

P. Starič, E. Margan:

Wideband Amplifiers

Part 5:

System Synthesis And Integration

Any sufficiently advanced technology is indistinguishable from magic.

Arthur C. Clarke

(Profiles of the Future: An Inquiry
into the Limits of the Possible, 1973)

(blank page)

Contents	5.3
List of Figures	5.4
List of Tables	5.6

Contents:

5.0 'The product is greater than the Sum of its parts'	5.7
5.1 Geometrical Synthesis of Inductively Compensated Multi-stage Amplifiers — A Simple Example	5.9
5.2 High Input Impedance Selectable Attenuator with a JFET Source Follower	5.25
5.2.1 Attenuator High Frequency Compensation	5.26
5.2.2 Attenuator Inductance Loops	5.37
5.2.3 The 'Hook Effect'	5.41
5.2.4 Improving the JFET Source Follower DC Stability	5.43
5.2.5 Overdrive Recovery	5.47
5.2.6 Source Follower with MOSFETs	5.48
5.2.7 Input Protection Network	5.51
5.2.8 Driving the Low Impedance Attenuator	5.53
5.3 High Speed Operational Amplifiers	5.57
5.3.1 The Classical Opamp	5.57
5.3.2 Slew Rate Limiting	5.61
5.3.3 Current Feedback Amplifiers	5.62
5.3.4 Influence of a Finite Inverting Input Resistance	5.69
5.3.5 Noise Gain and Amplifier Stability Analysis	5.72
5.3.6 Feedback Controlled Gain Peaking	5.79
5.3.7 Improved Voltage Feedback Amplifiers	5.80
5.3.8 Compensating Capacitive Loads	5.83
5.3.9 Fast Overdrive recovery	5.89
5.4 Improving Amplifier Linearity	5.93
5.4.1 Feedback and Feedforward Error Correction	5.94
5.4.2 Error Reduction Analysis	5.98
5.4.3 Alternative Feedforward Configurations	5.100
5.4.4 Time Delay Compensation	5.103
5.4.5 Circuits With Local Error Correction	5.104
5.4.6 The Tektronix M377 IC	5.117
5.4.7 The Gilbert Multiplier	5.123
Résumé of Part 5	5.129
References	5.131

List of Figures:

Fig. 5.1.1: A two-stage differential cascode 7-pole amplifier	5.10
Fig. 5.1.2: The 7 normalized Bessel–Thomson poles and their characteristic circles	5.11
Fig. 5.1.3: The 3-pole stage realization	5.14
Fig. 5.1.4: The 4-pole L+T-coil stage	5.15
Fig. 5.1.5: Polar plot of the actual 7 poles and the 2 real poles	5.18
Fig. 5.1.6: Frequency response comparison	5.19
Fig. 5.1.7: Step response comparison	5.19
Fig. 5.1.8: Two pairs of gain normalized step responses	5.20
Fig. 5.1.9: Two pairs of step responses with different gain	5.21
Fig. 5.1.10: Two pairs of step responses with similar gain	5.21
Fig. 5.1.11: The influence of the real pole on the step response	5.23
Fig. 5.1.12: Making the deflection plates in sections	5.24
Fig. 5.2.1: A typical conventional oscilloscope input section	5.26
Fig. 5.2.2: The 10:1 attenuator frequency compensation	5.26
Fig. 5.2.3: Attenuator frequency response (magnitude, phase and envelope delay)	5.31
Fig. 5.2.4: Attenuator step response	5.33
Fig. 5.2.5: Compensating the 50 Ω signal source impedance	5.34
Fig. 5.2.6: Switching the three attenuation paths (1:1, 10:1, 100:1)	5.35
Fig. 5.2.7: Attenuator with no direct path	5.36
Fig. 5.2.8: Inductances owed to circuit loops	5.37
Fig. 5.2.9: The attenuator and the JFET source follower inductance loop damping	5.40
Fig. 5.2.10: Step response of the circuit in Fig. 5.2.9	5.41
Fig. 5.2.11: The ‘Hook effect’	5.42
Fig. 5.2.12: Canceling the Hook effect in common PCB material	5.42
Fig. 5.2.13: Simple offset trimming of a JFET source follower	5.43
Fig. 5.2.14: Active DC correction loop	5.44
Fig. 5.2.15: The principle of separate low pass and high pass amplifiers	5.45
Fig. 5.2.16: Elimination of the high pass amplifier	5.46
Fig. 5.2.17: The error amplifier becomes an inverting integrator	5.46
Fig. 5.2.18: High amplitude step response non-linearity	5.47
Fig. 5.2.19: A typical n-channel JFET structure cross-section and circuit model	5.48
Fig. 5.2.20: A typical n-channel MOSFET cross-section and circuit model	5.50
Fig. 5.2.21: Input protection from electrostatic discharge	5.52
Fig. 5.2.22: Input protection for long term overdrive	5.52
Fig. 5.2.23: JFET source follower with a buffer	5.54
Fig. 5.2.24: The low impedance attenuator	5.55
Fig. 5.3.1: The classical opamp, simplified circuit	5.58
Fig. 5.3.2: Typical opamp open loop gain and phase compared to the closed loop gain	5.60
Fig. 5.3.3: Slew rate in an opamp with a current mirror	5.62
Fig. 5.3.4: A current feedback amplifier model derived from a voltage feedback opamp	5.62
Fig. 5.3.5: A fully complementary current feedback amplifier model	5.63
Fig. 5.3.6: Cross-section of the Complementary Bipolar Process	5.64
Fig. 5.3.7: The four components of C_T	5.64
Fig. 5.3.8: Current feedback model used for analysis	5.65
Fig. 5.3.9: Current on demand operation during the step response	5.68
Fig. 5.3.10: Comparison of gain vs. bandwidth for a conventional and a CF amplifier	5.68
Fig. 5.3.11: Non-zero inverting input resistance	5.69
Fig. 5.3.12: Actual CFB amplifier bandwidth owed to non-zero inverting input resistance	5.72
Fig. 5.3.13: VFB and CFB amplifiers with capacitive feedback	5.73
Fig. 5.3.14: Noise gain definition	5.73
Fig. 5.3.15: An arbitrary example of the phase–magnitude relationship	5.74
Fig. 5.3.16: VFB amplifier noise gain derived	5.75
Fig. 5.3.17: CFB amplifier noise impedance derived	5.76
Fig. 5.3.18: CFB amplifier and its noise impedance equivalent	5.77
Fig. 5.3.19: Functionally equivalent circuits using VF and CF amplifiers	5.78
Fig. 5.3.20: Single resistor bandwidth adjustment for CFB amps	5.79

Fig. 5.3.21: Frequency and step response of the amplifier in Fig. 5.3.20	5.79
Fig. 5.3.22: Improved VF amplifier using folded cascode	5.80
Fig. 5.3.23: Improved VF amplifier derived from a CFB amp	5.81
Fig. 5.3.24: The 'Quad Core' structure	5.82
Fig. 5.3.25: Output buffer stage with improved current handling	5.83
Fig. 5.3.26: Capacitive load adds a pole to the feedback loop	5.83
Fig. 5.3.27: Amplifier stability driving a capacitive load	5.86
Fig. 5.3.28: Capacitive load compensation by inductance	5.86
Fig. 5.3.29: Capacitive load compensation by separate feedback paths	5.87
Fig. 5.3.30: Adaptive capacitive load compensation	5.88
Fig. 5.3.31: Amplifier with feedback controlled output level clipping	5.89
Fig. 5.3.32: Amplifier with output level clipping using separate supplies	5.90
Fig. 5.3.33: CFB amplifier output level clipping at Z_T	5.90
Fig. 5.4.1: Amplifiers with feedback and feedforward error correction	5.94
Fig. 5.4.2: Closed loop frequency response of a feedback amplifier	5.95
Fig. 5.4.3: Optimized feedforward amplifier frequency response	5.98
Fig. 5.4.4: Grounded load feedforward amplifier	5.100
Fig. 5.4.5: 'Error take off' principle	5.102
Fig. 5.4.6: 'Current dumping' principle (Quad 405)	5.103
Fig. 5.4.7: Time delay compensation of the feedforward amplifier	5.104
Fig. 5.4.8: Simple differential amplifier	5.105
Fig. 5.4.9: DC transfer function of a differential amplifier	5.106
Fig. 5.4.10: Test set up used to compare different amplifier configurations	5.108
Fig. 5.4.11: Simple differential cascode amplifier used as the reference	5.108
Fig. 5.4.12: Frequency response of the reference amplifier	5.109
Fig. 5.4.13: Step response of the reference amplifier	5.109
Fig. 5.4.14: Improved Darlington	5.110
Fig. 5.4.15: Frequency response of the improved Darlington differential amplifier	5.110
Fig. 5.4.16: Step response of the improved Darlington differential amplifier	5.110
Fig. 5.4.17: Differential amplifier with feedforward error correction	5.111
Fig. 5.4.18: Differential amplifier with double error feedforward	5.111
Fig. 5.4.19: Frequency response of the differential amplifier with double feedforward	5.112
Fig. 5.4.20: Step response of the differential amplifier with double feedforward	5.112
Fig. 5.4.21: Cascomp (compensated cascode) amplifier	5.113
Fig. 5.4.22: Frequency response of the Cascomp amplifier	5.113
Fig. 5.4.23: Step response of the Cascomp amplifier	5.113
Fig. 5.4.24: Differential cascode amplifier with error feedback	5.114
Fig. 5.4.25: Modified error feedforward with direct error sensing and direct correction	5.114
Fig. 5.4.26: Cascomp evolution with feedback	5.115
Fig. 5.4.27: Frequency response of the Cascomp evolution	5.115
Fig. 5.4.28: Step response of the Cascomp evolution	5.115
Fig. 5.4.29: Differential cascode amplifier with output impedance compensator	5.116
Fig. 5.4.30: Frequency response of the amplifier with output impedance compensator	5.116
Fig. 5.4.31: Step response of the amplifier with output impedance compensator	5.116
Fig. 5.4.32: Basic wideband amplifier block of the M377 IC	5.118
Fig. 5.4.33: The basic M377 block as a compound transistor	5.119
Fig. 5.4.34: The M377 amplifier with fast overdrive recovery	5.120
Fig. 5.4.35: Simulation of the M377 amplifier frequency response	5.121
Fig. 5.4.36: Simulation of the M377 amplifier step response	5.121
Fig. 5.4.37: Gain switching in the M377 amplifier	5.122
Fig. 5.4.38: Fixed gain with $R-2R$ attenuator switching	5.122
Fig. 5.4.39: The Gilbert multiplier development	5.124
Fig. 5.4.40: DC transfer function of the Gilbert multiplier	5.126
Fig. 5.4.41: The Gilbert multiplier has almost constant bandwidth	5.126
Fig. 5.4.42: Another way of developing the Gilbert multiplier	5.127
Fig. 5.4.43: The four quadrant multiplier	5.127
Fig. 5.4.44: The two quadrant multiplier used in M377	5.128

List of Tables:

Table 5.1.1: Comparison of component values for different pole assignments	5.22
Table 5.3.1: Typical production parameters of the Complementary-Bipolar Process	5.64

5.0 'The Product Is Greater Than The Sum Of Its Parts'

... and that can be true in both the mathematical and technological sense! Well, in math, at least as long as we are dealing with numbers greater than two; but in technology the goal might not be so straightforward and neither are the means of achieving it.

Electronics engineering is a difficult and demanding job. Most electronics engineers pay attention to the input and output constraints imposed by the real world as a matter of course. Many will also silently nod their head when the components logistician tells them to use another part instead of the one originally specified, just because it's cheaper. A number of them will also agree to take it into account when the marketing manager tells them that the customer wants and expects from the product a feature which was not foreseen initially as one of the design goals. And almost all will shrug their shoulders when the chief director announces that the financial resources for their project have been cut low or even that the project has been canceled. But almost all will go mad when the mechanical engineer or the enclosure designer casually stops by and asks if a switch or a pot could be moved from the left side of the printed circuit board to the right. Fortunately for the electronics engineer, he has on his side the most powerful argument of all: "Yeah, maybe I could do that, but probably the final performance would suffer!" Electronics engineering is a difficult and demanding job, indeed.

In the past 50 years electronics engineers have been delivering miracles at an ever increasing rate. Not just the general public, but also other people involved within the electronics industry have become accustomed to this. In the mid 1980s no one ever asked you if something could be done; instead, the question was 'for how much and when?'. Today, no one asks 'how much' either—it has to be a bargain (between brothers) and it had better be ready yesterday!

How many of you could give a name or two of engineers who became rich or just famous in the last 50 or so years? Let us see: William R. Hewlett was famous, but that was probably due more to his name in the Hewlett-Packard firm's title and less to the actual recognition of his work by the general public. Then there is Ray Dolby, known for his noise reduction system. Gordon Moore of Intel is known for his 'law'. Errr... who else? Oh, yes, Bill Gates is both rich and famous, but he is more famous because he is so rich and far less for his own work!

Do you see what we mean? No doubt, many engineers have started a very profitable business based on their own original circuit ideas. True, most commercially successful products are the result of a team effort; still, the key solution is often born in the mind of a gifted individual. But, frankly speaking, is there a single engineer or scientist who could attract 50,000 delirious spectators to a stadium every Sunday? OK, 5,000? Maybe Einstein could have been able to do so, but even he was considered to be 'a bit nuts'. Well, there you have it, the offer-demand economy.

In this book we have tried to pay a humble tribute to a small number of amplifier designers. But our readers are engineers themselves, so we are preaching to the already converted. Certainly we are not going to influence, let alone reverse, any of those trends.

So, if tomorrow your boss tells you that he will be paying you 5% less, shrug your shoulders and go back to your drawing board. And do not forget to stop by the front-panel designer to tell him that he is asking too much!

(blank page)

5.1 Geometrical Synthesis of Inductively Compensated Multi-Stage Amplifiers — A Simple Example

The reader who has patiently followed the discussion presented in previous chapters is probably eager to see all that theory being put into practice.

Before jumping to some more complex amplifier circuits we shall give a relatively simple example of a two-stage differential cascode amplifier, by which we shall illustrate the actual system optimization procedure in some detail, using the previously developed principles in their full potential.

Since we want to grasp the ‘big picture’ we shall have to leave out some less important topics, such as negative input impedance compensation, cascode damping, etc.; these are important for the optimization of each particular stage which, once optimized, can be idealized to some extent. We have covered that extensively enough in [Part 3](#), so we shall not explicitly draw the associated components in the schematic diagram. But, at the end of our calculations, we shall briefly discuss the influence of those components to final circuit values.

A two-stage amplifier is a ‘minimum complexity’ system for which the multi-stage design principles still apply. To this we shall add a 3-pole T-coil and a 4-pole L+T-coil peaking networks, discussed in [Part 2](#), as loads to each stage, making a total of 7 poles. There is, however, an additional real pole, owed to the Q_1 input capacitance and the total input and signal source resistance. As we shall see later, this pole can be neglected if its distance from the complex plane’s origin is at least twice as large as that of the system real pole set by $-1/R_a C_a$.

Such an amplifier thus represents an elementary example in which everything that we have learned so far can be applied. The reader should, however, be aware that this is by no means the ideal or, worse still, the only possibility. At the end of our calculations, when we shall be able to assess the advantages and limitations offered by our initial choices at each stage, we shall examine a few possibilities of further improvement.

We shall start our calculations from the unavoidable stray capacitances and the desired total voltage gain. Then we shall apply an optimization process, which we like to refer to as the **geometrical synthesis**, by which we shall calculate all the remaining circuit components in such a way that the resulting system will conform to the 7-pole normalized Bessel–Thomson system. The only difference will be that the actual amplifier poles will be larger by a certain factor, proportional (but not equal) to the upper half power frequency ω_H . We have already met the geometrical synthesis in its basic form in [Part 2, Fig. 2.5.3](#) when we were discussing the 3-pole T-coil circuit. The name springs from the ability to calculate all the peaking network components from simple geometrical relations which involve the pole’s real and imaginary components, given, of course, the desired pole pattern and a few key component values which can either be chosen independently or set by other design requirements. Here we are going to see a generalization of those basic relations applied to the whole amplifier.

It must be admitted that the constant and real input impedance of the T-coil network is the main factor which allows us to assign so many poles to only two stages. A cascade of passive 2-pole sections could have been used, but those would load each

other and, as a result, the bandwidth extension factor would suffer. Another possibility would be to use an additional cascode stage to separate the last two peaking sections, but another active stage, whilst adding gain, also adds its own problems to be taken care of. It is, nevertheless, a perfectly valid option.

Let us now take a quick tour of the amplifier schematic, [Fig. 5.1.1](#). We have two differential cascode stages and two current sources, which set both the transistor's transconductance and the maximum current available to the load resistors, R_a and R_b . This limits the voltage range available to the CRT. Since the circuit is differential the total gain is a double of each half. The total DC gain is (approximately):

$$A_0 = 2 \frac{R_a}{R_{e1}} \cdot \frac{R_b}{R_{e3}} \quad (5.1.1)$$

The values of R_{e1} and R_{e3} set the required capacitive bypass, $C_{e1}/2$ and $C_{e3}/2$, to match the transistor's time constants. In turn, this sets the input capacitance at the base of Q_1 and Q_3 , to which we must add the inevitable C_{cb} and some strays.

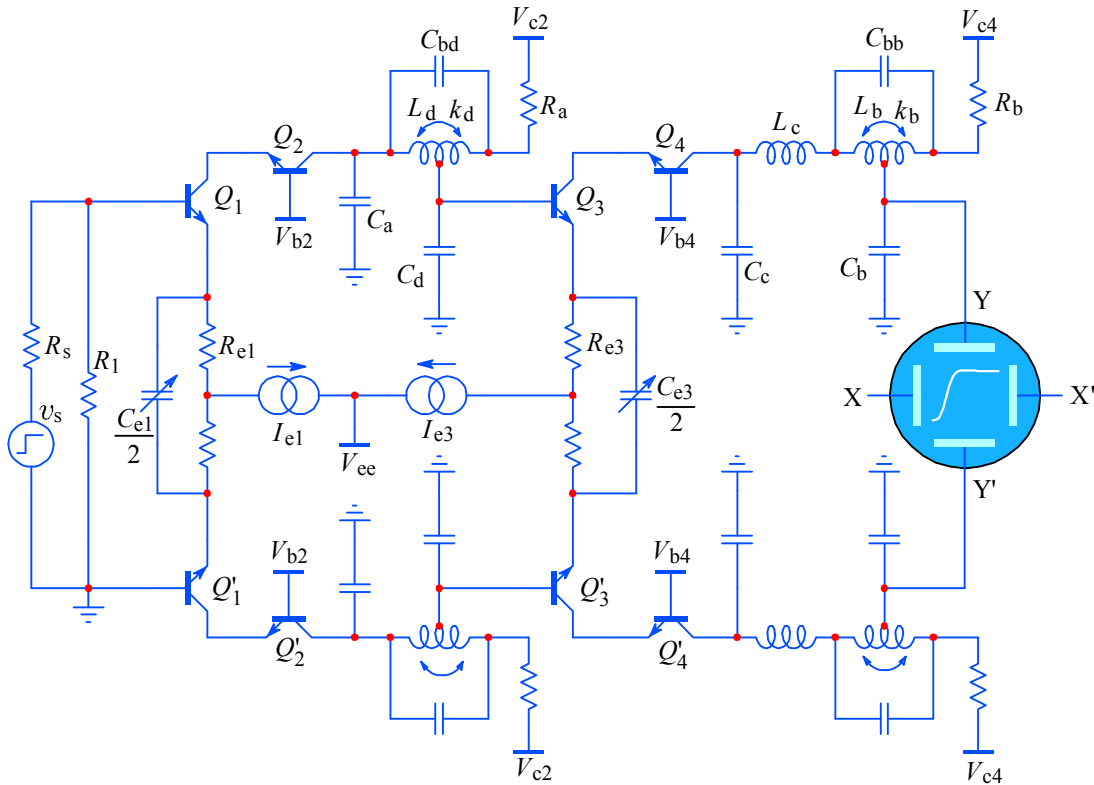


Fig. 5.1.1: A simple 2-stage differential cascode amplifier with a 7-pole peaking system: the 3-pole T-coil inter-stage peaking (between the Q_2 collector and the Q_3 base) and the 4-pole L+T-coil output peaking (between the Q_4 collector and the vertical plates of the cathode ray tube). The schematic was simplified to emphasize the important design aspects — see text.

The capacitance C_d should thus consist of, preferably, only the input capacitance at the base of Q_3 . If required by the coil 'tuning', a small capacitance can be added in parallel, but that would also reduce the bandwidth. Note that the associated T-coil L_d will have to be designed as an inter-stage peaking, as we have discussed in [Part 3, Sec. 3.6](#), but we can leave the necessary corrections for the end.

The capacitance C_b , owed almost entirely to the CRT vertical plates, is much larger than C_d , so we expect that R_a and R_b can not be equal. From this it follows that

it might be difficult to apply equal gain to each stage in accordance with the principle explained in [Part 4, Eq. 4.1.39](#). Nevertheless, the difference in gain will not be too high, as we shall see.

Like any other engineering process, geometrical synthesis also starts from some external boundary conditions which set the main design goal. In this case it is the CRT's vertical sensitivity and the available input voltage, from which the total gain is defined. The next condition is the choice of transistors by which the available current is defined. Both the CRT and the transistors set the lower limit of the loading capacitances at various nodes. From these the first circuit component R_b is fixed.

With R_b fixed we arrive at the first 'free' parameter, which can be represented by several circuit components. However, since we would like to maximize the bandwidth this parameter should be attributed to one of the capacitances. By comparing the design equations for the 3-pole T-coil and the 4-pole L+T-coil peaking networks in [Part 2](#), it can be deduced that C_a , the input capacitance of the 3-pole section, is the most critical component.

With these boundaries set let us assume the following component values:

$$\begin{aligned} C_b &= 11 \text{ pF} && (9 \text{ pF of the CRT vertical plates, } 2 \text{ pF stray}) \\ C_a &= 4 \text{ pF} && (3 \text{ pF from the } Q_2 \text{ } C_{cb}, 1 \text{ pF stray}) \\ R_b &= 360 \Omega && (\text{determined by the desired gain and the available current}) \end{aligned} \quad (5.1.2)$$

The pole pattern is, in general, another 'free' parameter, but for a smooth, minimum overshoot transient we must apply the Bessel–Thomson arrangement. As can be seen in [Fig. 5.1.2](#), each pole (pair) defines a circle going through the pole and the origin, with the center on the negative real axis.

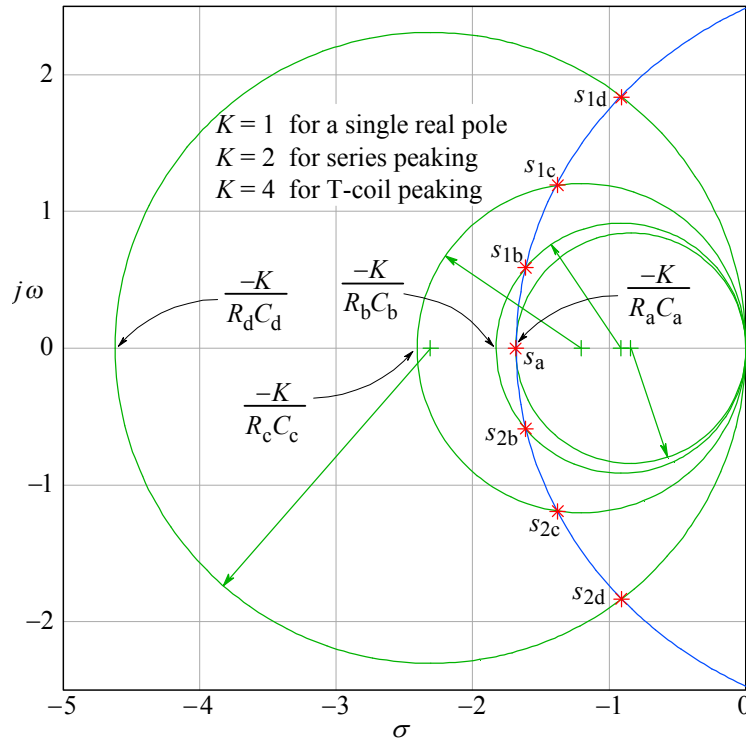


Fig. 5.1.2: The 7 normalized Bessel–Thomson poles. The characteristic circle of each pole (pair) has a diameter determined by the appropriate RC constant and the peaking factor K , which depends on the type of network chosen.

The poles in [Fig. 5.1.2](#) bear the index of the associated circuit components and the reader might wonder why we have chosen precisely that assignment.

In a general case the assignment of a pole (pair) to a particular circuit section is yet another ‘free’ design parameter. If we were designing a low frequency filter we could indeed have chosen an arbitrary assignment (as long as each complex conjugate pole pair is assigned as a pair, a limitation owed to physics, instead of circuit theory).

If, however, the bandwidth is an issue then we must seek those nodes with the largest capacitances and apply the poles with the lowest imaginary part to those circuit sections. This is because the capacitor impedance (which is dominantly imaginary) is inversely proportional both to the capacitor value and the signal frequency.

In this light the largest capacitance is at the CRT, that is, C_b ; thus the pole pair with the lowest imaginary part is assigned to the output T-coil section, formed by L_b and R_b , therefore acquiring the index ‘b’, s_{1b} and s_{2b} .

The real pole is the one associated with the 3-pole stage and there it is set by the loading resistor R_a and the input capacitance C_a , becoming s_a .

The remaining two pole pairs should be assigned so that the pair with the larger imaginary part is applied to that peaking network which has a larger bandwidth improvement factor. Here we must consider that $K = 4$ for a T-coil, whilst $K = 2$ for the series peaking L-section (of the 4-pole L+T-section). Clearly the pole pair with the larger imaginary part should be assigned to the inter-stage T-coil, L_d , thus they are labeled s_{1d} and s_{2d} . The L-section then receives the remaining pair, s_{1c} and s_{2c} .

We have thus arrived at a solution which seems logical, but in order to be sure that we have made the right choice we should check other combinations as well. We are going to do so at the end of the design process.

The poles for the normalized 7th-order Bessel–Thomson system, as taken either from [Part 4, Table 4.4.3](#), or by using the [BESTAP \(Part 6\)](#) routine, along with the associated angles, are:

$$\begin{aligned}
 s_a = \sigma_a &= -4.9718 & \theta_a &= 180^\circ \\
 s_b = \sigma_b \pm j\omega_b &= -4.7583 \pm j1.7393 & \theta_b &= 180^\circ \mp 20.0787^\circ \\
 s_c = \sigma_c \pm j\omega_c &= -4.0701 \pm j3.5172 & \theta_c &= 180^\circ \mp 40.8316^\circ \\
 s_d = \sigma_d \pm j\omega_d &= -2.6857 \pm j5.4207 & \theta_d &= 180^\circ \mp 63.6439^\circ
 \end{aligned} \tag{5.1.3}$$

So, let us now express the basic design equations by the assigned poles and the components of the two peaking networks.

For the real pole s_a we have the following familiar proportionality:

$$s_a = \sigma_a = D_a = -4.9718 \propto \frac{-1}{R_a C_a} \tag{5.1.4}$$

At the output T-coil section, according to [Part 2, Fig. 2.5.3](#), we have:

$$D_b = \frac{\sigma_b}{\cos^2 \theta_b} = \frac{-4.7583}{0.8821} = -5.3941 \propto \frac{-4}{R_b C_b} \tag{5.1.5}$$

For the L-section of the L+T output network, because the T-coil input impedance is equal to the loading resistor, we have:

$$D_c = \frac{\sigma_c}{\cos^2 \theta_c} = \frac{-4.0701}{0.5725} = -7.1094 \propto \frac{-2}{R_b C_c} \quad (5.1.6)$$

And finally, for the inter-stage T-coil network:

$$D_d = \frac{\sigma_d}{\cos^2 \theta_d} = \frac{-2.6857}{0.1917} = -13.6333 \propto \frac{-4}{R_a C_d} \quad (5.1.7)$$

From these relations we can calculate the required values of the remaining capacitances, C_c and C_d . If we divide [Eq. 5.1.5](#) by [Eq. 5.1.6](#), we have the ratio:

$$\frac{D_b}{D_c} = \frac{-\frac{4}{R_b C_b}}{-\frac{2}{R_b C_c}} = \frac{2 C_c}{C_b} \quad (5.1.8)$$

It follows that the capacitance C_c should be:

$$C_c = \frac{C_b}{2} \cdot \frac{D_b}{D_c} = \frac{11}{2} \cdot \frac{-5.3941}{-7.1094} = 4.1730 \text{ pF} \quad (5.1.9)$$

Likewise, if we divide [Eq. 5.1.4](#) by [Eq. 5.1.7](#), we obtain:

$$\frac{D_a}{D_d} = \frac{-\frac{1}{R_a C_a}}{-\frac{4}{R_a C_d}} = \frac{C_d}{4 C_a} \quad (5.1.10)$$

Thus C_d will be:

$$C_d = 4 C_a \frac{D_a}{D_d} = 4 \cdot 4 \cdot \frac{-4.9718}{-13.6333} = 5.8349 \text{ pF} \quad (5.1.11)$$

Of course, for most practical purposes, the capacitances do not need to be calculated to such precision, a resolution of 0.1 pF should be more than enough. But we would like to check our procedure by recalculating the actual poles from circuit components and for that purpose we shall need this precision.

Now we need to know the value of R_a . This can be readily calculated from the ratio D_a/D_b :

$$\frac{D_a}{D_b} = \frac{-\frac{1}{R_a C_a}}{-\frac{4}{R_b C_b}} = \frac{R_b C_b}{4 R_a C_a} \quad (5.1.12)$$

resulting in:

$$R_a = \frac{R_b}{4} \cdot \frac{C_b}{C_a} \cdot \frac{D_b}{D_a} = \frac{360}{4} \cdot \frac{11}{4} \cdot \frac{-5.3941}{-4.9718} = 268.5 \Omega \quad (5.1.13)$$

We are now ready to calculate the inductances L_b , L_c and L_d . For the two T-coils we can use the [Eq. 2.4.19](#):

$$L_b = R_b^2 C_b = 360^2 \cdot 11 \cdot 10^{-12} = 1.4256 \mu\text{H} \quad (5.1.14)$$

and

$$L_d = R_a^2 C_d = 268.5^2 \cdot 5.8349 \cdot 10^{-12} = 0.4206 \mu\text{H} \quad (5.1.15)$$

For L_c we use [Eq. 2.2.26](#) to obtain the proportionality factor of the RC constant:

$$L_c = \frac{1 + \tan^2 \theta_b}{4} R_b^2 C_c = \frac{360^2 \cdot 4.1730 \cdot 10^{-12}}{4 \cdot 0.8821} = 0.1533 \mu\text{H} \quad (5.1.16)$$

The magnetic coupling factors for the two T-coils are calculated by [Eq. 2.4.36](#):

$$k_b = \frac{3 - \tan^2 \theta_b}{5 + \tan^2 \theta_b} = \frac{3 - 0.1336}{5 + 0.1336} = 0.5584 \quad (5.1.17)$$

and likewise:

$$k_d = \frac{3 - \tan^2 \theta_d}{5 + \tan^2 \theta_d} = \frac{3 - 4.0738}{5 + 4.738} = -0.1183 \quad (5.1.18)$$

Note that k_d is negative. This means that, instead of the usually negative mutual inductance, we need a positive inductance at the T-coil tap. This can be achieved by simply mounting the two halves of L_d perpendicular to each other, in order to have zero magnetic coupling and then introduce an additional coil, L_e (again perpendicular to both halves of L_d), with a value of the required positive mutual inductance, as can be seen in [Fig. 5.1.3](#). Another possibility would be to wind the two halves of L_d in opposite direction, but then the bridge capacitance C_{bd} might be difficult to realize correctly.

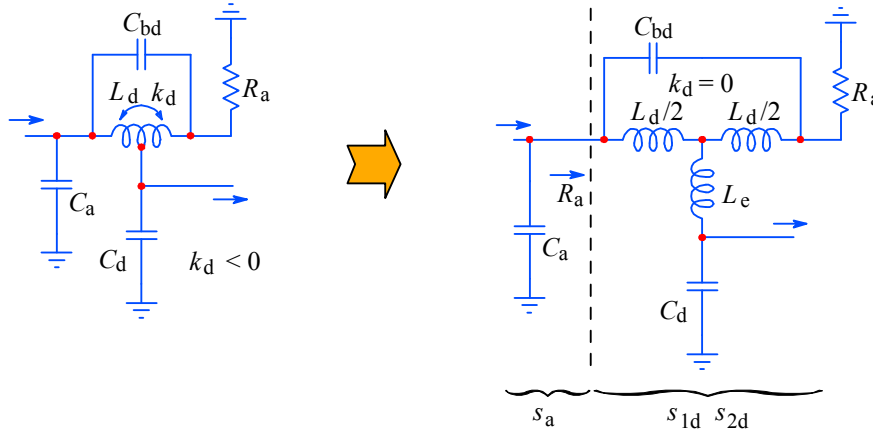


Fig. 5.1.3: With the assigned poles and the resulting particular component values the 3-pole stage magnetic coupling k_d needs to be negative, which forces us to use non-coupled coils and add a positive mutual inductance L_e . Even with a negative k_d the T-coil reflects its resistive load to the network input, greatly simplifying the calculations of component values.

The additional inductance L_e can be calculated from the required mutual inductance given by the negative value of k_d . In [Part 2, Eq. 2.4.1–2.4.5](#) we have defined the T-coil inductance, its two halves, and its mutual inductance by the relations repeated in [Eq. 5.1.19](#) for convenience:

$$\begin{aligned}
L &= L_1 + L_2 + 2 L_M \\
L_1 &= L_2 = \frac{L}{2(1+k)} \\
L_M &= -k \sqrt{L_1 L_2}
\end{aligned} \tag{5.1.19}$$

Thus, if $k = 0$ we have:

$$L_{1d} = L_{2d} = \frac{L_d}{2} = \frac{0.4206}{2} = 0.2103 \mu\text{H} \tag{5.1.20}$$

and:

$$L_e = -k_d \sqrt{\frac{L_d}{2} \cdot \frac{L_d}{2}} = -k_d \frac{L_d}{2} = 0.1183 \frac{0.4206}{2} = 0.025 \mu\text{H} \tag{5.1.21}$$

If we were to account for the Q_3 base resistance (discussed in [Part 3, Sec. 3.6](#)) we would get k_d even more negative and also $L_{1d} \neq L_{2d}$.

The coupling factor k_b , although positive, also poses a problem: since it is greater than 0.5 it might be difficult to realize. As can be noted from the above equations, the value of k depends only on the pole's angle θ . In fact, the 2nd-order Bessel system has the pole angles of $\pm 150^\circ$, resulting in a $k = 0.5$, representing the limiting case of realizability with conventionally wound coils. Special shapes, coil overlapping, or other exotic techniques may solve the coupling problem, but, more often than not, they will also impair the bridge capacitance. The other limiting case, when $k = 0$, is reached by the ratio $\Im\{s\}/\Re\{s\} = \sqrt{3}$, a situation occurring when the pole's angle $\theta = 120^\circ$.

In accordance with previous equations we also calculate the value of the two halves of L_b :

$$L_{1b} = L_{2b} = \frac{L_b}{2(1+k_b)} = \frac{1.4256}{2(1+0.5584)} = 0.4574 \mu\text{H} \tag{5.1.22}$$

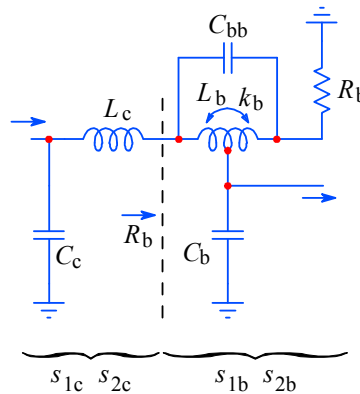


Fig. 5.1.4: The 4-pole output L+T-coil stage and its pole assignment.

The last components to be calculated are the bridge capacitances, C_{bb} and C_{bd} . The relation between the T-coil loading capacitance and the bridge capacitance has been given already in [Part 2, Eq. 2.4.31](#), from which we obtain the following expressions for C_{bb} and C_{bd} :

$$C_{bb} = C_b \frac{1 + \tan^2 \theta_b}{16} = 11 \frac{1 + 0.1336}{16} = 0.7793 \text{ pF} \quad (5.1.23)$$

and:

$$C_{bd} = C_d \frac{1 + \tan^2 \theta_d}{16} = 5.8349 \frac{1 + 4.0738}{16} = 1.8503 \text{ pF} \quad (5.1.24)$$

This completes the calculation of amplifier components necessary for the inductive peaking compensation and thus achieving the Bessel–Thomson system response. We would now like to verify the design by recalculating the actual pole values. To do this we return to the relations which we have started from, [Eq. 5.1.3](#) to [Eq. 5.1.7](#) and for the imaginary part using the relations in [Part 2, Fig. 2.5.3](#). In order not to confuse the actual pole values with the normalized values, from which we started, we add an index ‘A’ to the actual poles:

$$\sigma_{aA} = -\frac{1}{R_a C_a} = -\frac{1}{268.5 \cdot 4 \cdot 10^{-12}} = -931.1 \cdot 10^6 \text{ rad/s} \quad (5.1.25)$$

$$\sigma_{bA} = -\frac{4 \cos^2 \theta_b}{R_b C_b} = -\frac{4 \cdot 0.8821}{360 \cdot 11 \cdot 10^{-12}} = -891.0 \cdot 10^6 \text{ rad/s}$$

$$\omega_{bA} = \pm \frac{4 \cos \theta_b \sin \theta_b}{R_b C_b} = \pm \frac{4 \cdot 0.9392 \cdot 0.3433}{360 \cdot 11 \cdot 10^{-12}} = \pm 325.7 \cdot 10^6 \text{ rad/s}$$

$$\sigma_{cA} = -\frac{2 \cos^2 \theta_c}{R_b C_c} = -\frac{2 \cdot 0.5725}{360 \cdot 4.1730 \cdot 10^{-12}} = -762.2 \cdot 10^6 \text{ rad/s}$$

$$\omega_{cA} = \pm \frac{2 \cos \theta_c \sin \theta_c}{R_b C_c} = \pm \frac{2 \cdot 0.7566 \cdot 0.6538}{360 \cdot 4.1730 \cdot 10^{-12}} = \pm 658.5 \cdot 10^6 \text{ rad/s}$$

$$\sigma_{dA} = -\frac{4 \cos^2 \theta_d}{R_a C_d} = -\frac{4 \cdot 0.1917}{268.5 \cdot 5.8349 \cdot 10^{-12}} = -489.5 \cdot 10^6 \text{ rad/s}$$

$$\omega_{dA} = \pm \frac{4 \cos \theta_d \sin \theta_d}{R_a C_d} = \pm \frac{4 \cdot 0.4439 \cdot 0.8961}{268.5 \cdot 5.8349 \cdot 10^{-12}} = \pm 1015.6 \cdot 10^6 \text{ rad/s}$$

If we divide the real amplifier pole by the real normalized pole, we get:

$$\frac{\sigma_{bA}}{\sigma_b} = \frac{-931.1 \cdot 10^6}{-4.9718} = 187.3 \cdot 10^6 \quad (5.1.26)$$

and this factor is equal for all other pole components. Unfortunately, from this we cannot calculate the upper half power frequency of the amplifier. The only way to do that (for a Bessel system) is to calculate the response for a range of frequencies around the cut off and then iterate it using the bisection method, until a satisfactory tolerance has been achieved.

Instead of doing it for only a small range of frequencies we shall, rather, do it for a three decade range and compare the resulting response with the one we would get from a non-compensated amplifier (in which all the inductances are zero). Since to this point we were not interested in the actual value of the voltage gain, we shall make the comparison using amplitude normalized responses.

The non-compensated amplifier has two real poles, which are:

$$s_{1N} = -\frac{1}{R_a(C_a + C_d)} \quad \text{and} \quad s_{2N} = -\frac{1}{R_b(C_b + C_c)} \quad (5.1.27)$$

Consequently, its complex frequency response would then be:

$$F_N(s) = \frac{s_{1N} s_{2N}}{(s - s_{1N})(s - s_{2N})} \quad (5.1.28)$$

with the magnitude:

$$|F_N(\omega)| = \frac{\sqrt{s_{1N} s_{2N}}}{\sqrt{(\omega^2 - s_{1N}^2)(\omega^2 - s_{2N}^2)}} \quad (5.1.29)$$

and the step response:

$$\begin{aligned} g(t) &= \mathcal{L}^{-1} \left\{ \frac{s_{1N} s_{2N}}{s(s - s_{1N})(s - s_{2N})} \right\} \\ &= 1 + \frac{s_{2N}}{s_{1N} - s_{2N}} e^{s_{1N}t} - \frac{s_{1N}}{s_{1N} - s_{2N}} e^{s_{2N}t} \end{aligned} \quad (5.1.30)$$

The rise time is:

$$\tau_r = 2.2 \sqrt{\frac{1}{s_{1N}^2} + \frac{1}{s_{2N}^2}} \quad (5.1.31)$$

and the half power frequency:

$$f_h = \frac{\sqrt{s_{1N} s_{2N}}}{2\pi} \quad (5.1.32)$$

In contrast, the complex frequency response of the 7-pole amplifier is:

$$F_A(s) = A_0 \frac{-s_{aA} s_{1bA} s_{2bA} s_{1cA} s_{2cA} s_{1dA} s_{2cA}}{(s - s_{aA})(s - s_{1bA})(s - s_{2bA})(s - s_{1cA})(s - s_{2cA})(s - s_{1dA})(s - s_{2dA})} \quad (5.1.33)$$

and the step response is the inverse Laplace transform of the product of $F_A(s)$ with the unit step operator $1/s$:

$$g(t) = \mathcal{L}^{-1} \left\{ \frac{1}{s} F_A(s) \right\} = \sum \text{res} \left(\frac{1}{s} F_A(s) e^{st} \right) \quad (5.1.34)$$

We shall not attempt to solve either of these functions analytically, since it would take too much space, and, anyway, we have solved them separately for its two parts (3rd- and 4th-order) in [Part 2](#). Because the systems are separated by an amplifier (Q_3, Q_4), the frequency response would be a simple multiplication of the two responses. For the step response we now have 8 residues to sum (7 of the system poles, in addition to the one from the unit step operator). Although lengthy, it is a relatively simple operation and we leave it as an exercise to the reader. Instead we are going to use the computer routines, the development of which can be found in [Part 6](#).

In [Fig. 5.1.5](#) we have made a polar plot of the poles for the inductively compensated 7-pole system and the non-compensated 2-pole system. As we have learned in [Part 1](#) and [Part 2](#), the farther from origin the smaller is the pole's influence on the system response. It is therefore obvious that the 2-pole system's response will be dominated by the pole closer to the origin and that is the pole of the output stage, s_{2N} . The bandwidth of the 7-pole system is, obviously, much larger.

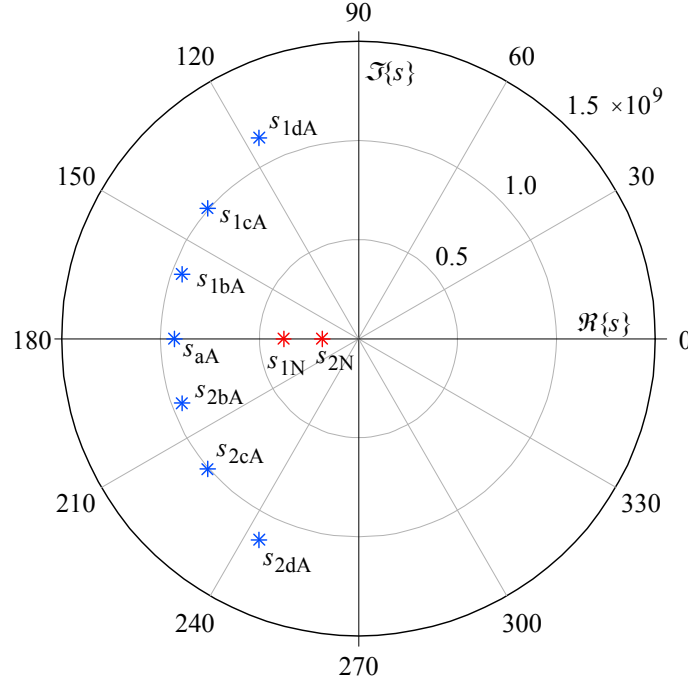


Fig. 5.1.5: The polar plot of the 7-pole compensated system (poles with index 'A') and the 2-pole non-compensated system (index 'N'). The radial scale is $\times 10^9$ rad/s. The angle is in degrees.

The pole layout gives us a convenient indication of the system's performance, but it is the magnitude vs. frequency response that reveals it clearly. As can be seen in [Fig. 5.1.6](#), the non-compensated system has a bandwidth of less than 25 MHz. The compensated amplifier bandwidth is close to 88 MHz, more than 3.5 times larger.

The comparison of step responses in [Fig. 5.1.7](#) reveals the difference in performance even more dramatically. The rise time of the non-compensated system is about 14 ns, whilst for the compensated system it is only 3.8 ns, also a factor of 3.5 times better; in addition, the overshoot is only 0.48 %.

Both comparisons show an impressive improvement in performance. But is it the best that could be obtained from this circuit configuration? After all, in [Part 2](#) we have seen a similar improvement from just the 4-pole L+T-coil section and we expect that the addition of the 3-pole section should yield a slightly better result at least.

One obvious way of extending the bandwidth would be to lower the value of R_b , increase the bias currents, and scale the remaining components accordingly. Then we should increase the input signal amplitude to get the same output. But this is the 'trivial' solution (mathematically, at least; not so when building an actual circuit).

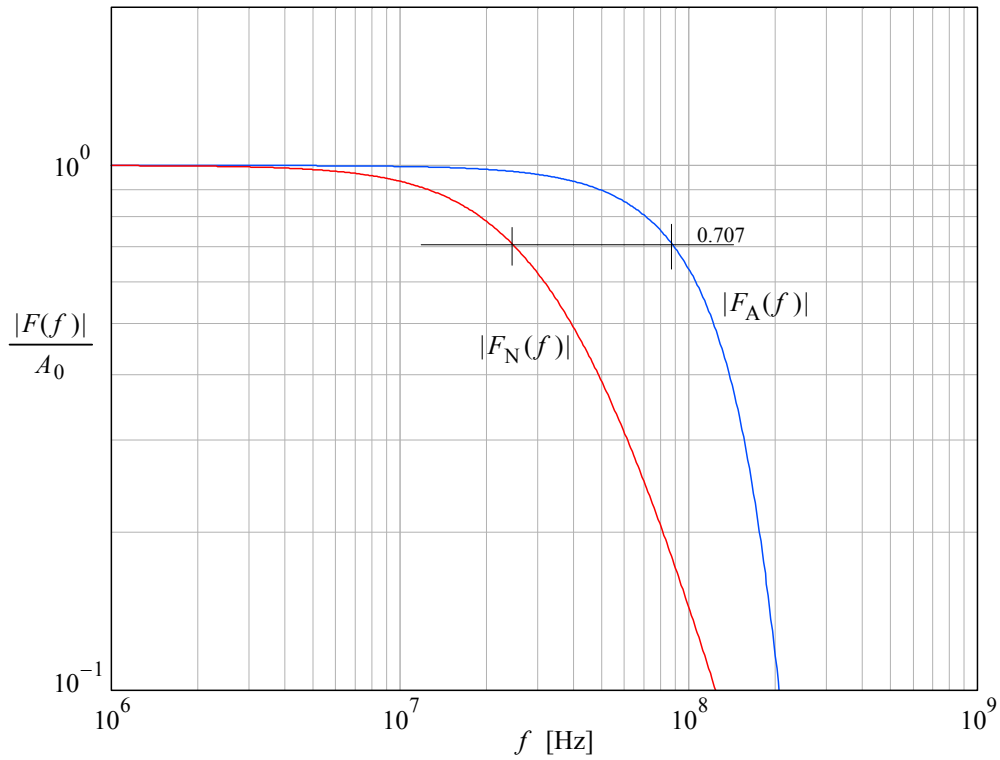


Fig. 5.1.6: The gain normalized magnitude vs. frequency of the 7-pole compensated system $|F_A(f)|$ and the 2-pole non-compensated system, $|F_N(f)|$. The bandwidth of F_N is about 25 MHz and the bandwidth of F_A is about 88 MHz, more than 3.5 times larger.

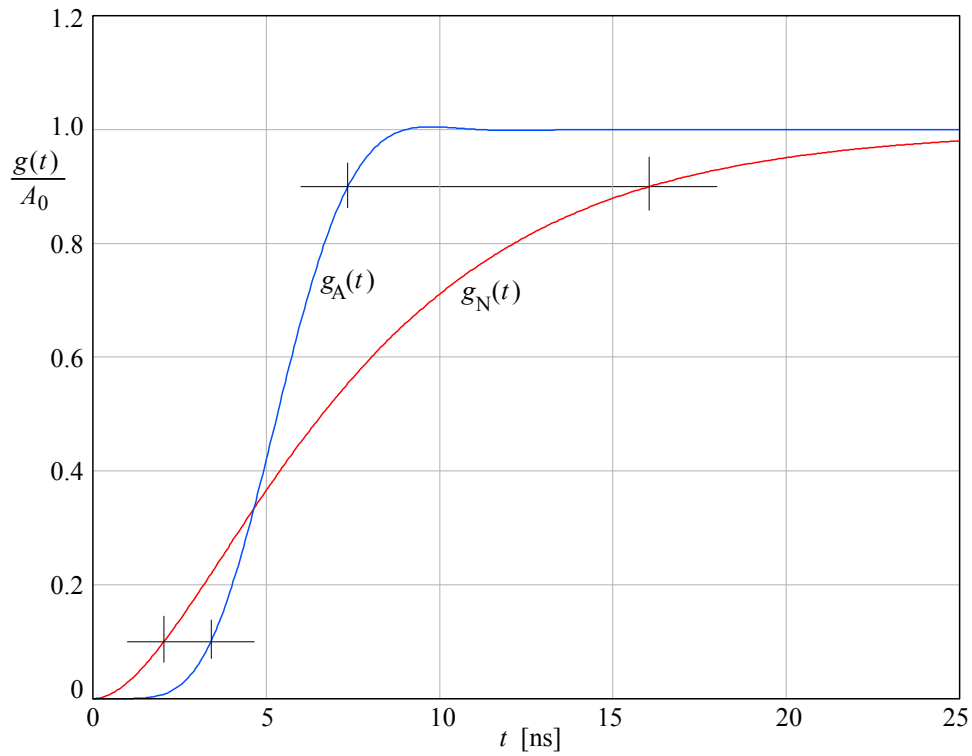


Fig. 5.1.7: The gain normalized step responses of the 7-pole compensated system $g_A(t)$ and the 2-pole non-compensated system $g_N(t)$. The rise time is 14 ns for the $g_N(t)$, but only 3.8 ns for $g_A(t)$. The overshoot of $g_A(t)$ is only 0.48 %.

By a careful inspection of the amplifier design equations and comparing them with the analysis of the two sections in [Part 2](#), we come to the conclusion that the most serious bandwidth drawback factor is the high value of the CRT capacitance, which is much higher than C_a or C_d . But if so, did we limit the possible improvement by assigning the poles with the lowest imaginary part to the output? Should not we obtain a better performance if we add more peaking to the output stage?

Since we have put the design equations and the response analysis into a computer routine, we can now investigate the effect of different pole assignments. To do so we simply reorder the poles and run the routine again. Besides the pole order that we have described, let us indicate it by the pole order: **abcd** ([Eq. 5.1.3](#)), we have five additional permutations: **abdc**, **acbd**, **adbc**, **acdb**, **adcb**. The last two permutations result in a rather slow system, requiring a large inductance for L_b and large capacitances C_c and C_d . But the remaining ones deserve a look.

In [Fig. 5.1.8](#) we have plotted the four normalized step responses and, because there are two identical pairs of responses, we have displaced them vertically by a small offset in order to distinguish them more clearly.

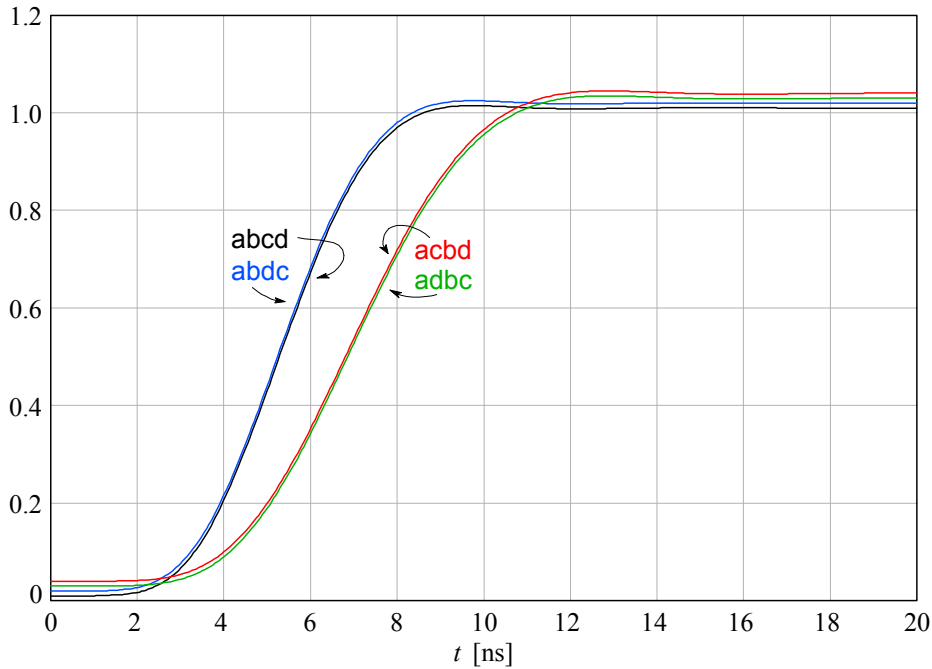


Fig. 5.1.8: The normalized step responses of the four possible combinations of pole assignments. There are two pairs of responses, here spaced vertically by a small offset to allow easier identification. One of the two faster responses (labeled ‘abcd’) is the one for which the detailed analysis has been given in the text.

If the pole pairs s_c and s_d are mutually exchanged the result is the same as our original analysis. But by exchanging s_b with either s_c or s_d the result is sub-optimal.

A closer look at [Table 5.1.1](#) reveals that both of the two slower responses have $R_a = 354 \Omega$ instead of 268Ω . The higher value of R_a means actually a higher gain, as can be seen in [Fig. 5.1.9](#), where the original system was set for a gain of $A_0 \approx 10$, in contrast with the higher value, $A_0 \approx 13$. The higher gain results from a different ‘tuning’ of the 3-pole T-coil stage, in accordance with the different pole assignment.

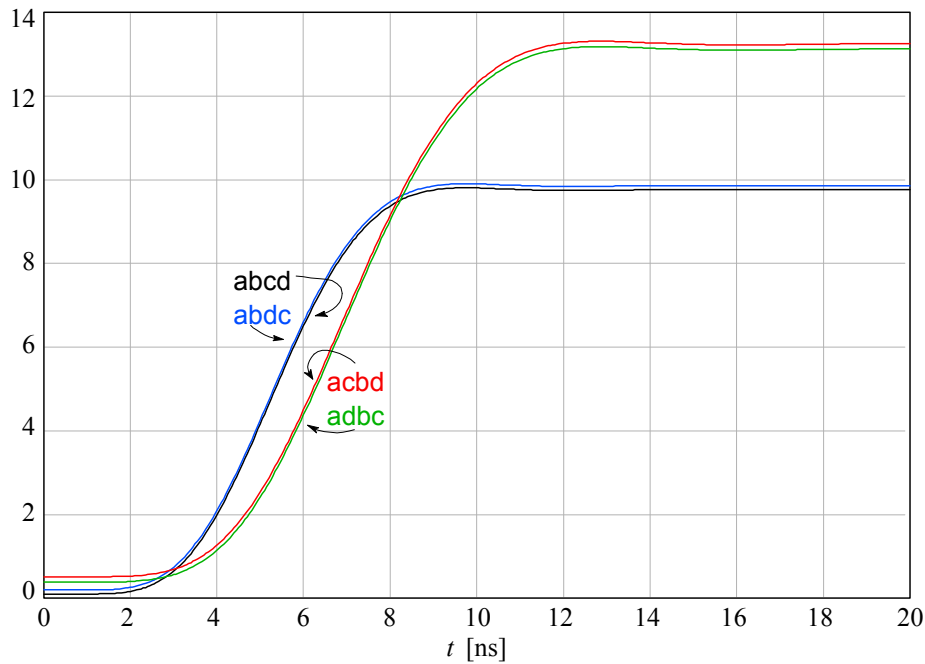


Fig. 5.1.9: The slower responses of Fig. 5.1.8, when plotted with the actual gain, are actually those with a higher value of R_a and therefore a higher gain.

Since our primary design goal is to maximize the bandwidth with a given gain, let us recalculate the slower system for a lower value of R_b . If $R_b = 316 \Omega$ (from the E96 series of standard values, 0.5 % tolerance), the gain is restored. Fig. 5.1.10 shows the recalculated responses, labeled 'acbd' and 'adbc', whilst the 'abcd' and 'abdc' responses are the same as before.

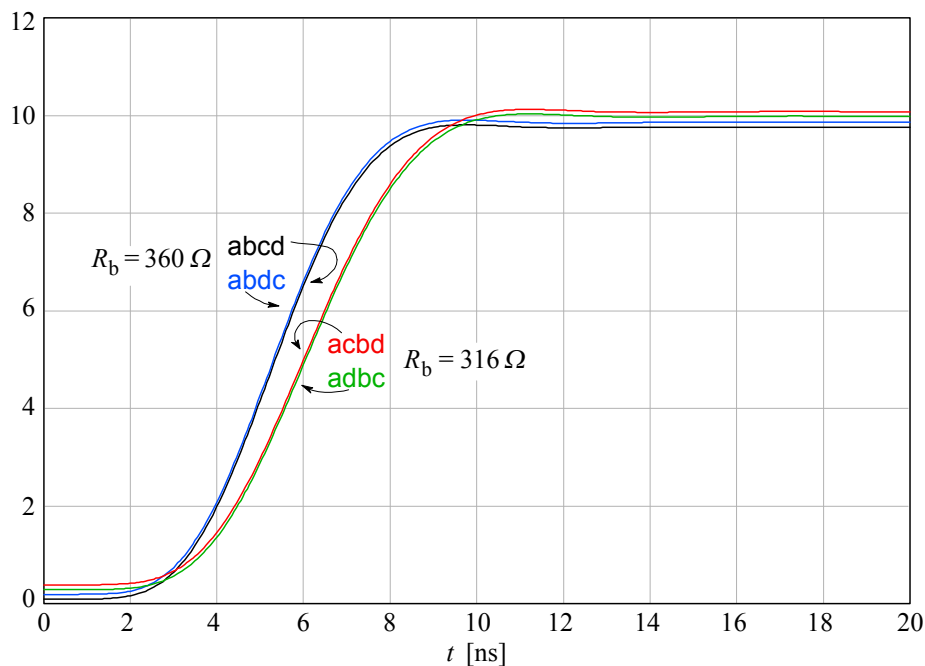


Fig. 5.1.10: If the high gain responses are recalculated by reducing R_b from the original 360Ω to 316Ω , the gain is nearly equal in all four cases. However, those pole assignments which put the poles with the higher imaginary part at the output stage still result in a slightly slower system.

The difference in rise time between the two pairs is much smaller now; however, the recalculated pair is still slightly slower. This shows that our initial assumptions of how to achieve maximum bandwidth (within a given configuration) were not guessed by sheer luck.

In [Table 5.1.1](#) we have collected all the design parameters for the four out of six possible pole assignments. The systems in the last two columns have the same pole assignments as in the middle two, but have been recalculated from a lower R_b value, in order to obtain the total voltage gain nearly equal to the first system. From a practical point of view the first and the last column are the most interesting: the system represented by the first column is the fastest (as the second one, but the latter is difficult to realize, mainly owing to low C_c value), whilst the last one is only slightly slower but much easier to realize, mainly owing to a lower magnetic coupling k_b and the non-problematic values of C_c and C_d .

Table 5.1.1

R_b [Ω]	360	360	360	360	316	316
pole order:	abcd	abdc	acbd	adbc	acbd	adbc
A_0	9.667	9.667	12.74	12.74	9.817	9.817
R_a [Ω]	268.5	286.5	353.9	353.9	310.7	310.7
C_c [pF]	4.173	2.177	2.870	7.249	2.870	7.249
C_d [pF]	5.838	11.19	14.75	5.838	14.75	5.838
C_{bb} [pF]	0.779	0.779	1.201	1.201	1.201	1.201
C_{bd} [pF]	1.851	1.222	1.045	1.851	1.045	1.851
L_b [μ H]	1.426	1.426	1.426	1.426	1.098	1.098
L_c [μ H]	0.153	0.080	0.162	0.410	0.125	0.316
L_d [μ H]	0.421	0.807	1.847	0.731	1.423	0.563
k_b	0.558	0.558	0.392	0.392	0.392	0.392
k_d	-0.118	0.392	0.558	-0.118	0.558	-0.118
η_b	3.57	3.57	2.35	2.35	2.84	2.84
η_t	3.55	3.55	2.33	2.33	2.81	2.81

Table 5.1.1: Circuit components for 4 of the 6 possible pole assignments. The last two columns represent the same pole assignment as the middle two, but have been recalculated for $R_b = 316 \Omega$ and nearly equal gain. The first column is the example calculated in the text and its response is one of the two fastest. The other fast system (second column) is probably non realizable (in discrete form), because $C_c \approx 2$ pF. The last column (adbc) is, on the other hand, only slightly slower, but probably much easier to realize (T-coil coupling and the capacitance values). The bandwidth and rise time improvement factors η_b and η_t were calculated by taking the non-compensated amplifier responses as the reference.

The main problem encountered in the realization of our original ‘abcd’ system is the relatively high magnetic coupling factor of the output T-coil, k_b . A possible way of improving this could be by applying a certain amount of emitter peaking to either the Q_1 or Q_3 emitter circuit. Then we would have a 9-pole system and we would have to recalculate everything. However, the use of emitter peaking results in a negative input impedance which has to be compensated (see [Part 3, Sec. 3.5](#)), and the compensating network adds more stray capacitance.

A 9-pole system might be more easily implemented if, instead of the 3-pole section, we were to use another L+T-coil 4-pole network. The real pole could then be provided by the signal source resistance and the Q_1 input capacitance, which we have chosen to neglect so far. With 9 poles both T-coils can be made to accommodate those two pole pairs with moderate imaginary part values (because the T-coil coupling factor depends only on the pole angle θ), so that the system bandwidth could be more easily maximized. A problem could arise with a low value of some capacitances, which might become difficult to achieve. But, as is evident from [Table 5.1.1](#), there are many possible variations (their number increases as the factorial of the number of poles), so a clever compromise can always be made. Of course, with a known signal source an additional inductive peaking could be applied at the input, resulting in a total of 11 or perhaps even 13 poles, but then the component tolerances and the adjustment precision would set the limits of realizability.

Finally, we would like to verify the initial claim that the input real pole s_1 , owed to the signal source resistance and base spread resistance and the total input capacitance, can be neglected if it is larger than the system real pole s_a . Since the input pole is separated from the rest of the system by the first cascode stage, it can be accounted for by simply multiplying the system transfer function by it. In the frequency response its influence is barely noticeable. In the step response, [Fig. 5.1.11](#), it affects mostly the envelope delay and the overshoot, whilst the rise time (in accordance with the frequency response) remains nearly the same.

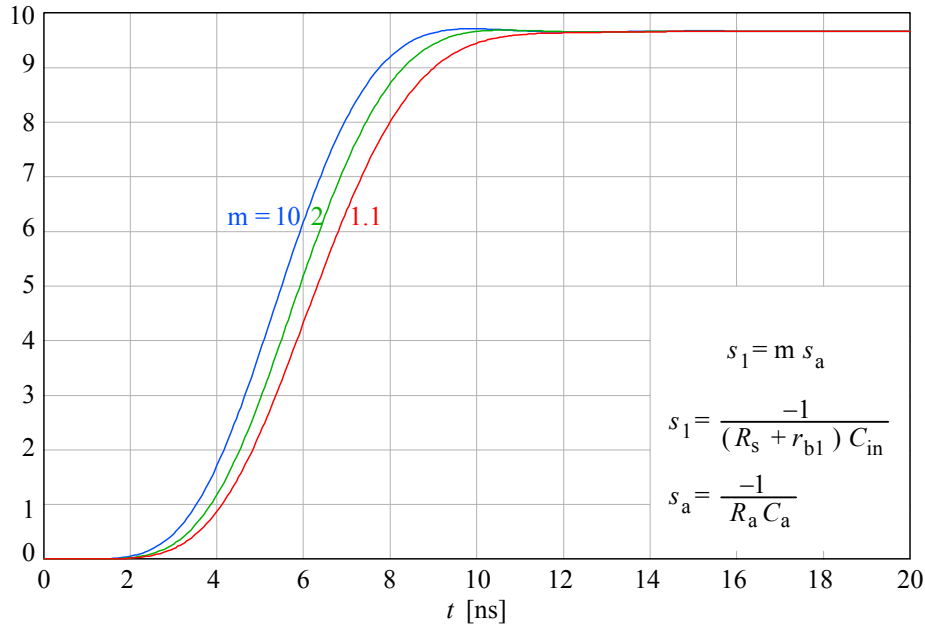


Fig. 5.1.11: If the real input pole s_1 is at least twice as large as the system's real pole s_a , its influence on the step response can be seen merely as an increased envelope delay and a reduced overshoot, whilst the rise time remains nearly identical.

Usually the signal source's impedance is $50\ \Omega$ or less; an input capacitance of several pF would still ensure that the input pole is high above the system real pole. However, an oscilloscope needs an input buffer stage and a preamplifier, with variable gain and attenuation, to adapt the signal amplitude to the required level. This preamplifier's output impedance, driving the input capacitance of our amplifier, can be high enough, forcing us to account for it. In such cases, as already stated before, it

might become feasible to replace the 3-pole peaking network by another 4-pole L+T-coil network and make the input pole the main system real pole.

As already mentioned, the relatively high capacitance of the CRT vertical deflection plates is the dominant cause for the amplifier bandwidth limitation.

To avoid this problem the most advanced CRTs from the analog 'scope era have had their deflection plates made in a number of sections, connected externally by a series of T-coils (see [Fig. 5.1.12](#)), thus reducing the capacitance seen by the amplifier to just a fraction of the original value. At the same time, the T-coils have provided a delay required to match the signal propagation to the electron velocity in the writing beam (compensating for the electron's finite travel time by the deflection plates, as well as some non negligible relativistic effects! — see [Appendix 5.1](#)), thus aiding to a better beam control.

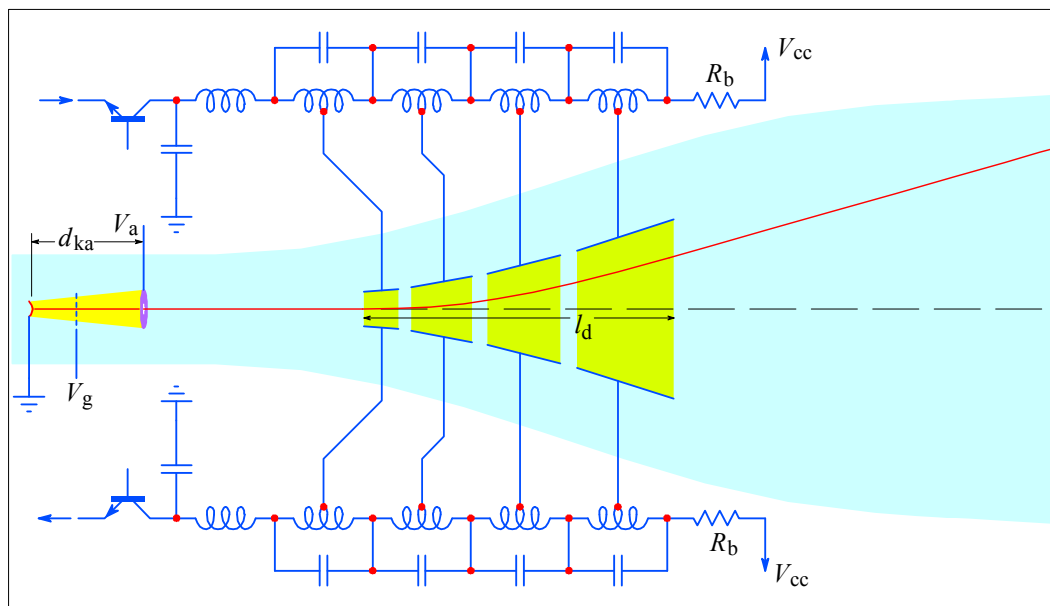


Fig. 5.1.12: If the CRT deflection plates are made in a number of sections (usually between 4 and 8), connected by a series of T-coil peaking circuits, the amplifier would effectively be loaded by a much smaller capacitance, allowing the system cut off frequency to be several times higher. The T-coils also provide the time delay necessary to keep the deflecting voltage (as seen by the electrons in the writing beam) almost constant throughout the electron's travel time across the deflecting field. For simplicity, only the vertical deflection system is shown, but a similar circuit could be used for the horizontal deflection, too (such an example can be found in the 1 GHz Tektronix 7104 model; see [Appendix 5.1](#) for further details). Note that, owing to the increasing distance between the plates their length should also vary accordingly, in order to compensate for the reduced capacitance. Fortunately, the capacitance is also a function of the plate's width, not just length and distance, so a well balanced compromise can always be found.

5.2 High Input Impedance Selectable Attenuator with a JFET Source Follower

A typical oscilloscope vertical input must incorporate a number of passive signal conditioning functions:

- 1) a selectable $50\ \Omega/1\ \text{M}\Omega$ input resistance, with low reflection coefficient on the $50\ \Omega$ setting;
- 2) a selectable DC–GND–AC coupling;
- 3) a selectable 1:1/10:1/100:1 or similar attenuation with a $1\ \text{M}\Omega \parallel 10\ \text{pF}$ input impedance (independent of the selected attenuation; resistance tolerance 0.1 %, capacitance between 10–20 pF, since the external probes are adjustable, but its dielectric properties must be constant with frequency and temperature);
- 4) a 2 kV electrostatic discharge spark gap, able to protect the input from a human body model discharge (200 pF, 15 kV);
- 5) a 400 V (DC + peak AC) continuous protection of the delicate input amplifier at no input attenuation setting (except for the $50\ \Omega$ setting).

To this we must add the following requirements:

- 6) the upper cut off frequency at least twice higher than the system's bandwidth;
- 7) the upper cut off frequency should be independent of any of the above settings;
- 8) the gain flatness must be kept within 0.5 % from DC to 1/5 of the bandwidth;
- 9) the protection diodes must survive repeating 1–2 A surge currents with $< 1\ \text{ns}$ rise and $50\ \mu\text{s}$ decay, their leakage must be $< 100\ \text{pA}$ and capacitance $< 1\ \text{pF}$;

To preserve a high degree of signal integrity the stray capacitances and inductances must be kept low throughout the input stage, which means small components with a small size of soldering pads and the traces as short as possible.

In addition, the unity gain JFET buffer stage performance should include:

- 10) a $> 100\ \text{M}\Omega$ input resistance;
- 11) a $< 1\ \text{pF}$ input capacitance;
- 12) a $50\ \Omega$ output resistance or the ability to drive such loads;
- 13) the bandwidth at least twice higher than the rest of the system;
- 14) $< 5\ \text{nV}/\sqrt{\text{Hz}}$ input noise density (see [Ref. 5.66] for low noise design);
- 15) $< 0.5\ \text{mV}_{\text{pp}}$ wideband noise (at $5\ \text{mV}/\text{div.}$; important for digitizing 'scopes);
- 16) gain close to unity, flat within 0.5 %, up to 1/5 of the system bandwidth;
- 17) low overshoot and undershoot, any resonance up to $10\times$ the system's bandwidth should be well damped to reduce ringing;
- 18) fast step settling time to within 0.1% of the final signal value;
- 19) recovery from input overdrive as short as possible;
- 20) ability to handle signals in a wide range, from $8\ \text{mV}_{\text{pp}}$ ($1\ \text{mV}/\text{div.}$) up to $40\ \text{V}_{\text{pp}}$ ($5\ \text{V}/\text{div.}$, with input attenuation), with a DC offset of $\pm 1/2$ screen at least;

This is an impressive list, indeed. Especially if we consider that for a 500 MHz system bandwidth the above requirements should be fulfilled for a 1 GHz bandwidth.

A typical input stage block diagram is shown in Fig. 5.2.1. The attenuator and the unity gain buffer stage will be analyzed in the following sections.

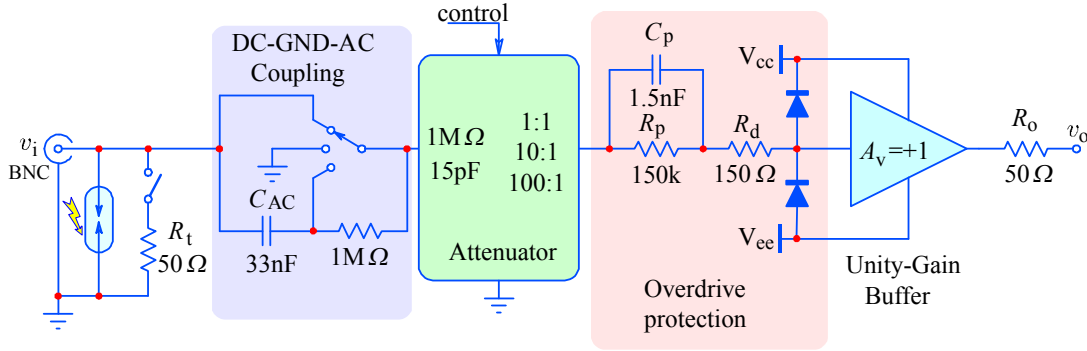


Fig. 5.2.1: A typical conventional oscilloscope input section. All the switches must be high voltage types, controlled either mechanically or as electromagnetic relays (but other solutions are also possible, as in [Ref. 5.2]). The spark gap protects against electrostatic discharge. The R_t 50Ω resistor is the optional transmission line termination. The $1\text{M}\Omega$ resistor in the DC–GND–AC selector charges the AC coupling capacitor in the GND position, reducing the overdrive shock through C_{ac} in presence of a large DC signal component. The attenuator is analyzed in detail in Sec. 5.2.1–3. The overdrive protection limits the input current in case of an accidental connection to the 240V_{ac} with the attenuator set to the highest sensitivity. The unity gain buffer/impedance transformer is a $>100\text{M}\Omega$ R_{in} , 50Ω R_o JFET or MOSFET source follower, analyzed in Sec. 5.2.4 and 5.2.5.

5.2.1 Attenuator High Frequency Compensation

A simple resistive attenuator, like the one in Fig. 5.2.2a is too sensitive to any capacitive loading by the following amplifier stage. For oscilloscopes the standard input impedance $R_a = R_1 + R_2$ is $1\text{M}\Omega$, so for a 10:1 attenuation the resistance values must be $R_1 = 900\text{k}\Omega$ and $R_2 = 100\text{k}\Omega$. With such values the output impedance, which equals the parallel connection of both resistances, would be about $90\text{k}\Omega$. Assuming an amplifier input capacitance of only 1pF , the resulting system bandwidth would be only 1.77MHz [$f_h = (R_1 + R_2)/(2\pi C_i R_1 R_2)$].

Therefore high frequency compensation, as shown in Fig. 5.2.2b, is necessary if we want to obtain higher bandwidth. The frequency compensation, however, lowers the input impedance at high frequencies.

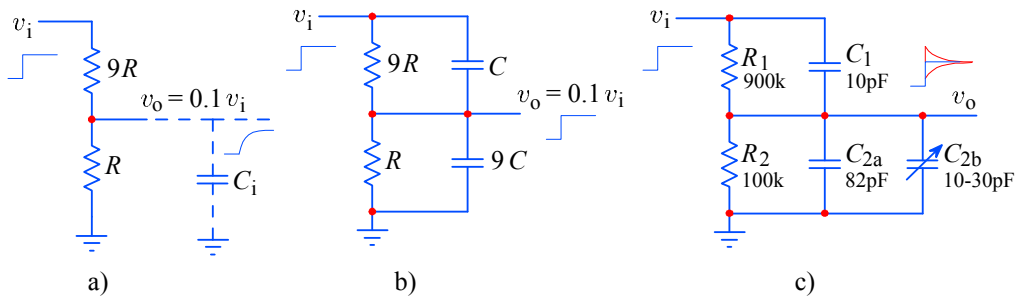


Fig. 5.2.2: The 10:1 attenuator; **a) resistive:** with $R = 100\text{k}\Omega$, the following stage input capacitance of just 1pF would limit the bandwidth to only 1.77MHz ; **b) compensated:** the capacitive divider takes over at high frequencies but the input capacitance of the following stage of 1pF would spoil the division by 1%; **c) adjustable:** in practice, the capacitive divider is trimmed for a perfect step response.

In general, at DC the signal source is loaded by the total attenuation resistance R_a ; for an attenuation factor A , the values of resistor R_1 and R_2 must satisfy the following equations:

$$R_1 + R_2 = R_a \quad (5.2.1)$$

The current through the resistive path is:

$$i_i = \frac{v_i}{R_1 + R_2} = \frac{v_o}{R_2} \quad (5.2.2)$$

so, from the last two expressions, the attenuation is:

$$\frac{v_o}{v_i} = \frac{1}{A} = \frac{R_2}{R_1 + R_2} \quad (5.2.3)$$

and the required resistance relation is:

$$R_1 = (A - 1)R_2 \quad (5.2.4)$$

Thus, for an $R_a = 1 \text{ M}\Omega$ and $A = 10$:

$$R_1 = 900 \text{ k}\Omega \quad \text{and} \quad R_2 = 100 \text{ k}\Omega \quad (5.2.5)$$

The high frequency compensation consists of a capacitive divider having the same attenuation factor as the high impedance resistive divider in parallel with it, as in [Fig. 5.2.2b](#). In order to achieve a precise attenuation, resistors with 0.1% tolerance are used, giving a maximum error of 0.2%. However, capacitors with a comparably tight tolerance are not readily available, and, even if they were, the layout strays would dominate. So in practice the capacitive divider is made adjustable, [Fig. 5.2.2c](#). Trimming of C_1 should be avoided, in order to reduce the circuit size (and thus stray inductances and capacitances); a much better choice is to trim only some 20–30% of C_2 . Care should be taken to connect the variable plate to the ground, otherwise the metal tip of the adjusting screwdriver would modify the capacitance by contact alone.

For a well trimmed attenuator, the capacitive reactance ratio at high frequencies must match the resistor ratio at DC and low frequencies:

$$\frac{R_1}{R_2} = \frac{X_{C1}}{X_{C2}} = \frac{\frac{1}{j\omega C_1}}{\frac{1}{j\omega C_2}} = \frac{C_2}{C_1} \quad (5.2.6)$$

which also implies that the two RC constants must be equal:

$$R_1 C_1 = R_2 C_2 = \tau_a \quad (5.2.7)$$

The input impedance now becomes:

$$\begin{aligned} Z_a = Z_1 + Z_2 &= \frac{1}{\frac{1}{R_1} + j\omega C_1} + \frac{1}{\frac{1}{R_2} + j\omega C_2} \\ &= \frac{R_1}{1 + j\omega C_1 R_1} + \frac{R_2}{1 + j\omega C_2 R_2} = (R_1 + R_2) \frac{1}{1 + j\omega \tau_a} \end{aligned} \quad (5.2.8)$$

In the latter expression we have taken into account [Eq. 5.2.7](#). This is the same as if we would have a single parallel $R_a C_a$ network:

$$Z_a = R_a \frac{1}{1 + j\omega C_a R_a} \quad (5.2.9)$$

where:

$$R_a = R_1 + R_2 \quad \text{and} \quad C_a = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2}} \quad (5.2.10)$$

By substituting $C_2 = (A - 1)C_1$ the input capacitance C_a relates to C_1 as:

$$C_a = \frac{1}{\frac{1}{C_1} + \frac{1}{(A - 1)C_1}} = C_1 \frac{A - 1}{A} \quad (5.2.11)$$

The transfer function can then be calculated from the attenuation:

$$F(j\omega) = \frac{1}{A} = \frac{v_{out}}{v_{in}} = \frac{Z_2}{Z_1 + Z_2} = \frac{1}{1 + \frac{R_1}{R_2} \cdot \frac{1 + j\omega C_2 R_2}{1 + j\omega C_1 R_1}} \quad (5.2.12)$$

Obviously, the frequency dependence will vanish if the condition of [Eq. 5.2.7](#) is met. However, the transfer function will be independent of frequency only if the signal's source impedance is zero (we are going to see the effects of the signal source impedance a little later).

The transfer function of an unadjusted attenuator ($R_1 C_1 \neq R_2 C_2$) has a simple pole and a simple zero, as can be deduced from [Eq. 5.2.12](#). If we rewrite the impedances as:

$$Z_1 = \frac{1}{\frac{1}{R_1} + sC_1} = R_1 \frac{\frac{1}{R_1 C_1}}{s - \left(-\frac{1}{R_1 C_1}\right)} = R_1 \frac{-s_1}{s - s_1} \quad (5.2.13)$$

and

$$Z_2 = \frac{1}{\frac{1}{R_2} + sC_2} = R_2 \frac{\frac{1}{R_2 C_2}}{s - \left(-\frac{1}{R_2 C_2}\right)} = R_2 \frac{-s_2}{s - s_2} \quad (5.2.14)$$

where s_1 and s_2 represent the poles in each impedance arm, explicitly:

$$s_1 = -\frac{1}{R_1 C_1} \quad \text{and} \quad s_2 = -\frac{1}{R_2 C_2} \quad (5.2.15)$$

The transfer function is then:

$$\frac{v_{out}}{v_{in}} = \frac{Z_2}{Z_1 + Z_2} = \frac{R_2 \frac{-s_2}{s - s_2}}{R_1 \frac{-s_1}{s - s_1} + R_2 \frac{-s_2}{s - s_2}} \quad (5.2.16)$$

By solving the double divisions, the transfer function can be rewritten as:

$$\frac{v_{out}}{v_{in}} = \frac{-s_2 R_2 (s - s_1)}{-s_1 R_1 (s - s_2) - s_2 R_2 (s - s_1)} \quad (5.2.17)$$

We can replace the products $s_i R_i$ by $1/C_i$:

$$\begin{aligned} \frac{v_{out}}{v_{in}} &= \frac{\frac{1}{C_2}(s - s_1)}{\frac{1}{C_1}(s - s_2) + \frac{1}{C_2}(s - s_1)} = \frac{C_1(s - s_1)}{C_2(s - s_2) + C_1(s - s_1)} = \\ &= \frac{C_1(s - s_1)}{s(C_1 + C_2) - s_2 C_2 - s_1 C_1} = \frac{C_1}{C_1 + C_2} \cdot \frac{(s - s_1)}{s - \frac{s_2 C_2 + s_1 C_1}{C_1 + C_2}} \end{aligned} \quad (5.2.18)$$

This can be simplified by defining a few useful substitutions: the capacitive divider attenuation:

$$A_C = \frac{C_1}{C_1 + C_2} \quad (5.2.19)$$

the system zero:

$$s_z = s_1 = -\frac{1}{R_1 C_1} \quad (5.2.20)$$

and the system pole:

$$s_p = \frac{s_2 C_2 + s_1 C_1}{C_1 + C_2} \quad (5.2.21)$$

Further, the system pole can be rewritten as:

$$\begin{aligned} s_p &= \frac{s_2 C_2 + s_1 C_1}{C_1 + C_2} = \frac{-C_2 \frac{1}{R_2 C_2} - C_1 \frac{1}{R_1 C_1}}{C_1 + C_2} = -\frac{\frac{1}{R_2} + \frac{1}{R_1}}{C_1 + C_2} \\ &= -\frac{R_1 + R_2}{R_1 R_2} \cdot \frac{1}{C_1 + C_2} \end{aligned} \quad (5.2.22)$$

From the system pole we note that the system time constant is equal to the parallel connection of all four components.

We will also define the resistance attenuation as:

$$A_R = \frac{R_2}{R_1 + R_2} \quad (5.2.23)$$

and we then rewrite the system pole as:

$$s_p = -\frac{R_1 + R_2}{R_2} \cdot \frac{1}{R_1(C_1 + C_2)} = -\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)} \quad (5.2.24)$$

With all these substitutions the complex frequency response is:

$$\boxed{F(s) = \frac{v_{out}}{v_{in}} = A_C \frac{s - s_z}{s - s_p}} \quad (5.2.25)$$

Again, it is obvious that the frequency dependence vanishes if $s_p = s_z$.

From $F(s)$ we derive the magnitude:

$$M(\omega) = |F(j\omega)| = \sqrt{F(j\omega) \cdot F(-j\omega)} = A_C \sqrt{\frac{j\omega - \sigma_z}{j\omega - \sigma_p} \cdot \frac{-j\omega - \sigma_z}{-j\omega - \sigma_p}} \quad (5.2.26)$$

which results in:

$$M(\omega) = A_C \sqrt{\frac{\omega^2 + \sigma_z^2}{\omega^2 + \sigma_p^2}} \quad (5.2.27)$$

The phase angle is the arctangent of the imaginary to real component ratio of the frequency response $F(j\omega)$:

$$\varphi(\omega) = \arctan \frac{\Im\{F(j\omega)\}}{\Re\{F(j\omega)\}} = \arctan \frac{\Im\left\{\frac{j\omega - \sigma_z}{j\omega - \sigma_p}\right\}}{\Re\left\{\frac{j\omega - \sigma_z}{j\omega - \sigma_p}\right\}} \quad (5.2.28)$$

First we must rationalize $F(j\omega)$ by multiplying both the numerator and the denominator by the complex conjugate of the denominator ($-j\omega - \sigma_p$):

$$\frac{j\omega - \sigma_z}{j\omega - \sigma_p} = \frac{(j\omega - \sigma_z)(-j\omega - \sigma_p)}{(j\omega - \sigma_p)(-j\omega - \sigma_p)} = \frac{\omega^2 + j\omega\sigma_z - j\omega\sigma_p + \sigma_z\sigma_p}{\omega^2 + \sigma_p^2}$$

and then we separate the real and imaginary part:

$$\frac{\omega^2 + j\omega\sigma_z - j\omega\sigma_p + \sigma_z\sigma_p}{\omega^2 + \sigma_p^2} = \frac{\omega^2 + \sigma_z\sigma_p}{\omega^2 + \sigma_p^2} + j \frac{\omega(\sigma_z - \sigma_p)}{\omega^2 + \sigma_p^2}$$

The phase angle is then:

$$\varphi(\omega) = \arctan \frac{\omega(\sigma_z - \sigma_p)}{\omega^2 + \sigma_z\sigma_p} = \arctan \frac{\frac{\sigma_z - \sigma_p}{\omega}}{1 + \frac{\sigma_z\sigma_p}{\omega^2}} \quad (5.2.29)$$

By using the identity:

$$\arctan x - \arctan y = \arctan \frac{x - y}{1 + xy}$$

we can write:

$$\varphi(\omega) = \arctan \frac{\sigma_z}{\omega} - \arctan \frac{\sigma_p}{\omega} \quad (5.2.30)$$

With $\sigma_p = \sigma_z$ the phase angle is zero for any ω .

The envelope delay is the frequency derivative of the phase:

$$\begin{aligned} \tau_d &= \frac{d\varphi}{d\omega} = \frac{d}{d\omega} \left(\arctan \frac{\sigma_z}{\omega} - \arctan \frac{\sigma_p}{\omega} \right) \\ &= \frac{1}{1 + \left(\frac{\sigma_z}{\omega}\right)^2} \left(-\frac{\sigma_z}{\omega^2} \right) - \frac{1}{1 + \left(\frac{\sigma_p}{\omega}\right)^2} \left(-\frac{\sigma_p}{\omega^2} \right) \end{aligned} \quad (5.2.31)$$

So the result is:

$$\tau_d = \frac{-\sigma_z}{\omega^2 + \sigma_z^2} - \frac{-\sigma_p}{\omega^2 + \sigma_p^2} \quad (5.2.32)$$

Again, note that for $\sigma_p = \sigma_z$ the envelope delay is zero.

We have plotted the magnitude, phase and envelope delay in [Fig. 5.2.3](#). The plots are made for the matched and two unmatched cases in order to show the influence of trimming the attenuator by C_2 (± 10 pF).

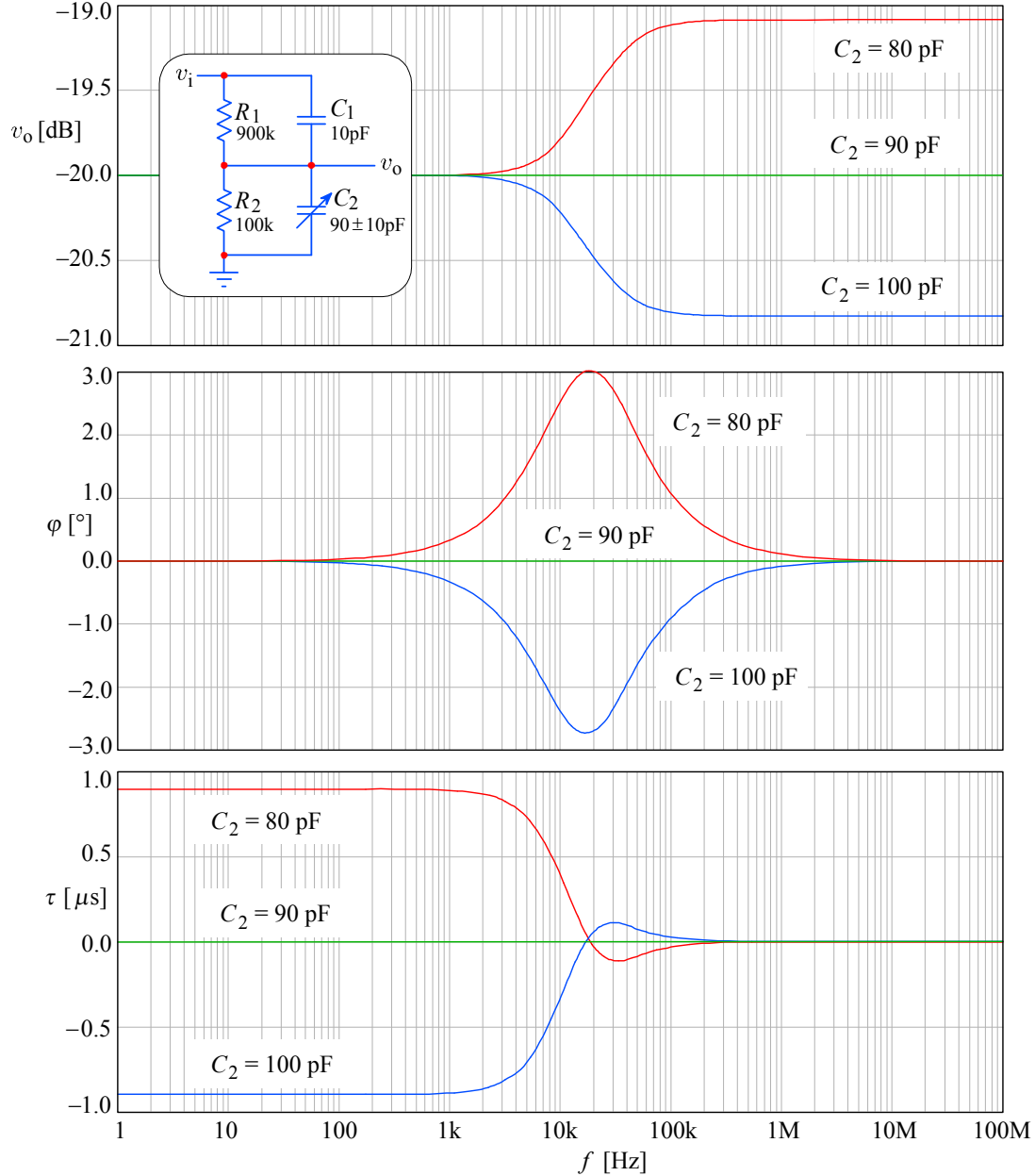


Fig. 5.2.3: The attenuator magnitude, phase, and envelope delay responses for the correctly compensated case (flat lines), along with the under- and over-compensated cases (C_2 is trimmed by ± 10 pF). Note that these same figures apply also to oscilloscope passive probe compensation, demonstrating the importance of correct compensation when making single channel pulse measurements and two channel differential measurements.

The step response is obtained from $F(s)$ by the inverse Laplace transform, using the theory of residues:

$$\mathcal{L}^{-1}\left\{\frac{1}{s} F(s)\right\} = \frac{A_C}{2\pi j} \oint_C \frac{s - s_z}{s(s - s_p)} e^{st} ds = A_C \sum_{s=0}^{s=s_p} \text{res}\left\{\frac{s - s_z}{s(s - s_p)} e^{st}\right\}$$

We have two residues. One is owed to the unit step operator, $1/s$:

$$\begin{aligned} \text{res}_1 &= A_C \lim_{s \rightarrow 0} s \left\{ \frac{s - s_z}{s(s - s_p)} e^{st} \right\} = A_C \lim_{s \rightarrow 0} \left\{ \frac{s - s_z}{s - s_p} e^{st} \right\} \\ &= A_C \frac{-s_z}{-s_p} = A_C \frac{-\frac{1}{R_1 C_1}}{-\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)}} \\ &= A_C A_R \frac{R_1(C_1 + C_2)}{R_1 C_1} = A_C A_R \frac{C_1 + C_2}{C_1} = A_C A_R \frac{1}{A_C} \\ &= A_R = \frac{R_2}{R_1 + R_2} \end{aligned} \quad (5.2.33)$$

As expected, the residue for zero frequency (DC) is set by the resistance ratio.

The other residue is due to the system pole, s_p :

$$\begin{aligned} \text{res}_2 &= A_C \lim_{s \rightarrow s_p} (s - s_p) \left\{ \frac{s - s_z}{s(s - s_p)} e^{st} \right\} = A_C \lim_{s \rightarrow s_p} \left\{ \frac{s - s_z}{s} e^{st} \right\} = \\ &= A_C \frac{s_p - s_z}{s_p} e^{s_p t} = A_C \frac{-\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)} - \left(-\frac{1}{R_1 C_1}\right)}{-\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)}} e^{s_p t} \\ &= A_C \frac{\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)} - \frac{1}{R_1 C_1}}{\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)}} e^{s_p t} \\ &= A_C \left(1 - A_R \frac{C_1 + C_2}{C_1}\right) e^{s_p t} = A_C \left(1 - \frac{A_R}{A_C}\right) e^{s_p t} \\ &= (A_C - A_R) e^{s_p t} \end{aligned} \quad (5.2.34)$$

The result is a time decaying exponential, with the time constant set by the system pole, s_p , and the amplitude set by the difference between the capacitive and resistive divider.

The step response is the sum of both residues:

$$\begin{aligned} f(t) &= \sum \text{res} = A_R + (A_C - A_R) e^{s_p t} \\ &= A_R + (A_C - A_R) e^{-\frac{1}{A_R} \cdot \frac{1}{R_1(C_1 + C_2)} t} \end{aligned} \quad (5.2.35)$$

So the explicit result is:

$$f(t) = \frac{R_2}{R_1 + R_2} + \left(\frac{C_1}{C_1 + C_2} - \frac{R_2}{R_1 + R_2} \right) e^{-\frac{R_1 + R_2}{R_1 R_2} \cdot \frac{1}{C_1 + C_2} t} \quad (5.2.36)$$

When $A_C = A_R$, the exponential function coefficient is zero, thus:

$$f(t) = \frac{R_2}{R_1 + R_2} \quad (5.2.37)$$

The system's time constant, as we have already seen in [Eq. 5.2.24](#), is:

$$\tau_a = -\frac{1}{s_p} = \frac{R_1 R_2}{R_1 + R_2} (C_1 + C_2) \quad (5.2.38)$$

For a well compensated attenuator, the following is true:

$$\tau_a = \frac{R_1 R_2}{R_1 + R_2} (C_1 + C_2) = R_1 C_1 = R_2 C_2 \quad (5.2.39)$$

We have plotted the step response in [Fig. 5.2.4](#). The plots are made for the matched and two unmatched cases in order to show the influence of trimming the attenuator by C_2 (± 10 pF), as in the frequency domain plots.

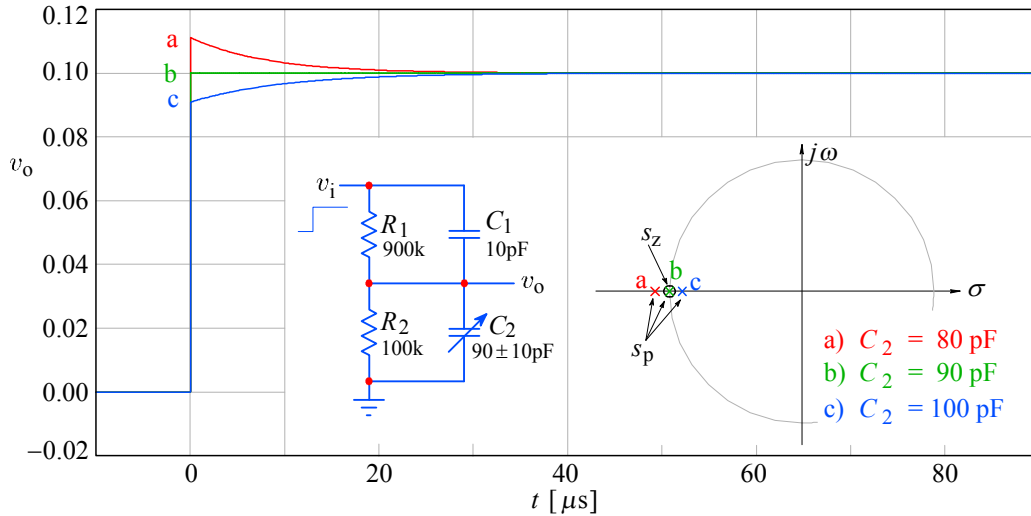


Fig. 5.2.4: The attenuator's step response for the correctly compensated case, along with the under- and over-compensated cases (C_2 is trimmed by ± 10 pF). Note that by changing C_2 the system pole also changes but the system zero remains the same.

Now we are going to analyze the influence of the non-zero source impedance on the transfer function. Since we can reuse some of the results we shall not need to recalculate everything.

The capacitive divider presents a relatively high output capacitance to the following amplifier, and the amplifier input capacitance appears in parallel with C_2 , changing the division slightly, but that is compensated by trimming.

However, the attenuator input capacitance C_a ([Eq. 5.2.10](#)) is smaller than C_1 (for an attenuation of $A = 10$, the input capacitance is $C_a = \frac{9}{10} C_1$). The actual values

of C_1 and C_2 are dictated mainly by the need to provide a standard value for various probes (compensated attenuators themselves, too). Historically, values between 10 and 20 pF have been used for C_a . Although small, this load is still significant if the signal source internal impedance is considered.

High frequency signal sources are designed to have a standardized impedance of $R_g = 50 \Omega$ (75 Ω for video systems). The cable connecting any two instruments must then have a characteristic impedance of $Z_0 = 50 \Omega$, and it must always be terminated at its end by an equal impedance in order to prevent signal reflections. As shown in Fig. 5.2.5a and 5.2.5b, the internal source resistance R_g and the termination resistance R_t form a $\div 2$ attenuator (neglecting the 1 M Ω of the 10:1 attenuator):

$$\frac{v_o}{v_g} = \frac{R_t}{R_g + R_t} = \frac{50 \Omega}{50 \Omega + 50 \Omega} = \frac{1}{2} \quad (5.2.40)$$

Therefore the effective signal source impedance seen by the attenuator is:

$$R_{ge} = \frac{R_g R_t}{R_g + R_t} = \frac{2500}{100} = 25 \Omega \quad (5.2.41)$$

With a 9 pF equivalent attenuator input capacitance ($C_a = 0.9 C_1$) a pole at $s_h = -1/R_{ge}C_a$ is formed, resulting in an $f_h = 1/(2\pi R_{ge}C_a) = 707$ MHz cut off.

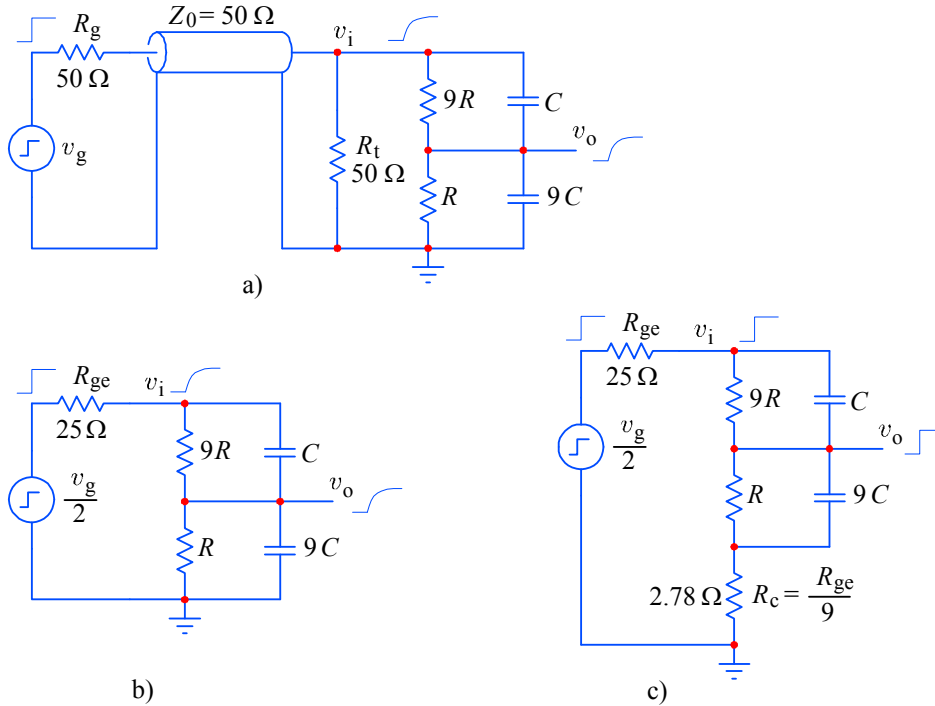


Fig. 5.2.5: **a)** When working with 50 Ω impedance the terminating resistance must match the generator internal resistance, forming a $\div 2$ attenuator with an effective output impedance of 25 Ω ; **b)** With a 9 pF attenuator input capacitance, a HF cut off at 707 MHz results; **c)** The cut off for the $\div 10$ attenuator can be compensated by a 25/9 Ω resistor between the lower end of the attenuator and the ground.

Owing to the high resistances involved, we can neglect the attenuator's resistive arms and consider only the equivalent signal source impedance and the capacitive divider (assuming the attenuator is correctly compensated). If $F_0(s)$ is the

attenuator transfer function for a zero signal source impedance (Eq. 5.2.25), the transfer function for the impedance $R_{ge} = 25 \Omega$ is:

$$F_1(s) = \frac{1}{2} R_{ge} \cdot \frac{-\frac{1}{R_{ge}C_a}}{s - \left(-\frac{1}{R_{ge}C_a}\right)} \cdot F_0(s) \quad (5.2.42)$$

The resulting pole s_h can be compensated by inserting a resistor R_c of $25/9 = 2.78 \Omega$ between the lower attenuator end and the ground, as in Fig. 5.2.5c; with the equivalent input resistance of $R_g = 25 \Omega$, the R_c provides the 10:1 division at highest frequencies, as required. The effective output resistance at these frequencies is only 2.5Ω , so the pole formed by it and the unity gain buffer total input capacitance of, say, 2 pF would be well beyond the bandwidth of interest (~ 32 GHz). However, the bandwidth of the passive part of the input circuitry is impaired by other effects, as we shall soon see.

A typical attenuator consists of three switch selectable sections, 1:1, 10:1, and 100:1, as shown in Fig. 5.2.6. This allows us to cover the required input amplitude range from millivolts to tens of volts (intermediate 2:1 and 5:1 attenuation settings are usually implemented after the input buffer stage with low value resistors). Because the attenuator's compensation is adjustable, the input capacitance changes, so it has to be 'standardized' by an additional compensating capacitor to make the input capacitance of all sections equal.

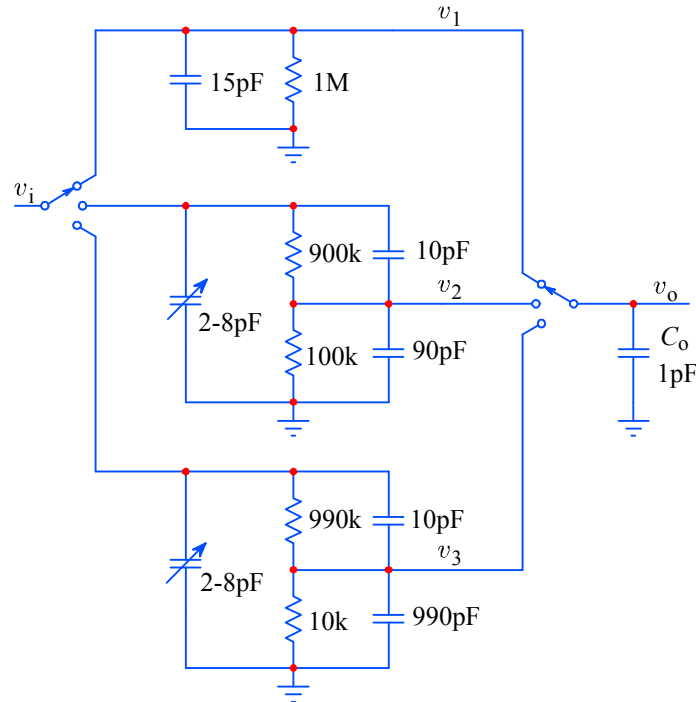


Fig. 5.2.6: The direct and the two attenuation paths are switched at both input and output, in order to reduce the input capacitance. For low cross-talk, the input and output of each unused section should be grounded (not shown here). The variable capacitors in parallel with the two attenuation sections are adjusted so that the input capacitance is equal for all settings. Of course, other values are possible, e.g., $\div 1$, $\div 20$, $\div 400$ (as in Tek 7A11), with the highest attenuation achieved by cascading two $\div 20$ sections. The advantage is that the parasitic serial inductance of the largest capacitance in the highest attenuation section is avoided; a disadvantage is that it is very difficult to trim correctly.

Unfortunately, for attenuator settings other than 10:1 the resistive compensation shown in Fig. 5.2.5 can be very difficult to achieve. For a 100:1 attenuation the resistance required between the attenuator and the ground would be only $0.2525\ \Omega$ and such a low value might be present in the circuit already (in the form of a ground return trace if it is not wide enough). Even if we are satisfied by a 1% tolerance we still need a well controlled design, taking care of every 10 m Ω .

The most critical is the direct (1:1) path. In order to present the same input load to an external 10:1 probe this path must have the same input resistance and capacitance as the higher attenuation paths. However, the direct path has no attenuation, so **the direct path can not be compensated** by a low value resistor.

It is for this reason that many designers avoid using the direct path altogether and opt for a 2:1 and 20:1 combination instead. Such a circuit, showing also the resistive compensation, is drawn in Fig. 5.2.7. As a bonus, the amplifier input current limiting and input overdrive protection circuitry is easier to realize (no 1:1 path), needing lower serial impedance and thus allowing higher bandwidths.

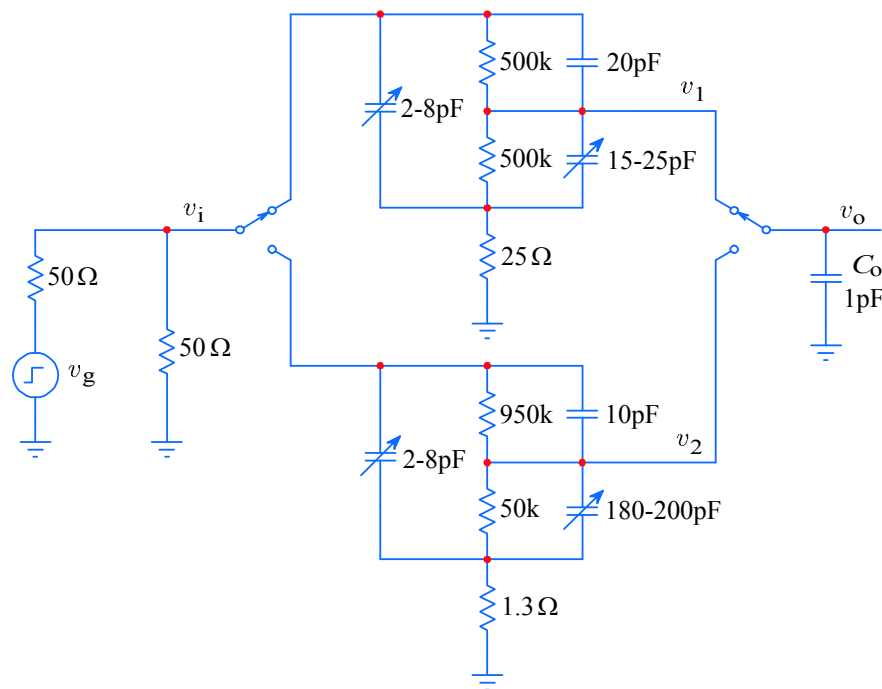


Fig. 5.2.7: The attenuator with no direct path, in which the $25\ \Omega$ effective source impedance compensation can be used for both settings. Low ground return path impedance is necessary.

On the negative side, by using the $\div 2$ and $\div 20$ attenuation, the amplifier must provide for another gain of two, making the system optimization more difficult. Fortunately, for modern amplifiers, driving an AD converter, the gain requirement is low, since the converter requires only a volt or two for a full range display; in contrast, a conventional 'scope CRT requires tens of volts on the vertical deflecting plates. Thus a factor of at least 10 in gain reduction (and a similar bandwidth increase!) is in favor to modern circuits.

Whilst the gain requirements are relaxed, modern sensitive circuits require a higher attenuation to cover the desired signal range. But obtaining a 200:1 attenuation can be difficult, because of capacitive feed through: even a 0.1 pF from the input to the

buffer output, together with a non-zero output impedance, can be enough to spoil the response. If we can tolerate a feed through error of one least significant bit of an 8 bit analog to digital converter, the 200:1 attenuator would need an effective isolation of $20 \log_{10}(200 \times 2^8) = 94 \text{ dB}$, which is sometimes hard to achieve even at audio frequencies, let alone GHz. A cascade of two sections could be the solution.

5.2.2 Attenuator Inductance Loops

Designer's life would be easy with only resistances and capacitances to deal with. But every circuit also has an inductance, whether we intentionally put it in or desperately try to avoid it. As we have learned in [Part 2](#), in wideband amplifiers, instead of trying to avoid the unavoidable, we rather try to put the inductance to use by means of fine tuning and adequate damping.

In [Fig. 5.2.8](#) we have indicated the two inductances associated with the attenuator circuit. Because of the high voltages involved, the attenuator circuit can not use arbitrarily small components, packed arbitrarily close together. As a consequence, the circuit will have loop dimensions which can not be neglected and, since the inductance value is proportional to the loop area, the inductance values can be relatively large (for wideband amplifiers).

As for stray capacitance, the value of stray inductance can not be readily predicted, at least not to the precision required. Each component in [Fig. 5.2.8](#) will have its own stray inductances, one associated with the internal component structure and the other associated with the component leads, the soldering pads, and PCB traces. These will be added to the loop inductance.

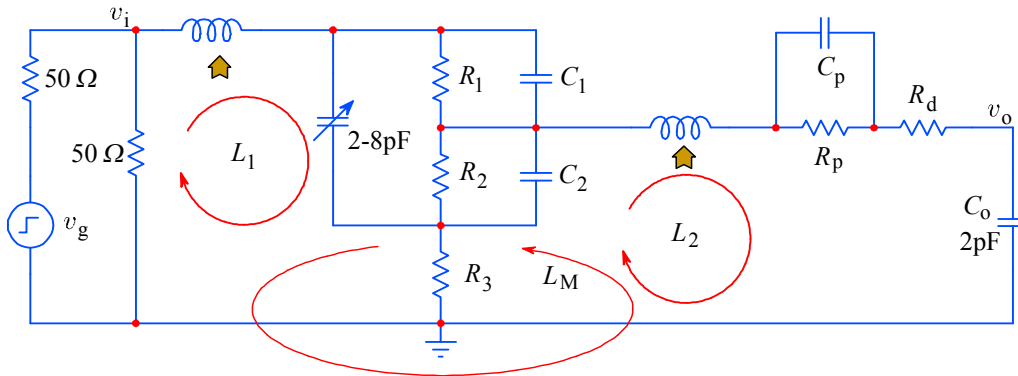


Fig. 5.2.8: Inductances owed to circuit loops can be modeled as inductors in series with the signal path. Note that in addition to the two self inductances there is also a mutual inductance between the two. The actual values depend on the loop's size, which in turn depends on the size of the components and the circuit's layout. Smaller loops have less inductance. Mutual inductance can be reduced by shielding, although this can increase the stray capacitances.

Nevertheless, it is relatively easy to estimate both loop inductances, at least to an order of magnitude. Basically, a single loop current I causes a magnetic flux Φ with a density B within the loop area S , so the self inductance is:

$$L = \frac{\Phi}{I} = \frac{BS}{I} = \frac{\mu HS}{I} = \frac{\mu_0 \mu_r HS}{I} \quad (5.2.43)$$

The current I and the magnetic field strength H are proportional: $H = I/2r$ for a single loop, where r is the loop radius. In a linear non-magnetic environment (with the relative permeability $\mu_r = 1$) I and B are also proportional because $B = \mu H$. Furthermore, μ_0 is the free space magnetic permeability, also known as the ‘induction constant’, the value of which has been set by the SI agreement about the Ampere: $\mu_0 = 4\pi \cdot 10^{-7} [\text{Vs A}^{-1} \text{m}^{-1}]$. This means that a current of 1 A encircling once a loop area of 1 m^2 causes a magnetic field strength of 1 Vs. Because for a circular loop $S = \pi r^2$, our loop inductance equation can be reduced to:

$$L = \frac{\mu_0 A}{2r} = \frac{\mu_0 \pi r^2}{2r} = \frac{\mu_0 \pi r}{2} = k r \quad (5.2.44)$$

where $k = 2\pi^2 \cdot 10^{-7} \text{ H/m}$. The inductance of a 1 m^2 loop ($r = 0.5642 \text{ m}$) is then $\approx 1.14 \cdot 10^{-6} \text{ H}$ (the unit of inductance is ‘henry’, after *Joseph Henry*, 1791–1878; $[\text{H}] = [\text{Vs A}^{-1}]$).

As a more practical figure, a loop of 10 cm^2 ($\approx 0.0178 \text{ m}$ radius circle) has an inductance of $\approx 35 \text{ nH}$. This does not look much, but remember that in our circuit the loop inductance L_1 is effectively in series with the signal source and is loaded by the attenuator’s input capacitance, forming a 2nd-order low pass filter with a cut off frequency $f_h = 1/(2\pi\sqrt{L_1 C_a}) \approx 268 \text{ MHz}$, assuming $L_1 = 35 \text{ nH}$ and $C_a = 10 \text{ pF}$. With such values the step response rings long, since the equivalent signal source resistance ($R_{ge} = 25 \Omega$) is not high enough to damp the resonance (such damping would be adequate for $L_1 < 5 \text{ nH}$).

The above inductance estimation is based on a circular loop model, whilst our loops will usually be of a square form (thus increasing L), with additional stray inductances owing to the internal geometry of the components (capacitors) and their leads, or just the PCB traces if surface mounted components are used.

The loop inductances can, of course, be measured. If we replace C_1 , C_2 and R_3 (Fig. 5.2.8) by a wire of the same total length, the input resistance and L_1 form a high pass filter, whose cut off frequency can be measured. Next, by removing the wire and also R_3 and replacing R_p , R_d and C_o by another wire, we can measure $L_1 + L_2$. Obtaining L_2 is then a matter of simple subtraction. Finally, by applying a signal to the input, shorting R_3 , and measuring the signal induced in L_2 , we can calculate the mutual inductance L_M . Note that a thin wire will have a somewhat larger inductance than a wide PCB trace.

The best way to reduce the loop area (and consequently L) is to use a 3-layer PCB, and make the middle layer a ‘ground plane’. In addition, using surface mounted components reduces the circuit size and also allows us to place them on both sides of the board. However, this technique also increases the stray capacitances and can also cause reflections if the ‘microstrip’ trace impedances are not well matched to the circuit. Therefore, a careful PCB design is needed, with wider ground clearance around sensitive pads and using a material with low ϵ_r . The most sensitive node in this respect is the attenuator output.

Another way of reducing the effect of stray inductance is to employ the same technique as we did for the low value resistors. This means that the inductance in the ground path (the signal return path) should not be too small, as it would be in the ground plane case; rather, the return path inductance should be kept in the same ratio

to the L_1 as the attenuation ratio. Precision in this respect is difficult, but not impossible to achieve. Our inductance expression [Eq. 5.2.44](#) does not show it, but inductance is also inversely proportional to trace width. Powerful finite element numerical simulation routines will be required for the job.

However, the same trick can not be used for L_2 (no attenuation in this loop!). Fortunately, as will become clear from the analysis below, the input inductance L_1 is more critical than L_2 , since the latter is loaded by a much smaller capacitance (C_o) and can be suitably damped with a larger resistance (R_d , which is already in the circuit because it is required for the FET gate protection).

We shall analyze the attenuator loops by assuming perfectly matched time constants, $R_1C_1 = R_2C_2$, matched also to the other attenuator paths, so that the variable capacitor in parallel is not needed. Also, we shall replace the two $50\ \Omega$ resistors with a single $25\ \Omega$ one, representing the effective signal source resistance R_s in series with the input, with $v_i = v_g/2$. The loop inductances are represented by discrete components, L_1 and L_2 in the forward signal paths, as drawn in [Fig. 5.2.8](#).

In the second loop the first thing to note is that C_2 is many times larger than C_o ($10\text{--}500\times$, depending on the attenuation setting) and the same is true for C_p , which means that their reactance will be comparably low and can thus be neglected. Likewise, the resistances R_2 and R_p in parallel with these capacitances are large in comparison with their reactances. What remains is the loop inductance L_2 in series with $R_d + R_3$, driving the amplifier input capacitance C_o . If the attenuated input voltage is v_i/A , the output voltage will be:

$$v_o = \frac{v_i}{A} \cdot \frac{1}{sC_o} \cdot \frac{1}{sL_2 + R_d + R_3 + \frac{1}{sC_o}} \quad (5.2.45)$$

So we have a 2nd-order transfer function:

$$F_2(s) = \frac{v_o}{v_i} = \frac{1}{A} \cdot \frac{\frac{1}{L_2C_o}}{s^2 + s\frac{R_d + R_3}{L_2} + \frac{1}{L_2C_o}} \quad (5.2.46)$$

Since R_3 is fixed and of quite low value, R_d is used to provide the desired damping.

The input loop analysis is similar. Here we have the equivalent source resistance $R_s + R_3$ in series with L_1 , driving the equivalent input attenuator capacitance C_a ([Eq. 5.2.9](#); the attenuator resistance $R_1 + R_2$ can be neglected at high frequencies). At the top of the attenuator we have:

$$v_i = \frac{v_g}{2} \left(R_3 + \frac{1}{sC_a} \right) \frac{1}{sL_1 + R_s + R_3 + \frac{1}{sC_a}} \quad (5.2.47)$$

which results in the following second-order transfer function:

$$F_1(s) = \frac{2v_i}{v_g} = \frac{\frac{1}{L_1C_a} + s\frac{R_3}{L_1}}{s^2 + s\frac{R_s + R_3}{L_1} + \frac{1}{L_1C_a}} \quad (5.2.48)$$

The numerator can be written as:

$$\frac{1}{L_1 C_a} + s \frac{R_3}{L_1} = \frac{1}{L_1 C_a} (1 + s C_a R_3) \quad (5.2.49)$$

It is clear that the frequency of the zero, $1/C_a R_3$, is much higher than the frequency of the pole pair, $1/\sqrt{L_1 C_a}$. Also, if $L_1 \approx L_2$ and C_a is at least 5 to 10 times larger than C_o , then C_a will dominate the response. Fortunately, as discussed above, with a clever layout of components and a suitable ground plane, L_1 can be broken into L_{1a} and L_{1b} , so that L_{1b} is in the ground return path. If we can make $L_{1a} = 9L_{1b}$ we would achieve an effective inductance compensation in this loop.

We are thus left with the L_2 loop and its transfer function, [Eq. 5.2.49](#). However, this 2nd-order function will be transformed by the pole of the JFET source follower into a 3rd-order function, owing to its capacitive loading.

Although the inductance is always caused by a current loop, the inductance of a straight PCB trace can be estimated as some 7–10 nH/cm (length), depending on the trace width. In [\[Ref. 5.16\]](#) a good empirical approximation is offered:

$$L = 0.2 l \left[\ln \left(\frac{2l}{w+h} \right) + 0.2235 \left(\frac{w+h}{l} \right) - 0.5 \right] \quad (5.2.50)$$

where the trace length l , width w and thickness h are all in mm, resulting in the inductance in nH (no ground plane in this case!). With surface mounted components, by using capacitors with low serial inductance, and using miniature relay switches in the attenuator, the inductance L_2 can be reduced to less than 10 nH, making the pole (pair) at $1/\sqrt{L_2 C_o}$ high, compared to the source follower real pole (set by the damping resistance R_d and the source follower loading capacitance C_L (see the JFET source follower discussion in [Part 3, Sec. 3.9](#)). However, by making L_2 somewhat larger, say, 30–50 nH, we can achieve a 3rd-order Bessel pole pattern, improving the bandwidth and reducing the rise time. In [Fig. 5.2.9](#) we see the attenuator circuit of the $A = 10$ section, followed by a JFET source follower.

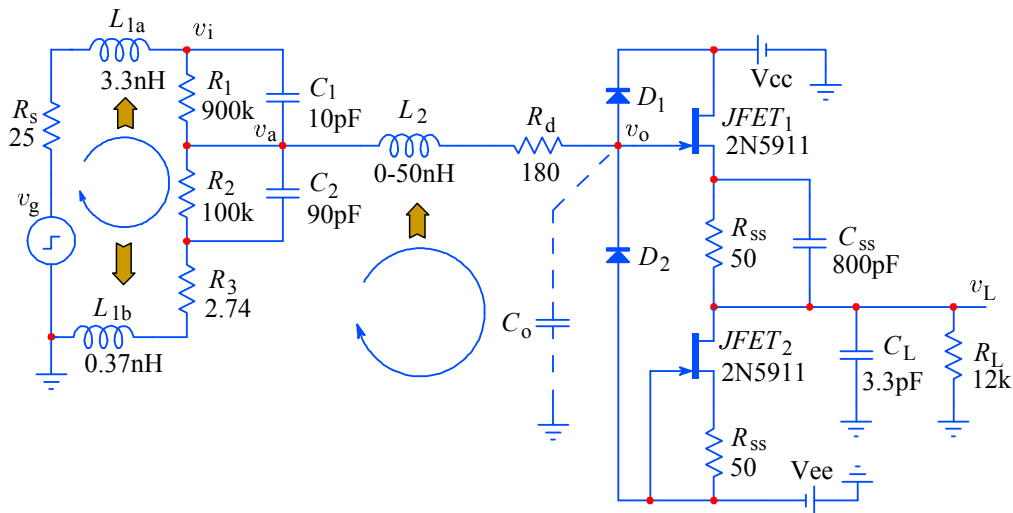


Fig. 5.2.9: The attenuator and the source follower JFET₁ (JFET₂ acts as a constant current source bias for JFET₁). The input loop inductance L_1 should be low, but the attenuation can be compensated by L_{1b} . The inductance L_2 of the second loop can be ‘tuned’ and damped by an appropriate value of R_d to provide a Bessel step response, as seen in [Fig. 5.2.10](#).

Note that here we have not drawn the protecting components C_p and R_p , but since a 325 V (peak value of the 230 V AC-mains) at the input results in a 32.5 V at the attenuator output, these components are absolutely necessary. Also, C_p should be a high voltage type (500 V), in order to survive the 325 V in the direct path (and still 163 V for a $\div 2$ attenuator); therefore, it will be of larger dimensions, so its internal serial inductance will have to be taken into account.

Note also that for high bandwidth a low value of C_o must be ensured. Since the negative input impedance compensation network (as in [Part 3, Sec. 3.9](#)), as well as R_d , D_1 , D_2 , C_{GD} , and C_L are present at the v_o node, C_o will tend to be high.

We have analyzed the step response in [Fig. 5.2.10](#) for two values of L_2 (10 and 50 nH; R_d has been chosen for a correct Bessel damping).

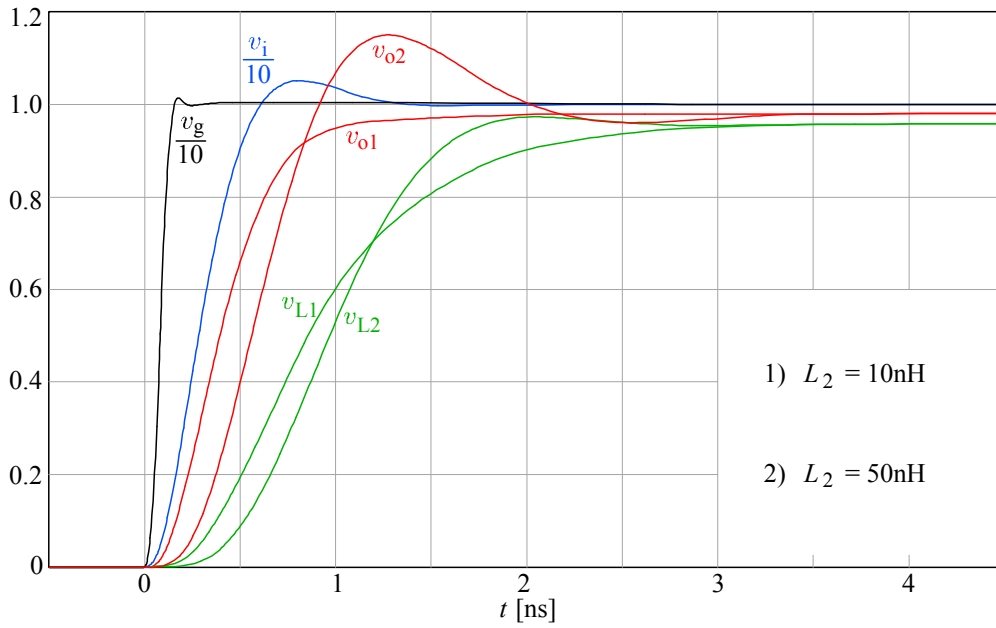


Fig. 5.2.10: Step response of the circuit in [Fig. 5.2.9](#). With a low L_1 , a correctly damped L_2 , and a good JFET, a 350 MHz bandwidth (v_{L2} rise time ≈ 1 ns), can be easily achieved. The source follower gain is a little less than one. v_o and v_L are drawn for the two L_2 cases.

5.2.3 The ‘Hook-Effect’

The discussion about high impedance attenuators would not be complete without mentioning the so called ‘hook-effect’. The name springs from the shape of the step response signal, which, owing to a sag in the 10–300 kHz region, resembles a hook at slower time base values ([Fig. 5.2.11](#)). The effect is caused by the frequency dependent relative permittivity, ϵ_r , of the PCB material (the standard glass epoxy FR4, FR stands for ‘flame resistant’, has an average $\epsilon_r = 4.5$, but it changes with frequency and temperature considerably).

The capacitance in farads of a parallel plates capacitor is expressed as:

$$C = \epsilon_0 \epsilon_r \frac{S}{d} \quad [F] = [As/V] \quad (5.2.51)$$

where S is the plate area [m²], d is their distance [m], $\epsilon_0 = 8.85 \times 10^{-12}$ [As/Vm] is the permittivity of the free space (vacuum) and ϵ_r is the relative permittivity of the

dielectric between the plates. A pad on the PCB thus has some small stray capacitance towards the ground (large if a ground plane is used). This capacitance changes with frequency proportionally with ε_r . Also, the material is porous and the fibres are long, extending to the edge of the board, allowing moisture in (water $\varepsilon_r = 80$), which causes long term changes. The problem is not specific to this material only, it is encountered with all traditional PCB materials (as well as many other insulators).

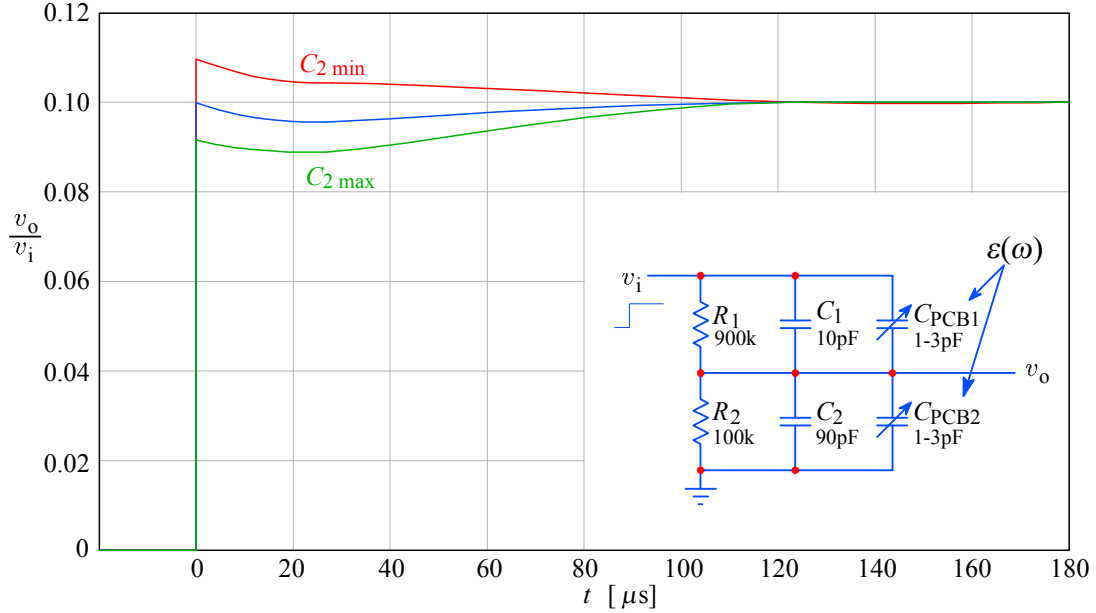


Fig. 5.2.11: The ‘hook-effect’ is most noticeable in the frequency range 10–300 kHz. Because the relative permittivity, ε_r , of a common PCB material is not exactly constant with frequency, the high impedance attenuator will exhibit a hook in its step response, which can not be trimmed out by the usual adjustment of C_2 . The C_{PCB} stray capacitance can vary by some 10–30 %, depending on the actual topology involved. Since C_1 is small, it is affected by a few percent. The lower attenuator leg is less affected, due to a larger value of C_2 .

To solve this problem, special Teflon[®] based material is used for instrument front end, but it is expensive and not readily available. If it can not be obtained, one possible solution could be to implement two large pads on a two sided PCB, in parallel to C_1 and C_2 , with their areas in the same ratio as the attenuation factor required [Ref. 5.68]. Then, the effect would be equally present in both legs, canceling out the hook, Fig. 5.2.12. Even some trimming can be done by drilling small holes in the larger pad pair (in contrast to cutting a pad corner, drilling removes the dielectric, thus lowering both A and ε).

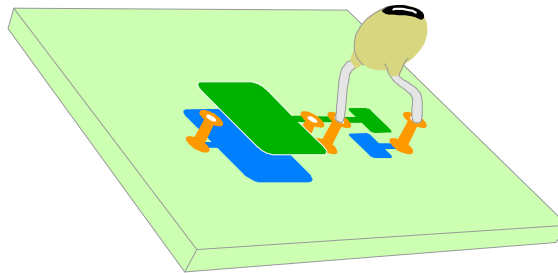


Fig. 5.2.12: Canceling the hook-effect in the common PCB material is achieved by intentionally adding two capacitances in form of large PCB pads, with areas in the same ratio as required by the attenuation (since the area is proportional to the square of the linear dimensions, for a 9:1 capacitance ratio, a 3:1 dimension ratio is needed). Trimming is possible by drilling small holes in the larger pad.

The main problem with this solution is that the use of external probes will expose the hook again, although to a lesser extent (owing to the large capacitance of the probe compensation).

5.2.4 Improving the JFET Source Follower DC Stability

The DC performance of a JFET source follower is far from perfect. Even if we use a dual JFET in the same case and on the same substrate, i.e., the Siliconix 2N5911 as in Fig. 5.2.9, their characteristics will not match perfectly. The 2N5911 data sheet state a V_{GS} mismatch of 10 mV maximum and a temperature drift of $20 \mu\text{V/K}$. The circuit in Fig. 5.2.13 offers moderate DC stability; the resistor R_T is trimmed for a zero v_L to v_{in} DC offset.

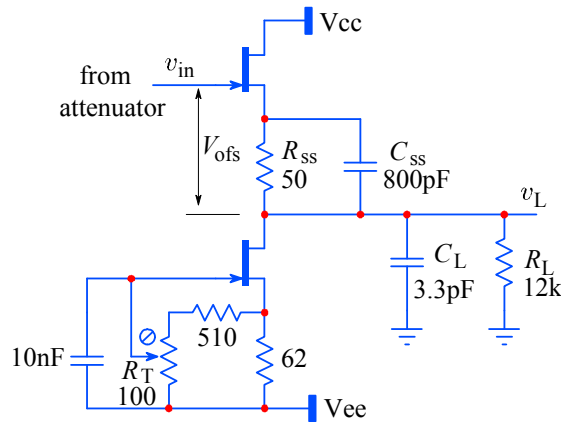


Fig. 5.2.13: Simple offset trimming of a JFET source follower.

Traditionally, oscilloscopes have a ‘vertical position’ control on the front panel (one for each channel), which is adjusted by the user in accordance with the particular measurement conditions, which differ from one situation to another, so the offset and drift (if not too high) are not of particular concern.

However, in modern instrumentation some additional features are becoming important, such as automated measurement, where we can not rely on the presence of a human operator to make adjustments every so often. Also it is not uncommon to find digital oscilloscopes with an 8 bit resolution (1:256) at high sampling rates, but capable of 12 bit (1:4096) or even 16 bit (1:65536) resolution at low speed, and in digital equipment it is expected that DC errors are of the order of ± 1 LSB.

By trimming the current source, we reduce the DC offset, but the temperature drift will remain. The gate current of a JFET, although normally in the $< 100 \text{ pA}$ range, is also temperature dependent and approximately doubles with every 10 K. The source follower input sees an attenuation dependent source resistance (from $1 \text{ M}\Omega$ to $10 \text{ k}\Omega$), so an additional offset component will be present owing to the gate current and the attenuator output impedance. A typical oscilloscope input has a maximum sensitivity of 5 mV/div. , or a 40 mV full screen range; the 1 LSB resolution for an 8 bit sampling is $40/256 = 0.15 \text{ mV}$, therefore the simple trimming circuit is inadequate for digital equipment, and an active offset compensation technique is required to keep the DC error below some $200 \mu\text{V}$.

Basically, there are three ways of achieving a low DC error, each having its own advantages and drawbacks. While DC performance is not of primary interest in this book, it should be implemented so that high frequency performance is preserved.

The first technique is suitable for microprocessor controlled equipment, where the input can be temporarily switched to ground, the offset measured, and the error either adjusted by a digital to analog converter or subtracted from the sampled signal in memory. But this operation should not be repeated too often or take a considerable amount of time, otherwise the equipment would be missing valid trigger events or, worse still, introduce errors by loading and unloading the signal source with the instrument's input impedance. This is a rather inelegant solution and it should be taken as the last resort only.

A better way, shown in [Fig. 5.2.14](#), is to use a good differential amplifier to monitor the difference between the Q_1 gate and the output, integrate it and modify the Q_2 current to minimize the offset. Note that this technique works well only while the input is within the linear range of the JFET; when in the non-linear range or when overdriven, the integrator will develop a high error voltage, which will be seen as a long 'tail' after the signal returns within the linear range. Also, owing to the presence of R_1 and R_2 , the attenuator lower arm resistors will need to be readjusted.

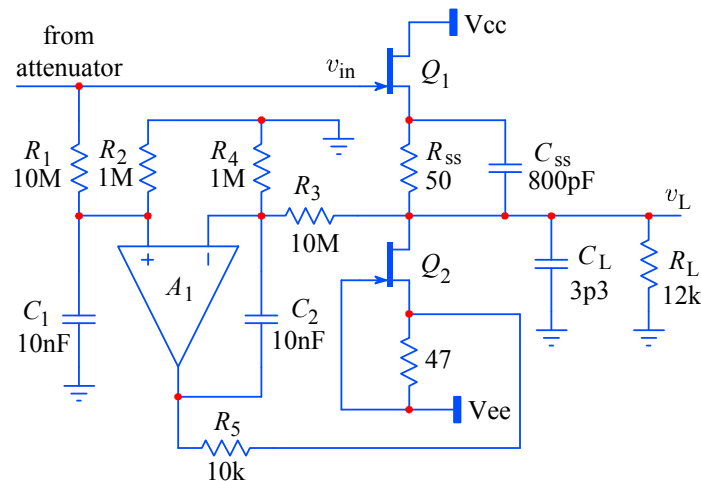


Fig. 5.2.14: Active DC correction loop. The amplifier A_1 amplifies and integrates the difference between the Q_1 gate and the output, driving through R_5 the source of Q_2 and modifying its current to minimize the offset. The resulting offset is equal to the offset of A_1 , multiplied by the loop gain $(1 + R_3/R_4)$. The differential amplifier with a very low offset will usually have its input bias current much larger than the JFET input current, therefore resistors R_2 and R_4 provide a lower impedance to ground. C_2 is the integration capacitor, whilst C_1 provides an equal time constant to the non-inverting input. The feedback divider, R_3 and R_4 should be altered to compensate for the system gain slightly lower than one (this is achieved by adding a suitably low value resistor in series with R_4). For a low error the amplifier A_1 must have a high common mode rejection up to the frequency set by C_2 and $R_3||R_4$.

But the most serious problem is owed to the amplifier A_1 : in order to minimize the system offset it should have both low offset and low input bias current itself. Although A_1 can be a low bandwidth device, the low input error requirements can easily put us back to where we started from.

The example in [Fig. 5.2.14](#) is relatively simple to implement. However, for a low error we must keep an eye on several key parameters. Ideally we would like to get

rid of the resistor R_2 (and R_4) to avoid the DC path to ground, because it alters the attenuator balance.

Unfortunately, the input common mode range of the error amplifier is limited and, more importantly, amplifiers with a low DC offset are usually made with bipolar transistors at the input, so their input bias current can be in the nA range, much higher than the JFET gate's leakage (< 20 pA). The bias current would then introduce a high DC offset over R_1 (and R_3). Here R_2 and R_4 come to the rescue, by conducting the large part of the bias current to ground over their lower resistance. On the other hand, the amplifier input offset voltage is then effectively amplified by the DC loop gain, $1 + R_3/R_4$. The amplifier is selected so that the total offset error is minimized:

$$V_{\text{ofs}} = \left(1 + \frac{4 \Delta R}{R}\right) \left(V_{A1\text{ofs}} \frac{R_3 + R_4}{R_4} + I_{A1\text{ofs}} \frac{R_1 R_2}{R_1 + R_2}\right) \quad (5.2.52)$$

where $\Delta R/R$ is the nominal resistor tolerance and $V_{A1\text{ofs}}$ and $I_{A1\text{ofs}}$ are the amplifier's voltage and current input offset, respectively.

An industry standard amplifier, the OP-07, has $V_{\text{ofs}} = 30 \mu\text{V}$ and $I_{\text{ofs}} = 0.4 \text{ nA}$ typical, so by taking the resistor values as in [Fig. 5.2.14](#) (with a 1 % tolerance), we can estimate the typical total system offset to be within $\pm 728 \mu\text{V}$, which is slightly larger than we would like. The offset can be reduced using a chopper stabilized amplifier, such as the Intersil's ICL-7650 or the LTC-1052 from Linear Technology, which have a very low voltage offset ($< 5 \mu\text{V}$) and low current offset (< 50 pA), but their switching noise must be filtered at the output; also their input switches are very delicate and must be well protected from over-voltage. Therefore, we can not do without R_2 and R_4 and consequently the attenuator must be corrected by increasing the lower resistance appropriately. See [\[Ref. 5.2\]](#) for more examples of such solutions.

The third technique involves separate low pass and high pass amplifier paths and summing their outputs.

The example in [Fig. 5.2.15](#) is made on the assumption that the sum of the two outputs restores the original signal in both phase and amplitude. As the readers who have tried to build loudspeaker crossover networks will know from experience, this can be done correctly only for simple, first-order RC filters (with just two paths; for higher order filters a third, band pass path is necessary).

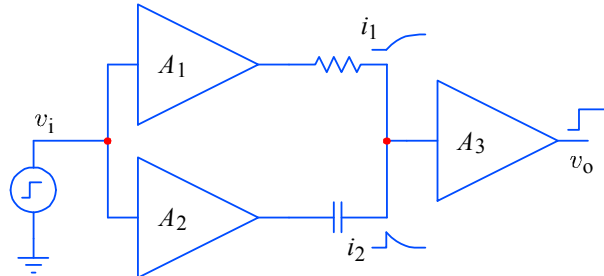


Fig. 5.2.15: The principle of separate low pass and high pass amplifiers.

Here the main problem is with the input of the low pass amplifier A_1 , which must have an equally low input bias current as the high pass A_2 , but should also have a very low voltage offset. Although in A_1 we don't need to worry about the high

frequency response, we are essentially again at the start, since JFETs and MOSFETs, which have low input current, have a high offset voltage and vice versa for the BJTs.

But we can combine [Fig. 5.2.14](#) and [5.2.15](#), and, by putting the RC network in front of the source follower, we can eliminate the amplifier A_2 . [Fig. 5.2.16](#) shows a possible implementation.

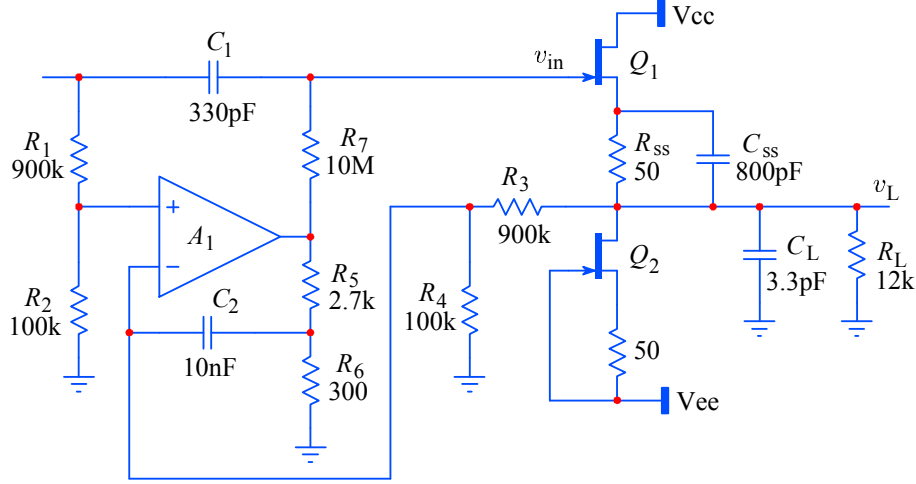


Fig. 5.2.16: With this configuration we can eliminate the need for a separate high pass amplifier. The DC correction is now applied to the Q_1 gate through R_7 . The error integrating amplifier A_1 must have a gain of 10 in order to compensate for the $1 + R_1/R_2$ and $1 + R_3/R_4$ attenuation. Resistors R_1 and R_2 now provide the $1\text{ M}\Omega$ input impedance for all attenuation settings, and this requires the compensated attenuators in front to be corrected accordingly.

Furthermore, instead of using a single differential amplifier we can invert the output by another low offset amplifier and rearrange the error amplifier into an inverting integrator, as in [Fig. 5.2.17](#). We can also self bias Q_1 by bootstrapping the resistor R_7 . This increases the input impedance by a very large factor, allowing us to reduce C_1 and thus further limit the current under overload conditions. However, be aware of the possibility of leakage currents from the protection diodes and the JFET gate itself, now that its DC input impedance has been increased.

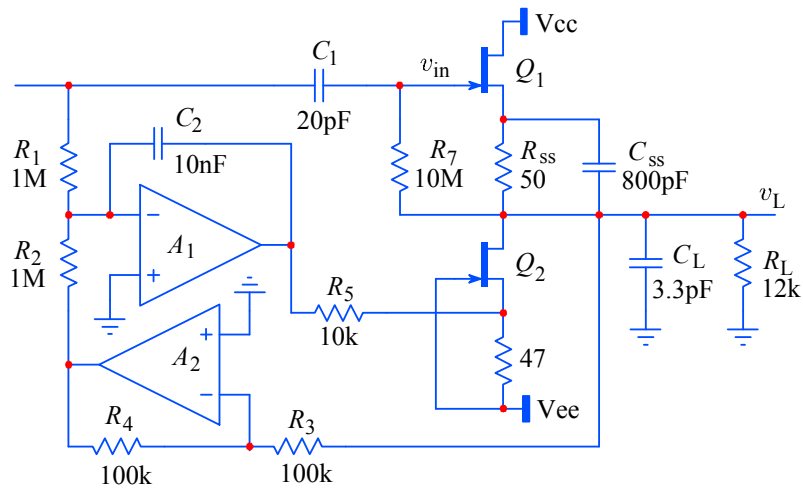


Fig. 5.2.17: By inverting the output the error amplifier becomes an inverting integrator and the offset correction is independent from the attenuator settings. The bootstrapping of R_7 produces an effective input resistance of about $2.4\text{ G}\Omega$.

Of course, now the DC error correction path must be returned to the current source Q_2 . The input resistor R_1 must be increased to $1\text{ M}\Omega$, since now the input of A_1 is at the virtual ground; likewise R_2 must be equal to R_1 . Note that both A_1 and A_2 offsets add to the final DC error.

Further evolution of this circuit is possible by combining a DC gain switching (R_2 or R_4 adjusting) with input attenuation. A very interesting result has been described in [Ref. 5.2], where also all input relays have been eliminated (using 3 source followers with the switching at their supply voltages by PIN diodes).

5.2.5 Overdrive Recovery

The integration loop will reduce the DC error only if the output follows the input. However, under a hard overdrive the JFET will saturate and the integrator will build up a charge proportional to the input overdrive amplitude and duration. When the overdrive is removed, the loop will reestablish the original DC conditions, but with the integration time constant, so the follower will exhibit a very long ‘tail’.

This is one of the most annoying properties of modern instrumentation, because we often want to measure the settling time of an amplifier and a convenient specification is the time from start of the transient to within 0.1 % of the final value. With a good old analog ‘scope we would simply increase the vertical sensitivity and adjust the vertical position so that the final signal level is within the screen range. But with modern DC compensated circuits this is not possible, and in order to avoid the post-overdrive tail we must use a specially built external limiter, [Ref. 5.6], to keep the input signal within the linear range of the ‘scope. The quality and speed of this limiter will also influence the measurement.

Note that simple follower circuits, like the one in Fig. 5.2.13, would also exhibit a small but noticeable post-overdrive tail, mainly owed to thermal effects. Also, high amplitude step response can be nonlinear, as shown in Fig. 5.2.18, owing to the variation of the JFET gate to channel capacitance with voltage (Eq. 5.2.18–19), but the time constant involved here is relatively small.

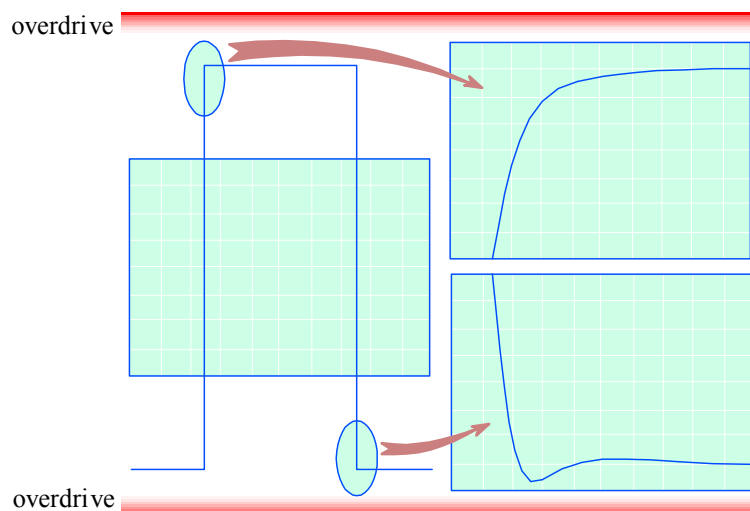


Fig. 5.2.18: Step response for large signals (but still below overdrive) is nevertheless nonlinear, caused by the variation of the JFET gate–drain capacitance with voltage.

5.2.6 Source Follower with MOSFETs

For a very long time, ever since semiconductors replaced electronic tubes in instrumentation, JFETs were the only components used for the source follower input section. Even today, JFETs outshine all other components in all performance aspects but one—shear speed. Unfortunately, BJT input impedance is much too low for the $1\text{ M}\Omega$ required. And MOSFETs, although having higher DC input resistance than JFETs, can have (depending on their internal geometry) higher input leakage current, are notoriously noisy, and their gate is easily damaged by overdrive.

If, however, we are ready to accept the design challenge to help the MOSFET by external circuitry, we might be rewarded with a faster follower. Also MOSFETs lend themselves nicely to integration, and this is where the experience gained from the design of high speed digital circuits can help. Circuit area reduction minimizes the stray capacitance and inductance, and new IC processing and semiconductor materials (e.g., GaAs, SiGe) increase charge mobility.

Note that for source follower applications a depletion type MOSFET is needed in order to achieve the required drain–source conductance with zero gate–source voltage. With appropriate doping, the supply voltage can be reduced to only 2 or 3 V (in contrast to several tens of volts required by conventional circuits), whilst retaining good high frequency operation. This also reduces the power dissipation and, more importantly, with low system supply voltage, the need for voltage gain is lower.

As with BJTs and JFETs, the parasitic capacitances of MOSFETs are also voltage dependent, but only partially, as will become evident from the following comparison with JFETs.

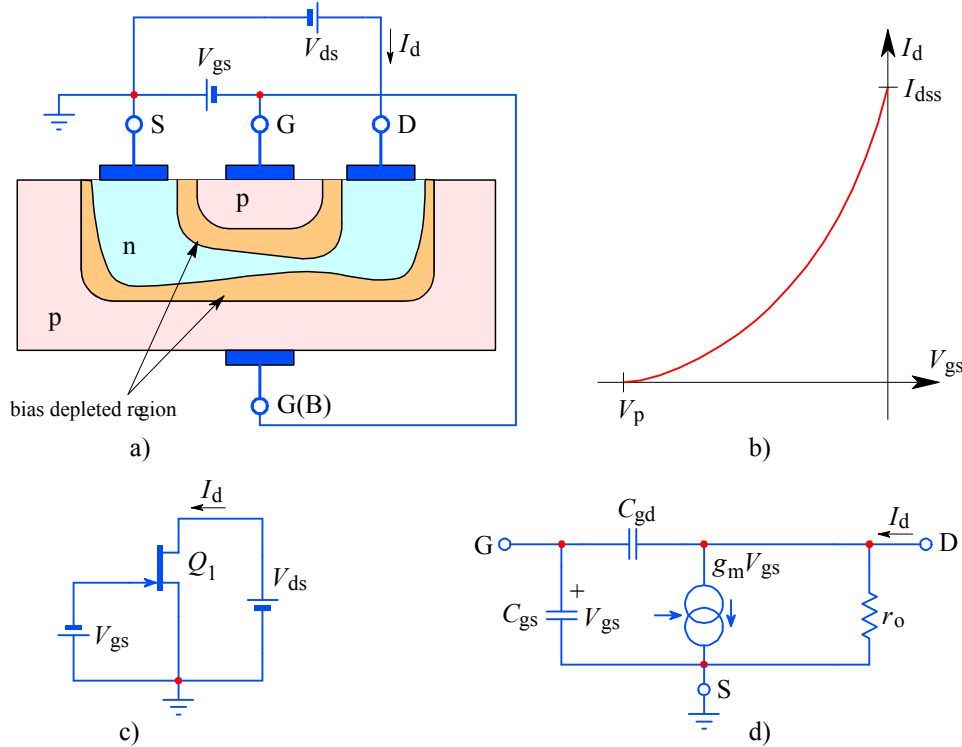


Fig. 5.2.19: a) A typical n-channel JFET structure cross-section under the bias condition. The p-type substrate is in contact with the p-type gate. The n-type channel is formed between the source and the drain. The bias voltage depletes the channel. b) The V_{gs} – I_d characteristic. c) The symbolic circuit. d) The equivalent circuit model.

The JFET capacitances C_{gd} and C_{gs} are voltage dependent:

$$C_{gd} = C_{gd0} \left(1 + \frac{V_{gd}}{V_{bi}} \right)^{-n_J} \quad (5.2.53)$$

$$C_{gs} = C_{gs0} \left(1 + \frac{V_{gs}}{V_{bi}} \right)^{-n_J} \quad (5.2.54)$$

where:

n_J is the junction grading coefficient (1/2 for abrupt and 1/3 for graded junctions; most JFETs are built with a graded junction);

$V_{bi} = \frac{k_B T}{q_e} \ln \frac{N_A N_D}{n_i^2}$ is the intrinsic zero bias built in potential;

N_A is the acceptor doping density in p -type material ($\approx 10^{21}$ atoms/m³);

N_D is the donor doping density in n -type material ($\approx 10^{22}$ atoms/m³);

n_i is the intrinsic Si charge density (1.5×10^{16} electrons/m³ at 300 K).

The built in potential V_{bi} relates to the JFET pinch off voltage parameter V_P as:

$$V_P = a^2 \frac{q_e N_A}{2 \varepsilon_{Si}} \left(1 + \frac{N_A}{N_D} \right) - V_{bi} \quad (5.2.55)$$

where:

a is the channel thickness ($\approx 2 \times 10^{-6}$ m);

ε_{Si} is the silicon dielectric permeability ($= 1.04 \times 10^{-10}$ F/m).

With the typical values above, $V_{bi} \approx 0.64$ V and $V_P \approx 3.4$ V.

In integrated circuits, a JFET would also have a gate to substrate capacitance C_{gss} , which, accounting for an abrupt junction, can be expressed as:

$$C_{gss} = C_{gss0} \left(1 + \frac{V_{gss}}{V_{bi}} \right)^{-\frac{1}{2}} \quad (5.2.56)$$

A typical zero bias range of values for these capacitances is:

$$C_{gd0} = 0.3\text{--}1 \text{ pF}$$

$$C_{gs0} = 1\text{--}4 \text{ pF}$$

$$C_{gss0} = 4\text{--}8 \text{ pF}$$

So a JFET with a transconductance $g_m = 2.5 \times 10^{-3}$ A/V and a total gate capacitance $C_T = C_{gd} + C_{gs} + C_{gss} = 4$ pF (under appropriate bias) would yield a cut off frequency:

$$f_T = \frac{1}{2\pi} \cdot \frac{g_m}{C_T} = \frac{2.5 \times 10^{-3}}{6.28 \times 4 \times 10^{-12}} \approx 100 \text{ MHz} \quad (5.2.57)$$

For MOSFETs, the situation is slightly different. [Fig. 5.2.20](#) shows a typical n -channel MOS transistor cross-section and the equivalent circuit model.

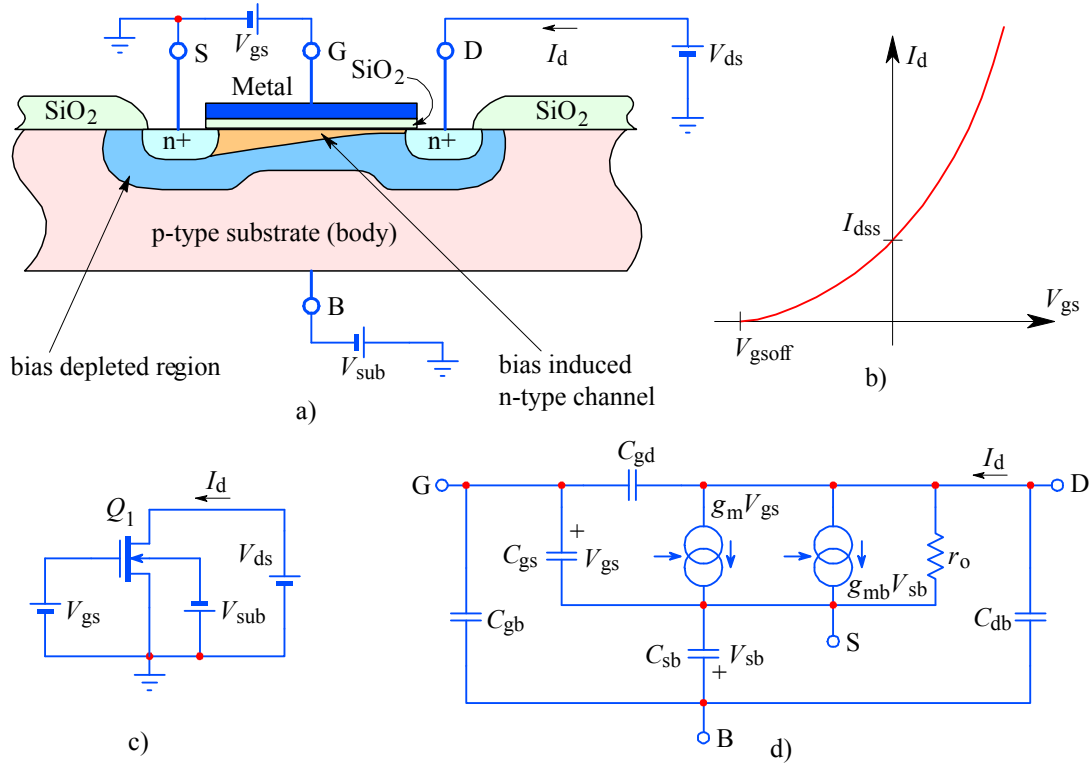


Fig. 5.2.20: **a)** A typical n-channel MOSFET structure's cross-section under bias condition. Two heavily doped n+ regions (source and drain) are manufactured on a p-type substrate and a metal gate covers a thin insulation layer, slightly overlapping the n+ regions. The bias voltage depletes a thick region in the substrate, within which an n-type channel is induced between the source and the drain. **b)** The V_{gs} – I_d characteristic. **c)** The symbolic circuit. **d)** The equivalent circuit model has two current sources, one owed to the usual mutual transconductance g_m and the gate–source voltage V_{gs} ; the other is owed to the so called ‘body effect’ transconductance g_{mb} and the associated source–body voltage V_{sb} . The g_{mb} is typically an order of magnitude lower than g_m .

From the MOSFET structure cross-section it can be deduced that C_{gb} is small, owing to the relatively large depleted region. Ordinarily its value is about 0.1 pF and it is relatively constant. Likewise the depletion region capacitances C_{sb} and C_{db} are also small (they are proportional to the gate and source area), but they are voltage dependent:

$$C_{sb} = C_{sb0} \left(1 + \frac{V_{sb}}{V_{bi}} \right)^{-\frac{1}{2}} \quad (5.2.58)$$

$$C_{db} = C_{db0} \left(1 + \frac{V_{db}}{V_{bi}} \right)^{-\frac{1}{2}} \quad (5.2.59)$$

The capacitances C_{gs} and C_{gd} are owed to the SiO_2 insulation layer between the gate and the channel. If S_g is the gate area and C_x is the unit area capacitance of the oxide layer under the gate, then the total capacitance is:

$$C_{gs0} + C_{gd0} = S_g C_x \quad (5.2.60)$$

Most MOSFETs are built with symmetrical geometry, thus the total zero bias capacitance is simply split in half. But in the saturation region the channel narrows, so

the drain voltage influence is small, resulting in a nearly constant C_{gd} whose value is essentially proportional to the small gate–drain overlapping area. Thus typical C_{gd} values range between 0.002 and 0.020 pF.

C_{gs} is larger, typically some 2/3 of the $S_g C_x$ value, or about 1–2 pF.

Although MOSFETs' g_m is typically lower than in JFETs, it is the very small capacitances, in particular C_{gd} and C_{gb} , which are responsible for the wider bandwidth of a MOSFET source follower. Cut off frequencies of many GHz are easily achieved.

5.2.7 Input Protection Network

The input protection network is needed for two distinct real life situations. The first one is the (occasional) electrostatic discharge, the second one is a long term overdrive.

Imagine a technician sitting on a well insulated chair, wearing woolen or synthetic clothes, and rubber plated shoes, repairing a circuit on his bench. For a while he rubs his clothes on the chair by reaching for the schematic, the spare parts, some tools, etc., thus quickly charging himself up to an average 500 V. Suddenly, he needs to put a 1:1 'scope probe somewhere on the rear panel and he stands up, touching the probe to identify its trace by the characteristic capacitive AC mains pickup. By standing up, he has increased his average distance from the chair by a large factor, say 30, but the charge on the chair and his clothes remains unchanged. This is equivalent to charging a parallel plates capacitor and then increasing the plates distance, so that the capacitance drops inversely to the distance ([Eq. 5.2.51](#)). Because $V = Q/C$, his effective voltage would increase 30 times, reaching some 15 kV!

The average capacitance of the human body towards the surroundings of an average room is about 200 pF. So, when our technician touches the probe tip, he will discharge the 15 kV of his 200 pF right into the input of the poor 'scope. And such a barbaric act can be repeated hundreds of times during an average repairing session.

At the instant the probe tip is touched the effective input voltage falls for the first 5 ns (the propagation delay of the 1 m long probe cable) to a level set by the resulting capacitive divider $a = 1/(1 + C_{cable}/C_{body})$, so if $C_{cable} = 100$ pF/m, $V = a V_{body} \approx 10$ kV. Here we assume a signal propagation velocity of 0.2 m/ns (about 2/3 of the speed of light). Also, note that the probe cable is made as a lossy transmission line (the inner conductor is made of a thin resistive wire, about 50 Ω /m).

After 5 ns the cable capacitance is fully charged and the signal reaches the spark gap. The spark gap fires, limiting the input voltage to its own breaking voltage (1500–2000 V), providing a low impedance path to ground and discharging $C_{cable} + C_{body}$. Some 25 ns later the voltage falls below the spark threshold.

Now the total capacitance $C_{cable} + C_{body} + C_{in}$ is discharged into the remaining input resistance. With the attenuator set to highest sensitivity (1:1), the input resistance is equal to the 1 M Ω of R_{in} , in parallel to the series connection of the damping resistor R_d and one of the protection diodes (depending on the voltage polarity). The diode must withstand a peak current $I_{dpk} = V_{spark}/R_d$; if $R_d = 150$ Ω , then $I_{dpk} = 2000/150 = 13.3$ A! Fortunately, the peak current is lowered also by the loop inductance. The spark discharges the capacitance in less than 30 ns and then the

current falls exponentially as the total capacitance is discharged through R_d , which lasts another 250 ns. At this time the voltage is lower than $V_{cc} + V_{D1}$ and the capacitance is discharged through R_{in} .

In Fig. 5.2.21 we have plotted the first 250 ns of the discharge event, along with the schematic, showing only the most important circuit components.

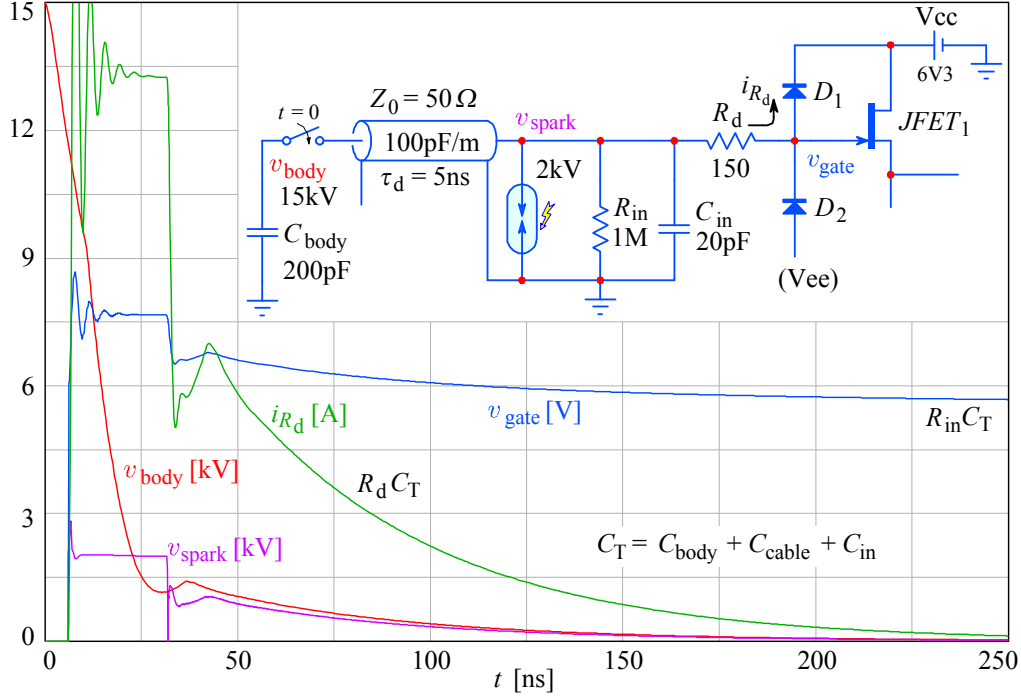


Fig. 5.2.21: A human body model of electrostatic discharge into the oscilloscope input. About 5 ns after touching the probe tip the probe cable is charged and the voltage reaches the spark gap. The spark gap fires and limits the voltage to its firing threshold. The arc provides a low impedance path discharging the body and cable capacitance until the voltage falls below the firing threshold (~25 ns). The remaining charge is fed through R_d and one of the protection diodes, until the voltage falls below $V_{cc} + V_{D1}$ (~250 ns). Finally, the capacitance is discharged through R_{in} .

A different situation occurs in case of a long term overdrive. Fig. 5.2.22 shows the protection network.

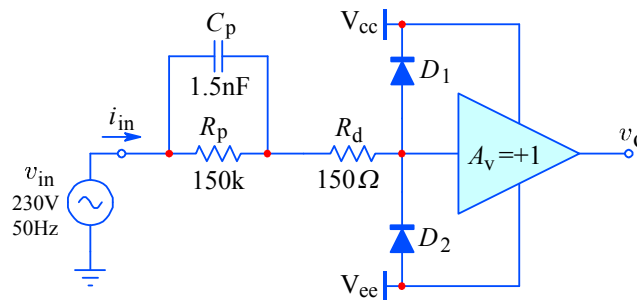


Fig. 5.2.22: Input protection network for long term overdrive: R_p limits the DC and LF current; at HF it is bypassed by C_p , leaving R_d to limit the input current.

The most severe long term input overdrive occurs when the oscilloscope input is on its highest sensitivity setting (no attenuation) and the user inadvertently connects a 1:1 probe to a high voltage DC or AC power supply. A typical highest sensitivity setting of 2 mV/div or ± 8 mV range is brutally exceeded by the 230 V_{eff}, 650 V_{pp} AC mains voltage. Since with a well designed instrument nothing dramatic would happen

(no flash, no bang, no smoke), the user might realize his error only after a while (a few seconds at best and several minutes in the evening at the end of a long working day). The instrument must be designed to withstand such a condition for indefinitely long.

With component values as in [Fig. 5.2.22](#) the peak current through R_p is:

$$I_{Rpk} = (V_{inpk} - V_{cc})/R_p = (325 - 10)/150 \times 10^3 = 2.1 \text{ mA}$$

and the peak current through C_p :

$$I_{Cpk} = (V_{inpk} - V_{cc})\omega C_p = (325 - 10)(2\pi \times 50 \times 1.5 \times 10^{-9}) = 0.15 \text{ mA}$$

Of course, I_C leads I_R in phase by $\pi/2$, so the total current through R_p is the vector sum, $\sqrt{I_C^2 + I_R^2}$, and its value is essentially that of I_R , since the mains frequency (50–60 Hz) is much lower than the network cutoff, $1/(2\pi C_p R_p)$ or 707 Hz.

One could easily be unimpressed by such low current values, however we must not forget the transient conditions. With abundant help from Mr Murphy, we shall make the connection at the instant when the mains voltage is at its peak. And Mr Gauss will ensure a 50% probability that the instantaneous voltage will be above the effective value. Then the input current is limited by R_d only (2.1 A peak!). Fortunately the current falls exponentially with the $R_d C_p$ time constant (225 ns), so the transient is over in about 1 μ s. The value of C_p should not be too low, either; note that it forms a capacitive divider with the JFET input capacitance and ground strays. If these are about 1.5 pF, the high frequency gain will be lower than at DC by about 0.1 %. All these components must be specified to survive voltage transients of at least 500 V, so their larger physical dimensions will increase the circuit size, and as a consequence the parasitic loop inductance and stray capacitances. We should also not forget to account for the JFET's negative input impedance compensation components, as discussed already in [Part 3](#) (see also [\[Ref. 5.69\]](#)).

Note also that by using a $\div 2$ basic high- Z input attenuation, the 150 k Ω and 1.5 nF can be safely omitted because the attenuator takes over their function.

5.2.8 Driving the Low Impedance Attenuator

The high impedance attenuator, discussed in [Sec. 5.2.1](#), is almost exclusively implemented as a two or three decade switch. The intermediate attenuation and gain settings of the 1–2–5 sequence vertical sensitivity selector are usually realized in the stages following the FET source follower. For highest bandwidth the 1–2–5 attenuator is designed as a 50 Ω resistive divider and there are some advantages (regarding the linear signal handling range) if this attenuator is put immediately after the FET source follower. However, the FET by itself can not drive such a low impedance load and additional circuitry is required to help it to do so.

An interesting solution is shown in [Fig. 5.2.23](#), patented by *John L. Addis* in 1983 [\[Ref. 5.21\]](#).

The input FET Q_1 is biased by the constant current source Q_2 , as we have seen in [Fig. 5.2.13](#). It is also actively compensated for large signal transient nonlinearity (by C_2 and Q_4) and bootstrapped by Q_3 , which reduces the input capacitive loading by

keeping the voltage at C_{gd} nearly constant. The output complementary emitter follower is driven from the Q_1 source and R_3 (R_3 should be equal to R_4 to reduce the DC offset). The Q_3 bootstrap is provided by DZ_2 , which, together with DZ_3 , also bootstrap the bias circuit ($R_{9,10}$ and $D_{3,4}$) for the bases of Q_5 and Q_6 , lowering in this way the load to Q_1 source.

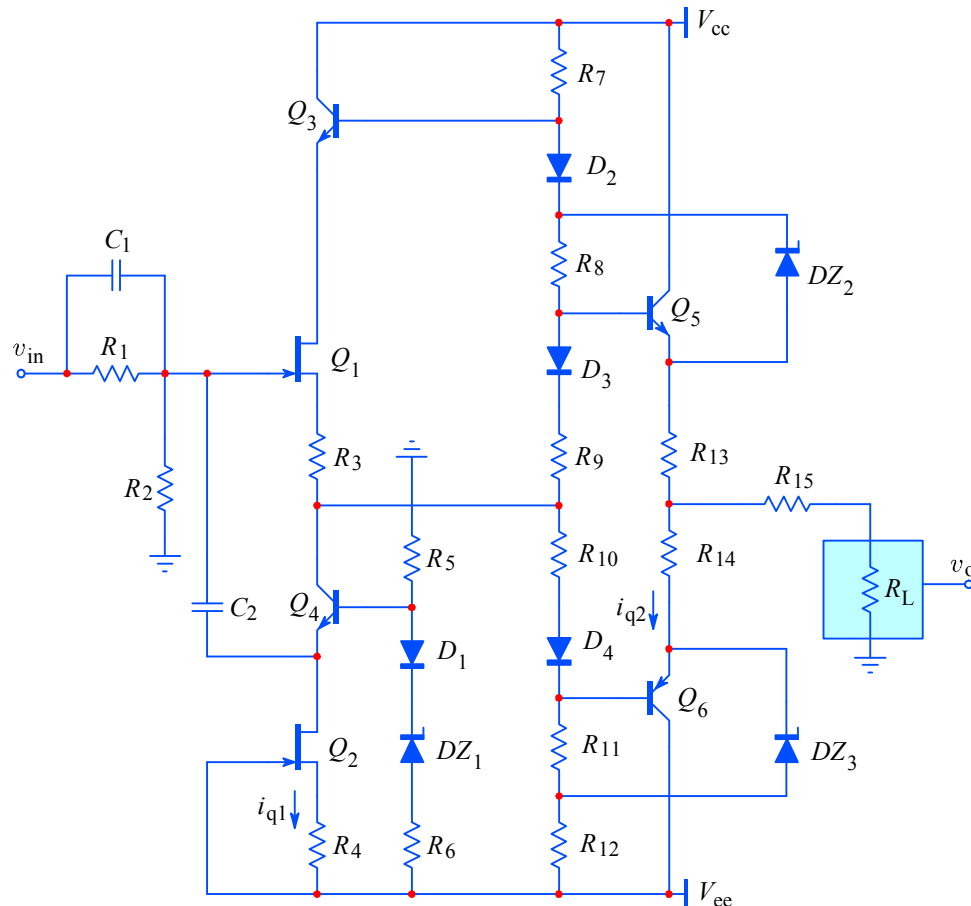


Fig. 5.2.23: The FET source follower with a buffer (patented by J. Addis in 1983) has a dynamic nonlinearity compensation and is capable of driving a low impedance load.

Bootstrapping increases the DC and low frequency impedance seen by Q_1 , but note that its use will make sense at high frequencies only if the $Q_{5,6}$ and Q_3 are substantially faster than the FET itself. Otherwise, the bootstrap circuitry would only increase the parasitic capacitances and thus increase the rise time.

With enough driving capability made available by Q_5 and Q_6 the load resistor R_L can now be realized as the low impedance three step attenuator with a direct path and two attenuated paths, $\div 2$ and $\div 4$. Besides the usual maximum input sensitivity of 5 mV/div., this attenuator will provide the next two settings of 10 and 20 mV/div. The following lower sensitivity settings (50 mV/div., etc.) are achieved by switching in the $\div 10$ and $\div 100$ sections of the high impedance input attenuator, achieving the lowest sensitivity of 2 V/div. An external $\div 10$ probe will decrease this to 20 V/div.

[Fig. 5.2.24](#) shows two possible implementations of the low impedance divider, having $50\,\Omega$ impedance at both input and output. The first attenuator design is based

on the τ -type network and the other on the π -type network. If the input signal is a current, the series $50\ \Omega$ in the $\div 1$ branch can be omitted.

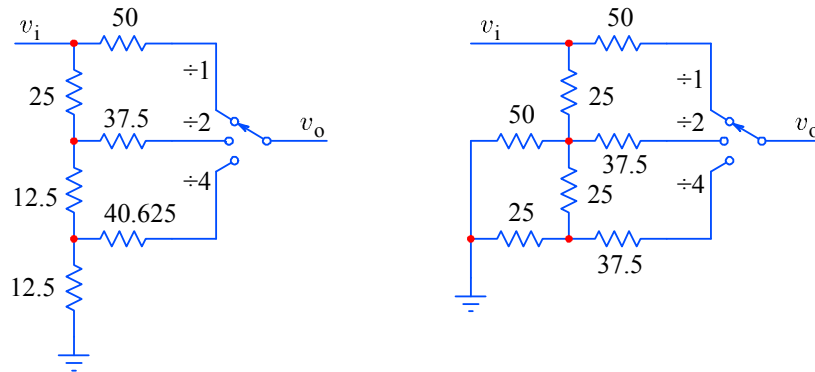


Fig. 5.2.24: The low impedance attenuator ($50\ \Omega$ input and output) can be built as a straight τ -type ladder or a π -type ladder. If driven by a current source the series $50\ \Omega$ in the $\div 1$ branch can be omitted.

Here we assume that the input impedance of the following amplifier stage is high enough and its input capacitance is low enough for the division factor to remain correct at each setting and the bandwidth does not change. It is also important to keep the switch capacitive cross-talk low and preserve the nominal impedance by designing it as a microstrip transmission line.

Unfortunately, placing the low impedance attenuator immediately after the unity gain buffer would seriously degrade the noise performance at some attenuator settings. It is therefore better to move the attenuator after a further amplifying stage. But the ability to drive a $50\ \Omega$ load is nevertheless useful, e.g., in Tektronix 11A32 and 11A34 the input buffer had to drive a 25 cm long coaxial cable to the main board where the next stages were located (the M377 chip, described in [Sec. 5.4.6](#), where we shall also see other ways of implementing the gain/sensitivity switching).

In addition to the discrete step attenuation, oscilloscopes, as well as other high speed instruments, often need a continuously variable attenuation (or gain), although within a restricted range (a range of 0.3 to 1 is often enough). A passive potentiometer is, of course, an obvious solution and it was used extensively in early days. However, this potentiometer is usually placed somewhere in the middle of the amplifier and its control shaft has to be brought to the instrument front panel, which can often be a mechanical nightmare. Also, its variable impedance causes the bandwidth to vary and this is a very undesirable property. An electronically controlled amplifier gain with constant bandwidth would therefore be welcome. We shall examine such circuits at the end of [Sec. 5.4](#).

(blank page)

5.3 High Speed Operational Amplifiers

From about 1980 we have been witnessing both the development of a radically different operational amplifier topology and a major improvement in complementary semiconductor devices' technology, resulting in a steep rise in performance. At the same time, the market's hunger for higher bandwidth has been met by a massive production increase, so that the prices have remained fairly low. With the accompanying development in digital technology, both in terms of switching speed and circuit complexity, the techniques which have previously been forbiddingly expensive and too demanding to realize suddenly became feasible and within reach.

At the turn of the century opamps with the unity gain \times bandwidth product of about 1 GHz or more (such as the Burr–Brown OPA-640) became available at a price comparable to that of a couple of good discrete high frequency transistors. Add to this a relatively good DC performance and, using surface mounting devices, a circuit area of $\sim 1 \text{ cm}^2$, a low power consumption, and a noise level comparable to the thermal noise of a 100Ω resistor, the advantages are obvious.

Clearly, we have come a long way from the ubiquitous $\mu\text{A}741$.

However, in order to better evaluate the performance and the design requirements of the new devices we shall first expose the weak points of the classical configuration.

5.3.1 The Classical Opamp

The name 'operational amplifier' springs from the analog computer era, in which amplifier blocks could be combined with passive components to perform various mathematical operations, from simple signal addition and subtraction to integration and differentiation.

The main performance limitation of early IC opamps was imposed by the integration technology itself: whilst fairly good NPN transistors could be easily produced, PNP transistors could be made along with NPN ones only as very slow, 'lateral' structures. This has restricted their use to only those parts of the opamp, where the needed bandwidth, gain, and load could be low. In practical terms, the only such place in a typical opamp is the so called 'middle stage level translator', as shown in [Fig. 5.3.1](#). Even so, the opamp open loop bandwidth was almost always below 100 Hz, mainly owing to the Miller effect and the need to provide enough phase margin at low closed loop gain to ensure unconditional system stability.

On the other hand, for general purpose applications, the important parameter was a high negative feedback factor, used to minimize the circuit performance variations owed to the transistor parameters (which in the early days were difficult to control) and instead rely on passive components which could be easily produced with a relatively tight tolerance. It was this ability to deliver predictable performance by a simple choice of two feedback resistors which made the opamp a popular and widely used circuit component. And not only a well defined gain, but also a broad range of other signal conditioning functions is made possible by combining various passive and active components in the feedback loop.

The feedback concept is so simple and works so well that too many people take it for granted; and equally many are surprised to discover that it can cause as much trouble as the solutions it offers.

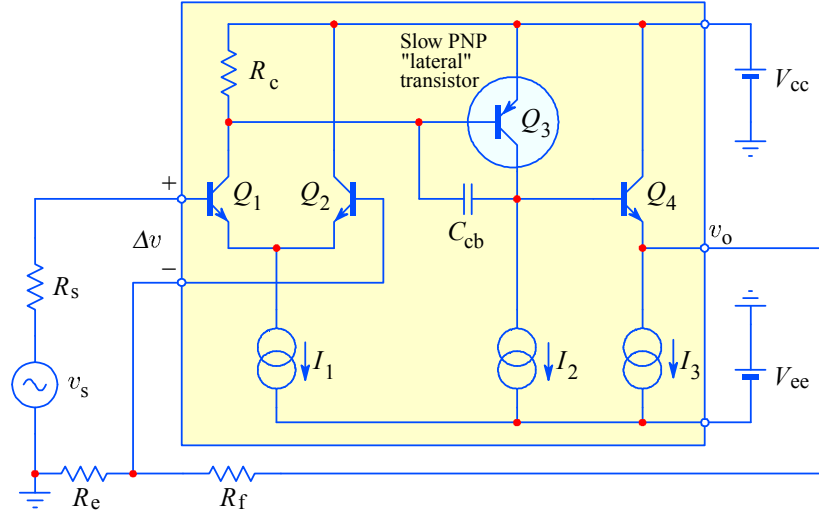


Fig. 5.3.1: The classical opamp, simplified. The input differential pair Q_1 and Q_2 subtract the feedback from the input signal, driving with the difference the ‘level translator’ stage Q_3 , which in turn drives the output emitter follower Q_4 , which provides a low output impedance. The feedback voltage is derived from the output voltage v_o by dividing it as $\beta = R_e/(R_e + R_f)$. If the opamp open loop gain A_o is much higher than $1/\beta$, the closed loop gain is $A_{cl} \approx 1/\beta$ (see text).

In the (highly simplified) opamp circuit in [Fig. 5.3.1](#) the open loop gain is equal to the gain of the differential pair Q_1 and Q_2 , multiplied by the gain of the level translator Q_3 . The output emitter follower Q_4 has a unit voltage gain. However, all the three stages have their gain frequency dependent, as was explained in [Part 3](#). Fortunately, the three poles are far apart (all are real) and the poles of the first and the third stage can be (and usually are) easily set high enough for the amplifier open loop frequency response to be dominated by the second stage pole (which was in turn named ‘the dominant pole’).

The dominant pole of the circuit in [Fig. 5.3.1](#) is set by the Q_1 collector resistor R_c and the Miller capacitance C_M :

$$s_0 = - \frac{1}{R_c C_M} \quad (5.3.1)$$

where we have neglected the input resistance of Q_3 in parallel with R_c , which we can do if R_c is small.

C_M appears effectively in parallel with R_c and its value is equal to the Q_3 collector to base capacitance C_{cb} , multiplied by the Q_3 gain:

$$C_M = C_{cb}(1 + A_3) \quad (5.3.2)$$

The gain A_3 is set by the Q_3 transconductance and the loading resistance, in the form of the equivalent (shunt) input resistance at the base of Q_4 :

$$A_3 \approx g_{m3} R_{b4} = \frac{q_e I_2}{k_B T} R_{b4} \quad (5.3.3)$$

The equivalent input resistance of Q_4 is approximately equal to the amplifier loading resistance reflected into the Q_4 base:

$$R_{b4} = R_L(1 + \beta_4) \quad (5.3.4)$$

where β_4 is the Q_4 current gain.

The gain of the input differential pair is set by the Q_1 collector load resistor R_c and the transconductance of both Q_1 and Q_2 (again neglecting the Q_3 input resistance):

$$A_1 \approx (g_{m1} + g_{m2}) R_c = \frac{q_e I_1}{k_B T} R_c \quad (5.3.5)$$

where we have assumed both transconductances to be equal, owing to the current I_1 being equally divided into I_{c1} and I_{c2} .

As a result the open loop transfer function can be written as:

$$A(s) = A_0 \frac{-s_0}{s - s_0} = A_1 A_3 \frac{-s_0}{s - s_0} \quad (5.3.6)$$

Now we can derive the closed loop transfer function. By considering that the feedback factor β is set by the feedback resistive divider:

$$\beta = \frac{R_e}{R_e + R_f} \quad (5.3.7)$$

the voltage at the inverting input is equal to the output voltage multiplied by the feedback factor:

$$v_i = v_o \beta \quad (5.3.8)$$

The signal being amplified is the difference between the source voltage and the voltage provided by feedback:

$$\Delta v = v_s - v_i \quad (5.3.9)$$

This voltage is amplified by the amplifier open loop transfer function, $A(s)$, to give the output voltage:

$$v_o = A(s) \Delta v \quad (5.3.10)$$

By considering [Eq. 5.3.6](#) and [5.3.9](#), we can write:

$$v_o = A_0 \frac{-s_0}{s - s_0} (v_s - v_i) \quad (5.3.11)$$

and, since v_i is a feedback scaled v_o :

$$v_o = A_0 \frac{-s_0}{s - s_0} (v_s - v_o \beta) \quad (5.3.12)$$

By rearranging this into:

$$v_o \left(1 + \beta A_0 \frac{-s_0}{s - s_0} \right) = A_0 \frac{-s_0}{s - s_0} v_s \quad (5.3.13)$$

we can obtain the explicit expression for v_o :

$$v_o = \frac{A_0 \frac{-s_0}{s - s_0}}{1 + \beta A_0 \frac{-s_0}{s - s_0}} v_s \quad (5.3.14)$$

or:

$$v_o = \frac{1}{\frac{1}{A_0 \frac{-s_0}{s - s_0}} + \beta} v_s \quad (5.3.15)$$

From this last expression it is obvious that if the open loop gain A_0 is very high the amplifier gain v_o/v_s is reduced to the familiar $1/\beta$, or $(R_f + R_e)/R_e$.

Likewise, for a finite value of A_0 , the frequency dependent part increases, thus lowering the closed loop gain at higher frequencies.

Take, for example, the $\mu A741$ opamp, [Fig 5.3.2](#): owing to its dominant pole, the open loop cut off frequency is at about 10 Hz, whilst the open loop gain at DC is about 10^5 . The unity gain crossover frequency f_1 is therefore about 1 MHz.

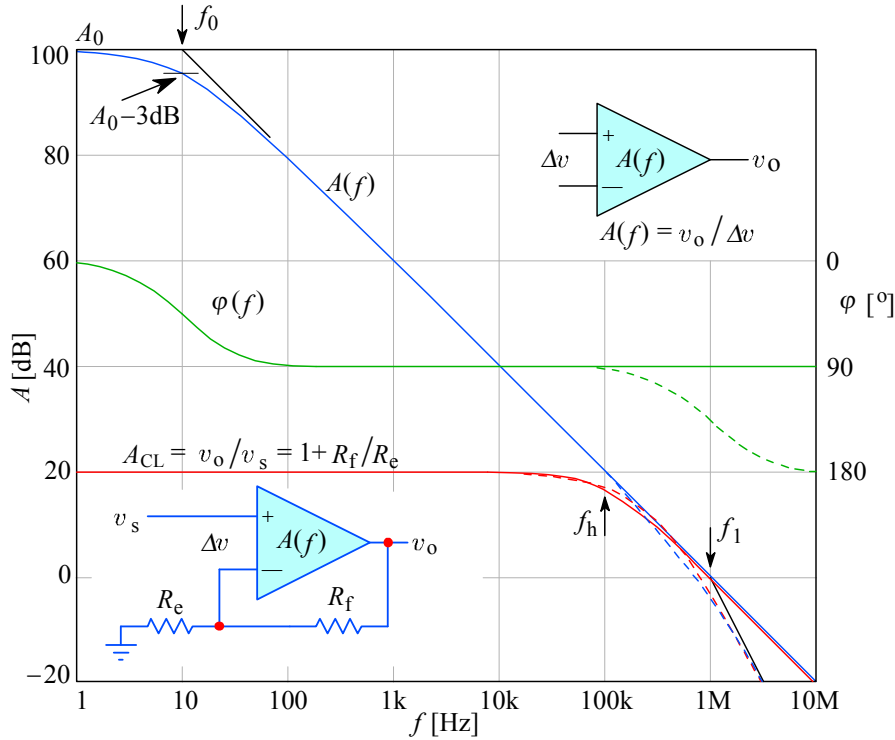


Fig. 5.3.2: A typical opamp open loop gain and phase compared to the closed loop gain. The dashed lines show the influence of a secondary pole (usually the input differential stage pole), which, for stability requirements, must be set at or above the unity gain transition frequency, $f_1 = 1$ MHz. f_h is the closed loop cutoff frequency.

For a closed loop gain of 10, $\beta = 0.1$; since the frequency dependence term is a ratio, the factor 2π can be extracted and canceled, leaving $f_0/(jf + f_0)$, where f_0 is the open loop cutoff frequency. By putting this into [Eq. 5.3.15](#), we see that the amplifier will be making corrections to its own non-linearity by a factor 10^4 (80 dB) at 1 Hz, but only by a factor of 10^2 (40 dB) at 1 kHz; and at 100 kHz there would be only

3 dB of feedback, resulting in a 50% gain error. This means that for a source signal of 0.1 V there would be a Δv of 0.05 V, resulting in an output voltage of $v_o = 0.5$ V (instead of the 1 V as at low frequencies). Above the closed loop cutoff frequency the amplifier has practically no feedback at all.

An additional error is owed to the phase shift: at 100 kHz a single pole amplifier would have the output at 90° phase lag against the input. An amplifier with an additional input differential stage pole at 1 MHz would shift the phase by 135°, so there would be only a 45° phase margin at this frequency and the circuit would be practically at the edge of closed loop stability. If we were to need this amplifier to drive a 2 m long coaxial cable (capacitance 200 pF), by considering the amplifier output impedance of about 75 Ω the additional phase shift of 5° would be enough to turn the amplifier into a high frequency oscillator.

5.3.2 Slew Rate Limiting

The discussion so far is valid for the small signal amplification. For large signals the bandwidth would be much lower than the small signal one. This is owed to the Miller capacitance causing Q_3 to act as an integrator. For a positive input step larger than $2 k_B T / q_e$ ($+I_1 R_{e1}$ if the input differential pair has emitter degeneration resistors), the transistor Q_1 will be fully open, while Q_2 will be fully closed. Therefore, the maximum current available to charge C_M will be equal to the tail current I_1 . The voltage across C_M will increase linearly until the input differential stage will be out of saturation. Consequently, the slew rate limit is:

$$SR = \frac{dv}{dt} = \frac{I_1}{C_M} \quad (5.3.16)$$

Usually I_1 is of the order of 100 μ A (or even lower if low noise is the main design goal). Also, owed to the gain of Q_3 the Miller capacitance C_M can be large; say, with $C_{cb} = 4$ pF and $A_3 = 50$, C_M will be about 200 pF, giving a slew rate $SR = 0.5$ V/ μ s. We know that for a sine wave the maximum slope occurs at zero crossing, where the derivative is $dv/dt = d(V_p \sin \omega t)/dt = \omega V_p \cos \omega t$; at zero crossing $t = 0$ and $\cos(0) = 1$, so the slew rate equation can be written as:

$$SR = \omega V_p = \frac{I_1}{C_M} \quad (5.3.17)$$

For a supply voltage of ± 15 V, the signal amplitude just before clipping would probably be around 12 V, so the maximal full power sine wave frequency would be $f_{\max} = I_1 / 2\pi C_M V_p$, or approximately 6.5 kHz only!

The frequency at which the sine wave becomes a linear ramp, with a nearly equal peak amplitude, is slightly higher: $f_r = 1/4t_r = SR/4 V_p = 10$ kHz (note that the SR of the circuit in [Fig. 5.3.1](#) is not symmetrical, since C_M is charged by I_1 and discharged through R_c ; in an actual opamp circuit, such as in [Fig. 5.3.3](#), R_c is replaced by a current mirror, driven by the collector current of Q_2 , giving a symmetrical slew rate).

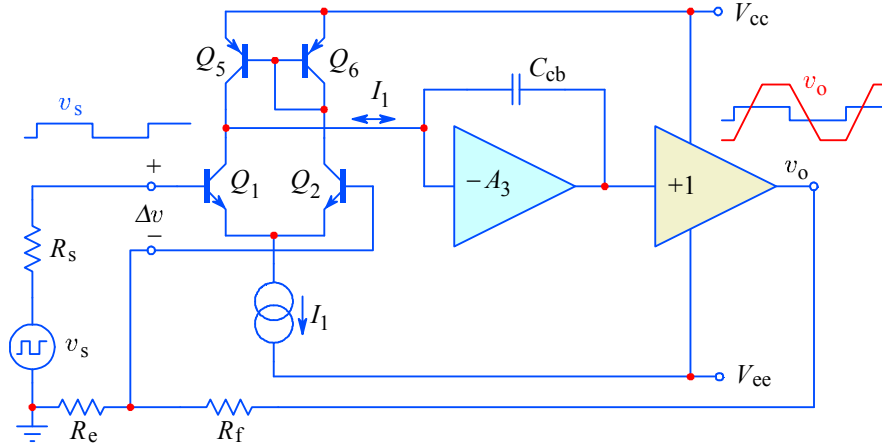


Fig. 5.3.3: Simplified conventional opamp circuit with the current mirror as the active load to Q_1 . The second stage is modeled as a Miller integrator with large gain. This circuit exhibits symmetrical slew rate limiting. The dominant pole $s_0 = -g_m/C_M$, where g_m is the differential amplifier's transconductance and $C_M = C_{cb}(1 + A_3)$.

5.3.3 Current Feedback Amplifiers

The circuit in [Fig. 5.3.1](#) could be characterized as a 'voltage feedback' amplifier and in the previous analysis we have shown its most important performance limitations. Instead the circuit in [Fig. 5.3.4](#) is characterized as a 'current feedback' amplifier, since the feedback signal is in the form of a current, which, as will become evident soon, offers several distinct advantages.

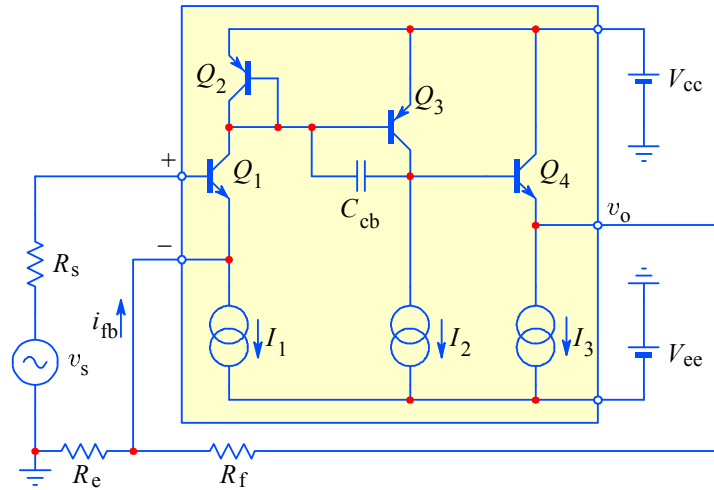


Fig. 5.3.4: Current feedback opamp, derived from the voltage feedback opamp ([Fig. 5.3.1](#)): we first eliminate Q_2 from the input differential amplifier and introduce the feedback into the Q_1 emitter (low impedance!). Next, we load the Q_1 collector by a diode connected Q_2 , forming a current mirror with Q_3 . Finally, we use very low values for R_f and R_e . The improvements in terms of speed are two: first, for large signals, the current available for charging C_{cb} is almost equal to the feedback current i_{fb} , eliminating slew rate limiting; second, C_{cb} is effectively grounded by the low impedance of Q_2 , thus avoiding the Miller effect. A disadvantage is that the voltage gain is provided by Q_3 alone, so the loop gain is lower. Nevertheless, high frequency distortion can be lower than in classical opamps, because, for the equivalent semiconductor technology, the dominant pole is at least two decades higher, providing more loop gain for error correction at high frequencies.

The amplifier in Fig. 5.3.4 would still run into slew rate limiting for high amplitude signals, owing to the fixed bias of the first stage current source I_1 . This is avoided by using a complementary symmetry configuration, as shown in Fig. 5.3.5. Of course, the complementary symmetry can be used throughout the amplifier, not just in the first stage.

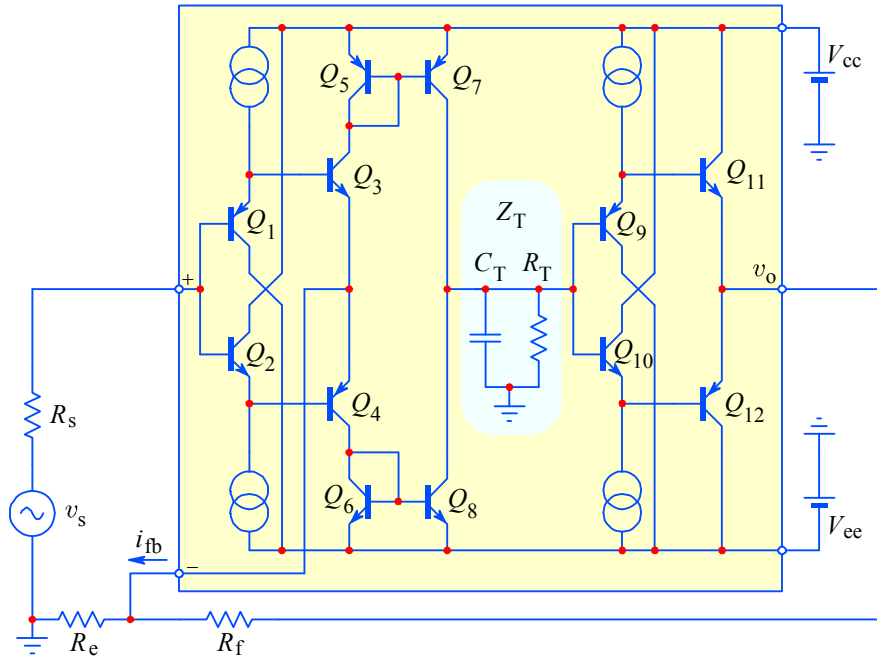


Fig. 5.3.5: A fully complementary current feedback amplifier model. It consists of four parts: transistors Q_{1-4} form a unity gain buffer, the same as Q_{9-12} , with the four current sources providing bias; $Q_{5,7}$ and $Q_{6,8}$ form two current mirrors. In contrast to the voltage feedback circuit, both of whose inputs are of high impedance, the inverting input of the current feedback amplifier is a low impedance output of the first buffer. The current flowing in or out of the emitters of $Q_{3,4}$ is (nearly) equal to the current at the $Q_{3,4}$ collectors. This current is reflected by the mirrors and converted into a voltage at the $Q_{7,8}$ collectors, driving the output unity gain buffer. The circuit stability is ensured by the transimpedance Z_T , which can be modeled as a parallel connection of a capacitor C_T and resistor R_T . The closed loop bandwidth is set by R_f and the gain by R_e (the analysis is presented later in the text). One of the first amplifiers of this kind was the Comlinear CLC400.

Perhaps, it would be more correct to label the structure in Fig. 5.3.5 as a ‘current on demand’ type of amplifier, owing to the fact that the feedback current, which is proportional to the input–output error, feeds the dominant pole capacitance. The larger the error, the larger the current, which practically eliminates the slew rate limiting. The slew rate will be limited nevertheless, due to secondary effects, which will be discussed later, but the maximum current charging C_T is usually much greater than in conventional amplifiers. Also, C_T is small (not affected by the Miller effect).

Another name often found in the literature is the ‘transimpedance amplifier’, after the transimpedance equation (Z_T , see the analysis below). Owing to historical reasons, we shall keep the name used in the section title.

The complementary symmetry nature of the circuit in Fig. 5.3.5 would have been difficult to realize with the available opamp integration technology of the late 1960s and early 1970s, owing to the different characteristics of PNP and NPN transistors. A major technological breakthrough, made between 1980 and 1985 at

Comlinear Corporation and Elantec, later followed by Analog Devices, Burr–Brown and others, enabled PNP transistors to have their f_T almost as high as the NPN had. [Fig. 5.3.6](#) shows a typical chip cross-section and [Table 5.3.1](#) presents the typical values of the most important production parameters improving over the years.

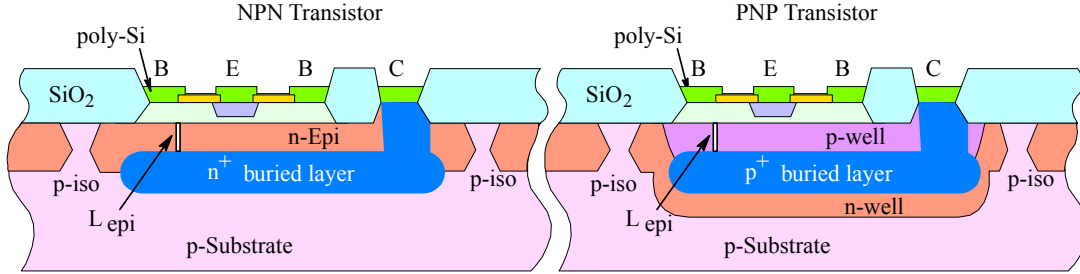


Fig. 5.3.6: Cross-section of the Complementary-Bipolar process.

Process (yr)	VIP1 (1986)		VIP2 (1988)		VIP3 (1994)		VIP10 (2000)		Units
Parameter	NPN	PNP	NPN	PNP	NPN	PNP	NPN	PNP	—
i_c/i_b	250	150	250	80	150	60	100	55	—
Early V_A	200	60	150	40	150	50	120	40	V
f_T	0.4	0.2	0.8	0.5	3	2.5	9	8	GHz
C_{js}	2.0	2.2	1.5	1.8	0.5	0.8	0.005	0.007	pF
E width	15		11		2		1		μm
Area	20000		18000		2400		300		μm^2
$V_{ce \text{ max}}$	36		36		32		12		V

Table 5.3.1: Typical production parameters of the Complementary Bipolar process [[Ref. 5.33](#)]

Although the same technology is now used also for conventional voltage feedback amplifiers, the current feedback structure offers further advantages which result in improved circuit bandwidth, as put in evidence by the following discussion.

The stability of the amplifier in [Fig. 5.3.5](#) is ensured by the so called transimpedance network, Z_T , which can be modeled as a parallel $R_T C_T$ network. Note that the compensating capacitor, C_T (consisting of 4 parts, $C_{T1}-C_{T4}$), is effectively grounded, as can be seen in [Fig. 5.3.7](#), since C_{cb} of $Q_{9,10}$ are tied to the supply voltages directly, whilst the C_{cb} of $Q_{7,8}$ are tied to the supply by the low impedance CE path of $Q_{5,6}$.

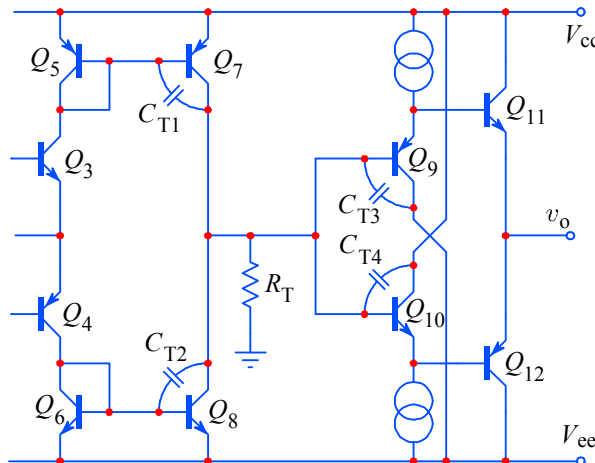


Fig. 5.3.7: The capacitance C_T consists of four components, all effectively grounded.

Therefore in this configuration the Miller effect is substantially eliminated. This means that, for the same driving current, this circuit is capable of much higher bandwidth, compared to the conventional opamp.

Also, owing to the two current mirrors, the current which charges and discharges C_T is equal to the current injected by the feedback network into the inverting input (the first buffer output). Since this current is feedback derived, it is proportional to the input–output error; thus for a fast input voltage step there would initially be a large input–output error, causing a large current into the inverting input and an equally large current will charge C_T , so its voltage, and consequently the output voltage, will increase fast. As the output voltage increases, the error is reduced, reducing the error and lowering the current.

To analyze the circuit operation we shall at first assume ideal buffers and current mirrors, as modeled in [Fig. 5.3.8](#). Later, we shall see how the real circuit parameters limit the performance.

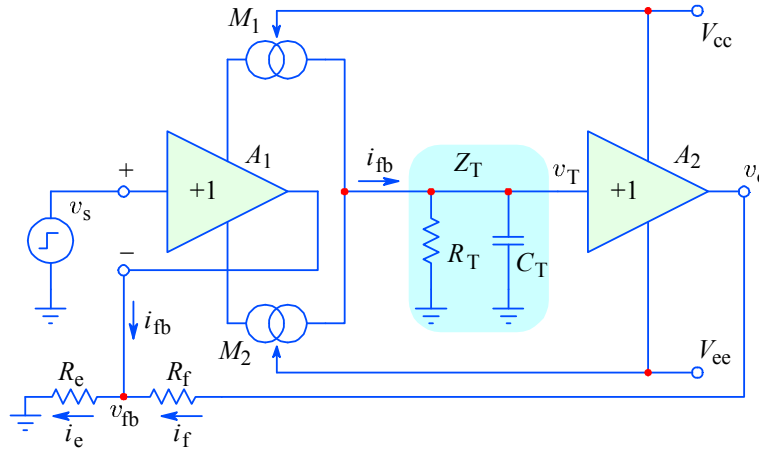


Fig. 5.3.8: Current feedback amplifier model used for the analysis.

Imagine for a moment that R_f is taken out of the circuit. Essentially this would be an open loop configuration, the gain of which can be expressed by the ratio of two resistors, R_T/R_e . The gain is provided by the current mirrors M_M ; their output currents are summed, so the two mirrors behave like a single stage; consequently the gain value, compared with that of a conventional opamp, is relatively low (in practice, maximum gains between 60 and 80 dB are common). It is important to note, however, that the open loop (voltage) gain does not play such an important role in current feedback amplifiers. As the name implies, it is more important how the feedback current is processed.

Let us now examine a different situation: we put back R_f and disconnect R_e . If there were to be any voltage difference between the outputs of the two buffers, a current would be forced through R_f , increasing the output current of the first buffer, A_1 . The two current mirrors would reflect this onto the input of the second buffer, A_2 , in order to null the output voltage difference. In other words, **the output of the first buffer A_1 represents an inverting current mode input** of the whole system.

If we now reconnect R_e it is clear that the A_1 output must now deliver an additional current, i_e , flowing to the ground. The current increase is reflected by the mirrors into a higher v_T , so the output voltage v_o would increase, forcing the current

i_f (through R_f) into the A_1 output. By looking from the A_1 output, i_e flows in the direction opposite to i_f , so the total current i_{fb} of the A_1 output will be equal to their difference. Thus with R_e and R_f a classical feedback divider network is formed, but the feedback signal is a current. As expected, the output of A_2 must now become $(R_f + R_e)/R_e$ times higher than the output of A_1 to balance the feedback loop.

Deriving the circuit equations is simple. The transimpedance equation (assuming an ideal unity gain buffer A_2 , thus $v_T = v_o$) is:

$$v_o = Z_T i_{fb} \quad (5.3.18)$$

The feedback current (assuming an ideal unity gain buffer A_1 , thus $v_{fb} = v_s$) is:

$$i_{fb} = \frac{v_s}{R_e} - \frac{v_o - v_s}{R_f} = v_s \left(\frac{1}{R_e} + \frac{1}{R_f} \right) - \frac{v_o}{R_f} \quad (5.3.19)$$

The closed loop gain (from both equations above) is:

$$\frac{v_o}{v_s} = \left(1 + \frac{R_f}{R_e} \right) \frac{1}{1 + \frac{R_f}{Z_T}} \quad (5.3.20)$$

We see that the equation for the closed loop gain has two terms, the first one resulting from the feedback network divider and the second one containing the transimpedance Z_T and R_f , but not R_e ! This is in contrast to what we are used to in conventional opamps.

If we now replace Z_T by its equivalent network, $1/(s C_T + 1/R_T)$, then the closed loop gain [Eq. 5.3.20](#) can be rewritten as:

$$\frac{v_o}{v_s} = \left(1 + \frac{R_f}{R_e} \right) \frac{1}{1 + \frac{R_f}{R_T} + s C_T R_f} \quad (5.3.21)$$

We can rewrite this to reveal the system's pole, in the way we are used to:

$$\frac{v_o}{v_s} = \frac{1 + \frac{R_f}{R_e}}{1 + \frac{R_f}{R_T}} \cdot \frac{\left(1 + \frac{R_f}{R_T} \right) \frac{1}{C_T R_f}}{s - \left[- \left(1 + \frac{R_f}{R_T} \right) \frac{1}{C_T R_f} \right]} \quad (5.3.22)$$

By comparing this with the general single pole system transfer function:

$$F(s) = A_0 \frac{-s_1}{s - s_1} \quad (5.3.23)$$

we note that the term:

$$s_1 = - \left(1 + \frac{R_f}{R_T} \right) \frac{1}{C_T R_f} \quad (5.3.24)$$

is the closed loop pole, which sets the closed loop cutoff frequency: $f_h = |s_1|/2\pi$.

We also note that the system closed loop gain is:

$$A_0 = \frac{1 + \frac{R_f}{R_e}}{1 + \frac{R_f}{R_T}} \quad (5.3.25)$$

Since R_f is normally much smaller than R_T , the term R_T/R_f represents the open loop gain, so R_f/R_T represents the closed loop gain error (in analogy with the finite open loop gain error at DC in classical amplifiers).

If, for example, we have an amplifier with $R_T = 300 \text{ k}\Omega$ and $C_T = 2 \text{ pF}$ and we form the feedback with $R_f = 330 \text{ }\Omega$ and $R_e = 110 \text{ }\Omega$, the amplifier would have a closed loop bandwidth of about 240 MHz and a gain of 4. Moreover, its loop gain would be $R_T(R_f + R_e)/R_f R_e \approx 3000$ and flat up to some 265 kHz, more than three orders of magnitude higher than the usual 100 Hz in conventional voltage feedback operational amplifiers.

Thus, we find that the closed loop bandwidth depends mainly on R_f (but not on R_e), whilst the gain, once R_f has been chosen, can be set by R_e alone. With current feedback the amplifier designer has independent control over the two most important circuit parameters and must only watch for possible second-order effects.

The benefits of the current feedback amplifier are all due to two main points:

- a) since there is only one voltage gain stage (the two current mirrors, working effectively in parallel) and only one internal high impedance node (Z_T), this structure is inherently a single pole system (since the integration technology allows the poles of both buffers to be much higher). As a consequence, the system may always be made unconditionally stable, whilst not compromising the available system bandwidth;
- b) since the feedback is entered in the form of a current the system bandwidth depends on R_f but not on R_e , so that if R_f is held constant while R_e is varied, the bandwidth remains (almost) independent of gain. As a bonus there is practically no slew rate limiting mechanism, because the feedback current drives a grounded C_T and the larger the input voltage step, the larger will be the current charging C_T . So the amplifier's step response will always look like that of a low pass $R_f C_T$ network for any signal amplitude up to the clipping level.

It might be interesting to return to [Eq. 5.3.19](#) with the result of [Eq. 5.3.21](#) and express the current which charges C_T as a function of input voltage:

$$i_{C_T} \approx i_{fb} = v_s \left(\frac{1}{R_f} + \frac{1}{R_e} \right) \left(1 - \frac{1}{1 + R_f/R_T + s C_T R_f} \right) \quad (5.3.26)$$

Therefore we can express i_{C_T} (see the transient response in [Fig. 5.3.9](#)) as:

$$i_{C_T} \approx i_{fb} = v_s \left(\frac{1}{R_f} + \frac{1}{R_e} \right) \frac{s}{s + \left(1 + \frac{R_f}{R_T} \right) \frac{1}{C_T R_f}} \quad (5.3.27)$$

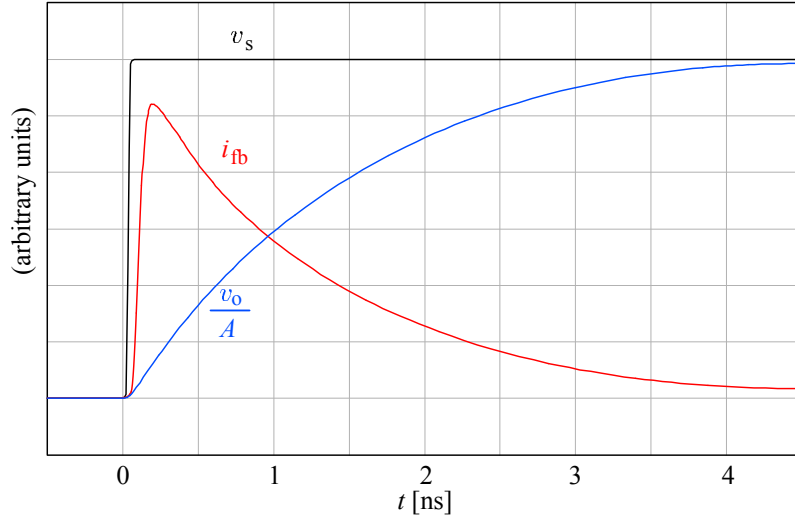


Fig. 5.3.9: ‘Current on demand’: The step response reveals that the feedback current is proportional to the difference between the input and output voltage, essentially a high pass version of the output voltage, as shown by [Eq. 5.3.27](#).

In [Fig. 5.3.10](#) we compare the cut off frequency vs. gain of a voltage feedback and a current feedback amplifier. The voltage feedback amplifier bandwidth is inversely proportional to gain; in contrast, the current feedback amplifier bandwidth is, in principle, independent of gain. This property makes current feedback amplifiers ideal for wideband programmable gain applications.

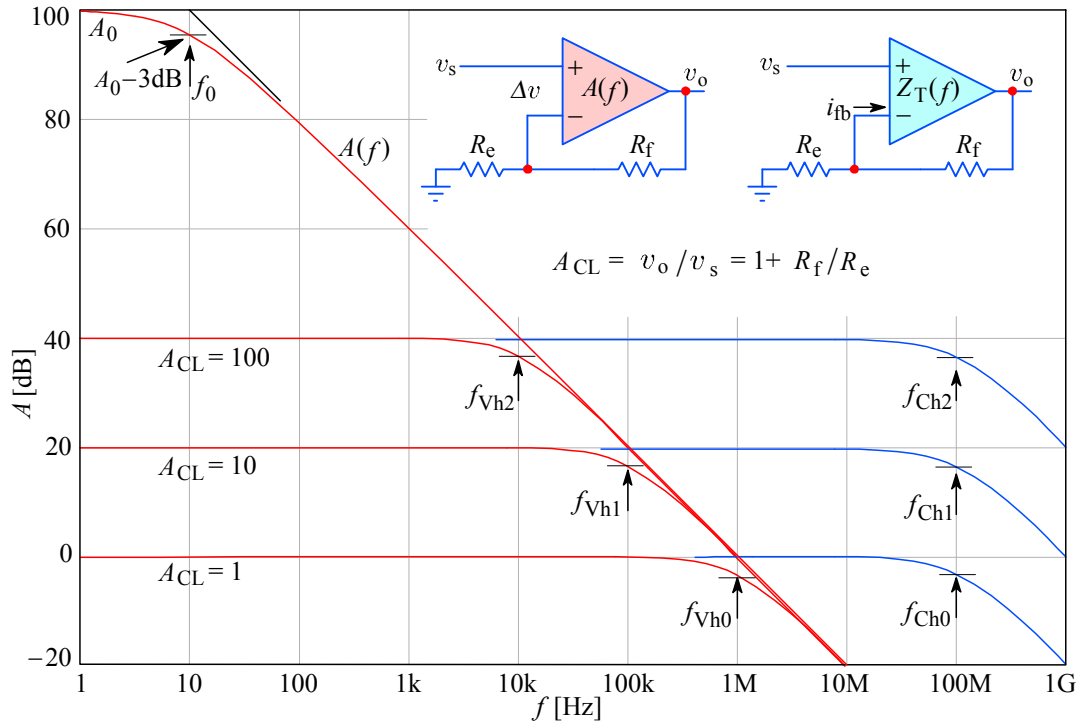


Fig. 5.3.10: Comparison of closed loop cut off frequency vs. gain of a conventional and a current feedback amplifier. The conventional amplifier has a constant GBW product (higher gain, lower cut off). But the current feedback cut off frequency is (almost) independent of gain.

Of course, a real amplifier will have some second-order effects, which we have not discussed so far, so its performance will be somewhat less than ideal.

The main causes for non-ideal performance are:

- the small but finite inverting input resistance (non-zero output impedance of A_1 in Fig. 5.3.8), which causes a voltage error between the non-inverting and the inverting input and, consequently, a lower feedback current through R_f ;
- the non-zero output impedance of A_2 , which, combined with the output load forms an additional feedback divider;
- the asymmetry and non-linearity of the two current mirrors, which directly influences the transfer function;
- the finite current gain of the output stage A_2 , which, if too low, would allow the amplifier load to be reflected at the input of A_2 and influence Z_T ;
- the secondary poles at high frequencies, owed to the finite bandwidth of the transistors within the amplifier.

The last four points are equally well known from conventional amplifiers and their influence is straightforward and easy to understand, so we shall not discuss them any further. The first point, however, deserves some attention.

5.3.4 Influence of a Finite Inverting Input Resistance

The current feedback amplifier requires a low (ideally zero) impedance at the inverting input in order to sense the feedback current correctly. This, in addition to the manufacturing imperfections between the transistors Q_{1-4} (Fig. 5.3.11), results in a relatively high input offset, owed to both DC voltage errors and current errors.

The offset is reduced by using the current mirror technique for the biasing current sources (Q_{a-d}), making the currents of $Q_{1,2}$ equal. Further reduction is achieved by adding low value resistors, R_{1-4} (a value of about $10r_e$ is usually sufficient) to Q_{1-4} emitters. This, however, increases the inverting input resistance.

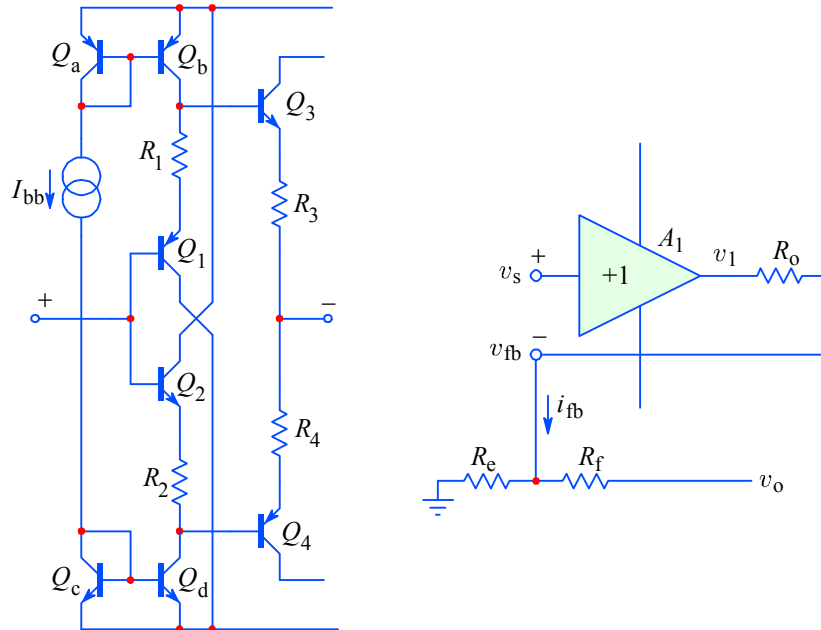


Fig. 5.3.11: The resistors R_{1-4} used to balance the inputs are the cause for the non-zero inverting input resistance, $R_3 \parallel R_4$, modeled by R_o ; this causes an additional voltage drop, reducing the feedback current (see analysis).

A typical circuit of the first buffer implementing DC offset reduction by current mirror biasing and using emitter degeneration resistors is shown in [Fig. 5.3.11](#); the equivalent inverting input resistance $R_3 || R_4$ is modeled by R_o . It causes an additional voltage drop which reduces the feedback current.

For the analysis we assume that the buffer has a unity gain, thus $v_{o1} \approx v_s$. Since the feedback current flows through R_o , the voltage at the inverting input, v_{fb} , will be lower than v_s by $i_{fb}R_o$:

$$v_{fb} = v_s - i_{fb}R_o \quad (5.3.28)$$

By summing the currents at the v_{fb} node we have:

$$\frac{v_o - v_{fb}}{R_f} = \frac{v_{fb}}{R_e} - \frac{v_s - v_{fb}}{R_o} \quad (5.3.29)$$

Note that the last term in this equation is the feedback current from [Eq. 5.3.28](#):

$$i_{fb} = \frac{v_s - v_{fb}}{R_o} \quad (5.3.30)$$

and from the transimpedance equation [Eq. 5.3.18](#) the feedback current required to produce the output voltage v_o is:

$$i_{fb} = \frac{v_o}{Z_T} \quad (5.3.31)$$

By substituting v_{fb} and i_{fb} in [Eq. 5.3.28](#), we obtain the transfer function:

$$\frac{v_o}{v_s} = \frac{\left(1 + \frac{R_f}{R_e}\right)}{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{Z_T} + \frac{R_f}{Z_T}} \quad (5.3.32)$$

If we express Z_T by its components:

$$Z_T = \frac{1}{\frac{1}{R_T} + sC_T} \quad (5.3.33)$$

we obtain:

$$\frac{v_o}{v_s} = \frac{\left(1 + \frac{R_f}{R_e}\right)}{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T} + sC_T \left[R_f + R_o \left(1 + \frac{R_f}{R_e}\right) \right]} \quad (5.3.34)$$

which we reorder in the usual way to separate the DC gain from the frequency dependent part:

$$\frac{v_o}{v_s} = \frac{1 + \frac{R_f}{R_e}}{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T}} \cdot \frac{\frac{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T}}{C_T \left[R_f + R_o \left(1 + \frac{R_f}{R_e}\right) \right]}}{s + \frac{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T}}{C_T \left[R_f + R_o \left(1 + \frac{R_f}{R_e}\right) \right]}} \quad (5.3.35)$$

Again, a comparison with the general normalized first-order transfer function:

$$F(s) = A_0 \frac{-s_1}{s - s_1} \quad (5.3.36)$$

reveals the DC gain:

$$A_0 = \frac{1 + \frac{R_f}{R_e}}{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T}} \quad (5.3.37)$$

and the pole:

$$s_1 = - \frac{1 + \left(1 + \frac{R_f}{R_e}\right) \frac{R_o}{R_T} + \frac{R_f}{R_T}}{C_T \left[R_f + R_o \left(1 + \frac{R_f}{R_e}\right) \right]} \quad (5.3.38)$$

The bandwidth is calculated from the pole s_1 :

$$f_h = \frac{1}{2\pi} |s_1| \quad (5.3.39)$$

The DC closed loop gain A_0 contains the desired gain A_{cl} :

$$A_{cl} = 1 + \frac{R_f}{R_e} \quad (5.3.40)$$

and an error term ε :

$$\varepsilon = A_{cl} \frac{R_o}{R_T} + \frac{R_f}{R_T} \quad (5.3.41)$$

which is small since R_T is usually 100 k Ω or higher, whilst R_o is between 5 and 50 Ω and R_f is between 100 Ω and 1 k Ω ; thus, even if A_{cl} is 100 or more, ε rarely exceeds 10^{-3} . The transfer function can therefore be expressed as:

$$\frac{v_o}{v_s} = \frac{A_{cl}}{1 + \varepsilon} \cdot \frac{\frac{1 + \varepsilon}{C_T(R_f + R_o A_{cl})}}{s + \frac{1 + \varepsilon}{C_T(R_f + R_o A_{cl})}} \quad (5.3.42)$$

Whilst the gain error ε is small, the bandwidth error can be rather high at high gain; e.g., if $R_f = 1\text{ k}\Omega$ and $R_o = 10\text{ }\Omega$, with $A_{cl} = 100$, the time constant would double and the bandwidth would be halved. In Fig. 5.3.12 we have plotted the bandwidth reduction as a function of gain for a typical current feedback amplifier. Although not constant as in theory, the bandwidth is reduced far less than in voltage feedback opamps (about $50\times$ less for a gain of 100).

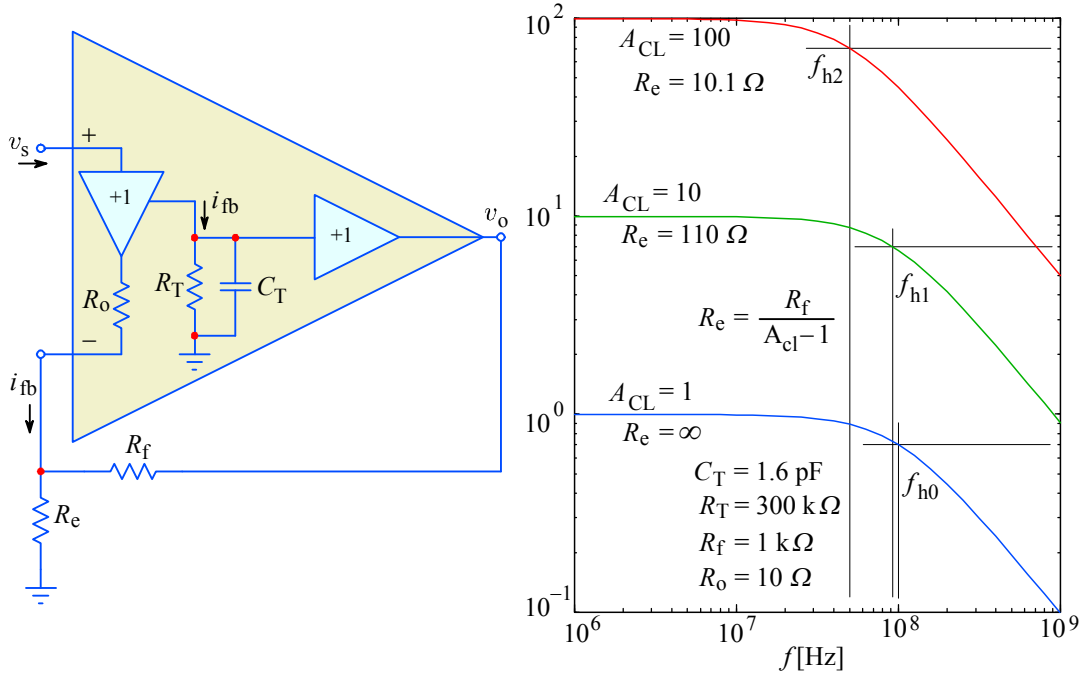


Fig. 5.3.12: The bandwidth of an actual current feedback amplifier is gain dependent, owing to a small but finite inverting input resistance R_o . The nominal bandwidth of 100 MHz at unity gain is reduced to 90 MHz at the gain of 10 and to only 50 MHz at the gain of 100 if R_o is $10\text{ }\Omega$. Nevertheless, this is still much better than in voltage feedback amplifiers.

From these relations we conclude two things: first, both the actual closed loop gain and bandwidth are affected by the desired closed loop gain $A_{cl} = 1 + R_f/R_e$; second and more important, for a given R_o , we can reduce R_f by $R_o A_{cl}$ and recalculate the required R_e to arrive at slightly modified values which **preserve both the desired gain and bandwidth!**

Note that in the above analysis we have assumed a purely resistive feedback path; additional influence of R_o will show up when we shall consider the effect of stray capacitances in the following section.

5.3.5 Noise Gain and Amplifier Stability Analysis

A classical voltage feedback unity gain compensated amplifier (for which any secondary pole lies above the open loop unity gain crossover) usually remains stable if a capacitor C_f is added in parallel to R_f , as in Fig. 5.3.13a. Because C_f lowers the bandwidth, it is often used to prevent problems at and above the closed loop cutoff. In contrast, a capacitor C_e in parallel with R_e , as shown in Fig. 5.3.13b, would reduce the feedback at high frequencies, leading to instability.

A different situation is encountered with current feedback amplifiers, which become unstable in the presence of any capacitance in parallel with either resistor of the feedback loop. Therefore the behavior of the circuit in [Fig. 5.3.13a](#) in the case of a current feedback amplifier is at odds with what we were used to.

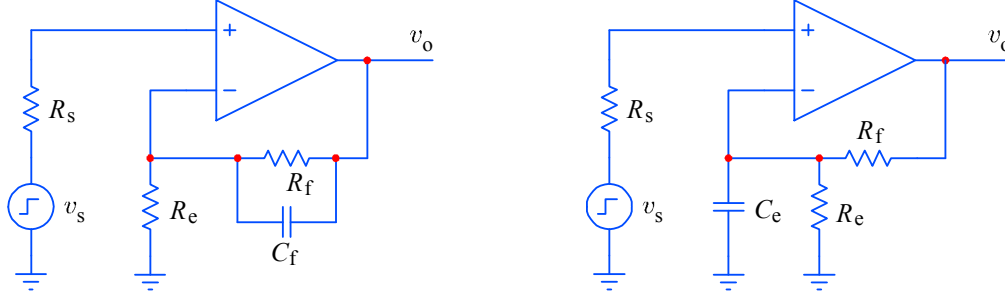


Fig. 5.3.13: a) A unity gain compensated voltage feedback amplifier remains stable with capacitive feedback, whilst in b) it is unstable. In contrast, the stability of a current feedback amplifier is upset by either a) C_f in parallel with R_f , or b) C_e in parallel with R_e .

Before we attempt to explain the unusual sensitivity of current feedback amplifiers to capacitively affected feedback, let us introduce an extremely useful concept, called **noise gain**. In contrast with the name, the noise gain is not used just to evaluate the circuit noise, but the circuit stability as well. It can be applied to all kinds of amplifiers, not just current feedback ones.

The noise, generated by the amplifier input stage, undergoes the full open loop amplification, so the input stage noise dominates over the noise of other stages. Therefore a noisy amplifier can be modeled as a noise generator in series with the input of a noiseless amplifier, as in [Fig. 5.3.14](#), regardless of the actual signal amplification topology, be it inverting or non-inverting.

Any signal within the amplifier feedback loop is processed in the same way as the input stage noise. Thus by grounding the signal input and by analyzing the noise gain within the amplifier and its attenuation in the feedback network, we shall be able to predict the amplifier behavior.

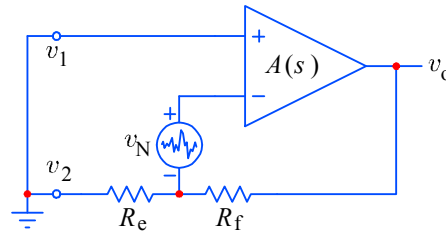


Fig. 5.3.14: Noise gain definition: A noisy amplifier is modeled as a noise generator v_N in series with the input of a noiseless amplifier. The inverting signal gain is $v_o/v_2 = -R_f/R_e$; the non-inverting signal gain is $v_o/v_1 = 1 + R_f/R_e$; the noise gain is $v_o/v_N = -(1 + R_f/R_e)$. The noise generator polarity is indicated only as a reference for the noise gain polarity inversion.

The noise gain is inverting in phase, but equal in value to the non-inverting signal gain:

$$A_N = \frac{v_o}{v_N} = - \left(1 + \frac{R_f}{R_e} \right) \quad (5.3.43)$$

For the voltage feedback amplifier the closed loop bandwidth is equal to the unity gain bandwidth frequency and the noise gain:

$$f_{cl} = \frac{f_1}{|A_N|} \quad (5.3.44)$$

If the feedback network is purely resistive the noise gain is independent of frequency; reactive components (usually capacitances) within the feedback loop will cause the noise gain to change with frequency.

To see this, we usually draw the asymptotes of the transfer function magnitude (absolute value) in a log-log *Bode* plot, with the breakpoints representing the poles and zeros, each pole adding a slope of $-20 \text{ dB}/10f$ and each zero a $+20 \text{ dB}/10f$. We then approximate the phase angle at each breakpoint and in the middle of the linear section (this is a simple and straightforward process if the breakpoints are far apart — see an arbitrary example in [Fig. 5.3.15](#) illustrating some most common possibilities). From this we can evaluate the feedback loop phase margin and, consequently, the amplifier stability.

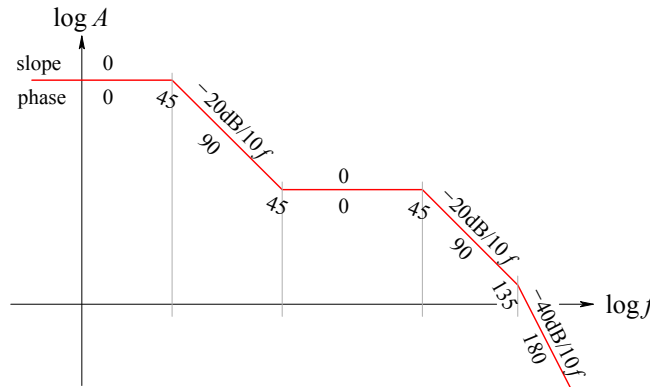


Fig. 5.3.15: An arbitrary example of the phase angle estimated from the gain magnitude, its slopes and various breakpoints. If two breakpoints are relatively close the phase would not actually reach the value predicted from the slope value, but an intermediate value instead.

In the same manner, along the amplifier open loop gain asymptotes, we draw the noise gain, as in [Fig. 5.3.16](#), and we look at the crossover point of these two characteristics. Two important parameters can be derived from this plot: the first is the amount of gain at the crossover frequency f_c ; the second is the relative slope difference between the two lines at f_c , which also serves as an indication of their phase difference.

If the available gain at f_c is greater than unity the phase difference determines the amplifier stability. The feedback can be considered ‘negative’ and the amplifier operation stable if the loop phase margin is at least 45° ; this means that, if a 360° phase shift is ‘positive’, the maximum phase shift within the feedback loop must always be less than 315° (if $A(f_c) > 1$). Since the inverting input provides 180° , the total phase shift of the remaining amplifier stages (secondary poles) and the feedback network should never exceed 135° . Note also that a phase margin of 90° or more results in a smooth transient response; for a phase margin between 90° and 45° an increasing amount of peaking would result.

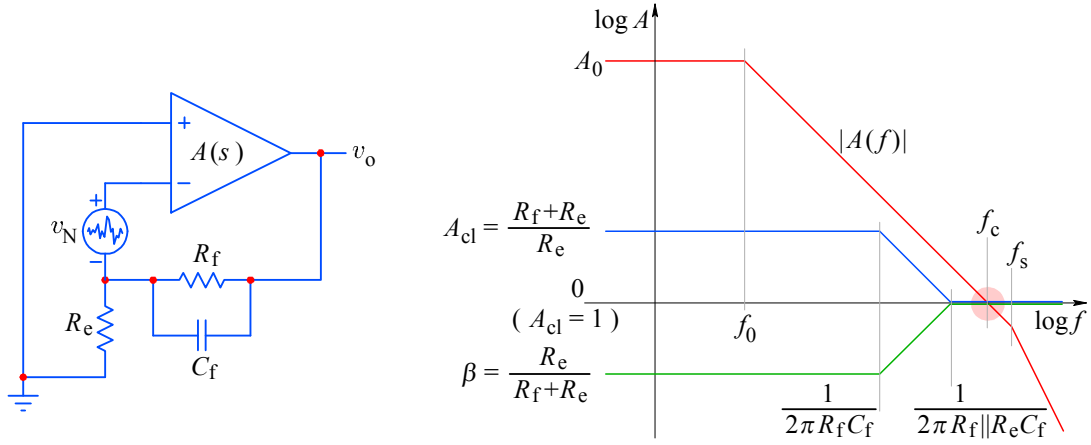


Fig. 5.3.16: Voltage feedback amplifier noise gain is derived from the equivalent circuit noise model (a noise generator in series with the input of a noiseless amplifier). The Bode plot shows the relationships between the most important parameters. Here $|A(f)|$ is the open loop gain magnitude, f_0 is the dominant pole, and f_s is the secondary pole, owed to the slowest internal stage. The inverse of the feedback attenuation β is the noise gain and it is equivalent to the amplifier closed loop gain A_{cl} . Note that the noise gain is flat up and beyond the open loop crossover f_c , owed to the zero at $1/(2\pi R_f C_f)$. The amplifier is stable since the noise gain crosses the open loop gain at a point where their slope difference is 20 dB/10f. If the amplifier open loop gain was higher, the gain at the secondary pole (at f_s) would be higher than unity and the slope difference would be 40 dB/10f. Then, the increased phase (135° at f_s and approaching 180° above), along with the 180° of the amplifier inverting input, would make the feedback positive ($\rightarrow 360^\circ$) and the amplifier would oscillate.

Now, let us find the noise gain of the voltage feedback amplifier in [Fig. 5.3.16](#).

Note that while there is some feedback available the amplifier tries to keep the difference between the inverting and non-inverting input as small as its open loop gain allows; so, with a high open loop gain, the input voltage difference tends to be zero (plus the DC voltage offset).

Note also that, in order to keep track of the phase inversion by the amplifier, we have added polarity indicators to the noise generator. If the '+' side of the noise generator v_N tries to push the inverting input positive, the output voltage must go negative to compensate it.

With C_f in parallel with R_f we shall have:

$$\frac{v_o}{v_N} = - \left(1 + \frac{R_f}{R_e} \cdot \frac{\frac{1}{C_f R_f}}{s + \frac{1}{C_f R_f}} \right) \quad (5.3.45)$$

which we can also rewrite as:

$$\frac{v_o}{v_N} = - \left(\frac{\frac{1}{C_f R_f}}{s + \frac{1}{C_f R_f}} + \frac{s + \frac{1}{C_f R_e}}{s + \frac{1}{C_f R_f}} \right) \quad (5.3.46)$$

Here we have a pole at $1/C_f R_f$ and a zero at $1/C_f R_e$. [Eq. 5.3.46](#) is the noise gain (and also the closed loop gain A_{cl} ; see the two distinct breakpoint frequencies in [Fig. 5.3.16](#)). The inverse of this is the feedback attenuation β .

With the open loop gain as shown in [Fig. 5.3.16](#) the amplifier is stable, since the noise gain crosses over the open loop gain at f_c , where the open loop and closed loop slope difference is 20 dB/decade, and the associated phase shift is (nearly) 90° . In addition to the 180° of the amplifier inverting input, the total phase angle is then 270° . The minimum phase margin for a stable amplifier would be 45° and here we have 90° ($360 - 270$), so the feedback can still be considered "negative".

However, if the open loop gain was higher (and if the poles remain at the same frequencies) the gain at f_s (the frequency of the secondary pole) can be greater than unity. In this case at the crossover of the noise gain and open loop gain the slope difference will be 40 dB/10f, with the associated phase of 135° at f_s and approaching 180° above. The feedback will become 'positive' and, with the gain at f_s greater than unity, the amplifier would oscillate.

In the case of the current feedback amplifier in [Fig. 3.5.17](#) we first note that instead of gain our Bode plot shows the feedback **impedances** and the amplifier **transimpedance**, all as functions of frequency.

Intuitively speaking, a capacitance C_f in parallel to R_f would reduce the impedance of the feedback network at high frequencies, thus also reducing the closed loop gain. However, intuition is misleading us: since the current feedback system bandwidth is inversely proportional to the feedback impedance in the f-branch (as demonstrated by [Eq. 5.3.24](#)), the addition of C_f increases the bandwidth. By itself this would be welcome, but note that at the crossover frequency f_c the slope difference between the transimpedance Z_T and the 'noise transimpedance' (in analogy with noise gain) is 40 dB/10f, causing a phase shift of 180° . This means that at f_c the feedback current becomes positive and the amplifier will oscillate.

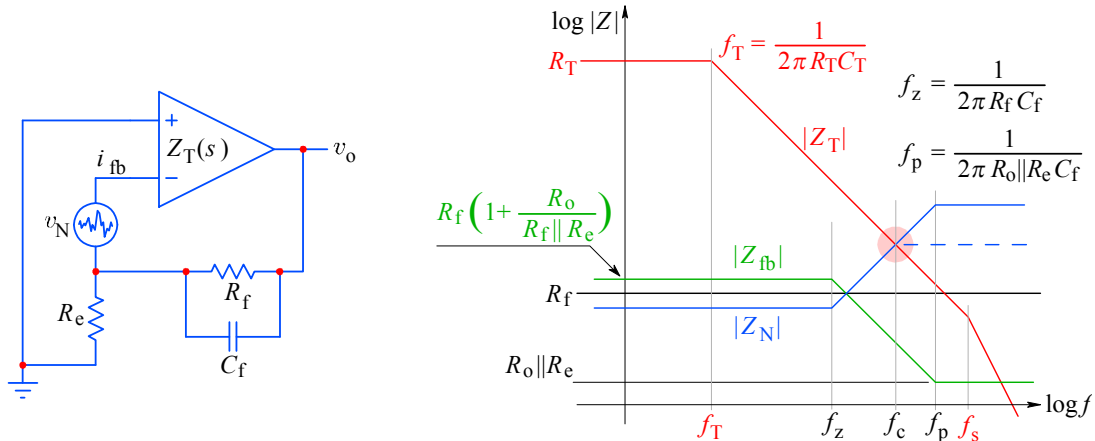


Fig. 5.3.17: For the current feedback amplifier we draw the impedances, not gain. The feedback network impedance $|Z_{fb}|$, as seen from v_o , is slightly higher than R_f at DC (owing to R_o , the inverting input resistance) and falls to $R_o || R_e$ at high frequencies; its inverse (about R_f) is the amplifier noise transimpedance, $|Z_N|$. The feedback network pole becomes the zero of the noise transimpedance: $s_z = -1/C_f R_f$ ($f_z = |s_z|/2\pi$); likewise, the feedback zero becomes the noise transimpedance pole $s_p = -1/C_f (R_o || R_e)$, ($f_p = |s_p|/2\pi$). At f_c the crossover with $|Z_T|$, the slope difference is 40 dB/10f and the relative phase angle is 180° ; the amplifier will inevitably oscillate, even if the secondary pole is far away and its Z_T breakpoint is well below R_f . The dashed line is the transimpedance required for stability, realized by an R in series with C_f .

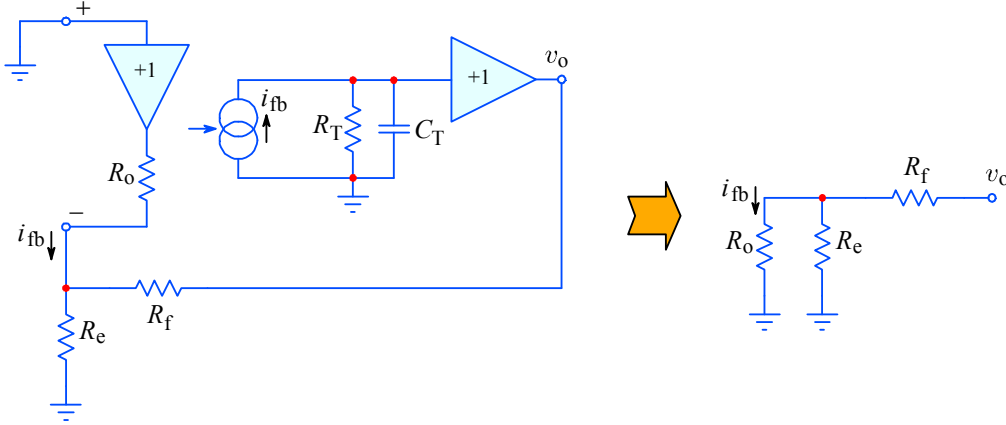


Fig. 5.3.18: The current feedback amplifier and its 'noise transimpedance' equivalent, v_o/i_{fb} .

We can find the noise transimpedance as simply as we found the noise gain for voltage feedback amplifiers. By assuming that the feedback current is noise generated, from the equivalent circuit in [Fig. 5.3.18](#) we calculate the ratio of the output voltage and the feedback current:

$$\frac{v_o}{i_{fb}} = R_o + R_f \left(1 + \frac{R_o}{R_e} \right) \quad (5.3.47)$$

By adding a capacitance C_f in parallel with R_f the noise transimpedance becomes:

$$\frac{v_o}{i_{fb}} = R_o + R_f \frac{\frac{1}{C_f R_f}}{s + \frac{1}{C_f R_f}} \left(1 + \frac{R_o}{R_e} \right) \quad (5.3.48)$$

and this equation is represented by $|Z_N|$ in [Fig. 5.3.17](#).

In most practical cases there will be stray capacitances in parallel to both R_f and R_e , and in addition between both inputs, as well as from the non-inverting input to ground, and also from output to ground. A real world situation can be rather complicated.

As we have shown, current feedback amplifiers are extremely sensitive to any capacitances within the negative feedback loop. This means that whole families of circuits (such as integrators, differentiators, some filter topologies, current amplifiers, I to V converters, logarithmic amplifiers, etc.) can not be realized in the same way as with conventional amplifiers. Fortunately, there are alternative ways of performing the same functions and some of the most common ones are shown in [Fig. 5.3.19](#) for a quick comparison:

- an inverting integrator can be implemented using two current feedback (CFB) opamps, with the bonus of providing both the inverting and non-inverting configuration within the same circuit;
- a single pole inverting filter amplifier can be implemented by exploiting the internal capacitance C_T and the external feedback resistor R_f (useful for high frequency cut off; for lower frequencies a high value of R_f is impractical since it would cause a large voltage offset, owing to a large input bias current);

- for filters the Sallen–Key non-inverting configuration is recommended for use with CFB amplifiers. This configuration can be easily cascaded (using second- and third-order sections) to realize multi-pole high order filters, in the same way as the ‘multiple feedback’ inverting configuration can be cascaded;
- current sources, such as some digital to analog converters, photo-diodes, etc., have a relatively large capacitance in parallel. If a CFB amplifier is used as an inverting current to voltage converter, the source capacitance must be compensated. The gain setting resistors R_f and R_e are in series with the compensation capacitor, thus preventing instability at high frequencies. Also, these resistors can be used to scale the compensation capacitor value to more practical values of 10–20 pF, instead of 1 pF or less which would normally be needed for high speed response.

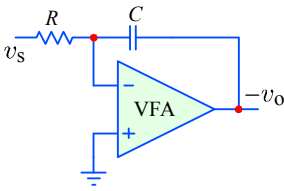
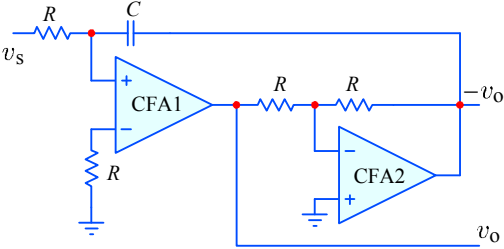
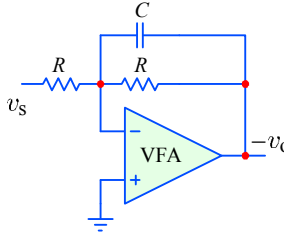
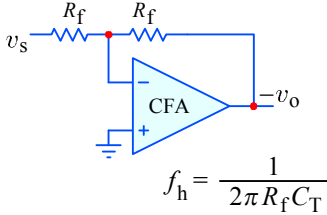
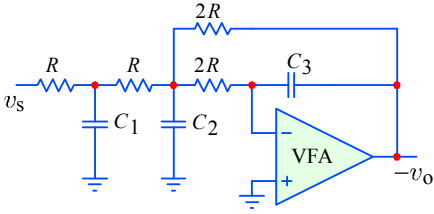
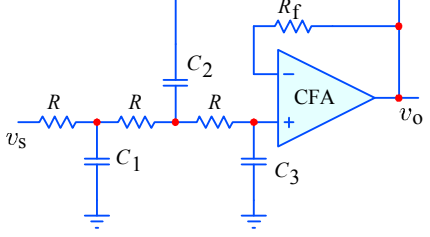
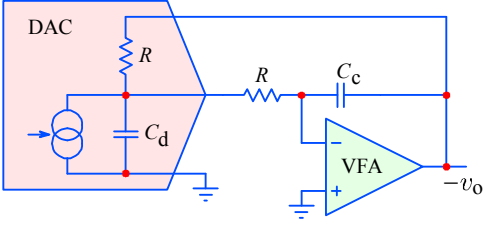
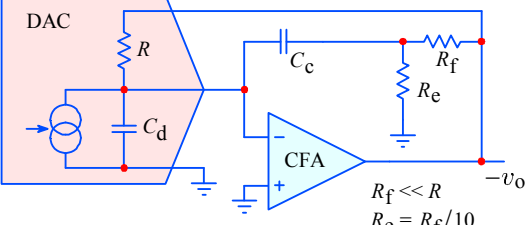
Voltage Feedback Amplifiers	Current Feedback Amplifiers
 <p>Integrator (inverting)</p>	 <p>Integrator (inverting and non-inverting)</p>
 <p>Single-Pole Amplifier (inverting)</p>	 $f_h = \frac{1}{2\pi R_f C_T}$ <p>Single-Pole Amplifier (inverting)</p>
 <p>Multiple Feedback 3-pole Filter (inverting)</p>	 <p>Sallen–Key 3-pole Filter (non-inverting)</p>
 <p>I to V Conversion DAC C_d Compensation</p>	 <p>I to V Conversion DAC C_d Compensation</p> $R_f \ll R$ $R_e = R_f/10$

Fig. 5.3.19: Functionally equivalent circuits with conventional and current feedback amplifiers. Integrators, filters and current to voltage converters in inverting configurations cannot be achieved using a single CF amplifier. However, two-amplifier circuits can provide inherent amplifier pole compensation, which is very important at high frequencies. Filters can be realized in the non-inverting configuration. And feedback capacitance can be isolated by a resistive divider.

5.3.6 Feedback Controlled Gain Peaking

As we have just seen, the current feedback amplifier is sensitive to capacitive loading of its inverting input. But stray capacitances are unavoidable and they can cause significant gain peaking and even oscillations in high speed low gain designs. Fortunately, the current feedback topology offers a simple way of controlling this by choosing such values of R_f and R_e which would set the bandwidth and the gain to the optimum. Although CFB amplifier performance is usually optimized for a specific value of the bandwidth defining resistance R_f , ample trade-off range is often possible.

However, as we have learned in [Part 4](#), in multi-stage amplifiers it is necessary to optimize the system as a whole and not just each stage individually. The possibility of controlling the gain peaking with feedback resistors lends itself nicely to our purpose. In practice we would have to iteratively adjust both R_f and R_e to obtain the desired gain, bandwidth, and peaking. What we would like is to be able to adjust the bandwidth and peaking by a single resistor, without affecting the gain. The circuit in [Fig. 5.3.20](#) does just that.

The feedback resistors R_f and R_e should be chosen for the gain required, but with the lowest possible values, which would not overload the output stage. Then the resistor between the feedback divider and the inverting input, R_b , should be adjusted for the required response, assuming a fixed value of the stray capacitance.

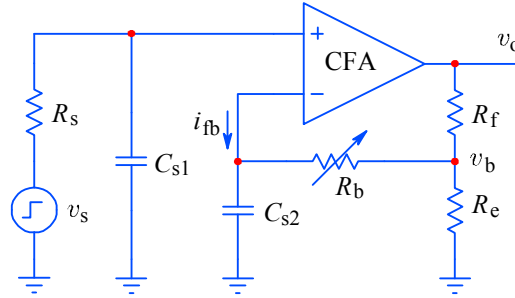


Fig. 5.3.20: This circuit exploits the ability of current feedback amplifiers to adjust the bandwidth and gain peaking independently of the gain. The price to pay is the lower slew rate limit. See the frequency and the step response below.

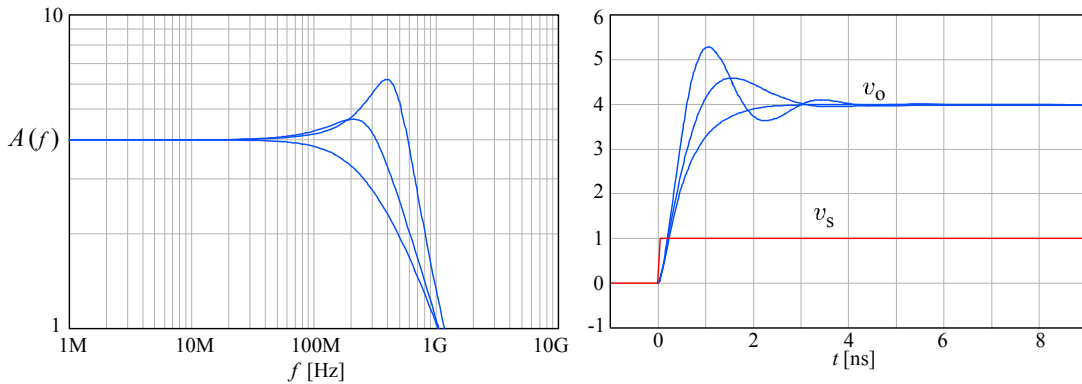


Fig. 5.3.21: **a)** Frequency response; **b)** Step response of the amplifier in [Fig. 5.3.20](#). The closed loop gain $A_{cl} = 1 + R_f/R_e = 4$, $R_f = 150 \Omega$, $R_e = 50 \Omega$, the source resistance $R_s = 50 \Omega$, the stray capacitances, $C_{s1,2} = 1 \text{ pF}$, the amplifier transcapacitance $C_T = 1 \text{ pF}$, while R_b is varied from 150Ω for highest peaking to 750Ω for lowest peaking.

Note, however, that in this way we lose the current on demand property of the CFB amplifier, since R_b will reduce the slew rate.

In a similar manner as was done for passive circuits in [Part 2](#) and in [Sec. 5.1](#), the resulting gain peaking can be used to improve the step response of a multi-stage system. As shown in [Fig. 5.3.21](#), the gain peaking reveals the amplifier resonance, which decreases with increasing R_b , while the DC gain remains almost unchanged.

5.3.7 Improved Voltage Feedback Amplifiers

The lessons learned from the current feedback technology can be used to improve conventional voltage feedback amplifiers.

Besides the improved semiconductor manufacturing technology, basically there are two approaches: one is to take the voltage feedback amplifier and modify it using the techniques of current feedback to avoid its weak points. One such example is shown in [Fig. 5.3.22](#). The other way, like the circuit in [Fig. 5.3.23](#), is to take the current feedback amplifier and modify it to make it appear to the outside world as a voltage feedback amplifier.

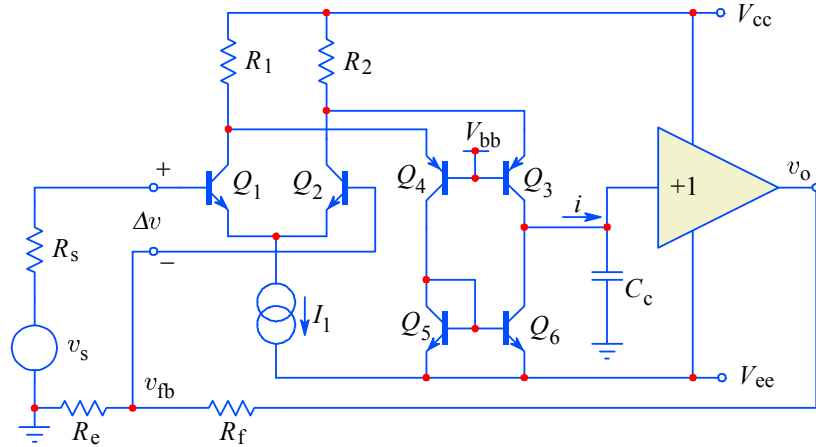


Fig. 5.3.22: The voltage feedback amplifier, improved. The transistors Q_{1-4} form a differential ‘folded’ cascode, which drives the current mirror $Q_{5,6}$. In this way the input is a conventional high impedance differential, but the dominant pole compensation capacitor C_c is grounded, eliminating the Miller effect. This circuit still exhibits slew rate limiting, although at much higher frequencies. Typical examples of this configuration are Analog Devices’ AD-817 and Burr–Brown’s OPA-640.

The differential folded cascode Q_{1-4} and the current mirror $Q_{5,6}$ of [Fig. 5.3.22](#) can be modeled by a transconductance, g_m , driven by the input voltage difference, $\Delta v = v_s - v_{fb}$. Here v_s is the signal source voltage and v_{fb} is the feedback voltage, derived from the output voltage v_o and the feedback network divider, $R_e/(R_f + R_e)$. The current $i = \Delta v g_m$ drives the output buffer and the capacitance C_c :

$$v_o = i \frac{1}{s C_c} = g_m (v_s - v_{fb}) \frac{1}{s C_c} = g_m \left(v_s - v_o \frac{R_e}{R_f + R_e} \right) \frac{1}{s C_c} \quad (5.3.49)$$

From this we obtain:

$$v_o \left(1 + \frac{R_e}{R_f + R_e} \cdot \frac{g_m}{s C_c} \right) = v_s \frac{g_m}{s C_c} \quad (5.3.50)$$

and, finally, the transfer function:

$$\frac{v_o}{v_s} = \left(1 + \frac{R_f}{R_e}\right) \frac{\frac{R_e}{R_f + R_e} \cdot \frac{g_m}{C_c}}{s + \frac{R_e}{R_f + R_e} \cdot \frac{g_m}{C_c}} \quad (5.3.51)$$

If we compare this with a general first-order amplifier transfer function:

$$\frac{v_o}{v_s} = A_{cl} \frac{-s_1}{s - s_1} \quad (5.3.52)$$

we see that the closed loop gain is, as usual, $A_{cl} = 1 + R_f/R_e$, whilst the amplifier closed loop pole is $s_1 = -R_e g_m / C_c (R_f + R_e)$, and therefore the cut off frequency is an inverse function of the closed loop gain, just as in voltage feedback amplifiers.

A similar situation is encountered in [Fig. 5.3.23](#), where the current i_b charging C_T is set by the input voltage difference and R_b : $i_b = \Delta v / R_b$.

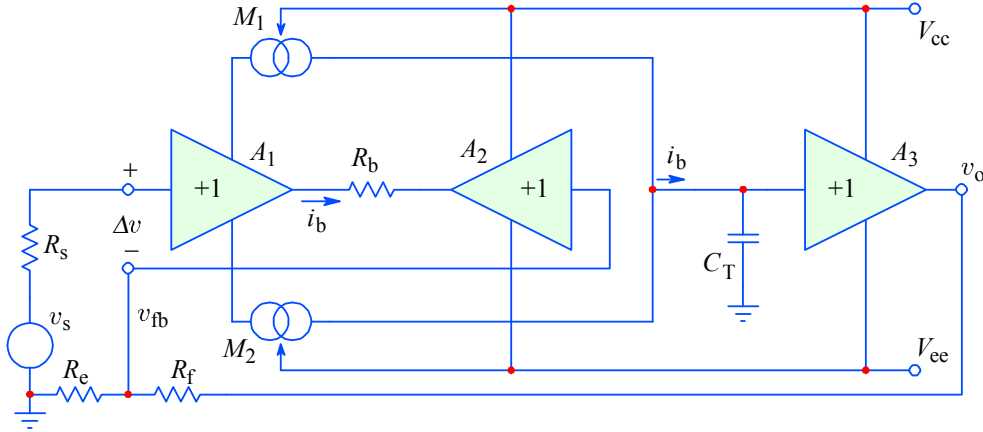


Fig. 5.3.23: The basic current feedback amplifier (A_1, A_3, M_1, M_2) is improved by adding another buffer, A_2 , which presents a high impedance to the feedback divider, R_f and R_e ; an additional resistor, R_b , now takes the role of converting the voltage feedback into current and provide bandwidth setting. Like the original current feedback amplifier, this circuit is also (almost) free from slew rate limiting. However, the closed loop bandwidth is, as in voltage feedback amplifiers, gain dependent. A typical representative of this configuration is Analog Devices' OP-467.

The output voltage is:

$$v_o = i_b \frac{1}{s C_T} = \Delta v \frac{1}{s C_T R_b} = \left(v_s - v_o \frac{R_e}{R_f + R_e} \right) \frac{1}{s C_T R_b} \quad (5.3.53)$$

so the transfer function is:

$$\frac{v_o}{v_s} = \left(1 + \frac{R_f}{R_e}\right) \frac{\frac{R_e}{R_f + R_e} \cdot \frac{1}{C_T R_b}}{s + \frac{R_e}{R_f + R_e} \cdot \frac{1}{C_T R_b}} \quad (5.3.54)$$

The closed loop gain is the same as in the previous case, whilst the pole is $s_1 = -R_e / C_T R_b (R_f + R_e)$, so the closed loop cutoff frequency is again an inverse function of the closed loop gain.

One of the important parameters in integrated circuit design is the available bandwidth vs. quiescent current. The average technology achievement around the year 1990 was about 10 MHz/mA and around the year 2000 it was already about 100 MHz/mA; the figure is steadily rising. With the ever increasing number of transistors on a silicon chip it is important to keep this value high. The implementation of structures which convey the supply current efficiently to the signal helps to reduce the waste of power.

An example, named the ‘Quad Core’ [Ref. 5.32], is shown in Fig. 5.3.24. This is an interesting combination of circuits in Fig. 5.3.22 and 5.3.23, where the two input buffers, formed by $Q_{1,2}$ and $Q_{3,4}$, combine their currents by the current mirrors, Q_{5-8} , driving the following gain stage $Q_{9,10}$ in a differential push-pull mode.

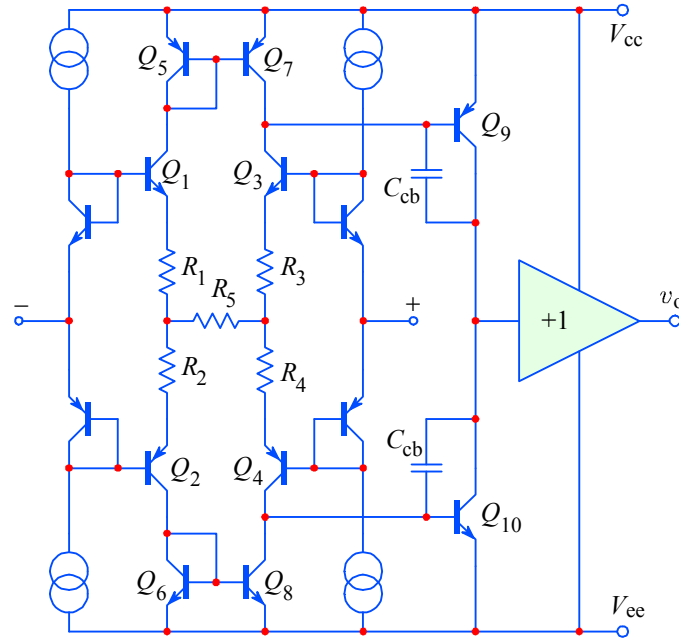


Fig. 5.3.24: An interesting combination of circuits in Fig. 5.3.22 and 5.3.23 is the so called ‘Quad Core’ structure, [Ref. 5.32]. Here both the inverting and the non-inverting input buffer currents are combined by the current mirrors to drive the C_{cb} of $Q_{9,10}$. The non-labeled transistors provide the V_{be} compensation for Q_{1-4} . Typical representatives are Analog Devices’ AD-8047, AD-9631, AD-8041 and others.

The current available to charge the C_{cb} capacitances of $Q_{9,10}$ is set by the input voltage difference and R_5 . This current is effectively doubled by the input structure, thus increasing the bandwidth, the loop gain, and linearity. A further bandwidth improvement is achieved by the low impedance of $Q_{7,8}$, which are practically fully open and so provide a tight control of the $Q_{9,10}$ base voltages, reducing the Miller effect considerably. The circuit behaves as a voltage feedback amplifier with the advantages of low offset and high loop gain and with a bandwidth and slew rate limiting close to that of current feedback amplifiers.

The output buffer stage can also be improved for greater current handling efficiency. An example is shown in Fig. 5.3.25.

Here the collectors of $Q_{2,3}$ and $Q_{1,4}$ are summed and mirrored by $Q_{5,7}$ and $Q_{6,8}$, respectively, and finally added to the output load current. With appropriate bias this scheme allows a reduction of the quiescent power supply current to just one third

of the conventional buffer, whilst not compromising the full power bandwidth. At the same time, the circuit has a comparable loading capability and offers better linearity.

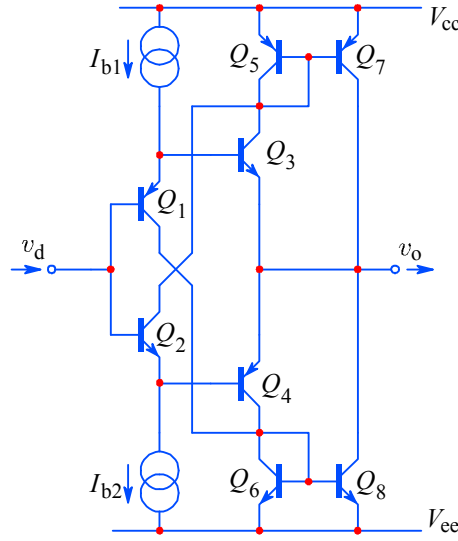


Fig. 5.3.25: Output buffer stage with improved current handling.

5.3.8 Compensating Capacitive Loads

Another very important property of high speed opamps is their ability to drive capacitive loads. To the amplifier in [Fig. 5.3.26](#), because of its non-zero output resistance R_o , the capacitive load C_L adds a high frequency pole within the feedback loop. The feedback becomes frequency dependent and the phase margin is lowered, thus compromising the stability.

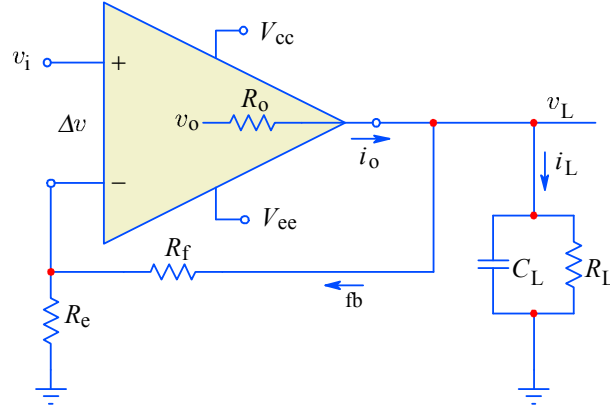


Fig. 5.3.26: Owing to the non-zero output impedance a capacitive load adds another pole within the feedback loop. If the closed loop gain is too low the resulting increase in phase can make the feedback positive at high frequencies, instead of negative, destabilizing the amplifier.

The load voltage v_L can be expressed as a function of the internal voltage v_o (seen when no load is present), the factor D_R , and the frequency dependence:

$$v_L = v_o D_R \frac{-s_L}{s - s_L} \quad (5.3.55)$$

Here D_R is the resistive divider formed by the output resistance R_o , the load resistance R_L and the total resistance of the feedback divider $R_f + R_e$:

$$D_R = \frac{\frac{R_L(R_f + R_e)}{R_L + R_f + R_e}}{R_o + \frac{R_L(R_f + R_e)}{R_L + R_f + R_e}} \quad (5.3.56)$$

The pole s_L is formed by the load capacitance C_L and the equivalent resistance seen by it, R_q , whilst ω_L is the appropriate cut off frequency:

$$s_L = -\frac{1}{R_q C_L} ; \quad \omega_L = |s_L| \quad (5.3.57)$$

R_q is simply the parallel combination of all the resistances at the output node:

$$R_q = \frac{1}{\frac{1}{R_o} + \frac{1}{R_L} + \frac{1}{R_f + R_e}} \quad (5.3.58)$$

The internal output voltage, v_o , is a function of the input voltage difference, Δv , and the amplifier open loop gain A , which, in turn, is also a function of frequency, $A(s)$:

$$v_o = A(s) \Delta v \quad (5.3.59)$$

The input voltage difference is, of course, the difference between the signal source voltage and the output (load) voltage, attenuated by the feedback resistors:

$$\Delta v = v_s - v_L \frac{R_e}{R_f + R_e} \quad (5.3.60)$$

The open loop gain $A(s)$ is defined by the DC open loop gain A_0 and the frequency dependent term owed to the amplifier dominant pole at the frequency ω_0 :

$$A(s) = A_0 \frac{\omega_0}{s + \omega_0} \quad (5.3.61)$$

With this in mind, we can express the internal output voltage:

$$v_o = A_0 \frac{\omega_0}{s + \omega_0} \left(v_s - v_L \frac{R_e}{R_f + R_e} \right) \quad (5.3.62)$$

and by inserting this back in [Eq. 5.3.53](#) we have the load voltage:

$$v_L = A_0 \frac{\omega_0}{s + \omega_0} \left(v_s - v_L \frac{R_e}{R_f + R_e} \right) D_R \frac{\omega_L}{s + \omega_L} \quad (5.3.63)$$

which we solve for v_L explicitly:

$$v_L \left(1 + \frac{R_e}{R_f + R_e} A_0 \frac{\omega_0}{s + \omega_0} D_R \frac{\omega_L}{s + \omega_L} \right) = v_s A_0 \frac{\omega_0}{s + \omega_0} D_R \frac{\omega_L}{s + \omega_L} \quad (5.3.64)$$

Now we can write the transfer function:

$$\frac{v_L}{v_s} = \frac{A_0 D_R \frac{\omega_0 \omega_L}{(s + \omega_0)(s + \omega_L)}}{1 + \frac{R_e}{R_f + R_e} A_0 D_R \frac{\omega_0 \omega_L}{(s + \omega_0)(s + \omega_L)}} \quad (5.3.65)$$

where we separate the closed loop gain term:

$$\frac{v_L}{v_s} = \frac{R_f + R_e}{R_e} \cdot \frac{\frac{R_e}{R_f + R_e} A_0 D_R \frac{\omega_0 \omega_L}{(s + \omega_0)(s + \omega_L)}}{1 + \frac{R_e}{R_f + R_e} A_0 D_R \frac{\omega_0 \omega_L}{(s + \omega_0)(s + \omega_L)}} \quad (5.3.66)$$

and, by multiplying the numerator and the denominator by $(s + \omega_0)(s + \omega_L)$, which we expand into a polynomial, we obtain the expression for the transfer function. Clearly it is a second-order function of frequency:

$$\frac{v_L}{v_s} = \frac{R_f + R_e}{R_e} \cdot \frac{\frac{R_e}{R_f + R_e} A_0 D_R \omega_0 \omega_L}{s^2 + s(\omega_0 + \omega_L) + \omega_0 \omega_L \left(1 + \frac{R_e}{R_f + R_e} A_0 D_R\right)} \quad (5.3.67)$$

The product of the poles, $s_1 s_2$, is a function of not just ω_0 and ω_L , but also of the open loop gain A_0 and the closed loop feedback dividers, D_R and $R_e/(R_f + R_e)$ (refer to [Appendix 2.1](#) to find the system poles of a 2nd-order function). Since the output resistance, R_o , is usually much lower than both the load resistance R_L and the feedback resistances $R_f + R_e$, the output divider D_R is usually between 0.9 and 1. The system's stability is therefore dictated by the amount of loop gain when $s \rightarrow \omega_L$. Thus close to ω_L the loop gain will be higher than 1 either if A_0 is very high, or if ω_L is relatively low and $R_f \rightarrow 0$, that is, if the closed loop gain approaches unity!

This is often counter-intuitive, not just to beginners, but sometimes even to experienced engineers. Usually, if we want to enhance the amplifier's stability, we increase the feedback at high frequencies by placing a capacitor in parallel to R_f . However, in the case of a capacitive load the amplifier will be turned into an oscillator by that procedure. In contrast, the stability will improve if we increase the closed loop gain (increase R_f or decrease R_e). This is illustrated in [Fig. 5.3.27](#), where the gain and the phase are plotted for the three values of R_f (∞ , $4R_e$ and 0) and the capacitive load is such, that the loop gain at $f_L \approx 2 \cdot 10^4 f_0$ is about 3.

Conventional opamp compensation schemes (consisting usually of a resistor or a series RC network, connected between both inputs, thus increasing the noise gain, without affecting the signal gain) increase the stability at the expense of either the gain, or the bandwidth, or both! Conventional compensation should be used as the last resort only, when the circuit must meet an unknown load capacitance, which can vary considerably.

In fixed applications, when the capacitive load is known, or is within a narrow range, it is much better to compensate the load by inductive peaking, as we have seen in [Part 2](#). The simplest approach is shown in [Fig. 5.3.28](#), where the load impedance appears real to the amplifier output, so that the feedback is not frequency dependent.

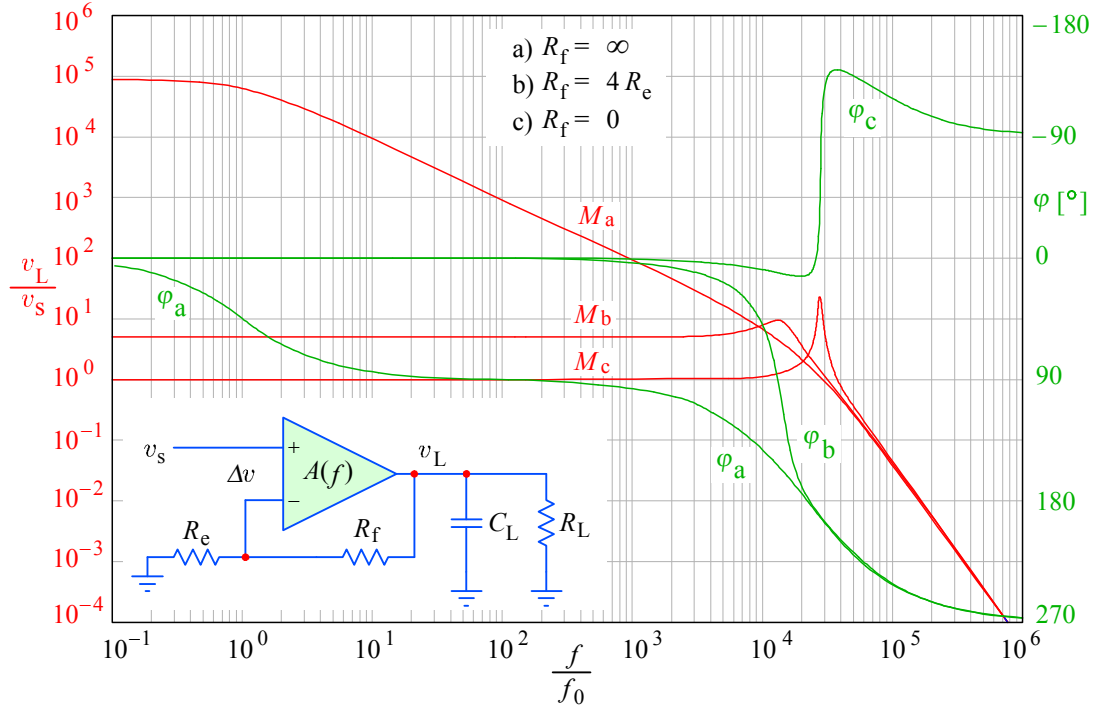


Fig. 5.3.27: An amplifier driving a capacitive load can become unstable if its close loop gain is decreased too much: a) with no feedback and b) with a gain of 5, the amplifier is stable, although in the latter case there is already a pronounced peaking; whilst in c) with the closed loop gain reduced to 1 the peaking is very high and the phase goes over 360° and oscillation is excited at the frequency at which the phase equals 360° .

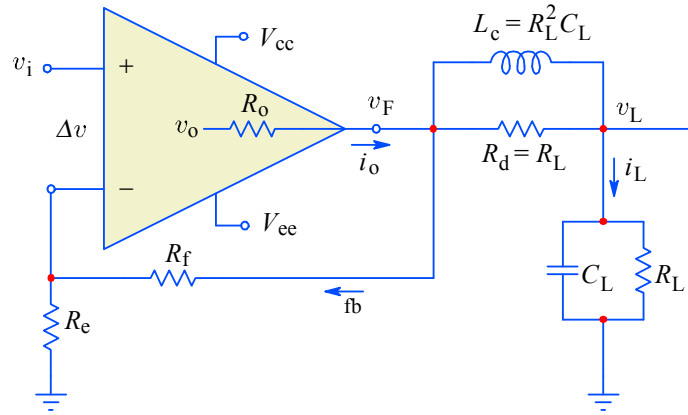


Fig. 5.3.28: Capacitive load compensation which makes the load to appear real and equal to R_L to the opamp. This works for fixed load impedances.

Here the inductance L_c and its parallel damping resistance, R_d , are chosen so that $R_d = R_L$ and $L_c = R_L^2 C_L$, and the amplifier sees an impedance equal to R_L from DC up to the frequency at which the coil stray capacitance starts to influence the compensation. With a careful inductance design the frequency at which this happens can be much higher than the critical amplifier frequency.

However, even with such compensation, the bandwidth can be lower than desired, since the compensation network forms a low pass filter with the load, and the value of the inductance is influenced by both the load resistance R_L and the load capacitance C_L .

The filter transfer function is:

$$\frac{v_L}{v_F} = \frac{\frac{1}{L_c C_L} + s \frac{1}{R_L C_L}}{s^2 + s \frac{2}{R_L C_L} + \frac{1}{L_c C_L}} \quad (5.3.68)$$

The cut off frequency is $\omega_h = 1/\sqrt{L_c C_L} = 1/R_L C_L$ and that is much lower than ω_L of [Eq. 5.3.57](#), which would apply if the amplifier could be made stable by some other means. If R_L and C_L can be separated, it is possible to build a 2-pole series peaking or a T-coil peaking, tuned to form a 3-pole system together with the pole associated with the amplifier closed loop cut off. This procedure is similar to the one described in [Part 2](#) and also in [Sec. 5.1](#), so we leave it as an exercise to the reader.

Another compensation method, shown in [Fig. 5.3.29](#), is to separate the AC and DC feedback path by a small resistance R_c in series with the load and a feedback bridging capacitance C_c :

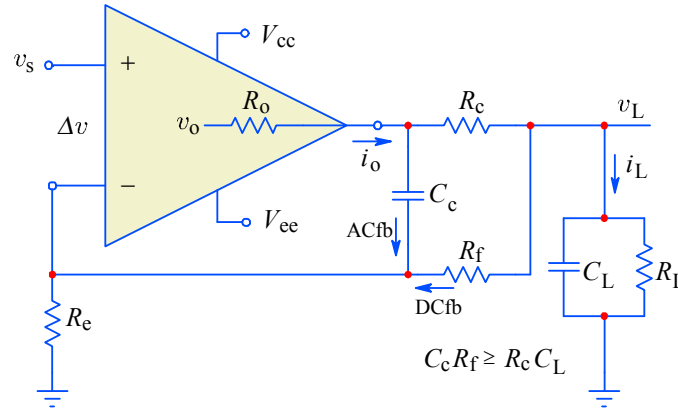


Fig. 5.3.29: The capacitive load is separated from the AC feedback path by a small resistor R_c in series with the output; owing to the capacitance C_c this type of compensation can be applied only to voltage feedback unity gain compensated amplifiers.

This type of compensation can be very effective, since very small values of R_c can be used (5–15 Ω or so), lowering the bandwidth only slightly; however, due to the feedback bridging capacitance C_c , it can be applied only to voltage feedback unity gain compensated amplifiers; it can not be used for current feedback amplifiers.

A more serious problem is the fact that, in some applications, the load capacitance would vary considerably; for example, some types of fast AD converters have their input capacitance code dependent (thus also signal level dependent!). It is therefore desirable to design the amplifier output stage with the lowest possible output resistance and employ a compensation scheme which would work for a range of capacitance values.

[Fig. 5.3.30](#) shows the implementation found in some CFB opamps, where the compensation network, formed by C_c and R_c , is in parallel with the output buffer. With a high impedance load the voltage drop on the output resistance R_o is small and the current through the compensation network is low; but with a capacitive load or other low impedance load the output current causes a high voltage drop on R_o , and consequently the current through C_c increases. Effectively the series combination of

C_c and C_L is added in parallel to C_T , thus lowering the system open loop bandwidth in proportion with the load and keeping the loop stable.

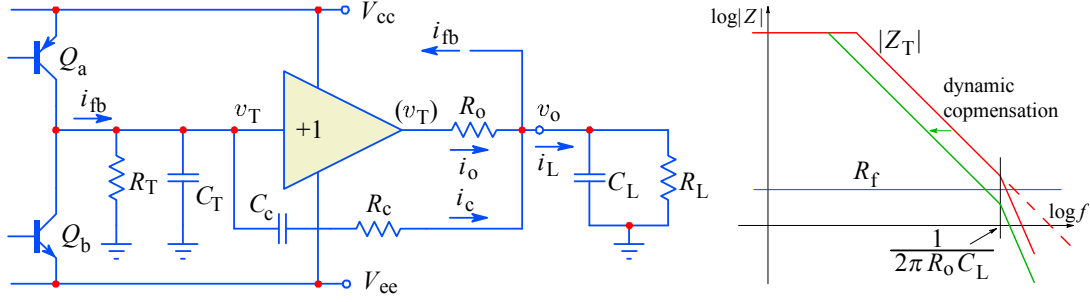


Fig. 5.3.30: The output buffer with a finite output resistance R_o would, when driving a capacitive load C_L , present an additional pole within the feedback loop (taken from v_o), which could condition the amplifier stability. The compensation network, formed by a serial connection of C_c and R_c , draws part of the feedback current to the output node, effectively increasing C_T in proportion to the load, reducing the transimpedance and preserving the closed loop stability.

In [Fig. 5.3.30](#) we have three RC impedances: Z_T is the CFB amplifier transimpedance consisting of R_T and C_T in parallel; Z_c is the compensation impedance consisting of R_c and C_c in series; and Z_L is the load impedance consisting of R_L and C_L in parallel. The output unity gain buffer is assumed to be ideal and the real circuit is modeled by its output resistance R_o .

If v_T is its input voltage, the output voltage v_o will be lower by $i_o R_o$, where i_o is the output current. Since Z_c is connected between the buffer input and the load, the voltage over Z_c is equal to $i_o R_o$, so:

$$v_T - v_o = i_o R_o \quad (5.3.69)$$

$$i_c = \frac{v_T - v_o}{Z_c} = \frac{i_o R_o}{Z_c} \quad (5.3.70)$$

The transimpedance Z_T is driven by the feedback current i_{fb} ; the voltage v_T , which in an ideal case would be equal to $i_{fb} Z_T$, is now lower, because part of the current is stolen by the compensation network Z_c :

$$v_T = (i_{fb} - i_c) Z_T \quad (5.3.71)$$

With a few simple substitutions we obtain:

$$v_o = i_{fb} Z_T - i_o R_o \left(\frac{Z_T}{Z_c} + 1 \right) \quad (5.3.72)$$

This equation shows that the original transimpedance equation ([Eq. 5.3.18](#)) is modified by the output current and the Z_T/Z_c ratio. Thus a capacitive load, which would draw high currents at high frequencies (or at the step transition), will automatically lower the system open loop cut off frequency. Consequently the gain at high frequencies is reduced so that the closed loop crossover remains well above the secondary pole (created by R_o and C_L).

Note that the distortion at high frequencies of the compensated amplifier will be worse than that of a non-compensated one.

5.3.9 Fast Overdrive Recovery

The ability to resume linear signal amplification after a prolonged large overdrive is one of the most important oscilloscope design considerations. The average oscilloscope user is often tempted to zoom in on a small detail of a relatively large signal to inspect or debug the performance of electronic equipment.

With the input sensitivity turned high, depending on the attenuation and gain settings, various stages of a multi-stage amplifier can be overdriven into their non-linear region or even saturated, whilst others can remain within their linear region; at some settings it is the output stage that is overdriven first, at others it will be so with the input stage (this is because the input attenuator is always varied in steps of $\times 10$, the following attenuation and gain steps are usually lower, $\times 2$ or $\times 2.5$). When overdriven, different amplifier stages will have different electrical and thermal histories, so that when the signal falls back within the linear range the circuit will not re-balance immediately, but will take some time before it regains its original accuracy, often with many different time constants, characterized by relatively long rising or falling ‘tails’.

For decades, Tektronix oscilloscopes excelled in fast recovery, far above all its competitors. Although the problem of overdrive recovery has never been easy to solve, in the old days of electronic tubes and discrete transistors the power supply voltage was always very high, allowing ample signal range. With modern ICs, having several transistors one on top of the other and with ever decreasing power supply voltages, the useful linear range is often only a volt or so. Therefore, special local limiting circuitry must be added to smoothly switch in and out with overdrive.

The overdrive problem is even more pronounced in many modern fast ADCs (e.g., the ‘flash type’, especially those with a two-stage pipeline architecture) whose input voltage range is only 1–2 V. When overdriven, they become slower; or they can generate large error codes, even if the overdrive level is just a few least significant bits (LSBs) above maximum. Such ADCs require a well controlled clipping amplifier to drive their input. With a voltage feedback opamp, the solution could be simple, by adding diodes shunting the feedback loop. Precise levels with sharp knees are needed, so a Schottky diode bridge with a biased Zener diode is often used. [Fig. 5.3.31](#) shows one such possible circuit.

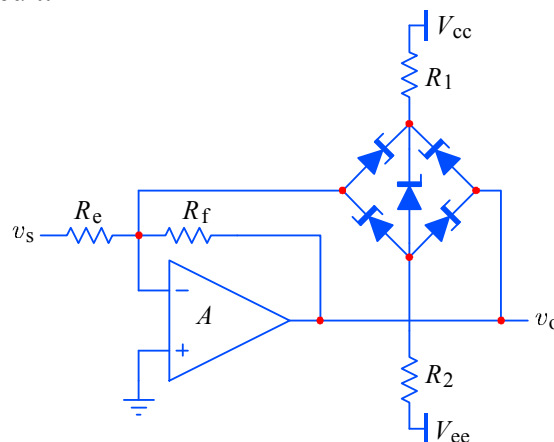


Fig. 5.3.31: The output level clipping is more precise if a biased Zener diode in a Schottky diode bridge is controlling the feedback. However, this circuit can be used only with voltage feedback unity gain compensated amplifiers.

Current feedback opamps do not like changes in feedback impedance, therefore a different output clipping scheme is used, having separate supply voltages for the output stage, as in Fig. 5.3.32. However, when the output reaches the level of its supply voltage the input stage loses feedback, so the input signal can overdrive the input differential stage. Usually we can prevent this by adding two anti-parallel Schottky diodes between the two inputs. Unfortunately, this would also increase the input capacitance.

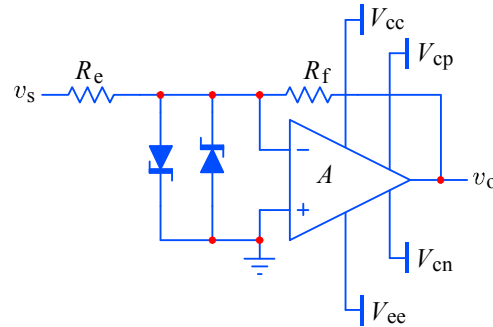


Fig. 5.3.32: The output buffer with a separate lower supply voltage can be used for output signal clipping with current feedback amplifiers. Since feedback is lost during clipping, the input stage must be protected from overdrive by anti-parallel Schottky diodes, which, inevitably, increase the input capacitance.

Another solution, often used with current feedback topology, is realized by limiting the voltage at the C_T internal node, using two normally closed voltage followers, as shown in Fig. 5.3.33. The addition of the voltage limiters increases the total capacitance at the C_T node, so, all other things being equal, limiting amplifiers tend to be slower. But with a careful design the bandwidth can still be quite high.

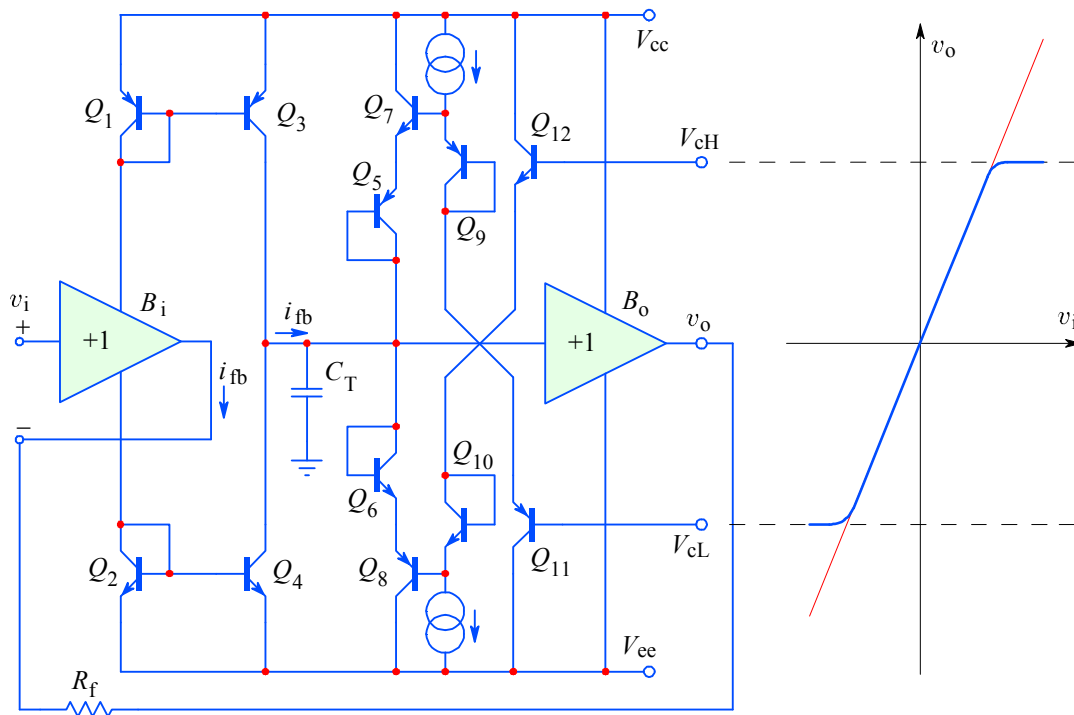


Fig. 5.3.33: Output signal clipping by limiting the internal C_T node of a current feedback amplifier. The transistors Q_{5-8} are normally reverse biased. For positive voltages $Q_{6,8}$ ($Q_{5,7}$ for negative) start conducting only when the voltage at C_T reaches V_{cH} (V_{cL}).

Transistors Q_{1-4} form the two current mirrors, which reflect the feedback current i_{fb} from the inverting input (the first buffer output) into the transimpedance node (at C_T). Transistors Q_5 and Q_7 are normally reverse biased and so are the B–E junctions of Q_6 and Q_8 . The transistors $Q_{5,7}$ start to conduct when the C_T voltage (and consequently the output voltage v_o) falls below V_{CL} . Likewise, the transistors $Q_{6,8}$ conduct when the C_T voltage exceeds V_{CH} . When either of these transistors conduct, they diverge the mirrored current i_{fb} to one of the supplies. Note that the voltages which set the clipping levels can be as high as the supply voltage. Also, they can be adjusted independently, as long as $V_{CH} > V_{CL}$. Since only two transistors at a time are required to switch on or off, the limiting action, as well as the recovery from limiting, can be very fast.

The most common misconception of overdrive recovery, even amongst more experienced engineers, is the belief that short switching times can be achieved only if the transistors are prevented from being saturated by artificially keeping them within a linear signal range. It often comes as a surprise if this does not solve the problem or, in some cases, can make it even worse.

It is true that adding Schottky diodes to a TTL gate makes it faster than ordinary TTL, and the inherently non-saturating ECL is even faster. Fast recovery is ultimately limited by the so called *minority carrier storage time* within the semiconductor material, and it depends on the type and concentration of dopants which determine the mobility of minority charge carriers. However, in analog circuits the problem is radically different from digital circuits, since we are interested in not just how quickly the output will start to catch up with the input, but rather how quickly it will follow the input to within 0.1 %, or even 0.01 %. In current state of the art circuitry, the best recovery times are < 5 ns for 0.1 % error and < 25 ns for 0.01 %.

In this respect it is the thermal ‘tails’ that are causing trouble. Wideband circuits need more power than conventional circuits, so good thermal balance is critical. Careful circuit design is required in order to keep temperature differences low, both during linear and non-linear modes of operation.

To some extent, we have been dealing with thermal problems in [Part 3](#). There we were discussing two-transistor circuits, such as the cascode and the differential amplifier. But the problem in multi-transistor circuits is that, even if it is differentially or complementary symmetrical, only one or two transistors will be saturated during overdrive, the rest of the circuit will either remain linear or cut off, which in this last case means cooling down. In saturation there is a low voltage across the device (usually a few tens of mV), so, even if the current through it is large, the power dissipation is low. Inevitably this results in different thermal histories of different parts of the circuit.

In integrated circuits the temperature gradients across the die are considerably lower than those between transistors in discrete circuits, but complex circuits can be large and heat conduction can be limited, so designers tend to reduce the power of auxiliary circuitry which is not essential for high speed signal handling. Therefore, hot spots can still exist and can cause trouble. Another important factor is gain switching and DC level adjustment, which must not affect the thermal balance, either because of amplifier working point changes or because of the control circuitry.

Circuits which rely on feedback for error correction can be inherently less sensitive to thermal effects (except for the input differential transistor pair!). However, the feedback stability, or, more precisely, the no feedback stability during overdrive, can cause identical or even worse problems. If there is insufficient damping during the transition from saturation back to the linear range, long ringing can result, impairing the recovery time. Such problems are usually solved by adding normally reverse biased diodes, which conduct during the saturation of a nearby transistor, allowing the remaining part of the circuit some control over critical nodes.

We will review a few possible solutions in the following section.

5.4 Improving Amplifier Linearity

The discussion of modern wideband amplifiers would not be complete without considering their linearity. In older wideband instrumentation a non-linearity of 1% was considered adequately low. In oscilloscopes this figure was comparable to the width of the luminous trace on the CRT screen.

In modern digitizing oscilloscopes the vertical resolution is limited by the resolution of the analog to digital converter (ADC), which for high sampling rate systems is rarely better than 8 bit ($1:2^8$, or $\sim 0.4\%$). At lower sampling rates the resolution can be 10 bit ($1:2^{10}$, or $\sim 0.1\%$) or 12 bit ($1:2^{12}$, or $\sim 0.025\%$). However, in such cases the digital display's resolution limits the readout accuracy (some new digital oscilloscopes have LCD screens with 1024 horizontal by 768 vertical pixels).

Nevertheless, the linearity issue is still important, because digital systems are often used when additional signal processing is required, either by the digital signal processor (DSP) within the 'scope itself or by an external computer to which the sampled signal is transferred. This processing can range from simple two-channel signal subtraction, addition, multiplication, etc., to averaging, convolution, or spectral analysis by fast Fourier transform (FFT). Note that spectral analysis, even if performed on 8 bit data, can offer an apparent 80 dB range (0.01 % resolution) if the signal in memory is long (1 MB signal acquisition length is not uncommon!) and if the FFT is performed on a large number of waveforms. It is therefore important to preserve all the linearity which the amplifier and the digitizer can offer.

At low frequencies the simplest way of improving linearity is by applying some sort of local corrective (negative) feedback at each amplifying stage, as we have seen in [Sec. 5.2](#). But at high frequencies the feedback can give more trouble than it can solve, owing to multiple high frequency poles and the total phase and time delay within the loop. Pole-zero compensation and phase correction can be used to a certain extent, but ultimately the amplifier's time delay sets the limit. With feedback the error can be reduced only and never eliminated, since the error reduction is proportional to the loop gain, which can not be infinite, and it also falls with frequency. So the error is small at low frequencies, increasing to its full value at the closed loop cut off.

At the highest frequencies the only error correction technique which can be made to work is the so called error 'feedforward'. This technique involves taking the driving signal error from an earlier circuit stage and then subtracting it from the output signal, so that the error is effectively canceled.

For the younger generation of engineers it is perhaps interesting to mention the historical perspective of feedback and feedforward error correction. Feedback is an omnipresent concept today, but it was not always so! In fact, both the feedback and the feedforward concepts were invented by *Harold S. Black* [[Ref. 5.39](#) and [5.40](#)]. However, feedforward was invented in 1923 and feedback in 1927, some 4 years later. While the patent for feedforward was granted in 1928, the feedback was such a "strange and counterintuitive" concept that the patent was granted almost 10 years after the application, in 1937! In spite of the fact that the feedback concept has been invented later and has been slow to catch up, once it happened it soon became the preferred method of error reduction, mostly owing to the work of *H. Nyquist* (1889–1976) and *H.W. Bode* (1905–1982), [[Ref. 5.67](#)]. On the other hand, feedforward was

almost forgotten, then ‘reinvented’ from time to time, only to be rediscovered by the broader engineering community in 1975, when the Quad 405 audio power amplifier came onto the market [Ref. 5.42 and 5.43]. Following the presentation article by the 405’s designer, *Peter J. Walker*, the idea was refined and generalized by a number of authors, among the first by *J. Vanderkooy* and *S.P. Lipshitz* [Ref. 5.44].

Later *M.J. Hawksford* [Ref. 5.46] showed that between the two extremes (pure feedback on one end and pure feedforward on the other) there is a whole spectrum of solutions combining both concepts. Moreover, such solutions can be applied both at system level (like the Quad 405 itself or similar solutions, as in [Ref. 5.48]) or down to the transistor level (as in [Ref. 5.47]).

It is interesting that while there were several examples of feedforward application in the field of RF communications (some even before 1975), most of the theoretical work was done with audio power amplification in mind. Apparently it took some time before the designers of high speed circuits grasped the full potential and the inherent advantages of the feedforward error correction techniques. In a way, this situation has not changed much, for even today we meet feedforward error correction mostly in RF communications systems and top level instrumentation. At the IC level we find feedforward only as a method of phase compensation (bypassing a slow inter-stage, not error correction), mostly in older opamps. From 1985 the advance in semiconductor processing has been extremely fast, discouraging amplifier designers from seeking more ‘exotic’ circuit topologies.

Before we examine the benefits of the feedforward technique for wideband amplification we shall first compare the feedback and feedforward principles from the point of view of error correction.

5.4.1 Feedback and Feedforward Error Correction

Fig. 5.4.1 shows the comparison of amplifiers with feedback and feedforward error correction. The feedforward amplifier is shown in its simplest form — later we shall see other possible realizations of the same principle.

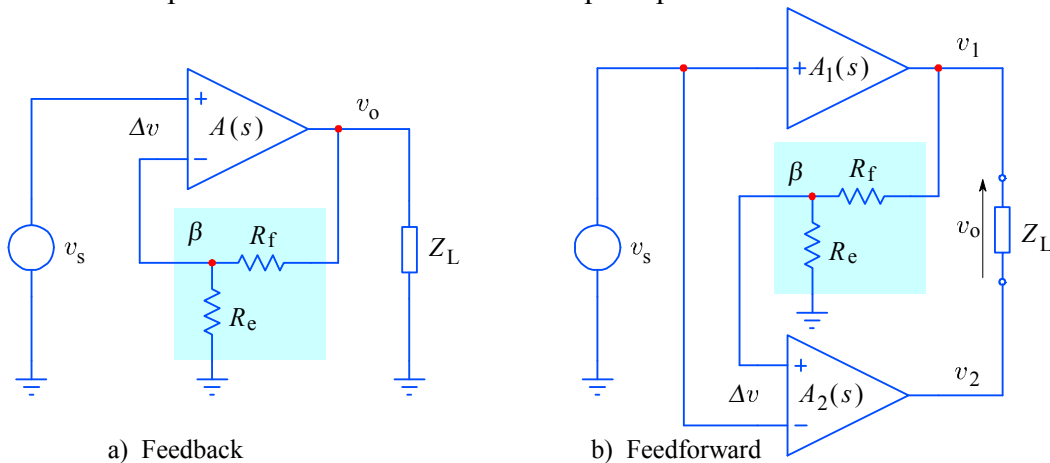


Fig. 5.4.1: Comparison of amplifiers with feedback and feedforward error correction. The feedback amplifier has excess gain, $A(s)$, which is reduced to the required level by taking the output voltage, suitably attenuated (β), to the inverting input of the differential amplifier (hence the name ‘negative feedback’). The feedforward case, in its most simple form, has two amplifiers: the main amplifier $A_1(s)$ is assisted by the auxiliary amplifier $A_2(s)$, which takes the difference between the input voltage and the attenuated main amplifier output forward to the load.

The analysis of the **feedback amplifier** has already been presented in [Sec. 5.3](#), but we shall repeat some expressions in order to correlate them with the feedforward amplifier. From [Fig. 5.4.1a](#) we can write:

$$v_o = (v_s - \beta v_o) \cdot A(s) \quad (5.4.1)$$

where $\beta = R_e / (R_f + R_e)$. By solving for v_o we have:

$$v_o = v_s \frac{A(s)}{1 + \beta A(s)} \quad (5.4.2)$$

The fraction on the right hand side is the amplifier closed loop gain G_{fb} ; it can be rewritten in such a form, from which it is easily seen that the gain expression can be approximated as $1/\beta$ if $A(s)$ is large:

$$G_{fb} = \frac{A(s)}{1 + \beta A(s)} = \frac{1}{\frac{1}{A(s)} + \beta} \approx \frac{1}{\beta} \Big|_{A(s) \rightarrow \infty} = 1 + \frac{R_f}{R_e} \quad (5.4.3)$$

Of course, this final simplification is valid at low frequencies only. Since $A(s)$ is finite and falls with frequency owing to the amplifier dominant pole s_0 :

$$A(s) = A_0 \frac{-s_0}{s - s_0} \quad (5.4.4)$$

the closed loop transfer function must also have a pole, but at a frequency s_h at which $A(s_h) = 1/\beta$. Since $f_h = |s_h|/2\pi$ and $f_0 = |s_0|/2\pi$ we can write:

$$\frac{1}{\beta} = A_0 \frac{f_0}{f_h + f_0} \quad (5.4.5)$$

and, considering that $\beta A_0 \gg 1$, the closed loop cut off frequency is:

$$f_h = f_0(\beta A_0 - 1) \approx f_0 \beta A_0 = f_0 \frac{A_0}{G_{fb}} \quad (5.4.6)$$

So the closed loop cut off frequency of a voltage feedback amplifier is inversely proportional to the closed loop gain.

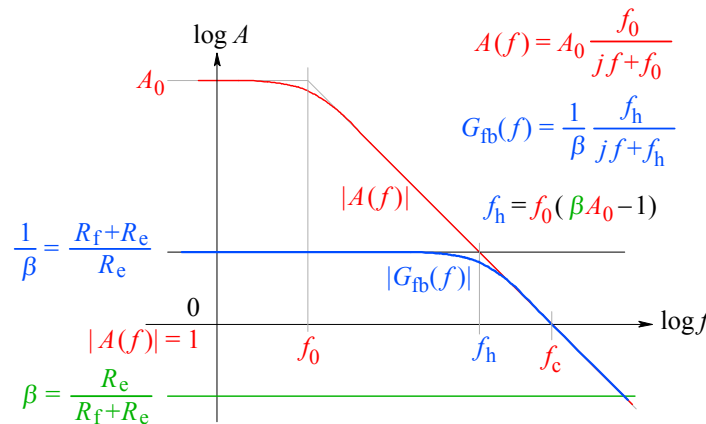


Fig. 5.4.2: For a voltage feedback amplifier the closed loop frequency response $G_{fb}(f)$ depends on the open loop gain A_0 , its dominant pole f_0 and the feedback attenuation factor β . The transition frequency f_c is equal to the amplifier gain bandwidth product (but only if the amplifier does not have a secondary pole close to or lower than f_c).

For the **feedforward amplifier**, [Fig. 5.4.1b](#), we must first realize that the load voltage, v_o , is the difference of the output voltages of individual amplifiers:

$$v_o = v_1 - v_2 \quad (5.4.7)$$

Since the main amplifier output is:

$$v_1 = v_s A_1(s) \quad (5.4.8)$$

and the auxiliary amplifier output is:

$$v_2 = (\beta v_1 - v_s) \cdot A_2(s) \quad (5.4.9)$$

it follows that the system output is:

$$\begin{aligned} v_o &= v_s A_1(s) - (\beta v_1 - v_s) \cdot A_2(s) \\ &= v_s [A_1(s) - \beta A_1(s) A_2(s) + A_2(s)] \end{aligned} \quad (5.4.10)$$

So by temporarily neglecting the frequency dependence, the closed loop gain of the feedforward amplifier is:

$$G_{ff} = A_1 + A_2 - \beta A_1 A_2 \quad (5.4.11)$$

The following reasoning is the most important point in feedforward amplifier analysis and it is probably difficult to foresee, but once we do, it becomes all too obvious. Let us say that, within the frequency range of interest, we can achieve:

$$\beta A_1 = 1 \quad (5.4.12)$$

Then:

$$G_{ff} = \frac{1}{\beta} + A_2 - \beta \frac{1}{\beta} A_2 = \frac{1}{\beta} \quad (5.4.13)$$

So, whatever the actual value of the auxiliary amplifier gain A_2 , the system's gain G_{ff} will be equal to $1/\beta$ if we can make $A_1 = 1/\beta$. Note that we have not requested any of the two gains to be very high, as we were forced to for feedback amplifiers, therefore this result is achieved without any approximation! True, if A_1 is frequency dependent and β is not, at high frequencies [Eq. 5.4.12](#) would not hold and, consequently, [Eq. 5.4.13](#) would not be so simple.

However, when A_1 starts to fall with frequency the factor $-\beta A_1 A_2$ is also reduced by the same amount, and the appropriate part of A_2 compensates the loss. This would be so as long as the gain A_2 remains constant with frequency (and even beyond its own cut off, provided that there still is enough gain for correction!).

Thus feedforward (in principle) achieves the dream goal:
a zero error, high cut off frequency gain, using non-ideal amplifiers!

We shall, of course, still have to deal with component tolerances, temperature dependencies, uncontrollable strays, time delays, etc., but with a manageable effort all these factors can be minimized, or, at least, kept below some predictable limit.

If you now think that there are no more surprises with feedforward amplifiers, consider the following points:

[Eq. 5.4.11](#) is, in a sense, symmetrical, thus $G_{ff} = 1/\beta$ (as in [Eq. 5.4.13](#)) can be achieved also if we decide to make:

$$\beta A_2 = 1 \quad (5.4.14)$$

However, the advantage of making $\beta A_1 = 1$ is that the input signal is canceled at the auxiliary amplifier differential input (remaining as a common mode signal only). In contrast with the ‘output balance’ condition represented by [Eq. 5.4.14](#), [Eq. 5.4.12](#) represents the so called ‘input balance’ condition, in which the auxiliary amplifier needs only a very low level amplitude swing at low frequencies (but rising to $1/2$ amplitude and higher at A_1 cut off and beyond, respectively). It therefore processes only the errors of the main amplifier, canceling them at the load and leaving only those of the auxiliary amplifier; as errors in processing the main amplifier error, those are secondary errors only.

It is also possible to make both gains equal to $1/\beta$:

$$A_1 = A_2 = 1/\beta \quad (5.4.15)$$

achieving in this way both input and output node balance.

In some instances there are good reasons to make A_1 larger than $1/\beta$, achieving also gain correction (in this case, the auxiliary amplifier will have to partly handle the input signal, too). Of course, in all these cases the gain matching has to be very precise, since in feedforward amplifiers the error cancellation is additive, as can be deduced from [Eq. 5.4.11](#), not multiplicative as in feedback amplifiers ([Eq. 5.4.2](#)).

Another nice property of the feedforward circuit is that it is feedback-free, so there are no loops causing potential instability, and consequently no stability criterion to satisfy. As a disadvantage, the output impedance is not lowered by the feedforward action (as it is in feedback amplifiers), so for high amplifier power efficiency it has to be made as low as possible within the frequency range of interest.

Let us return to the frequency dependence of the feedforward system gain. By defining:

$$A_1(s) = A_{01} \frac{-s_1}{s - s_1} \quad \text{and} \quad A_2(s) = A_{02} \frac{-s_2}{s - s_2} \quad (5.4.16)$$

the system gain is:

$$G_{ff} = A_{01} \frac{-s_1}{s - s_1} + A_{02} \frac{-s_2}{s - s_2} - \beta A_{02} A_{01} \frac{s_1 s_2}{(s - s_1)(s - s_2)} \quad (5.4.17)$$

where A_{01} and A_{02} are the main and auxiliary amplifier DC gain, respectively. By choosing $A_{01} = 1/\beta$ we get:

$$\begin{aligned} G_{ff} &= \frac{1}{\beta} \cdot \frac{-s_1}{s - s_1} + A_{02} \frac{-s_2}{s - s_2} \left(1 - \frac{-s_1}{s - s_1} \right) \\ &= \frac{1}{\beta} \cdot \frac{-s_1}{s - s_1} + A_{02} \frac{-s_2}{s - s_2} \cdot \frac{s}{s - s_1} \end{aligned} \quad (5.4.18)$$

Note that the auxiliary amplifier gain is effectively multiplied by the high pass version of the main amplifier's frequency dependence. If we now decide to make $A_{02} = 1/\beta$ also, we obtain:

$$\begin{aligned} G_{ff} &= \frac{1}{\beta} \left(\frac{-s_1}{s-s_1} + \frac{-s_2}{s-s_2} \cdot \frac{s}{s-s_1} \right) \\ &= \frac{1}{\beta} \left[\frac{-s_1}{s-s_1} \left(1 + \frac{s_2}{s_1} \cdot \frac{s}{s-s_2} \right) \right] \end{aligned} \quad (5.4.19)$$

The question is: is it desirable to make $s_1 = s_2$? Let us see:

$$G_{ff} = \frac{1}{\beta} \left(\frac{-s_1}{s-s_1} + \frac{-s_1 s}{(s-s_1)^2} \right) \quad (5.4.20)$$

The second fraction represents a second-order band pass response, which will add some gain peaking and extend the bandwidth by a factor of almost 3:

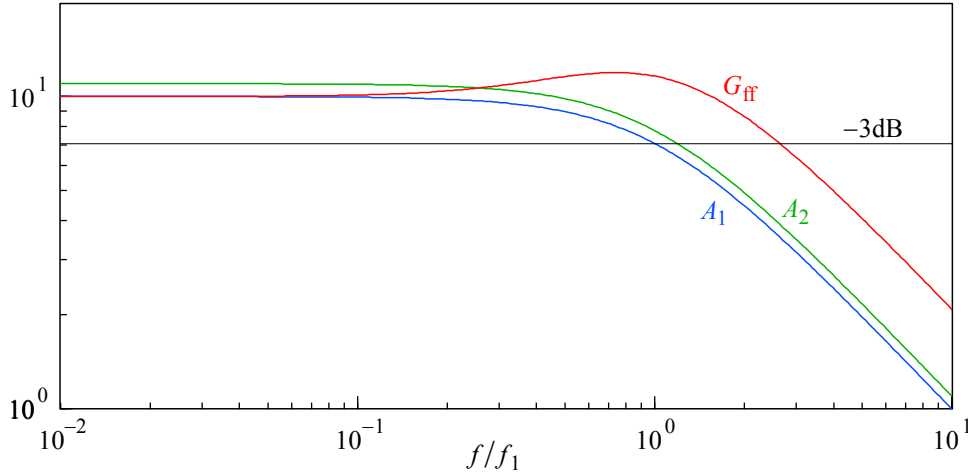


Fig. 5.4.3: The feedforward amplifier bandwidth G_{ff} is highest (and optimized in the sense of lowest gain \times bandwidth requirement of the auxiliary amplifier) if both amplifiers have the same bandwidth and the gains equal to $1/\beta$. In this figure $1/\beta = 10$, $A_{01} = 10$, and $A_{02} = 11$ (in order to distinguish A_2 from A_1 more easily). If $s_1 \neq s_2$ the G_{ff} bandwidth will be lower.

5.4.2 Error Reduction Analysis

In order to analyze the error reduction by both feedback and feedforward action we have to determine the system gain **sensitivity** on the amplifier gain variations. Generally, the concept of sensitivity of some system property, let us call it P , to variations of some subsystem parameter x is expressed as the x fraction of P , multiplied by the partial derivative of P on x :

$$S_x^P = \frac{x}{P} \cdot \frac{\partial P}{\partial x} \quad (5.4.21)$$

and it represents the amount of change in P for a unit change in x .

For the feedback amplifier we want to know the influence of variations in the amplifier open loop gain A to the closed loop gain G_{fb} , as defined by [Eq. 5.4.3](#):

$$\begin{aligned}
 S_A^{G_{fb}} &= \frac{A}{G_{fb}} \cdot \frac{\partial G_{fb}}{\partial A} = \frac{A}{\frac{A}{1 + \beta A}} \cdot \frac{\partial \frac{A}{1 + \beta A}}{\partial A} \\
 &= (1 + \beta A) \cdot \left[\frac{1}{1 + \beta A} - \frac{\beta A}{(1 + \beta A)^2} \right] = 1 - \frac{\beta A}{1 + \beta A} \\
 &= \frac{1}{1 + \beta A} \longrightarrow 0 \Big|_{A \rightarrow \infty}
 \end{aligned} \tag{5.4.22}$$

This means that the gain sensitivity is low only if A is very high. We also want to know the influence of variations of the feedback attenuation β :

$$\begin{aligned}
 S_\beta^{G_{fb}} &= \frac{\beta}{G_{fb}} \cdot \frac{\partial G_{fb}}{\partial \beta} = \frac{\beta}{\frac{A}{1 + \beta A}} \cdot \frac{\partial \frac{A}{1 + \beta A}}{\partial \beta} \\
 &= \frac{\beta(1 + \beta A)}{A} \cdot \left[-\frac{A^2}{(1 + \beta A)^2} \right] \\
 &= -\frac{\beta A}{1 + \beta A} \longrightarrow -1 \Big|_{A \rightarrow \infty}
 \end{aligned} \tag{5.4.23}$$

In the case of the feedforward amplifier, the influence of A_1 and A_2 on the system gain, as well as the influence of β , using [Eq. 5.4.11](#) for G_{ff} , is:

$$S_{A_1}^{G_{ff}} = \frac{A_1}{G_{ff}} \cdot \frac{\partial G_{ff}}{\partial A_1} = \frac{A_1(1 - \beta A_2)}{A_1(1 - \beta A_2) + A_2} = \begin{cases} 1 - \beta A_2 & \text{if } \beta A_1 = 1 \\ 0 & \text{if } \beta A_2 = 1 \end{cases} \tag{5.4.24}$$

$$S_{A_2}^{G_{ff}} = \frac{A_2}{G_{ff}} \cdot \frac{\partial G_{ff}}{\partial A_2} = \frac{A_2(1 - \beta A_1)}{A_2(1 - \beta A_1) + A_1} = \begin{cases} 0 & \text{if } \beta A_1 = 1 \\ 1 - \beta A_1 & \text{if } \beta A_2 = 1 \end{cases} \tag{5.4.25}$$

$$S_\beta^{G_{ff}} = \frac{\beta}{G_{ff}} \cdot \frac{\partial G_{ff}}{\partial \beta} = \frac{-\beta A_1 A_2}{A_1 + A_2 - \beta A_1 A_2} = \begin{cases} -\beta A_2 & \text{if } \beta A_1 = 1 \\ -\beta A_1 & \text{if } \beta A_2 = 1 \\ -1 & \text{if } A_1 = A_2 = 1/\beta \end{cases} \tag{5.4.26}$$

It is evident that the second condition in [Eq. 5.4.24](#) and the first in [Eq. 5.4.25](#), as well as the third condition in [Eq. 5.4.26](#), are the same as for the feedback amplifier. However, remember that for the feedback amplifier the results belong to the ideal case for which $A \rightarrow \infty$, so in practice they can be approximated only, whilst for the

feedforward amplifier they can be realized without any approximation (but within the limits of the specified component tolerances).

In a similar way we can determine the error reduction. For the feedback amplifier we have:

$$\frac{\varepsilon_{fb}}{\varepsilon_A} = \frac{1}{1 + \beta A} \longrightarrow 0 \Big|_{A \rightarrow \infty} \quad (5.4.27)$$

and again, zero distortion is achievable only in the idealized case of infinite gain.

In contrast, for the feedforward amplifier we have:

$$\frac{\varepsilon_{ff}}{\varepsilon_{A_1}} = 1 - \beta A_2 = 0 \Big|_{\beta A_2=1} \quad (5.4.28)$$

and this extraordinary result can be realized (not only approximated!) to whatever degree of precision we are satisfied with (accounting also for the technology cost).

Also, we must not forget that the open loop gain of the feedback amplifier decreases with frequency, so, for a given A_0 , the theoretically achievable maximum error reduction, $1/(1 + \beta A_0)$, is obtained only from DC up to the frequency of the dominant pole, f_1 ; beyond f_1 the error increases proportionally with frequency.

In contrast, feedforward amplifiers offer the same level of error reduction from DC up to the full feedforward system bandwidth and even beyond!

5.4.3 Alternative Feedforward Configurations

The main drawback of the feedforward amplifier in [Fig. 5.4.1b](#) is the ‘floating’ load (between the outputs of both amplifiers); in most cases we would prefer a ground referenced load, instead.

We have already noted that the output impedance of the main amplifier is not reduced by feedforward action; in fact, it does not need to be low in order to achieve effective error canceling. This leads to the idea of summing passively the two outputs, with that of the auxiliary amplifier inverted in phase:

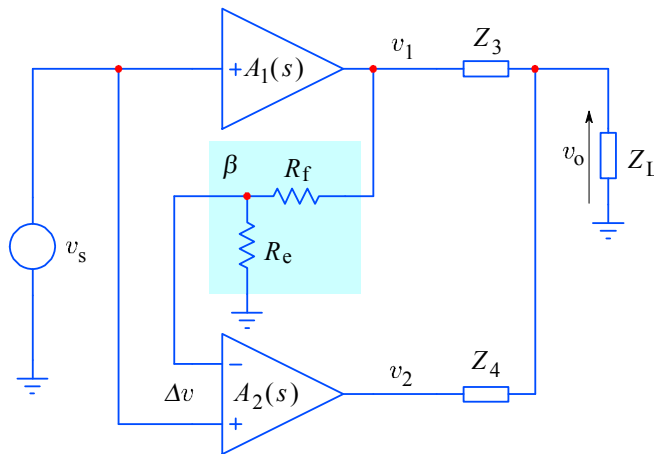


Fig. 5.4.4: Grounded load feedforward amplifier. Note the inverted input polarity of the auxiliary amplifier, compared with the circuit in [Fig. 5.4.1b](#). This allows passive signal summing over the output impedances Z_3 and Z_4 .

The impedances Z_3 and Z_4 should be lower than Z_L , but this condition is dictated mainly by amplifier power efficiency, not by the summing process. We have labeled Z_3 and Z_4 in accordance with tradition (emerging from the Quad 405 circuit, see [Fig. 5.4.6](#)) and also in accordance with a general form, in which Z_1 and Z_2 appear as feedback components to A_1 and A_2 . As implied by the Z symbol, these impedances can also be complex.

The output voltage v_o can be calculated by summing the two output currents:

$$i_1 = \frac{v_1 - v_o}{Z_3} \quad \text{and} \quad i_2 = \frac{v_2 - v_o}{Z_4} \quad (5.4.29)$$

so that:

$$v_o = Z_L(i_1 + i_2) \quad (5.4.30)$$

which results in:

$$v_o = \frac{Z_L}{Z_L + \frac{Z_3 Z_4}{Z_3 + Z_4}} \left(\frac{Z_3}{Z_3 + Z_4} v_1 + \frac{Z_4}{Z_3 + Z_4} v_2 \right) \quad (5.4.31)$$

By extracting the common attenuation factor, a :

$$a = \frac{Z_L}{Z_L + \frac{Z_3 Z_4}{Z_3 + Z_4}} \cdot \frac{Z_3}{Z_3 + Z_4} \quad (5.4.32)$$

the system's gain is:

$$G_{ff} = \frac{v_o}{v_s} = a \left[A_1 + \frac{Z_4}{Z_3} (A_2 - \beta A_1 A_2) \right] = \frac{a}{\beta} \left| \begin{array}{l} \text{when } \beta A_1 = 1 \\ \text{or } \beta A_2 = Z_3 / Z_4 \end{array} \right. \quad (5.4.33)$$

Because of the passive summing, the correct balance condition and error canceling for this circuit is achieved when the two output voltages are in the inverse ratio as the impedances:

$$\frac{v_2}{v_1} = \frac{Z_3}{Z_4} \quad (5.4.34)$$

If we assume that the output balance has been achieved, then:

$$v_2 = v_1 \beta A_2 \quad (5.4.35)$$

But the output balance condition is $\beta A_2 = 1$, therefore:

$$\frac{v_1 \beta A_2}{v_1} = \frac{Z_3}{Z_4} \Rightarrow \beta A_2 = \frac{Z_3}{Z_4} \quad (5.4.36)$$

On the other hand, the input balance condition, $\beta A_1 = 1$, because of [Eq. 5.4.34](#) and [5.4.36](#), results in the gain ratio equal to the impedance ratio:

$$\frac{A_2}{A_1} = \frac{Z_3}{Z_4} \quad (5.4.37)$$

The auxiliary amplifier will, under this condition, draw considerable current, even without the load. To reduce the current demand, we have to give up the input balance condition. If we set $v_2 = v_1$ then there would be no current if $Z_L = \infty$, and by choosing $Z_3 || Z_4 \ll Z_L$ the current demand is reduced for the nominal load:

$$v_2 = A_2 \left(\frac{v_1}{A_1} - \beta v_1 \right) = v_1 \quad (5.4.38)$$

and this means that:

$$\frac{A_2}{A_1} = \beta A_2 + 1 \quad (5.4.39)$$

which, considering [Eq. 5.4.35](#), results in:

$$\frac{A_2}{A_1} = \frac{Z_3}{Z_4} + 1 \quad (5.4.40)$$

and this should be compared with the ‘simple’ balance condition in [Eq. 5.4.37](#).

Another configuration, known as the ‘error take off’, by *Sandman* [[Ref. 5.49](#)], is shown in [Fig. 5.4.5](#). Here both the main and the auxiliary amplifier are of the negative feedback type; however, the auxiliary amplifier senses both the distortion and the gain error from the main amplifier feedback input and delivers it to the load in the same feedforward passive summing manner.

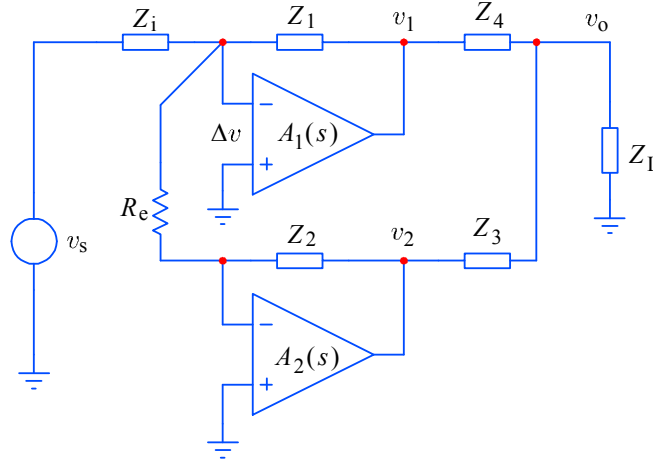


Fig. 5.4.5: ‘Error take off’ principle. The error Δv of the main amplifier, left by feedback, is taken by the auxiliary amplifier and fed forward to the load, where it is passively summed to the main output. The impedances Z_1 to Z_4 form the balancing bridge.

With an ideal main amplifier the voltage at its inverting input would be at a (virtual) ground potential; any signal Δv present at this point represents an attenuated version of the main amplifier error (gain error and distortion). If adequately amplified, it can be added to the main output to cancel the error at the load:

$$\Delta v = \beta \varepsilon = \frac{Z_i}{Z_i + Z_1} \varepsilon \quad (5.4.41)$$

To achieve effective error cancellation, we must set:

$$\frac{Z_2}{R_e} \cdot \frac{Z_i}{Z_i + Z_1} = \frac{Z_3}{Z_4} \quad (5.4.42)$$

A variation of this circuit is shown in [Fig. 5.4.6](#), which actually represents the original Quad 405 ‘current dumping’ amplifier configuration, and we now see how it follows from both the ‘error take off’ circuit and the pure feedforward circuit. If the balance condition is achieved A_2 must compensate whatever the amount of error at the A_1 output. It is important to realize that the input signal of A_1 can be taken from any suitable point within the circuit (the only condition is that it should, preferably, not be out of phase); the A_2 output represents just such a convenient point. The balance condition for the 405 is:

$$\frac{Z_3}{Z_4} = \frac{Z_2}{Z_1} \quad (5.4.43)$$

and the main amplifier error is canceled. Again, the impedances Z_n can be real or complex, whichever combination satisfies this equation.

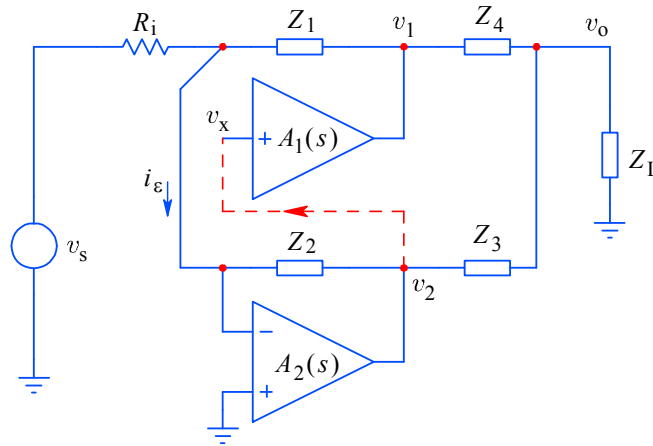


Fig. 5.4.6: Current dumping: by taking the main amplifier input signal (v_x) from the auxiliary amplifier (v_2), we obtain the original Quad 405 amplifier configuration. Although it is effective in error cancellation, its main disadvantages are the requirement for large voltage at the auxiliary amplifier output and a relatively low cut off frequency. In the Quad 405, Z_1 and Z_3 are resistors, Z_2 is a capacitor and Z_4 is an inductor.

A disadvantage of the current dumping scheme is that the auxiliary amplifier, although supplying relatively low current, must supply all the output voltage plus the error term; in such a condition the auxiliary amplifier's error can be rather high, and although it is a second-order error it can be significant. Also, in the classical 405 realization Z_2 (the feedback impedance of the auxiliary amplifier) is a capacitance (compensating the inductance Z_4), which results in a relatively slow system (high speed is not an issue at audio frequencies).

5.4.4 Time Delay Compensation

Let us return to the basic feedforward amplifier of [Fig. 5.4.1b](#). It is clear that if both the main and the auxiliary amplifier have limited bandwidth, time delays are inevitable. In order to work properly the feedforward scheme must include some time delay compensation, otherwise error cancellation close to and beyond the system cut off frequency would not occur.

For two sinusoidal signals a small time delay between them is transformed into a small amplitude error when summed and it might even be possible to compensate it

by altering the balance condition. However, for a square wave or a step function, large spikes, equal to full amplitude difference, would result and these can not be corrected by the balance setting components. Moreover, these spikes can overdrive the input of the auxiliary amplifier and saturate its output; error correction in such conditions is non-operating. Thus for high speed amplification some form of time delay compensation is mandatory.

In [Fig. 5.4.7](#) we see a general principle of time delay compensation. Since each amplifier has its own time delay acting on a different summing node, at least two separate time delay circuits are needed. Here, τ_1 compensates the delay of the main amplifier, allowing the auxiliary amplifier input to perform the difference between the input and β -attenuated signal with the correct phase. Likewise, τ_2 does so for the auxiliary amplifier delay for correct summing at the output.

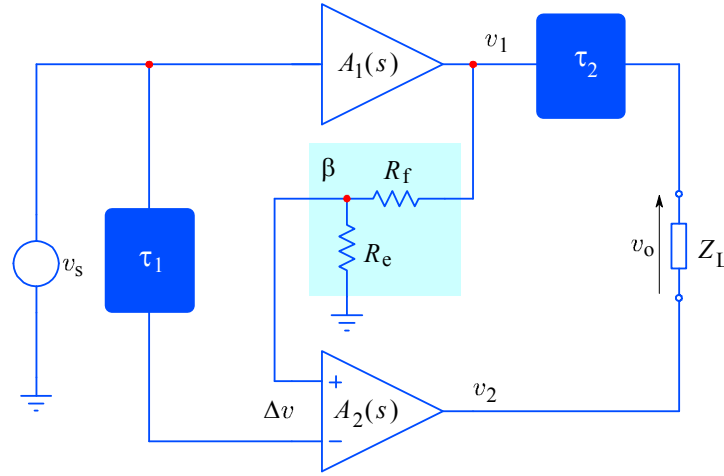


Fig. 5.4.7: Time delay compensation principle for the basic feedforward amplifier: τ_1 compensates the delay introduced by the main amplifier and τ_2 compensates the delay introduced by the auxiliary amplifier.

5.4.5 Circuits With Local Error Correction

In this section we shall briefly discuss a few simple circuits in which we shall employ our knowledge of feedback and feedforward error correction. We shall demonstrate how the same technique, which was developed at the system level, can also be applied at the local (subsystem) level. Local error correction often gives better results, since the linearization of individual amplifier stages lowers the requirements or indeed completely eliminates the need for system level correction. In many applications, such as oscilloscopes and adaptable data acquisition instrumentation, system level correction can be difficult to implement, owing to variable inter-stage conditions (variable attenuation, gain, DC level setting, trigger pick up delays, etc.).

The most interesting circuits to which the error correction is applied are the differential amplifier and the differential cascode amplifier. We have discussed them briefly in [Part 3, Sec. 3.7](#). Here we shall review the analysis with the emphasis on their non-linearity. The dominant non-linearity mechanism is the familiar exponential dependence of I_e to V_{be} of a single transistor:

$$I_e = I_s \left(e^{\frac{q_e}{k_B T} V_{be}} - 1 \right) \quad (5.4.44)$$

Here I_s is the saturation current (about 10^{-14} A in silicon transistors, depending on the dopant concentrations in the p-n junctions); the remaining symbols have the usual meaning (see [Part 3, Sec. 3.1](#)). Under normal operating conditions the DC emitter bias current I_{e0} exceeds I_s by at least 10^{11} , simplifying [Eq. 5.4.44](#) to:

$$I_e \approx I_s e^{\frac{q_e}{k_B T} V_{be}} \Big|_{I_{e0} \gg I_s} \quad (5.4.45)$$

For small signals, not altering the junction temperature considerably, we can assume $V_T = k_B T / q_e$ to be constant. So if we also neglect the dependence of the current gain and I_s on temperature and biasing, as well as the internal resistance and capacitance variations with the signal, we can express the non-linearity in form of the internal emitter resistance:

$$r_e = \frac{\partial V_{be}}{\partial I_e} \quad (5.4.46)$$

For the differential pair of [Fig. 5.4.8](#) the effective resistance seen by their base-emitter junctions is the sum:

$$r_{ed} = \frac{\partial V_{be1}}{\partial I_{e1}} + \frac{\partial V_{be2}}{\partial I_{e2}} \quad (5.4.47)$$

which, since one increases and the other decreases with the signal, varies much less over the much larger input signal range than in the single transistor case.

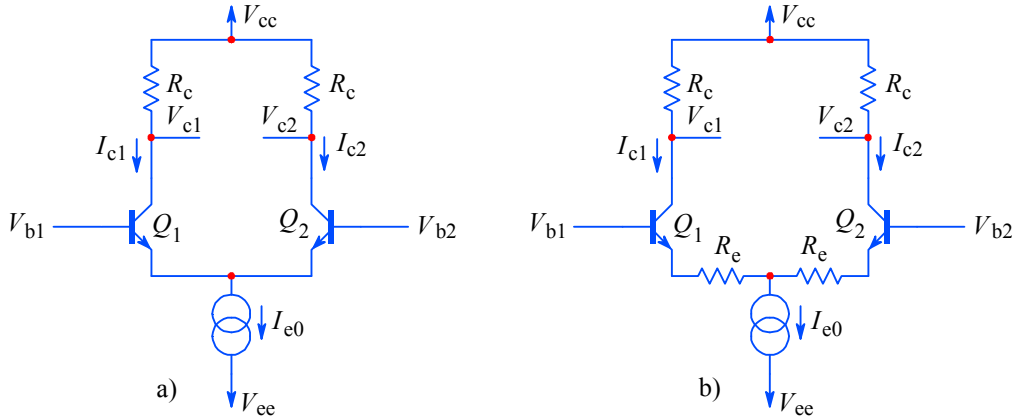


Fig. 5.4.8: **a)** A simple differential amplifier, showing the voltages and currents used in the analysis. **b)** The same, but with the emitter degeneration resistors R_e .

For the amplifier in [Fig. 5.4.8a](#) we must first realize that the differential input voltage is equal to the difference of the two V_{be} junction voltages:

$$V_{di} = V_{b1} - V_{b2} = V_{be1} - V_{be2} \quad (5.4.48)$$

We calculate the V_{be} junction voltages from [Eq. 5.4.45](#):

$$V_{be} = V_T \ln \frac{I_e}{I_s} \quad (5.4.49)$$

In an integrated circuit, $I_{s1} \approx I_{s2}$, so the ratio of emitter currents is:

$$\frac{I_{e1}}{I_{e2}} = \frac{e^{V_{be1}/V_T}}{e^{V_{be2}/V_T}} = e^{\frac{V_{be1}-V_{be2}}{V_T}} = e^{\frac{V_{di}}{V_T}} \quad (5.4.50)$$

The collector current $I_c = \alpha_F I_e$; also, the sum of emitter currents must be equal to the bias provided by the constant current source, I_{e0} . Thus:

$$I_{c1} + I_{c2} = \alpha_F(I_{e1} + I_{e2}) = \alpha_F I_{e0} \quad (5.4.51)$$

From the last two equations we obtain:

$$I_{c1} = \frac{\alpha_F I_{e0}}{1 + e^{-V_{di}/V_T}} \quad \text{and} \quad I_{c2} = \frac{\alpha_F I_{e0}}{1 + e^{+V_{di}/V_T}} \quad (5.4.52)$$

The collector voltage is equal to the potential drop $R_c I_c$ from the supply voltage:

$$V_{o1} = V_{cc} - R_c I_{c1} \quad \text{and} \quad V_{o2} = V_{cc} - R_c I_{c2} \quad (5.4.53)$$

Therefore the differential output voltage will follow a hyperbolic tangent function of the input differential voltage:

$$\begin{aligned} V_{do} &= V_{o1} - V_{o2} = R_c(I_{c2} - I_{c1}) \\ &= \alpha_F R_c I_{e0} \left(\frac{1}{1 + e^{+V_{di}/V_T}} - \frac{1}{1 + e^{-V_{di}/V_T}} \right) \\ &= \alpha_F R_c I_{e0} \tanh \frac{-V_{di}}{V_T} \end{aligned} \quad (5.4.54)$$

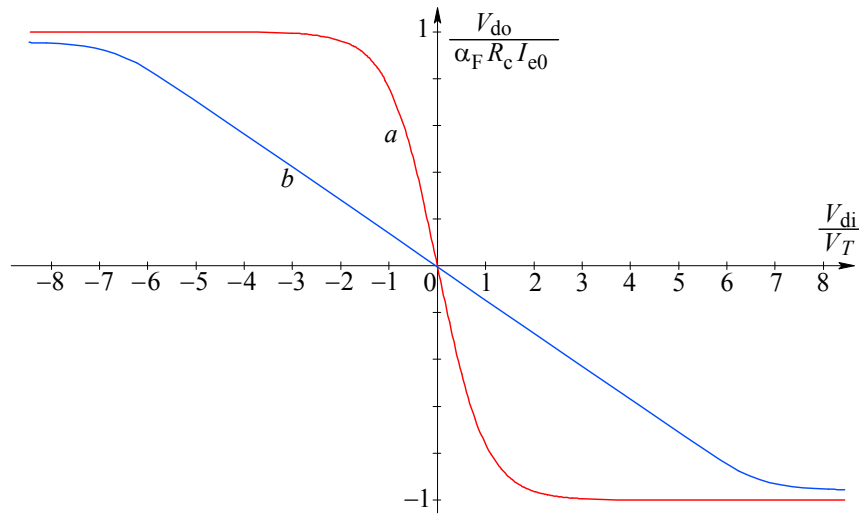


Fig. 5.4.9: **a)** The DC transfer function of the differential amplifier in [Fig. 5.4.8a](#): the input differential voltage is normalized to V_T and the output is normalized to $\alpha_F R_c I_{e0}$. **b)** With the emitter degeneration, as in [Fig. 5.4.8b](#), the transfer function is more linear, but at the expense of the gain (lower slope).

The system gain is represented by the slope of the plot, which, for $V_{di} \approx 0$, is:

$$\frac{V_{do}}{V_{di}} = \frac{\alpha_F R_c I_{e0}}{V_T} = \frac{\alpha_F R_c I_{e0}}{r_e I_{e0}} = \alpha_F \frac{R_c}{r_e} \quad (5.4.55)$$

Now, in wideband amplifier applications, the emitters are usually ‘degenerated’ by the addition of external resistors R_e . The degeneration resistor acts as a local current feedback, extending the linear part of the DC transfer function by $I_{e0}R_e + V_T$, instead of just V_T . Of course, this reduces the gain to $R_c/(R_e + r_e)$.

By considering one half of the differential pair and accounting for the bias current I_{e0} and signal current i flowing into R_e , the input voltage can be expressed from [Eq. 5.4.44](#) as:

$$V_{in} = (i - I_{e0})R_e + V_T \ln\left(\frac{i + I_{e0}}{I_s}\right) \quad (5.4.56)$$

and by differentiating this we obtain:

$$\partial V_{in} = R_e \partial i + \frac{V_T}{i + I_{e0}} \partial i \quad (5.4.57)$$

We separate the linear and non-linear components:

$$\partial V_{in} = \underbrace{\left(R_e + \frac{V_T}{I_{e0}}\right) \partial i}_{\text{linear}} - \underbrace{\frac{V_T i}{I_{e0}(i + I_{e0})} \partial i}_{\text{non-linear}} \quad (5.4.58)$$

We define the loading factor x as the ratio of the signal current to the bias current:

$$x = \frac{i}{I_{e0}} \quad (5.4.59)$$

So we can express the incremental non-linearity (INL) factor N as the ratio of the non-linear gain component to the linear one:

$$N(x) = \frac{-x}{1+x} \cdot \frac{V_T}{V_T + I_{e0}R_e} \quad (5.4.60)$$

Generally N can be (and usually is) a function of many variables, not just one.

In a similar way we can derive INL for the differential pair, where the differential input voltage is:

$$V_{in} = iR_e + (V_{be1} - V_{be2}) = iR_e + V_T \ln\left(\frac{I_{e0} + i}{I_{e0} - i}\right) \quad (5.4.61)$$

The linear and non-linear components are:

$$\partial V_{in} = \left(R_e + \frac{2V_T}{I_{e0}}\right) \partial i + \frac{2V_T i^2}{I_{e0}(i^2 - I_{e0}^2)} \partial i \quad (5.4.62)$$

and the INL:

$$N(x) = \frac{x^2}{1-x^2} \cdot \frac{2V_T}{2V_T + I_{e0}R_e} \quad (5.4.63)$$

This expression can be used to estimate the amount of error for a given signal and bias current, which an error correction scheme attempts to suppress.

In the following pages we are going to show a collection of differential amplifier circuits, employing some form of error correction, either feedback, feedforward or both. We are also going to present their frequency and time domain

performance to compare how the bandwidth has been affected as a result of increased circuit complexity (against a simple cascode amplifier).

For a fair comparison all circuits have been arranged to suit the test setup shown in [Fig. 5.4.10](#); the amplifiers were set for a gain of 2, using the same type of transistors (BF959) and biased by a 10 mA current source. The input signal was modeled by a 10 mA step driven current source, loaded by two $50\ \Omega \parallel 1\ \text{pF}$ networks. An equal network was used as the output load. Finally, all circuits were ‘tuned’ for a Bessel system response (MFED), using only capacitive emitter peaking (of course, inductive peaking can be used in the final design). Note that this setup offers only a relative indication of what can be achieved, not a final optimized design.

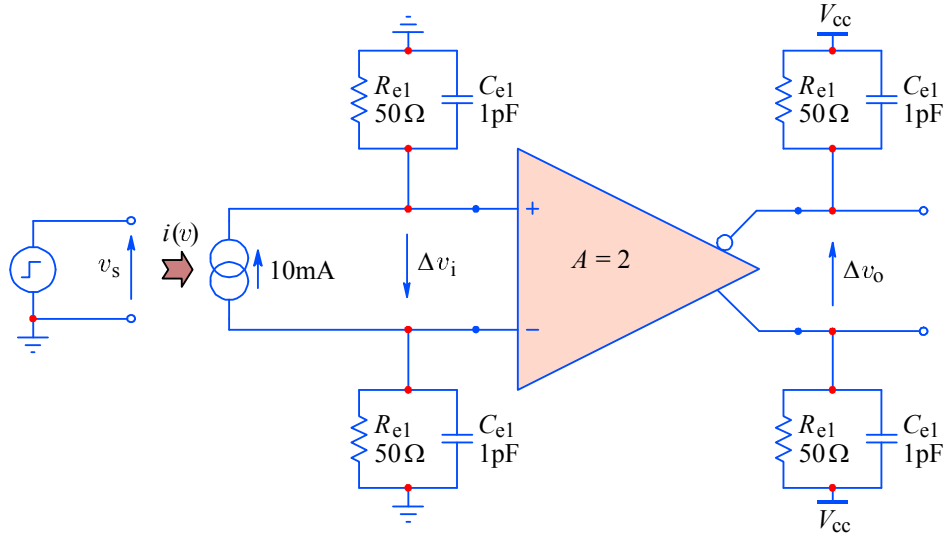


Fig. 5.4.10: Test set up used to compare different amplifier configurations.

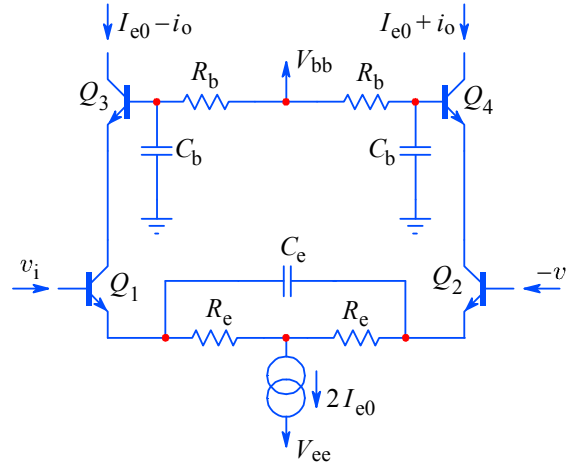


Fig. 5.4.11: This simple differential cascode amplifier, employing no error correction, is used in the test set up circuit of [Fig. 5.4.10](#), representing the reference against which all other amplifiers are compared. The emitter peaking and base impedance are adjusted for a MFED response.

The simple differential cascode amplifier of [Fig. 5.4.11](#), employing no error correction, represents a reference against which all other amplifiers will be compared. The $Q_{1,2}$ emitter peaking capacitor C_e and the $Q_{3,4}$ base network $R_b C_b$ are adjusted

for a MFED response. [Fig. 5.4.12](#) and [5.4.13](#) show the frequency domain and time domain responses, respectively.

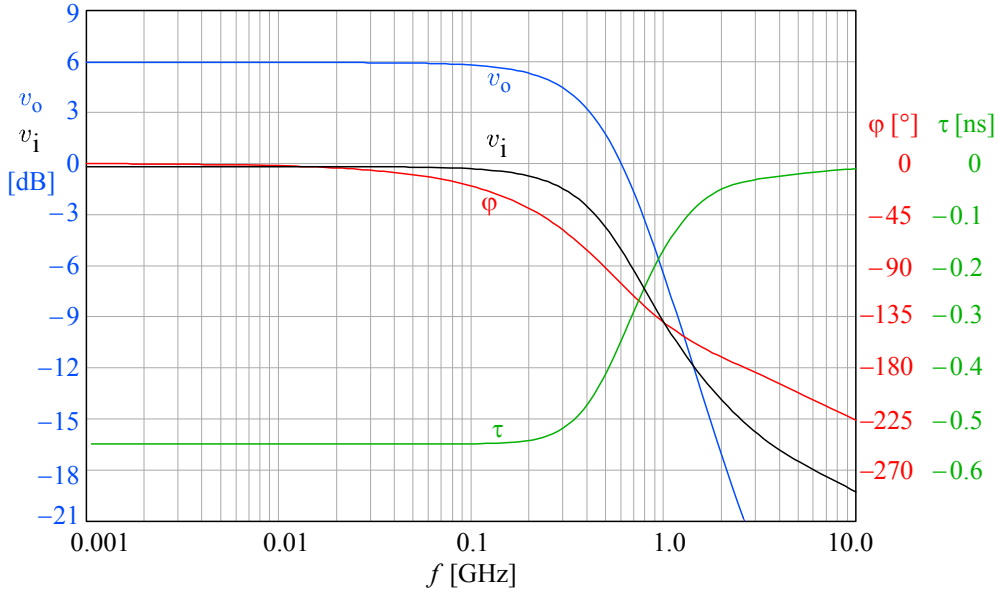


Fig. 5.4.12: Frequency domain performance of a simple differential cascode amplifier of [Fig. 5.4.11](#) (no error correction) used in the test set up circuit of [Fig. 5.4.10](#). This will be used as the reference for all other circuits. The bandwidth achieved is a little over 400 MHz.

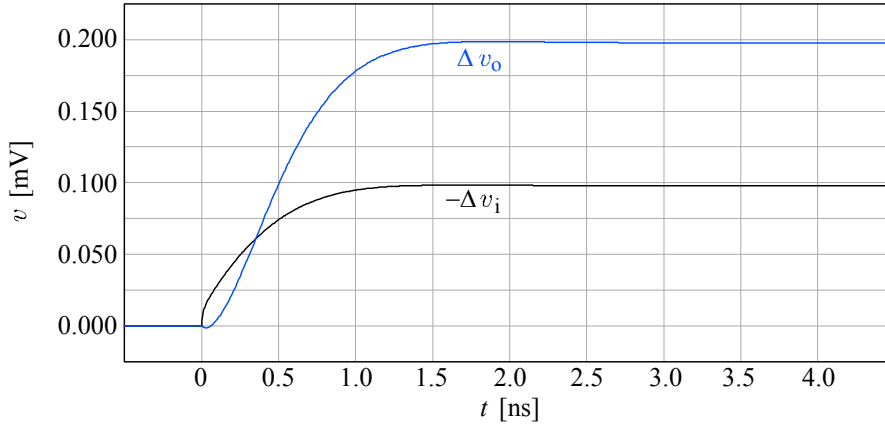


Fig. 5.4.13: Time domain performance of a simple differential cascode amplifier of [Fig. 5.4.11](#) (no error correction) used in the test set up circuit of [Fig. 5.4.10](#). This will be used as the reference for all other circuits. The input voltage indicates the input impedance dynamics in the first 100 ps and up to 1.5 ns. Note also the small undershoot at the output, owed to the cross-talk via C_{bc} of $Q_{1,2}$. The output rise time is less than 1 ns.

The first circuit to be compared with the reference is shown in [Fig. 5.4.14](#). The circuit is owed to *C.R. Battjes* [[Ref. 5.18](#)] and is functionally a Darlington connection (Q_1 , Q_2), improved by the addition of Q_3 . Used as the differential input stage of a cascode amplifier, it enhances the input characteristics and increases both the output current handling and the bandwidth. At a first glance it may seem that the diode connected Q_3 (shorted collector and base) can not do much. However, it allows Q_1 to carry a current much larger than the Q_2 base current, delivering it to the resistance R_e and lowering the impedance seen by the base of Q_2 , thus extending the bandwidth. The compound device has about twice the current gain of a single transistor.

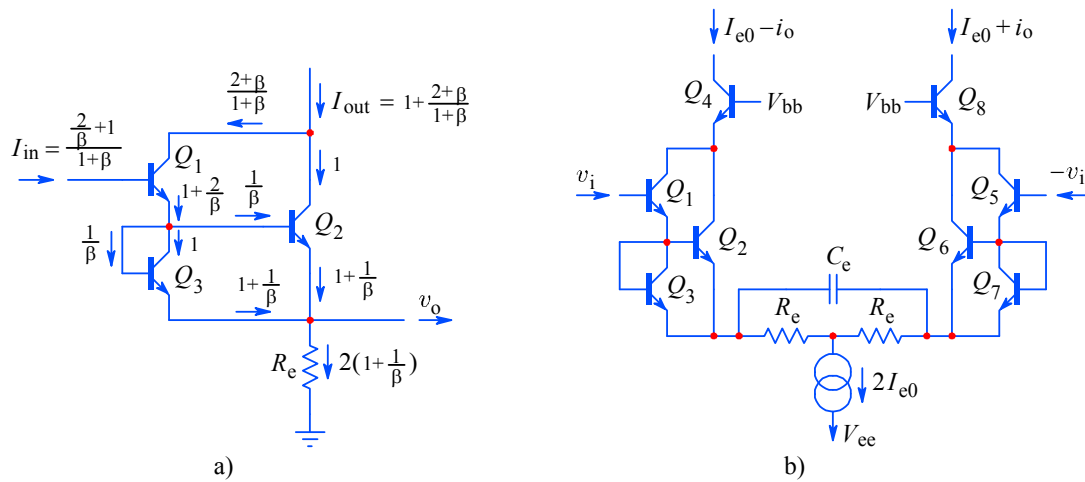


Fig. 5.4.14: a) Improved Darlington. b) Used as the input differential stage of the cascode amplifier — see the performance in the following figures.

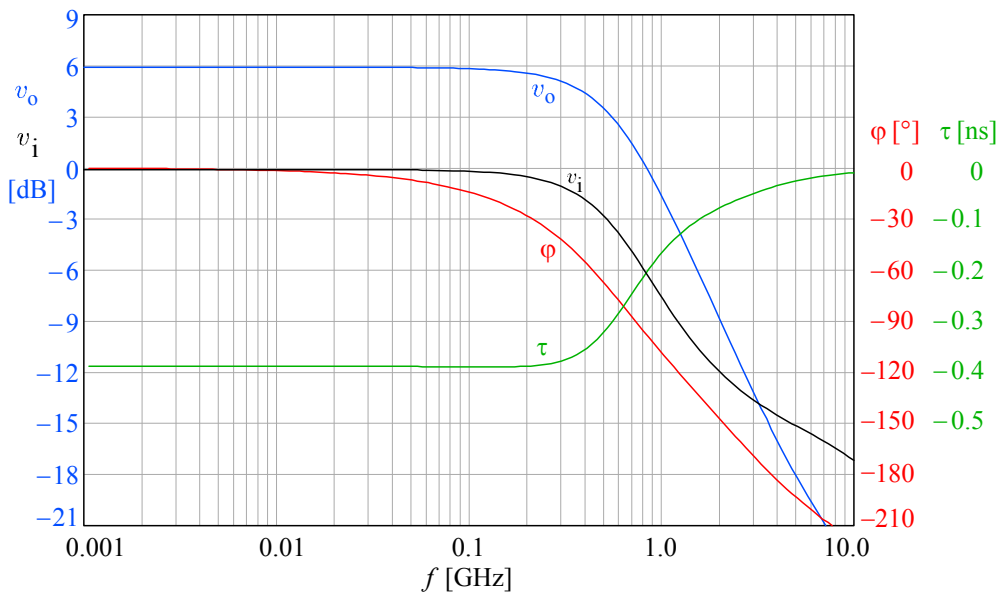


Fig. 5.4.15: Frequency domain performance of the differential cascode amplifier using the circuit of [Fig. 5.4.14b](#). The bandwidth is about 560 MHz. Note the input voltage changing slope above 2 GHz.

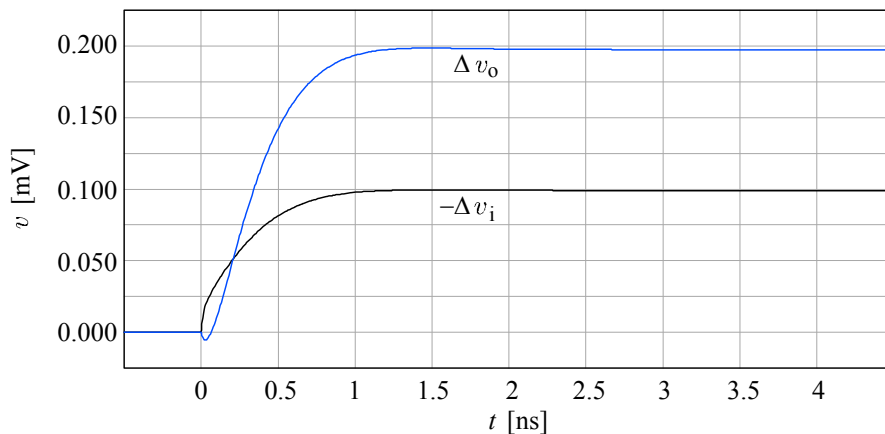


Fig. 5.4.16: Time domain performance of the differential cascode amplifier of [Fig. 5.4.14b](#). The undershoot has increased, but the rise time is less than 0.7 ns.

In [Fig. 5.4.17](#) Q_1 and Q_2 form the differential amplifier, whose error current i_1 is sensed at the resistor R_1 and is available at the collector of Q_3 for summing with the output current i_2 (error feedforward) further in the following circuit.

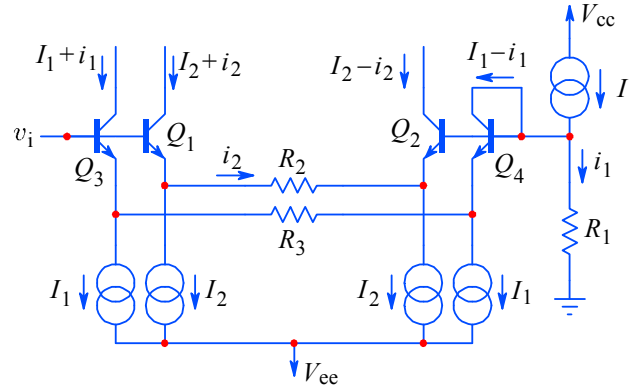


Fig. 5.4.17: A simple differential amplifier with feedforward error correction. Accurate matching of transistors is required only for DC error reduction, not for the feedforward linearization. Here i_2 is the differential current, whilst the error current, i_1 , sensed at R_1 , is available at the Q_3 collector to be added to the output current further in the circuit.

Two such circuits can form a differential amplifier, employing a double error feedforward correction, as shown in [Fig. 5.4.18](#). The error currents can now be summed directly with output currents, without further processing.

However, the main problem with this linearization technique is that R_1 must be relatively high for a suitable error sensing, so it can reduce the bandwidth considerably. In part the bandwidth can be improved by adding precisely matched capacitors in parallel to both R_2 and R_3 (emitter peaking), but then the input impedance can become negative and should be compensated accordingly. This negative input impedance compensation is easily implemented at $\pm v_i$ inputs, but, by adding it to the inputs connected to R_1 , the error sensing will be affected, since a part of i_1 would flow into the compensating networks, reducing error correction at high frequencies.

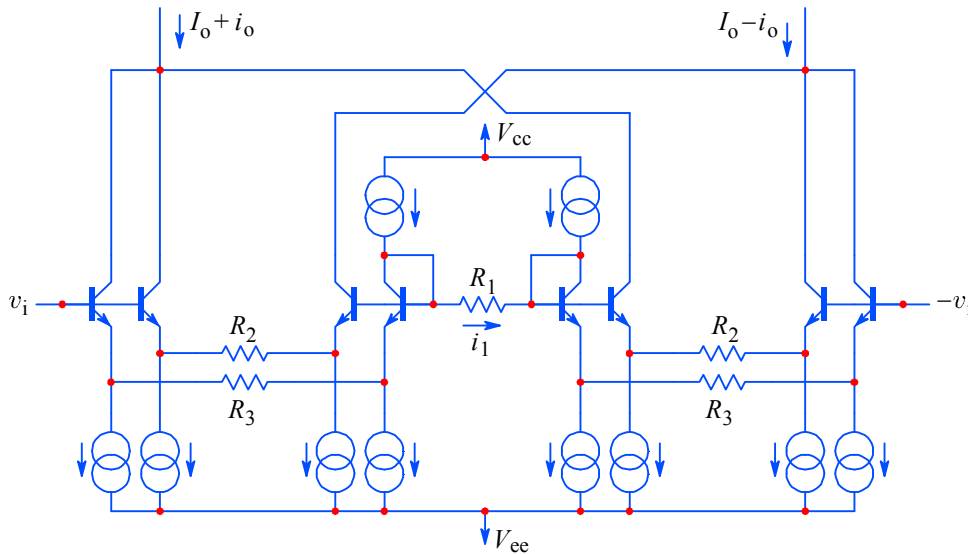


Fig. 5.4.18: Two circuits from [Fig. 5.4.17](#) can form a differential amplifier with a double error feedforward directly summed with the output.

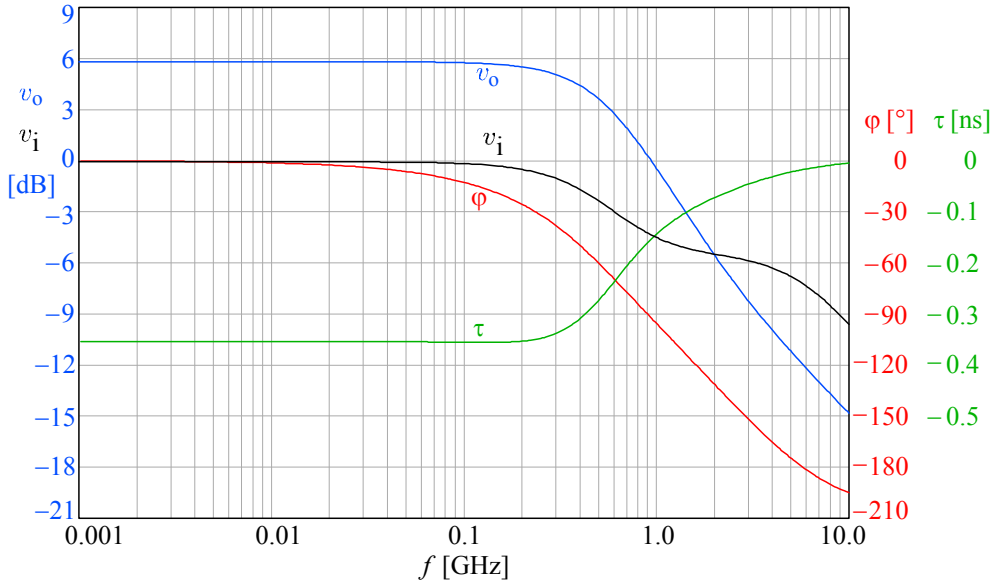


Fig. 5.4.19: Frequency domain performance of the circuit from Fig. 5.4.18. The bandwidth can be high (560 MHz), but for a suitable error sensing the required high value of R_1 would compromise it. The plot of v_i indicates the negative input impedance at high frequencies, which would need additional compensation networks at both inputs and at R_1 .

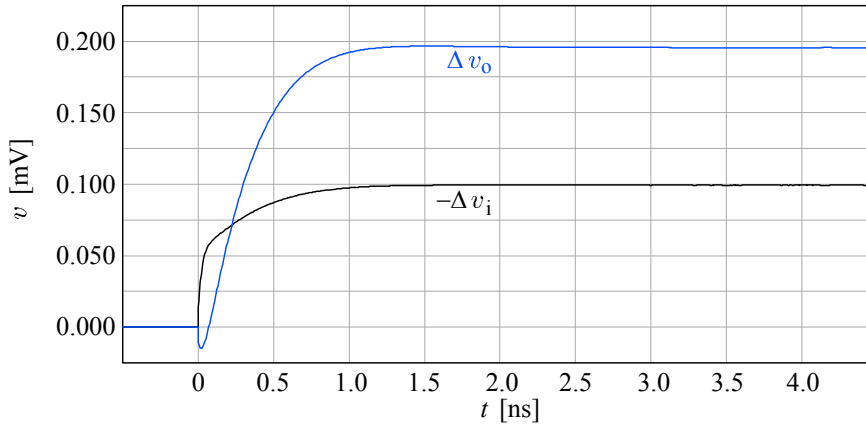


Fig. 5.4.20: Time domain performance of the circuit in Fig. 5.4.18. Note the jump of v_i in the first 100 ps and a pronounced undershoot in Δv_o .

A very interesting circuit, known as the ‘Cascomp’ (compensated cascode), shown in Fig. 5.4.21, was invented by *Patrick Quinn* [Ref. 5.53–54]. Here the usual differential cascode amplifier, Q_{1-4} , is enhanced by indirect error sensing and feedforward error correction. The error, generated at the emitters of $Q_{1,2}$, is also available at the emitters of $Q_{3,4}$, where it appears almost identical (owed to the same bias and signal currents and thus similar emitter resistances). Sensed by $Q_{5,6}$ and amplified adequately, the error is subtracted from the $Q_{3,4}$ collector currents. The addition of a further common base stage, $Q_{7,8}$, lowers the summing node impedance, which increases the summing precision, and at the same time improves the bandwidth of the error sensing amplifier, $Q_{5,6}$. Since the error signal voltage at the $Q_{3,4}$ emitters is low (several mV, or so), the emitter resistances R_{e2} of the error amplifier $Q_{5,6}$ can also be very low, or even eliminated completely, without degrading the linearity.

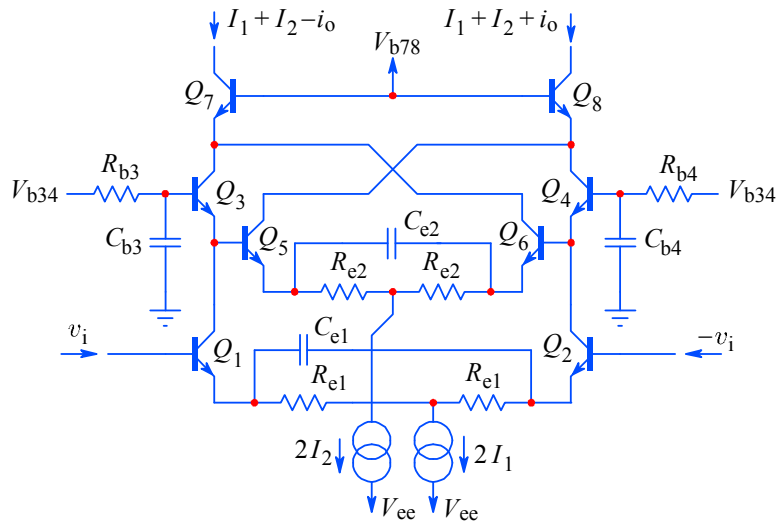


Fig. 5.4.21: The Quinn's 'Cascomp' employs indirect error sensing and error feedforward.

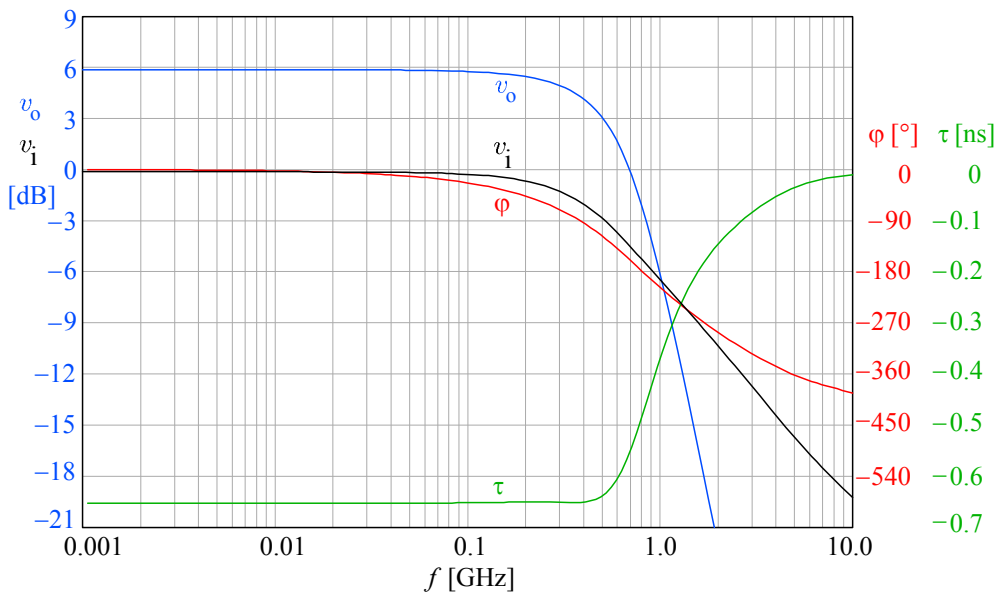


Fig. 5.4.22: Frequency domain performance of the 'Cascomp' amplifier. The bandwidth is about 500 MHz, but note also the high flatness of the envelope delay, right up to the bandwidth limit.

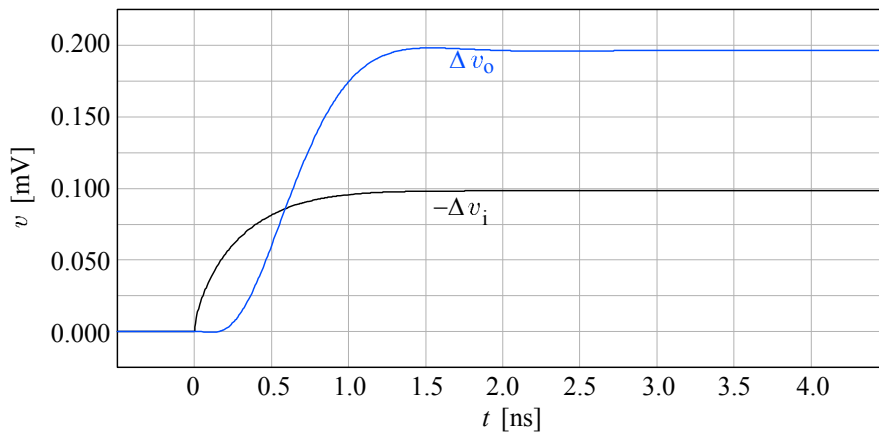


Fig. 5.4.23: Time domain performance of the 'Cascomp' amplifier. The rise time is about 0.7 ns and the initial undershoot is very low.

[Fig. 5.4.24](#) shows a similar circuit, but with a feedback error correction. The error signal is taken at the same point as before, but its amplified version is applied to the emitters of the input differential pair $Q_{1,2}$. Unfortunately, in spite of its attractive concept this configuration is not suitable for high frequencies, since capacitive emitter peaking can not be used (a capacitance in parallel to R_1 would short the auxiliary amplifier outputs, reducing the amount of error correction at high frequencies), thus the bandwidth is about 180 MHz only. But when bandwidth is not the primary design requirement this amplifier can be a valid choice. We shall not plot its performance.

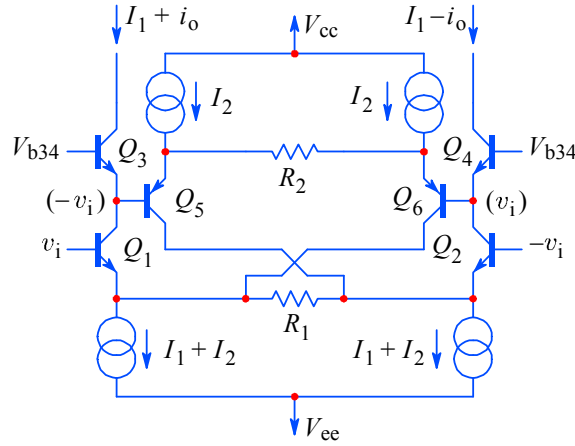


Fig. 5.4.24: Differential cascode amplifier with indirect error sensing and error feedback. Unfortunately this configuration tends to be rather slow.

The circuit in [Fig. 5.4.25](#) represents a modification of the feedforward error correction in which the correction current is summed at the same point where the error was generated. This configuration requires reasonably well matched devices with a high current gain β . Transistors $Q_{1,2}$ form the differential pair, $Q_{5,6}$ form the error sensing amplifier and $Q_{3,4}$ provide an additional V_{be} voltage to enhance the error amplifier dynamic range.

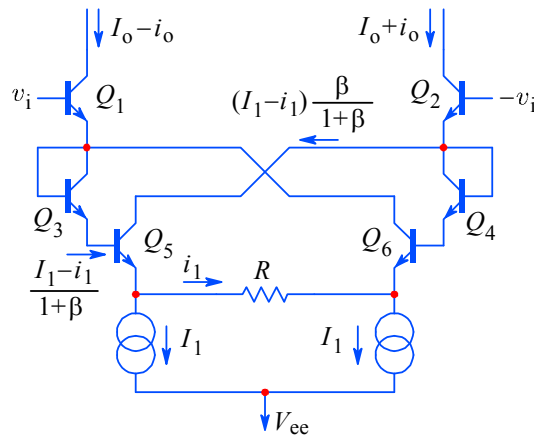


Fig. 5.4.25: A possible modification of error feedforward, employing direct error sensing and direct feedforward error correction.

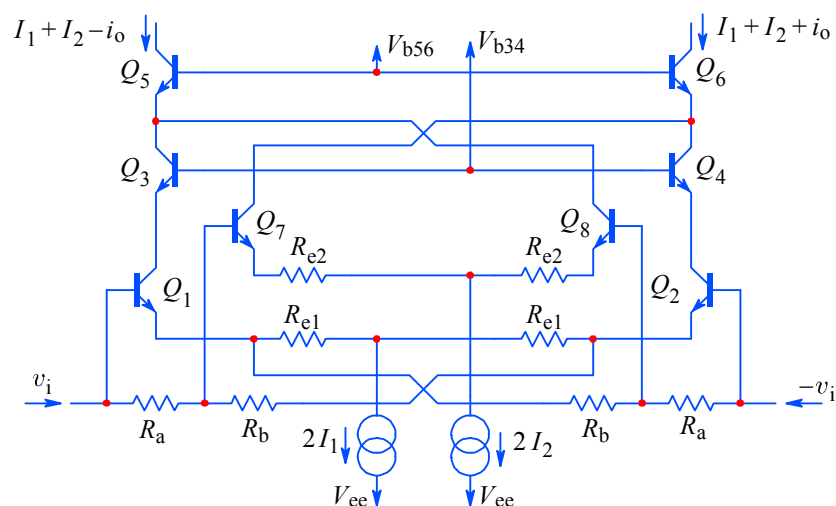


Fig. 5.4.26: Another evolution of the Cascomp is realized by feedback derived error sensing and feedforward error correction. The junction of R_a and R_b is at the ‘virtual ground’ at which the error of the $Q_{1,2}$ pair is sensed and amplified by the auxiliary amplifier $Q_{7,8}$. The error is subtracted from the output current at the emitters of $Q_{5,6}$.

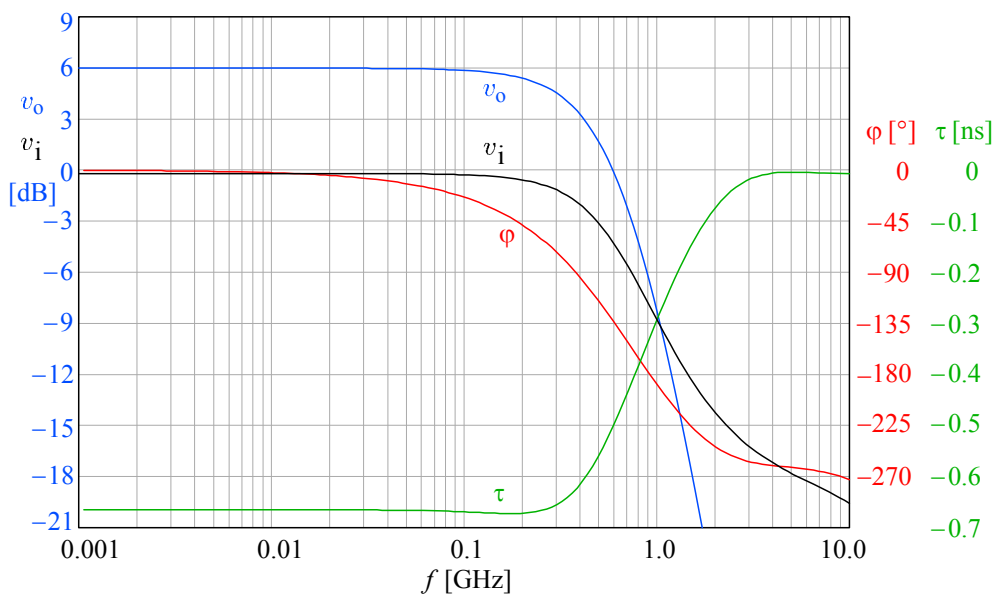


Fig. 5.4.27: Frequency domain performance of the Cascomp evolution amplifier. The bandwidth is a little over 400 MHz.

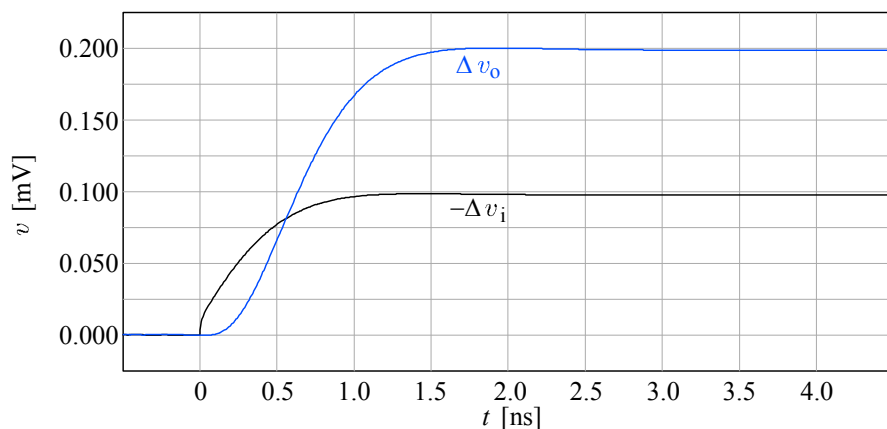


Fig. 5.4.28: Time domain performance of the Cascomp evolution amplifier.

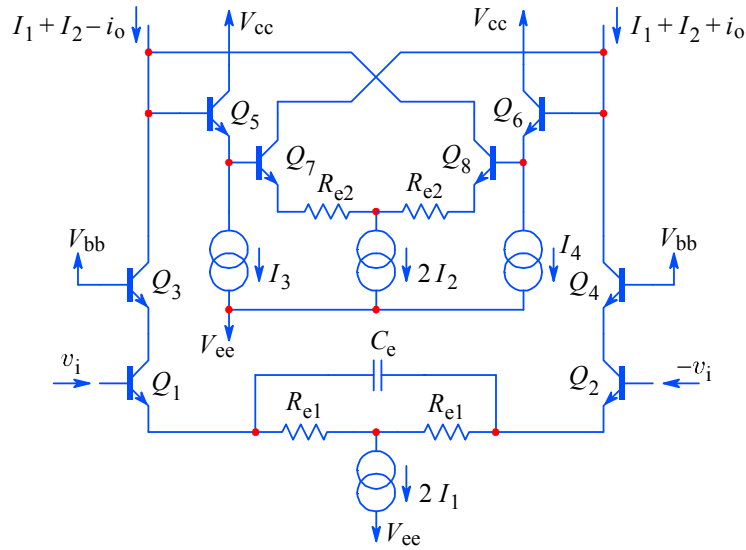


Fig. 5.4.29: This ‘output impedance compensation’, also patented by Pat Quinn, has direct error sensing and direct feedforward error correction, performed by Q_{5-8} .

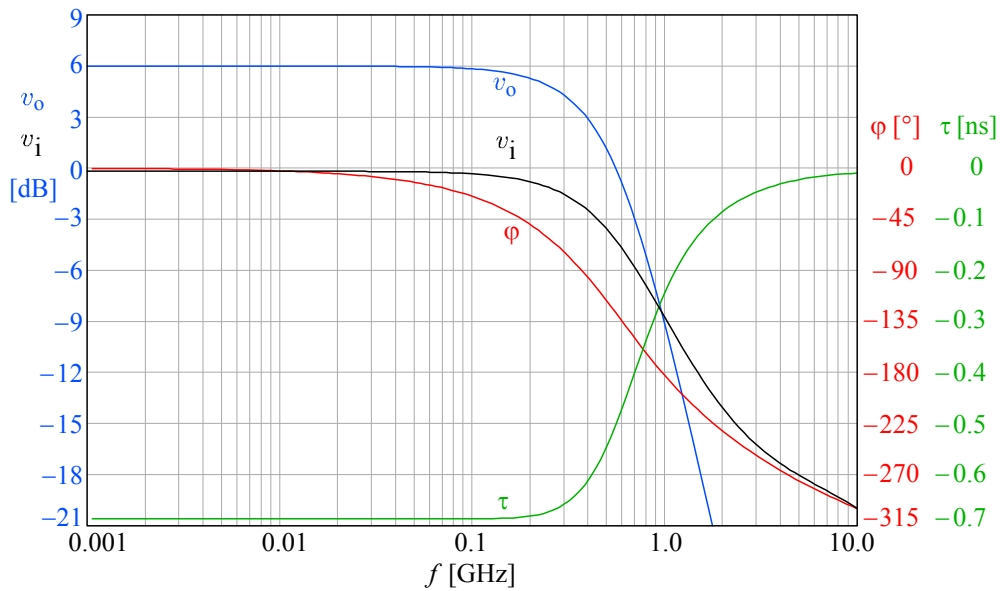


Fig. 5.4.30: Frequency domain performance of the amplifier in [Fig. 5.4.29](#). The bandwidth is about 400 MHz.

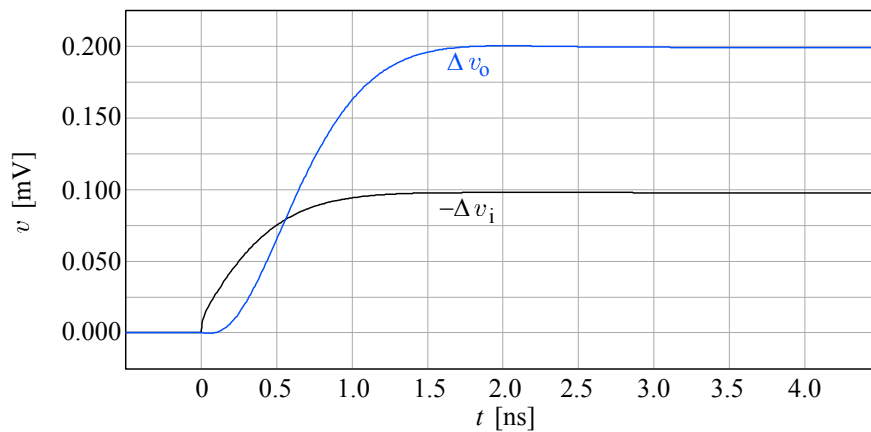


Fig. 5.4.31: Time domain performance of the amplifier in [Fig. 5.4.29](#).

5.4.6 The Tektronix M377 IC

In this section we shall have a brief discussion of the M377 IC, made by Tektronix for their 11 000 series oscilloscope, later also employed in many other models. The M377 was described in [Ref. 5.50 and 5.51] by its designer *John Addis*.

When it was designed, back in mid 1980s, this integrated circuit started a revolutionary trend in oscilloscope design which is still evolving today in the first decade of the XXI century. One of the important design goals was to eliminate the need for manual adjustments as much as possible. Classical oscilloscopes, made with discrete components, required a lot of manual adjustment; for example, the Tektronix 7104 oscilloscope with its two single-channel plug-ins required 32 manual adjustments for correcting the thermal effects only, many of which needed iterative corrections in combination with another one or several other settings. This caused long calibration periods and a lot of bench testing by experienced personnel, increasing the production cost considerably. In contrast, the M377 needs only one manual adjustment (for optimizing the transient response), all other calibration procedures are performed at power-up by a microprocessor, which varies the settings using several DC voltage control inputs. Of course, the calibration can also be done upon the user's request by pressing a push button on the front panel. Some settings, such as the high impedance attenuator calibration, is done in production by laser trimming of the components.

The elimination of almost all electromechanical adjustments and their replacement by electronic controls resulted in circuit miniaturization, reducing strays and parasitics, thus improving bandwidth. But it also increased the circuit complexity (the number of active devices) and density, resulting in higher power dissipation, and consequently higher temperatures, requiring careful thermal compensation.

The first step in this direction was the 'Cascomp', [Ref. 5.53–5.54], which we have met already in Fig. 5.4.21. The first instrument to use the Cascomp was the Tektronix 2465 oscilloscope. Although it represented a significant improvement in precision over a simple cascode differential pair, it also had some limitations. First, the addition of the error amplifier, $Q_{5,6}$, helps to reduce both the nonlinearity and, if the operating points are chosen carefully, also the thermally induced tails. But since both the main and the error amplifier have nearly equal gain, the system gain can not be high, otherwise the error amplifier's non-linearity would show up.

Another problem is the stack of three transistor pairs in the signal path, which results in an increased gain sensitivity to α variations. The gain, as a function of temperature, increases as α^3 , about 225 ppm/K for a Cascomp using transistors with $\beta = 80$, compared to some 150 ppm/K of the standard cascode. But a standard cascode also has a counteracting temperature dependent gain term, ≈ -185 ppm/K (at $I_e = 20$ mA, $T_j = 60^\circ\text{C}$ and $R_e = 40\ \Omega$) owed to the dynamic emitter resistance r_e , which in the Cascomp is compensated. To some degree α effects can be compensated by adding resistors $R_{b3,4}$ to the bases of $Q_{3,4}$, but these resistors can never be made large enough, owing to the transient response requirement for low base resistance (see Part 3, Fig. 3.4.6 and the associated analysis).

The three transistor pair stack (and not forgetting the current source, I_1) further disadvantage the IC against the discrete design, since for a given supply voltage the output linear voltage range is severely reduced. Also, any level shifting back down

to the negative supply voltage, needed by an eventual subsequent stage, requires a greater level shift than in a conventional cascode.

Finally, the Cascomp has a limited ability to handle overdrive signals. The emitters of $Q_{3,4}$ do not 'see' the whole signal during overdrive, thus the error amplifier signal is clipped off at peaks, and as a result the main and the error amplifier experience different thermal histories. Additional circuitry is needed to ensure correct input signal clipping and acceptable thermal behavior.

All these limitations dictated a different approach in the M377 IC. The basic amplifier block (shown in Fig. 5.4.32 and a differential version in Fig. 5.4.33) is already inherently insensitive to thermal influence, and uses local feedback, resulting in high linearity. It can be viewed (simplified) as a compound transistor, with the base represented by the Q_1 base, the emitter by the Q_3 emitter, and the collector by the Q_3 collector, respectively. Since Q_3 is current driven, its V_{be} variation with temperature is not important. Compared with a single transistor, operating at the same current, such a compound transistor has greater g_m and β , an excellent linearity, and no thermals.

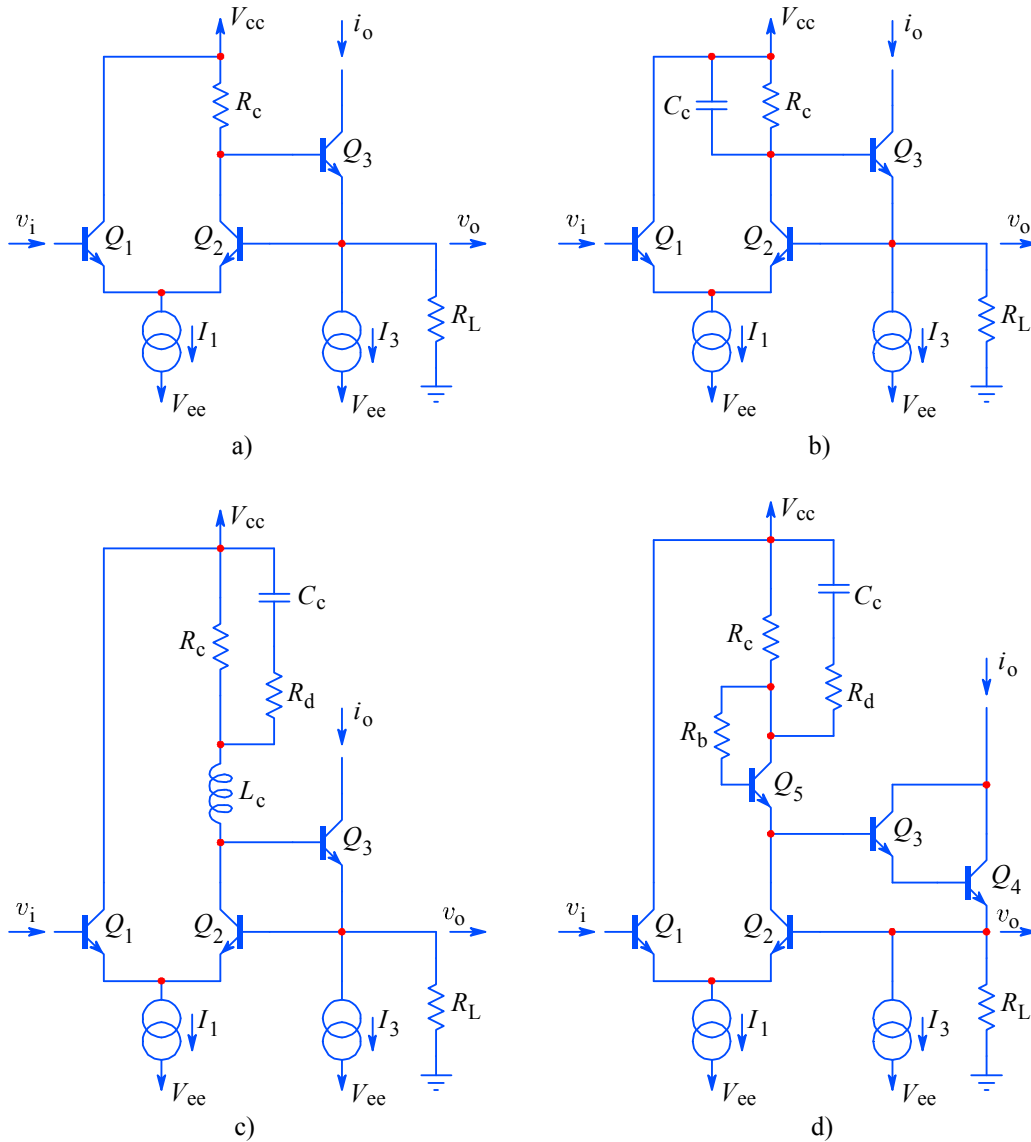


Fig. 5.4.32: a) The M377 IC main amplifier block, basic scheme. b) Dominant pole compensation. c) Inductive peaking. d) Inductance created by the Q_5 emitter impedance.

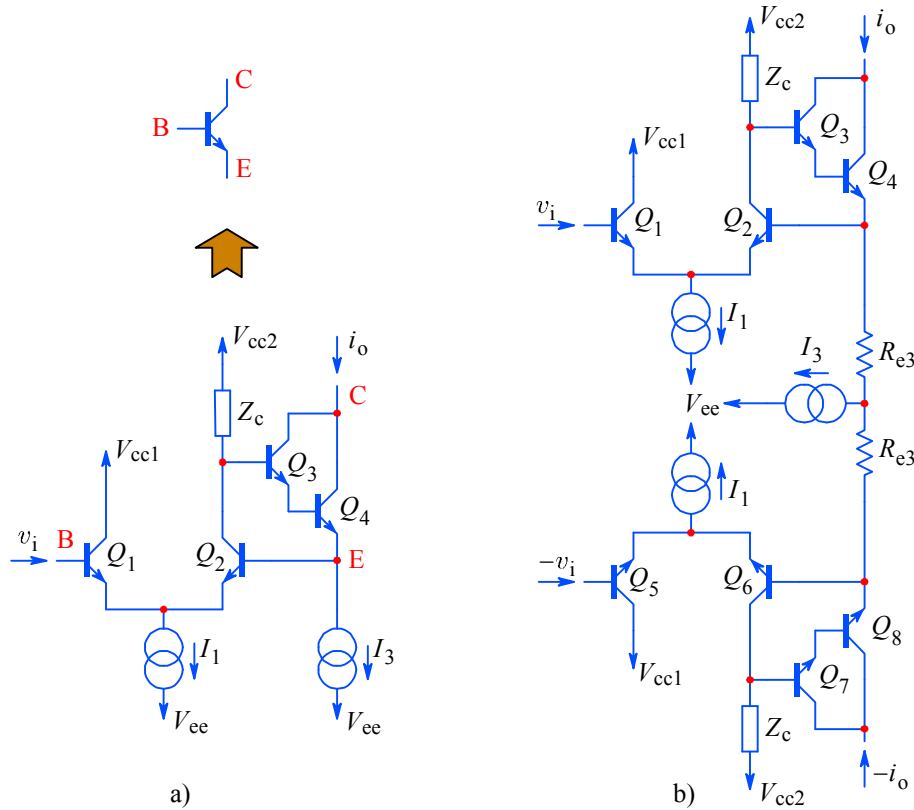


Fig. 5.4.33: **a)** The basic wideband amplifier block of the M377 IC can be viewed as a compound transistor, in which the Q_1 base, the Q_3 emitter, and the Q_3 collector represent the base, emitter and collector of the compound device. **b)** Two such blocks form a differential amplifier. An improved design results if Q_3 is of a Darlington configuration, Q_3+Q_4 . To a high precision the output current is $i_o = v_i/R_{e3}$. The impedance Z_c is the compensation shown in [Fig. 5.4.32d](#).

However, the α of Q_3 is not improved, neither by Q_1 , nor Q_2 . This could be corrected if, for example, the Q_1 collector were to be connected to the Q_3 emitter, which, besides improving α , would also bootstrap the collector of Q_1 , lowering the input capacitance. Unfortunately, the low collector voltage and, consequently the operating point of Q_1 , would reduce its cut off frequency. Also, owing to feedback through the capacitance C_{cb} of Q_1 , such a circuit can have a negative input impedance at very high frequencies, compromising stability. The best solution for a high f_T (≈ 8.5 GHz in the M377) is to use a Darlington for Q_3 (Q_3+Q_4 in [Fig 5.4.33](#)).

One of the most important parameters of a high speed amplifier is the overdrive recovery time. Following the ever increasing requirements for speed and precision we usually specify the time (in ns) needed to settle to, say, 0.1% or even 0.01% of the final level, following an overdrive level of many times the maximum level over a relatively long overdrive period. [Fig. 5.4.34](#) shows the configuration used in the M377, which recovers to 0.04% in just 6 ns after a 2 V overdrive; recovery to 0.01% is about 25 ns. The current sources and Schottky diodes, added to the original circuit of [Fig. 5.4.33b](#), allow separate feedback paths under overdrive. It is important to realize that in this circuit, if not compensated, only the half with the negative input voltage would be overdriven. The positive part then takes all the current provided by the I_3 source. Under this condition D_5 and D_6 cut off, while D_4 conducts owing to I_2 and I_4 , allowing the feedback of the lower circuit half to remain operative, preserving the delicate thermal balance.

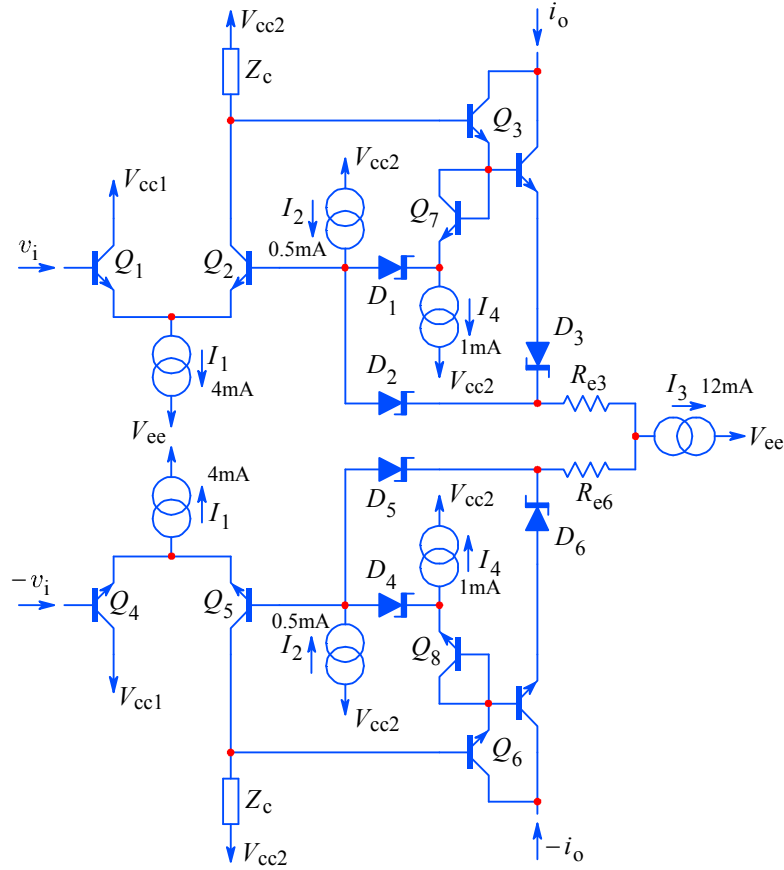


Fig. 5.4.34: The M377 amplifier with the components for high speed overdrive recovery.

The frequency domain and time domain responses of the circuit in [Fig. 5.4.34](#) are shown in [Fig. 5.4.35](#) and [5.4.36](#), respectively. Note that the compensating impedance Z_c was adjusted to the needs of the transistors used for circuit simulation (BF959, as in all previous circuits, thus allowing comparison), therefore the graphs do not represent the true M377 performance capabilities.

Although it can be argued that the simulation has been performed using a simplified basic version of the circuit, there are a few points to note, which are nevertheless valid. First, there is the potential instability problem (owed to feedback), indicated by the phase plot turning upward and the envelope delay going positive above some 4 GHz. If proper care is not taken, especially compensating the parasitics and strays in an IC environment, the step response might display some initial waveform aberrations, or even ringing and oscillations.

Another point of special attention is the parasitic capacitance to the substrate of the Schottky diodes. Being within the feedback loop, these capacitances can be troublesome. Proper forward bias for low impedance is needed to move those unwanted poles (transfer function zeros) well above the cutoff frequency. High bias would result in high temperatures, which, in a densely packed IC such as this one, can be problematic. Also, since noise increases with temperature, the bias can not be as high as one would like.

As an advantage, judging by the constant slope of the input voltage plot, the circuit input impedance is well behaved, thus the loading of a previous stage should not be critical. Likewise, the active inductive peaking, realized by the base resistance

R_b and transformed as inductance at the Q_5 emitter (as shown in Fig. 5.4.32d), offers a simple way of adjusting the frequency compensation network.

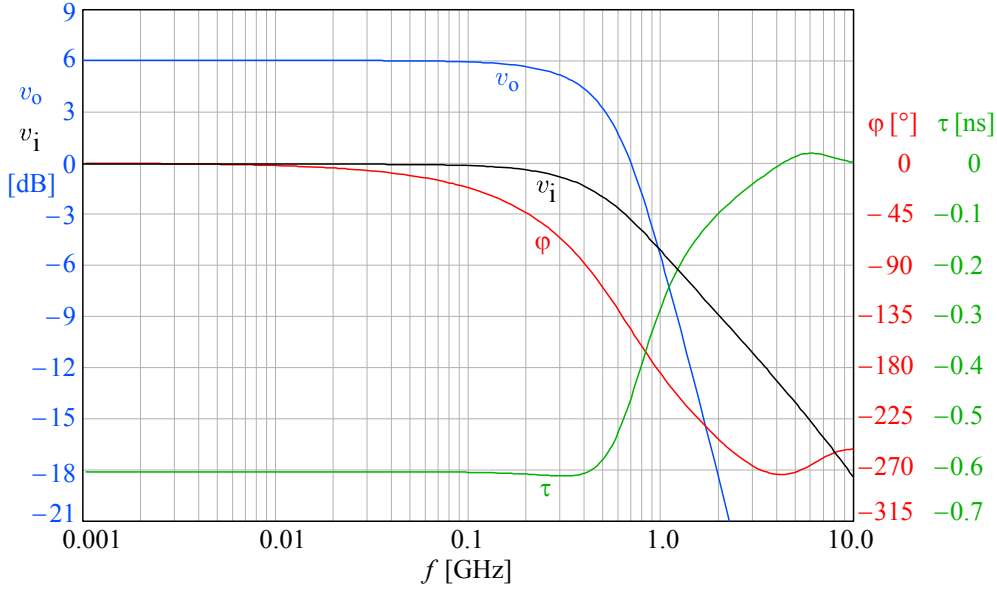


Fig. 5.4.35: Frequency domain performance of the amplifier in Fig. 5.4.34. The bandwidth of the simulated circuit is about 500 MHz, but this is owed to the transistor used (BF959) and the frequency compensation network adjusted in accordance, therefore the graph is not representative of the actual M377 IC performance.

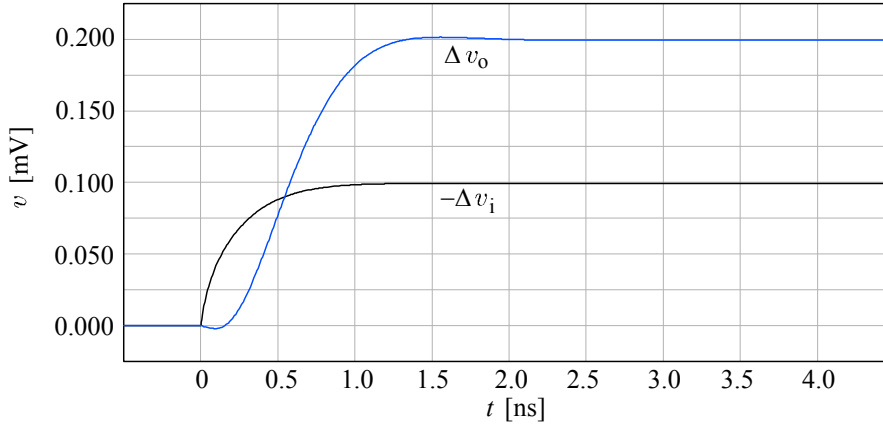


Fig. 5.4.36: Time domain performance of the amplifier in Fig. 5.4.34. The comment in the caption of Fig. 5.4.35 also holds here.

Now take a close look at the circuit in Fig. 5.4.34; in particular, the diode pairs $D_{2,3}$ and $D_{5,6}$, the resistors $R_{e3,6}$ and the current source I_3 ; if another such block is added in parallel (with different values of resistors $R_{e3,6}$), and if the current sources are switched on one at a time, a gain switching in steps can be achieved. The bandwidth would change only slightly with switching. Fig. 5.4.37 shows such a circuit with two gain values, but several more can easily be added.

Another way of changing the vertical sensitivity is to use a fixed gain amplifier and switch the attenuation at its output, as shown in Fig. 5.4.38. In this way the bandwidth is preserved, but attenuation switching has its own weak points, such as a reduced signal range and higher noise at higher attenuation.

As a point of principle, switching the amplifier gain is preferred to fixed gain with switched attenuation. Although it alters the bandwidth, gain switching preserves the signal's dynamic range at all settings, whilst the system with fixed gain and an attenuator will have a comparable dynamic range only with no attenuation; at any other attenuation setting the dynamic range would be proportionally reduced. Also, gain switching systems will preserve the same noise level, whilst the fixed gain systems will have the lowest signal to noise ratio at maximum attenuation.

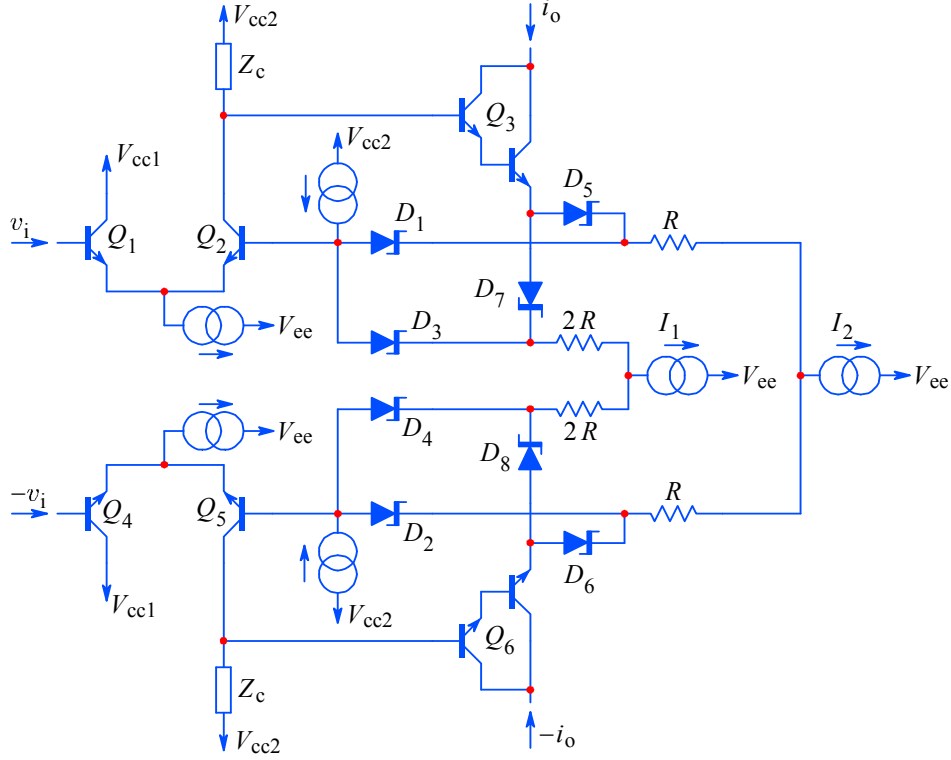


Fig. 5.4.37: Gain switching in steps was achieved by adding one or more emitter current sources (I_1, I_2, \dots) with appropriate resistor values and Schottky diode pairs and switching on one current source at a time.

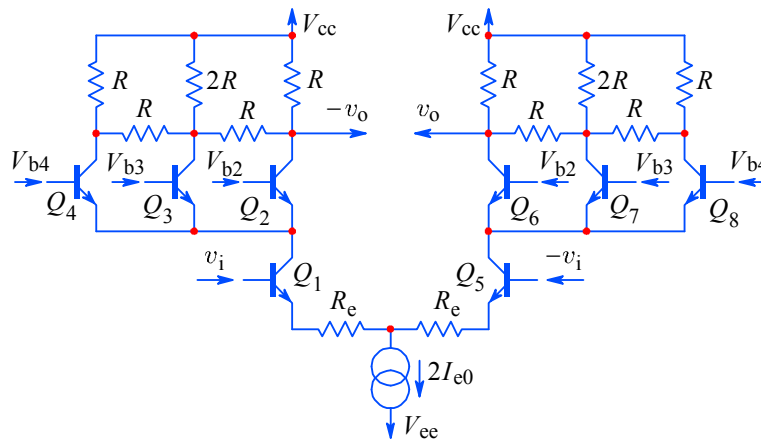


Fig. 5.4.38: With an R - $2R$ network the attenuation can be switched in steps by applying a positive DC voltage $V_{b2,3,4}$ to $Q_{2,6}$, $Q_{3,7}$, $Q_{4,8}$, one pair at a time; at each collector the load is $R \parallel 2R = 2R/3$. With $Q_{2,6}$ on the gain is $A = 2R/3R_e$, with $Q_{3,7}$ on $A = R/3R_e$, and with $Q_{4,8}$ on $A = R/6R_e$, effectively halved by each step. A similar circuit was used in the Tek 2465.

For small gain or attenuation changes of, say, 1:4, as commonly found in oscilloscope amplifiers, these differences can be small. However, in M377, the gain

switching had a 50:1 range. With such a high gain range the frequency compensation needed to be readjusted (for the highest gain no compensation was needed).

5.4.7 The Gilbert Multiplier

A similar problem of compensation readjustment as a function of gain is encountered with a continuously variable gain. Although used only occasionally, a continuously variable gain is a standard feature of almost all oscilloscopes and no manufacturer dares to exclude it, even in digital instruments (although there it is done in very small steps).

In older analog instruments a simple potentiometer was used. This worked well up to some 20 MHz. For higher frequencies the 'pot' size and the variable impedance at the slider represented major difficulties, even if the required gain change was within a relatively small range, ordinarily about 3:1. At Tektronix, an ingenious wire pot was used, having a bifilar winding to cancel the inductance, but its parasitic capacitance, which also varied with the setting, was causing too much cross-talk at higher frequencies. Finally, there was also the mechanical problem of placing the pot at the correct point in the circuit and still being able to bring its axis on the front panel, aligned with the main attenuator switch.

A much more elegant choice is to use some sort of electronic gain control, by using either a voltage or a current controlled amplifier (VCA or ICA). Such an amplifier modulates (ideally) only the signal amplitude, a process which can be mathematically described as multiplication of a signal by a DC voltage; thus we often refer to those amplifiers as multipliers or modulators. Of course, electronic gain control has its own problems and great care is needed to make it linear enough and fast enough, as well as not too noisy. But it solves the problem of mechanical pot placement, since it now has to handle only a DC control signal, so the pot can be placed at any convenient place. In digital systems, the pot is replaced by a digital to analog converter (DAC; in lower speed instruments, a multiplying DAC can be used to attenuate the signal directly, replacing the VCA altogether).

Oscilloscopes do not need to exploit the full modulation range, as RF modulators normally do. In contrast to RF modulators, which are four-quadrant devices (both the carrier and the modulation are AC signals), the gain control in oscilloscopes needs to work in two quadrants only (AC signal and DC control); four quadrants would allow simple gain inversion, but this is more accurately done by a switch. Therefore the modulation cross-talk or the common mode rejection at HF is not an issue. On the other hand, DC stability is important since it directly affects measurement accuracy. Whilst RF modulators operate over a limited frequency range, for oscilloscopes the wideband gain flatness at all gain settings is also very important.

The simple differential amplifier in [Fig. 5.4.8a](#) can perform the variable gain control by varying the emitter current. If we assume that the modulation voltage is $v_M = V_M - V_{be} - V_{ee}$, the modulation current is:

$$2I_e = \frac{v_M}{R_e} \quad (5.4.64)$$

By inserting this into the gain equation of the differential amplifier the multiplication function results:

$$v_o = v_{bb} \frac{2R}{2r_e} = v_{bb} \frac{R I_e}{V_T} = v_{bb} \cdot v_M \frac{R}{2R_e V_T} \quad (5.4.65)$$

Unfortunately the bandwidth is also proportional to the emitter current and with the usual values of R and stray capacitances, the dominant pole at low currents can be very low. In addition the output common mode level also changes with current.

Almost all wideband multipliers are based on one of the variations of the basic circuit, now known as the Gilbert multiplier, after its inventor *Barrie Gilbert* (see [Ref. 5.56–5.62]). The circuit development can be followed from Fig. 5.4.39a by noting that if the output is to be a linear function the input has to be nonlinear.

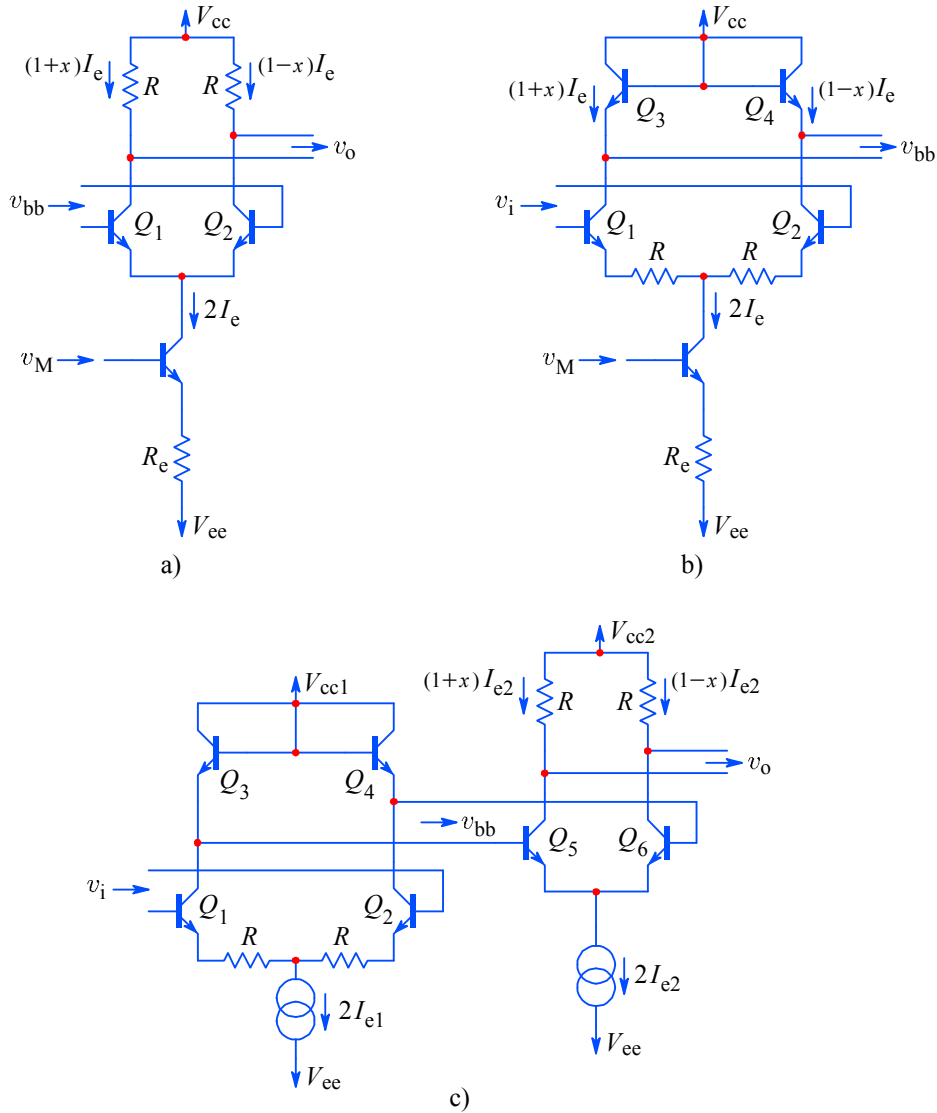


Fig. 5.4.39: The Gilbert multiplier development. **a)** The gain of an ordinary differential amplifier is proportional to the emitter current, but so is the bandwidth. Also, for a linear output a nonlinear input is required. **b)** Inverting the function by linearizing the g_m stage and loading it by simpler p-n junctions gives a nonlinear output, such as required by the circuit in a) to give a linear output. **c)** The simple combination of b) and a) gives the so called 'translinear' stage.

Starting from the exponential I_e – V_{be} relationship:

$$I_e = I_s (e^{V_{be}/V_T} - 1) \quad (5.4.66)$$

and considering that the intrinsic current $I_s \approx 10^{-14}$ A, then even for currents as low as 1 nA we can say that $I_e \gg I_s$; so we are not making a big mistake if we use:

$$I_e \approx I_s e^{V_{be}/V_T} \quad (5.4.67)$$

Since the differential pair was made by the same IC process, we can expect that the devices will be reasonably well matched, so $I_{s1} \approx I_{s2}$ and their temperature dependence will also be well matched if the transistors are at the same temperature:

$$\frac{I_{e1}}{I_{e2}} = \frac{I_{s1}}{I_{s2}} e^{(V_{be1}-V_{be2})/V_T} \approx e^{(V_{be1}-V_{be2})/V_T} \quad (5.4.68)$$

In order to achieve a linear output current the expected input voltage should follow the logarithmic function:

$$v_{bb} = \Delta V_{be} = V_T \ln \frac{I_{e1}}{I_{e2}} \quad (5.4.69)$$

If i_e is the modulation component superimposed on the DC current I_{e0} , we can write:

$$I_{e1} = I_{e0} + i_e \quad \text{and} \quad I_{e2} = I_{e0} - i_e \quad (5.4.70)$$

Let x be the AC to DC component ratio:

$$x = \frac{i_e}{I_{e0}} \quad (5.4.71)$$

Then the currents are:

$$I_{e1} = I_{e0}(1 + x) \quad \text{and} \quad I_{e2} = I_{e0}(1 - x) \quad (5.4.72)$$

and, returning to [Eq. 5.4.69](#), we obtain the required nonlinear input function:

$$v_{bb} = V_T \ln \frac{I_{e0}(1 + x)}{I_{e0}(1 - x)} = V_T \ln \frac{1 + x}{1 - x} \quad (5.4.73)$$

How can this nonlinear relationship at the differential amplifier's input be realized? By inverting the [Fig. 5.4.39a](#) circuit's function, such as in [Fig. 5.4.39b](#); that is, by using relatively large emitter degeneration resistors, resulting in a linear voltage to current conversion and loading the collectors by similar p-n junctions as V_{be} , would produce exactly such a nonlinear relationship as the original circuit needs at its input to produce a linear output.

[Fig. 5.4.39c](#) is thus a simple combination of the [b](#) and [a](#) circuits, but with some very interesting properties. First, it is compensated and thus quite linear within the entire input range ($-1 \leq x \leq +1$). Also the v_{bb} swing is small (less than V_T) owing to the very low impedance ($\approx V_T/I_e$) of $Q_{3,4}$, which means that charging and discharging of stray capacitances is minimal, so the bandwidth limit is owed to the collector impedances of $Q_{5,6}$ and the transistors' f_T . Finally, the gain is entirely 'current mode', with low temperature dependence (both V_T and I_s are canceled in the expression for v_o) and the gain control is set by the current ratio I_{e2}/I_{e1} .

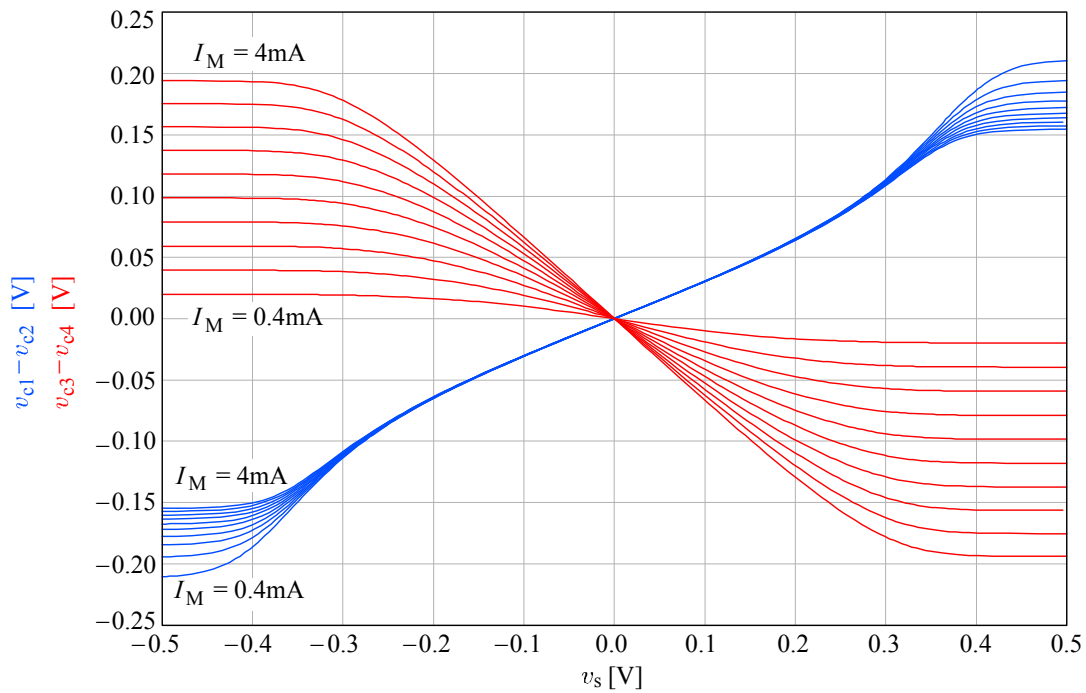


Fig. 5.4.40: DC transfer function of the Gilbert multiplier of Fig. 5.4.39, for $2I_{e1} = 4 \text{ mA}$ and modulation current $2I_{e2} = I_M = 0.4\text{--}4 \text{ mA}$. The signal source (v_s) range is $\pm 0.5 \text{ V}$.

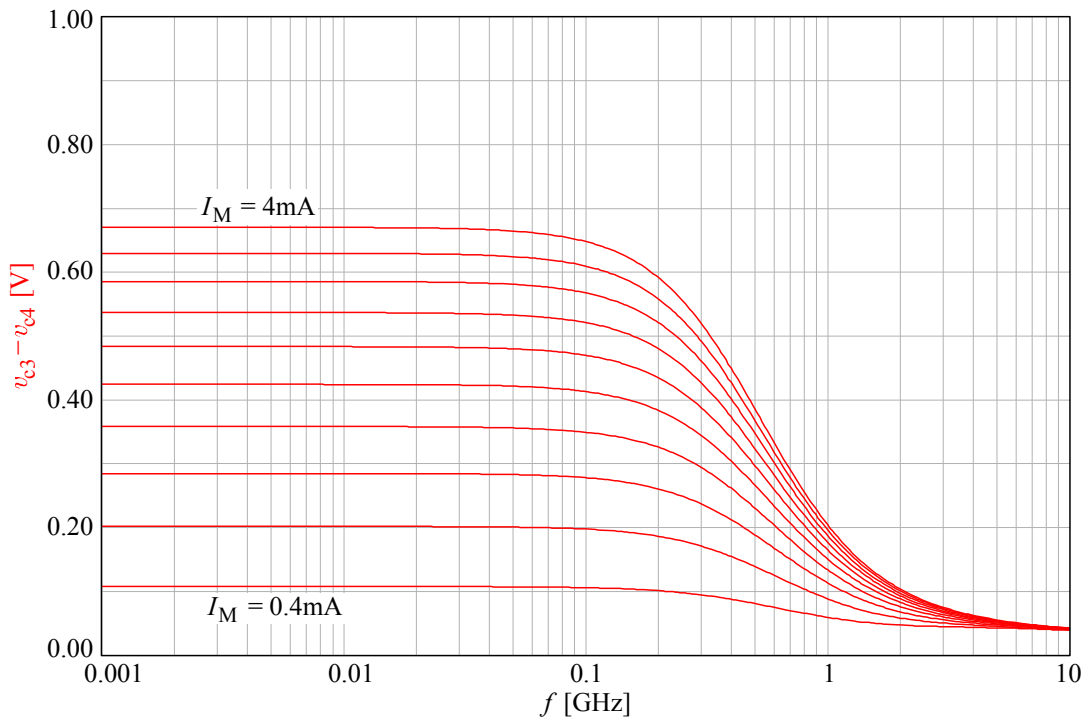


Fig. 5.4.41: The Gilbert multiplier bandwidth is almost constant over the 10:1 modulation current range.

Another way of developing this ‘translinear gain cell’ can be followed by observing Fig. 5.4.42. Two current mirrors, with the current gain proportional to the emitter area A , can be interconnected by breaking the two emitters carrying the mirrored currents and biasing them by a current source, thus forming a differential

amplifier, whose input nonlinearity is compensated by the nonlinearity of the remaining two transistors.

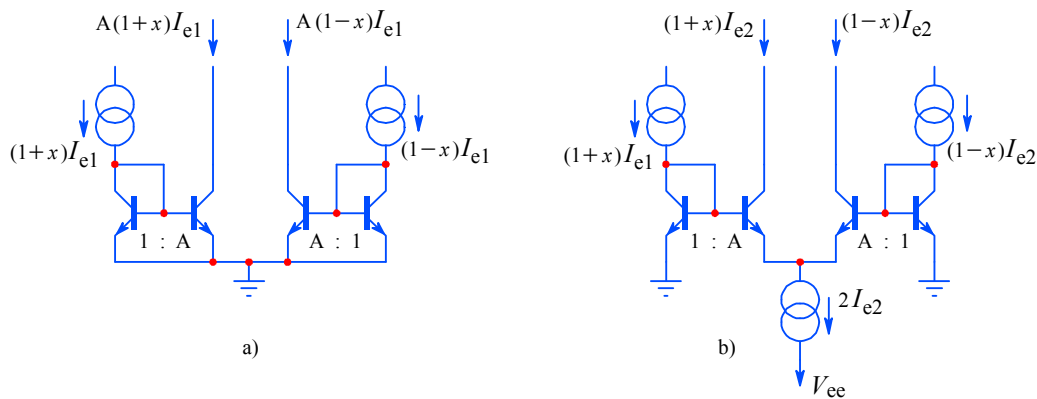


Fig. 5.4.42: Another way of developing the Gilbert multiplier is by interconnecting two current mirrors into a differential amplifier, whose input nonlinearity is compensated by the nonlinearity of the two grounded transistors.

Once the basic multiplier is developed it is relatively easy to construct a four quadrant multiplier, [Fig. 5.4.43](#), by adding another differential pair with inputs in parallel and the collectors cross-coupled.

A further differential amplifier can be used to perform the voltage to current conversion and splitting the current source ($4I_e$) into the required signal currents. The gain is now controlled by biasing the compensation transistors with current.

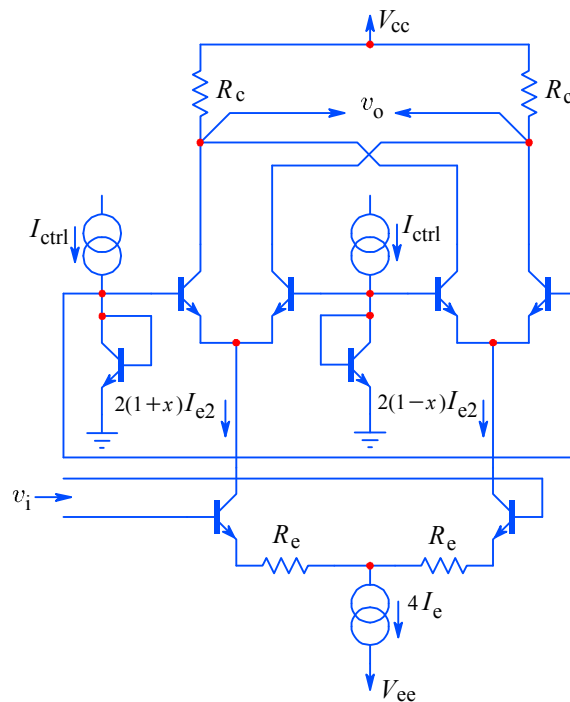


Fig. 5.4.43: A four quadrant multiplier developed from the previous circuit. The gain is controlled by current biasing the compensation transistors. Four quadrant operation is obtained from the fact that the cross-coupled collectors cancel out if the two tail currents are equal and the third differential amplifier allows it to distribute the tail currents in a symmetrical manner about the mid-bias value. Thus both the input and the control can be AC signals and can also be mutually exchanged, not compromising the bandwidth or the linearity.

By cross-coupling the collectors of the two differential pairs we have achieved effective output current cancellation if the two control currents are equal and the two emitter currents are equal. Varying any pair of currents about this mid-point changes the polarity as well as the gain of the multiplier for the other input. Thus, a nice byproduct of four-quadrant configuration is that the ‘signal’ and the ‘control’ inputs can be mutually exchanged, without compromising the bandwidth or the linearity.

Returning to the M377 design, where only a two-quadrant multiplication is needed, the multiplier in [Fig. 5.4.44](#) represents the circuit used. It is almost identical to the four-quadrant multiplier, except that the collectors are not cross-coupled and the output is taken from a single pair. The differential circuit symmetry was retained merely because of its good thermal balance and DC stability. For the same reason, all four collectors must be equally loaded.

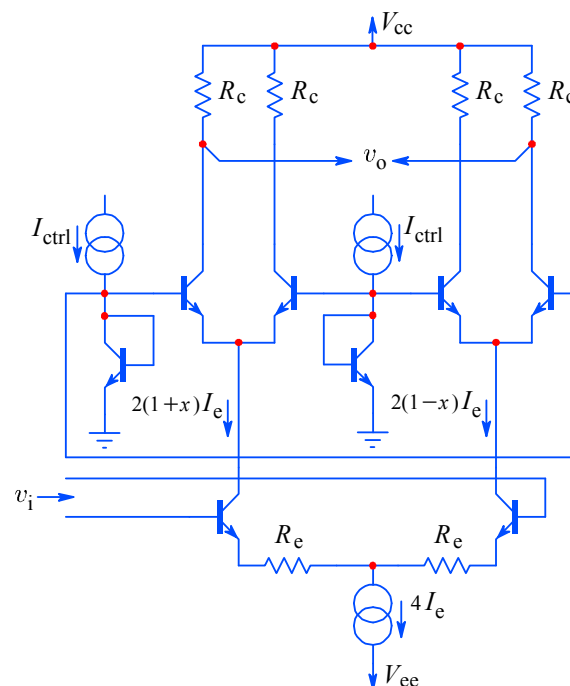


Fig. 5.4.44: Two quadrant multiplication is sufficient for the oscilloscope continuously variable gain control; however, the same differential symmetry of the four quadrant multiplier has been retained for the M377 because of good thermal balance and DC stability.

The multiplier circuits shown represent the basic linearization principle. In actual implementations, a number of additional linearization techniques are used, most of them also patented by Barrie Gilbert while he was at Tektronix¹, and some later when at Analog Devices, [[Ref. 5.56–5.61](#)].

¹ It might be interesting to note that Barrie Gilbert published an article [[Ref. 5.56](#)] describing his multiplier before Tektronix applied for a patent. Motorola quickly seized the opportunity and started producing it (as the MC1495). Tektronix claimed the primarity and Motorola admitted it, but nevertheless continued the production, since once published the circuit was ‘public domain’. Barrie’s misfortune gave the opportunity to many generations of electronics enthusiasts (including the authors of this book) to play with this little jewel and use it in many interesting applications. Thanks, Barrie!

Résumé of Part 5

In this part we have briefly analyzed some of the most important aspects of system integration and system level performance optimization, with a special emphasis on system bandwidth.

We have described the transient response optimization by a pole assignment process called ‘geometrical synthesis’ and showed how it can be applied using inductive peaking. We have discussed the problems of input signal conditioning, the linearization and error reduction and correction techniques, employing the feedback and feedforward topologies at either the system level or at the local, subsystem level. We have also revealed and compared some aspects of designing wideband amplifiers using discrete components and modern IC technology.

On the other hand, we have said very little about other important topics in wideband instrumentation design, such as adequate power supply decoupling, grounding and shielding, signal and supply path impedance control by strip line and microstrip transmission line techniques, noise analysis and low noise design, and the parasitic impedances of passive components. But we believe that those subjects are extensively covered in the literature, some of it also cited in the references, so we have tried to concentrate on the bandwidth and transient performance issues.

We have also said nothing about high sampling rate analog to digital conversion techniques, now already established as the essential ingredient of modern instrumentation. While there are many books discussing AD conversion, most of them are limited to descriptions of applying a particular AD converter, or, at most, to compare the merits of one conversion method against others. Only a few of them discuss ADC circuit design in detail, and even fewer the problems of achieving top sampling rates for a given resolution, either by an equivalent time or a real time sampling process, time interleaving of multiple converters, combining analog and digital signal processing and other techniques, which today (first decade of the XXI century) allow the best systems to reach sampling rates of up to 20 GSps (Giga Samples per second) and bandwidths of up to 6 GHz.

Just like many other books, this one, too, ends just as it has become most interesting (the reader might ask himself whether there is really nothing more to say or whether the authors simply ran out of ideas? — since most of the circuits presented are not of our origin, and electronics certainly is an art of infinite variations, we the authors can be, one hopes, spared the blame). As already said in the Foreword, the most difficult thing when writing about an interesting subject is not what to include, but what to leave out. Although we discuss the effects of signal sampling in [Part 6](#) and a few aspects of efficient system design combining analog and digital technology in [Part 7](#), this book is about amplifier design, so we leave the fast ADC circuit design discussion for another opportunity.

(blank page)

References:

- [5.1] *D.L. Feucht*, Handbook of Analog Circuit Design, Academic Press, Inc. San Diego, 1990
See also the latest compact disk version at <<http://www.innovatia.com/>>
- [5.2] *S. Roach*, Signal Conditioning in Oscilloscopes and the Spirit of Invention, (*J. Williams*, [Editor], The Art and Science of Analog Circuit Design, Part 2) Butterworth–Heinemann, Boston, 1995
- [5.3] *P.R. Gray & R.G. Meyer*, Analysis and Design of Analog Integrated Circuits, John Wiley, New York, 1969
- [5.4] *J. Williams*, Composite Amplifiers, Linear Technology Application Note AN-21, July 1986
- [5.5] *P. Starič*, Wideband JFET Source Follower, Electronic Engineering, August 1992
- [5.6] *J. Williams*, Measuring 16-bit Settling Times: the Art of Timely Accuracy, EDN Magazine, Nov. 19, 1998, <<http://www.ednmag.com/>>
- [5.7] *W. Kester*, High Speed Design Techniques, Analog Devices, 1996, <<http://www.analog.com/>>
- [5.8] *A.D. Evans* (Editor), Designing with Field Effect Transistors, Siliconix, McGraw-Hill, 1981
- [5.9] *J.E. Lilienfeld*, Method and Apparatus for Controlling Electric Current, US Patent 1 745 175, Jan. 28, 1930
(the FET patent, apparently predating the BJT patent by about 23 years)
- [5.10] FET Design Catalog, Siliconix, 1982, <<http://www.siliconix.com/>>
- [5.11] IC Applications Handbook, Burr-Brown, 1994, <<http://www.burr-brown.com/>>, <<http://www.ti.com/>>
- [5.12] A New Approach to OpAmp Design, Application Note AN 300-1, Comlinear Corporation, March, 1985
- [5.13] *S. Franco*, Design with Operational Amplifiers and Analog ICs, McGraw-Hill, 1988
- [5.14] *F.A. Muller*, High Frequency Compensation of RC Amplifiers, Proceedings of the I.R.E., August, 1954, pp. 1271–1276.
- [5.15] *B. Orwiller*, Vertical Amplifier Circuits, Tektronix, Inc., Beaverton, Oregon, 1969.
- [5.16] *F.W. Grover*, Inductance Calculation, (Reprint) Instrument Society of America, Research Triangle Park, N. C. 27 709, 1973.
- [5.17] *J. Williams*, High Speed Amplifier Techniques, Application Note AN-47, Linear Technology, March, 1985, <<http://www.lt.com/>>
- [5.18] *C.R. Battjes*, Monolithic Wideband Amplifier, US Patent 4 236 119, Nov. 25, 1980²

² Note : For US Patents go to <<http://www.uspto.com/>> and type the patent number in the Search pad. Patent figures are in TIFF graphics format, so a TIFF Viewer software is recommended (links for downloading and installation are provided within the USPTO web pages).

- [5.19] *J.L. Addis, P.A. Quinn*, Broadband DC Level Shift Circuit With Feedback, US Patent 4 725 790, Feb. 16, 1988
- [5.20] *J.L. Addis*, Precision Differential Amplifier Having Fast Overdrive Recovery, US Patent 4 714 896, Dec. 22, 1987
- [5.21] *J.L. Addis*, Buffer Amplifier, US Patent 4 390 852, Jun. 28, 1983
- [5.22] *J.L. Addis*, Feedbeside Correction Circuit for an Amplifier, US Patent 4 132 958, Jan. 2, 1979
- [5.23] *T.C. Hill, III*, Differential Amplifier with Dynamic Thermal Balancing, US Patent 4 528 516, July 9, 1985
- [5.24] *J. Woo*, Wideband Amplifier with Active High Frequency Compensation, US Patent 4 703 285, Oct. 27, 1987
- [5.25] *Wakimoto, Tsutomu, Azakawa, Yukio*, Wideband Amplifier, US Patent 4 885 548, Dec. 5, 1989
- [5.26] *H. Weber*, A Method of Predicting Thermal Stability, Motorola, Application Note AN-128
- [5.27] *R. Ivins*, Measurement of Thermal Properties of Semiconductor Devices, Motorola, Application Note AN-226
- [5.28] *B. Botos*, Nanosecond Pulse Handling Techniques in IC Interconnections, Motorola, Application Note AN-270
- [5.29] Field Effect Transistors In Theory and Practice, Motorola, Application Note AN-211A
- [5.30] *N. Freyling*, FET Differential Amplifier, Motorola, Application Note AN-231
- [5.31] *R.W. Anderson*, s-Parameter Techniques for Faster, More Accurate Network Designs, Hewlett-Packard, Application Note AN-95-1
- [5.32] *R. Gosser*, Wideband Transconductance Generator (Quad-Core Amplifier), US Patent 5 150 074, Sep. 22, 1992
- [5.33] New Complementary Bipolar Process, National Semiconductor, Application Note AN-860
- [5.34] *J. Bales*, A Low Power, High Speed, Current Feedback OpAmp with a Novel Class AB High Current Output Stage, IEEE Journal of Solid-State Circuits, Vol. 32, No. 9, Sept. 1997
- [5.35] Development of an Extensive SPICE Macromodel for Current Feedback Amplifiers, National Semiconductor Application Note AN-840, July, 1992
- [5.36] Topics on Using the LM6181 – A New Current Feedback Amplifiers, National Semiconductor Application Note AN-813, March, 1992
- [5.37] *T.T. Regan*, Designing with a New Super Fast Dual Norton Amplifier, National Semiconductor Application Note AN-278, Sept., 1981
- [5.38] Simulation SPICE Models for Comlinear OpAmps, National Semiconductor Application Note OA-18, May, 2000

- [5.39] *H.S. Black*, Translating System,
U.S. Patent 1 686 792, Oct. 9, 1928
- [5.40] *H.S. Black*, Wave Translation System,
U.S. Patent 2 102 671, Dec. 21, 1937
- [5.41] *H.S. Black*, Inventing the Negative Feedback Amplifier,
IEEE Spectrum, vol. 14, pp. 55–60, Dec., 1977
- [5.42] *P.J. Walker*, Current Dumping Audio Amplifier,
Wireless World, vol. 81, pp. 560–562, Dec., 1975
- [5.43] *P.J. Walker, M.P. Albinson*, Distortion-Free Amplifiers,
U.S. Patent 3 970 953, Jul. 20, 1976
- [5.44] *J. Vanderkooy, S.P. Lipshitz*, Feedforward Error Correction in Power Amplifiers,
Journal of the Audio Engineering Society, vol. 28, No. 1/2, pp. 2–16, Jan./Feb., 1980
- [5.45] *M.J. Hawksford*, Distortion Correction in Audio Power Amplifiers,
Journal of the Audio Engineering Society, vol. 29, No. 1/2, pp. 27–30, Jan./Feb., 1981
- [5.46] *M.J. Hawksford*, Distortion Correction Circuits for Audio Amplifiers,
Journal of the Audio Engineering Society, vol. 29, No. 7/8, pp. 503–510, Jul./Aug., 1981
- [5.47] *M.J. Hawksford*, Low Distortion Programmable Gain Cell Using Current Steering Cascode Topology, Journal of the Audio Engineering Society, vol. 30, No. 11, pp. 795–799, Nov., 1982
- [5.48] *S. Takahashi, S. Tanaka*, Design and Construction of a Feedforward Error Correction Amplifier, Journal of the Audio Engineering Society, vol. 29, pp. 31–37, Jan./Feb., 1981
- [5.49] *A.M. Sandman*, Reducing amplifier distortion by error add-on,
Wireless World, vol. 79, pp. 32, Jan., 1973
- [5.50] *J.L. Addis*, Versatile Analogue Chip for Oscilloscope Plug-Ins,
Electronic Engineering, Aug., 1988 (Part I), Sept., 1988 (Part II)
- [5.51] *J.L. Addis*, Good Engineering and Fast Vertical Amplifiers,
(*J. Williams*, [Editor], Analog Circuit Design, Art, Science and Personalities, Part 1)
Butterworth-Heinemann, Boston, 1991
- [5.52] *C. R. Battjes*, Who Wakes the Buglar?,
(*J. Williams*, [Editor], The Art and Science of Analog Circuit Design, Part 2)
Butterworth-Heinemann, Boston, 1995
- [5.53] *P.A. Quinn*, Feed-Forward Amplifier,
US Patent No. 4 146 844, Mar. 27, 1979
- [5.54] *P.A. Quinn*, Feed-Forward Amplifier,
US Patent No. 4 146 844 (Reissue 31 545), Mar. 27, 1984
- [5.55] *P.A. Quinn*, Differential Impedance Neutralization Circuit,
US Patent No. 4 692 712, Sep. 8, 1987
- [5.56] *B. Gilbert*, A Precise Four-quadrant Multiplier with Subnanosecond Response,
IEEE Journal of Solid-State Circuits, pp. 365–373, Dec. 1968

- [5.57] *B. Gilbert*, Multiplier Circuit,
US Patent No. 4 156 283, May 22, 1979
- [5.58] *B. Gilbert*, High accuracy four–quadrant multiplier also capable of four–quadrant division,
US Patent No. 4 586 155, April 29, 1986
- [5.59] *B. Gilbert*, Synchronous logarithmic amplifier,
US Patent No. 5 298 811, March 29, 1994
- [5.60] *B. Gilbert*, Single supply analog multiplier,
US Patent No. 6 074 082, June 13, 2000
- [5.61] *B. Gilbert*, Circuit having dual feedback multipliers,
US Patent No. 6 456 142, September 24, 2002
- [5.62] *B. Gilbert*, Where Do Little Circuits Come From?,
(*J. Williams* [Editor], Analog Circuit Design, Part 1)
Butterworth-Heinemann, Boston, 1991
- [5.63] *P. Gamand, C. Caux*, Semiconductor device comprising a broadband and high gain
monolithic integrated circuit for a distributed amplifier,
US Patent No. 5,386,130, January 31, 1995
- [5.64] *K. Nakahara, Y. Sasaki*, Distributed amplifier and bidirectional amplifier,
US Patent No. 5,414,387, May 9, 1995
- [5.65] *R.W. Chick*, Non-uniformly distributed power amplifier,
US Patent No. 5,485,118, January 16, 1996
- [5.66] *C.D. Motchenbacher, F.C. Fitchen*, Low Noise Electronic Design,
John Wiley, New York, 1973, ISBN 0-471-61950-7
- [5.67] *H.W. Bode*, Network Analysis And Feedback Amplifier Design,
Van Nostrand, New York, 1949
- [5.68] *E. Margan*, RC Attenuator Distortion,
Electronics World + Wireless World, Sept., 1992, Circuit Ideas, pp. 765
- [5.69] *E. Margan*, Amplifier Instability,
Electronics World + Wireless World, Apr., 1998, pp. 311–312