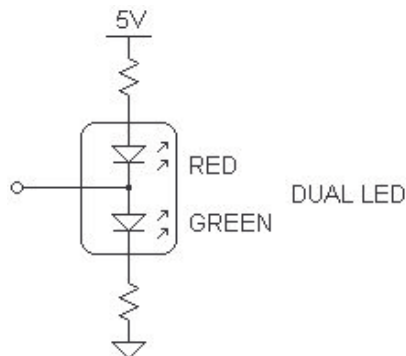


Figure 7.25 Dual LED control.

7.4.7 Remote control circuits

Remote control circuits are the circuits (typically) held inside a foot pedal. A reasonable demand for a remote control circuit is an ability to switch effects or channels on and off and to show an indication of toggled switches – usually with LEDs. Some remote control circuits can do even more; like control the amount of certain effect – technically this functionality is not a very hard to implement either.

A very crude and simple way to wire up a footswitch remote is to use individual nodes for V_{CC} and common plus for all the switch and potentiometer returns. In the example schematic potentiometers control luminosity of LDR's LED element by varying the DC current flow through it. Switching is done either with latching switches or non-latching pulse triggering. An indicator LED is illuminated when a switch is actuated. The switches shown in the schematic example are simplified although an ordinary DPDT switch inside the pedal would work perfectly in the concerned application.

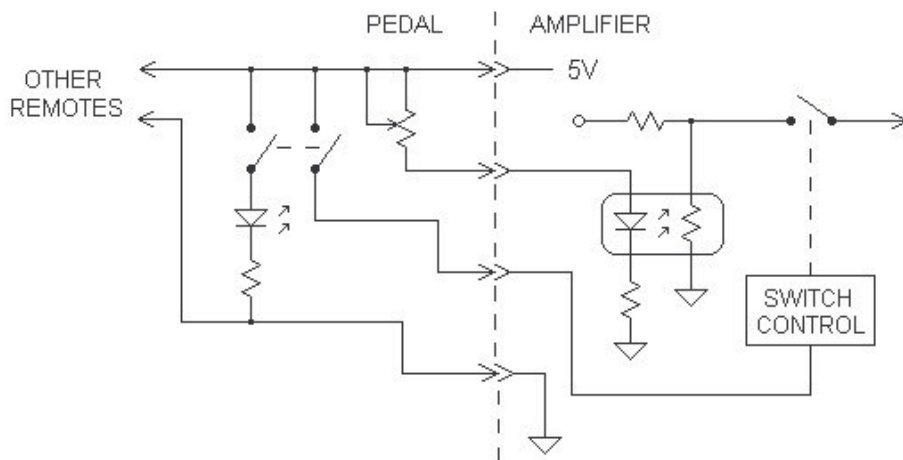


Figure 7.26 A simplified schematic of a toggle/effect volume remote pedal control.

In theory there is nothing wrong with this very simple arrangement and it is even advisable to use it with systems that do not require a vast amount of control. The real drawback comes from the amount of required conductors: V_{CC} and common require two and after that each control (be it a switch or a potentiometer) requires at least one. A standard TRS phone jack has only three conductors, XLR has four (if the case is separated) and an 8-pin DIN connector has eight. It is quite common to use 5-pin DIN connector when only few switches are required. However, this connector type is beginning to be a little “flimsy” and even less-reliable connectors are the RS232-type - although these can provide connection for 25 conductors. If finding a proper connector is not a problem, then at least finding a proper and reliable cord that has multiple wires will be.

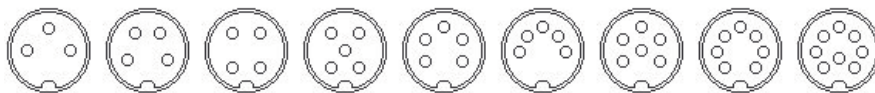


Figure 7.27 DIN Connector types.

There are ways to get around the requirement for multiple conductor wires – like the voltage comparator circuit used in some Fender amplifiers: The footswitch has only two wires, one for a suitable voltage and one for the common, yet it can accommodate four switches. How is this possible? Let's examine the circuit: Instead of ordinary DC voltage, the footswitch is fed with AC. The switches are SPST-type and the switched function is considered being on when the diode and indicator LED are not bypassed.

The series diodes and LEDs create a different voltage drop in each switch configuration. These are (for positive half wave): S1 and S2 on, S1 and S2 off, S1 on and S2 off and S2 on and S1 off. Switches S3 and S4 vary the voltage drop of the negative half wave in similar manner.

The amplifier itself has to be equipped with another, similar circuit to compensate the situations when the remote switching pedal is not used. (This circuit is marked as “compensation circuit” in the schematic). Switching between pedal and the compensating circuit can be done easily by using an ordinary “shunt-tip” TRS jack: Without a plug inside the footswitch jack the tip connector of the jack is shunted to the input of the compensating circuit.

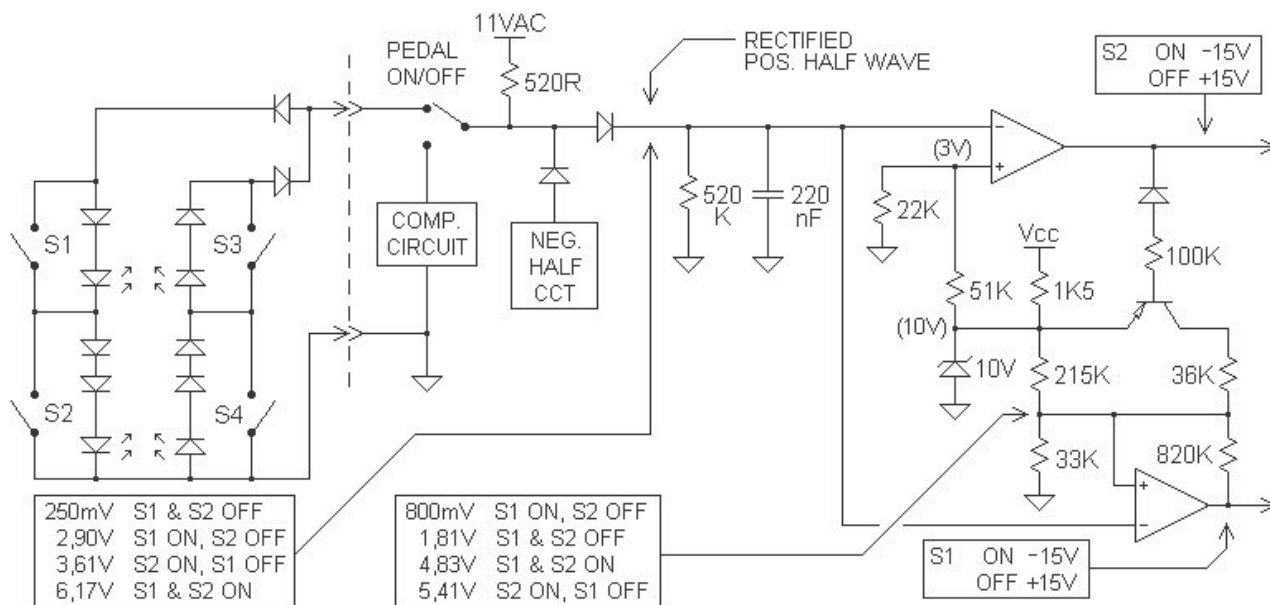


Figure 7.28 Fender-style arrangement for 4 switches and 2 conductor nodes.

The signal from the footswitch is *half wave rectified* and filtered to steady DC with a 220 nF capacitor. The value seems small but works fine since the impedances are high and the currents involved very low. Using only half wave rectifying is essential for the operation of the circuit. The 520 kilo-ohm resistor provides the important common reference. The DC signal is fed to a moderately complex comparator network that determines what switches are on based on the voltage potential over the foot pedal circuit. Thus the output of each of the two comparators corresponds the state of an individual footswitch. The outputs still have to be configured to drive a proper switching arrangement and possibly an onboard indicator LEDs or lamps. Note that *only the positive half wave side of the switching “detector” circuit is shown*. The negative side of the half wave will naturally require a similar one.

Note: Fender has also utilized a 2-button version of this “rectifying”-type remote control. The circuit in question is naturally somewhat simpler. Example of the arrangement can be found, for example, from the circuit diagram of “Deluxe 112” amplifier, which, at the time of writing, is available for download (along with many other amplifier schematics) at Fender’s website under the “Support” section.

Is something like this worth the trouble? Well, if you feel that substituting an arrangement requiring six conductor wires with an arrangement requiring only two - thus having the possibility to use standard guitar cords and TRS jacks between the footswitch and amplifier - is beneficial then yes. The circuit is indeed complex but

you have to remember that today most circuit boards are not even assembled manually. Aside cost, added complexity increases the risk of troubles and faults. The circuitry also requires board space. Many manufacturers still resort to using two separate footswitches if the node amount exceeds five – the number obtainable with a standard 5-pin DIN connector.

7.5 A little bit about MIDI

Musical Instrument Digital Interface, or MIDI, is a standardized communication protocol that was originally devised for use in synthesizers. The protocol simply allows devices to digitally communicate with each other and exchange system data. An internal timer can synchronize the data transfer to follow sequences. The MIDI time code (MTC) is an implementation of the SMPTE time code and cannot carry user-controlled bits.

A guitar amplifier practically receives three kinds of MIDI commands:

1. Program changes. These are a part of standard notation for “channel” messages and tell to switch from one device to another.
2. Controller changes. These are also a part of standard notation for “channel” messages. Controllers allow changing device parameters in real-time.
3. System exclusive, or “SysEx”, messages. These are reserved for all commands that MIDI doesn’t support. SysEx messages differ from standard “channel” messages by being exclusive to certain device or devices.

According to standard, in the receiving end (which in this case is the amplifier), the MIDI message data is sent to UART (Universal Asynchronous Receiver/Transmitter) unit that interprets the messages and acts accordingly. The UART may also send MIDI messages but there is hardly use for that in conventional applications where amplifier simply acts as a “slave”. However copying the input signal to another port (“MIDI thru”), thus allowing device chaining, is a useful and common feature. Note that the standard suggests (or expressed better, requires) that the input is *galvanically isolated* from the outputs and UART to prevent ground loops. The outputs are buffered so that they can drive cables without noticeable losses. The standard for highest cable length is 15 meters.

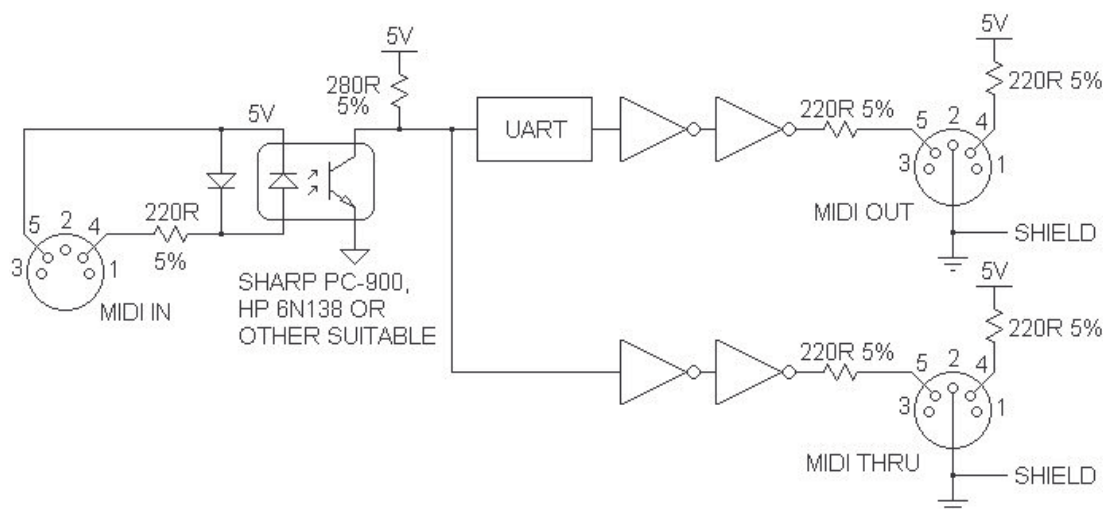


Figure 7.29 Standard MIDI electrical specification diagram.

In practical application, the UART (and often the MIDI out port) are omitted and replaced with a microcontroller that interprets the MIDI commands and controls the state of its outputs according to them (in a programmed manner). In complex systems the microcontroller can, aside controlling solid-state switching, control DSP circuits and alike as well. Even more exciting functions are possible: For example, as the SysEx messaging allows virtually any data to be sent and received it can be used for exchanging “tone patch” – or even emulation model - data with a computer. The possibilities are virtually endless but in practice, out of the topic of this book.

MIDI protocol has a bit rate of 31250 bps (bits per second). The data is sent in strings consisting of a start bit (which is zero), eight data bits and a stop bit (another zero). There is no parity correction. System exclusive commands make up for all commands that MIDI standard doesn't support. A SysEx command is marked with a header string 11110000 (between start and stop bits - as always). This is 0F in hexadecimal format. After this the command continues in syntax: 0iiiiiii, 0ddddddd, ...0ddddddd. The 0iiiiiii stands for 7-bit manufacturer's ID code; if the device recognizes this as its own it continues to interpret the rest of the SysEx message (0ddddddd...etc.), otherwise the message is ignored. The system exclusive command is terminated with string 11110111 (F7 in hexadecimal format).

It is highly recommendable to get acquainted with MIDI standard and the list of manufacturer ID codes to avoid conflicts and incompatibilities. Since the SysEx command standard is quite loose, only containing definition for start and termination string plus manufacturer ID, it is possible to implement a parity check into it. As a matter of fact, quite many manufacturers have done so since it significantly improves the reliability of the concerned application.

Some commercial MIDI foot controllers use specific messages suitable for controlling MIDI instruments as well. These are usually patch or program change messages directed on a specific channel. Browsing through a list of MIDI commands and finding out their binary or hex counterparts is a mandatory task in software design.

Another useful MIDI message is a system real-time message of timing clock. This message is sent 24 times per quarter note when synchronization is required. The message is 11111000 (F8 in hexadecimal format). It is – naturally - located between start and stop bits. Calculating the interval between receiving this message allows calculating the song tempo, which can then be used to set effect parameters like reverb/echo decay/delay times etc.

7.6 Some useful interfaces for microprocessors

Although some guitar players have welcomed digital technology quite sourly it will undoubtedly see more and more use in the future. Needless to say, digital signal processing presents merely just a fraction of this technology. A great field of applications lies in the “interfacing” between the user and the analog circuitry. Nowadays it is common to implement features like push-button volume controls or effects that can be selected and adjusted using a single potentiometer. Such interfaces are usually easier to realize with the aid of digital circuitry. As has been mentioned before, this book will not discuss topics such as programming of microcontrollers, however, even digital circuits require analog parts to function and some of these are worth of discussion.

One of the most basic interfaces is a switch, which can be either a latching or a momentary one. The circuit to detect switching action is rather simple and should not require extensive examination: When the switch is pressed down the microcontroller unit's (MCU) input is grounded and consequently receiving a "low" input. Typical switches are *always* troubled by oscillation that is caused by bouncing of the switch contacts during switching. The phenomenon is known as *chatter*. Rocker switches generally have worse behavior in this respect than push-button types. This oscillation is hardly an issue in analog circuits, but since digital ones deal with only two states the effects become considerable. Using inputs with hysteresis can solve the problem. The hysteresis can also be introduced within the program code, which is the method what designers understandably tend to prefer. Whatever the solution is, the most important thing is to acknowledge that such problem exists.

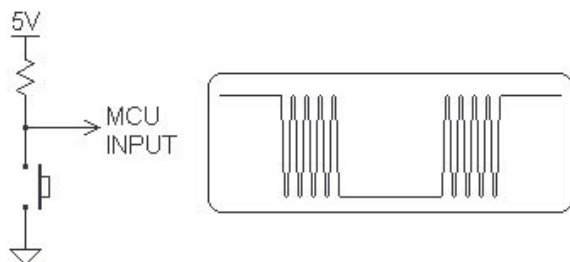


Figure 7.30 An interface for a momentary switch.

A simple RC circuit provides a way to create many useful interfacing circuits. The circuit a) in Figure 7.31, for example, allows measuring of capacitance C_X or resistance R_X : The circuit is based on charging the capacitor C_X through the resistor R_X . Depending on the application the resistor can be a VDR, LDR, NTC, PTC – or just a plain potentiometer. Time to charge the capacitor so that level U_{REF} is exceeded (which changes the state of the comparator) is naturally defined by the RC time constant. Since either R or C is known, the unknown variable can be calculated based on the time constant.

To make this application any practical the calculation naturally has to be done over and over again, which means that a way to reset the RC circuit is required: This is done by opening the MOSFET, which effectively drains the capacitor. The MOSFET is controlled by a square wave oscillator, which is realized within the program code. (State of one of the outputs is switched between low and high at a certain time interval). When the drive signal switches to "low" state - thus MOSFET becomes closed - the program begins to calculate the time it takes to receive a "high" input from the comparator. The longer the time delay, the higher the capacitance or resistance is.

Circuit b) is a variation of the scheme: It employs a constant RC but a variable reference voltage U_X . Naturally the time to charge the RC circuit will be longer the higher the variable reference voltage is – and vice versa.

A practical mean to utilize the circuit is also depicted: Based on the measured capacitance, resistance (or voltage) the microprocessor calculates a proper pulse width modulated output signal. In the example this is filtered, buffered and then used in driving a voltage-controlled amplifier (VCA). Essentially the schematic depicts a volume control. However, we no longer have to worry about scratchy or corroded potentiometers on the signal path – and even more: Various taper functions can be implemented within the program code.

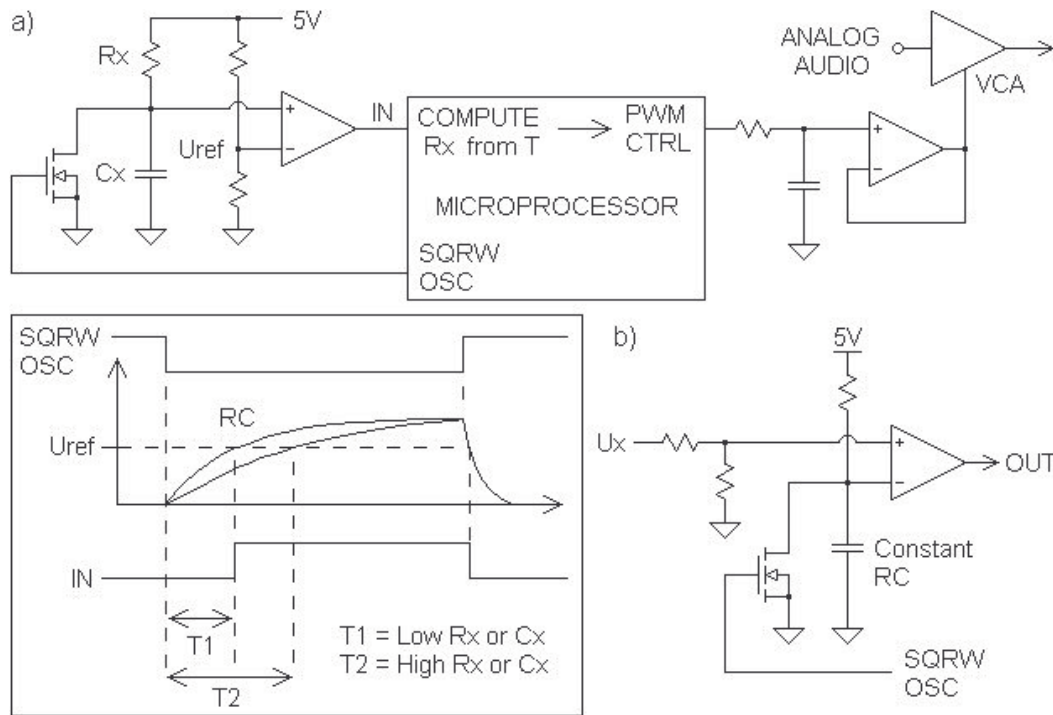


Figure 7.31 RC Interfaces. a) Analog comparator interface for a mathematical measuring of resistance or capacitance. b) Interface for mathematical measuring of voltage.

Here are few other examples how the circuits could be used:

- Tone control: Similar to volume control but the PWM controls a voltage controlled filter or a set of them.
- “Multitasking” special effect control: Resistance of a potentiometer is measured. When resistance exceeds (or falls below) certain values another effect is switched on. The amount of resistance also determines the “level” or “amount” of the concerned effect.
- Single potentiometer that provides a means to select one from many channels and also controls its volume. Basically very similar to previous configuration.
- Volume or tone control for stereo or multi-channel systems that need simultaneous adjustment. Single potentiometer replaces a dual-gang one.
- Loudness compensation: Volume potentiometer’s resistance value is measured and based on this a VCA and a vast set of filters is controlled with PWM.
- Bias circuit: Bias voltage is formed using PWM, measured and corrected if necessary. This arrangement may also employ a NTC or PTC resistor.
- Variation to previous scheme: To improve efficiency the bias is decreased to minimum level when a certain time limit when no audio signal is present becomes exceeded. (A patent protects a similar arrangement).
- A “ring-around” LED display for potentiometer controls. Dual potentiometers are used, one half works normally controlling volume, tone et cetera. The other half is used for measuring the set resistance value and this information controls a LED array that replaces the conventional numeric dial. Variations to them are horizontal and vertical array bars.

A conventional potentiometer can be a cause of many problems: During the time it will get dirty and corroded. It is also vulnerable to DC current, which ultimately ruins the resistive strip consequently causing increase in noise and “crackling” sound when potentiometer is adjusted. It is understandable that we wish to get rid of such a

nuisance. Push-button adjustment is of course one considerable option but the digital circuitry provides a convenient alternative for it as well:

An optical encoder is a device that converts motion into set of digital pulses. By decoding these pulses an absolute or relative measurement of position can be determined. The encoders come in both linear and rotary form. The rotary one is the most common and its physical appearance is very similar to that of an ordinary potentiometer. An internal structure is shown simplified in Figure 7.32, which depicts an incremental encoder: A slotted glass or plastic disc is attached to a rotary shaft mechanism. In a *break-beam* encoder rotating the shaft creates a set of pulses each time the IR beam is interrupted from reaching the sensor. There are also *reflective* encoders, which tend to have inferior performance, though.

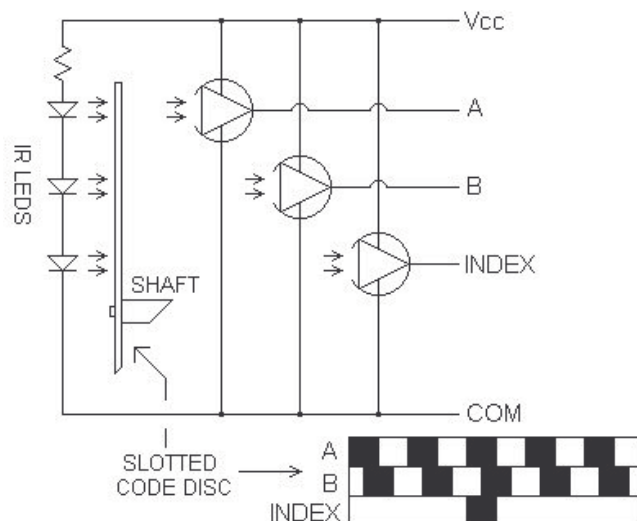


Figure 7.32 Optical rotary incremental encoder.

The rotary encoders come in two forms: An *absolute* encoder is designed to produce a set of bits that distinguishes a certain number of different positions of the shaft. An *incremental* (or relative) encoder, on the other hand, is designed for measuring the relative position of the shaft. It is much simpler in construction, usually having only two or three tracks of slots in a staggered (quadrature phase) pattern. The direction of rotation is determined by judging which channel (A or B) leads the other. Often a third “index” channel, which gives one pulse per revolution, is added since it’s extremely useful in defining the “zero position” or in counting the number of turns (unlike potentiometers, most rotary encoders do not have a limited travel). The figure shows channels A and B using separate tracks but a more common configuration is to use a single track and simply align sensors A and B to a quadrature offset.

7.7 Relays

Relays are switches that can be actuated by magnetizing a coil, which in turn pulls the switching arm. This either makes or breaks the connection of one or several fixed contacts. Generally there are two classes of relays - latching and non-latching. When the magnetizing force (current) of a non-latching relay falls below the *holding level* the arm is returned into its initial position with the aid of a spring or sometimes even gravity. This is the principle of operation most of us has become familiar with. A latching relay works a bit differently: When detecting a current pulse the arm flips its orientation and holds it until another pulse occurs. Thus the hold current of a latching

relay is zero. All relay types are manufactured to operate quickly, mainly to prevent noise in low-voltage applications and arcing in high-voltage applications.

The relay contacts have a certain amount of power they can safely handle. In the worst case, exceeding the power limit can weld the contacts together. As a rule of thumb, direct current stresses the relay contacts more than an alternating current. Usually the higher current the relay can switch, the higher the current consumption of the coil is. Latter is easy to calculate since usually the operating voltage and coil resistances are stated. Note that conventional non-latching relays have a higher “pulling” current than “holding” current – typically the latter is around 20 % to 50 % of the first. The common voltage ratings for the coil circuit are 5 V (logic), 6 V (tube filament systems), 9 V (batteries), 12 V (car batteries), 24 V and 48 V.

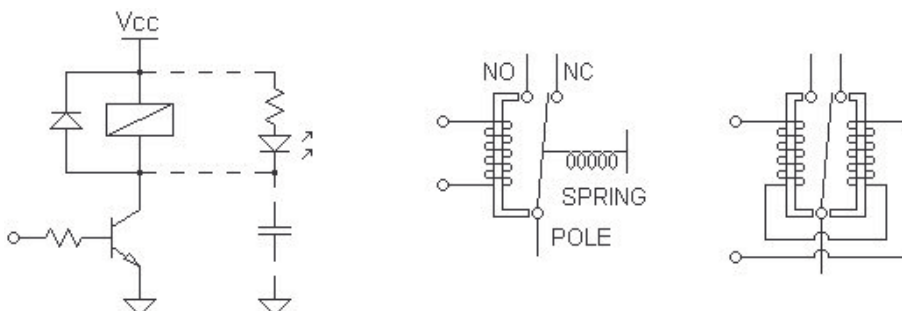


Figure 7.33 The basic collector load relay-driving configuration. Additional components are shown with dashed lines.

Like all mechanical switches, relays are troubled by contact bouncing. Since relay coil is an inductive element the switching will also create back EMF - transient that has an opposite polarity than the supply. This may - and likely will - reverse bias and destroy the switching transistor, thus it has to be damped by a parallel “flywheel” suppression diode that shunts the transient to supply and protects the transistor from excessive forward bias. The idea is exactly the same as in protecting output transistors from back EMF created by a speaker.

Relays should be kept well away from strong magnetic fields that are generated by transformers, chokes or speakers.

8 Earthing and common circuits

8.1 Introductory to grounding and its terminology

Very few aspects of circuit design cause as much debate and confusion as grounding does. The importance of efficient grounding scheme is very important as lack of it may cause unwanted behavior and degradation of performance, for example in the form of hum, oscillation or perhaps distortion of the signal. Many technical books and articles focus on the circuitry and devote less attention into how it should be grounded, which furthermore stirs up more confusion. A circuit diagram is not an effective tool in presenting a suitable component layout or a grounding scheme, which are two things that tremendously interact with each other. This creates an obvious problem: The “ground” is an essential part of the circuit; however, we lack a proper way to depict it in a technically standard and effective way.

Then there’s semantics: *The standard is to use terms “Earth” and “Ground” only when referring to Safety Earth of the mains wiring.* The mains Neutral is *not* the same thing as Ground or Earth either; it is a *load return* node of a three-phase supply system that connects to the point where the phase currents cancel each other and form a zero voltage potential. Nor should the secondary side “ground” be called as ground; it should be called *common*. You need to realize that “common” is a plain voltage reference point for the secondary side and thus far from the original concept of Earth/ground. Also, a “common” node of a real-life circuit never has an ideal connection to the common point; *it merely has a current return path to it* – or in short: A common return. This concept is perhaps the most important thing to adopt. The presented terminology is the standard; unfortunately, hardly anyone (including the author) follows it. Add to the mess the habit of British people referring to common return as “Earth”. While terminology is a totally unimportant issue when it comes to actual implementation of an effective grounding scheme it is worthwhile to understand that using incorrect terms can often be extremely confusing.

The reason why terms “common” and “ground” are often (and mistakenly) used in an interchangeable manner is the fact that we (usually) *wish* the potential of the common reference to be zero volts, which is ideally the potential of Earth/ground as well. To ensure this the common is simply tied to Safety Earth. If common is not tied to ground this way, its voltage potential will float in reference to it, thus we have what is mistakenly called a “floating ground”. Ground/Earth potential basically should not be floating but the floating of common reference is perfectly acceptable. It makes no difference for the circuit’s operation since the voltage potential between common and other secondary nodes will stay constant no matter how much the common point has drifted from zero volts. So, *it is not necessary to tie common and Earth together.* However, since fault conditions in circuits using a floating common can create a situation where the common’s potential may be virtually anything (for example, several hundreds of volts) it is preferable that a (floating) common is isolated from the user.

Furthermore, “grounds” can be divided into several categories: First and foremost we have the real and *only* ground, which is of course the Earth potential – also known as safety ground/earth. According to standards, its schematic symbol is a triangle formed from horizontal lines. The device frame (chassis) has a schematic symbol, which looks

a bit like a rake. Common's schematic symbol is a plain triangle. Sometimes you see common connection marked with a horizontal line dissecting with a vertical wire. This is actually a schematic symbol of a supply rail.

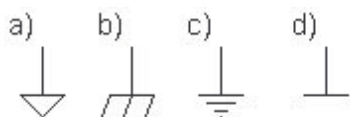


Figure 8.1 Schematic notation standards: a) Common, b) frame (chassis), c) Earth and d) rail node.

Then (as always) we mix-up the standard: Since common can be considered as being one of the supply rails (which it also is), it may share the same schematic symbol (d). Also, you may see the Earth's schematic symbol (c) used as an indication of common connection since nowadays it's clearer to mark an Earth connection with the schematic symbol of a frame (b). Since common returns are (sometimes) divided into signal and supply returns it is quite typical that both triangular symbols are used in depicting each type respectively. The same logic is even more often used when the schematic depicts common current returns for both digital (switching circuitry) and analog (signal and supply) circuits. Understandably there is no standard that would state which one of the symbols should be used in each occasion. If the schematic is professional quality it documents to which category each common return belongs. For example, you might see a number inside the triangle or tags under the symbol. Anyway, schematics are just presentations of real circuits and as was pointed out, a circuit diagram is not an effective tool to present a layout or grounding scheme.

8.2 Common return current routing and grounding topologies

A proper routing of common return currents is one of the most crucial design aspects in analog circuits that deal with high AC currents. It is essential that one starts to treat common as a circuit itself: A common return is – as name implies – merely a route to the *real* common reference point (wherever that is), thus it will have its share of un-ideal properties such as inductance, capacitance *and resistance*. Always imagine wires and circuit board traces as resistors! (Or better yet, inductors). Since stage currents use this route to return to the common point the resistance will (according to Ohm's laws) create a voltage drop - and where is voltage drop there must be voltage as well. Thus our “common point” is not as ideal anymore. Since gain stages usually try to reference their inputs to the common point imagine what happens when a high return current from another stage creates a significant voltage fluctuation at the reference point. A typical example is induced hum from connecting the returns of a high current supply filter capacitor and a high gain, low-voltage amplifier stage together.

Since only one *real* common point can exist in the system it becomes obvious that, in order to avoid any interference, each common return should connect this point individually. Now, add to the rule that the returns should not run parallel with each other (in order to eliminate capacitive effects) and you have created a “grounding” topology known as “star ground”. In theory, this configuration should prevent all anomalies caused by return current induced errors. In practice, an attempt to create an effective star ground to a modestly complex circuit will create an enormous tangle of wiring, which will most likely act as an antenna and catch all kinds of interference. For this sole reason the star ground topology is not very effective, especially in RF applications where *wires practically become inductors*.

Fortunately, quite many of the common return nodes can be safely bundled together without the risk of introducing any notable interference – assuming this is done properly, of course. The bundling is not “art” or “black magic”, as some people tend to think, but follows a certain logic: You simply need to separate stage returns that might interfere with each other. Refer to bulletin below.

- Signal and power supply returns should be completely separated from each other. Thus we have a concept of individual signal and supply (power) “grounds”. Considerably high current pulses often charge the filter capacitors; the effects of this pulsating become more significant the higher the charging currents in the concerned capacitors are. Also, ripple voltage and switching transients are ever-present annoyances in many power supplies. You do not want any of these to interfere with the signal. The common returns of signal and supply should connect each other only at the real common point.
- Transient pulses infest common returns of digital circuitry: Operation of digital circuits is based on switching voltage on and off (usually at a very high frequency). This always creates transient spikes and you do not want them to interfere with the signal nor with the blameless operation of the power supply. The common returns of digital and analog circuitry should connect each other only at the real common point.
- High currents should not return through stages using lower currents.

The last rule brings us to the concept of “ground bus”: Rather than just bundling all suitable common return nodes together and connecting them individually to the common point it is much more tidy to connect them to a “collecting rail” (often referred to as “bus”) that connects to the common point at one end. This can be done without harm if one follows the last rule in the bulletin: *High currents should not return through stages using lower currents*. In practice, the common return carrying the highest currents should connect the bus so that the connection point is physically closest to the common point. Other connections should be made in ascending order so that the bus connection of the return with lowest current is physically furthest from the common point. Signal, supply and digital circuitry returns should (naturally) use separate buses.

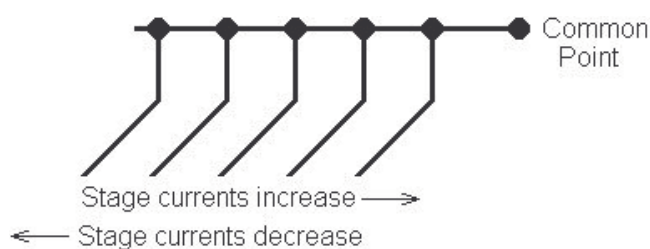


Figure 8.2 Concept of ground bus.

It is fairly safe to bundle the signal returns of an individual gain stage together. In a typical transistor gain stage this would mean bundling together base and emitter resistors plus the emitter bypass capacitor’s return, assuming there is one of course. This bundle is then connected to a proper place in the ground bus. Returns from passive circuits can be bundled together with the returns of the following active stage. Best practice is to use a separate return for all stages that use very high currents, such as the load (speaker) and power transistors. If the circuit utilizes a common referenced feedback loop (like differential input stage) the feedback and input should share the same common reference thus bundling them basically becomes mandatory.

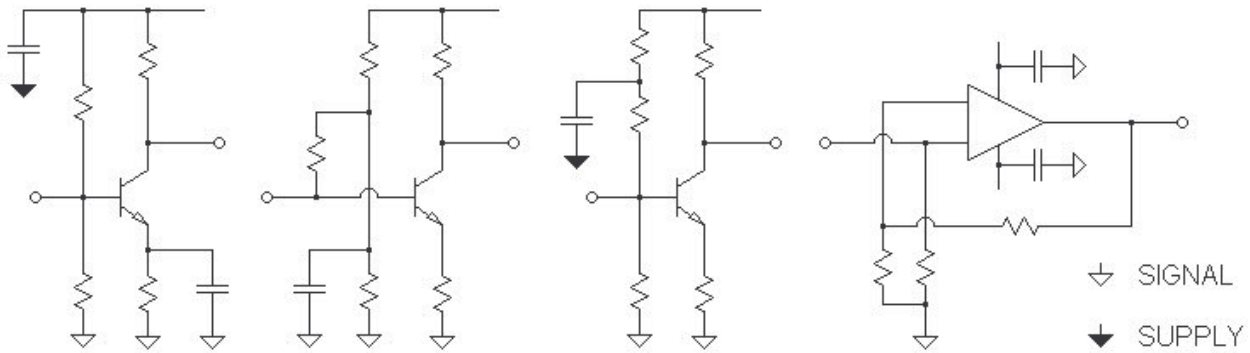


Figure 8.3 Example of typical common return types.

Figure 8.3 depicts which common returns are considered to belong to either signal or supply category. The parallel resistor-capacitor combination in the circuit second from the left is on the “gray zone”: Basically it is a supply filter, however it is located on the signal path since the transistor’s base is referenced to common through it. The best practice is to treat it as a signal’s common return.

Zobel network can share the common return of the speaker since it is a passive stage - and most importantly, because Zobel is a part of the actual load. The common return of the current sense resistor(s) of a mixed mode feedback amplifier should obviously be treated as a high current speaker return. However, any low-current resistors in the mixed mode feedback loop should be referenced to input’s return.

Another option for bundled commons and ground buses is to form a large “ground plane”: The effect of this very old design topology is based on creating a low impedance current return path to common point with least inductance as possible. In theory, the electrons spread evenly and cover the whole plane area. In practice, *the electrons always find the least resistive path to the single common point*. Basically the idea of ground plane concept is to hook as much surface area to common reference as possible – thus logical (and sometimes long) routing of return paths can be avoided. Sometimes the concept works, sometimes it doesn’t. Mostly it has value in RF applications. In multi-layered circuit boards one layer is dedicated to act solely as a large ground plane, in single-sided circuit boards the common area is made to cover as much board space as possible. Sometimes you see circuit boards earthed through the metal bolts that mount them to the chassis. If common and Earth are hooked together the chassis may serve as a “common point” for plenty of components. The ground plane topology can work amazingly well (properly done it is in some occasions much more effective than using ground buses), however, I do not advice using it unless you are willing to experiment with several different layouts and circuit board prototypes. I have also seen a commercial product that raped this topology in the worst possible way: It used a *single thin wire* to connect the “commons” of two circuit boards that were located on separate metal panels (front and rear). This practically violated every rule listed in the bulletin above. Unsurprisingly, the amplifier also had *serious* problems with hum.

The selection of a proper point for common reference is also related to return current issues: Diode rectifiers create switching noise and momentarily large current surges when feeding the reservoir capacitors. The effect of switching distortion is worst in the close proximity of the diodes and the surges affect the whole current loop running from the rectifier to the reservoir capacitors and back. Basically, neither the rectifier nor any part of the wiring between the rectifier and main filter capacitors is suitable for acting as a common reference point. The worst mistake one can make is to

“ground” both rectifier and filter capacitors; this will create an enormous ground loop. So far I have found the following arrangements to work perfectly:

- a) Single supply scenario #1 (Transformer has no common connection): Rectifier connects the main filter capacitor(s) and capacitor(s) connect (recommended) or form the common point. All other common returns connect only to the common point.
- b) Single supply scenario #2 (Center-tapped transformer and two rectifier diodes): Transformer’s common-referenced center tap connects the main filter capacitor(s), which connect (recommended) or form the common point. All other common returns connect only to the common point.
- c) Dual-supply: Center tap connects the main filter capacitor(s), which connect (recommended) or form the common point. All other common returns connect only to the common point.

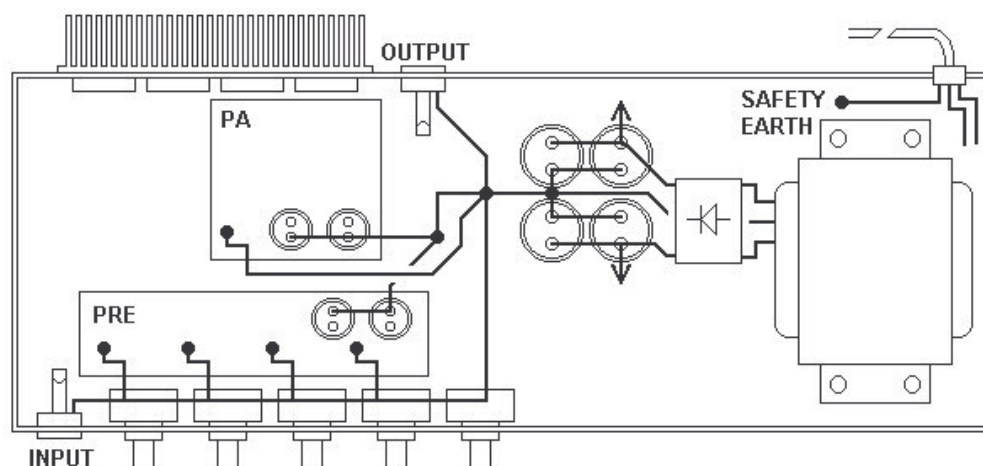


Figure 8.4 An example of common return routing.

Figure 8.4 Depicts an example of typical routing scheme for common returns. Notice how speaker load return is brought to common point separately and how the audio and supply returns of both preamplifier and power amplifier boards are separated. In this example the supply returns are bundled together since this can be done fairly safely – the most important thing is to separate them from the signal returns. One minor detail is very common to some older amplifiers: The signal return bus is hooked to potentiometers. Tying potentiometer cases to common can offer some additional shielding against EMI and RF interference; be sure that the cases do not have a chassis connection though.

The transformer’s center tap does not connect the bridge rectifier. It is connected to the main filter capacitors with a thick wire. The common point is “separated” from the aforementioned “ground” pins of the main filter capacitors to isolate the “charge-current turbulence” caused by transient-like charging of the capacitors. The connection between the capacitors (“reservoir ground”) and the common point is made with heavy-gauge wire and kept as short as possible. Also notice the wiring in the main filter capacitors: This scheme divides the wiring resistance more evenly. Even better “wiring” solution would be to use a large ground plane that hooks up all the common returns of the main filter capacitors together. In this case the “ground” plane could fairly safely be used as the common point without fear of charge currents affecting the circuit’s operation. A thick aluminum or copper plate is perfect for the job but in order to use one the filter capacitors must be equipped with screw terminals.

8.3 Safety ground considerations

The “Earth” wire *may* be tied to the neutral at the household’s breaker box (panel board), other alternative is that the connection is done locally inside the mains socket. Note that in the latter case, breaking the solidness of the neutral return from the device to the breaker box will lift the device’s neutral to mains potential. The safety earth was designed to protect from faults such as these; during a fault it acts as an alternative conductor that replaces the broken neutral. If the mains socket is only locally “earthed” the chassis will be in the mains potential as well (that is, assuming the device has “earthed” chassis of course).

These days, ground fault current interrupters (GFCI) have become very popular. They are also known as Residual Current Devices (RCD), Residual Current Circuit Breakers (RCCB) or Appliance Leakage Current Interrupters (ALCI). A device such as this basically consists of a differential transformer that monitors the current flow in Live and Neutral wires. During normal operation the currents have an opposite phase and they cancel each other out in the sensing element of the protection device. If for some reason the current in one of the conductors grows higher than the current in the other, the protection device will sense a difference and release the relay hooked to live supply wire (L). Usually the fault current required to trip the device (within few milliseconds) is set to be about 2 – 6 mA. However, if *only equipment* is being protected the trip current may be as high as 10 – 300 mA. Using locally earthed mains sockets (Neutral is connected to Earth inside the mains receptacle) in conjunction with correctly earthed mains supply line will trip these protectors since the currents between L and N get unbalanced (because current now returns also through the safety earth wire).

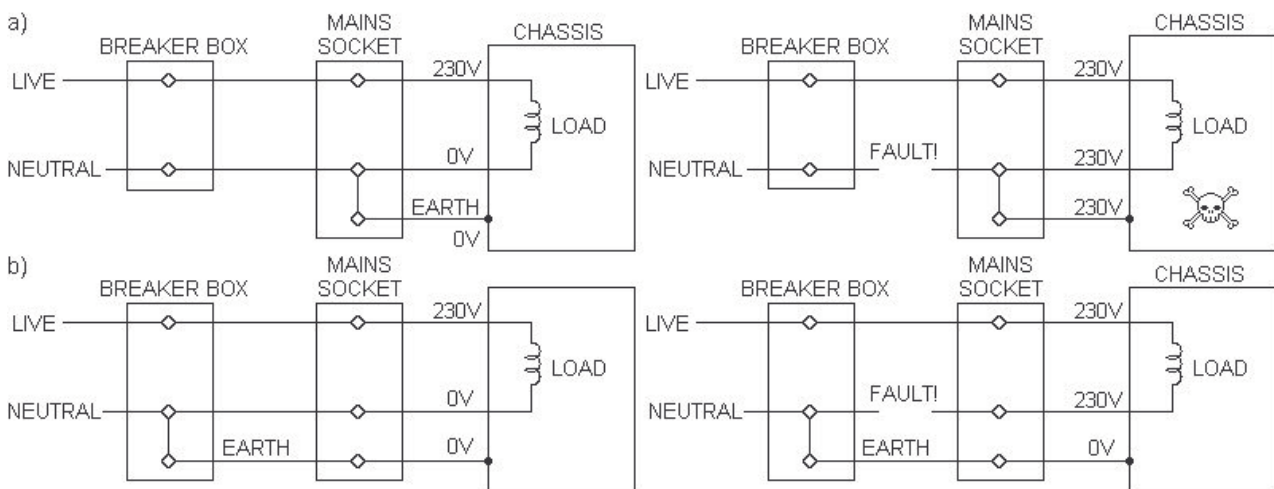


Figure 8.5 a) "Zeroed" and b) safety grounded systems in normal (left) and fault condition (right) where neutral wire has broken.

One possible fault mechanism is that the live potential shorts to the chassis. Although direct shorts are quite rare a very typical example of a milder “short” is leakage from an aged “grounding capacitor”. These are also known as “death caps”. The grounding capacitors were used in older equipment to create a high-impedance connection between chassis and neutral, basically by hooking the chassis to ground potential at AC level. The capacitors were typically used in conjunction with a “ground switch” that allowed to select the proper orientation of the capacitor in relation to AC line. This was done in order to acquire the minimum amount of hum. Nowadays the safety earth wire replaces the function of the grounding capacitor.

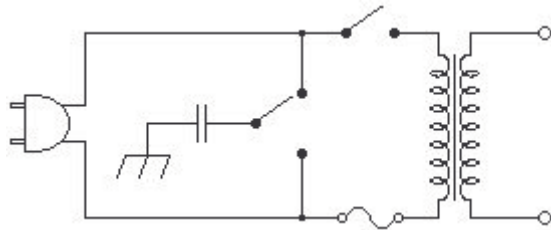


Figure 8.6 These days, potentially lethal circuits like this should no longer be needed.

Unless the chassis has an earthing connection it will, during faults, just float in the mains potential (or any other fault potential) and therefore cause all kinds of risks. If the chassis is earthed and the earth wiring is intact all the way down to the breaker box (or other proper grounding point) an electrical short to chassis will just trip the fuse and cut off the supply.

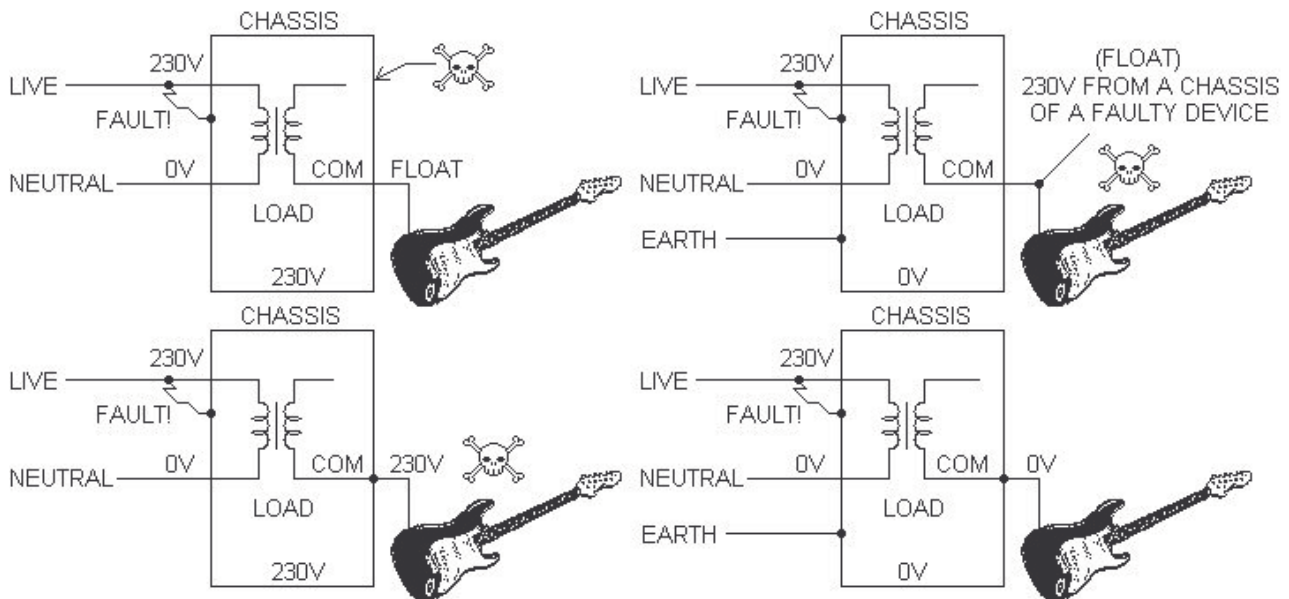


Figure 8.7 Fault scenario where mains supply shorts to chassis in two- and three-prong corded equipment.

Another thing causing troubles is leaving the secondary side common potential to float: Being isolated from the mains and chassis the floating common is virtually bullet-proof against faults where the chassis gets shorted into a high-voltage potential. However, it is virtually helpless against faults where the common potential itself hooks up to a high voltage potential - for example, through another faulty piece of equipment. The common point does not care whether it has a potential of zero volts or a potential of 400 volts as long as the voltage potential differences within the load current loop are not interfered. The musician may notice a little tingling sensation but nothing will happen until he touches a device or an object that has a significant voltage potential difference (for example, the grounded case of a microphone). At that point he becomes a conductor for high current.

The above scenario is basically what will happen as well in a case when the earth wiring has no continuity to neutral point; in this case common, along with the guitar and chassis, will float in an undefined potential. To prevent accidents all inputs should be (safety) earthed and the earth wire should be intact.

Sometimes (actually very often) the Earth wiring is totally absent from the mains system, in which case all earthing configurations inside the device are naturally rendered useless.

8.4 Earthing and ground loops

Ground loops are created by voltage potential difference between two ground points. Connecting two or more safety grounded devices together from more than a single point will give birth to a ground loop that usually manifests itself by having the ground current amplified. Since the ground potentials are different a current forms a loop between the devices using the common (or sometimes a shield) of a signal cable as a conductor. This interferes with the return current of the common and thus gets amplified.

Since the impedance of the ground circuit is usually very low even small voltage differences can cause a very high current flow. One-volt difference in a wire that has resistance of 0.1 ohms will create a current flow of 10 amperes! The best ways to minimize ground potential differences are to power all devices from the same mains line – preferably from the same receptacle - and to avoid using long mains wires.

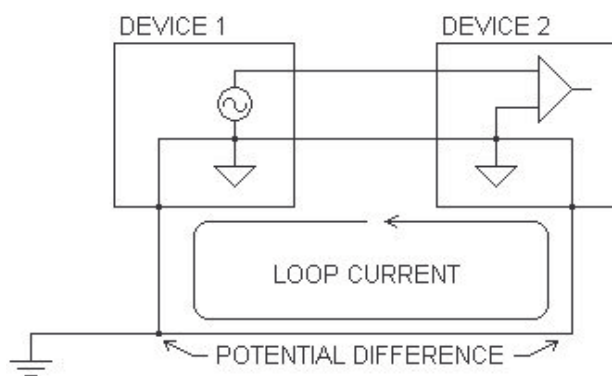


Figure 8.8 The formation of ground loop.

Breaking the connection between common and ground can break the ground loop. Since this procedure effectively decreases the safety of the equipment it should *never* be used as the solution to solve a ground loop problem. There are more effective ways to battle ground loops than risking one's safety: Careful configuration to power each of the devices in a way that minimizes voltage differences between grounds, galvanically isolating the interconnects or introducing current limiting resistance between common and ground to name a few.

8.5 “Loop breaker” circuits

Loop breaker circuits are used in preventing ground loop currents from reaching the common reference. A loop breaker circuit usually consists of a moderately high-power and low-value resistor (e.g. 5 W, 3.3 – 10 ohms) in parallel with a low-value capacitor. For a ground loop current even a low resistance is considerable, thus the amplitude of hum passing from earth circuit to common side will be notably attenuated. The capacitance (usually about 100 nanofarads) is a bypass for RF. Note that the circuit is worthless unless interconnects are isolated from the chassis. Transformer body, and other similar parts of the circuitry, must have a direct connection to earth and should not be connected through a loop breaker circuitry.

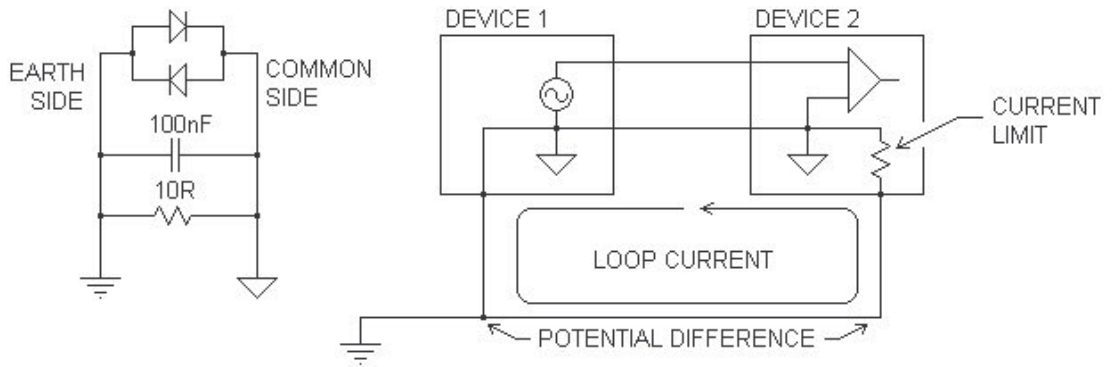


Figure 8.9 Typical loop breaker configuration.

Since there is a fair chance that the conventional loop breaker circuitry burns open in a certain fault condition (i.e. high current flow from common to earth) it is usual to bypass the circuit with anti-parallel high-current, high-power diodes that will conduct the fault currents. A high-power bridge rectifier can also substitute them. Diodes usually fail to short circuit so the protection likely stays effective even if the diodes are damaged. Note the words *usually* and *likely*! Fact remains that a loop breaker circuit compromises the protection offered by direct safety earthing and in some countries circuits like this are therefore illegal.

9 Practical component choices and design aspects

9.1 Resistors

When choosing resistors one's first thought should not be what "tone" the resistor is claimed to have. Resistors are chosen (and sold) mostly according to their ratings of tolerance (precision) and power dissipation. Other considerable matter is their inductance: A resistor formed from a coil of resistive wire (a wirewound resistor) is basically an *inductor* with a large series resistance and small inductance. Resistors with a resistive material inside the case are less inductive. The resistive material is usually either carbon composition, nickel chromium (metal film), mixture of metal and glass (metal glaze), or mixture metal and insulating oxide (metal oxide).

Carbon composite material is extremely noisy and *intolerant against DC current*. It actually becomes "corroded" and noisy after a small period of time under the influence of DC. The uneven surface of carbon composite material in cheap resistors causes excessive heating and gradual changes in resistance. Also, resistance of carbon composite resistors may drift a great deal when heated (tolerance may change 3 %), the resistance decreases according to applied voltage (2.5 %) and humidity conditions may increase the resistance as much as 10 %! The decrease of resistance in high-voltage applications can create a compressive effect (in some circuits gain decreases) and this may count for all "mojo" that carbon composite resistors possess. The required voltage to acquire this effect is several hundreds of volts so typical solid-state designs will not produce it. For a solid-state design the use of carbon composite resistors is nothing but trouble; it is not a good choice if you wish to ensure a consistent operation of the circuit.

Wirewound resistors have a poor tolerance as well but mostly it stays on an acceptable level. They are also highly stable, very quiet and boast remarkable power dissipation ratings. Problem is their inductivity, although it sometimes is a useful character in ballast resistors of the output stage. For values below 10 kilo-ohms, even non-inductively wound resistors usually have an inductance of about 20 μH .

Metal film resistors are quiet and stable. They have very good tolerance ratings but can have only medium power dissipation ratios.

Special purpose parts are "fusing resistors" that *deliberately open circuit* when failing due to exceeding the rated power dissipation. These resistors are often non-flammable as well, which means they can toast without the risk of starting a fire. At a time the non-flammable resistors were marked with a body color of either white or blue. Today, you cannot trust this rule anymore. In some applications, fusing resistors are a better solution than ordinary fuses. Zero-ohm resistors (marked with a black stripe or a dot in SMD parts) are jumper wires in a package that can be handled by the industrial assembling machines.

All resistors age and drift in value, which is mainly a concern in precision equipment. An interesting side note concerning this is that metal film resistors have a burn-in period of one week at rated power - during which the resistance value drifts the most [35].

9.2 Capacitors

When selecting capacitors the most important factors (aside capacitance, voltage rating and dielectric strength of course) are physical size, polarity (or lack of it) endurance and equivalent series resistance (ESR). Although the dielectric insulator between capacitor's plates is no near perfect the effects this might introduce (phase shift, non-linearity and so on) generally affect at frequencies that are way above the hearing threshold. Most of the literature concerning non-linearity of different types of capacitors is aimed for people working with HF and UHF designs – or for people designing equipment that requires extreme precision. Capacitor non-linearity is undeniable, though. The most prominent effect of capacitor not being a perfect component is the resonant “tank” circuit formed because of the self-inductance that all capacitors possess at some extent. This may sometimes cause “ringing” oscillation at a certain frequency way above the hearing threshold – usually way above the amplifier's bandwidth as well. Practically the self-inductance (or series inductance) is lowest in surface mount, ceramic or tantalum capacitors.

Some capacitors, for example ceramic types, drift in capacitance according to applied frequency. This is the reason why they are not preferred in filter circuits – note that filters in this concept actually means filters *of precision equipment*. Little drift doesn't mean a lot in audio applications and while some people eagerly like point out the non-linearity of certain capacitor types they tend to forget that resistors are hardly perfect in value either and therefore equally affect the RC filter circuit. It is somewhat funny that many people make a big number about picking the most “linear” capacitors and then parallel them with carbon comp resistors that have a 10 % tolerance. In the midst of all this wealth of information the common sense of how to use it seems to be lost sometimes. Some capacitor types exhibit non-linearity that appears as a change of capacitance in relation to capacitor's temperature. This is known as temperature coefficient and it will not necessarily change in a linear proportion to temperature. Some capacitors types exhibit a non-linearity that appears as a change of capacitance in relation to applied voltage. This effect is known as voltage coefficient and is most prevalent in ceramic and high-voltage capacitors.

To put this in perspective, the effects of reasonably high ESR (common in aged capacitors) are much more perceptible than the effects of various coefficients. Think of an ideal capacitor in series with an ideal resistor: At high frequencies the resistor's resistance mainly forms the total impedance of the circuit. Thus we can conclude that capacitors with high ESR will likely attenuate higher frequencies a bit more. The internal resistance also means a power loss, which scales linearly with the ESR.

But let's ignore the non-linearities for a while because capacitors have way more important parameters from the design point of view: These are *power dissipation factor and ripple current rating*. Dissipation factor is closely tied to ESR because higher series resistance causes more power loss. This in turn causes the capacitor to heat up more and increased heat causes drift in capacitance and decrease in lifespan. Ripple voltage is the AC voltage variation seen by the capacitor, thus *ripple current means the current going to and coming from the capacitor*. Ripple current rating is the maximum continuous ripple the capacitor can have if its life expectancy is to be achieved. The rating quoted in datasheets is for operation in ambient temperature of 85° C. The value increases if operating temperature is decreased. The parameter especially concerns the main filter capacitors and (perhaps even more importantly) the speaker load coupling capacitor, as these are the components seeing highest AC

current variations. Ripple current should not be exceeded, as it means that the component will work under serious stress and its lifespan is therefore decreased. Large capacitors generally have a higher ripple current rating than the smaller ones, but since larger capacitors usually have a smaller surface area per unit capacitance as well using two small capacitors to replace a single large one is often beneficial. This way ESR will be lower, the ripple current rating higher and surface area increased. It is often a cheaper combination as well.

Another important parameter of capacitors is their voltage rating. This is sometimes quoted in a “confusing” form of “WV”, which simply is an abbreviation from “Working Voltage”. All capacitors, especially electrolytic ones, have a maximum voltage that their plate insulation can safely tolerate. This maximum is also the quoted voltage rating so the capacitor should really be operated at about 75 % of it in order to leave some tolerance for transients. In most cases the capacitor will not fail if you operate it at the quoted voltage (or few volts above it) but in this case you are really pushing your luck. Too low operating voltages in turn cause the electrolyte to gradually “dry out”. This deterioration consequently decreases the capacitor’s working voltage rating to about the same level where the capacitor is currently operated. Thus it is recommended not to operate the capacitor in a voltage that is less than 33 % of WV. Overall, the 75 % is a very convenient rule of thumb.

All capacitors, to some extent, exhibit behavior known as “soakage” or dielectric absorption, which is described as memory effect of a capacitor that is reluctant to discharge completely: After a short term discharge the capacitor will slowly regain some of its charge back. The amount of recovery is proportional to *temperature*, initial voltage and soakage characteristics of the concerned capacitor. The effect is also known as “dielectric hysteresis”. In conventional audio applications the soakage will likely never cause problems or perceivable changes in the circuit’s operation. However, it may easily become detrimental in circuits where voltage shifts are moderately high and happen fast; such circuits are, for example, switches and comparators as well as sample and hold circuits.

If you examine how capacitors are marketed (excluding ridiculous advertisements trying to sell you overpriced high-end “audio-grade” capacitors) you will see that the strongest advertising arguments are *HF* performance, tolerance, size and durability. For example, so called “computer grade” capacitors are very durable and withstand high ambient temperatures. This is a significant attribute since temperature is the most important factor affecting the lifetime of a capacitor. As a rough rule of thumb the lifetime is reduced to half per each 10° Celsius rise above ambient 25° C. Some capacitors are also more tolerant against humidity than others.

Most likely you’ll want to concentrate on improving the lifetime and reliability of your guitar amplifier. After all, guitar amplifiers are in most cases considered as instruments that should have a long lifetime. This rule should dictate your choice over capacitors (and any other components as well). To ease the future servicing, you should try to limit the amount of capacitors that are known to be unreliable or to have a limited lifespan. In short, minimize the amount of electrolytic capacitors and *never* use tantalum capacitors unless you really have to!

Capacitors can be separated into three categories according to their dielectric material, these are ceramic, plastic film and electrolytic. The following table presents a short introductory to each.

Table 9.1 Capacitor attributes

Dielectric material:		Properties
Air gap		Typically “Stacked plate” variable capacitors. Used in RF applications. Low dielectric loss and non-polar. Impractical for any capacitance values above few hundred picofarads.
Printed circuit board and wiring		Capacitance between traces, wires and/or copper layers - either deliberate or unintentional. Low capacitance. Mainly used in HF applications. Stable.
Ceramic	Generic	Excellent HF performance. Nonlinear, affected by temperature. Sometimes exhibits piezoelectric effects, thus being microphonic.
	Monolithic	Low impedance, good for bypassing, Excellent HF performance, nonlinear.
	Silvered mica	Excellent linearity, good for RF circuits, very small dielectric losses. Expensive.
	“Stacked”	Manufactured by connecting several monolithic capacitors in parallel.
Plastic film	Polystyrene	Excellent general-purpose capacitor. Linear and stable, highly sensitive to heat and solvents. Highly inductive.
	Polyester (Mylar)	Popular and suitable for all audio applications and bypassing. Stable and reliable. Low voltage ratings.
	Polycarbonate and Polysulfone	Best suited for critical timing and coupling applications. High voltage ratings – especially with Polysulfone, poorer tempco than with Polystyrene, larger than Mylar.
	Polypropylene	Excellent for passive crossovers. Moderate voltage ratings. Superior to all other plastic types. Highly inductive.
Electrolytic	Aluminum	The standard material. Polarized and nonlinear. Extremely high capacitances, vulnerable to heat. Drifts with age. Leaky. Normal lifespan approximately 10 to 30 years.
	Bipolar	Two polarized in series. Good for passive crossovers. Mostly very nonlinear, unreliable and leaky. Good ones are expensive. Avoid using.
	Oil/paper	Polarized and very nonlinear. Obsolete. Rugged and self-healing. Will not fail as short. Drift highly with age and sensitive to moisture. Low capacitance.
	Tantalum	Polarized. Small and high capacitance. Very nonlinear! Very unreliable!

Always use reliable capacitors in the power supply and in the stages that use high currents. At least the (large) reservoir capacitors should be computer grade. All capacitors that you have to place near warm parts should be computer grade as well. The extra money put to these components is an investment that ensures a long and rugged life before any service is needed. Never use recycled electrolytic capacitors when fixing or building an amplifier, especially if those capacitors have been on a shelf for few years! If you see capacitor recycling as an effective way to save money

then do think again: An exploding capacitor that fails with a short circuit and takes down the whole power amplifier stage is not my idea of a way to cut expenses.

In most cases, large coupling capacitances (typical for vintage transistor designs) are unnecessary. Using high coupling values is a “HiFi design rule” that is based on minimizing phase shifts of the RC filter circuit and the non-linearity of electrolytic capacitors. The latter attribute becomes significant at the “sloping” frequencies of the RC hi-pass filter, thus extending the corner frequency way below our hearing range can improve the filter’s linearity. However, as opposed to HiFi amplifiers, guitar amplifiers benefit from not having an extended low frequency response and consequently with a proper design one should be able to effectively couple all small signal stages using less than 1 μF of capacitance. This pretty much eradicates the need for electrolytic capacitors in the preamplifier - except in its power supply of course. If board space is not an issue then for slightly better audio quality one can substitute ceramic capacitors with polyester or polypropylene ones. Stay away from expensive “high-end” audio capacitors: Their effect to the tone in comparison with cheaper capacitors is questionable.

Some high-end “gurus” suggest stripping off the plastic cover of electrolytic capacitors. It is doubtful whether this has any benefits sonically; at least I have never seen any valid proof of it. When considering that 1 dB change in sound pressure (which corresponds to about the slightest difference in loudness that an average human can perceive) correlates to a voltage level change that is about one tenth in magnitude, the truth of such claim should be quite easy to measure. Yet, I have never have seen any published results of such measurements. I personally regard the suggestion as utter nonsense. It is said that moving your head one inch will make more difference to the sound than some obscure tweaks suggested by so-called gurus. Removing the plastic cover *will make hell out of any future service of the amplifier* since it gets complicated to determine the correct capacitor orientation and even more complicated to determine the correct capacitance without measurements. It *may* slightly improve heat dissipation, though. If you want to do this then at least do yourself (and others) a favor and write down the capacitance and voltage rating to the side of the capacitor with a permanent marker pen.

9.3 Potentiometers and trimmers

Potentiometers wear out fast in guitar amplifiers since they are in a constant use. Also, in order to minimize signal losses the conventional circuits deliberately skimp on using protective series resistors. Since the wiper gradually wears its way through the resistive strip the best choice is to select high-grade potentiometers that have a long lifespan. It is not recommendable to use PCB-mounted parts since very likely the potentiometers will require changing on few year intervals. Naturally, high quality parts will outlast the cheaper ones and offer less noisy performance. An ultimate solution would be to replace all potentiometers with digitally controlled optoelectronic devices. One may of course argue about the lifespan of the complex circuitry this solution involves.

Potentiometer actuators come in few different forms, rotary actuators being the most common. Multi-turn rotary potentiometers have a high accuracy and therefore they tend to be expensive. This type is more common amongst trimmer potentiometers. Linear actuator potentiometers, also known as faders, are much more expensive than ordinary rotary potentiometers. High quality “slide potentiometers” use conductive

plastic strip and are very expensive. Their price grows higher along with the accuracy, which in turn correlates to actuator's length. Also, potentiometers can be "dual"; these come in two forms: Common "dual gang" potentiometer has two individual resistive strips and wipers plus a single, shared actuator. The rare "dual concentric" form has also separate actuators, which are attached to the common shaft mechanism. This type is also known as "stacked" potentiometer. Dual gang potentiometers tend to be more expensive than the ordinary ones but they do serve their purpose in quite many applications.

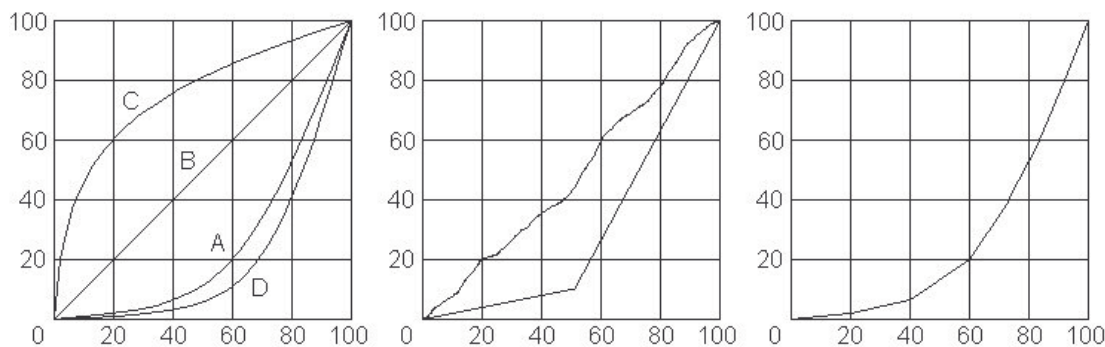


Figure 9.1 Left: An example of various tapers or "laws". Middle: A typical audio taper potentiometer is constructed from two linear resistive elements. There are no perfectly linear resistive strips. Right: Expensive, high precision potentiometers are constructed from several resistive elements.

Potentiometers have various tapers or resistance "laws" that define how the resistance advances in the potentiometer's resistive strip. The two most typical taper functions are linear and logarithmic ("audio"). More uncommon resistance laws are, for example, the "antilog", "reverse audio" and the "accentuated audio" tapers. These are mostly out of production and can be sometimes found from older equipment. Manufacturers are naturally willing to construct potentiometers with almost any unusual taper function - assuming the price is not an issue. Typically it is. The taper's type is commonly expressed with a one-letter code but unfortunately the coding system has too many conflicts that render the whole idea nearly *useless*. A table listing typical variations can be found from the end of this book. As situation is what it is, your best bet is to always measure the resistance from one end to the wiper when wiper is located at 0, 25, 50, 75 and 100 % of the turn. This will indicate what taper type is used.

As the attached figure shows, a typical logarithmic potentiometer is not really that logarithmic as it's usually just constructed from two linear resistive elements (it is cheaper that way). Higher quality potentiometers may have more resistive elements than two. Even linear tapers are not perfectly linear and therefore manufacturers quote a tolerance figure (in percentage), which is measured from at least 70 % of the effective rotation range. Most of this is pretty useless information since guitar amplifiers are not precision devices; yet knowing these shortcomings may explain some seemingly strange quirks in performance.

The selection of trimmer potentiometers is also based on their durability. Since trimmers are mainly used only for calibrating the circuit at the end of the assembly process they must be components that are highly durable and will not drift. If you read the paragraph concerning resistors you might remember that metal film resistors drift the most during a burn-in period that lasts *one week* at the rated power. This is not a very good property for a trimmer potentiometer that usually requires precision.

Carbon composite trimmers are not an advisable selection either since they get noisy very quickly.

Table 9.2 Potentiometer properties

Material	Properties
Carbon composition	Most common material. Cheap or average quality. Average lifespan. Noisy and vulnerable to DC. Mediocre power dissipation.
Metal film (Cermet)	Used in high quality trimmers. Low noise and high stability. Very limited lifespan. Medium power dissipation.
Conductive plastic	High quality. Long lifespan, accurate. Low noise. Mediocre power dissipation.
Wirewound	Almost indefinite lifespan. Low noise. Inaccurate, no smooth transition. High-power dissipation.

Trimmers come in many shapes and you should select one, which offers the best protection for the resistive element. While “vintage”-style trimmers with exposed strips may look cool stay away from them: The exposed strip gets dirty and corrodes easily. They are surely easier to clean than the sealed types but if you buy a high quality trimmer that is housed in a proper casing you will *never* have to clean it anyway. Usually old trimmer are also carbon composite-types. In most cases one should prefer multi-turn trimmers since they offer the highest accuracy – a property needed for example in the delicate adjustment of bias current. Lock the trimmers with nail polish (or similar substrate) after adjustment.

9.4 Switches and jacks

Switches and jacks should all be as sturdy as possible. The most important parameters for selecting a switch should be voltage and current ratings. Even most of the “mini” switches are rated for 250 VAC and 3 amperes but their real ability to withstand large currents is more than questionable. Choose a proper “mains switch”-type switch for concerned task and preferably stay away from models that have a metal lever or other metal parts that can be touched. These are even illegal in some countries. I recommend that switches that are intended to be board-mounted should not be used; they have a habit of breaking their solder joints even under medium stress. Exception to rule are switches with multiple contacts: Hand-wiring them is untidy and space-consuming – and usually the several solder joints spread on a large board area divide the tension making the switches sturdier as well. If you, for some reason, have to use board-mounted switches make sure they are at least mounted to chassis or other adequate support to handle the tension.

I encourage no to use board-mounted jacks for the same reasons that applied to switches. If you have to, make sure they are mounted in a way that is adequate to withstand the tension. The tension should never, ever, be directed to the circuit board alone. Make sure of this and your board-mounted part will not likely break any of its solder joints and go intermittent. However, after the part has worn out the service procedure will be painstaking and hazardous to the circuit board. *Board-mounted jacks and switches without any external support tend to destroy their solder joints within a year.* Furthermore, push them hard enough and they may even destroy the circuit board.

Some switches and jacks have gold-plated connectors. The benefit of having a golden surface on a connector is that the contact area will not corrode and thus introduce noise or intermittency. Gold-plated connectors, however, are useful only in extremely rare occasions. In continuous use the golden surface will wear out very quickly. From top of my head I can come up with only one application in guitar amplifiers that would benefit from using gold-plated connectors: The RCA jacks for reverb send and return of an “inboard” spring reverb tank. They usually stay “hard-wired” throughout the whole lifetime of the amplifier.

An important detail is that there is no standard for the diameter of phone jacks. Although the jack’s internal diameter itself is 6.3 mm (1/4”) the diameter measured from the jack’s threaded surface may vary even few millimeters. There is at least one considerable difference between metal-body “open case”-styled jacks and plastic-body jacks - whether they are isolated or not: Plastic bodied jacks usually have a larger diameter so they are not interchangeable with metal jacks. And please, do not refer to *isolated* jacks as “plastic jacks”: Not all jacks made out of plastic are isolated.

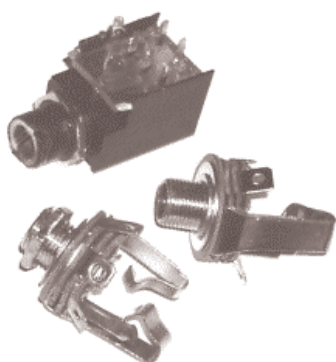


Figure 9.2 Jack types. Top: Insulated plastic jack with internal DPDT switch and bottom: un-insulated metal-body “Switchcraft” jacks with internal tip-shunt switch.

Since the topic is phone jacks I feel it’s reasonably fair to suggest that you *never, ever* use the “miniature” 3.5 mm (1/8”) phone jacks for anything: These jacks and plugs were only designed to be cheap and compact, not reliable or professional quality. About everything is wrong in these components: First of all, the plug is too tiny to support the weight of the cord, which leads to *intermittent jacks*. A solder joint of a sturdy 6.3 mm jack is much more reliable even though the 6.3 mm plug and a thick cord weight a lot more. Secondly, the plug is too tiny to support the weight of the cord, which leads to *intermittent wiring*: This is a very sneaky problem and not something you would suspect at first. However, in my experience it is the most common problem with the 3.5 mm

interconnects. Broken wiring is also annoying to fix because most 3.5 mm plugs have a “molded” construction and therefore cannot be disassembled. Because the plugs themselves (although they’re small) are quite sturdy, it is not surprising that interconnects tend to break just right outside the molded part where the plug usually has only cheesy rubber enforcement. Anyway, only a fool would even dream about fixing these jacks when they break (which happens in about one week): This stuff is generally made as cheaply as possible and not even intended to be reliable or repairable.

9.5 Semiconductors

Choosing semiconductors is one of the hardest aspects of the design process. With high- and semi-power devices one has to consider phase angle and maximum ratings for voltage, current and power dissipation. Other considerable parameters are the physical size and shape of the semiconductor. H_{FE} figures are important as well, especially when it comes to substituting or matching devices. In practice, selecting complementary pairs does most matching one will ever need with NPN and PNP pairs. Many small signal transistors tend to operate similarly in conventional amplifier circuits so concentrating on clipping characteristics is a bit vain, especially since

varying the gain and frequency response of the circuit can provide way more drastic alteration of clipping tone. FETs are trickier since they exhibit a lot of variation between devices. Using proper bias configuration is a professional way to provide more tolerance against device variation but the devices used in the most critical applications should really be tested and matched. The matching parameter is either the $I_{DS}-V_{GS}$ curve or transconductance Y_{DS} . Other important parameters are gate cutoff voltage $V_{GS(OFF)}$ and maximum drain-source current I_{DSS} . Sometimes it is best to use just the least noisy device. Always base your selection of semiconductors to a fact how they interact and perform with the concerned circuit – not to a fact that they were used in some famous design.

When you finally have all the above figured out, expect to find out that the devices you planned to use are temporarily out of stock or gone obsolete. This is why I usually design around standard “workhorse” devices such as TIP142/147 or MJE3055/2955 for power, BD139/140 for semi-power and BC546/BC556 for small signal stages. Even these are sometimes out of stock. Using esoteric components will surely result into troubles when sourcing for parts. I advice that you make sure that your design can cope with few substitutes that are not perfectly identical to original devices.

When selecting operational amplifiers one should not care about their claimed “tone”. Typical operational amplifiers have THD figures that are better than 0.001 % and if you can hear any difference in tone at those levels (note that difference in output signals must be distortion) it is either product of your imagination or the operational amplifier is unstable and modulating the output signal in some obscure way. When clipping, operational amplifiers have no “tone” either; the huge amount of distortion would mask it anyway. Operational amplifiers operate best as linear devices and it is not good engineering to deliberately drive them to clipping. The only way an operational amplifier can practically have “tone” is when you create an impedance mismatch (and the consequent tone loss) by using BJT input operational amplifiers in an application where you really should use FET input.

Do not be too kind on subjective results of “OpAmp picking” either: In order to make fair comparisons between various models each of them should be optimized specifically. Some OpAmps have very unique requirements for this and these are hardly met in conventional OpAmp picking experiments. Also notice that while the IC socket is a tremendous ease for repair jobs it is ironically also the cause for many troubles: Some IC sockets are extremely poor quality and develop intermittency or high resistance after a very short period of time. Many high-speed OpAmps are also very sensitive to stray capacitance the sockets introduce. Generally, decision whether to use sockets or not is a compromise between easy serviceability and reliability. Sockets should be used *only* with expensive ICs, ICs likely getting damaged (meaning mainly those interfering with the outside world, i.e. the input and output stages) and naturally for ICs that need to be upgraded now and then.

The most important parameters in the selection of operational amplifiers (aside input device type) should be: Bandwidth, CMRR, PSRR, THD, input common voltage range (the range for input signal), offset current and voltage, input and output resistances, open loop voltage gain, output voltage swing (greatest voltage swing V_{PP}), slew rate (speed), supply current without load, power dissipation and *noise*. The last parameter is very important when selecting operational amplifiers to critical places of the circuit: These are typically the stages where the signal amplitude is very low (thus SNR is a major concern) or where the gain is fairly high. Such stages really do benefit

from using high quality operational amplifiers. Correspondingly, you don't have to use them in every imaginable stage. In some applications you need to focus on current delivering capacity as well. Examples of such applications are headphone amplifier circuits and reverb tank drivers. Note that many OpAmps can't properly handle load impedances that are lower than 600 ohms.

9.6 Power supply parts

The selection of power supply parts is essentially based on three parameters: Voltage, current and power dissipation rating. Diodes and bridge rectifiers should always be generously overrated since currents during capacitor charging, surges and inrush can easily rise ten times higher than what the nominal load current draw is. A reasonable rule is to at least double the expected current rating. Having that said, 35A rectifier bridges cost next to nothing and I suggest using one in every design that has a higher output power than 50 watts. If you are pinching pennies in designs that go to large production runs you likely know already what rectifiers to use.

Mains line filter capacitors and such should be generously overrated as well in order to make them handle voltage surges. They also have to be *specifically rated for AC applications*. Such capacitors are generally categorized into various grades such as X1, X2, Y1 and Y2. Class X capacitors are designed for applications where failure cannot lead to an electric shock, such as across-the-line use, while class Y capacitors are for applications where failure could lead to an electric shock - like in line-to-safety ground use. These capacitors are "self clearing", which means that whenever a flaw in the dielectric develops a short, the "avalanche" of electrons generates enough heat to vaporize the thin metallic electrode. The vaporization removes the short and the capacitor continues to work normally. The property is also known as *self-healing*. Class X1 is mainly intended for industrial use and tolerates supply spikes up to 4 kV. X2 is for commercial, residential and light industrial use and tolerates only 2.5 kV. Class Y1 capacitors are impulse tested to 8 kV and Y2 to 5 kV. X3, Y3 and Y4 are generally for lower voltages and not accepted by safety standards.

The mains transformer has a tremendous effect on amplifier's output power so care should be taken in selecting it. Novices usually look only the voltage rating of the secondary but this is just a single parameter, albeit an important one. However, *a certain secondary voltage figure does not make transformers equal!* For example, 12 V transformers can exist in sizes of an average pencil sharpener or an average brick. Another equally important parameter is the VA rating and for some reason this parameter seems to be harder to understand. VA rating is a function of maximum current and voltage obtainable from the transformer. Consequently this means that a transformer with a maximum VA rating of 10 VA cannot supply no more than about 10 watts of power – even if the secondary voltage would be 100 volts or so. The rating simply follows Ohm's law: Power is voltage times current, thus $V \cdot A$.

A power amplifier circuit that could push out hundred watts of power cannot give out hundred watts if it's hooked to a supply with a 10 VA transformer; it will likely give out only about 10 watts. Peak power rating, as well as the average power for a period of, say, 200 ms might be higher but this is just a way of saying that the amplifier is not as efficient (read: not as *powerful*) as an amplifier that can provide the same average power *continuously*. In a nutshell: The circuit is as good as its weakest link. For example, *the popular "60W mod" to Vox Cambridge Reverb does not magically convert a 30-watt amplifier to a 60-watt amplifier*: Although the LM3875 power

OpAmp chip (in comparison to the initial TDA2050) could deliver all the required power, the wimpy power supply just can't keep up – and that's the end of story.

Since class-AB amplifiers are quite inefficient the VA rating should be considerably higher than the *intended* output power of the amplifier. In practice, the VA rating and supply capacitance are factors that really determine the amplifier's *continuous average* output power, which is the only parameter worth of stating. No one, except the marketing department, cares if the amplifier can supply 1000 watts for a period of one microsecond.

Basically, the minimum VA rating would equal the output power of the amplifier – yet this rating fails to consider *the efficiency of the amplifier*, which is typically about 60 %. Therefore a good rule of thumb is to use a transformer with a VA rating that is *at least two times higher* than the intended output power. Going higher than three times the output power will likely provide no benefit but trouble in the form of increased inrush current. After ensuring the transformer can provide the power that the circuit needs it gets more beneficial to concentrate on increasing the supply capacitance. If the supply voltages will not sag (or if the amount of sagging is decreased) it means the amplifier will run “cleaner”, give out higher power and exhibit less distortion. Increased capacitance also reduces the capacitor ripple current, which furthermore improves reliability. Also, a high power transformer working at 50 % of its capacity will run significantly cooler than a lower power one working at 95 %. For sure the first one will also have a longer lifespan.

Sufficient VA ratings are typically very large. As an example, let's examine an amplifier that provides an average power of 60-watts into a 4-ohm load: The peak power of such amplifier is 120 watts to four ohms. A voltage potential required to create this power to a four-ohm load is about 22 volts and can be calculated from the square root of P (120 W) times R (4 ohms). We must add the saturation voltage drops of the output devices and the voltage drops over the rectifier diodes to this result. It is reasonable to assume that they are at least five volts in total so the end result is 27 volts. This is multiplied with two (giving 54) to acquire the required rail-to-rail voltage. This is a peak value and must be converted to RMS by multiplying with 0.707 (result is 38.2). We now know the required supply voltage but in order to calculate the required VA rating we also need to know the peak load current, which is calculated from square root of P divided by R . The result is about 5.5 amperes. When 5.5 A is multiplied with 38.2 V we get 210.1 watts. We still need to take the poor efficiency (60 %) into account and therefore multiply the result with 1.4, thus the final result is 294 watts! The rating is about five times greater than the output power! In practice, this is the “overkill” VA rating and with the assist of huge supply capacitance - and with some toleration against a greater amount of sag – one can live with lower VA rating values.

9.7 Fusing

Fuses have three parameters to consider: Current rating, voltage rating and time delay. The rule of thumb is to use a fuse that has a current rating that is 125 percent of the usual circuit current. For protection against direct shorts only, the rating should be 150 percent. However, this will prove to be tricky in case of mains fuses that are subjected to high inrush current. The voltage rating of a fuse *does not* indicate the voltage the fuse can withstand while carrying current. Instead it indicates the ability to extinguish (and prevent) arcing after the fusing element has melted. The rating states the

maximum RMS value. Fuses should have the same, or preferably higher, voltage rating than what the voltage of the circuit is.

Fuses have three time delay ratings (US): Type F, Type T and FF. Type F is the normal fuse type and has alternative names such as type Q (“quick”), type QB (“quick blow”), standard and fast-acting. A type F fuse will carry 135 % of its rated current for about an hour but will blow within one second at 200 %.

Type T is a delayed-type fuse. The alternative names are time lag, anti-surge, slo-blo, or the European marking “gL”. A special and extremely rare type is the “semi-delay” that is sort of a mix between type F and type T.

Type FF is rapid or ultra-fast type and responds up to 100 times faster than type F or type T fuses. They are often called semiconductor fuses because they have a *slight* chance of saving semiconductors during faults.

The rule of thumb for sizing mains fuse is to start from the value equal to peak current derived from VA. Advance by increasing the value until the inrush current or power down stops blowing the fuse. Read 5.5 for more details. Due to high inrush currents mains fuses are usually slow blow type. Rail fuses are usually fast acting but secondary fuses located before the rectifier or the filter caps should likely be delayed type due to current turbulence.

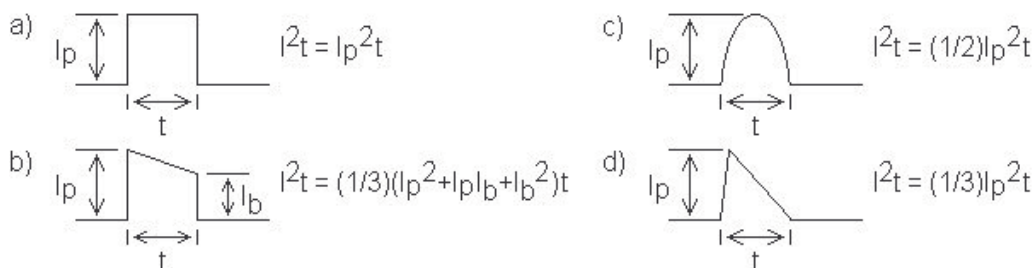


Figure 9.3 Various pulse shapes and formulas for calculating I^2t .

Because fuses often cannot be selected solely on the basis of normal operating current and ambient temperature (mainly because there is no “normal” current draw in the first place) it is necessary to figure out a parameter known as nominal melting I^2t rating. This is a measure of energy that is required to melt the fuse and expressed as “Ampere Squared Seconds” [70]. Usually procedure of calculating the variable involves estimating what wave shapes (“pulses”) will occur during transient current draw. I^2t is equal to $I_{RMS}^2 t$, thus I_{RMS} is $I_2 t$ squared.

9.8 Printed circuit board design

It has been said that a PCB that is designed to work with RF works well on lower frequencies as well but not vice versa. Although guitar amplifier circuits do not operate at high frequencies many design principles learned from applications that do can enhance performance and reduce the risk of oscillation. Even though the design practices might seem unnecessary it is best to adopt and try to follow them since it never does any harm. The following is not a comprehensive guide but a presentation of few conventional tricks that are beneficial even in low frequency applications. Note that most of the methods and techniques presented here apply equally to other circuit board construction methods – be the board veroboard, stripboard, perfboard or a turret board.

Note: If you haven't read the chapter 8, which discusses earthing and common circuits, read it now. Theory presented in this section is more beneficial once you have assimilated the aforementioned information.

First, something about units: Component spacing relies on imperial units, one *thou* equals 1/1000 of an inch and one *mil* equals one thou. Therefore, mil should not be confused with millimeter. A typical practice is to use inches when thou value is higher than 100. 100 thous is a general reference measure and vast amount of component spacing is based on multiples of this unit. Thous should be used when the measures are related to grids, tracks, pads and spacing, (or generally anything that is related to layout), while metric units can be used with board dimensions, hole diameters and generally with anything related to mechanical design.

The units become especially important when a third-party company is used in manufacturing the boards as these companies will quote the lowest track width/spacing resolution they are capable of manufacturing (in thous). Going below the quoted limits just assures the manufacturing process will create a bunch of boards with shorted or open traces. Usually companies also list a typical resolution value – and charge more money if thinner tracing or smaller spacing is required. Real word track/space resolution is between 10/10 and 8/8. IPC standard recommends 4 thous as the lower limit. Naturally soldering (manually) is easier the wider the traces, and the spaces in between them, are. Typical spacing between pins of an IC in DIP package is 100 thous (2.54 mm). This is also the pin spacing of an ordinary transistor. With radial “small” capacitors the center-to-center spacing between pins (also known as “pitch”) ranges from 200 thous (smallest ceramics) to 600 thous (large electrolytic capacitors or polyester film capacitors). With radial capacitors and very large electrolytic ones the pin spacing can be just about anything. Typical spacing for a 1/2 W resistor is 500 thous.

Some manufacturers offer a standard “drill size” (hole diameter) and charge extra if other sizes are used. Some manufacturers offer a standard set of drill sizes but charge more for any non-standard size. Typical component lead diameters are 0.63, 0.7 mm and 0.8 mm (for ordinary components, 0.7 mm being most common), 1 mm (high power resistors and transistors) and 1.5 mm (trimmers, potentiometers etc.). The area surrounding the hole is called a pad. Pad size (or “annular ring width”) is both manufacturability and soldering issue: The thinner the pad is the easier it breaks while drilling the holes. A standard requirement for pad size is a 5-mil annulus. However, 10-mil annulus is often a reasonable requirement for solderability.

Now that the basics have been cleared out we can delve into more “artistic” aspects. You will likely find out that there are a great deal of rules and practices. All of them are difficult to follow at once and the final PCB design is usually more or less a compromise between numerous things. Judging when a certain rule should be neglected in order to create the most efficient solution is always advisable and often necessary as well. Sometimes the decision you make is poor and the layout degrades the performance; just learn to live with it and accept the fact that you may have to create several prototype boards before the design works optimally. Divide your circuit to board modules so that you may experiment with alternative layouts of different circuit parts without having to repopulate a large circuit board again and again. Remember to minimize the wire length between the modules.

First few general rules:

- Equal components coming from different manufacturers tend to have different packages and physical sizes. Make sure the PCB is compatible with this variation. Especially capacitors have a lot of variation in their pitch.
- Consider all traces as wires that have resistance and determine what wire gauge would be suitable for a specific current. Make sure the trace width matches this wire gauge.
- Begin the layout design process by placing components that *must* have a specific location. Route signal traces (that including their common returns) first using smallest trace lengths and least amount of jumper wires as possible, then route power supply traces (again common returns included) - after that, route everything else.
- Leave “slots” for compensating components and all other optional and experimental “extras” that you think might become useful. Do this at least in few of the first prototype board designs.
- Lay out components in logical order and try to keep the loop lengths as small as possible. (You do remember that current flows in loops?) Realize that a loop is essentially an inductor.
- Make sure that you know the return current paths and do not let them interfere with other circuitry.
- Oscillations are usually caused by long feedback loops, which introduce considerable amounts of inductance and capacitance.
- Keep heat sensitive parts well away from heating parts.
- Keep input well away from output in high current circuitry.
- If you are handy with soldering iron use surface mount components in places where trace lengths have higher importance than power ratings.
- If you use a PCB design software *turn off the auto-router!* You are smart; an auto-router is not.
- Double and triple check the board layout. Sleep over the night and then double and triple check the board layout again. Hand out the board design to someone who checks it out as well. You are blind to your own mistakes!

Electricity doesn't care whether the board looks neat and “professional”. If you have to use messy-looking rounded traces or other anomalies considered “ugly” in order to achieve the best layout and the most optimal routing then do it! In most cases this means that you have to etch the board by yourself so start from that initial assumption. There are some PCB design guides that tell you to keep the board neat and pretty but you have to remember that aesthetics are not the first priority in PCB design. A neat circuit board can be easier to troubleshoot but usually it is not as good as performer. Keep in mind that no one will care about the looks of the circuit board once it's inside the chassis. Although retaining symmetry of the design in stereo or other multi-channel applications might seem like an aesthetic decision it is not: PCB layout is a component and some components should be kept as identical as possible.

Producing double-sided boards can be complex but they have some benefits: The top layer can be dedicated to sole grounding purpose (“ground plane”), which increases the mechanical strength of the board, lowers the impedance of all ground connections, suppresses noise by adding distributed capacitance to every net in the circuit and acts as a shield [34]. However, in order to do this the plane has to be as solid as possible. Ground planes have drawbacks as well: You have very little control where return currents actually flow (read 8.2) and disruptions that break the solidness of the ground

plane (there usually will be plenty) cause return currents to flow using very strange paths, which may also turn the PCB into an antenna. Improper use of ground plane arrangement can even contribute to addition of noise, oscillation and hum. Double-sided boards are also useful in constructing complex circuits that still require thick traces. Such circuits are, for example, power amplifiers. In many cases cramming all the required traces together in a single-sided board is a sheer impossibility.

In most cases there is no such thing as too much copper: From electrical point of view, using wider traces usually has nothing but benefits. However, wide traces require more space and often do not fit well with modern, very small components. Since a proper clearance between tracks is also important it isn't always the best practice to use thick traces – use them only in cases when they are required. Since wide tracks are not always an option a far better practice is to use boards with a *thick layer of copper*, as the only downside of this approach is merely the larger consumption of acid when boards are etched. A thick layer of copper is also a lot sturdier than a wide but thin trace. Usually, all unneeded board space can be left un-etched, which also saves acid and increases the board's mechanical strength. However, don't leave floating copper "islands", either remove them completely or connect them to supply or common potential. Be sure that extra copper is not interfering with the way you want loops to run.

All parallel traces capacitively couple each other to some extent. *Note that you have parallel traces on both horizontal and vertical levels.* The higher the frequency of the signal in the trace, the more effective the coupling becomes. Do not overlap digital and analog traces. Most digital circuits can be so fast that even a low capacitance between the traces is enough to couple the digital signal. Even ordinary signal lines should not run in parallel with each other – unless the capacitive effect between them is a desired feature. *In conventional analog designs the trace capacitance can and will become a very significant factor if impedances are high:* A signal trace leading to high impedance resistor that runs in parallel with another signal trace will likely trigger oscillation. As the traces essentially form a capacitor, increasing the distance of traces can minimize the capacitive effect.



Figure 9.4 Overlapping of digital and analog traces creates a coupling capacitor.

A pretty important point to notice is how easily inconsiderately routed traces can create an antenna: The formation of loop antenna (in single-sided boards) is easy to detect and avoid but much more harder is to visualize the formation of a *slot* antenna. The important rule of thumb is that *return currents always try to follow their source path, as this is the least inductive route for them.* Thus in a double-sided PCB, where the top layer is a dedicated ground plane, the return current finds its way directly above the source trace. If this path is interrupted, a slot antenna is created. The cutout for the IC also forms another loop antenna. The only way to prevent these incidents is to use a solid ground plane. In board c) the cutout still exists but it does not interrupt the return current flow.

Since return currents try to follow their source conductors it is beneficial *not to* separate the traces of supply and its common return on opposite sides of the board but

to keep them in a close proximity to each other. This lowers the loop inductance, improves capacitive decoupling of the supply and allows an effective bypassing with shorter trace lengths. When designing supply routing, design the common return layout first – then make the supply layout follow the scheme. Do not compromise the routing of the common returns for the benefit of the supply: Power system noise can always be decoupled with filters – common return noise can be not. Apply the same principle to routing of signal traces as well. Remember: If low DC currents are involved you can always “jumper” the supply to a convenient point on the board with ordinary “flying leads”. The only reason why this method is not very popular nowadays is that it increases the amount of manual labor in the manufacturing process.

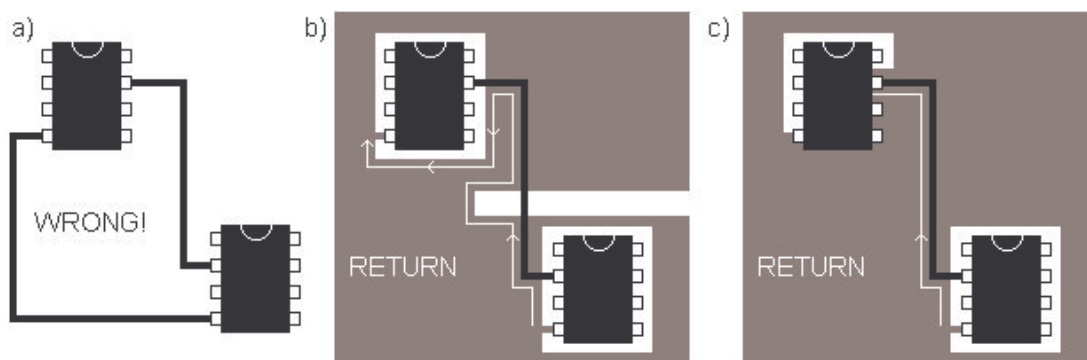


Figure 9.5 Traces creating an antenna: a) Loop antenna, b) slot antenna and c) eliminating the effect.

Straight angles in the traces can create trace reflections: In the apex of the turn the trace width is increased, which affects the distributed capacitance and self-inductance of the trace. The result of this is a reflection [34]. Sharp corners are relic from days when most PCB boards were created using tape masks; one should always avoid them since they increase trace lengths and enhance the reflective effect. Trace angles that are smaller than 90° may create “acid traps” and the reflective effect in them is even higher. Avoid them at all cost. This is another place where rounded corners show their superiority: They create less trace reflections and the acid less likely traps at the corners. The only reason why round-cornered PCB traces are not commonly used is *cost*: They don’t blend into CAD software very easily.

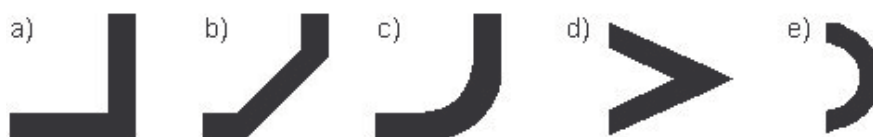


Figure 9.6 Trace angles in PCB: a) Wrong, b) better and c) the correct way. d) Definite NO: Sharp angles create an acid trap and reflections. e) The preferable method to create a sharp angle.

Circuits should have their supply fed *from* their respective filter capacitors and these are usually “sub-categorized”; for example, main filter capacitors feed the whole circuit, main preamplifier capacitors feed the preamplifier circuit and local supply decoupling capacitors feed certain stages. *Never* use a “T” configuration that places the capacitors and power consuming stage in “parallel” but a series connection of capacitors and supply pins of the stage. In other words, incoming tracks must go directly to capacitor terminals and outgoing tracks must leave from these terminals. Any other layout can increase hum and create transient pulses to the supply lines [3]. Using the correct configuration is extremely important with main filter capacitors.

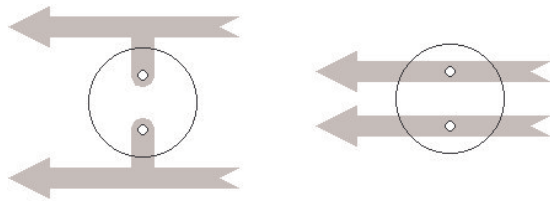


Figure 9.7 Supplying power to a circuit. Left: “T”: The wrong way (“parallel”) and right: The correct way (“series”).

Counter-intuitively, usually the best way to suppress oscillation of an operational amplifier chip is decreasing the value of the close proximity power supply decoupling capacitance. This is because low value capacitors have best high frequency characteristics. Typically a parallel combination of both low- and high value capacitors is needed: The low value capacitors should be physically close to the power pins of the individual chips while higher value capacitors can decouple many chips simultaneously. The decoupling capacitor’s distance to *both* power pin and common return has an equal importance, although several articles falsely claim otherwise. In fact, the inappropriate length of the common return *loop* creates the most severe inductive effect. Remember that the inductance of the trace/wire limits how much current can be delivered in case of a sudden and fast demand for it. Thus the inductance creates *local* sagging at the terminals, which is in turn battled by introducing some *local* capacitance. If there is a notable trace length between the current drawing terminal and the capacitor the resulting inductance of the trace will render the capacitor insignificant. *If the trace distance is more than 1 centimeter you can practically ditch the capacitor as useless.*

It may come as a surprise but the bypass capacitor currents form a loop *with the load*, which means the best routing method for capacitor returns is the straightest path to common return node of the load. This theory is consistent with *all* class-AB output stages, including high-power OpAmps we know as power amplifiers.

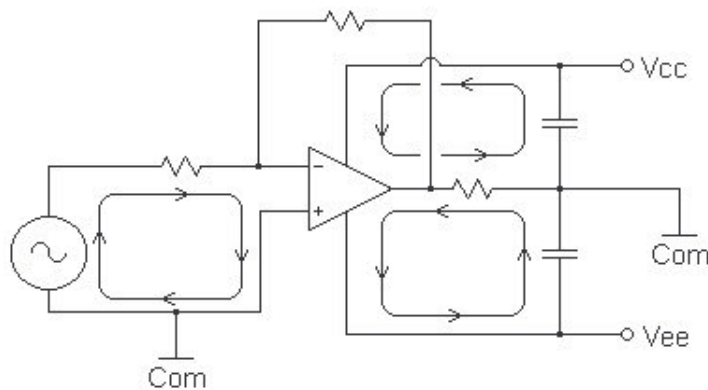


Figure 9.8 Current loops of an inverting operational amplifier circuit.

The proper configuration for parallel-connected parts has already been mentioned in this book but as it concerns the PCB design quite evidently it’s beneficial to mention it again: *In order to have the current divide evenly between the parallel components the track’s impedance properties should be equal for all components.* The picture shows a parallel connection of transistors in a configuration that attempts to keep the track length to each transistor equal. The concept is the same for parallel capacitors and resistors as well. Naturally, this rule has more importance with high-power applications and long trace lengths so scarifying layout to follow the rule is not recommendable in small signal stages. However, memorize the theory and if you can follow it, while still retaining a compact layout, do it.

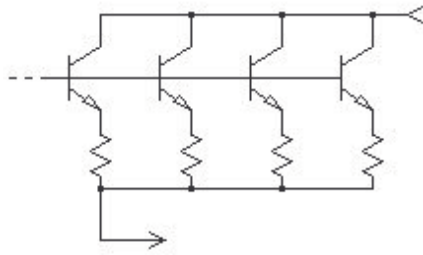


Figure 9.9 Ideal parallel connection of transistors.

The copper tracks can be used for many tricks; quite useful ones are ring guarding and Faraday shielding. A guard ring surrounds a certain circuit part with a “guard track” that is at the exact same potential as the sensitive node. This will effectively protect the guarded area against capacitive effects and leakage sources. Usually the guard junction is a direct connection but in most sensitive applications a low impedance buffer might even be used to drive the guard ring. I doubt one will ever have to resort into that in typical audio applications. The figure shows how a guard ring can be applied to inputs of conventional operational amplifier stages. Another variation of ring guarding ties the ring into common or Earth; however, this may induce some disturbances existing at those potentials.

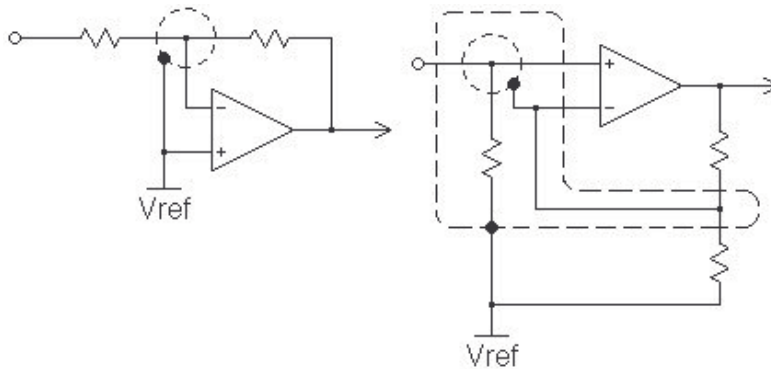


Figure 9.10 Guard ring potentials for inverting and non-inverting operational amplifiers.

A Faraday shield operates like putting a metal plate in front of an inductive or capacitive source. The shield effectively divides the coupling capacitance C and diverts it away by shunting it to the common of the feeding circuit. Note that in this case coupling of the high frequencies is unintentional and one wants to get rid of it. To be fully effective, the shield must return the displaced currents back to their source and completely block the electric field between the noisy source and the protected circuit [35]. A Faraday shield should never be left floating as this can exaggerate the coupling problem.

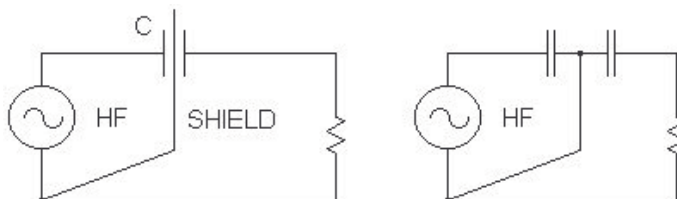


Figure 9.11 Concept of Faraday shielding shows how to prevent unintentional coupling. The equivalent schematic is shown at right.

Few words about other board types are in place: A proper layout is as much a demand on veroboard, stripboard, perfboard and turret boards. If you cannot make a good layout on PC board you most certainly cannot do it on one of these boards either. PC

boards allow arbitrary routing of traces; the “prototype” boards do not - at least not without using a vast amount of jumper wires. There is nothing “easy” in using alternative boards. The first three board types are quite good for prototyping as they allow *slight* changes and small additions of additional components. However, a skillful designer, who is handy with soldering gun, has already realized the need for extra slots for the “optional” and experimental components and can even solder these to the leads of other components if necessary. Notice that all prototype boards are as vulnerable to damage caused by overheating as conventional PC boards. This is where turret boards (and soldering stands) show their claws – however, they usually do not allow making the most compact and consistent layout – and more importantly, cannot support modern multi-pin ICs and components. Requirement for great deal of manual labor in assembling process is also a big issue if mass production is considered. Realize that because of small clearance the tracks of prototype boards cannot be tinned (which means a decrease in current handling capacity) and the small clearance also contributes to capacitive effects between tracks. The only valid reason for not using PC boards is the lack of tools to make them.

In some cases it is worthwhile to consider if using a circuit board is really necessary. Power supply circuits usually deal with high currents so extremely thick and wide traces are required. The filter capacitors come in various shapes and sizes and the rectifier bridge usually requires a heat sink so the best practice is to mount the thing to the chassis anyway. In practice, point-to-point construction of this circuit might prove wiser than using a PCB. The same thing also applies to high-power amplifiers that spread the transistors all over the heat sink and require bulky ballast resistors. These usually require a vast amount of board space and generate a lot of heat so a form of point-to-point construction is once again a considerable option. Another good practice is to run separate wiring from PC boards to jacks, potentiometers and switches – basically to all components that wear out in use or become substantially weaker if mounted directly to PC board.

9.9 Component layout and heat dissipation

The power dissipation rating of components is true *only for a specific mounting orientation*. Mounting resistors “vertically” requires less board space and can reduce trace lengths but realize that this will also decrease the resistor’s power dissipation capability. On top of that, typical 3 to 5-watt power resistors cannot handle the mechanical stress they are subjected to in the vertical mounting configuration and therefore their leads will typically break off. If you *must* orientate the power resistors vertically then use so-called vertical mount resistors that are specially designed for this purpose. If you use conventional components then retaining the horizontal mounting configuration and lifting the resistor body off the board is a recommended way to mount resistors that are expected to get moderately warm. Never place heating components under the board and if you can, improve the air circulation around warm parts by adding ventilation holes to the PCB as well. *If the board is mounted vertically a proper orientation of components will exploit the aid of natural convection in the cooling process.*

Capacitors dissipate most of their heat through the flat (metal) top; therefore it is not advisable to mount them “upside down”. Because capacitors that tend to heat the most are usually the big and heavy ones this would be rather silly anyways. However, note that some capacitor clamps actually heat sink the component. Biggest filter capacitors usually have a different construction since they are designed to be mounted terminals

facing upwards and using either clip (clamp) or stud mounting. With clip mounting it is preferable to have some clearance to the chassis to aid cooling. Most capacitors of this type have a little vent near the terminals. This “safety vent” should not be obstructed and it should have a reasonable clearance to objects in close vicinity. If the capacitor fails the safety vent allows the capacitor to “open” in a controlled way. If the vent is blocked the pressure has no path to escape and the capacitor may explode violently. The same rule applies to capacitors that have the “cut” metal top as well: The cut deliberately weakens the canned structure allowing the capacitor to destroy itself in a controlled manner. Never block the cut surface.

While tightly packed “capacitor banks” are pretty compact and ideal from the layout point of view they are also horrible considering heat dissipation. Capacitors should always have some clearance to each other to assure that heat can be dissipated properly. Every 10° C decrease doubles the capacitor’s lifetime. If cooling really is an issue then it can be improved by heat sinking the capacitor either the conventional way or by using heavy bus bars at the terminal ends. Increasing the cooling airflow will work as well. Of course, the best ways to fight heat are *preventive*. This means increasing voltage and ripple current ratings of the capacitors and power ratings of the resistors: For example, when dissipating the same amount of power a 5-watt resistor stays significantly cooler than a 2-watt resistor. Similarly, a capacitor with a higher voltage rating heats up less than a capacitor with a lower voltage rating. Another preventive measure is to use parallel components to divide the heat.

9.10 About epoxy

Epoxy has earned itself a bad reputation since it was once eagerly used as tool to hide circuitry from unwanted viewers. Other than “annoyingly” hiding the vital trade secrets of manufacturers, (who, to be fair, just tried to earn their living and protect the design they invested a great deal of time and money), the epoxy cover unfortunately made servicing a total nightmare as well. Using a layer of epoxy to cover circuitry - or parts of it - is *extremely* beneficial in some applications but, in my opinion, using epoxy just for the sake of concealing the circuitry is a pretty poor idea. Reputedly these secretive practices have even led to downfall of some major companies. You should think carefully if you wish to choose this path.

Despite its bad reputation epoxy has many great advantages: Epoxy resins are excellent insulators and aside that they also protect from arcing, dust and moisture. In RF applications (and in other circuits that require an extremely tight tolerance for component values) slight bending of components can even detune a resonant circuit and stop it from working completely. Covering the components with a layer of epoxy can naturally prevent this. In TVs and monitors the high-voltage cascades are usually potted in epoxy as well. Even though the aforementioned properties can be extremely beneficial in demanding RF, medical, aero, marine and industrial applications they are less important in such a plain application as amplifying musical instruments.

A very common use for epoxy resin is potting (encapsulating) transformers and inductors, which in turn reduces noise caused by magnetostriction and other magnetic effects. Potting the coils with epoxy resin is also recommendable because *epoxy is much better heat conductor than air*. The aforementioned property also improves temperature sharing among components that critically depend on it (i.e. differential transistor pairs, current mirror transistor pairs, temperature tracking diodes coupled to transistors etc). In some designs the enhanced temperature sharing, introduced by a

drop of epoxy over certain components, can in fact create huge difference in performance. Epoxy can also substitute glue in securing heavy components in place and therefore prevent vibration from breaking their solder joints. This is extremely useful with heavy, high-power, cement resistors and large main filter caps.

9.11 SPICE simulations

SPICE simulation is an effective way to verify a design before it even goes to breadboard level. However, there are certain aspects of SPICE that have to be acknowledged: First of all, simulations are only as accurate as the used component models and the modeled circuit. In practice, the effect of real-time temperature variation within devices cannot be accurately simulated - neither the significant capacitive, inductive and resistive effects of wiring and PC board. Well, the latter could be modeled but creating a realistic mock-up would be extremely difficult and computing times would grow painstakingly long. SPICE also does not model variations that happen in the common node but computes it as an ideal zero voltage potential. This is understandable but also quite unrealistic. To model the common current flow the direct ground connections must be replaced with low value inductors/resistors that emulate the actual wiring. The “wire models” should construct a mockup of the actual common circuit and finally connect together in a simulated common point, which then hooks up to SPICE ground through low resistance. There are naturally a plenty of other means to make the simulation more realistic, however, as the circuit grows more and more complex the simulation time also increases.

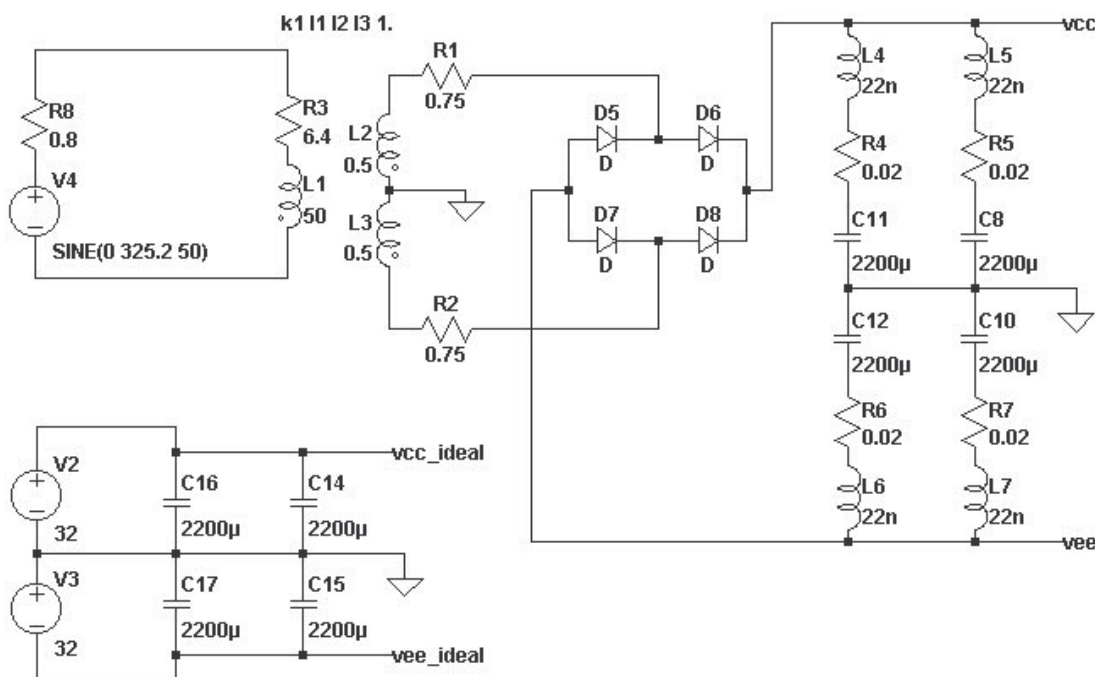


Figure 9.12 SPICE simulations of ± 32 V 250 VA power supply circuit using toroidal transformer. Top: “Realistic” and bottom: Ideal.

There are also “proven” methods to invalidate practically *all* simulation results. First one is using ideal semiconductor models and another one is shown in Figure 9.12. An ideal supply circuit is a nice and *fast* tool for testing various circuits but unfortunately it is also *ideal* and therefore never exists in real life. When the circuit is replaced with a simulation of the mains transformer (including series resistance), rectifier element and filter circuit (each capacitor can be made to include ESR and series inductance) it is quite capable of bringing most circuits down to their knees by the means of induced

parasitic oscillations and - even more so - due to power supply sag. In practice, simulating a circuit using only the ideal supply model tells *virtually nothing* about its true behavior. Adding some series resistance to ideal supply will help a little but there is no denying that the other model behaves quite differently. More or less similar behavior you can also expect in real-life.

The same principle also applies to constant current sources. Again, they are fast and easy tools that were intended for quick testing of circuit ideas but cannot provide realistic results. For sensible results, use as accurate circuit as you possibly can. In practice, there is a limit where simulation of accurate component models becomes too slow and vain. For instance, it is not practical to replace every capacitor, inductor and resistor with a “realistic” model – unless you have an access to a supercomputer. What you need to do is to focus on the areas that really can make some difference. The behavior of the power supply circuit plus the behavior of constant current- and voltage sources are some of these areas. In some places capacitive and inductive properties of board or some components can make a difference as well. Remember: SPICE is just a tool that helps you to get rid of nasty surprises that might occur when you actually build the circuit.

Another frighteningly common mistake is simulating the amplifier’s behavior using *ideal* resistive loads. In practice, no one really plugs his amplifier into a purely resistive load. Unfortunately, the load can have a drastic effect on amplifier’s behavior – even more so if the amplifier is designed poorly to begin with. The chapter concerning speakers will present a method to simulate a speaker (and its enclosure) in SPICE. This method can be used to simulate nearly any dynamic speaker – as long as some of its parameters are known. Respected manufacturers should provide these on a datasheet of the speaker.

In reality, semiconductors are not ideal either and even specimens of a certain model differ from each other. However, SPICE usually provides just a one sample of a certain model. In demanding circuits, such as current mirrors or differential amplifiers, where component matching is important it is usually best practice to make few “worst case” simulations: One of the transistors can be substituted by a higher or lower H_{FE} version (A, B or C) – or even with a completely different model. Circuits should be tweaked to give satisfactory results with varying component parameters as this minimizes the labor of “handpicking” and guarantees admirable results even when components do not exactly match. Some SPICE softwares feature a so-called Monte Carlo analysis that can simulate the component variation.

As a close-up, I present a small summary on how I create a power amplifier design using LTspice/Switcher CAD III software: Usually, at the start of the process I have a pretty good idea of what semiconductors I’m going to use and what closed loop gain the circuit should have. I then start by drafting the basic circuit and run some initial tests using ideal supply to make sure the circuit outputs a perfect sine wave on proper amplitude. At this point I test the circuit with both resistive, ideal load and with a simulation of a speaker load. If the circuit clips too early or oscillates I work to cure the problem with proper methods. It is advisable to leave tampering the feedback circuit or amount of Miller compensation alone and figure out other ways to solve problems. However, do note that amplifiers with LTP input stage usually require a small amount of Miller compensation in order to work – even in such a perfect environment as circuit simulation. During all testing phases I run random AC analyses

to verify the performance at frequency domain (bandwidth and phase shift). I also test the square wave response and see if the circuit exhibits any excessive “ringing”.

Once the initial testing is done and I’m satisfied with the results it’s time to start tweaking: First task is usually balancing the input stage currents. To do this I switch the simulation mode to “DC operating point” and tweak the input stage values so that the DC offset at output is as close to zero volts as it can possibly get. The initial assumption before all this is of course that the resistance tying the input to common is equal to resistance that connects the feedback path to output. The currents are usually pretty well balanced when the DC offset is small. This is also a good moment to make worst-case simulations of varying component parameters.

After the circuit is proven to work optimally with an ideal supply it is time to switch to using a simulation model of a realistic supply. Now, at this stage it is very common that the circuit will experience a notable power loss as the rail voltages sag – it is also very common that the earlier clipping triggers parasitic oscillations. In other words, simulation with realistic supply model will usually reveal all the weaknesses of the design. Unsurprisingly, best way to fight these is to rely on constant current source loading, current mirroring and other proven techniques. Increasing Miller capacitance or high frequency feedback should be the last resort. All simulations are run with varying frequencies, amplitudes and using both resistive and simulated loads. Square wave response is tested again with various frequencies. Note that quite illogically the circuit can sometimes be stable when it’s run near full power but running it on lower input signal amplitude may trigger oscillation. Sometimes the circuit oscillates without input signal or when it clips to rails heavily enough. It is naturally impossible to simulate every scenario but one should try few selected ones. Circuits of startup delays and protection against DC or short circuits are also tested appropriately during this stage.

If the circuit has rail-sticking problems or if it exhibits ringing while clipping the stage currents are tweaked. In the latter case usually some low-pass filtering of the rails feeding the lower current stages helps a lot. For example, feeding the input and VAS stages through a 100-ohm – 22 μ F supply rail RC filter may do wonders for stability.

The final tests exploit a great feature of LTspice, which is that a wave file (in “wav” format) can be assigned as the voltage or current source. Similarly, a wave file can be “recorded” according to voltage (or current) at a specific node (or component). I have some dedicated test files for the purpose that are intended to simulate anything from clean acoustic guitar playing with greatly varying dynamics and transients to “chugging” palm muted death metal riffs with the majority of the signal content being square waves. At this point it is usual to find out that the amplifier *will* color the sound to some extent, as it cannot possibly reproduce the signal with total accuracy. However, the main purpose of this testing is to make sure that the realistic signal content will not trigger any oscillation– and that the small amount of coloration is (if required) either unperceivable by ear or otherwise “pleasant”.

9.12 How to learn more?

Learning circuit theory and audio design is a lifetime task, thus “how to learn more” is an extremely good question. In essence, learning is a very individual process and different methods work better with different people. Some people learn more by reading while some people prefer “hands on” approach. While the latter certainly is

important it must be recognized that the “conventional” method of learning by studying theory books (and alike) usually brings up aspects and issues that you might not even come across on your own.

A vast amount of technical literature has been written to cover the basics of audio power- and preamplifiers, filters, power supply design, signal switching and aspects of circuit board layout and design. Practically, all you likely ever need to know about electronics is without doubt covered in a selection of various tomes. However, you need to bear in mind that majority of this wealth of information is not written from the aspect of such a marginal application as guitar signal amplification. Aside books, good sources of information are application notes, technical white papers, journals and (in some cases) magazines, as the accurateness of these is often reviewed by a selection of people who are experts in the concerned field. Other mentionable sources of information are patents, Wikipedias and websites. These often prove to be extremely useful and educative – however, a good deal of criticism towards their content is required because neither of these has an obligation to provide accurate information. You may also find various (audio) electronics-related Internet forums invaluable.

Extremely useful sources of information, that deserve a special mention, are schematics and service manuals: While doing research for this book I studied literally hundreds of schematics and simulated all interesting and unconventional circuits with SPICE software. For me this approach has proven invaluable. Computer simulation is an extremely handy learning tool because virtually nobody has time (or money) to breadboard, scope and tweak all unordinary circuits they bump into. I don’t exaggerate if I say that schematics have taught me at least as much as ordinary books and alike: In many cases they really open your mind out to new circuit ideas and topologies – plus give you ideas about alternative methods to apply the more common ones.

Concerning schematics, I most warmheartedly recommend studying old schematics of Lab Series, Roland, GMT, Gallien-Krueger, ADA and Acoustic Control amplifiers. These tend to include topologies and circuit solutions that are quite novel and unordinary. Their amplifiers also have a great reputation. Other good schematics to study are those of modern Fender, Randall, Gallien-Krueger, Hughes & Kettner, Peavey, Vox and Roland amplifiers. Sometimes many tube amplifiers employ nice circuit ideas too.

Don’t make the mistake of limiting your selection to guitar amplifier schematics only: PA, HiFi and instrument amplifiers (for basses, keyboards etc.) have often been miles ahead in technology since guitarists (and the people designing guitar amplifiers) tend to be extremely conservative and resentful towards new ideas. If it were a choice of some people we would still be playing only Fender Stratocasters (or Gibson Les Pauls) with amplifiers that were *at least designed* in the 1950’s. Not all of us wish to do so.

10 Speakers

10.1 Power ratings and other misunderstandings related to speakers

When people are dealing with speaker selection it's common that a particularly irrelevant question is asked now and then; that is "can a x -watt amplifier drive a y -inch speaker?" The answer is: *The speaker's physical size and the amplifier's output power rating have no correlation to each other whatsoever*. The only factor affecting the amplifier's output power is the impedance of the speaker, and cone size has nothing to do with that. An amplifier will deliver its rated power to any loudspeaker regardless of the cone size and after that, it's up to the loudspeaker to use the delivered power as efficiently as possible. Thus we come to *efficiency*, a topic dealt later in this chapter.

Power ratings are another subject causing a great deal of confusion, especially when it comes to output power of the amplifier; the latter issue is discussed later in this chapter as well since it strongly goes hand in hand with the speaker selection. When it comes to speakers, *the power rating only indicates how much power the voice coil can withstand without being overheated and damaged*. This is usually expressed in terms of average power – although manufacturers that use the value as a marketing gimmick may also state the peak power rating, as it is higher. Some "vintage" speakers used to quote "British" and "American" wattage, where the "American" rating stood for peak power and "British" rating, more conservatively, for average power. *The power rating is not the same thing as the efficiency of the speaker*, although they often tend to have a correlation due to structure of the speaker. Another form of the power rating involves stating the power required to exceed the maximum voice coil displacement. This is a trickier variable since it is practically useless unless signal frequency and enclosure volume are specified as well.

10.2 Efficiency and SPL ratings

Loudspeaker is a transducer that converts electrical energy to acoustic energy (sound), which simply means varying air pressure. Typical (dynamic) speakers are *extremely inefficient*, that means they can exploit only about few percents of the power fed to them; the remainder is simply converted to *heat*. Matching of acoustic impedance between the loudspeaker (also known as driver unit) and air is especially difficult at lower frequencies due to longer wavelengths. The aforementioned fact should explain why high frequency speaker elements usually have much better sensitivity (efficiency) ratings and require less power.

The unit that expresses efficiency of a loudspeaker is called sound pressure level or SPL. It simply means the air pressure that the speaker produces when it's being fed with a certain power input. The unit has a logarithmic decibel scale. The loose unit definition inevitably causes problems: First of all, the input power must be a product of AC because with a DC signal the speaker exhibits poorest efficiency and just produces heat. But if the efficiency is related to AC power what is the measuring frequency and if the efficiency is related to AC power what is the measuring power? Also, how much distortion can be tolerated in the measurement and what is the distance from the driver to measure the air pressure?

There are no standard answers to the aforementioned questions but fortunately some basic rules have been defined: The sound pressure level is *usually* measured one meter away from the driver, on-axis and using an input signal that is sine wave and produces a watt of (electrical) output power to the nominal load in question. The rule still lacks a definition for signal's frequency (very important) and the rule for using an on-axis direction of measurement is not absolute. One can assume that this direction is chosen since it is typically optimal considering speaker's directivity (explained later). Unfortunately the measurement frequency is often used as a gimmick: The manufacturer can simply state the highest SPL value they managed measure at a certain frequency. Some reputable manufacturers calculate the average value of efficiencies measured throughout the speaker's *effective bandwidth*. Some manufacturers use pink noise input instead of sine waves. The important point is that one can never be completely sure that SPL ratings are comparable. The variation of sound pressure levels throughout the effective bandwidth can also be large enough to render the quoted SPL as quite a useless variable. In other words, a graph that shows the frequency response of a driver tells a lot more about the actual loudness – and tone. This will be discussed later in greater detail. It also seems that the amount of distortion produced by the speaker is never quoted - probably because it's very difficult to measure. One can only assume that it's likely quite high.

The logarithmic decibel scale of air pressure and its relation to output power are also causing a lot of confusion: The reference level is 0 dB (SPL), which correlates to quietest sound (change in pressure) that an average human can perceive. This equals 20 micropascals. 1 dB *increase* in sound pressure stands for about the smallest change in loudness that an average human can perceive. 3 dB counts for about 1.23x increase in perceived loudness, 6 dB counts for about 1.5x increase and 10 dB means doubling up the perceived loudness. Conversely –3 dB converts to a decrease of 0.812x, –6 dB to 0.66x and –10 dB to 0.5x.

To produce a 3 dB increase in sound pressure an amplifier must double its output power. For a 6 dB increase an amplifier needs four times more output power and for a 10 dB increase (which means doubled loudness) it needs ten times more output power. The above readings are (to some portion) subjective since all people perceive loudness at different frequencies a bit differently. Bear in mind that the aforementioned rules are also “averaged”. Real perception of loudness follows the non-linear curves initially defined by Fletcher and Munson (refer to Figure 1.10).

The amount of confusion, that all the variables involved tends to create, is unfortunate since it has lead to a situation where many “experts” spread what they think is correct information. As is often mistakenly claimed, a 6 dB increase in SPL *does not* count for doubling up the loudness – it only counts for *doubling up the sound pressure*. Another common false belief is that the loudness is doubled with only a 3 dB increase in SPL. 3 dB doubles something all right, however, the variable in question is the *signal power*, not loudness. Third common false belief is that the same amount of change in decibels is somewhat louder at different SPL levels (e.g. change from 0 dB to 6 dB compared to a change from 94 dB to 100 dB). Wrong again: In the aforementioned example the *sound pressures* rise from about 20 micropascals to 40 micropascals and from about 1 pascal to 2 pascals. One can notice that the ratio of change is equal in both cases (pressure is doubled), yet the difference in actual units of pressure is astounding. However, *due to our logarithmic hearing the perceived change in loudness is just the same*.

As a general rule, (if using identical speakers) an amplifier with an output power of 100 watts is not ten times louder than a 10-watt amplifier; it is only about two times louder. An amplifier that is four times louder than a 10 W amplifier would have an output power of 1000 watts! Getting ten times louder than that would mean an output power requirement of 21000 W! The last figure stands for about 33.3 dB increase. The important point to learn here is that the output power of an amplifier is essentially quite a *worthless* variable when considering loudness. The amount of output power *will* have a great effect on the amount of distortion and clean headroom though. This is because reproducing the dynamic peaks of the signal requires a vast amount of power. Even with a low power amplifier the nominal sound level can be extremely loud very easily, yet at the same time the dynamic range is usually limited and the signal distorted during peaks. Now, depending on the clipping characteristics this is either quite difficult or very easy to detect. If you want a loud amplifier system the best practice is *always* to use speakers that have the highest SPL ratings as possible. If you *also* want a *clean* and linear system drive your speakers with a high-power amplifier. In all cases it is worthwhile to use a soft limiter to tame the highest signal peaks.

Table 10.1 Speaker SPL and power

Speaker SPL (dB)	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103
Power	Actual SPL																	
1 Watt	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103
2 Watts	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106
4 Watts	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109
8 Watts	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
16 Watts	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115
32 Watts	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118
64 Watts	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121
128 Watts	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124
256 Watts	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127
512 Watts	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130
1024 Watts	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133

Doubling up the cone area (e.g. by using two identical drivers) provides a 3 dB increase in SPL and sound power level. If the output power is *also* doubled the increase will be 6 dB. To get a 6 dB increase just by increasing the cone area would require a fourfold increase. However, these are pretty generalized rules so do not to take them too literally. Different speakers are far from being identical so judging their efficiency based on cone area is useless. However, if we simplify things by assuming that everything is identical and ideal, it can be observed that the relation of cone area and SPL is similar to relation of power and SPL. Thus a half stack would be about 1.5 times louder than a single-driver combo if both used identical speakers. However, this is again just a very vague rule since the combo cabinet may have an open back structure while the half stack may have a closed back (or vice versa). Issues like this will inevitably have a huge effect on efficiency and perceived loudness. There are also many other variables affecting the “loudness” rules. These are dealt later in this book.

Since the SPL is quoted at 1 watt of output power it is fairly easy to calculate the SPL of a speaker with a certain power input. It is:

$$10 \times \text{LOG}(P_{\text{OUTPUT}}) + \text{SPL} \quad \text{(Equation 85)}$$

Watt of acoustical power is equal to 112 dB @ 1 m so a perfect transformation of electrical energy into sound pressure would yield a speaker with a SPL rating of 112 dB. Thus the efficiency of a driver (in percent) is:

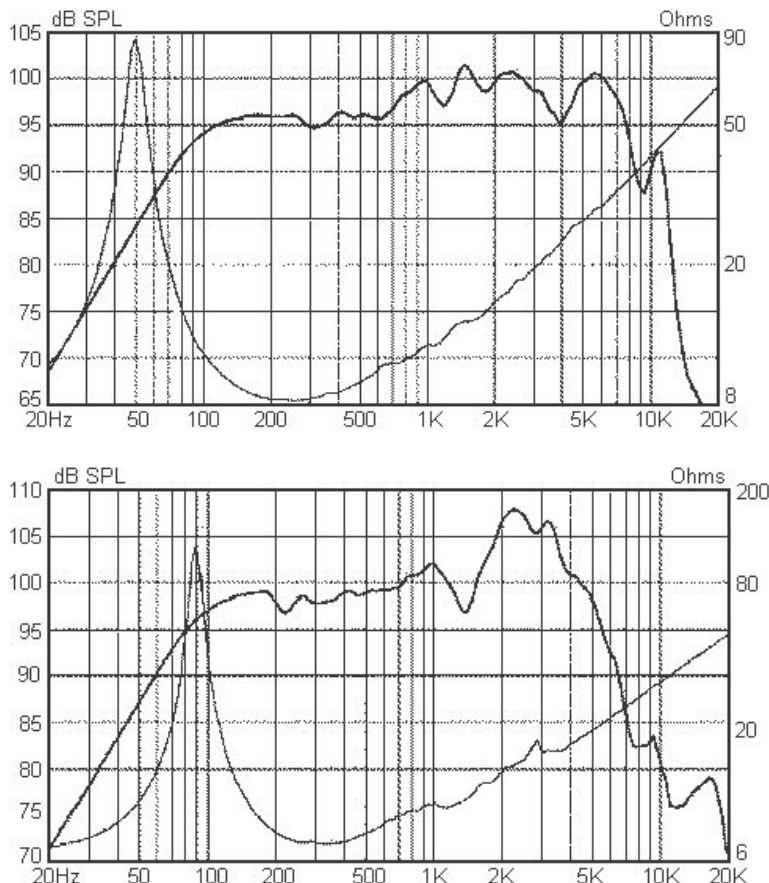
$$100 \times 10^{(SPL-112)/10}$$

(Equation 86)

Doubling the distance from the driver (remember that measuring starts from one meter away) creates an approximately 6 dB drop in the sound pressure level, thus 4 meters away the sound pressure has dropped 12 dB and 8 meters away 18 dB.

10.3 Impedance and frequency response

Guitar speakers color the sound a lot. Due to laws of physics it is simply impossible to construct a speaker that would have a flat response on a wide band of frequencies. The speakers that tend to have good reproduction of bass frequencies usually cannot reproduce the high frequencies and vice versa. Even the effective bandwidth of a typical loudspeaker is not very flat but full of peaks and notches. If the speaker is designed so that these non-linearities do not shift more than 3 dB it can be considered being “high end” quality. Most “flat” speaker systems, used as studio monitors, are therefore usually either two- or three-way designs equipped with complex and linearizing crossover circuits. From guitar speakers, however, one does not expect such linearity and bandwidth.



Figures 10.1 a) and b). Plottings of speaker’s frequency response and impedance. a) Eminence BETA-12LTA and b) Eminence Legend V128.

The plots in Figures 10.1 a) and b) show the frequency response and impedance of Eminence’s BETA-12LTA and Legend V128 speakers. The first one is intended for use as a mid-range or woofer element in PA systems while the latter one is intended for use in guitar amplification. Both have a 12-inch cone. The plotted graphs actually tell a lot: One can see that the frequency response of the guitar speaker rolls off

heavily after 5 kHz, this is characteristic for guitar speakers and actually very typical for all 8" to 15" drivers intended for woofer or mid-range duty. Some bass elements even roll off as early as 2 kHz. The BETA-12LTA is actually quite remarkable speaker to possess such a broad effective bandwidth. The "early" high frequency cut is actually very beneficial in guitar amplification since it attenuates harsh sounding frequencies caused by signal clipping. Some modeling amplifiers that rely solely on the preamplifier as the main source of tone processing are reputed to use "flat" speakers that have a broader bandwidth towards high frequencies and in the case of *acoustic* amplifiers it's even fairly common to see utilization of crossovers with horn or tweeter elements.

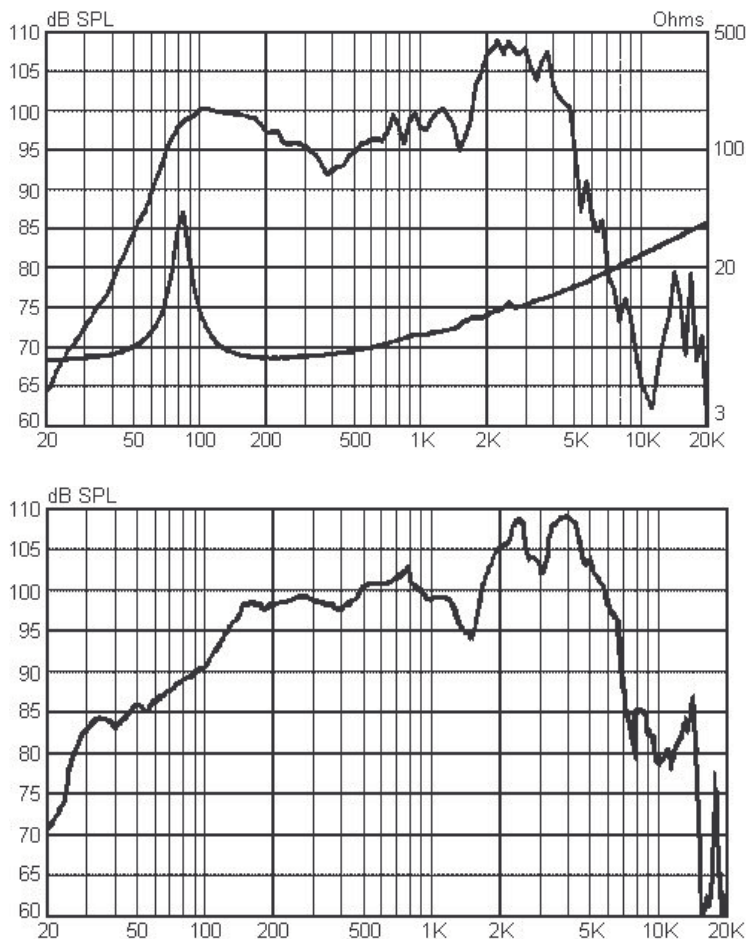
There is another drop in bandwidth at approximately 80 Hz; this is not by coincidence very close to the fundamental frequency of E₂ (open low E string). Although the frequency of the lowest open string is not necessarily the lowest frequency that one can squeeze out of the guitar, limiting all frequencies lower than 80 Hz can add a great deal of "clarity" to the sound. Most midrange woofers having the same cone size tend to have quite identical acoustic impedance so they roll off around 80 Hz as well. Basically one needs a proper bass speaker to reach lower than this. There is a significant difference between typical PA- and guitar speakers that have the same low frequency roll-off point, though. This difference is visible in the Figures 10.1 and will be discussed later.

The Eminence's V128 speaker also strongly emphasizes frequencies from 2 kHz to 3 kHz. They are about 10 dB louder than other frequencies within the effective bandwidth, which converts to a doubled perceived loudness. This seems to be a sought after characteristic since the same emphasize is pretty common for all drivers that are intended for guitar use, it can be seen in another set of example plots shown in Figures 10.2 as well. The emphasized frequencies are basically second and third order harmonics for higher notes of the guitar. The PA speaker doesn't introduce such effect although it has a similar tendency to turn more non-linear towards the high frequencies.

There is a mechanical reason for high frequency non-linearity, which is that the speaker cone simply cannot move quickly enough. Before the response starts to roll off heavily all speakers exhibit *cone breakup modes*, which is distortion that artificially extends the effective bandwidth. Breakup modes are seen as increased amount of "rupturing" in the frequency response plot. HiFi and PA speaker systems generally try to employ crossover filters that roll off before the cone breakup begins to occur (it's distortion after all). Guitar amplifiers, on the other hand, use the cone breakup as an additional effect: Cone breakup is difficult to detect from a clean guitar signal that has little high frequency content. However, the breakup modes become prominent with distorted signals (that are rich in high order harmonics) and introduce *a significant voicing filter*. It is highly subjective and driver-related issue when this effect is beneficial and when it is not. Sometimes it is solely responsible for unpleasant "smearing" of the distorted tones – and unfortunately, instead of the speaker, the amplifier is the one most commonly taking the blame.

What basically should be learned from the graphs showing the plotted frequency responses is that *well alone the fact of swapping the speaker into another model usually makes a greater difference than swapping the amplifier, both in terms of loudness, frequency response and tone*: One can, for example, observe how Celestion's Alnico Gold rolls off approximately *half the perceived loudness* when

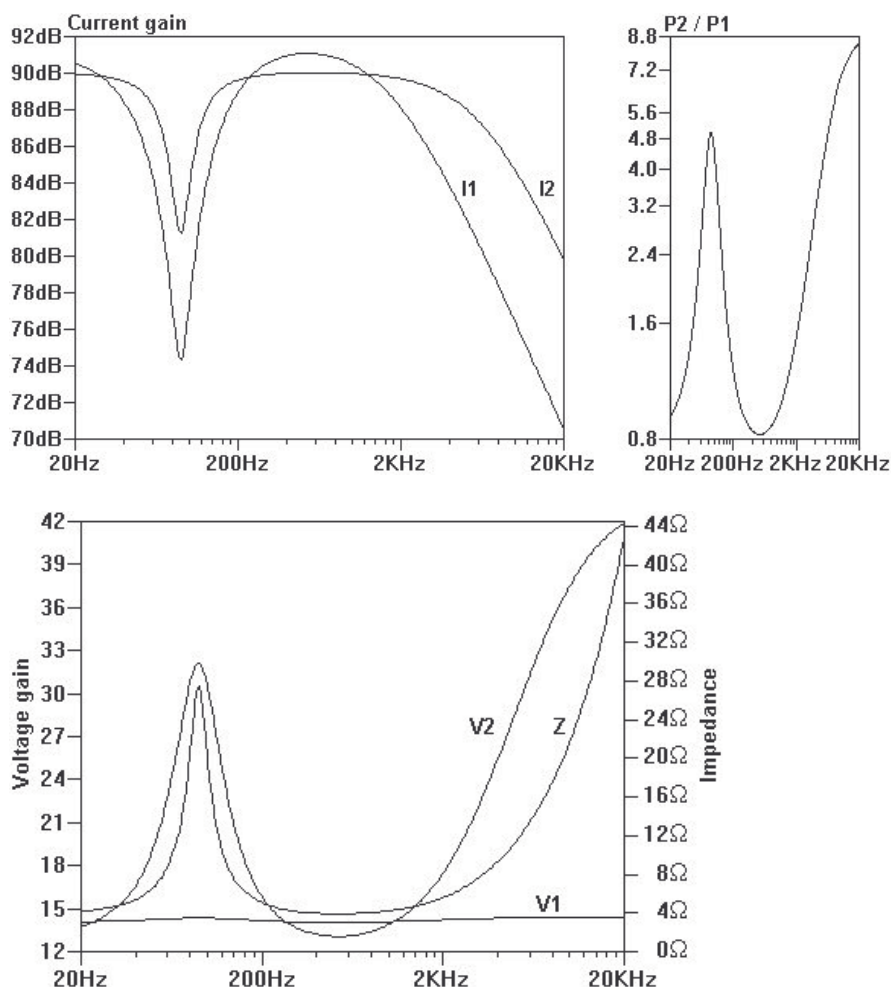
switching over from open G string (196.0 Hz) to open E (82.4 Hz) and how Jensen's Alnico P12Q introduces a similar mid-range scoop that some metal guitarists tweak the tone control circuit to produce. However, I feel this has been enough of tone evaluation. The intention of this chapter is not to discuss subjective preferences over speaker tone but to show that (due to many reasons) speakers are very far from being linear reproducers of sound. What I find a bit frustrating is that instead of publishing meaningful and scientific data (like plots of frequency responses), guitar cabinet manufacturers typically retain to imprecise "tone vocabulary" as a means to describe their products.



Figures 10.2 a) and b). Plottings of speaker's frequency response and impedance. a) Jensen Alnico P12Q and b) Celestion Alnico Gold.

What was shown in most of the plotted graphs as well was the hugely varying impedance of the speaker. This plot is not included in Celestion's case for a simple reason: They do not publish any data that would show, or allow calculating, it – however, they do tell that the speaker's (free air) resonant frequency is 75 Hz, which is a pretty typical value. Since only *current* can magnetize the voice coil the effect of impedance will create more non-linearity to the frequency response. Most solid-state amplifiers are voltage driven – meaning that they try to keep the voltage gain constant regardless of the output impedance. This is beneficial since *ideal* voltage amplifier has *zero-ohm* output impedance and decreasing the output impedance increases the amplifier's efficiency. Unfortunately generating a constant voltage over higher impedance requires less current, which in turn leads to lesser amount of magnetomotive force in the voice coil. A perfect current driven amplifier would keep the load current constant regardless of the impedance but that would lead to huge output voltage when the load impedance is high and the output power requirement for

such an amplifier would be enormous. Increasing the output impedance of a voltage driven amplifier shifts its operation closer to current drive (since in series with a high resistance, changes in the load impedance have less effect) but this also means wasting more power due to increased losses.



Figures 10.3 The effect of load impedance in ideal voltage driven and mixed-mode feedback amplifiers. The load is a simulation of Eminence Legend 875 based on Thiele-Small parameters stated by the manufacturer.

Graphs in the Figures 10.3 can shed a lot of light into what actually happens due to effect of load impedance. The graphs plot outputs of two ideal amplifiers; the first one is voltage driven and the second one utilizes mixed-mode feedback, meaning it varies its output impedance by feeding back some load current. The mixed-mode feedback topology was chosen to this example over pure current driven amplifier since it is realistically more applicable. More about the topology is explained in 3.6.2. Both amplifiers are set up to have a voltage gain of about 14 at 1 kHz. One can note how the voltage gain of mixed-mode feedback amplifier (V_2) tries to follow the load impedance (Z) while the voltage gain of a voltage driven amplifier (V_1) stays nearly linear. The graphs also show that at the resonant frequency the mixed-mode feedback produces over two times higher voltage gain – this is a *serious* requirement since with a $2 V_{PEAK}$ input signal the amplifier's output signal should swing ± 66 volts. The rightmost graph shows that during resonance the requirement for output power is around five times higher in a mixed-mode feedback amplifier; tough demand if one wishes to build a 100 watts at 1 kHz amplifier. Note that this is just one example of how to set up a mixed-mode feedback design and maybe not a very smart one either.

What one can gain by this approach is a higher load current to higher load impedances. This is directly proportional to magnetomotive force that is a product of current and number of turns in the voice coil. In other words, the speaker's cone is moved more efficiently on resonant frequency and on higher frequencies where the impedance starts to rise again. Various people have stated that this brings a character to the tone described as "punch". Vice versa, voltage driven amplifiers with low output impedance will be less efficient. There is sense in both approaches: Underpowered mixed-mode feedback amplifiers (that many of modern guitar amplifiers are) tend to "fart out" when they hit the resonance. Also, at the resonant frequency the speaker's cone movement is naturally the most effective, thus it requires less power than usually. This is another reason for "farting". Plenty of good sounding guitar amplifiers have been made using both topologies and even though most people tend to favor mixed-mode feedback, it would be wrong to state that one of the topologies is superior over another. Most speakers are optimized for voltage drive anyway. Now, if you consider the difference of having the resonant frequency either around 110 Hz to 80 Hz (open A to open low E) or around 60 Hz to 30 Hz it makes a lot of sense why speaker designers tune the PA and guitar speakers differently.

Note: An IEC standard (IEC60268-3) defines that the speaker impedance may exceed the nominal (rated) load but may never fall below it more than 20 % at any frequency, including DC.

The effects of non-linear impedance will become more pronounced when speakers are wired in series; likewise, a parallel connection will decrease those effects. If four *identical* speakers are wired in series-parallel or in parallel-series configuration the impedance properties of a single speaker are retained – at least in theory. In reality, two completely identical speakers will not exist.

Table 10.2 Frequencies of open guitar strings

Open	Frequency	12th fret	Frequency	24th fret	Frequency
E ₂	82.4 Hz	E ₃	164.8 Hz	E ₄	329.6 Hz
A ₂	110.0 Hz	A ₃	220.0 Hz	A ₄	440.0 Hz
D ₃	146.8 Hz	D ₄	293.7 Hz	D ₅	587.3 Hz
G ₃	196.0 Hz	G ₄	392.0 Hz	G ₅	784.0 Hz
B ₃	246.9 Hz	B ₄	493.9 Hz	B ₅	987.8 Hz
E ₄	329.6 Hz	E ₅	659.3 Hz	E ₆	1318.5 Hz

Table 10.2 lists frequencies of guitar strings ranging from low E to high E on 24th fret. Bear in mind that *although these frequencies represent the fundamental note of the string they do not mean the absolute frequency limit of the guitar*. Vibrating string adds at least a portion of second and third harmonics to the tone and techniques like pinch harmonics, palm muting, hitting the guitar body or scraping the pick along the strings can create tones that contain even more harmonics. Output of constant frequency without harmonics would mean a sine wave output.

10.4 Directivity, diffraction and dispersion

Dispersion is another key element in defining the tone of a speaker system and also the main reason why recorded or amplified music never sounds exactly like acoustic music. A diminishing dispersion decreases the beam area of the cone, which in turn enhances loudness because the sound's intensity is proportional to power and beam area. As the air is moving inside a smaller space the sensitivity is increased. This is

true only inside the beam area of course. The relation of intensity, power and area can be expressed with equation:

$$I = P/A \quad \text{(Equation 87)}$$

Sounds with different wavelengths (frequencies) disperse differently, and in practice higher frequencies disperse into a much narrower area. One can easily experience this by standing behind a closed back guitar cabinet, or just moderately aside it. The greatest difference between closed and open back cabinets is created by dispersion since open back cabinets also radiate backwards. One disadvantage of dispersion is that its effects cannot be corrected with an equalizer.

High frequency sounds tend to be more directional since they do not diffract as much. In fact, the diffraction is the factor that causes sound waves from the loudspeaker to spread out, not just go straight ahead. Small loudspeakers cannot efficiently couple signals with large wavelengths but since they will spread those much more the equal loudness patterns of high and low frequencies become much more pronounced. Therefore smaller speakers typically have a poorer off-axis response at higher frequencies. Diffraction also allows sound waves to pass through very small openings, which essentially form another sound source. This is one of the key reasons why sealing the loudspeaker cabinets is an extremely important issue.

10.5 Speaker structures

10.5.1 Dynamic speakers

The cross-sectional diagram shows the structure of a typical dynamic loudspeaker. The diaphragm or cone is attached to voice coil former and suspended by surround and spider. The suspension also keeps the cone centered and returns it to neutral position after movement. A permanent magnet magnetizes the frame that forms the top, bottom and pole pieces. There is a gap between the voice coil (which is wrapped over the former) and the frame, which allows uninterrupted movement of the cone structure. Usually the gap is very small since larger gaps mean a less-efficient speaker. The gap can have either a magnetically symmetrical or a magnetically asymmetric structure. Asymmetric structures promote creation of even order harmonic distortion components while symmetric structures mainly create only odd ones. Depending on the relation of gap and voice coil length the coil is either overhung (coil is longer than the gap) or underhung (coil is shorter than the gap). The pole piece may or may not have an air vent. Dust cap protects the delicate gap area from getting dirty.

Both techniques, overhanging and under-hanging the coil, try to achieve the same thing, which is linear force on the coil. An overhung voice coil tries to achieve linearity by keeping the number of windings within the gap area equal, thus retaining the same magnetic strength throughout the whole cone travel. An underhung coil tries to keep the magnetic flux constant at the cost of reduced linear cone movement area. Both methods have their advantages and disadvantages: An overhung design offers a softer non-linearity when the cone begins to exceed its maximum linear displacement. This behavior is known as *compression*. Speakers with overhung coils are also efficient and have greater power handling capacity. Underhung coils are less efficient and require more powerful magnets but their advantage is the greater linearity within linear displacement operating region. When it is exceeded the distortion increases

more rapidly. The linear displacement region is the region of the voice coil travel where cone movement follows the input current linearly.

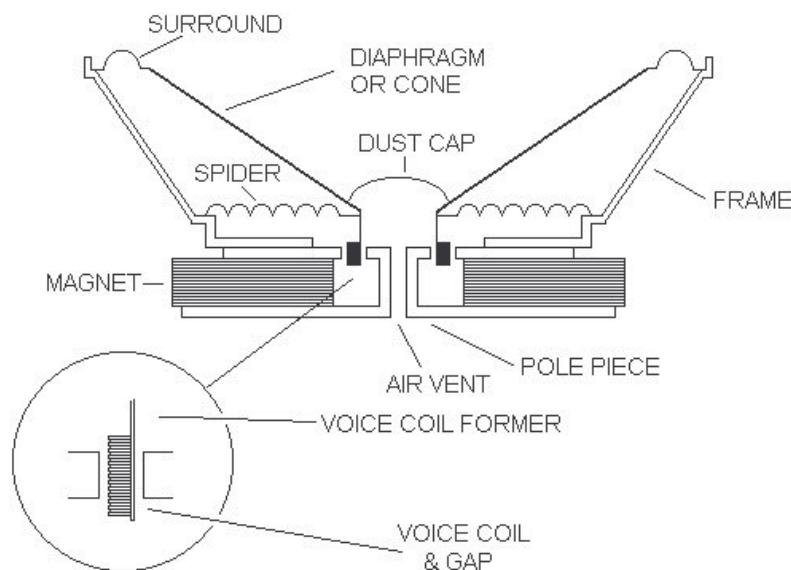


Figure 10.4 Cross-section of a dynamic loudspeaker.

The cone material makes difference as well. Thicker cone material reduces (distortion from) cone deformation under high-level input signals and adds mass, which lowers the speaker's resonant frequency. However, the distortion becomes harsher when the speaker is driven hard. Another way to affect the resonant frequency is making the suspension more or less flexible. Stiff suspension restricts the cone movement and increases the natural resonant frequency while limp suspension cannot even support a lighter cone mass. The latter allows the cone to vibrate at lower frequencies and lowers the resonant frequency. Naturally, larger cones with greater mass tend to have a lower resonant frequency than the smaller cones. One can think of the cone and its suspension as spring that has some mass tied to it.

Even the dust cap has an effect to speaker's tone: The small space between dust cap and pole piece is a resonant cavity with a high Q. A conventional dust cap is made from material that is as acoustically inert as possible but if the material is changed to being acoustically active the dust cap begins to alter high-frequency response: The top end of the signal is now produced by the breakup modes in the dust cap. The harder the dust cap material, the higher the "Q" of the resonant modes is. Aluminum dust caps are used in some occasions since they can radiate some heat of the voice coil away. However, *avoid* them in guitar amplifier applications: An aluminum dust cap is hard material and enhances discordant harmonics (generally all harmonics above 7th), which tends to cause nasty overtones at frequencies above 5 kHz.

A domed dust cap structure may smooth out higher frequencies by radiating on a wider dispersion. A conical dust cap will diminish the dispersion of high frequency signals, thus making the speaker sound harsher from direct angles. One can consider the dust cap structure equal to very inefficient tweeter. Some full-range speakers utilize the similar idea by using a "whizzer cone" that is attached to the voice coil former and diaphragm. A whizzer cone is a smaller, stiff cone structure that enhances the high frequency reproduction. Again, these are not typical in conventional guitar applications.

Note: A tweeter isn't a specific speaker structure type. It is a common name for speaker elements that are dedicated to reproduce high frequencies. As all loudspeaker, tweeters can be dynamic, piezoelectric etc.

Three typical magnet materials in speakers are ceramic, neodymium, and aluminum-nickel-cobalt - or simply AlNiCo. Very often the latter substance also holds a mix of copper and iron. Actually there are many AlNiCo substances and the one used in loudspeakers is usually type 5. Ceramic magnets are made of *sintered* composite of strontium carbonate and iron oxide so they are essentially ferrites. AlNiCo magnets are usually *cast* instead of sintering. This causes them to have higher magnetic field than ceramic magnets and therefore they can be much smaller and lighter. Neodymium magnets are made of combination of neodymium, iron and boron. Out of the three, this material exhibits the highest magnetic field strength. However, neodymium is a rare earth material and thus quite expensive. The greatest advantages of neodymium are the size and weight: The weight of a neodymium magnet is usually about half of that of an equal ceramic magnet.

Using AlNiCo as magnet material is a topic people like to fuzzle about. It is pretty difficult to provide a quick, meaningful and scientific insight to this topic since calculating the repulsive and attractive forces between two magnets is an extremely complex task depending on shape, orientation, magnetization and separation of the magnets. However, one common nominator can be alienated from all equations involved and this is variable "B" – or magnetic flux density, which is higher with AlNiCo magnets due to higher remanence. The outcome of this particular property we already know: An equal AlNiCo magnet is smaller than a ceramic one. In practice, the magnet material has less effect than most people believe. At the time when speaker manufacturers switched from AlNiCo magnets into using ceramic ones they also made a great deal of other changes that more likely explain the differences between tones of "modern" and "vintage" speakers. Cone material and its thickness, suspension of the cone (in both spider and surround), gap length, voice coil size and voice coil material were changed as well and they most certainly affect the speaker's tone more than the magnet material ever will.

10.5.2 Field coil speakers

A field coil speaker is essentially a dynamic loudspeaker that just lacks the permanent magnet. Instead the magnet is generated with DC current flow through a "field coil"; hence the name. The invention was popularly used in between the 1920's and 1930's when permanent magnets were expensive and all materials scarce. Practically, field coil speakers are something one would see in very old vintage tube amplifiers and using one in a modern guitar amplifier would be quite absurd.

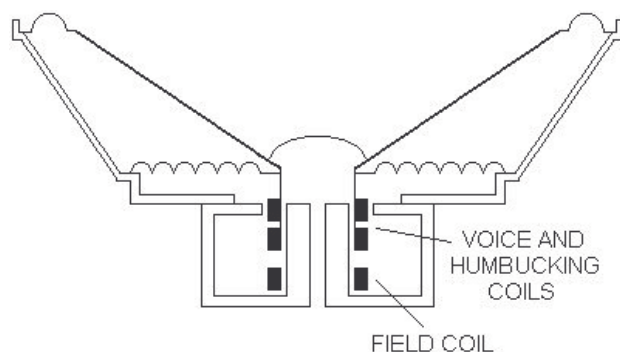


Figure 10.5 Cross-sectional diagram of a field coil speaker.

The DC current to magnetize the field coil was usually generated by making the field coil serve double-duty as a power supply choke – another benefit in generating the most inexpensive *tube* amplifier. Unfortunately, the practice also induces mains frequency hum to the speaker's output. Winding the voice coil in series with another coil in hum-bucking configuration was a common technique to reduce this effect.

10.5.3 Piezoelectric and electrostatic speakers

Piezoelectric speakers are based on piezoelectric crystals that generate voltage under mechanical stress. The piezoelectric effect also works reversibly and when voltage is applied to a piezoelectric crystal it either expands or shrinks, thus being capable of generating sound. Since piezoelectric speakers have no voice coil they do not have inductance either – a fact that allows very efficient high frequency coupling. Piezoelectric speakers are also resistant against overloads that would burn the voice coils of conventional speakers. Due to being highly capacitive, the elements can often be used without a crossover circuit as well. However, piezoelectric elements have limited excursion and they are not loud enough at low frequencies so their use is restricted to tweeter applications. Even in this use a piezoelectric speaker cannot reproduce as low frequencies as conventional tweeter types can. Also, the frequency response is very odd when compared to more expensive dynamic speaker solutions. The capacitive loading may cause oscillation in some systems.

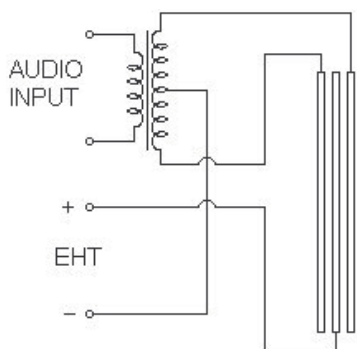


Figure 10.6 Concept of electrostatic speaker.

Up to my knowledge, at the time of writing this, no guitar amplifier has yet used an electrostatic loudspeaker – probably because they tend to be rather delicate, expensive and complex systems and do not give remarkable performance in the concerned application. An electrostatic speaker uses thin and flat diaphragm that is usually plastic and coated with conductive material. The diaphragm is fed with extra high voltage (EHT) and sandwiched between two stators (grids) leaving a small air gap in between. The voltage potential of diaphragm is usually *several kilovolts* in relation to stators. The stators are driven by

audio signal so that front and rear stators are in counter phase. This forms a uniform electrostatic field that moves the diaphragm causing a fluctuation in air pressure. To generate sufficient field strength a transformer, which lifts the signal at stators to higher amplitudes, is required – the same transformer also serves as an impedance matching element since typical electrostatic speakers have rather high impedance. This type of speaker can have a very linear frequency response and it usually exhibits a very prominent directivity. Since the structure is essentially a dipole without an enclosure the bass response is lacking. Also, the speakers are very sensitive to humidity. None of the properties is really ideal for guitar amplifier application.

10.5.4 Ribbon and planar speakers

Planar speakers – or preferably *planar magnetic* speakers, since all considerably flat speakers are considered planar - are pretty similar to electrostatic speakers but do not require high-voltage stators. The diaphragm is stretched non-conductive material and has a metal plate bonded to it. It is placed between plates that have an array of

magnets, which induce the motion of the diaphragm. The structure is dipole so the bass response is once again quite poor.

A ribbon speaker has a metal film suspended in magnetic field. When signal is applied to ribbon it vibrates, and due to its small weight it can do so very fast. However, the ribbon is usually very fragile and can be damaged by large air pressure (similarly to ribbon microphones). They also require very strong magnets and show up as very low impedance thus requiring a step-down transformer. Practically the properties of ribbon speakers are far from ideal to be used in guitar amplifier applications.

10.5.5 Flat panel and NXT

Flat panel speakers are dynamic speakers that replace the cone structure with a flat panel. Because panel is a more flexible structure than a cone it has to be made out of stiffer material. It may also exhibit very odd resonance causing the distortion to rise higher. Typical flat panel structure omits the side suspension, thus the radiating panel is held on its place from the edges, which causes the panel to “flex” in relation to voice coil’s movement. The most common material used in the panels has been Polystyrene (Styrofoam) as it is quite stiff, yet light.

These odd types of speakers were actually used in some guitar amplifiers – mainly in some exciting Yamaha models from the 60’s since the company in question also manufactured the flat panel *Flexion* or “*Natural Sound*” speakers. Fender licensed one of the speaker models to be used in the Bantam Bass amplifier and some electric organs (e.g. Yamaha Electone) used them as well. The speaker type never caught popularity, which could be explained by the fragility of the Styrofoam panels: The speakers blew often. The newer implementation of the structure, NXT, uses a flexible panel material and a vibration triggering “exciter” mounted in off-center position. Both types of panel speakers can exhibit a very wide directivity and they can be reasonably loud as well.

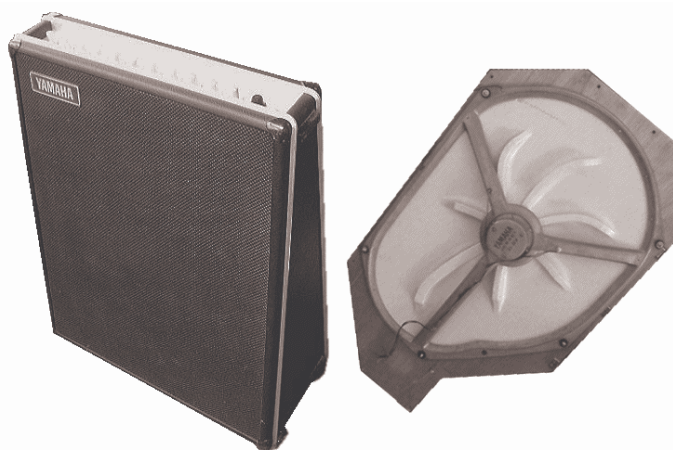


Figure 10.7 Yamaha TA-30 and a Yamaha flat panel speaker.

10.6 How a speaker “dies”

Normal aging deteriorates the speaker’s cone, spider and surround materials. Exposure to moisture or direct sunlight speeds up these effects. In the end, the materials become so fragile that they cannot offer a passable suspension, which causes the cone to travel further than it should. A deteriorated cone may tear up and it creates buzzing sounds.

Besides dying to old age, a speaker can be damaged by exceeding its power rating. This type of damage will *hardly ever* tear the cone apart or off from its suspension. In fact, speaker death due to exceeding the excursion is very rare. In a typical scenario the excessive power overheats the voice coil. This results into two alternative breakdown modes: Either the coil's insulation is melt and the coil is therefore partially shorted – or, even more likely, the voice coil just deforms. The tolerance between voice coil and gap is so small that even small deformations will cause the voice coil to rub itself against the gap – or just jam into it. The first will cause annoying buzz, the latter stops the cone movement completely, thus muting the speaker. Both damages can be verified by gently pushing the speaker cone (which in turn pushes the voice coil). You will likely hear or feel the coil rubbing against the gap or push it loose. In the latter case the speaker will likely function “normally” for a while but jams again in a short time.

One must acknowledge that it is almost too easy to kill a speaker with coil deformation: If we have a 100-watt amplifier driving a typical speaker the speaker can likely convert only 5-watts of its input power to mechanic movement. The rest must be dissipated as heat. Now, think of the size of a 100-watt power resistor in comparison to thin voice coil of a speaker; the equation seems almost impossible. Luckily the poor speaker must work at full power only very small periods of time.

Many PA and instrument speakers are actually equipped with “invisible” protection methods against the roughest abuse: Voice coil wire insulation, coil's support, as well as all kinds of adhesives that the speaker structure requires, are typically made from materials that withstand very high temperatures. The voice coil wire may be thicker than one used in HiFi drivers and the coil may be cooled with the aid of passive convection or the speaker frame. In latter case the voice coil is spaced close to pole piece and top plate to maximize the heat transfer. Some configurations utilize forced cooling that is achieved with the movement of the spider. Forced cooling is more common in high power PA speakers though. Typically the voice coil has less winding at the ends of the cone travel, which means that at the extremes of excursion the force directed to piston is decreased because the coil leaves the linear part of the magnetic field. A stiff suspension may be used to limit the cone travel as well. Special materials that endure the stresses caused by large displacements are utilized.

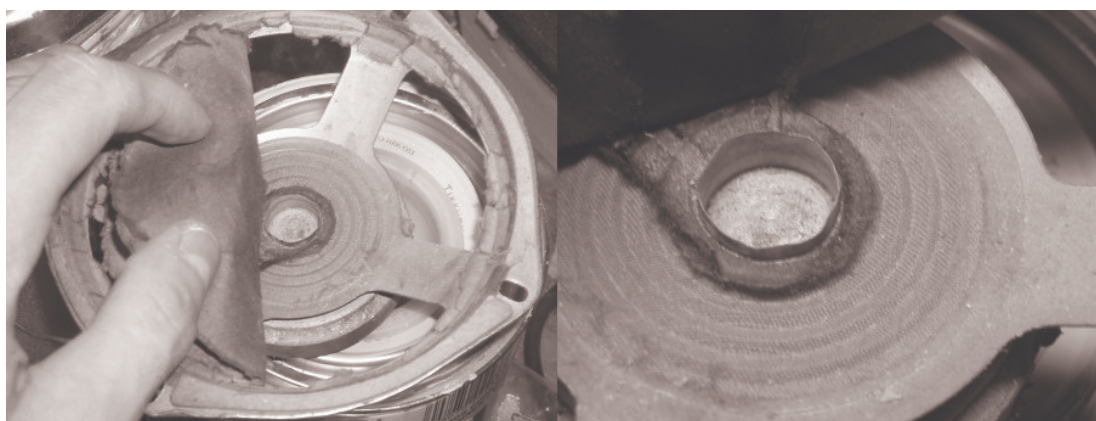


Figure 10.8 The inner structure of a dynamic speaker is much more delicate than the cone and its surround. Voice coil (and its former) must be able to move freely around the pole piece. This is often prevented by voice coil's deformation.

HiFi speakers typically lack many of the aforementioned protections and therefore will not usually last more than a week if used in guitar amplification. This is not to say that HiFi drivers are poor – just that they are designed by different principles. For

instance, some of the protection methods would create too much non-linearity, which in HiFi applications is quite intolerable. HiFi drivers are not optimized for providing a great efficiency and power handling capacity but a linear performance at medium levels of sound pressure- and power where chances to exceed the limits of excursion or temperature are rare anyway.

10.7 Electrical speaker model and Thiele-Small parameters

In the 1960's Neville Thiele published a paper that described how to simulate mechanical parameters of a dynamic loudspeaker with the aid of an equivalent electric circuit model. In 1971 Richard Small enrolled a PhD that expanded and refined the premises presented by Thiele. The theory behind Thiele-Small circuit models is based on dynamical analogies that allow modeling of acoustic and mechanic systems with corresponding electrical components. In short, a *rectilinear mechanic model* of a speaker is a spring that has some mass attached to it. The movement of the spring and mass is further damped by losses of the system. This rectilinear model can be converted into an analogous electrical circuit (or "lumped circuit model"): Spring (compliance) becomes a capacitor, as the component has a property of storing energy similarly as a compressed spring, and mass (usually in respect to earth) converts to inductance. Resistor represents losses caused by friction or viscosity. The resultant circuit can be utilized in SPICE simulations, as it provides a fairly accurate model of speaker's impedance.

The thing causing confusion in lumped circuit modeling is using both *mobility- and impedance-type analogies*. A mobility-type (also known as "inverse") analogy substitutes velocity with voltage and force with current. The concept is usually more suited for modeling mechanic behavior and therefore one can derive the mechanic impedance from $Z_M = F / V$. An impedance-type or "direct" analogy substitutes force with voltage and velocity with current so it's practically an opposite of the mobility-type and passes as a more suitable analogy to acoustic systems. One can derive acoustic impedance from $Z_A = V / F$. One analogy is *dual* of the other, which means that in conversion process between lumped circuit models

- Elements that are in series in one analogy are in parallel in the other and vice versa
- Capacitive elements become inductive, Inductive elements become capacitive and resistive elements become responsive ($1/R$)
- Turns ratios of transformers are inverted
- Voltage sources become current sources and vice versa.

Practically a complete *inversion* of the circuit is created. Note that this inversion considers only electromechanical and electro acoustic models; therefore the voice coil resistance R_E and inductance L_E stay unaltered with both types. It is also practical to learn a way of considering the voltage sources as "constant drop generators" because in an analogous circuit the modeled quantity is independent of what is connected to generator. Same thing applies to current sources, which should be considered as "constant flow generators".

The major elements in the (simplified) lumped speaker model are the real electrical components; voice coil resistance R_E and inductance L_E . They are in series with a branch that represents the driver's rectilinear model consisting of equivalent for cone mass C_{MES} , equivalent for suspension losses R_{ES} , equivalent for suspension

compliance L_{CES} and radiation impedance X_{RS} . As one can note, the analogy is an inverse (mobility) type, which is useful as no circuit element is required to perform the impedance inversion. The effect of X_{RS} element is so small that it can be left out of further equations.

Table 10.3 Conversion between impedance- and mobility-type analogies

Type	Rectilinear mechanic			Acoustic				
	Force	Velocity	Mass	Spring	Losses	Duct	Tank	Losses
Direct	U	I	L	C	R	L	C	R
Inverse	I	U	C	L	1/R	C	L	1/R

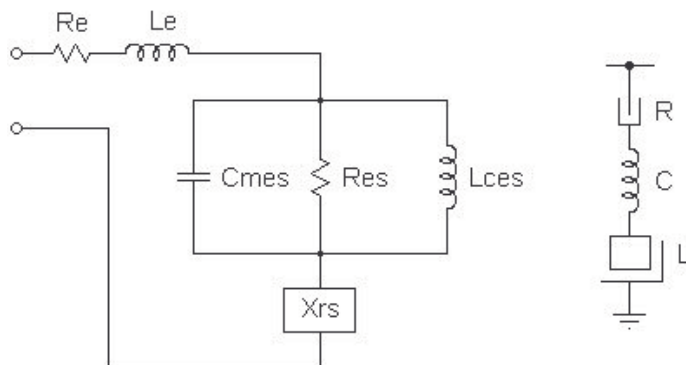


Figure 10.9 Mobility-type lumped model of a loudspeaker.

The electrical values of the components can be derived from their physical counterparts, which are mechanical suspension compliance C_{MS} in m/N, mechanical moving mass compliance M_{MS} in kg and mechanical suspension losses R_{MS} in Ns/m. One also needs to know the Bl product (or “force factor”), which is the product of wire length l immersed in the magnetic field B . This is measured in either N/A or T/M. These values are listed in any respectful datasheet of the concerned speaker. Thus one can calculate the component values using formulas:

$$L_{CES} = C_{MS} \times (Bl)^2 \quad \text{(Equation 88)}$$

$$C_{MES} = M_{MS} / (Bl)^2 \quad \text{(Equation 89)}$$

$$R_{ES} = (Bl)^2 / R_{MS} \quad \text{(Equation 90)}$$

Let’s take a closer look at each derivation. In the formula of mechanical suspension compliance C_{MS} the defining parameter is C_{AS} or acoustic compliance of driver suspension:

$$C_{MS} = C_{AS} / S_D^2 \rightarrow L_{CES} = (C_{AS} \times Bl^2) / S_D^2 = C_{MS} \times Bl^2$$

In the equation of mechanical moving mass compliance M_{MS} the defining parameter is M_{AS} or acoustic mass of driver diaphragm and air load:

$$M_{MS} = M_{AS} \times S_D^2 \rightarrow C_{MES} = (M_{AS} \times S_D^2) / Bl^2 = M_{MS} / Bl^2$$

The final parameter is R_{MS} or mechanical suspension losses. Here the defining parameter is R_{AS} or acoustic resistance of driver suspension.

$$R_{MS} = R_{AS} \times S_D^2 \rightarrow R_{ES} = Bl^2 / (R_{AS} \times S_D^2) = Bl^2 / R_{MS}$$

The lumped model is simplified but in practice it works well in the concerned application. For deeper understanding of the topic recommended reading are of course the original papers by Neville Thiele and Richard Small. However, the basics of analogous circuits are better described in such books as Harry Olson's "Dynamical Analogies" and Leo Beranek's "Acoustics".

The following list introduces some common Thiele-Small parameters:

R_E	Voice coil's resistance	Ohms	Listed in datasheet	Usually about $\frac{3}{4}$ of the nominal impedance.
L_E	Voice coil's inductance	Henries	Listed in datasheet	
F_S	Resonant frequency of the <i>driver</i>	Hz	$1 / [2\pi * \sqrt{(C_{MS}M_{MS})}]$	Point where the combination of moving mass M_{MS} and suspension compliance C_{MS} cause a resonance.
F_C	Resonant frequency of the <i>system</i>	Hz		System resonance for sealed box systems.
F_B	Resonant frequency of the <i>system</i>	Hz		Same as F_C but used for bass reflex systems.
Q_{MS}	Mechanical Q factor		$[2\pi F_S M_{MS}] / R_{MS}$	Losses in the suspension.
Q_{ES}	Electrical Q factor		$[2\pi F_S M_{MS} R_E] / (Bl)^2$	Electrical damping factor. Opposing current caused by voice coil movement.
Q_{TS}	Total Q factor		$(Q_{ES} * Q_{MS}) / (Q_{ES} + Q_{MS})$	Total electric and mechanical damping. Inverse of damping ratio. $Q_{TS} < 0.4$ vented enclosures, $Q_{TS} < 0.7 > 0.4$ sealed enclosures and $Q_{TS} > 0.7$ free-air or infinite baffle.
M_{MS}	Mechanical moving mass	kg	$1 / [(2\pi F_S)^2 C_{MS}]$	Mass of moving speaker parts <i>and acoustic load</i> .
C_{MS}	Mechanical compliance	m/N	$V_{AS} / (\rho_0 c^2 S_D^2)$	
R_{MS}	Mechanical suspension losses	Ns/m	$1 / (2\pi F_S C_{MS} Q_{MS})$	
Bl	Force factor	Wb/m or T/m	$\sqrt{[(2\pi F_S R_E M_{MS}) / Q_{ES}]}$	
V_{AS}	Equivalent acoustic (or compliance) volume	Liters or m^3	$[\rho_0 c^2 S_{DS}] * 1000$ (In liters)	Requirement for air volume to acquire the same stiffness as that of the speaker's suspension when motivated by a piston having an area equal to that of the cone.

X_{MAX} or X_C	Maximum linear displacement	Meters	Listed in datasheet	The distance that voice coil can move in and out while still maintaining the same number of windings in the magnetic gap.
η_O or R_{EF}	Reference efficiency	%	$(4\pi^2 / c^3) * [(f_S^3 * V_{AS}) / Q_{ES}]$	More reliable indicator of efficiency than SPL rating stated by manufacturers.
S_D	Diaphragm area or effective piston area	m^2	Listed in datasheet	The larger the cone, the more air it moves on a given excursion.
V_D	Peak diaphragm displacement volume	Liters	$S_D * X_{MAX}$	The maximum amount of moved air. High values often mean inefficiency.
R_{ES}	Electrical losses	Ohms	$(Bl)^2 / R_{MS}$	
EBP	Efficiency bandwidth product		F_S / Q_{ES}	A variable defining whether the driver should be used in vented or sealed enclosures. ≈ 100 vented, $50 \approx$ sealed.
F3	-3 dB point	Hz		Frequency where sound pressure has fallen 3 dB from the reference level.
Pe	Nominal power	W	$[V_G / (R_{SOURCE} + R_E)]^2 * R_E$	Defines the nominal power. V_G is the open circuit output voltage from source

When doing calculations remember to convert all units accordingly. Some commonly needed conversion rules are:

- 1 l = 0.001 m^3
- 1 cu ft = 28.3168 l
- 1 cc = $cm^3 = 1 \times 10^{-6} m^3 = 0.000001 m^3$
- 1 cc = 0.001 liters
- 1" = 2.54 cm = 0.0254 m
- 1 F = 1000000 μF
- 1 $\mu F = 10 \times 10^{-6} F$

11 Cabinet design

11.1 Materials

The most important rule in cabinet material selection is to avoid particleboard (“chip board”) like plague: It is cheap and lightweight but sucks humidity like a sponge, swells and becomes soft. If soaked badly, the particleboard crumbles and will literally disintegrate in higher sound pressures. Usually before this the particleboard has rattled its joints and screw holes loose. Even worse is the situation when the particleboard is coated with a paint or similar “protective” material: The board will still suck all moisture from air but due to coating it cannot evaporate it and consequently will rot even faster. If you, for some reason, have to use particleboard then impregnate and coat it with water resistant sealant, then use as much bracing in the construction as possible. This will naturally eliminate the “weight” advantage so you still may wish to consider the material choice. Corners of particleboard cabinets get dented very easily. Worse yet, some particleboards have so weak binder substance that one can carve through the cabinet walls with bare hands! Such material is not durable enough to withstand weary life on the road. The only reasonable application for particleboard I can think of (besides cheap mass production) is using it as prototype cabinet material.

Fiberboard is made out of food fibers so it is more durable than simple particleboard. Because the amount of binder in relation to wood fibers (density) is higher than in particleboard the fiberboard is also a lot more resistant against humidity. For the same reason the fiberboard also weights a lot more. MDF is not a specific material but an abbreviation of Medium Density Fiberboard – so, fiberboard and MDF essentially mean the same thing. MDF is also known as custom wood or craft wood. Structurally and acoustically MDF is very good since it’s consistent on its strength and density - which is a thing that cannot be said of all materials. The consistency helps in creating a “neutral” sound without additional resonances or coloration. MDF is also denser than plywood of equal thickness and thus resonates even less. However, it weights a lot and is still somewhat vulnerable to moisture, which initially causes swelling, warps and expansion. Even worse, one of the binder substances is a known carcinogen that irritates eyes and lungs. MDF is also quite hard on saw plates. Serious HiFi people prefer even denser, non-resonant materials like hardboard (HDF) or stone. The weight of these materials is outrageous.

Plywood is made out from sheets of wood (veneer) that are stacked on each other so that the directions of ply’s grains are in a 90° angle against each other. The structure is very sturdy and resistant to bending, cracking, twisting and warping that solid wood structures exhibit. It also tolerates moisture a lot better than fiberboard materials. However, be aware that the internal structure of plywood is not necessarily very consistent and there may be voids that are just hidden by the top plies. This is quite common especially with *cheaper* grade plywoods. The poor quality of plywood has a habit of manifesting itself as increased cabinet resonance (known as “sympathetic vibration”). Typical - and cheapest - plywood material is birch.

Building cabinets out of wood planks was a common technique up to the 1950’s – mostly because materials like plywood or MDF were either unheard of or too expensive. It is essential to notice that as soon as these alternative materials became cheaper (or equally cheap) as common wood the manufacturers joyfully adopted them.

There really isn't that much difference between tones of wooden and artificially wooden cabinets so one should really consider if it's worthwhile to build a solid wood cabinet. Wood has a nasty habit of warping and cracking if it is assembled "wet" and most of the stuff bought from an average lumberyard *is* wet. The proper percentage of moisture is about 4 % (how many have an access to moisture meter) and to retain this proper level the cabinet must be "sealed" with paint or lacquer both internally and externally. If this is not done the wood will shrink or expand in relation to humidity.

Generally, all solid wood cabinet structures have to be braced to prevent drying wood from cracking openings between planks and preferably the planks should be attached together using tongue-and-groove joints. The maximum width of the wooden planks is very limited as well. Solid wood cabinets also tend to be heavy, yet resonant, which is not exactly the very finest combination. However, properly built they are very durable and can look very pretty if you like the oiled or lacquered wooden look, (which unfortunately won't last a second in equipment that sees frequent gigging). A common material choice has been (and still is) pine. Historically pine was used *because it was cheap* (it still is) – not because of its acoustic properties, which are more than a bit lacking when compared to some dedicated "instrument woods". There *is* one significant acoustic property, though: Pine is very homogenous in density and therefore creates similar "dead" and "uncolored" sound as fiberboard materials.

My personal choice for fully wooden cabinets would be *thermotimber*: This is wood that has been artificially aged by "baking" it and in the process it has both dried almost completely and changed in cellular level: When thermotimber ages or becomes subjected to changes in humidity it exhibits no warping or cracking. Thermotimber is also lighter than typical lumberyard woods. Unfortunately, it is also hard to acquire and therefore very expensive. The artificial aging also makes the wood substantially weaker, which means it will get dented easier. Since the wood has dried almost completely it is also quite fragile as well.

With all materials, the thickness in relation to density is an essential parameter that affects both the resonance of the cabinet and its strength. The less internal bracing is used, the thicker the board material should be.

Never use glass wool as acoustic stuffing. The material is light, has very good damping properties and is easy to sculpt. Practically it is nearly perfect acoustic stuffing material - except that the high air pressure, created by the speaker cone movement, will dust off the tiny and spiky glass fibers, which will then be carried into the ambient air and inhaled. This is not a major concern in HiFi applications where sound pressures stay low but once again an example of how one has to design based on different principles. Never, ever, use rock wool either: This material contains magnetic particles that eventually wind up into the voice coil thus blocking the gap. It is best to use sheep wool or some synthetic fiber brands (like "Dacron", "AcoustaStuf", Poly-Fil").

Hardware, like covering material, corners, handles and casters, is usually sold by same places that sell auto HiFi stuff – the downside is that these places also charge a high price on quite inexpensive materials. Hardware stores may sometimes act as another source and usually their prices for the same stuff are much lower. All recessed handles should preferably be made out of metal. If they are "bail handles", make sure they are spring-loaded so that they will not rattle. There is not much difference between the strength of metal and plastic corner pieces as long as they have been made properly.

Some plastics will wear out and crumble; especially in cold temperatures most plastics tend to lose their flexibility and break easier. Anyway, the task of a corner piece is to protect cabinet corners from direct blows (as the corners are mainly the weakest parts in the structure) and this is a task that all of the stuff sold out there should be able to accomplish perfectly. In most cases the choice of hardware is purely aesthetic or driven by economics.

Tolex (and other plastic cover materials) are *overrated* and used mainly because they are *cheap*. You will see this type of covering used in most of the equipment but you will also see that in equipment that has been in real use the covers are all torn up: The material can't handle life on the road. The best deals of vinyl cover materials are likely found from local textile stores rather than from the stores that are specialized for selling Auto-HiFi or guitar amplifier hardware. Tweed is more durable than tolex but also wears out pretty badly. It is also a lot more expensive and harder to find. The textile store might again prove to be the best source. Note that tweed is usually impregnated with shellac or lacquer that stiffens the fibers and keeps the cloth together when tears occur. The lacquer also protects the cloth from spilled liquids and dust so varnishing is a good practice for other fabrics as well. Because most garments are inherently "fuzzy" a china bristle brush may prove to be a better tool for applying the varnish than an aerosol – at least at first. If the shabby look and extreme durability is a concern to you, use carpet materials – similar to ones seen in bass amplifiers. They are extremely tough. Vintage freaks will naturally hate the idea that the amplifier looks like an average doormat. As far as I know, the most durable coverings are polyurethanes that are spray painted on the cabinet, (e.g. "DuraTex", "Warnex" or "Poly Tuff"). These look a lot like tolex but will not tear. Because they are practically painted in place you don't have to worry about seam work either.

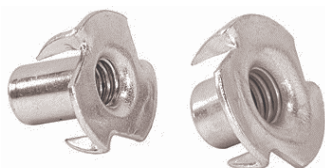


Figure 11.1 T-nuts should be used for secure mounting.

All bolts, used in mounting speakers and handles, should be secured with T-nuts. Never use ordinary wood screws in the task, especially if the cabinets are heavy. Ordinary woodscrews get torn and rattled loose – especially from particle- and fiberboard materials. A bolt secured with a T-nut spreads the force applied to the nut on a larger area and therefore is a lot more durable than an ordinary woodscrew. It also cannot rattle completely loose unless the board it has been mounted on is in an extremely bad condition.

Standard ¼-inch phone jacks are not the best option as speaker connectors: The effective contact surface of the tip connector is extremely small and although most of the jacks exhibit a current handling capacity of 10 to 15 amperes one of the main concerns is still their reliability: A typical phone jack does not have a locking mechanism so the plug easily gets pulled out of its socket. For a tube amplifier this can be disastrous, for a solid-state amplifier usually just extremely annoying – especially in the middle of a show. Phone jacks with locking mechanism exist but finding them (with a moderate price) might prove to be extremely difficult. It should also be noted that typical TRS jacks and plugs are prone to periodic shorts when they are being plugged in or pulled out. This was the sole reason why many old guitar amplifiers used peculiar speaker jacks. It is also common that inexperienced people may use guitar cords with them; these have an inadequate current carrying capacity.

Typical banana plugs together with binding post connectors have a larger contact area and a current rating of about 15 amperes. A banana plug is fast to connect but has the

same problem with plugs accidentally pulling out. Binding post connectors are better for fixed installations but in a long-term use, where connections are made and broken regularly, they are not very reliable and wear out the (stripped) speaker cables.

Speakon connectors are more reliable solution as they provide more contact surface together with a locking mechanism. They also force people to use dedicated speaker wires. However, a company called Neutrik holds the patent for the connector and therefore they are quite expensive. Speakon jacks come in 2-, 4- and 8-contact versions. These are, respectively, NL2MP, NL4MP and NL8MP. Model variations include, for example, NL4MPR (round panel) and NL4MP-ST, which has “set-screw” termination points instead of spade or solder termination points. The two-contact point plug will fit into both 2- and 4-contact jacks (NL2FC and NL4FC) but the 4-contact plug will work only with the four-contact jack. Therefore, you are better off using only NL4MP. Typical configurations have no need for jacks and plugs with eight contact points. The labeling of Speakon is also a bit puzzling: The standard 4-conductor jack has four pins labeled +1, -1, +2 and -2. In typical configurations you wish to use the set of “1” because these are the ones that are compatible with NL2FC plugs. Note that some systems may use very random wiring configuration for the four terminals. Since 6.33 mm (1/4”) phone connector is still somewhat a “standard” it is convenient to have one wired in parallel with the Speakon connector. Neutrik has also developed a jack called “Speakon Combo” that supports connection of both Speakon plug and standard quarter inch *TS* phone jack. Unfortunately, this is only compatible with 2-contact connector (NL2FC).

11.2 Construction techniques

Joinery and bracing are aspects of construction that have a major impact on how sound and reliable the cabinet will be in a long run. A fairly solid design without an excessive amount of bracing requires good woodworking skills and tools because reliable and durable joints are rather difficult to cut and make. The simple butt joint itself is extremely weak when exposed to direct pressure – in practice, using almost as simple miter joints is always a better solution. However cutting them will require a special angle jig because cuts have to be precise. Remember, that in 99 % of applications the butt joint construction has to be coupled with bracing. Although miter joints have a neater exterior look they are nearly as weak as butt joints, which means they also require bracing.

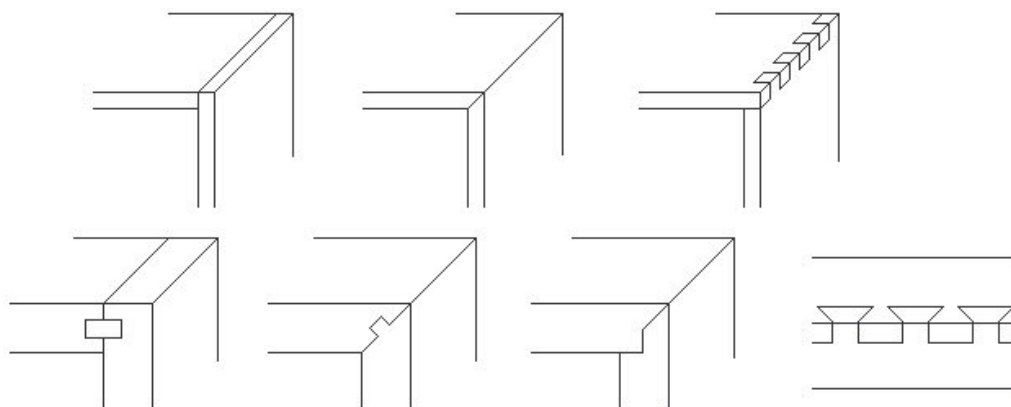


Figure 11.2 Joinery techniques. Top, from left to right: Butt joint, miter joint, and finger joint. Bottom, from left to right are sub structures: Butt joint with added dowels, tongue and groove miter joint, combination of miter and butt joint and dovetailed finger joints.

The requirement for bracing can be avoided with more durable jointing techniques – such as utilization of dowels and finger- or dovetailed finger joints. These joints are lot more tolerant against bending forces and therefore require less bracing. Dowels, tongue and groove structure or combined miter and butt joint are other examples of better jointing techniques. However, for these, good woodworking skills and jigs are mandatory. Even these fairly durable techniques benefit from use of external bracing, which becomes more and more a demand as sound pressures increase.

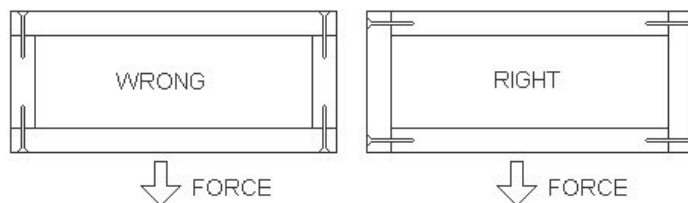


Figure 11.3 Proper structure will not target all forces to bottom and top pieces when the cabinet is lifted.

When butt joints without external bracing are used, the construction detail presented above should be obvious: The force of gravity is directed downwards which means that in the left configuration (assuming the carrying handle is mounted on the top) the top plate carries all the weight and the bottom plate has to support the weight of the chassis. Such is not a durable configuration! The correct structure can divide the weight more evenly by applying the supporting force on the side panels of the cabinet. And speaking about carrying handles: These are often put to places that are aesthetically pleasing but not very ideal functionally and ergonomically. In all other but the lightest amplifiers, the top-placed handles are generally pretty useless if the option is to use handles mounted on the sides of the cabinet. Two handles also share the stress more evenly, thus the configuration is much more durable. Single handle is of course cheaper than two of them but in many cases even a head cabinet is easier to carry from one side-mounted handle than from the handle located at the top. The aesthetics are also pretty much the only reason why side-mounted handles are mounted vertically straight. In fact, a cabinet is usually much more easier to carry (at least by a single person) if the handles are a bit “slanted”.

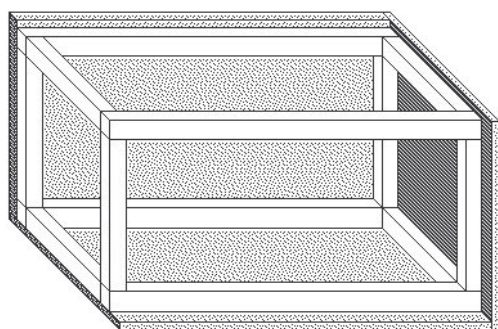


Figure 11.4 An example of bracing.

Bracing is crucial as it (in most designs) is the part that is actually holding the whole cabinet structure together as well as providing damping against vibration. The most common technique for small and open-backed cabinets is to brace just few of the corners – however, as the sound pressures increase the cabinet will also begin to require bracing to the middle. A very common and effective technique is to insert a brace that connects the center of the baffle board to the rear panel of the cabinet. In practice, the higher the sound pressures and the lower the produced frequencies, the more bracing the cabinet requires. If there is a need for a lot of external bracing the

braces should be divided on *odd* distances because any symmetry will enhance standing waves.

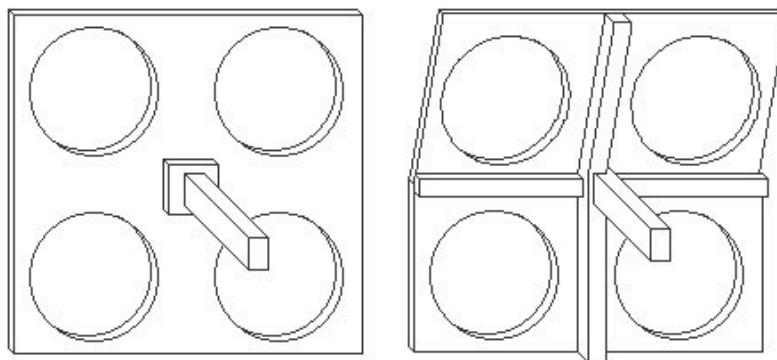


Figure 11.5 Baffle board construction and bracing techniques. Left: Straight baffle and four drivers. Right: Slanted baffle and four drivers.

A durable construction requires a vast amount of wood screws since in practice it is not wise to trust that the glue holds pieces firmly together. All screw holes should be pre-drilled in order to prevent the wood from cracking up when the screw is fastened. As was already recommended, use T-nuts and washers to increase surface area and durability of all bolted connections.

Front and rear mounted speakers behave differently: Rear mounting a speaker (this means that the speaker is mounted on the rear side of the baffle board) will create a cavity in front of the speaker. Since the radiated sound waves diffract from any angles (this includes cabinet corners as well) the cavity will cause an odd *diffraction* pattern. If closed cabinet structures are used the front mounting usually permits creating a better-sealed cabinet as the baffle board doesn't need to be removable. Another method to mount speakers is *flush* mounting. This technique embeds the speaker frame's rims at the same level as the baffle board. The aim is to reduce diffractions even more and one can see this method used in many HiFi speakers. In order to make this work the baffle board must be quite thick to compensate the cut.

Few words about finishing the cabinet are maybe in place: If the cabinet is painted or lacquered be sure that the surface is sanded extremely even: The paint or lacquer will reveal every dent and scratch mercilessly - and so will tolex, by the way. To prepare the cabinet for painting, sand it with a fine-grit sanding paper and then clean it thoroughly from dust. After this you may also stain the cabinet if you wish to do so. The following is important: The first layer of lacquer or paint should be as thin as possible and the paint/lacquer used in it should be *heavily diluted* with the proper diluting substance. The following layers can be painted or varnished without diluting the paint or lacquer (although this is merely for comfort and better results are gained with slight diluting). Before lacquering the following layers, *always* sand the shiny surface smooth with a fine-grit paper. Paint doesn't need sanding but usually it won't hurt either. The important thing is to never use too much paint or lacquer: It will either pour or "bubble" and look very sloppy.

11.3 About acoustic design of guitar cabinets

It hasn't been a long ago when guitar amplifier manufacturers finally started paying more attention to acoustical design of speaker cabinets. For a long time the guitar cabinet design has been evolving around choosing a set of speakers that are stuffed inside a random enclosure without sacrificing too much time to think what acoustic

effects this causes. Fender's most famous design is perhaps a *bass* amplifier with an *open back cabinet* housing four 10-inch speakers that had a *low end cut at 90 Hz!* The speakers were screwed to a *very thin baffle board!* Unsurprisingly it was a horrible design for bass but happened to work great with guitar. Jim Marshall has openly admitted that to one of their designs they simply chose the smallest cabinet dimensions that could house four 12-inch speakers. Four speakers were used instead of two only because the efficient amplifiers were constantly blowing the low-power drivers and more power sharing was required. Also, the novel slanted cabinet design was a result of trying to make the cabinet look more "technical". In some respect guitar cabinet design has been very irrational and many esteemed designs are products of dumb luck. Only quite recently have guitar amplifier manufacturers started to get a grasp on techniques that HiFi and PA speaker builders have known for decades.

Most acoustic design ideas are based on either tuning the bass response or adjusting the dispersion of higher frequencies. Some designs try to improve the efficiency of the signal coupling between the driver and air. When it comes to acoustics, guitar cabinets are not extremely demanding applications because their response does not have to reach very low or very high. In fact, even poor designs can produce at least acceptable results. Basically, large portion of the design principles presented in this chapter are more related to design of bass guitar cabinets and alike. However, it has been noticed in several occasions that tuning the acoustical properties of a guitar cabinet can yield better results than building cabinets with random dimensions and without a clue of acoustics. Understanding basic principles allows fine-tuning the designs to exhibit certain characteristics and ensures that they will also sound good. Some design principles are simply universal and concern all applications.

This chapter also represents some lumped circuit models based on Thiele-Small parameters. Be noted that these should be interpreted with a grain of salt: First of all, Thiele-Small parameters are designed mainly for low frequency reproducing systems that use sealed or vented cabinets and operate the speaker at its *linear* area. Practically, the design theory based on Thiele-Small parameters has its greatest use in designing subwoofer enclosures and alike. On frequencies where guitar amplifiers operate most Thiele-Small parameters pose only a minimal concern and for designs that use open backed cabinets or operate speakers at the limits of non-linear excursion they are virtually useless. The parameters cannot effectively model the interference of adjacent sound sources or structural design of the cabinet - like the effect of tilting - either. However, the models work within certain limits and they can be used for checking if certain enclosures are appropriate. In some cases they can be even used in fine-tuning the design.

11.4 Resonance

Since loudspeaker cabinets in guitar amplifier applications usually act as sound *producing* elements their design is based on different principles than the design of loudspeaker cabinets intended for sound *reproducing* (the latter are mainly used in PA, domestic or HiFi applications). PA and HiFi applications require that the cabinet is as transparent as possible and that resonance originating from it is damped. Guitar cabinets *sometimes* benefit from introduced resonance and harmonics created by it. This is a matter of taste and each cabinet designer tends to have his own opinion about it. Some people prefer hearing only the sound of the speakers and therefore try to completely dampen the cabinet. My opinion is that too much resonance converts into rattle - especially when dealing with bass guitar amplifier cabinets – and rattling

cabinets sound horrible. Little amount of resonance, especially that originating from the baffle board, can be beneficial.

People usually fall for marketing hype that compares the used cabinet materials to a soundboard of an instrument. In extremes, the hype is used just to sell you expensive solid wood cabinets. An even more ludicrous statement made by some boutique cabinet manufacturers is that the covering dampens the material's natural resonance and using it is equal to covering a Stradivarius violin with tolex. This statement is utter nonsense: Material wise, *stiffness - or more precisely, "elasticity" - and mass* have the greatest impacts on the amount of resonance. Thin and light baffle board and cabinet walls resonate more than thick and heavy ones. If the material is thick and considerably heavy (such as 1-inch thick plywood board) damping it will require a lot more mass than what is needed to dampen a 1/8" plywood board. If the cabinet walls were as thin as the woods in violins (and some other instruments) the covering indeed would have effect on tone but as we all know – they aren't.

When the cabinet walls are thick enough and adequately braced the greatest amounts of overtones are created by the resonances of the baffle board and the cabinet as a *whole*. Greatest factor affecting the latter is the cabinet's *mass*. Even when certain parts of the cabinet resonate the covering has no effect on that since it's simply too light in respect to mass of the walls and therefore rather ineffective in damping anything. Unless the covering is a *very* heavy carpet or a thick pad of foam, that makes the amplifier "float" in respect to floor it stands on, you can safely ignore its effect. Even the more easily resonating baffle boards are thick enough not to suffer from effects of covering, which can be perceived by stripping off the grille cloth: This makes no difference for the tone either, except by removing an obstacle that diffracts higher frequencies. In practice, you do not have to use expensive instrument woods and other gimmicks to make a great sounding cabinet. In fact, such materials might even make the whole design process unintentionally difficult.

Essentially, even the simplest cabinets have numerous ways to resonate: These can be divided into three main categories, which are 1) panel resonance, 2) air resonance and 3) driver resonance. Let's discuss them:

Each panel of the cabinet exhibits its individual resonant frequency and a series of its overtones. These are mainly dependent on such factors as panel material's density and modulus of its elasticity - as well as panel's dimensions and the way they are fixed on the edges or braced. The fundamental resonant frequency is controlled with stiffness while the overtones are controlled with mass. Principally, the stiffer the material, the higher the resonant frequency is (and if stiffness is sufficient the resonant frequency is raised above the cabinet's pass band). Similarly, the heavier the panel, the more damped the overtones are. Panel resonance is generated by the cabinet pressure (rear wave of the driver), which means the resonance is in anti-phase with the signal radiated from the front of the cone. Resonance always stores energy and in case of the baffle panel, the stored energy results in a poor transient response of the system. Bracing can increase the stiffness without effecting the mass too considerably. A brace that connects the centers of two opposing panels is so effective because it suppresses the fundamental mode of resonance. As a side note, solid wood is not uniform so its elastic modulus depends on direction, being greatest along the grain.

Air resonance will affect both the panels and loudspeaker's cone. Since the cone is significantly thinner than the cabinet walls the HiFi approach is to compensate this by

damping the panels with some acoustic treatment. The largest resonances appear in reflex cabinets since they are designed to act as Helmholtz resonators anyway. The resonant frequency of a reflex system is so low that the panels are essentially stiffness controlled. This means that panels must be thick enough so that the resonant frequency will not be affected too much.

The loudspeaker resonates on its mounting panel and the magnet resonates on the speaker's frame. An equal force to that subjected to the voice coil is also affecting the magnet. However, because speakers have a very poor efficiency the force applied to the magnet is many times greater than the generated sound pressure. Thus, the baffle becomes subjected to serious vibration. One mode of vibration resonates the baffle along with the heavy mass of the driver unit. An effective solution to reduce this vibration is to increase the stiffness of the baffle - usually by setting the baffle width only slightly wider than the diameter of the driver. This solution also improves radiation characteristics. A more difficult mode of magnet vibration flexes the speaker's frame. About the only way to suppress this vibration is to brace the magnet tightly to the rear panel.

11.5 Cabinet shape: Directivity, diffraction, beam blocking etc.

Cabinet shapes usually utilize the dispersion effect to some extent: Slanted and tilted cabinets radiate more directly toward guitarist's ears (result of this is high frequency "emphasize") as well as decrease the effect of bass emphasize that is caused by lower frequencies reflecting back from the floor. Both characteristics divide people to prefer either slanted or straight cabinets. Some players use slanted cabinets turned on their sides for horizontally wider dispersion. At the "experimental" era of guitar amplifiers (from the 1940's to the 1960's) some cabinets were actually even designed to have such an unconventional form. These designs are now making a comeback in some acoustic amplifiers.

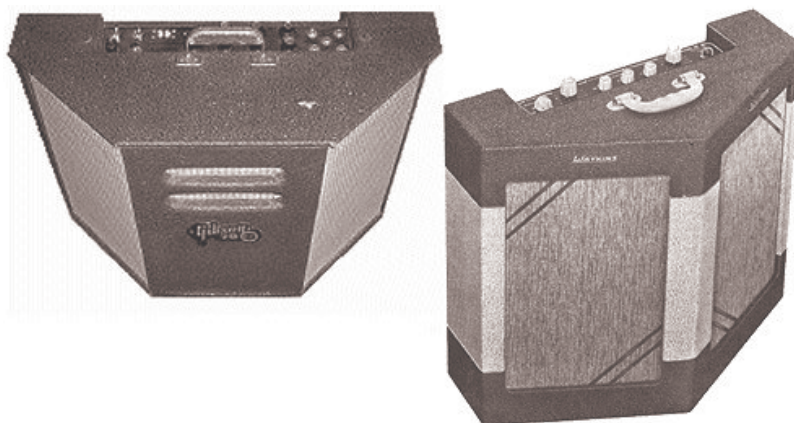


Figure 11.6 Slanted cabinets of Gibson GA-79RV Stereo Reverb and Watkins Dominator from 60's. The 45° horizontal slant accentuates stereo separation but the lack of frontal dispersion creates a "dead zone" directly in front of the cabinet.

"Beam blockers" are used to deflect the high frequency dispersion that exhibits the most beaming effect. A beam blocker is an element that disperses high frequency signal material to alternative directions making it sound less harsh. The wider dispersion will also generalize the speaker system's tone regardless of the position the player stands in relation to speaker system. Since the high frequency sounds do not reach one's ear directly but through wall reflections and with a prominent time delay they are perceived quite differently. The beam blocker's shape and material (mainly

the question is it acoustically absorbent or not) define the effect since the attributes affect dispersion directions and amount of attenuation.

For example, dome-shaped beam blocker (with its “top” facing the speaker) can provide a very wide dispersion. In one test it was shown that beam blockers do not really affect the sound in 90° and 180° angles. That is, they do not “block” the sound. When measured few meters away in 45° angle it was shown that a beam blocker however, decreased the amount of high frequency attenuation (above 2 kHz) as much as 10 decibels! One must note that the effect is equal to doubling the perceived loudness (and making a tenfold increase in output power). Today, the need for beam blockers is as valid as it has always been but strangely the manufacturers tend to ignore them in cabinet the design process leaving room for third party companies to sell you beam blocker products with a *very* considerably price. If you look at baffle board designs from “vintage” era (the 1920’s to the early 1970’s) you can see that beam blockers were often an essential part of the baffle board structure. They also protected the speaker.

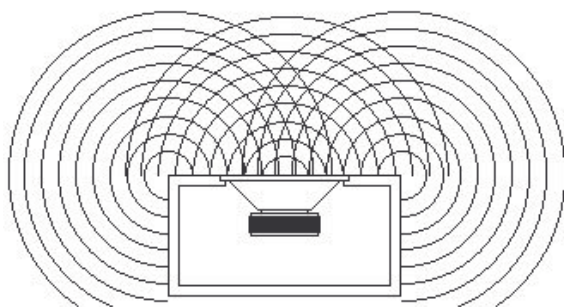


Figure 11.7 Diffraction: Edges create new radiation sources.

Diffraction happens when waves bend around obstacles, or spread out of openings that are smaller than the wavelength. Edges also create new sound sources, which may lead to very odd behavior of attenuating or emphasizing certain frequencies. In practice, a single speaker mounted to a cabinet cannot create a perfect spherical/conical pattern either. Rounding cabinet edges only reduces diffraction, it doesn’t completely eliminate it and usually the rounded sections need to have a fairly large radius (usually several inches) to be efficient enough. A better solution is to apply sound absorbing material between the driver and the relevant edges. This method was already covered in US Patent 4,167,985, issued in 1979.

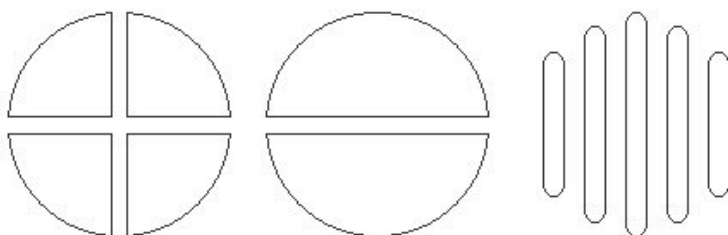


Figure 11.8 Baffle board’s speaker hole cut patterns from “vintage” era also served as beam blockers. The most usual brand to use the shape (from left to right): Vox, Ampeg, Gibson.

It is common knowledge that when two speakers are used side by side they will reinforce each other by summing the sound waves and therefore generate a louder sound (assuming that they are in correct phase of course). However, it is less well known that this arrangement also causes phase cancellation for sound waves that have a wavelength that is twofold the distance between the drivers. For example, if there is a distance of 14-inches from center of one driver to another (wavelength for 968 Hz in

normal conditions), the configuration will cause a dip in frequency response around 484 Hz (the distance between sources is half the wavelength of this signal). A half- or full stack configurations will have a very odd frequency response as they exhibit equal interference between all the drivers. Naturally, these effects take place in a three-dimensional space creating a very complex dispersion pattern.

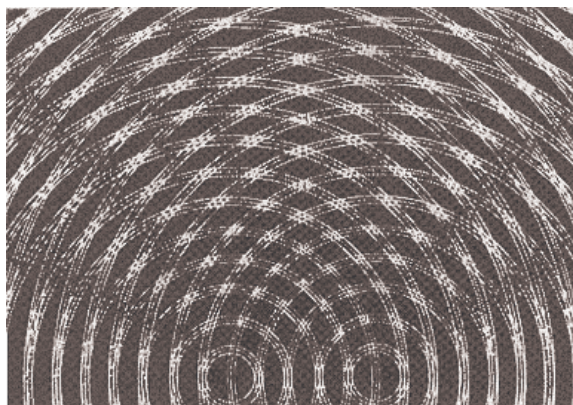


Figure 11.9 Two-dimensional plot showing the interference of two sound sources. Dark lines show cancellation.

In air and in normal conditions sound has a velocity v of 335 meters (or 1100 feet) in second. Thus one can calculate the relation between frequency and the wavelength λ in either meters or feet using the following formula:

$$\lambda = v/f \quad \text{(Equation 91)}$$

Quite illogically, horizontally paired speakers reinforce each other's waves in *vertical* direction while vertically paired speakers (like PA columns) reinforce each other's waves *horizontally*. The effects of both configurations begin to look more logical when dispersion patterns are drawn, as is done in Figure 11.10.

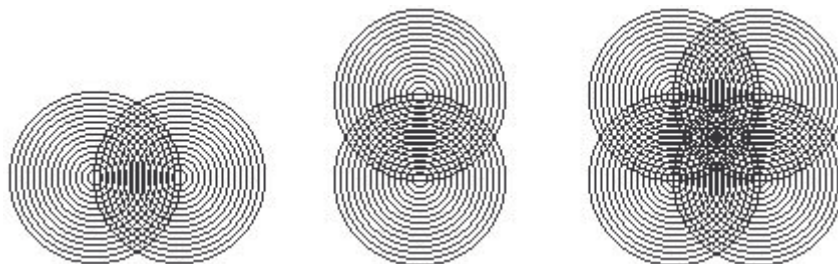


Figure 11.10 Dispersion patterns and interference. From left to right: Vertically paired sound sources, horizontally paired sound sources and a combination of both.

11.6 Open-backed cabinets

It would be more accurate to call open back cabinets as *folded baffle* cabinets. Since they have a “figure-of-eight radiation pattern” – meaning they radiate sound to both front and rear - they are essentially *dipoles*; this means that their bass response is often not as good as that of closed back cabinet structures. Due to structure, the sound is also highly dominated by the cabinet placement: Near walls, the waves radiating from the rear are highly emphasized and the cabinet exhibits a totally different response.

Open back cabinets do not really have dimensioning issues, except few: As the structure is basically *just a folded baffle board* it must be large enough to effectively

delay the rear sound wave. A rule of thumb in HiFi speaker design is that a simple baffle board only (without a “box” structure) would have to have a width that is at least one quarter of wavelength of the lowest frequency to be reproduced [36]. This is because a half wavelength delay on the rear wave reaching the front sets the phase equal with the front wave. As the low frequencies have a moderately long wavelength, folding the baffle board is the way to make the system more compact. Note that when the baffle board is folded its width starts to have little importance as the distance the waves travel equals to depth of the cabinet.

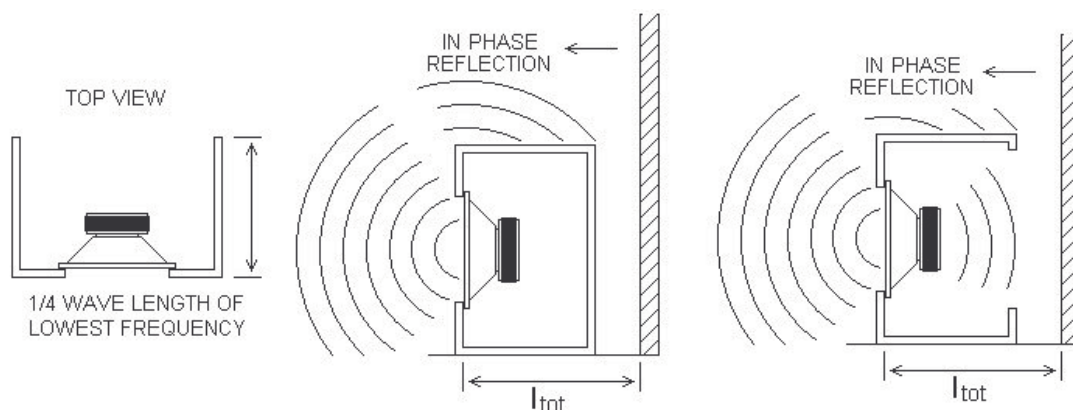


Figure 11.11 Cabinet dimensions and distance from obstacles can have a great effect on frequency response – especially with open back designs.

An open backed cabinet also has a resonant frequency, which, as expected, is equal to frequency where the depth of the cabinet approaches quarter wavelength. At this frequency the structure forms a resonant tube and more power is radiated from the rear side than at other frequencies [41]. As the rear wave may combine in phase with the front wave, the resonant frequency can create a notable emphasize. In HiFi design any change towards a non-linear response is considered bad but in guitar amplification it’s just another design parameter to consider. Generally, *most* open back cabinets seen in guitar amplifiers are too small as the wavelength of 80 Hz is as much as 169.5 inches – that equals 4.3 meters! Yes, the minimum cabinet depth providing a HiFi response would be about one meter. Usually when open back cabinets are built they are made “compact”. However, this is also the absolutely wrong way to design them: An open back cabinet should be *large*. This is one of the reasons explaining the poor and “boxy” sound of small and cheap combo amplifiers; the other reason is the poor speakers they tend to have.

11.7 Sealed cabinets

Sealed cabinet structures can be divided into infinite baffle or acoustic suspension configurations. According to Vance Dickason, the author of Loudspeaker Design Cookbook, if compliance of air (V_{AS}) divided by air in the enclosure (V_B) is greater than 3 the design can be classified as acoustic suspension. For infinite baffle designs the driver Q_{TS} should be greater than 0.3, smaller values are good for acoustic suspension design.

Infinite baffle configuration usually requires a very large enclosure because the internal pressure should not be allowed to affect the backward motion of the speaker cone. Enclosure is usually filled with foam, fiberglass or other similar acoustic stuffing material in order to artificially increase the volume. The acoustic wadding also attenuates internal reflections and standing waves. In practice, the driver’s mass

and compliance determine resonant frequency and damping, which affects the overall frequency response since the response falls off quite quickly below the resonant frequency. The larger the resonant peak, the lower frequencies the speaker can reproduce. However, the same variable also over-emphasizes the resonant peak.

Acoustic suspension limits the cabinet space, thus the air trapped inside the cabinet acts as a spring that works against the cone movement. Cabinets using this design can be quite small and they typically have a smooth low frequency response that decreases gradually. They also have less distortion than infinite baffle designs, especially at low frequencies and high sound pressure levels. However, the system is rather *inefficient* and requires greater cone excursion to produce same SPL than most other enclosure types.

Typical cabinet dimensions for sealed box design follow the “golden ratio”: 0.62:1:1.62. These are naturally *interior* dimensions. Boxes following the golden rule generate less internal standing waves and therefore they have smaller resonance peaks than cubical boxes with symmetric dimensions. Another common dimensioning pattern is 0.8:1:1.25. Scaling the dimensions of infinite baffle design by the factor of 0.88 usually provides proper dimensions for acoustic suspension design. Note that the golden rule applies to dimensions of bass reflex designs as well.

A word of warning about sealed cabinet combo design: An effective cooling of a device requires uninterrupted circulation of air. A cool air must have a path to flow to replace the warmer air and the warmer air must have a path away. Even a large heat sink coupled with a fan becomes ineffective if both are stuffed inside a sealed space where ambient temperatures will grow very high. As a rule of thumb: *Never* place the device chassis inside a sealed space. As we also have deduced, the speaker is a very inefficient device: Typically a 100-watt speaker is essentially just a 99-watt heater inside your speaker box – although it only spends very small fractions of time working at the full power. Nevertheless, your speaker is another considerable heat source that you preferably wish to keep as cool as possible.

Another issue is a “sealed” design where the chassis is partially inside the “sealed” space. These can create very tricky problems since air has a habit of finding all possible leakage paths and will most certainly leak away through the chassis ventilation holes or the very small openings at the sides of the chassis. A sealed cabinet should indeed be literally sealed (usually with a silicone rubber or similar substance) and such leakage holes should not exist. If there are holes they decrease the overall efficiency or in the worst case cause very nasty side effects that usually manifest themselves as whistling sounds that are generated by a similar fashion as blowing air inside a bottle. Also, do note that in some cases the *speaker’s surround can be a major source of air leak as well*. Such is especially the case with speakers that have treated paper or cloth accordion surrounds. These types of speakers are generally good only for open back designs.

11.8 Ducted/vented cabinets

The ducted cabinet structure – also known as bass reflex – was invented back in the 1930’s. Back then the invention was called “acoustical phase inverter”, but since it was patented technology it didn’t have an immediate success. There are two common ways to construct the reflex hole: Ported cabinets literally have a simple hole; vented cabinets use a duct. Unfortunately, these terms are usually mixed together and in

practice it is safe to assume that if a cabinet is said to be ported it is actually ducted. A rectangular duct is also commonly called a slotted port. Note that a simple port is still essentially *a duct* with a length equal to cabinet wall thickness so semantics really do not matter. The principle of the system's operation is fairly simple: A specific amount of air is entrapped in the port opening and the rear wave of the speaker pushes against this air setting motion to a wave that is in phase with the front wave [36]. The area and depth of the hole defines the amount of pushed air, thus ducting is an efficient way to increase the air volume. The port placement does not affect the performance in any way – *unless the airflow between driver and vent is obstructed*. It is nevertheless good practice to locate the vent opening at least two times its diameter away from the driver. In ducted cabinet designs the driver Q_{TS} should be between 0.2 – 0.5. Also, Efficiency Bandwidth Product ($EBP = F_S / Q_{ES}$) of the speaker should be about 90 or greater. EBP values lower than 50 mean that the concerned driver will likely perform better in a sealed enclosure.

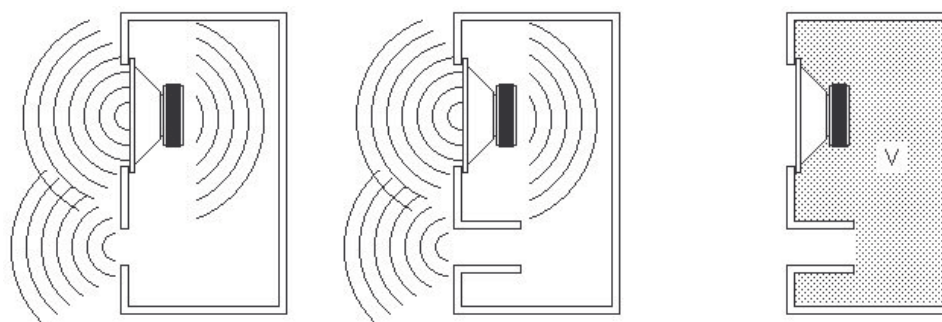


Figure 11.12 The principle of acoustical phase inverter's operation. Left: Ported cabinet and middle: Vented cabinet. The rightmost picture shows the effective air volume V_B in grey.

A vented cabinet can be 3 dB more efficient than a sealed cabinet of an equal size and its bass response can be extended an octave lower. There are also some major drawbacks: Vented cabinet is more sensitive to parameter variations and usually works only with one type of speaker. The transient response is reduced since a vented box falls 24 dB per octave instead of -12 dB. The impedance curve and excursion response are more complex and there is more phase shift and group delay. Unless a specific port structure (like 90° bend) or insulation of the cabinet is used, certain mid-range frequencies leak from the port in opposite phase to driver's radiation; this, in turn, causes attenuation. Ports that are too small in diameter may also begin to whistle when large amounts of air flow through them. Another common form of "port noise" is usually introduced if the port opening is not rounded.

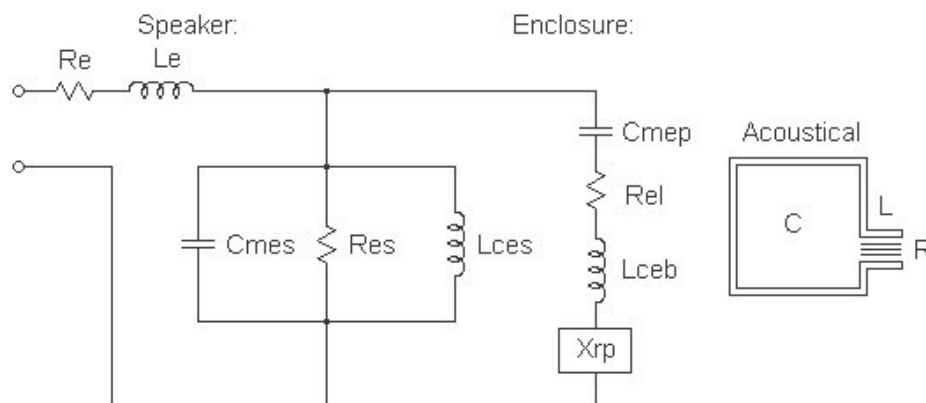


Figure 11.13 Mobility-type lumped model of a speaker in vented enclosure.

A vented enclosure modifies the lumped model of speaker with an addition of another parallel branch consisting of a series connection of capacitive equivalent C_{MEP} for acoustic port mass (M_{AP}), resistive equivalent R_{EL} for enclosure leakage resistance (R_{AL}), inductive equivalent L_{CEB} for enclosure compliance (C_{AB}) and port radiation impedance X_{RP} . The parameter X_{RP} is so insignificant that it can be left out of further equations. The electrical circuit elements are connected in series. This circuit is inversely analogous to *acoustic model* of a vented box as the conversion usually goes as follows: Volume of compressed air or "enclosure" (acoustic compliance) represents capacitance, duct filled with gas (acoustic mass) represents inductance and losses (acoustic resistance or acoustic responsiveness) represent resistance. The resistance is usually depicted with a symbol of fine-mesh screen. The component values are calculated using formulas:

$$L_{CEB} = \left((Bl)^2 \times V_B \right) / \left(\rho_0 \times c^2 \times S_D^2 \right) \quad \text{(Equation 92)}$$

$$C_{MEP} = \left(\rho_0 \times L_V \times S_D^2 \right) / \left((Bl)^2 \times S_V \right) \quad \text{(Equation 93)}$$

$$R_{EL} = \left((Bl)^2 \times V_B \times R_V \right) / \left(\rho_0^2 \times c^2 \times L_V \times S_D^2 \times S_V \right) \quad \text{(Equation 94)}$$

Where V_B is the enclosure's acoustic volume in cubic meters (V_B = enclosure's physical volume if no acoustic stuffing is used), S_D = effective cone area in square meters (see datasheet or measure), S_V = vent/port area in square meters, L_V = vent/port length in meters and R_V = all enclosure losses. Air density ρ_0 is approximately 1.205 kg/m³ and velocity of sound $c \approx 344.4$ m/s @ 25° C. Abbreviation V_B is basically used to indicate internal volume of a system with a vent or passive radiator. V_C is used when a sealed enclosure is in question. If you do not have the parameters to calculate R_{EL} assume that it is very small. The components form a series resonance circuit that mimics the operation of a Helmholtz resonator.

Today one has access to vast amount of software that is dedicated for enclosure design. However, it is worthwhile to take a peek at the basics: The old school method of ported cabinet tuning was to measure the impedance of the driver when it was mounted to the cabinet. Compared to the free air impedance curve, the impedance curve of a ducted cabinet has two peaks in its impedance that are a result of interaction between parallel and series resonance circuits. When the cabinet is properly tuned the series resonance shunts the parallel resonance. At 70's it was widely believed that the cabinet is optimally tuned when the two impedance peaks have equal height and amplitude. However, this is pretty incorrect generalization and commonly results into horrible performance.

The design process usually starts from determining a suitable enclosure volume and continues by calculating its optimum tuning frequency f_B . Unfortunately, if you do a search on the topic you will find out that each author presents his or her own theory of fitting equations. This is regrettable but also understandable from the point of view that the concerned equations are only approximations that merely try to *model* reality. Presented here are three variations: The first variation is based on equations presented by Brian Steele at The Subwoofer DIY Page, second variation is based on D.B. Keele's interpretations on Thiele's papers and the third variation is from R. Bullock's article in Speaker Builder 4/80. According to them, the optimal enclosure volume V_B is:

$$1. V_B = 20 \times V_{AS} \times Q_T^{3.3} \quad \text{(Equation 95)}$$

$$2. V_B = 15 \times V_{AS} \times Q_T^{2.82} \quad \text{(Equation 96)}$$

$$3. V_B = V_{AS} \times Q_T^{3.153} / 0.0569 \quad \text{(Equation 97)}$$

You will find out that the equations give pretty similar results, however in bass reflex design there is not much tolerance between parameters. The equations rely on assumption that Q_T and Q_{TS} are interchangeable. If they differ, Q_T is the proper variable to use. Remember that when a box is designed, the volume taken by the speaker, duct, bracing, handles, etc. should be added to the result in order to retain the effective volume inside the enclosure same as V_B . Next step is to calculate the tuning frequency.

$$1. f_B = (V_{AS} / V_B)^{0.31} \times f_S \quad \text{(Equation 98)}$$

$$2. f_B = 0.42 \times Q_T^{-0.9} \times f_S \quad \text{(Equation 99)}$$

$$3. f_B = 0.42 \times f_S / Q_T^{0.953} \quad \text{(Equation 100)}$$

V_{AS} , Q_{TS} (Q_T) and f_S are listed in the driver's datasheet along with other necessary Thiele-Small parameters.

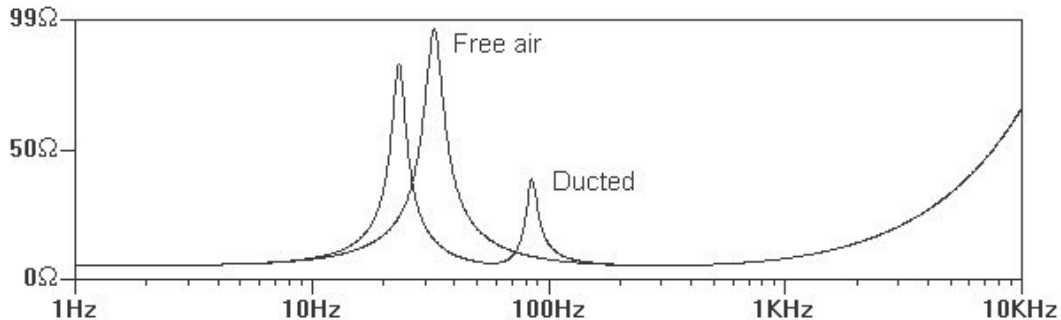


Figure 11.14 Impedance plots of ducted and open-backed cabinets acquired using a SPICE simulation.

The tuning frequency determines the required duct area and length. Before calculating the duct parameters it is worthwhile to check out that the configuration is providing adequate bandwidth. In order to do this one must find out -3dB point of the system – in Thiele-Small terms this is called $F3$ and after this point the response will roll off – 12 dB per octave.

$$1. F3 = (V_{AS} / V_B)^{0.44} \times f_S \quad \text{(Equation 101)}$$

$$2. F3 = 0.26 \times Q_T^{-1.4} \times f_S \quad \text{(Equation 102)}$$

$$3. F3 = 0.305 \times f_S / Q_T^{1.33} \quad \text{(Equation 103)}$$

Port dimension equations one will bump into are not identical either, as the port tuning length equation

$$L_v = \left[\frac{c^2}{(2\pi)^2} \right] \times \left[\frac{(S_v \times n_p)}{(f_B^2 \times V_B)} \right] - k 2\sqrt{(S_v/\pi)} \quad \text{(Equation 104)}$$

consists of

- Numerical constants (e.g. certain speed of sound, value of pi, etc.)
- Physical properties (enclosure and port dimensions, tuning frequency and number of the ports)
- The end correction factor k, which is defined by added length factor of effective port radius for generalized port shape and subtracted from the acoustic length of port [42].

In equation, c = speed of sound in air (m/s), S_v = inside area of the port (m^2), n_p = number of ports and k = end correction factor. The “standard” end corrections are 0.425 for flanged-end ports and 0.307 for free-end ports. The end corrections are summed together so the initial assumption is one flanged end and one free end (that is inside the enclosure) resulting into a figure of 0.732. Most differences between diverse equations you will see are result of using different constants for c and different procedures to calculate k for multiple ports. Note that the presented equation uses standard SI-units and has no simplified constants or unit conversions. This cannot be said of most other equations. Likely the best procedure is to type the equation into Excel (or similar program) and let the computer to do the time consuming calculations. This way you can vary the parameters and automatically get the result for L_v . The same spreadsheet can also calculate component values for the lumped circuit model.

As a general rule it is beneficial to make a longer duct than indicated by the math and then fine tune by *shortening* it to the right length. Logically, this cannot be done in opposite direction – or at least it would be extremely difficult. The primary concern is maintaining the sound velocity at the port relatively low so that port turbulence generates no noise. It has been experimentally determined that a correct velocity is about 5 % of speed of sound (0.05 Mach) [39]. Port whistling will commonly occur at Mach values greater than 0.16. The velocity of sound inside the port v can be calculated using the formula

$$v = p \times 4\pi \times r / (S_v \times \rho_o \times \omega) \quad \text{(Equation 105)}$$

where, v = velocity in m/s, p = sound pressure *in pascals*, r = distance from pressure’s measuring point, S_v = port area, ρ_o = air density $\approx 1.205 \text{ kg/m}^3$ @ 20° C and $\omega = 2\pi f$. The formula can be simplified even more since most measures are taken 1 meter away from the source (r = 1m). Effectively, the equation becomes:

$$v = p \times \pi / (S_v \times \rho_o \times \omega) \quad \text{(Equation 106)}$$

To turn the result into a Mach number we simply calculate its relation with speed of sound:

$$M = v/c \quad \text{(Equation 107)}$$

Note that equations do not take into account the use of multiple ports. The exact value has little importance since in practice the variable is varying according to atmosphere conditions. You may wish to plot the equations in Excel, as they are frequency dependent and greatly affected by the given SPL: More output power means higher port velocities. Naturally, you want to calculate the worst-case scenario. To convert ordinary SPL (dB) unit to Pascal use the following equation:

$$P_{(Pa)} = 20 \times 10^{-6} Pa \times 10^{SPL(dB)/20} \quad \text{(Equation 108)}$$

When the cabinet is tuned correctly it exhibits an extended low frequency response that should be quite linear. It is common that the response exhibits a small “saddle”.

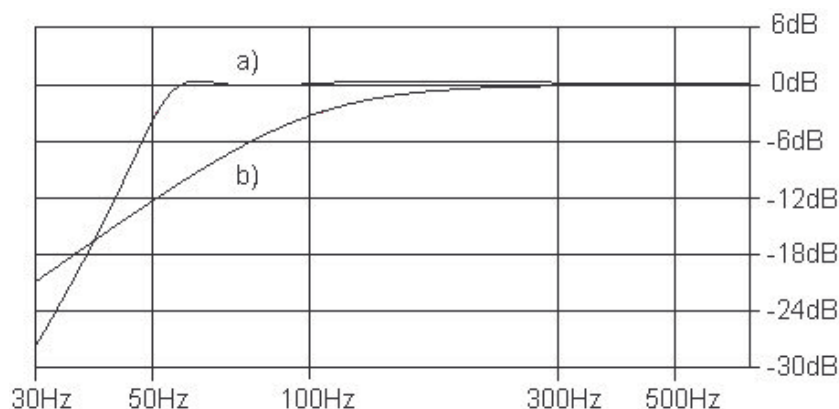


Figure 11.15 Frequency responses of a) ported and b) sealed 70-liter cabinet. Driver is Eminence Legend 875.

This has been a representation of the “pen-and-paper” tuning process. One can get away easier by using dedicated design software such as WinISD, Unified Box Model (UniBox), Speaker Workshop, LspCAD or Basta! These are just few examples, as there is a lot from where to choose. Most programs rely on Thiele-Small principles and theory of using lumped models so attaining a basic knowledge on how these things work is beneficial – especially because *only a small fraction of software actually states what equations they use in the calculations!* This is very regrettable, as the user has no way to evaluate the weaknesses of the software. It is also highly recommended to run a set of SPICE simulations of an amplifier design using a selection of various cabinet and speaker models as its load. This will tell you a lot about the stability of the design. Hopefully this chapter has given some insight on how to model a load like this. Remember though, that these loads are *very* simplified and *real* loads are even more complex.

A special design that is worth mentioning here is the *Speaker Projector Ring*, introduced by Fender in the early 60’s. This construction type is also known as “double baffle” or simply “tone ring”. What is less understood is that a projector ring cabinet is essentially a reflex structure: The internal reflection baffle in together with the front baffle forms a port that has its mouth around the driver. This is clearly visible from the diagram. While this type of structure is disobeying most HiFi design rules it is still quite appreciated in instrument cabinets. Typically there is about 2-inch clearing between the main baffle and the reflection baffle to which the tone ring is mounted. Due to ring, which is typically made out of sturdy metal, the speaker cone is elevated to the level of the front baffle, thus edge diffraction is reduced.

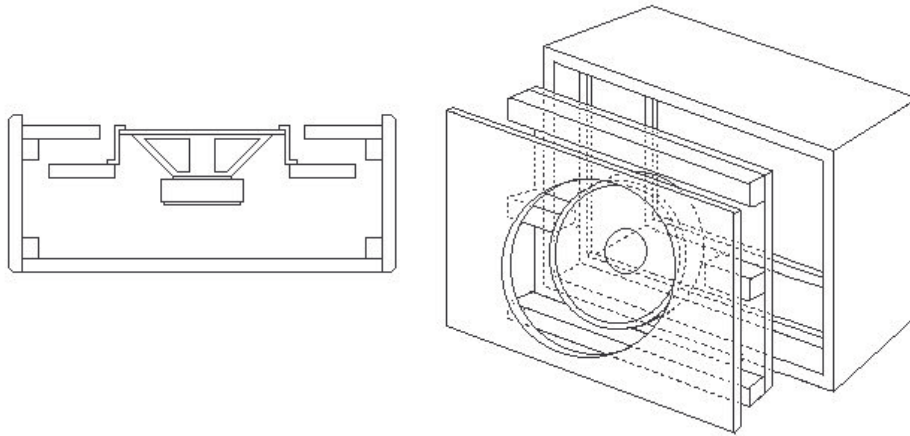


Figure 11.16 Fender's *Speaker Projector Ring* cabinets essentially have a reflex structure.

11.9 Passive radiators

Passive radiator systems are very similar to ported ones except that the port is replaced with a passive radiator cone, also known as “drone cone”. A drone cone looks very much like an ordinary speaker but it does not have a motor built out of a magnet and voice coil; instead the movement of the air inside the sealed enclosure operates it passively. Passive radiators can extend the low end that is obtainable from moderately small enclosures. In practice, passive radiator systems are pretty similar to ducted systems, except that they have a steeper roll off slope below tuning frequency, poorer transient response and a dip in the frequency response at drone cone's resonant frequency. The dip is usually way below the system's pass band so practically it can be ignored.

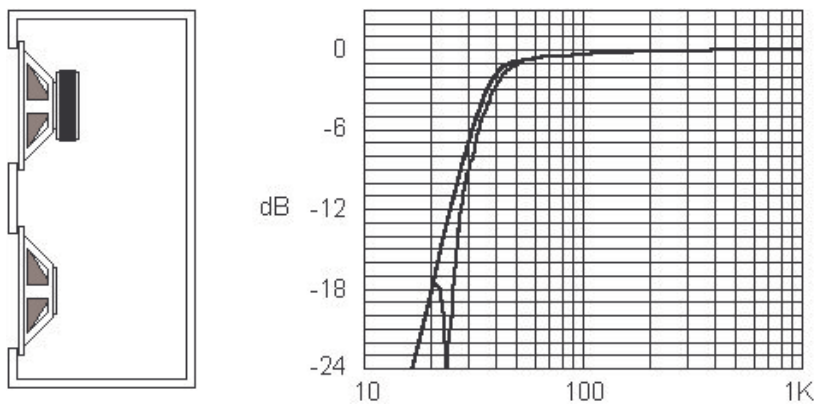


Figure 11.17 Passive radiator replaces the reflex duct. The roll off slope is steeper and a notch is added to the frequency response.

Passive radiators are usually utilized in situations when duct lengths would grow too long or when the port's resonance, turbulence or power compression induced by high airflow has to be eliminated. Weight of the radiator cone should be equal to weight of air that otherwise would fill the duct ($\text{mass} = \text{port volume} \times \text{density of air}$). Also, the passive radiator always has to be larger in diameter than the active driver is as the displacement requirement for the radiator is about 1.5 to 2 times higher. Sometimes few radiators are required to fulfill this rule.

11.10 Sealed enclosures revisited

Since the theory behind lumped circuit modeling of the cabinet has been discussed in this chapter it is worthwhile to revisit the sealed enclosure design and examine its

characteristics from electronics point of view. In very simplified form the port and its losses are omitted from the lumped circuit model leaving the inductive equivalent L_{CEB} for enclosure compliance (C_{AB}) as the sole design parameter. As a result, the circuit no longer forms a second resonator but introduces parallel inductive element that affects the driver's lumped circuit model, thus shifting the resonant frequency.

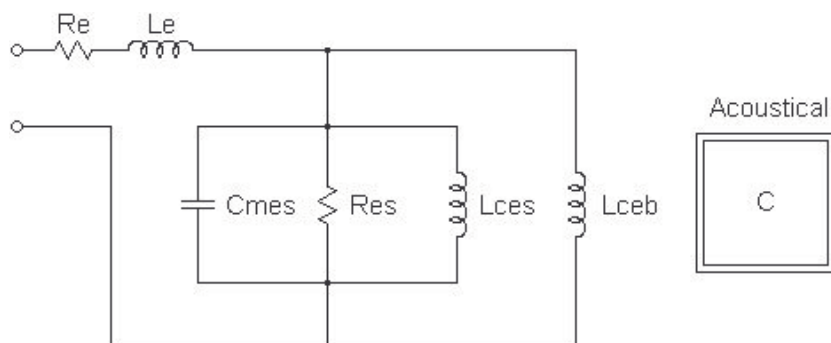


Figure 11.18 Lumped circuit model of a sealed cabinet and driver.

The previously presented equation for calculating L_{CEB} still suffices because the enclosure volume is the only variable that is related to the enclosure itself. Since L_{CEB} is in parallel with L_{CES} the total effective inductance will be smaller than the smallest of the inductances – similarly as in parallel connection of resistors. In practice, one will find out that the smaller the enclosure, the higher the resonant frequency is. To acquire effectively the same resonant frequency that the speaker would have in free air, the enclosure should be as large as possible – the extreme of this is naturally the free air volume of open-backed cabinet structure. In fact, the low-end bandwidth of a sealed cabinet will be worse than that of an open-backed cabinet – however, due to “cancellation” of rear wave the sound is more “focused” and the tone is affected less by cabinet placement.

11.11 Horn structures

Horn structures act as passive acoustic transformers that improve the coupling between driver and air thus resulting into increased overall efficiency. The horn structure has gradually (and usually exponentially) increasing cross sectional area, which is also known as flare. The “rate” of increase is known as *taper* and it usually follows a tractrix function. The smaller end of horn near the driver is called throat while the larger end is called mouth. The small cross-sectional area of the throat presents high impedance for the driver allowing it to develop a high pressure at given displacement. The taper of the flare gradually decompresses the sound hence at the mouth exists a large displacement and low pressure. The slower the flare rate is the lower frequencies the horn can reproduce at a given horn length.

Horn structures can roughly be divided into rear- and front-loaded ones. Front-loaded horns are often pretty basic and usually employ a sealed cabinet structure. These are commonly “bin” cabinets that are generally intended for use in bass amplification. A rear-loaded horn usually employs the driver in dual action: The front of the cone radiates higher frequencies directly while the lower frequencies are radiated with the aid of a horn. Since higher frequencies have a better impedance match with air the sound levels are kept equal. The mouth of the horn can either face forward or rear, as bass frequencies are rather omnidirectional. The most usual enclosure structure of this type is “scoop”.

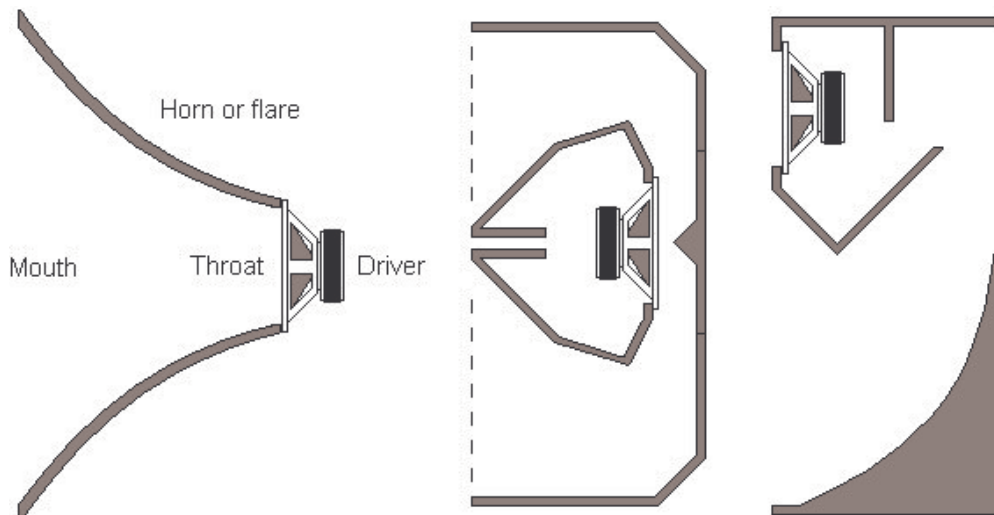


Figure 11.19 Left: Cross-section of a front-loaded horn structure. Enclosure is omitted from the picture. Middle: Front-loaded “W-bin” cabinet. Right: Rear-loaded “scoop” horn structure.

Making the throat area smaller than the actual cone diaphragm area is a technique known as “compression” or “horn loading” and the amount of it, (known as compression ratio), can be deduced by dividing cone area by throat area. The higher the compression the greater the efficiency is, as the ability to couple diaphragm to the air at horn’s mouth increases – at least up to a point where too small throat area begins to impede cone motion. For bass and mid-range frequencies typical compression ratio ranges from 2 to 1. High frequency tweeters can have compression ratios as high as 10.

There are innate problems in horn structures, for instance, the flare length for low frequencies has to be huge and for higher frequencies the driver (in most cases) has to be a special *compression driver* with high pressure and small displacement. As these drivers are very tough to design they are also *extremely expensive*. The length for a typical bass horn can be 2.5 meters or more so these types of structures are usually *folded*. Mouths of bass horns can easily have diameter as large as one meter. In practice, an average horn-loaded bass cabinet is larger than a typical washing machine – and likely weights as much as well. The above reasons limit the use of horn designs mainly to public address, bass and acoustic amplifier rigs. Electric guitars do not really benefit enough from the use of horns to justify the price and design efforts involved, however, if you are designing an acoustic guitar amplifier you might want to consider using a horn loaded high frequency element in it.

11.12 Transmission line enclosures

Dr A.R. Bailey and A.H. Radford described the principles behind transmission line topology in an article of *Wireless World* 1965 (p. 483 - 486). The reasoning was to absorb the loudspeaker’s rear wave as transparently as possible by channeling it into a pipe that is long enough to resist resonance below a desired frequency. Therefore, transmission line structure replaces the cabinet with an acoustical “labyrinth”. The length of this transmission line is quarter wavelength of the resonant frequency of the driver – however, often the length is artificially increased by filling the line with acoustic stuffing material that lowers the speed of sound. The transmission line may be tapered or simply have the same cross-sectional area throughout the whole line. The tapered structure, known as “tapered quarter-wave pipe” or TQWP, is a “reversed horn” that attempts to absorb the rear wave of the driver. The difference between

ported cabinets should be clear: The rear wave is absorbed instead of using it for reinforcement.

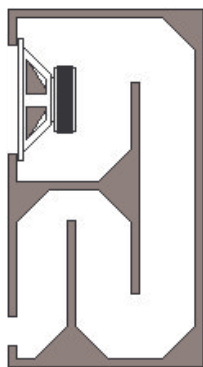


Figure 11.20 Cross-section of a transmission line structure.

Transmission line structure offers similar transient response characteristics as a sealed cabinet but with a shallower roll off (which is about -10 dB per octave). In practice they perform way better than ported cabinets but are much more difficult to design and build. Enclosure sizes are usually very large and although the topology is over thirty years old there are no certified design formulas for it. (Although lately some progress for developing such has been made). Solid construction of a transmission line cabinet definitely requires first-rate carpentering skills.

11.13 Compound or band-pass cabinets

A band-pass structure design is quite an old invention; the first patent for such was filed in 1934 by Andre d'Alton. In a band-pass design the ports radiate the entire output of the driver, which effectively creates an acoustic low-pass filter that limits the high frequency response. In comparison to conventional crossover filtering and ported cabinet structure the band-pass can provide substantial increase in efficiency and bass extension. At narrower bandwidths the efficiency at given bandwidth increases and at wider bandwidths the low-end frequency response in relation to demand of enclosure size is improved, however this happens at the expense of efficiency and transient response.

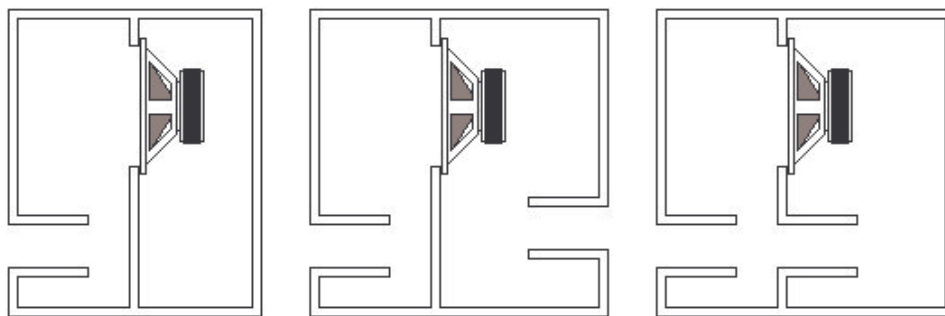


Figure 11.21 Ported enclosure structures. Left: 4th order single-reflex, middle: 6th order dual-reflex (“coupled cavity”) and right: 6th order series-tuned dual-reflex.

Band-pass enclosures are divided into two categories: Single-reflex and dual-reflex. In single-reflex structure one of the chambers is sealed and the other one is ported. Dual-reflex design ports both chambers. There is also a variation called “series-tuned” where a port connects the two chambers. Naturally the port elements can be substituted with passive radiators. Differences between single- and dual-reflex designs are similar to differences between sealed and ported enclosures: A dual-reflex has better efficiency but inferior transient response and steeper roll-off rate. Dual-reflex also has to be larger in size.

Even though a band-pass enclosure is an acoustic filter it is not a very effective one and therefore should also be accompanied by electronic crossover; without it the “out-of-band noise” can become annoying.

11.14 Isobaric loading

Isobaric is a name for loading method derived from Greek term “iso baric”, which means constant pressure. When two drivers are mounted one behind another, so that there is a chamber with constant pressure in between them and so that their cones are moving in equal mechanical phase, they couple together to work as one unit. Harry Olson introduced the technique in 1950 and as a result of it the demand for enclosure size is *theoretically* halved. The cost of this (aside the cost of an additional driver) is a -3dB loss in efficiency. Another advantages, besides the decreased space requirement, are increased linearity, as the drivers correct their nonlinearities, and increased power handling. In practice, the enclosure size is not technically halved, as the additional driver and the coupling enclosure require a significant amount of space.

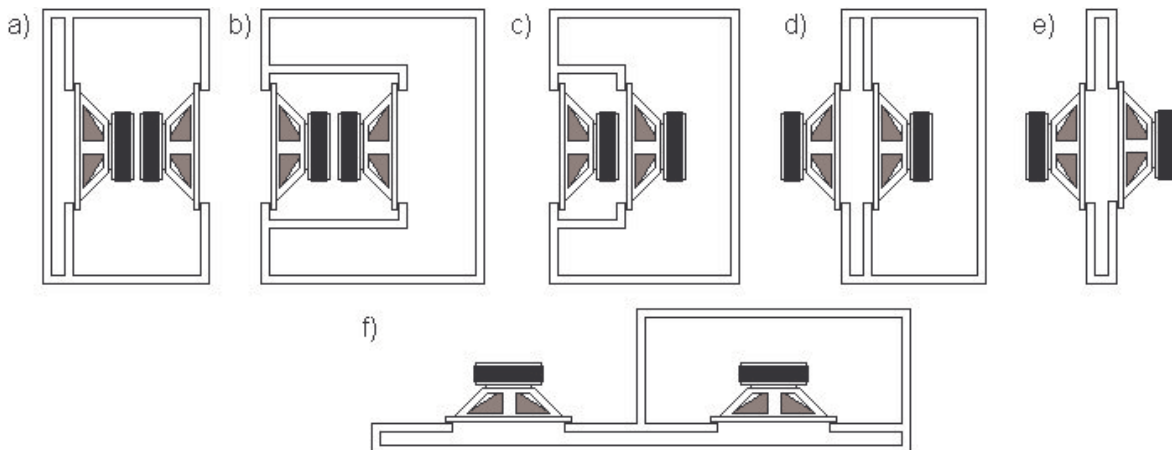


Figure 11.22 Isobaric loading configurations: a) Magnet-to-magnet, b) magnet-to-magnet tunnel load, c) magnet-to-cone or “piggyback” tunnel load, d) cone-to-cone or “clamshell”, e) clamshell dipole and f) cone-to-cone planar load.

Figure 11.22 depicts the most typical isobaric configurations. Note that as the isobaric configuration is merely a way to load the driver it can be applied into any enclosure design topology – thus the isobaric configuration can be used in reflex enclosures as well as in the sealed ones. The image also reveals the significant demand for space in tunnel loading configurations – which also is a fine reason to avoid them altogether. The tunnel-loaded configurations also weaken voice coil’s ability to cool down and the introduced “thermal distortion” is the smallest problem caused by this. Normally, when “clamshell” configurations are used you don’t see the acoustic space shown in the cross sections, as it is common to mount the speakers almost directly to each other. Typically just a small spacer is put in between to prevent cone rub. Because theoretically this space still exists it is shown for reference.

11.15 Isolation cabinets

Isolation cabinets are used when there is a need to capture a cranked guitar amplifier tone at reduced listening volumes. Their use is not mainly reserved solely to tube amplifiers, as it is often a case that a speaker system tends to sound better when it works on it’s extremes – which usually means that notable sound pressures are generated. In practice, an isolation cabinet is ideally a soundproofed, nearly anechoic and vibration-free enclosure that houses one or several microphone stands. The cabinet can either have a built-in speaker/cabinet system or it can be an isolation *box* that contains enough room to house an external cabinet. One step up from this is an isolation *booth* – or in extreme a complete room.

The frequency response of an isolation cabinet is naturally dependant on speaker and cabinet configuration, microphone type and its placement as well as on damping of the “recording” side of the cabinet. In fact, the latter is likely the single most important factor in the design of concerned cabinets, as the speaker system has to project into a very limited acoustic space. In practice, the enclosure configurations are nearly always very far from being ideal. Usually, in order to obtain enough control, the signal from the microphone has to be directed to a multi-band graphic equalizer anyway. Often a simple, single layer isolation cabinet does not provide a sufficient degree of external sound reduction either, as bass frequencies tend to leak out through the cabinet walls. For improved isolation, it is common to construct the cabinet walls out of several layers that have dead space in between. The structure is similar to “floating” live rooms of professional studios.

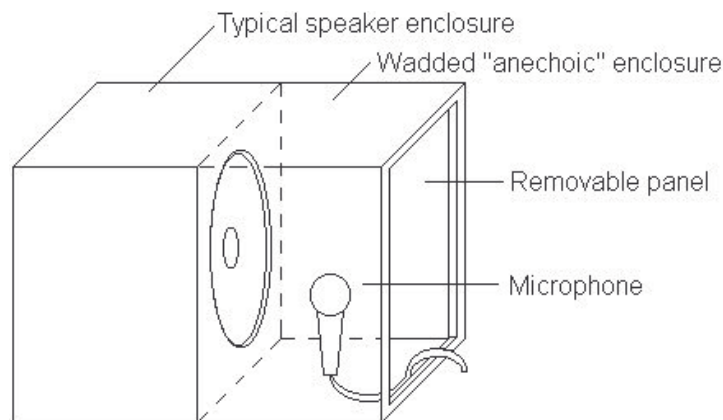
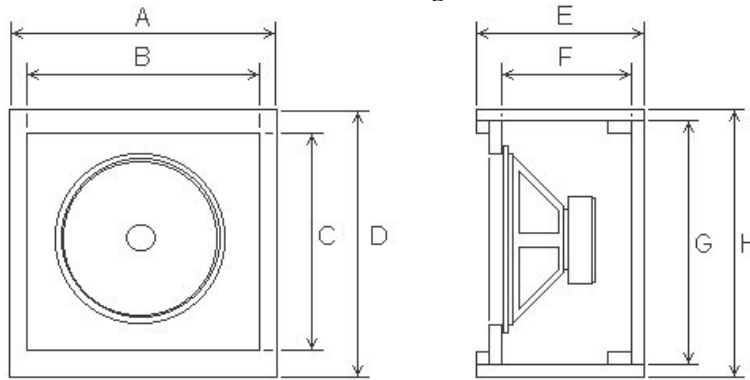


Figure 11.23 Isolation cabinet.

11.16 Examples of cabinet dimensions

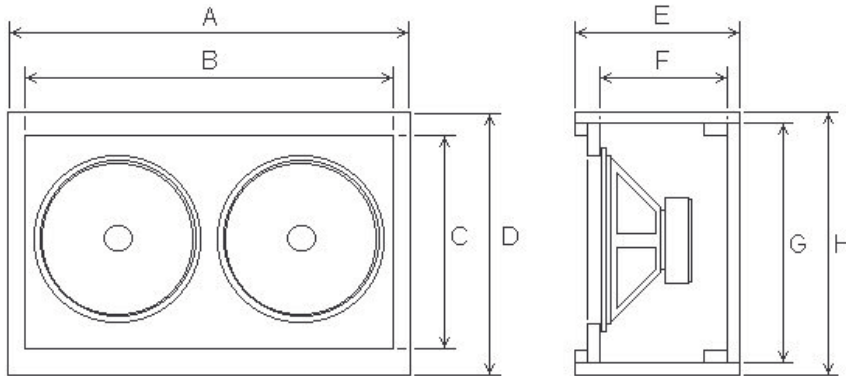
At closing it is worthwhile to present some examples of real-life cabinet dimensions. Note that the information is for some parts lacking – however, the missing pieces are easy to deduct since guitar speaker cabinet building is not rocket science. Most replica cabinet structures work perfectly well - even when they contain some errors in exact dimensions. Variable CT in the tables stands for “cabinet thickness”, which is the overall thickness of side, bottom and top panels. In some cases the rear panel is thinner than this to save weight – this has not been mentioned separately, as is not the thickness of baffle board either. All units are millimeters. Note that the dimensions are provided for *example*, not for exact guide of copying a cabinet design completely. The latter would require a complete article and set of drawings describing how the baffle board and grille cloth are set up and how the cabinet is braced internally.

Dimensions for cabinets that use a single driver.



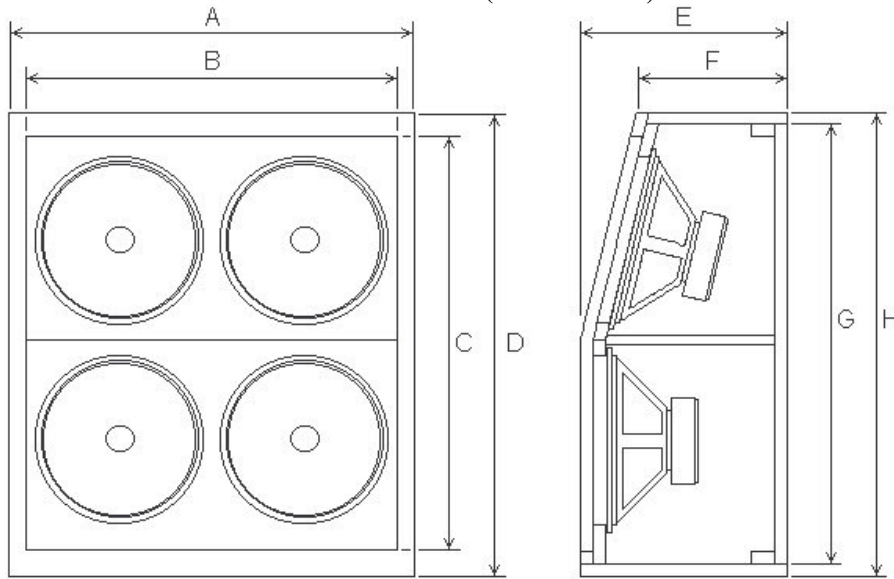
Cabinet/Amplifier	CT	A	B	C	D	E	F	G	H
Fender Bassman 1 x 15" (closed back)	18	806	-	-	527	292	-	491	527
Rivera R-30-112 1 x 12" (open back)	18	530	-	-	457	247	-	421	457
Marshall 1 x 12" (closed back)	16	500	428	398	470	290	238	438	470
Fender Champ 5C11 x 6" (open back)	-	305	264	162	279	181	-	-	279

Dimensions for cabinets with two drivers.



Cabinet/Amplifier	CT	A	B	C	D	E	F	G	H
Marshall 2 x 12" (closed back)	16	760	688	-	520	290	238	488	520
Marshall 2 x 12" (closed back)	16	760	688	528	600	310	258	568	600
Marshall 1922 2 x 12" (closed back)	18	670	-	-	520	250	-	484	520
Marshall 1936 2 x 12" (closed back)	18	750	-	-	600	310	-	564	600
Fender Bandmaster 2 x 12" (closed back)	18	806	-	-	527	292	-	491	527
Rivera R-100-212 2 x 12"	18	673	-	-	457	247	-	421	457

Dimensions for cabinets with four drivers (“Half stacks”).



Cabinet/Amplifier	CT	A	B	C	D	E	F	G	H
Marshall 4 x 12" (closed back)	16	760	688	688	760	360	260	728	760
Vox Beatle 4 x 12" (closed back)	18	686	-	-	1016	292	292	980	1016
Hagström 4 x 12" (closed back)	18	920	-	-	690	280	280	654	690
Fender Bassman '59 4x10" (open back)	18	568	-	-	597	267	267	561	597
Hagström 4 x 8" (closed back)	18	900	-	-	300	145	145	264	300

11.17 Typical speaker wiring configurations

Small but a quite important aspect of cabinet design is the speaker wiring. With cabinets using a single or dual speaker configurations the alternatives for wiring configurations are scarce but when the number of speakers is increased the amount of choices increases rapidly. At first, it is worthwhile to discuss configurations that use an odd number of speakers: These are hardly ever used due to a simple reason; they create an odd load impedance and often result to uneven power sharing. About the only usable configuration is three 16-ohm speakers wired in parallel, which creates a load of 5.33-ohms. This is still passable for belonging into “*nominally 4-ohms*”-category.

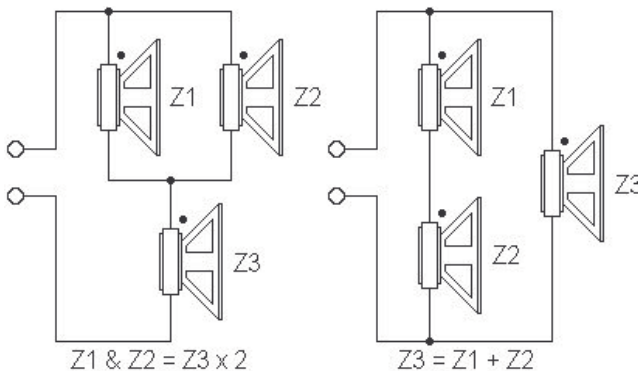


Figure 11.24 Using an odd number of speakers often results into uneven power sharing.

Adequate nominal impedances can be acquired by connecting two of the three speakers either in series or in parallel as shown by Figure 11.24. However, this produces a problem of uneven power sharing: One of the speakers handles half of the total output power while the other two handle only a quarter. Consequently, this

means that in order to maintain balance in loudness levels the lower power speakers must have a higher efficiency. In essence, matching the efficiencies is a good rule even in general and for that reason it is more recommendable to stick with configurations that use an even number of speakers – each also having the same nominal impedance.

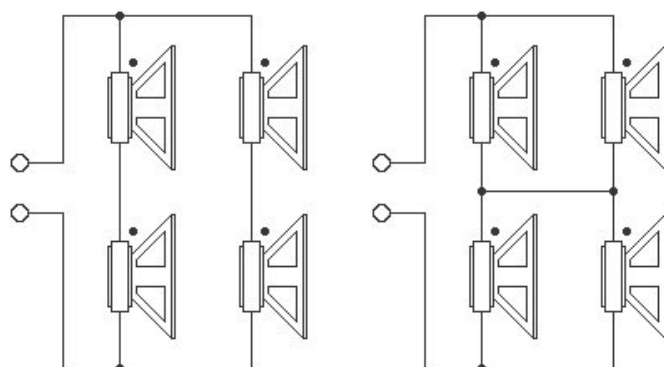


Figure 11.25 Configurations for four-speaker setups. Left: Series/parallel and right: Parallel/series.

Since majority of amplifiers is configured for either 8 or 4-ohm load impedance, systems with four speakers generally have only two realistic wiring options: Series-parallel or parallel-series. The resulting load impedance from both configurations is that of an individual driver (assuming all drivers have the same nominal impedance of course). Few years ago there was a lot of debate whether the two configurations have any difference. Strong arguments were presented both for and against. The truth is that differences indeed exist: When speakers are wired in series their (high) impedance at resonant frequency is likewise summed. On the other hand, a parallel configuration evens the impedance between the two speakers. Consequently, *if speakers used in series-parallel and parallel-series configurations are identical there is no difference in the load impedance of the two systems. However, when speakers are not identical the series-parallel configuration will make the differences more pronounced.* Impedance-wise, the greatest dissimilarities between speakers are different resonant frequencies and different heights of the resonant peak (in ohms). Thus differences between wiring configurations mainly show up at the resonant frequency.

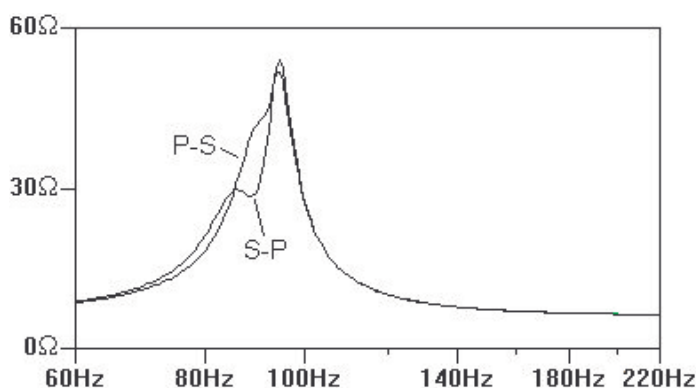


Figure 11.26 When non-identical speakers are used the resonant behavior of series-parallel and parallel-series configurations is different.

Attached plot (refer to Figure 11.26) shows an example of the differences. The graphs are from series-parallel and parallel-series systems that use only three identical speakers. Series-parallel configuration shows a clearer pronunciation of two individual resonant frequencies, it also shows a sharper “dominant” resonance peak, which also has higher impedance. In parallel-series configuration the combination of two

resonant frequencies is less clear and the peak impedance is lower. However, *it is highly questionable how audible these differences are*: In conjunction with a realistic *current sensing* (mixed mode feedback) amplifier, the aforementioned changes in impedance affect the frequency response in the order of *one decibel* (or less) to one way or the other – and this only at a very narrow band of frequencies. Systems that do not employ current sensing are naturally rather immune to these changes anyway. Hopefully, this insight has shed some light into this seemingly everlasting debate.

Cabinets can benefit a lot from some simple switching arrangements. The most common ones of these are shown in the Figure 11.27. Leftmost configuration depicts a series-parallel switch (along with an additional paralleling “daisy-chain” jack). For “automatic” toggling, the switch can naturally be integrated to a suitable input jack in the cabinet but this approach always has its drawbacks. A discrete switch usually offers a more “visual” indication of its state and, more importantly, often has a higher current handling capacity than the integrated jack switches. Since the speaker current can be several amperes the latter parameter should not be ignored. In this configuration, as well as in most of the following ones, it is extremely important to use jacks that have an insulated sleeve.

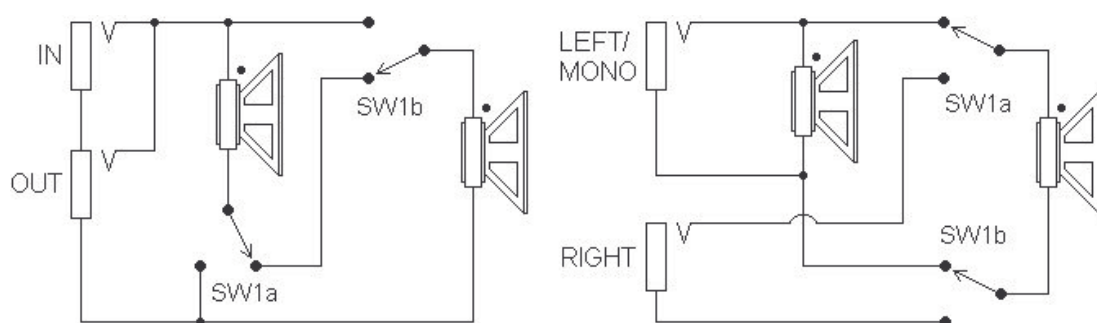


Figure 11.27 Special configurations. Left: Series-parallel selector for cabinet impedance. Right: Stereo-parallel mono selector.

Rightmost setup has two jacks as well, however, this time the other one serves as an additional channel input for stereo configurations. When the stereo mode is disabled the two speakers are wired in parallel and the input jack for the right channel is lifted off the circuit. Enabling the stereo mode doubles the nominal impedance but requires each speaker to be powered from its individual input jack. An integrated DPDT jack switch (fitted to “right” channel jack) can be used to automatically toggle the stereo operation.

As a final point, let’s take a look at some handy output arrangements for a combo amplifier. These are depicted in Figure 11.28. Generally, when versatility is a requirement combo amplifiers tend to call for a lot more demanding wiring. A reasonable expectation is that an external speaker cabinet can be used in conjunction with the combo’s internal speaker; however, another common expectation is that the internal speaker can be disabled and the combo can be used as a plain amplifier that feeds a separate speaker system. The leftmost configuration offers one solution: Switching jack #1 disables the branch of the internal speaker whenever a plug is inserted. Another switching jack (#2) is normally bypassed but allows an additional series connection, which is once again enabled by inserting a plug. If there is no plug in jack #1 the whole external speaker system branch is naturally disabled while the internal speaker is enabled. (This is the normal condition). Switching jack #3 works similarly to jack #2 and allows connecting an external cabinet in series with the internal speaker. Note that a series connection here is mandatory because the load

impedance should not fall below the amplifier's rated limit. Transformer coupled amplifiers naturally behave differently and therefore these configurations will not work with them.

(For reference: Typical mono jacks usually include a single N.C. contact, which is connected to the tip when no plug is in the socket and disconnected when plug is inserted. This is depicted with an arrow).

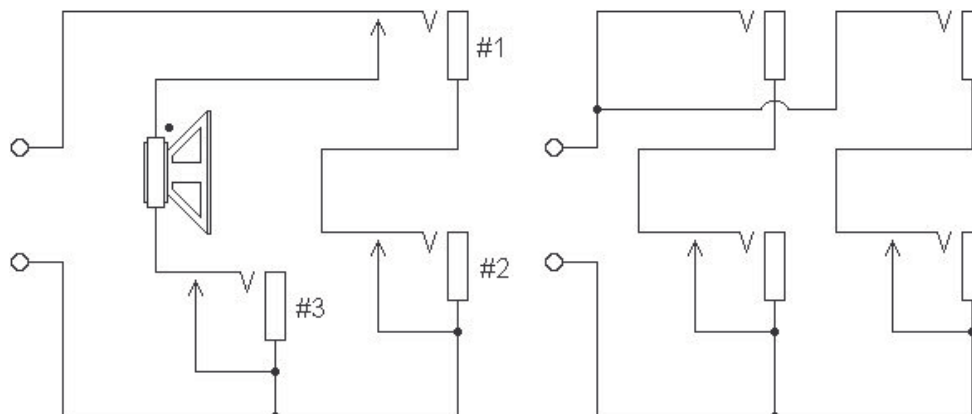


Figure 11.28 Output arrangements for a combo amplifier.

Of course we may take another route, similar to rightmost configuration: The internal speaker is hardwired to a conventional 6.3 mm phone plug (or some other suitable connector), which is, by default, connected to a suitable output jack from the selection of four. Removing the internal speaker plug thus will leave four output jacks free. These offer both parallel and series connections and lend themselves to just about any required configuration. As an additional benefit, the connection of the internal speaker no longer relies to the (sometimes unreliable) switching jack mechanism. A huge drawback is the fact that it is *too easy* to connect an external load in parallel with the internal speaker, thus amplifier overload is always a serious risk if connections are made uncaringly.

12 Chassis construction

12.1 Common chassis materials and types

These days one can walk into a suitable shop and buy a professional quality chassis with a reasonable price - also, a vast selection of professional rack-mount cases can be found from the Internet-based stores. In practice, acquiring a suitable enclosure will not require any extraordinary metalworking skills. A drill and a selection of different size *metal* drill bits is about the only requirement and even ordinary drill bits work quite perfectly if the chassis material is aluminum. More adventurous and DIY-spirited people will of course cannibalize and old chassis or build one by themselves from sheet metal. If the chassis is not very complex this will not require any special tools either (although using them can provide more successful outcome). The main aspect of this book is not to teach how to craft metal, but since it is a rather notable part in the amplifier building process a small insight is in place.

The most common materials for chassis are steel and aluminum. Compared to aluminum and copper, steel is a pretty poor conductor. This makes it a fine magnetic material and therefore useful in blocking magnetic fields. Steel is incredibly durable and strong – unfortunately the very same property makes it very heavy and difficult to machine as well. For a DIY-oriented person this is actually a *very* big drawback. In poor atmospheric conditions steel also rusts – unless it's stainless of course, in which case the machining is even more difficult.

Aluminum is light and soft and therefore easy to machine. In practice, it can be cut with an ordinary saw and odd-shaped holes (like square ones) can be shaped even with a file and without an excessive amount of work. Drilling requires no special drill bits and is very fast. These properties also make thin aluminum a poor choice for chassis that has to support heavy components, such as high-power transformers. It is common to see torn aluminum frames in high-power tube amplifiers. Thin aluminum chassis should always be braced to make it support heavy weights. Like copper, aluminum is a good conductor, thus a poor magnetic material and not useful for blocking magnetic fields. However, it is good in blocking electrical fields. Aluminum does not rust but it will still corrode and the corroded layer acts as an electrical insulator, which is a problem that has to be fought when chassis requires an electrical connection such as the safety ground. Aluminum may also be a tad more expensive than steel.

Copper is rarely used as a material for a complete chassis - mainly for good reasons: It is heavier than steel, although a bit easier to machine. Like aluminum, the material is a poor choice for blocking magnetic fields but a good choice for blocking electric. It is also an excellent conductor. Copper also corrodes forming an insulating layer. In practice, the greatest downside is the high cost that cannot be justified with any of the material's properties when compared to alternatives. Leave copper to be used as small, thin faceplate sheets or as heat spreaders for output devices.

For the record, it is my sole opinion that one should stay away from completely wooden chassis structures. Not only do they fail to provide protection against magnetic fields, RF and radiation they also introduce a serious risk of fire hazard and cannot be sufficiently safety earthed. If you wish to encase your amplifier into a completely wooden cabinet-chassis hybrid then at least built the electrical circuit on a

sturdy sheet of metal, which is properly safety grounded. Switches, jacks, potentiometers and other user-touchable parts should then have a solid electrical connection to this sheet. All flammable or moderately hot parts should also have a proper clearing from the wooden parts or a “firewall” in between. Layouts that allow even moderate amounts of dust to gather inside the chassis should be avoided as well: Igniting dust will not set a metal chassis on fire but it will do that for a wooden one. If you wish to neglect RF and magnetic shielding go ahead but at least do not neglect safety issues: Ignoring the shielding only makes up a poor sounding amplifier but ignoring the safety may set your house on fire. In my opinion, you might as well build an ordinary metal chassis since it’s equally difficult to make a wooden chassis that meets all the safety requirements.

One option for chassis material is plastic. This is not a very typical option for a guitar amplifier but undeniably it is used sometimes. Truthfully, *plastic is not a poor choice* but it has been characterized as being “cheap” and mechanically weak because it has been used in so many “toy-ish” designs. In reality, there are many types of plastics: Some of them are very strong, heat resistant and can provide electric shielding similarly to metals – and then again, some of them are exactly the opposite. In the PA field the use of sturdy, molded plastic cases – serving as both speaker cabinets and rack-gear boxes - is very common. However, likely the idea of introducing similar designs to the extremely conservative guitar world would be a commercial suicide. Molding a suitable plastic chassis is practically an impossible task for a common DIY-person and, in turn, the variety of commercial plastic boxes is really poor: Most of those are mechanically too weak to withstand the abuse introduced in any serious use. They also melt in heat and have to be spray-painted with conductive paint or coated with aluminum foil for sufficient shielding. About the only advantage is the inherent electrical insulation. However, do note that plastics also age and some of them may turn weaker or conductive in the process.

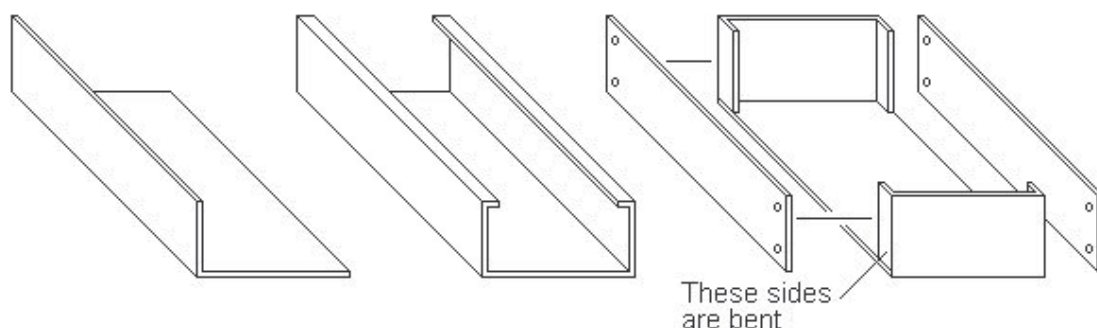


Figure 12.1 Simple chassis structures. Left: L-bar, middle: U-bar and right: “Bowl”.

Shown above are few basic chassis types that can fairly easily be constructed using ordinary tools. The L-bar “chassis” is about the simplest it can get. Compared to basic metal sheet, this construction is usually easier to mount on a cabinet and takes up considerably less space. The L-bar can be easily bent from a metal sheet using a wooden “angle jig” and a hammer – no other special tools are required. RF and EMI shielding properties of the “chassis” are, of course, nearly inexistent. The U-bar is one step forward, requiring little more material and skill. In practice, it can be bent similarly but added complexity is introduced if one wishes to make the 90° angled bends on each side for reinforcement (similar to ones shown at top). Since additional bends will make the chassis substantially stronger it is advisable to use them. Naturally the structure provides fairly improved shielding properties when compared to L-bar. This chassis type is likely the most common for small combo amplifiers. A metal “top plate” can be introduced within the cabinet structure.

Crafting the last chassis type begins to require serious construction skills and tools. The combined bottom and sidepiece is bent from a “distorted” H-shaped metal sheet. Of course this has to be cut first, which, without special tools, is actually a lot trickier than it may first seem. The rear and front plates are separate and can optionally be shaped so that the chassis can be fit to a 19” rack. For the best shielding properties, the separate rear and front sheets are substituted with an U-bar that also forms a top panel. The panels should be either riveted, screwed or bolted on place. If screws are used, a special tool for cutting a thread to the screw hole on the metal sheet is required. Getting all angles bent correctly and retaining correct dimensions is very difficult without special tools (bender/metal brake etc), yet it is utterly important.

Those planning to make a rack mount chassis may find the following information helpful: A 19-inch (482.6 mm) rack is a standardized mounting platform for all kinds of electronic devices. The equipment is mounted from its front panel to two parallel vertical metal bars (“rails”) - each of them is 0.625” (15.875 mm) wide. The distance between bars (measured from the “inside”) is 17.75 inches (450.85 mm), thus the overall width of *the whole rack* is 19 inches. The holes for mounting bolts have center-to-center distance of 18.3 inches (465.82 mm). The height of rack mount equipment is specified in a unit measure. One unit (1 U) is 1.75 inches (44.45 mm) from top to bottom. The holes in the vertical mounting bars are arranged (vertically) in repeating sets of three: Holes in each set have a center-to-center distance of 0.625 inches (15.875 mm). Center-to-distance between the “outer” holes of two sets is 0.5 inches (12.7 mm). Thus the pattern of the holes repeats every 1.75 inches – which also happens to be the height of one unit. The mounting arrangement is vertically symmetric.

12.2 Basics of metalworking

12.2.1 Cutting metal

In industry, sheet metal is usually cut with computer controlled (CNC) lasers or punch presses. However, a common DIY person usually has no access to such tools and has to rely on saws, routers and metal shears. Best way to cut metal precisely (aside using industrial machines) is to use a *fast* spinning saw equipped with a blade that is designed for cutting metal. You will also need some cutting fluid. There will be small metal shavings flying around so be sure to protect your eyes as well as the table saw: A nylon sock blocking the air intake holes of the saw’s motor is a good trick in the latter. Usually hacksaws and metal shears apply too much stress on the metal sheet causing it to warp awkwardly near the cut. Handheld metals shears are practically worthless for long cuts or cutting of thick metal sheets. A hacksaw is a decent tool for cutting thick metal profiles but compared to spinning saw it is too slow (I don’t mean its efficiency) and therefore leaves a very rough surface on the cut. In practice, all cuts made with a hacksaw have to be filed smooth.

12.2.2 Bending

In industry, special bending machines called *metal brakes* are used to bend sheet metal. Again, these are not necessarily something that an average DIY-spirited person can have an access to. A decent DIY bending trick is to shape a jig out of a wooden block and use it as a guide for bending. In practice, any sturdy object with a correct angle will do fine. Thick sheets are difficult to bend by hand so a hammer becomes an

essential tool. The trick is to hit the sheet very near the bending spot – otherwise you will be just punching dents to the sheet. A sturdy piece of wood can be used (and should be) to divide the force more evenly across the surface. The sheet should be attached to a vise that holds it firmly since any movement will create a non-precise line of bend. Use wooden pads between the metal and “jaws” of the vise to prevent scratching the sheet.

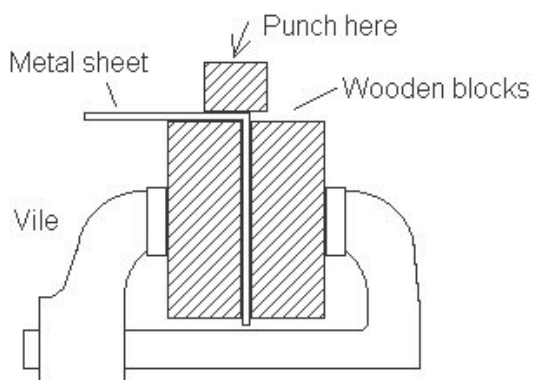


Figure 12.2 A crude jig for bending metal sheets.

12.2.3 Drilling

A large variation of various size drill bits and a good drill with variable speed control is the minimum requirement of tools for every one who is constructing their own chassis. Basic components like switches, jacks and potentiometers come in many shapes and sizes and with most certainty each of them requires a hole that has a different diameter than holes required by others components. Besides these, holes have to be drilled for fuse holders, circuit board stands, terminal strips, heat sink mounts, LED holders and for a vast deal of other parts and components. In practice, one can't have too many different-size drill bits. And of course it's good to have spare ones in case some will snap in two (very common with small drill bits) or plainly loose their edge. A powerful drill press (also known as pillar- or bench drill) is a recommended purchase since in most tasks it clearly surpasses a handheld drill in performance and accuracy. It is also *extremely* handy in drilling those very small holes of printed circuit boards.

Drilling metal generates a vast amount of heat that weakens the cutting power of the drill bit. When drilling requires an excessive amount of time the drilled area should be *oiled* regularly to reduce the heat and wear of the drill bit. Never drill large holes right away, instead start by drilling smaller diameter “guidance holes” - these will hold the drill more accurately in the correct position. When drilling a very large hole, the size should always be increased gradually. While this might seem like unnecessary extra work you will soon find out that it is not. The guidance holes vastly decrease the time required to drill through a sheet and provide a more precise result. When very accurate placement of drill holes is required it is recommendable to use a hammer and a nail to punch a small guidance slot before drilling the first guidance hole. The slot will hold the drill bit in its place throughout the period required for the bit to start biting through the metal. Of course, the most accurate results are acquired by using a drill press.

A word of warning: Occasionally the drill bit can “stuck” to the object you are drilling – likely causing it to spin wildly in an uncontrolled manner. If you are drilling a metal plate with sharp edges, something like this can do pretty gruesome damage if it hits you. At worst it can cut your precious fingers off. Always tie the drilled object firmly

on place with vise grips or some other appropriate way. Remember to wear gloves and suitable clothing that can absorb most of the cutting impacts far better than your skin does.

12.2.4 Polishing and painting

Rough and ugly metal surfaces can be polished shiny using fine grit sandpaper or fine metal wool. As a general principle, start with rougher grits and advance into finer ones. However, do note that too rough grit can easily ruin the surface leaving it full of nasty and deep scratches. It will take an extensive effort to sand and polish something like this to look good again. The final polishing is done with a cotton piece. When sanding the surface, it is *extremely* important to keep the sanding motion running into *one direction only* – you need a steady hand for this. The principle is difficult to explain on paper but you’ll figure it out very quickly once you actually start working on the sanding process. Sanding randomly into alternative directions creates a very uneven shine pattern that (in most cases) looks very unprofessional and ugly. The most talented craftsmen can use these various shine patterns as a very impressive effect but it requires a lot of skill.

Before painting it is important to sand the surface smooth from all roughness and (most importantly) *wash the chassis clean from any grease and dirt*. Soap water and acetone should usually be an adequate treatment (be careful not to leave any grease traces from the soap, though). After cleaning, handle the chassis only while wearing clean gloves and preferably hold it from places where a perfect paint surface is not an absolute requirement. For best finish, the chassis should absolutely be spray-painted. Apply only thin layers of paint in order to make it spread evenly and to prevent “bleeding”. Do not spray from a too close distance and always let the paint cure thoroughly before applying new layers. Do note that some paints require a primer for smooth finish. Sometimes it is useful to slightly sand the surface before applying another layer of paint. If you apply a too thick surface of paint without letting the other layers cure thoroughly the finish will be coated with hairline cracks and marks of air bubbles. For shiny finish, spray the chassis with lacquer. These days a paint called “Hammerite” seems to be common among guitar effect DIY community. It creates a “battered” surface - similar to ancient Fuzz faces - and requires no primer.

Text and other decals can be added with various methods: There is a vast amount of companies that offer silk-screening (or “screen-printing”) and laser-based “etching” services. These are techniques that are art by themselves to apply; yet it is possible. In etching, a bare metal chassis surface is required. It is then covered with a layer of suitable protective lacquer leaving only the areas where some decals are required visible. When acid is applied it eats through the unprotected areas. This creates a dull “matte” pattern that is either light or dark depending on the etching time. The method does not work with all metals - such as aluminum or stainless steel. A laser beam can be used for etching as well, however, this requires industrial machinery. In silk-screening a layer of paint is applied through a stencil. The original technique to create the stencils was to block areas of silk fabric with a non-permeable material, hence the name. However, polyester has been used instead of silk since the 1940’s. When paint is applied on top of the stencil screen the fine mesh of the cloth lets it through while the non-permeable portions of the stencil are left unpainted. A breakthrough was the invention of photo-emulsive materials that could form the non-permeable mask. In this technique, UV light applied through a film mask is used to harden the emulsion in a required pattern. Stencils manufactured this way can have very high levels of detail.

The aforementioned techniques are not extremely difficult but they do require a lot of dedication, time, special tools and most importantly, practice. Naturally one must learn the basics of these arts as well and teaching that is not the topic of this book. Most people abandon the idea of using these techniques in favor of simpler methods such as using of permanent marker pens or “letraset”. The latter is a common name for sheets that contain transferable alphabetic letters and numbers (in various font styles). However, letraset do not attach very firmly and one can easily scrape them off. Consequently they must be protected, for example, by applying a layer of lacquer.

12.3 Ventilation and other thermal-related considerations

When designing chassis ventilation one must note that the force of natural convection is very small, thus all sharp-angled turns will seriously impede the airflow. In practice, the best arrangement for ventilation would likely involve having intake air holes at the bottom of the chassis (with sufficient clearance of course) and exhaust air holes at the top. Fitting ventilation holes at the top of the *rear* plate (which is an extremely usual configuration) is always more or less a compromise, as the warm air has to escape making a 90° turn. Substituting simple round holes with tall and narrow slots can help this a little since it widens the angle. Unfortunately such slots are difficult to craft without a punch press. To retain the efficiency, all large, horizontally mounted circuit boards should have some holes to provide air paths as well. Such structure may create a weaker circuit board but it also enhances the cooling of board-mounted components.

Ventilation and heat sinking arrangements often go hand in hand. Obviously, fitting the heat sink inside the chassis may significantly impair the cooling airflow so I recommend avoiding such designs unless the arrangement is such that it provides an unobstructed path for airflow. However, the configuration does have some considerable benefits: The inherent simplicity is perhaps the main one since it allows coupling the output devices to the heat sink very efficiently. Other benefits are that the “delicate” heat sink is *isolated from the user* as well as *protected by the sturdy chassis*. If you ever cut yourself to the sharp fins of a heat sink or experienced how easy it is to dent that piece of metal you will appreciate these attributes. Normally the “internal heat sink” design appears in two forms: 1. The heat sink is “floating” (not attached to chassis *thermally*) and 2. The heat sink is attached to the chassis, thus chassis also serves as a heat sink. In some cases the “heat sink” part is just a “spreader block” that leaves majority of heat dissipation to responsibility of the chassis.

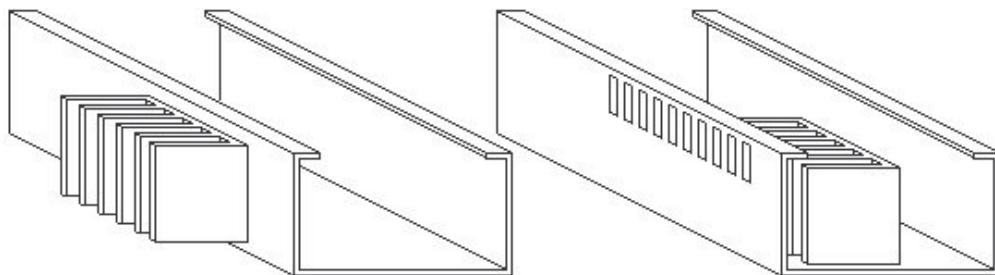


Figure 12.3 Cooling solutions: External and internal heat sinks.

The external heat sink mounting typically provides a less obstructed path for airflow but it also suffers from the issues mentioned earlier. Unless a hole is cut into the chassis wall it will also introduce another thermal resistance in between the semiconductors and the heat sink. In many cases utilizing this scheme also means that the semiconductors are exposed. Since *their cases typically have a direct electrical*

connection to the supply rails there must be a scheme to prevent user from touching them. This typically creates a need for hard-to-find special parts such as plastic transistor covers or metal grilles that fold over the heat sink.

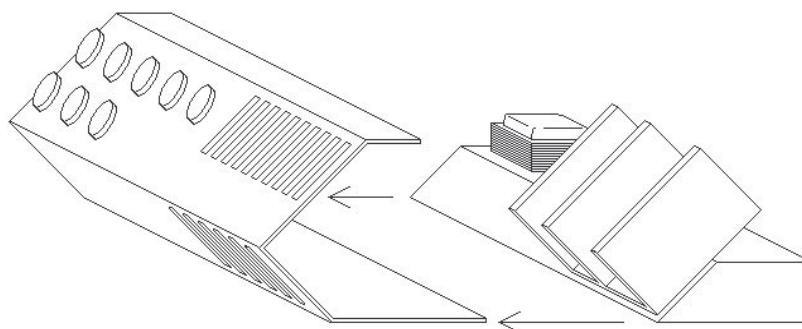


Figure 12.4 Cooling arrangement of the early solid-state Fender amplifiers isolated the heat sink and utilized natural convection with the aid of “chimneys”.

One of the earliest solutions to overcome some of the aforementioned problems was introduced by CBS (Fender) designer Paul B. Spranger in the late 60’s. The details of the invention are fully covered in U.S. Patent 3,462,553 (Solid-State Amplifier, And Control Panel Assembly Incorporated Therein, Spranger, P.B, August 1969). Essentially, the chassis is constructed from two parts that are the sheet metal front plate and the main chassis, which also serves as the mounting platform for the heat sink. Besides housing the controls, the front plate serves as a metal grille that protects the output transistors and the heat sink. It also forms a set of vertical heat tunnels, chimneys. A chimney boosts up the natural convection and prevents warm air from leaking into other sections of the amplifier. Since heat sink is located at the front side of the amplifier – as well as “elevated” due to angled panel shape - obstructing the airflow accidentally (for example, by placing the amplifier against a wall) is nearly impossible and the heat sink is always fed with “fresh”, cool air.

Various ways to mount the heatsink either into the top or bottom side of the chassis exist as well. These basically have the benefit of providing a bit more heat sink “protection” than the rear-mount arrangement – yet both configurations are usually somewhat compromises. The bottom-mounted heat sink configuration always suffers more or less from a “cooking” effect where the warm air, that is rising upwards, heats up the chassis above. (This problem is far more severe in tube amplifiers, though). The upward exit path for the warm air is also severely obstructed. A legion of variations of the theme exists. A rather novel arrangement was used, for example, in the construction of some Lab Series amplifiers: The cabinet has an air intake hole (for a fan) in one of its side panels. This usually provides a rather unobstructed source for cool air, which is then pushed (by the fan) along the horizontal heat sink all the way to the main filter capacitors and the power transformer (which are the secondary parts requiring of cooling). Warm air then exits from the open rear side of the cabinet. Without forced cooling the arrangement would be less ideal, yet quite typical. A metal grill covers the TO-3-case transistors, thus preventing the user from accidentally touching them.

If the configuration is flipped around, so that the heat sink is at the top of the chassis, we are naturally obstructing the convective intake airflow again. (However, if you have read the chapter about heat sinking you know this is a lot more feasible arrangement than the aforementioned one). Basically, the layout will work quite well assuming the warm exhaust air has a rather unobstructed way to escape and the cool intake air to flow in. Usually a semi-open backed cabinet structure fills these clauses

quite well. Again, variations exist in legions. One worth mentioning was used in some Randall amplifier heads: The heat sink is formed from the combination of the chassis and some vertical plates screwed into it. Transistors are then mounted to the top panel of the chassis in between the plates. The top panel of the chassis now acts as the base plate of a heat sink while the plates act as the fins. This is not a very effective arrangement when compared to a normal heatsink, though: For example, to serve as an adequate base plate the chassis must be very thick, there is a notable thermal resistance between the fin plates and the chassis and five rather thick plates (as used by Randall) dissipate heat quite poorly. However, one can't ignore the fact that the arrangement is *dirt-cheap* and can be built from a pile of scrap metal. While I do not recommend utilizing a configuration like this, it hopefully provided some food for thought. In the concerned Randall amplifier the TO-3-case transistors had plastic covers, which prevented accidental touching of the transistor cases.

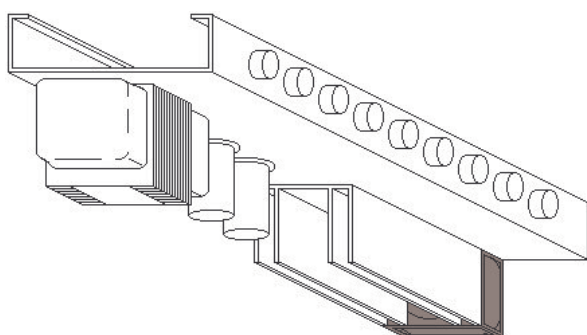


Figure 12.5 Lab Series chassis with a bottom-mounted horizontal heat sink and forced cooling.

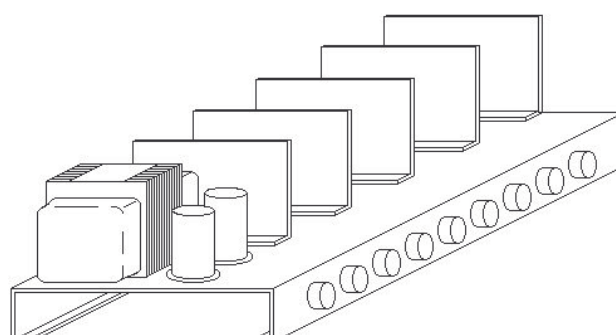


Figure 12.6 Randall forms the heatsink from the chassis and vertical “dissipation” plates attached to it. Transistors are mounted to chassis between the plates.

Since majority of designs use cooling solutions that are based on natural convection it is often useful to think about methods that would increase the force of airflow. A novel, as well as patented, technique is to harness the speaker cone's movement as a pump for cooling air. Inspiration for such designs can be acquired, for example, from US Patent 3,778,551 (Chicago Musical Instruments, Grodinsky, 1973). I believe such designs were used at least in some Gibson amplifiers. One should not forget various heat tunnel configurations either as they can be extremely effective.

12.4 Effects of mechanical vibration

Solid-state devices are not microphonic so there is no need to worry about resonance affecting the signal. *Sound pressure can have a significant effect on the resonance of the chassis* but this is a different matter. Generally, a solid-state amplifier is not a transducer; however, you do need to dampen the chassis because in the worst case the resonance can break solder joints - this usually happens on large components that are mounted improperly. Since large capacitors fill the clause some people mistakenly believe they are severely microphonic. Capacitors do indeed possess microphonic qualities since the distance of the plates affects the capacitance (the condenser microphones are based on this phenomenon) but in a typical capacitor the resonance causes a signal with maximum amplitude of only few microvolts. This would require an enormous gain to be amplified audible. A secure mounting is the sole reason why you see capacitors glued on place. If in doubt, you can test this claim by slightly tapping the capacitors. If you do hear any difference check the solder joints and the condition of the capacitor.

Exceptions to the above are some ceramic capacitors, especially ones using “high-K” dielectric material. They are susceptible to microphonics since the substrate possesses piezoelectric properties. The use of these materials in semiconductor devices (assumably in MOSFETs) is still on an engineering level and no such commercial product is yet available. There *are* some microphonic capacitors but in practice, one sees them only in very old equipment. Today, the manufacturing methods are so much better.

The chassis and some of the components inside it can make plenty of *mechanical* noise that can be misinterpreted (by some people) as being a part of the signal: Transformers tend to vibrate as current flows through them and this causes a very audible noise; the phenomenon is known as *magnetostriction*. For example, on quiet occasions I can hear the transformer of my HiFi stereo amplifier resonating at a frequency of 50 Hz from about one meter away from the amplifier. Similarly, on some occasions my computer’s switched mode supply makes a high-pitched ringing sound – similar to brief tinnitus. Being a high frequency, this is very audible even from larger distances. Some transformers are “potted” to stop these ringing effects. Generally, the mechanical noise is louder the more powerful the transformer is - and especially the more current flows through the windings – thus the sound of magnetostriction can also serve as a very good indicator of a short circuit, even mild one. During a short circuit the transformer usually emits a moderately loud, wailing noise. Toroidal transformers tend to create less mechanical noise than EI-cored.

Heat sinks create mechanical noise too. Vibration of large fins causes changes in the air pressure that we perceive as sound, the natural frequency of this is usually between 2 and 3 kHz. The vibrations are mainly born by two ways: An external source (like loudspeaker) resonates the heat sink or temperature changes cause the fins to vibrate. As a common and rather logical rule, heat sinks with large and thin fins will vibrate more noisily than heat sinks with small, stiff and thick fins. This is unfortunate since efficient heat sinks also tend to be the ones that are “ringing”. Common damping methods against heat sink ringing are fitting rubber bands around fins or rubber pads between them. Silicon paste or the hot glue used in mounting large capacitors has the same effect.

The sound originating from “ringing” transformers and heat sinks is not very loud but might impose a problem in “quiet” practice sessions or during recording with moderately low sound pressures. Compared to mains hum or picked up interference the aforementioned problems are still pretty minor concerns – yet they do exist and when they appear they are often extremely annoying. List of other, more noteworthy, noisemakers includes cooling fans and spring reverb tanks. At typical *stage volumes*, the latter is the one I consider as the only noteworthy source of very audible and unwanted noise.

Since last one on the list *is* a transducer it is also the only notable source that might introduce *a lot* of vibration-related noise to the actual output signal. Some people prefer the resonating effect of spring reverbs to some extend - after all one of the basic techniques in “modern” electric guitar playing is exploiting the sound of feedback. The resonant feedback can add some “life” to the generic guitar tone; it also serves as a special effect, similar to pinch harmonics or tricks with the tremolo arm. Feedback through guitar strings is mostly controllable, however, excessive feedback through the reverb tank is not, and thus it is hardly a welcomed feature: At worst, it can even be fatal to hearing and to the speaker system. Best practice is to mount the tank on a

surface that damps a lot of external vibrations. The tendency to feedback should be tested on multiple frequencies to assure the reverb tank will *never* start to resonate due to its own output signal. Important thing is to make sure the master volume controls the signal of the tank as well – otherwise a reverb tank catching feedback will feed the power amplifier with excessive signal amplitude *and the feedback will soon occur at full loudness of the system!* I have seen at least one amplifier that went on to this “self destruction” mode after the reverb potentiometer was turned past “9”. Preventing mechanic feedback is area where separate, “piggybacked” head-cabinet systems really show their worth.

Miscellaneous tables and information

Resistor 4-band colour codes

Colour	1. Band	2. Band	3. Band (multiplier)	4. Band (tolerance)
Black	0	0	1	
Brown	1	1	10	± 1 %
Red	2	2	100	± 2 %
Orange	3	3	1000	
Yellow	4	4	10 000	
Green	5	5	100 000	± 0.5 %
Blue	6	6	1000 000	± 0.025 %
Violet	7	7	10 000 000	± 0.1 %
Grey	8	8	100 000 000	± 0.05 %
White	9	9	1000 000 000	
Gold			0.1	± 5 %
Silver			0.01	± 10 %
No				± 20 %

Resistor 5-band colour codes

Colour	1. Band	2. Band	3. Band	4. Band (multiplier)	5. Band (tolerance)
Black	0	0	0	1	
Brown	1	1	1	10	± 1 %
Red	2	2	2	100	± 2 %
Orange	3	3	3	1000	
Yellow	4	4	4	10 000	
Green	5	5	5	100 000	± 0.5 %
Blue	6	6	6	1000 000	± 0.025 %
Violet	7	7	7	10 000 000	± 0.1 %
Grey	8	8	8	100 000 000	± 0.05 %
White	9	9	9	1000 000 000	
Gold				0.1	± 5 %
Silver				0.01	± 10 %
No					± 20 %

Zero-ohm resistors (jumper “wires” for automated assembly machines) are marked with a black stripe in the middle of the component’s body or, in the case of SMD-components, with a dot.

Letter coding of potentiometers

Type	Old marking	New marking		Alternatives
		USA / Japan	Europe	
Linear	A	B	A	Lin
Logarithmic	C	A	B	Log / Audio
Antilog or Log type 2	F	N.A. / C	C	No
Accentuated audio/various		D		

NOTE: For reliable info, consult the manufacturer’s datasheets.

IEC 3-digit coding of resistors and capacitors

1. Digit	2. Digit	3. Digit (multiplier or number of zeros)		Letter (tolerance EIA code)	
Value in picofarads or ohms	0	1	A	0.05 %	
	1	10	B	0.1 %	
	2	100	C	0.25 %	
	3	1000	D	0.5 %	
	4	10 000	E	0.5 %	
	5	100 000	F	1 %	
	6	No	G	2 %	
	7	No	H	3 %	
	8	0.01	J	5 %	
	9	0.1	K	10 %	
			L	15 %	
			M	20 %	
			N	30 %	
			P	+100 % / -0 %	
			S	+50 % / -20 %	
			W	+200 % / -0 %	
			X	+40 % / -20 %	
			Z	+80 % / -20 %	

NOTES: For example, 474M = 47 x 10,000 pF = 470 nF, 20%

EIA-rated voltage marking code

Voltage	4	6.3	10	16	20	25	35	50
Code	G	J	A	C	D	E	V	H

Capacitor 5-band colour codes

Colour	1 & 2	3. Band (multiplier)	4. Band (tolerance)	5. Band (voltage)
Black	0	1	± 20 %	
Brown	1	10		100 V
Red	2	100		250 V
Orange	3	1000		
Yellow	4	10 000		400 V
Green	5	100 000		
Blue	6			630 V
Violet	7			
Grey	8	0.01		
White	9	0.1	± 10 %	

NOTES: Value is in picofarads. In axial capacitors, the stripes are read starting from the one that is closest to the end of the body. In radial capacitors, the stripes are read starting from the side that is opposite to the side to which the leads are connected.

Capacitor dotted colour code:

This type of coding has not been in use for decades but you can still bump into it in older equipment. The dots are read from left to right and in case of a column break the reading continues again in a similar manner (from left to right). Thus the common reading pattern is following:

A	B	C	D	E
---	---	---	---	---

The three-dot code simply indicates the value in picofarads while a five-dot code also indicates the voltage rating and tolerance. Sometimes these may be printed on the other side of the component. Marking with six dots is a special case and read in an alternative manner:

A	B	B1
D	E	C

Interpretation table

Colour	A	B & B1	C (multiplier)	D (voltage)	E (tolerance)
Black	0	0	1	100 V	1 %
Brown	1	1	10	200 V	2 %
Red	2	2	100	300 V	3 %
Orange	3	3	1000	400 V	4 %
Yellow	4	4	10 000	500 V	5 %
Green	5	5	100 000	600 V	6 %
Blue	6	6	1000 000	700 V	7 %
Violet	7	7		800 V	8 %
Grey	8	8		900 V	9 %
White	9	9			
Silver					20 %
Gold				1000 V	
No				500 V	20 %

Capacitor polarity identification

Identification	Meaning
Plus	Positive
Minus	Negative
Side with stripe	Negative
Side with groove (Axial)	Positive
Arrow's direction (Axial)	Negative
Longer lead	Positive
White side of circle (PC boards)	Negative
Dot or stripe (SMD)	Negative
Black plate (schematics)	Negative
White Block (schematics)	Positive
Curved plate (schematics)	Negative
Half moon or half circle	Positive
Triangle	Negative

Alternative unit markings (standards in bold)

Marking	Unit	Meaning	Notes
mF	Millifarad	1×10^{-3}	Rare (conflicting)
µF or uF	Microfarad	1×10^{-6}	
mF	Microfarad	1×10^{-6}	USA (conflicting)
ufd	Microfarad	1×10^{-6}	USA
mfd or MFD	Microfarad	1×10^{-6}	Not used anymore
nF	Nanofarad	1×10^{-9}	
pF	Picofarad	1×10^{-12}	
mmF	Micromicrofarad	1×10^{-12}	USA

LED lead identification

Identification	Meaning	
Flat side	Cathode	Negative
Shorter lead	Cathode	Negative
Longer lead	Anode	Positive
NOTES:		
1. In anti-parallel bicolour LEDs (these have only two leads) the longer lead typically connects to cathode of the (typically) red LED, which is also the anode of the (typically) green LED.		
2. In shared-cathode bicolour LEDs all leads are typically of different length. Longest lead (typically) connects the shared cathodes, second longest lead connects the anode of the green LED and shortest lead connects the anode of the red LED.		

Bridge rectifier lead identification

Identification	Meaning
Plus	Positive out
Slanted corner	Not a valid indicator
Dot	Positive out
“Angled” lead	Positive out
Wave symbol	AC in
Longest leg	Positive out
White circle (PC boards)	Typically positive out

In discrete diodes the cathode (negative) side is marked with a stripe or a rounded end of the body.

Semiconductor markings – European (Pro-Electron)

<i>Two letters, [Optional letter], Serial number, [Suffix] i.e. BC506B</i>							
1. Letter (material)		2. Letter (type)			Suffix (version)		
A	Germanium	A	Diode, detector, switch, mixing			A	Low Hfe
B	Silicon	B	Capacitance diode			B	Medium
C	Gallium arsenide	C	Transistor, low power, LF			C	High
D	Indium antimonide	D	Transistor, high power, LF			R	Reverse
R	Hall / photoelectric	E	Tunnel effect diode			W	SMD
<i>Details:</i> <i>A: Material with a band gap of 0.6 – 1.0 eV</i> <i>B: Material with a band gap of 1.0 – 1.3 eV</i> <i>C: Materials with band gap that is higher than 1.3 eV</i> <i>D: Ceramic materials</i> <i>R: Compound materials</i>		F	Transistor, low power, HF				
		G	Combination				
		H	Magnetically sensitive				
		K	Hall-element, open				
		L	Transistor, high power, HF				
		M	Hall-element, closed				
		N	Optocoupler				
		P	Light or radiation sensitive				
		Q	Light or radiation emitting				
		R	Low power switch				
		S	Switching transistor, low power				
		T	Triggerable switching (Thyristor, Triac, etc.)				
		U	Switching transistor, high power				
		X	Diode, varactor, step recovery				
Y	Rectifier						
Z	Zener						
<p>NOTES:</p> <p>1. The optional (third) letter indicates if the device is indented for industrial, professional or military use. If it is used, the serial number has only three digits instead of four. Letter coding has no fixed meaning except, A=Triacs, B=HBT transistors, F=Fibre-optic, H=HEMT transistor, L=Lasers, M=Transistor drivers, O=Opto-triacs, R=Resistor network (SS), T=Tri-state dual colour (bicolour) LEDs and W=Transient voltage suppressor diodes.</p> <p>2. Version letters do not have a fixed meaning by standard. Typically A, B and C are used for denoting gain category. R indicates reversed polarity and W a surface-mount device. Sometimes suffix has two letters.</p>							

Semiconductor markings – Joint-electron Device (JEDEC)

<i>Digit, Letter, Sequential number, [Suffix] i.e. 2N2222A</i>		
Digit	Letter	Suffix
1 Diodes	Always "N"	Indicates an enhanced version. Consult the datasheet.
2 Bipolar transistors		
3 FET devices		
4 Optocouplers		
5 Optocouplers		
<p>NOTES: The digit used to mark number of junctions but this has become a non-standard marking.</p>		

Semiconductor markings – Japanese Industrial Standard (JIS)

<i>Digit, Two letters, Sequential number, [Suffix] i.e. 2SC504</i>				
Digit	Letters (type)		Suffix (gain)	
Number of junctions.	SA	PNP, HF	A	Low
	SB	PNP, AF	B	Medium
	SC	NPN, LF	C	High
1 = Diodes 2 = Transistors 3 = FETs	SD	NPN, AF		
	SE	Diode		
	SF	Thyristor		
	SG	Gunn		
	SH	UJT		
	SJ	FET/MOSFET, P-channel		
	SK	FET/MOSFET, N-channel		
	SM	Triac		
	SQ	LED		
	SR	Rectifier		
	SS	Signal diode		
	ST	Diode, avalanche		
	SV	“Varicap”		
	SZ	Zener		

Semiconductor markings – Manufacturer specific prefixes

MJ	Motorola, high power, metal case
MJE	Motorola, high power, plastic case
MPS	Motorola, low power, plastic case
MRF	Motorola, HF, VHF, microwave
RCA	Radio Corporation of America (RCA)
RCS	RCS
TIP	Texas Instruments, low power, plastic case
TIPL	Texas Instruments, high power, planar
TIS	Texas Instruments, low power, plastic case
ZT	Ferranti
ZX	Ferranti

References

1. Bussey, Stephen, W. – Haigler, Robert, M., Tubes versus transistors in electric guitar amplifiers, IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 6, p. 800 – 803, 1981
2. Hamm, Russell, O., Tubes versus Transistors: Is There An Audible Difference? 43rd Audio Engineering Society convention, 14. September 1972
3. Self, Douglas, Audio Power Amplifier Design Handbook, Third Edition, Newnes, 2002 (ISBN 0 7506 56360)
4. United States Patent 2,896,029, Semiconductor Amplifier Circuits, Lin, Chang, Huang, 29. July 1959
5. Darr, Jack, Electric Guitar Amplifier Handbook, Third Edition, Howard W. Sams & Co., Inc., the Ross-Merrill Co., Inc., 1973
6. Brown, Marty, Good Physical Layout Takes Black Magic Out Of Power-Supply Design, Fairchild Semiconductor whitepaper, http://www.techonline/community/ed_resource/feature_article/5430
7. Crown whitepaper, Crown's Grounded Bridge Output Topology
8. Duncan, Ben, High Performance Audio Power Amplifiers, Newnes, 1996
9. Bridge/Parallel Amplifier (BPA-200) Documentation, National Semiconductor whitepaper, 1997
10. DeCelles, John, Audio Amplifiers Utilizing: SpiKe Protection, National Semiconductor Application Note 898, August 1993
11. Meyer, Daniel, M/M/M Instrument Amplifier Part 1, Popular Electronics, Volume 28, Number 4, April 1968
12. Meyer, Daniel, M/M/M Instrument Amplifier Part 2, Popular Electronics, Volume 28, Number 5, May 1968
13. Linsley-Hood, John, Simple Class A Amplifier, Wireless World, April 1969
14. Linsley-Hood, John, The Evolution of Audio Amplifier Design, parts 1 – 5. Electronics In Action, February – June issues 1994
15. Ruehs, Richard, G., High Power Audio Amplifiers With Short Circuit Protection, Motorola Application Note AN485, Motorola, Inc. 1985
16. Kitchin, Charles, Avoiding OpAmp Instability Problems In Single-Supply Applications, Analog Dialogue, Volume 35, 2001
17. Leach, Marshall, Introduction to Electroacoustics and Audio Amplifier Design, Second Edition, Revised Printing, Kendall/Hunt, 2001
18. Elliott, Rod, Using Current Sources, Sinks and Mirrors in Audio, <http://sound.westhost.com/ism.htm>
19. Elliott, Rod, Variable Amplifier Impedance, <http://sound.westhost.com/project56.htm>
20. Current Limiting Application Note 9, Apex Microtechnology Corp., Revision B, February 1998
21. United States Patent 7,061,740, System and Method for Protecting An Audio Amplifier Output Stage Power Transistor, Mendenhall, Eric, 13. June 2006
22. Wilkins, Charles, Variable Damping Factor Control, Audio, September 1954
23. Audio Power Amplifier Output Stage Protection, Audio Engineering Society, Convention Paper 5695, October 5 – 8, 2002
24. Shin, Donglyoul, Thermal Design and Evaluation Methods for Heat Sink
25. Simpson, Chester, Linear and Switching Voltage regulator Fundamentals, National Semiconductor whitepaper
26. Barbour, Eric, The Cool Sound of Tubes, IEEE Spectrum, Volume 35, Number 8, August 1998
27. Hawksford, Malcolm, O., Distortion Correction in Audio Power Amplifiers, An Audio Engineering Society Preprint from 65th Convention, February 1980

28. Leach, Marshall, Feedforward Compensation Of the Amplifier Output Stage for Improved Stability With Capacitive Loads, IEEE Transactions on Consumer Electronics, Vol. 34, No. 2, p. 334 – 338, May 1988
29. Borbely, Erno, JFETS: The New Frontiers, Part 1, Audio Electronics, No. 5, p. 26 – 31 1999
30. DC Power Supply Handbook, Agilent Technologies, Application Note 90B, 1978
31. Vanderkooy, John – Lipshitz, Stanley, P., Feedforward Error Correction in Power Amplifiers, Journal of The Audio Engineering Society, January-February, Volume 28, Number 1/2, 1980
32. JFET Biasing Techniques, Siliconix, Application Note 102, March 1997
33. Marston, Ray, FET Principles and Circuits Part 2, Nuts & Volts Magazine, June 2000
34. Mancini, Ron, Op Amps For Everyone - Design Reference, SLOD006B, Texas Instruments, August 2002
35. Jung, Walt, Op Amp Applications Handbook, Analog Devices, Inc., Newnes, 2005 (ISBN 0-7506-7844-5)
36. Loudspeaker Enclosures – Their Design and Use, First Revision, Altec Sound Product Division, 1974
37. Small, Richard, Direct-Radiator Loudspeaker System Analysis, IEEE Transactions on Audio and Electroacoustics, vol. AU-19, p. 269 – 281, December 1971
38. Small, Richard, H., Closed-Box Loudspeaker Systems, Part I: Analysis, Journal of the Audio Engineering Society, Volume 20, Number. 10, p. 798 - 808, December 1972
39. Small, Richard, H., Vented-Box Loudspeaker Systems, Part II: Large-Signal Analysis, Journal of the Audio Engineering Society, Volume 21, Number 6, p. 438 – 444, July/August, 1973
40. Olson, Harry, F., Dynamical Analogies, D. Van Nostrand Company, Inc. 1943
41. Beranek, Leo, L., Acoustics, Acoustical Society of America, 1954, 1996 Edition (ISBN 0-88318-494-X)
42. Ahonen, Janne, About the ports and differences between various equations and design software calculations, Linearteam tech paper, December, 2004
43. Slone, Randy, G., High-Power Audio Amplifier Construction Manual, McGraw-Hill, 1999
44. Brosnac, Donald, Guitar Electronics for Musicians, Omnibus Press, 1983 ISBN-0-7119-0232-1
45. Pass, Nelson, Cascode Amp Design, Audio, March 1978
46. Leach, Marshall, The Leach Amp: The Output Stage, 2000
<http://users.ece.gatech.edu/~mleach/lowtim/output.htm>
47. Langford-Smith, Radiotron Designer's Handbook, Fourth Edition, Radio Corporation of America, 1953
48. United States Patent 5,467,400, Solid State Audio Amplifier Emulating A Tube Audio Amplifier, Keir, Bruce – Keynes, Milton, 14. November 1995
49. United States Patent 6,350,943 B1, Electric Instrument Amplifier, Suruga et al., 26. February 2002
50. United States Patent 3,536,958, Amplifier Protection Circuit, Sondermeyer, Jack, 27. October 1970
51. United States Patent 4,405,832, Circuit For Distorting An Audio Signal, Sondermeyer, Jack, 20. September 1983
52. United States Patent 4,439,742, Circuit For Simulating Vacuum Tube Compression In Transistor Amplifiers, Sondermeyer, Jack, 27. March 1984
53. United States Patent 4,811,401, Superdistorted Amplifier Circuitry With Normal Gain, Brown, Sr. et al., 7. March 1989
54. United States Patent 5,197,102, Audio Power Amplifier System With Frequency Selective Damping Factor Controls, Sondermeyer, Jack, 23. March 1993
55. United States Patent 5,524,055, Solid State Circuit For Emulating Tube Compression Effect, Sondermeyer, Jack, 4. June 1996

56. United States Patent 5,619,578, Multi-Stage Solid State Amplifier That Emulates Tube Distortion, Sondermeyer et al., 8. April 1997
57. United States Patent 5,647,004, Multi-Stage Solid State Amplifier That Emulates Tube Distortion, Sondermeyer et al., 8. July 1997
58. United States Patent 5,675,656, Power Amplifier With Clipping Level Control, Sondermeyer et al., 7. October 1997
59. United States Patent 5,032,796, Solid State Amplifier Circuit Simulating Vacuum Tube Distortion Characteristics, Tiers, Jerrold, S. – Kieffer, Thomas, E., 16. July 1991
60. Floyd, Thomas, L., Electronic Devices, Third Edition, Merrill, 1992
61. United States Patent 6,816,009 B2, Circuit And Method For A Transimpedance Audio Amplifier Providing Variable Gain While Maintaining Load Impedance, Hughes et al, 9. November 2004
62. United States Patent 3,952,256, Multi-Impedance Output For Transistor Power Amplifier, Kosinski et al, 20. April 1976
63. United States Patent 4,987,381, Tube Sound Solid-State Amplifier, Butler, Brent, K., 22. January 1991
64. United States Patent 5,022,305, Tube Overdrive Pedal Operable Using Low Voltage DC Battery Eliminator, Butler, Brent, K., 11. June 1991
65. United States Patent 5,343,159, Direct Box Employing Hybrid Vacuum Tube And Solid State Circuitry, Butler, Brent, K., 30. August 1994
66. United States Patent 5,705,950, Hybrid Solid-State/Vacuum Tube Audio Amplifier, Butler, Brent, K., 6. January 1998
67. United States Patent 6,507,240, Hybrid Audio Amplifier, Butler, Brent, K., 14. January 2003
68. Harwood, H., D. – Mathews, R., Factors in the design of loudspeaker cabinets, BBC Research Department, Engineering Division. January 1977
69. United States Patent 5,315,263, Push-Pull Amplifier Having Bootstrap Capacitor Isolation, Mudra et al, 24. May 1994
70. Electronic Designer's Guide EC101-J, Littelfuse, Inc. April 2005

Glossary of terms and abbreviations

I_{ABC}	Amplifier Bias Current	CT	Center Tap
AC	Alternating Current	CVS	Constant Voltage Source
ADC	Analog Digital Converter	DAC	Digital to Analogue Converter
AES	Audio Engineering Society	Damping factor	Amplifier's ability to slow down speaker cone movement caused by inertia
AGC	Automatic Gain Control	DC	Direct Current
Anode	An electrode through which positive electric current flows into a polarized device	DF	Damping Factor
beta	Current gain of a BJT	Dielectric	Insulating material used in a capacitor
Bias	A fixed voltage or current that defines an operating point	DIL	Dual-in-line
Bias current	Quiescent or idle current	DIN	Deutsches Institute fur Normung (German standards organisation)
Bi-filar	A winding style that cancels inductivity	DIP	Dual-in-line Package
BJT	Bipolar Junction Transistor	Discrete	Composed of individual components
Bleeder	A resistor that discharges stored charge	DMM	Digital Multimeter
BTL	Bridge-Tied Load	DPST	Double Pole Single Throw (switch)
Buzz	Sharp sounding oscillation. Typically second harmonic of mains frequency.	Driver	A loudspeaker or an (amplifying) transistor preceding a specific stage
Cascade	A series connection of circuits	DSP	Digital Signal Processing
Cascode	An arrangement that combines two amplifier circuits	Dummy load	A (silent) test load that substitutes the speaker (or any other specific load)
Cathode	An electrode through which positive electric current flows out of polarized device	EF	Emitter Follower
CCS	Constant Current Source	EIA	Electronic Industrial Association
CCW	Counter clockwise	EMF	Electro Motive Force
CE	Common Emitter	EMI	Electro Magnetic Interference
CFB	Current Feedback	EQ	Equalisation (tonal manipulation)
CFP	Complementary Feedback Pair	ESR	Equivalent (or Effective) Series Resistance
Chatter	Bouncing of contacts during switching	FB	Feedback
Chip	Synonym for integrated circuit	FET	Field Effect Transistor
Choke	An inductor	FFT	Fast Fourier Transform
CLG	Closed Loop Gain	Flywheel diode	Diode that protects from reverse polarization due to inductive spiking
Clipping	Overload, where signal peaks are flattened	FMV	"Fender-Marshall-Vox". A tone control circuit topology
CMRR	Common Mode Rejection Ratio	FP	Flame Proof
CO	Controlled Oscillator	FSK	Frequency-Shift Keying also known as MSK
Coil	An inductor	FX loop	Effect loop
Condenser	A capacitor	GFCI	Ground Fault Current Interrupter
Crosstalk	Leakage from one channel to another		

Heat sink	An object that absorbs and dissipates heat from another object using thermal contact	Pole	(Switches) Set of contacts that belong into a single circuit
H _{FE} , h _{FE}	Current gain (transistor's)	Pot (s.)	Slang term for potentiometer
Hiss	A category of noise, properly called white noise	Pot (v.)	To encapsulate with epoxy resin
HPF	High Pass Filter	PP	Push-pull
HT	High Tension (voltage)	PSRR	Power Supply Rejection Ratio
Hum	Audible oscillation at the frequency of mains supply	PSU	Power Supply Unit
IC	Integrated Circuit	PTC	Positive Thermal Coefficiency
IGBT	Isolated Gate Bipolar Transistor	PTP	Point-To-Point
Impedance	Resistance to AC signals	PWM	Pulse-width modulation
Inrush current	Current surge in mains that occurs when the device is switched on	Rail	A single voltage provided by power supply unit
LDR	Light Dependent Resistor	RCA	Radio Corporation of America
LED	Light Emitting Diode	RCD	Residual Current Device
LFO	Low Frequency Oscillator	Rectifier	A circuit or device that converts AC to DC.
LPF	Low Pass Filter	RF	Radio Frequency
LTP	Long-tailed Pair	Rheostat	An adjustable resistor (potentiometer)
MOSFET	Metal Oxide Field Effect Transistor	RMS	Root Mean Square: DC equivalent for AC voltage or current
MSK	Minimum-Shift Keying also known as FSK	RTA	Real Time Analyzer
NC	Normally Closed	Sag	Voltage drop under loading
NO	Normally Open	Schematic	A circuit diagram
NPO	Negative-Positive-Zero (negative and positive tempco:s are zero).	SCR	Silicon Controlled Rectifier (Thyristor)
NTC	Negative Thermal Coefficiency	SE	Single Ended
OLG	Open Loop Gain	Shunt	Connect to parallel i.e. present an alternative path
OpAmp	Operational amplifier	SIP	Single-in-line Package
OPS	Output Stage	Slew rate	The maximum rate of change of an amplifier's output voltage with respect to its input voltage
OT	Output transformer	SLS	Simulated Loudspeaker (a dummy load that emulates loudspeaker's characteristics)
OTA	Operational Transconductance Amplifier	SMD	Surface-Mount Device
OTL	Output Transformerless	SMT	Surface-Mount Technology
PA	(1) Public Address (2) Power Amplifier	SMPS	Switched Mode Power Supply
PCB	Printed Circuit Board	SNR	Signal to Noise Ratio
PCM	Pulse-Code modulation	S/N	Signal to Noise Ratio
PDM	Pulse-density modulation / frequency modulation	Snubber	A circuit that suppresses electrical transients
Pickup	A transducer that converts the vibration of either guitar's strings or its body to an electric signal	SOA	Safety Operating Area
PLL	Phase-Locked Loop		
PMPO	Peak Musical Power		

SPICE	Simulation Program with IC Emphasis (a software to simulate electronic circuits)
Spring reverb	A set of transducers that convert an electric signal to vibration of springs and the vibration back to electricity
SPL	Sound Pressure Level
SS	Solid-State
SSR	Solid-State Relay
TC	Temperature coefficient
tempco	Temperature coefficient
THD	Total Harmonic Distortion
Throw	(Switches) One of two or more positions a switch can adopt.
TIM	Transient Intermodulation Distortion
Tonestack	Circuitry consisting of stacked tone control circuits
Transducer	A device that converts one type of energy to another
Trimpot	A trimmer potentiometer
TRS	Tip-Ring-Sleeve, refers to a 1/4-inch stereo phone jack
T/S	Thiele-Small
UJT	Uni Junction Transistor
UL	Underwriter's Laboratories
Variac	A variable AC transformer
VAS	Voltage Amplifier Stage
V_F	Forward voltage
VFB	Voltage Feedback
WV	Working Voltage (capacitors)
XFMR	Transformer
Z_{IN}	Input impedance
Z_{OUT}	Output impedance