Erik Margan

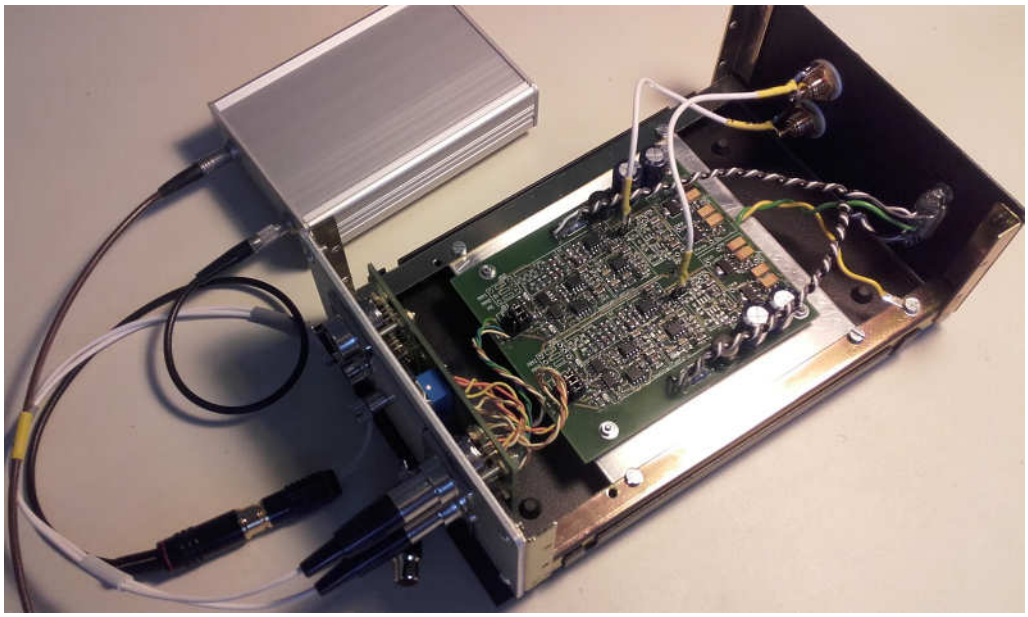
Ironing RIAA:

Der Wohltemperierte RIAA Verstärker

(A Well Tempered RIAA Amplifier)

(40 years too late, but finally correct)

Ljubljana, August 2013



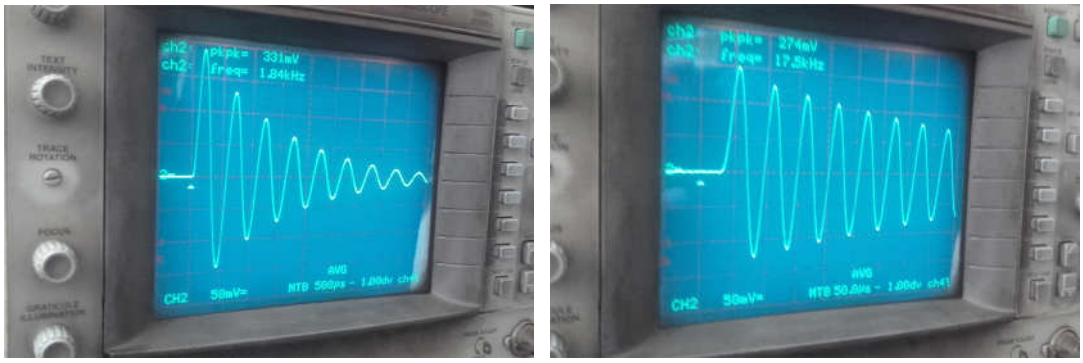
Author's circuit realization: a pair of miniXLR connectors at the input, with DIL switches setting the input impedance, another miniXLR pair for differential output, an RCA chinch pair at the rear for single-ended output, and a 9-pin sub-D connector for external DC power. The small box houses the inverse passive circuit for testing. An integral design feature is the complete lack of IP address for IoT connectivity, making the system virtually unhackable. For other details of the circuit design and performance estimation please see the text.

Abstract

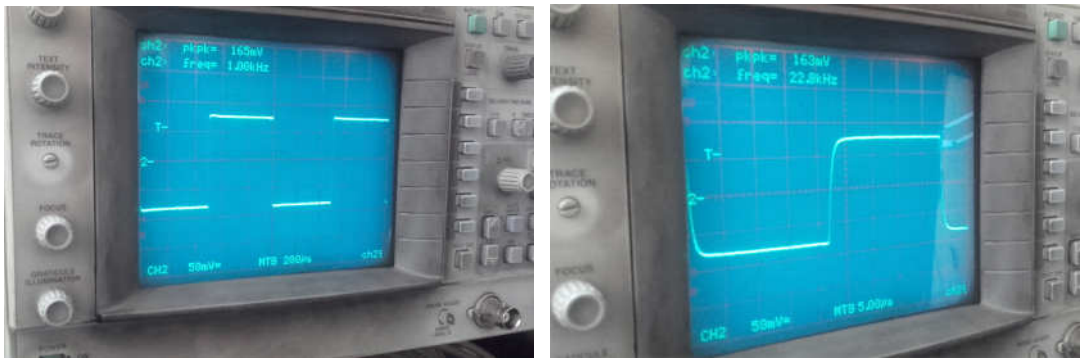
Described is a complete design procedure of a vinyl record playback amplifier equalizer with perfect conformance to the RIAA standard, with some additional features and possible circuit variations for obtaining lower noise and lower distortion, as well as cartridge loading impedance adjustment, making it adaptable for various cartridge types.

Explained also is the source and the cause of the most common design error often made when calculating the equalization component values and a calculation procedure is shown which results in a perfectly flat equalized frequency response. An easy way of checking the actual circuit performance is also demonstrated.

Performance Illustrations



The startup of an exponentially decaying sine wave at 2 kHz (left) and 18 kHz (right).



Square wave response at 1 kHz (left) and 23 kHz (right).

The oscillograms illustrate the performance of the actual RIAA equalization circuit as described in the article, driven from an inverse RIAA encoding passive circuit (as described in Appendix 1). In these tests the equalizer circuit was used in a ground referenced input signal mode, however it is also capable of accepting a signal from a floating source.

The exponentially decaying sine wave with a frequency of 2 kHz and 18 kHz with a 5 Hz repetition rate indicates both a clean transient start and a symmetrical exponential decay, testifying to the excellent stability of the DC correction integration loop. The square wave of 1 kHz and 23 kHz testify also of clean high frequency signal handling, as well as perfect equalization matching. With a suitable high quality cartridge the system bandwidth of up to 60 kHz is achievable, with noise level being down to 81 dB below the nominal 5 mV (at 1 kHz) input signal level.

Motivation

The temptation to give such an odd title to the article was simply irresistible.

On one hand, I used to be *Switched on Bach* [1] for quite a while. On the other hand, from about 1970 onward, when I began fiddling more seriously with Hi-Fi, I have seen so many mistuned RIAA playback equalization networks, both in literature and in marketed products, that the idea came up almost naturally. Regardless of whether the circuits had passive, active, or hybrid equalization, most of them were "tuned" by simply assigning most convenient RC values to get them numerically close to the required standard time constants, without bothering much about the resulting frequency response, which was often far off the usually tolerable ±0.5 dB band. Also, those rare circuits that were better and flatter in the audio mid band were often peaking or sagging at either the low or the high frequency extreme, most often both. And like it were not enough, many circuits exhibited excessive noise, low dynamic headroom, compromised bandwidth and slew rate, improper pickup loading, and even influence of poor feedback factor on the amplifier's input impedance.

In 1974 I got an HP-29C, a pocket calculator with 98 registers (!) of continuous memory. One of the first tasks I programmed was a routine for the optimization of the RIAA correction network to standard component values. At the time, precision low tolerance components were not as common as they are today, so it was necessary to combine measured components in parallel to obtain the required time constants (a technique still valid today). However, my first circuit built upon that optimization (using discrete transistors) was not accurate enough for reasons explained later in the text, and was also too noisy. My next circuit using integrated operational amplifiers was only marginally better. Only my third circuit, built in 1978, was worth the 'equalizer' title.

With the introduction of digital music media, my interest in audio slowly faded, although I did return to power amplifier design on a few rare occasions. Recently however, I was surprised to learn that some general interest in old analogue techniques is still relatively high. One might think that today there is not much to be said on the RIAA subject after all the work done by people like Peter J. Baxandall [2], John Linsley Hood [3], Stanley Lipshitz [4], and numerous others. But as a recent quick flyby over some audio web pages revealed, there are still many misconceptions and prejudices. Another surprise was a discovery that there are still many people willing to build such things by themselves, motivated essentially by the desire to learn, in spite of the cost of such an endavour being many times higher than buying a finished product. So here I present this text for all analogue enthusiasts, young and seasoned, and I hope some might find it entertaining to read and possibly useful to build.

The Reference

The RIAA directive (actually a silent agreement between several record manufacturers, implemented between 1954-1958, after a number of similar previous proposals) brought standardization to the conversion of the velocity encoded amplitude of the microgroove record cutter head into an equalized flat spectrum upon reproduction. The velocity encoding was chosen as a convenient technique to maximize the groove density on vinyl records in order to achieve 'long playing' time, hence the LP acronym for vinyl media (some interesting history can be found here [5]).

It is often stated in literature that the standard defines some particular 'corner frequencies' or '3 dB frequencies' to comply with the inverse of the cutter system response. That is not quite correct, as will be shown. In fact, the encoding standard only declares the following **three time constants**:

$$\begin{aligned}
\tau_1 &= 3183\,\mu\text{s} \ldots\ldots\ldots \text{response zero at } s_1 = -\tfrac{1}{\tau_1} = -\omega_1 \approx -2\pi 50\,\text{rad/s} \\
\tau_2 &= 318.3\,\mu\text{s} \ldots\ldots\ldots \text{response pole at } s_2 = -\tfrac{1}{\tau_2} = -\omega_2 \approx -2\pi 500\,\text{rad/s} \\
\tau_3 &= 75\,\mu\text{s} \ldots\ldots\ldots\ldots \text{response zero at } s_3 = -\tfrac{1}{\tau_3} = -\omega_3 \approx -2\pi 2122\,\text{rad/s}
\end{aligned}$$

Of course, the equalization network must have the poles in place of the zeros, and a zero in place of the pole. The approximate angular frequencies are stated here only for convenience, and it is important to note that those equivalent frequencies (50 Hz, 500 Hz, 2122 Hz) are **not** the frequencies at which the response deviates by 3 dB from the asymptotic response. That would be true only in a circuit with a single time constant, but not for the combination having the time constants relatively close to each other. This will become evident by calculating the transfer function magnitude within the frequency domain of interest, and at those particular frequencies.

The cutter system model equation based on the three time constants can be written in the complex Laplace space in the polynomial form:

$$F_3(s) = \frac{(s - s_1)(s - s_3)}{(s - s_2)} \tag{1}$$

where s denotes the complex frequency; see Appendix 1 for a complete circuit analysis. Replacing the zeros and the pole by their associated time constants, $s_i = -1/\tau_i$, it is possible to write:

$$F_3(s) = \frac{\tau_2}{\tau_1 \tau_3} \cdot \frac{(s\tau_1 + 1)(s\tau_3 + 1)}{(s\tau_2 + 1)} \tag{2}$$

It must be realized that the function defined like this has a response which diverges to infinity with increasing frequency. This implies infinite energy, and such a system is physically impossible to realize, because any real system will eventually encounter a bandwidth limit. In the actual implementation of the encoding function this limit is usually set at 50 kHz ($\tau_4 = 3.183\,\mu\text{s}$). Also, the disk cutting system has an implicit second pole at 50 kHz imposed by the cutter mechanics. Neither of these has ever been included in the official standard, even if being physically unavoidable. Likewise, the driving amplifier bandwidth, though much higher, at around 400 kHz ($\tau_5 \approx 0.4\,\mu\text{s}$), was also disregarded. Moreover in most cutting systems a 2^{nd}-order Butterworth low pass filter at 50 kHz is employed to prevent any possibility of high frequency overdrive and consequent adjacent groove contact or overlap.

For these reasons the development of the correct equalization must include at least one additional pole, and we are going to show that by including τ_4 into the reference inversion response makes the designer's life much easier. The new encoding equation with four time constants is:

$$F_4(s) = F_3(s)\frac{\tau_4}{s\tau_4 + 1} = \frac{\tau_2\tau_4}{\tau_1\tau_3} \cdot \frac{(s\tau_1 + 1)(s\tau_3 + 1)}{(s\tau_2 + 1)(s\tau_4 + 1)} \tag{3}$$

Similarly a function with five time constants can be written as:

$$F_5(s) = F_4(s)\frac{\tau_5}{s\tau_5 + 1} = \frac{\tau_2\tau_4\tau_5}{\tau_1\tau_3} \cdot \frac{(s\tau_1 + 1)(s\tau_3 + 1)}{(s\tau_2 + 1)(s\tau_4 + 1)(s\tau_5 + 1)} \tag{4}$$

Because of these additional poles, the response at high frequencies does not rise indefinitely, as is shown in Fig.1. This is important because most equalization amplifier configurations also deviate in this region, and so do the playback heads with their mechanical (needle, cantilever, magnet) and electrical (coil and its load) resonances (usually between 10–30 kHz), and those also need to be taken into consideration.

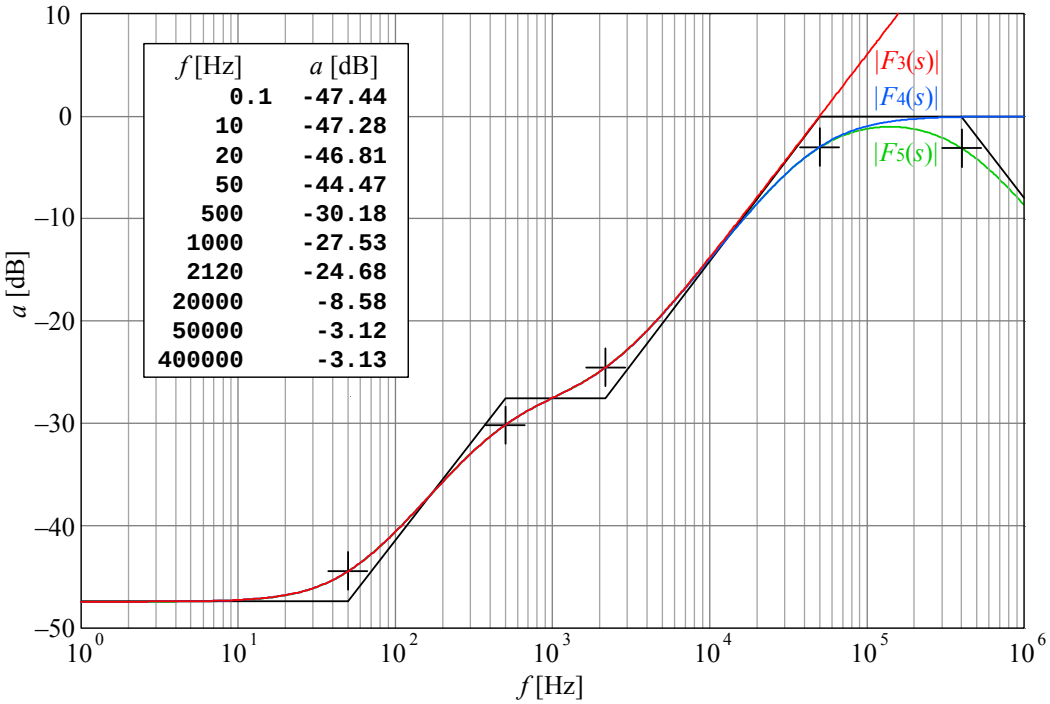


Fig.1: Absolute value of the transfer functions (magnitudes) normalized to the same value at DC. The tabulated numerical values are given for $|F_5(s)|$. At 20 kHz $|F_3(s)|$ is about 1 dB higher, and crosses 0 dB at 50 kHz. $|F_4(s)|$ approaches 0 dB at 1 MHz.

Because of the particular circuit topology of the equalization network, we shall base our discussion on $F_4(s)$, equation (3), using these four time constants:

$$\begin{aligned} \tau_1 &= 3183\,\mu s \\ \tau_2 &= 318.3\,\mu s \\ \tau_3 &= 75\,\mu s \\ \tau_4 &= 3.183\,\mu s \end{aligned} \tag{5}$$

We shall include other bandwidth limitations at a later stage, after the basic circuit response has been correctly established.

When plotting any complex function we usually calculate separately its magnitude and phase angle as functions of frequency (Bode plot, [8]). The phase is the arctangent of the imaginary to real part ratio, $\varphi = \arctan(\Im\{F(s)\}/\Re\{F(s)\})$, but we are not particularly interested in this here. We are more concerned with the magnitude (absolute value), which is the square root of the product of the function with its own complex conjugate (to denote this explicitly we set $s = j\omega = j2\pi f$):

$$|F(j\omega)| = \sqrt{F(j\omega) \cdot F(-j\omega)} \tag{6}$$

Note that equations (2), (3) and (4) have different multiplication factors owed to the different number of time constants, which effectively means different attenuations, so in order to make a fair comparison we shall normalize all three responses to the DC level ($s = 0$) of equation (4). This means that we shall use $\tau_2\tau_4/\tau_1\tau_3$ for all three expressions. Fig.1 shows $|F_3(s)|$, $|F_4(s)|$, and $|F_5(s)|$, along with the appropriate asymptotes (black line), with crosses marking the frequencies corresponding to the defined time constants.

From the tabulated values in Fig.1 we can see that the response at 50 Hz is about 3 dB higher than at 0.1 Hz, as expected. But at 500 Hz it is about 2.65 dB below the value at 1 kHz (not 3 dB, as is often assumed), and similarly at 2120 Hz it is about 2.85 dB higher than the value at 1 kHz. Then, at 20 kHz the value of $|F_4(s)|$ and $|F_5(s)|$ is only about 19 dB above the value at 1 kHz, instead of 20 dB as is often assumed. This simply means that using the 3 dB break points for calculating the component values of the equalization network will not yield a flat response.

When plotting the responses, and to be able to see their differences easier, we shall normalize the responses of both the inverse and the equalized RIAA network to their value at 1 kHz, and we can deal with the actual attenuation and the required gain later. The normalization is done simply by dividing the frequency response of a particular function by its value at $f_1 = 1$ kHz:

$$F_{i\mathrm{N}} = \frac{F_i(j2\pi f)}{F_i(j2\pi f_1)} \tag{7}$$

On the dB scale of Fig.1 this simply means shifting the response up by 27.53 dB. We shall append the index 'N' to a function normalized in this way. To avoid any possible confusion, we shall denote an inverse RIAA function as $F(s)$, whilst for the equalization network we shall write $G(s)$. So, for example, $F_{4\mathrm{N}}(s)$ will mean the inverse RIAA function with 4 time constants, with gain normalized to 1 kHz.

Important consideration: As we shall see soon, the response of the preferred equalization network has an implicit zero at high frequencies, which is owed not to an actual time constant but to the particular circuit topology, with the resulting response similar to the inverse $|F_4(s)|$ function. For this reason, a purely theoretical calculation of component values based on the equalization transfer function will produce an approximation of the required response, not an exact one. This is the source of the most important misconception, which to my knowledge has never been adequately explained in literature. To arrive at the exact response, and a flat frequency response after equalization, we shall have to implement a correction by including τ_4 artificially.

The Equalization

There are several ways to generate an equalization function. A passive equalization network placed between two amplifying stages, as shown in Fig.2, was inherited from the tube/valve era, and has experienced a revival since 1980s.

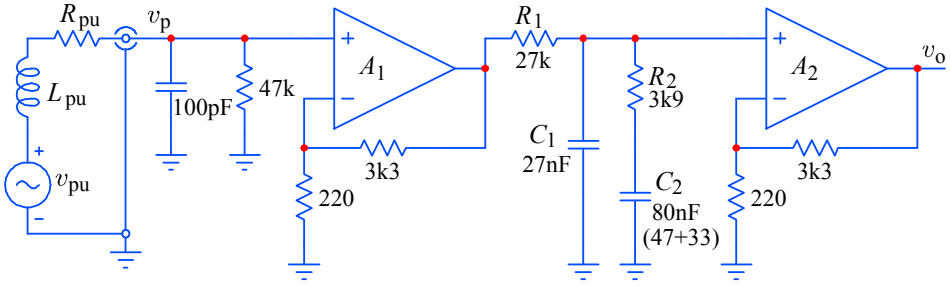


Fig.2: Typical passive RIAA correction network; component values as usually found in literature.

Passive equalization became popular again after some audio ‘gurus’ expressed their concern over amplifiers having capacitive feedback, supposedly sounding “all wrong” and “spoiling” the time constants. In fact, passive equalization is noisier, prone to clipping at high frequencies, and can suffer from slew rate distortion on transients, whilst time constant mismatch can be (often is) as bad as in active systems.

Conventional ‘active’ configurations employ a pair of RC components in the feedback of an amplifier, the most common circuits are shown in Fig.3 and Fig.4.

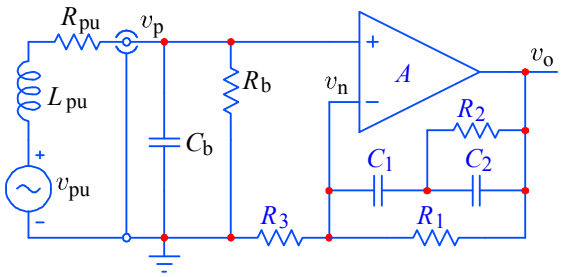


Fig.3: Typical type A active RIAA correction network.

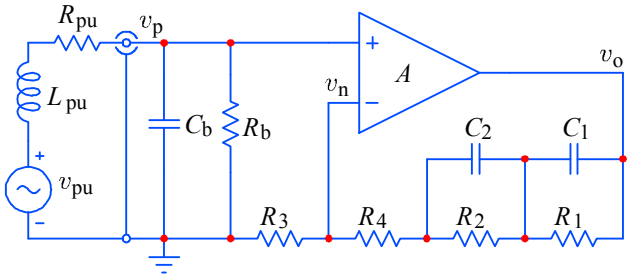


Fig.4: Typical type B active RIAA correction network.

Besides these, hybrid (passive + active) examples, such as in Fig.5, are also found in literature, though not so often in commercial products. Whilst breaking the equalization network in separate sections makes it easier to calculate the correct

component values, problems with high frequency overdrive, slew rate limiting, and noise, remain similar as in purely passive systems.

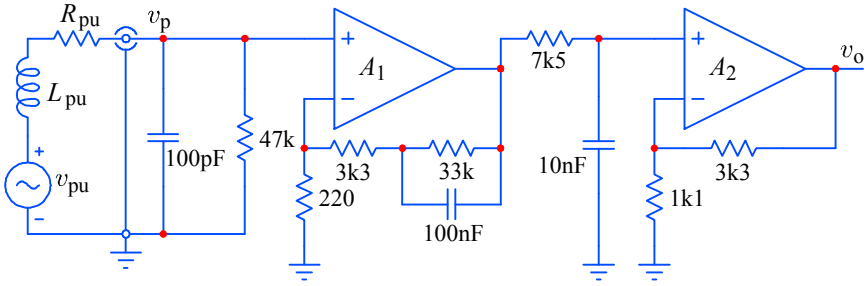


Fig.5: Hybrid (active and passive) RIAA correction network.

By any experienced electronics engineer all these circuits can be legitimately characterized as YATA (yet another trivial amplifier). However, a thorough analysis of their performance is anything but trivial. Consequently approximations and design shortcuts have been extensively used, resulting in compromised or indeed flawed designs to the bitter dissatisfaction of audiophiles, but also causing misjudgments and misconceptions, and endless debates in literature, as well as unjustly interpreting the problem as poor performance of a phonograph cartridge.

Here we present the analysis and response optimization of the circuit in Fig.4, with some additional improvements further on. At first we concentrate on getting the amplifier feedback set up correctly, neglecting the cartridge internal impedance, $R_{pu} = 0$ and $L_{pu} = 0$, thus also $v_{pu} = v_p$. Later we include the influence of the input impedance loading the cartridge internal impedance, and then show a few examples of various cartridge types and their response optimization.

For an ideal and not overdriven amplifier we can assume:

$$v_n \approx v_p \tag{8}$$

because any feedback amplifier with a high open loop gain will force the difference between the inverting and non-inverting input to a minimum (see Appendix 2 for the complete analysis). Consequently the current in the feedback loop is:

$$i_f = \frac{v_p}{R_3} \tag{9}$$

The impedance of the feedback branch is:

$$\frac{v_o - v_p}{i_f} = \frac{1}{sC_1 + \dfrac{1}{R_1}} + \frac{1}{sC_2 + \dfrac{1}{R_2}} + R_4 \tag{10}$$

The amplifier's transfer function is obtained by inserting (9) into (10):

$$R_3 \frac{v_o - v_p}{v_p} = \frac{R_1}{sC_1R_1 + 1} + \frac{R_2}{sC_2R_2 + 1} + R_4 \tag{11}$$

What follows is just simple arithmetic manipulation and reordering, but we shall write each step explicitly, to make the life easier for readers allergic to maths. We first divide all the terms in (11) by R_3:

$$\frac{v_o}{v_p} - 1 = \frac{1}{R_3} \cdot \frac{R_1}{sC_1R_1 + 1} + \frac{1}{R_3} \cdot \frac{R_2}{sC_2R_2 + 1} + \frac{R_4}{R_3} \tag{12}$$

and transfer the -1 term to the right hand side:

$$\frac{v_o}{v_p} = \frac{1}{R_3} \cdot \frac{R_1}{sC_1R_1 + 1} + \frac{1}{R_3} \cdot \frac{R_2}{sC_2R_2 + 1} + \frac{R_4}{R_3} + 1 \tag{13}$$

We want to put the two frequency dependent terms on the same denominator:

$$\frac{v_o}{v_p} = \frac{1}{R_3} \cdot \frac{R_1(sC_2R_2 + 1) + R_2(sC_1R_1 + 1)}{(sC_1R_1 + 1)(sC_2R_2 + 1)} + \frac{R_4}{R_3} + 1 \tag{14}$$

Since this function must be an inversion of (3), the denominator is already in the form expected as the numerator in (3), with the time constants $\tau_1 = C_1R_1$ and $\tau_3 = C_2R_2$. To find τ_2 we must reorder the numerator of (14) to separate the frequency dependent term, so we multiply each parenthesis by its appropriate R:

$$\frac{v_o}{v_p} = \frac{1}{R_3} \cdot \frac{(sC_2R_1R_2 + R_1) + (sC_1R_1R_2 + R_2)}{(sC_1R_1 + 1)(sC_2R_2 + 1)} + \frac{R_4}{R_3} + 1 \tag{15}$$

then we extract the common terms:

$$\frac{v_o}{v_p} = \frac{1}{R_3} \cdot \frac{s(C_2 + C_1)R_1R_2 + R_1 + R_2}{(sC_1R_1 + 1)(sC_2R_2 + 1)} + \frac{R_4}{R_3} + 1 \tag{16}$$

We need to make the additive term in the numerator equal to 1, instead of $R_1 + R_2$, so let us divide the numerator by $R_1 + R_2$ and multiply by it in front:

$$G_3(s) = \frac{v_o}{v_p} = \frac{R_1 + R_2}{R_3} \cdot \frac{s(C_1 + C_2)\dfrac{R_1R_2}{R_1 + R_2} + 1}{(sC_1R_1 + 1)(sC_2R_2 + 1)} + \frac{R_4}{R_3} + 1 \tag{17}$$

Now the numerator contains the explicit expression for the second system time constant τ_2, so we can assign:

$$\tau_1 = C_1R_1 \tag{18}$$

$$\tau_3 = C_2R_2 \tag{19}$$

$$\tau_2 = (C_1 + C_2)\frac{R_1R_2}{R_1 + R_2} \tag{20}$$

The transfer function (17) has, as expected, a second order polynomial in s in the denominator, and a first order polynomial in s in the numerator, so three time constants in total, which is the reason for labeling it $G_3(s)$. It also has a DC gain factor $(R_1 + R_2)/R_3$, as well as a frequency independent gain, $1 + R_4/R_3$, which will be dealt with a little later.

When calculating the component values we would like to have standard values wherever possible, so let us take the following convenient values for τ_3:

$$C_2 = 10\,\text{nF} \tag{21}$$

$$R_2 = 7.5\,\text{k}\Omega \tag{22}$$

Next, from (18) we express C_1:

$$C_1 = \frac{\tau_1}{R_1} \tag{23}$$

and similarly from (19) for C_2:

$$C_2 = \frac{\tau_3}{R_2} \tag{24}$$

With (23) and (24) we return to (20):

$$\tau_2 = \left(\frac{\tau_1}{R_1} + \frac{\tau_3}{R_2}\right)\frac{R_1 R_2}{R_1 + R_2} \tag{25}$$

We multiply all by $R_1 + R_2$ and obtain:

$$\tau_2(R_1 + R_2) = \tau_1 R_2 + \tau_3 R_1 \tag{26}$$

We now divide all by τ_2:

$$R_1 + R_2 = \frac{\tau_1}{\tau_2}R_2 + \frac{\tau_3}{\tau_2}R_1 \tag{27}$$

and group together the expressions containing the same R:

$$R_1\left(1 - \frac{\tau_3}{\tau_2}\right) = R_2\left(\frac{\tau_1}{\tau_2} - 1\right) \tag{28}$$

and we can now express R_1 as:

$$R_1 = R_2\frac{\tau_1 - \tau_2}{\tau_2 - \tau_3} \tag{29}$$

By inserting the appropriate values, we obtain:

$$R_1 = 88308\,\Omega \tag{30}$$

Finally, from (23) we get:

$$C_1 = 36.046\,\text{nF} \tag{31}$$

Before we calculate the response using (17), we must determine R_3 and R_4. For a moment, let us assume that $R_4 = 0$, and look for the value of R_3 which will give the circuit the same DC gain as is the attenuation in (3). So we can write:

$$\frac{R_1 + R_2}{R_3} + 1 = \frac{\tau_1\tau_3}{\tau_2\tau_4} \tag{32}$$

and from this R_3 is equal to:

$$R_3 = \frac{R_1 + R_2}{\dfrac{\tau_1\tau_3}{\tau_2\tau_4} - 1} \tag{33}$$

By inserting the appropriate numerical values we obtain:

$$R_3 = 408.34\,\Omega \tag{34}$$

Now we can check the system response by driving the equalized response $G_3(s)$ (17) from the output of the inverse function $F_4(s)$ (3). In the frequency domain this requires a simple multiplication of the two: $F_4(s) \cdot G_3(s)$. Of course, to plot the response magnitude we must take the absolute value of the total complex function.

In order to see both functions on the same plotting scale together with their product, we have normalized $F_4(s)$ to $F_{4N}(s)$ as in (6), and in the same way $G_3(s)$ to $G_{3N}(s)$, then multiplied $F_{4N}(s) \cdot G_{3N}(s)$, rewritten it as $F_{4N}(j\omega) \cdot G_{3N}(j\omega)$ and finally taken the absolute value to obtain the response magnitude as a function of frequency:

$$M(\omega) = |F_{4N}(j\omega) \cdot G_{3N}(j\omega)| = \sqrt{F_{4N}(j\omega) \cdot F_{4N}(-j\omega) \cdot G_{3N}(j\omega) \cdot G_{3N}(-j\omega)} \tag{35}$$

Analytically this requires a lot of hard work. But today most mathematical computer programs on the market (such as Matlab™, Mathematica™, MathCad™, and others) have the possibility to calculate the absolute value of a complex function for a range of frequencies directly. And likewise, all circuit simulator programs do the same for the AC response analysis. Of course, the actual function must be the same, regardless of the normalization.

In Fig.6 the result of encoding the response into Matlab™ is shown (see Appendix 5 for the complete calculation procedure), and Fig.7 shows the deviation from a flat response in high detail (±0.5 dB vertical scale).

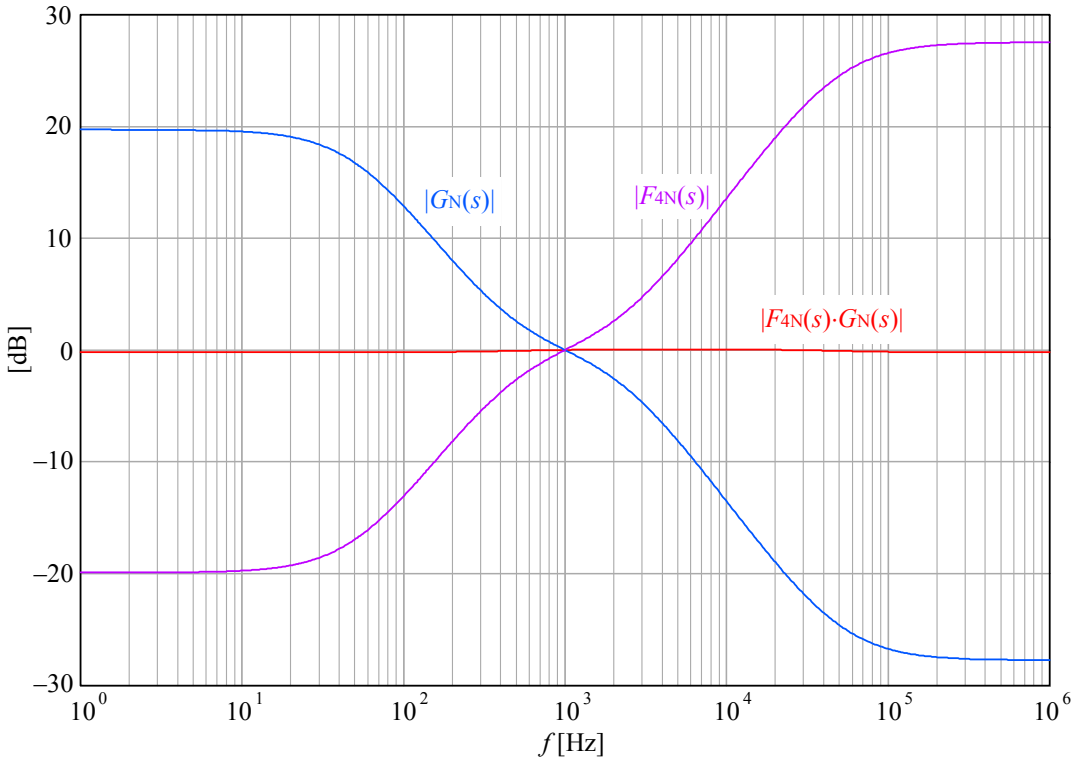


Fig.6: The inverse function $F_{4N}(s)$ multiplying the equalization function $G_N(s)$ should produce a flat response, $|F_{4N}(s) \cdot G_N(s)|$. At the 60 dB vertical scale the result looks like a straight line; a close up view in Fig.7 shows deviations at both low and high frequencies.

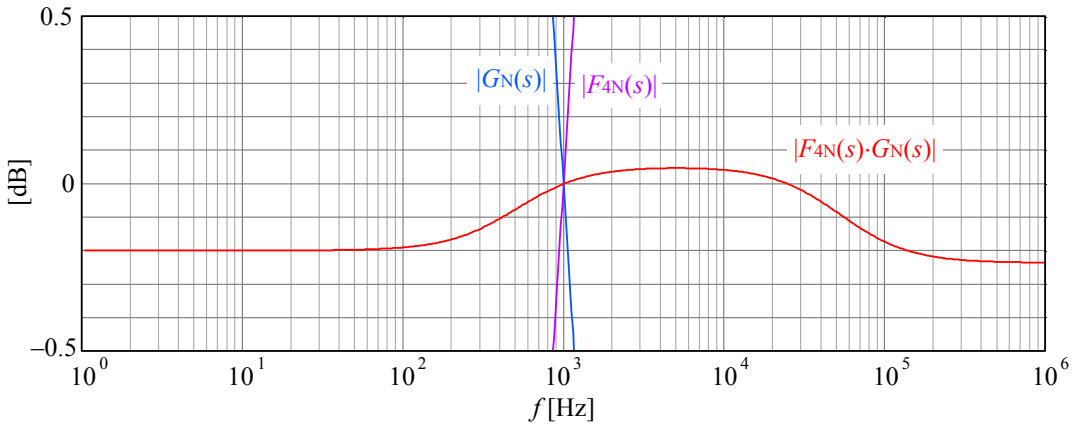


Fig.7: Expanding the vertical scale to ±0.5 dB reveals a deviation from flat response at both high and low frequencies. It may seem that being within 0.3 dB the response is acceptable, but we should and can do better.

The question that arises by looking at Fig.7 is: **since the response is not as flat as we expected, did we make a calculation error, and if so, where?**

The answer is: no error was made in the calculations, but in the original assumption! We have purposely done that in order to expose the most common error made in the past by many circuit designers.

As stated in a previous **note**, the deviation from flatness in Fig.7 stems from the equalization network not having a time constant at 50 kHz, unlike the inversion function $F_4(s)$. The $G_3(s)$ has only three time constants, but its response plot does mimic the inverse of $F_4(s)$ above 50 kHz like having an additional zero, yet **this is a consequence of approaching the unity gain caused by the circuit topology, not because of an actual *RC* time constant!** Consequently, the effective corner frequency is gain dependent, and the gain depends on the value of R_3, and we have calculated R_3 from the sum $R_1 + R_2$ multiplied by three time constants (29), where τ_4 is missing, which means that the obtained **value of R_1 was actually wrong**.

In order to correct the problem, we must recalculate R_1 by including the influence of τ_4 in the equation (29) artificially:

$$R_1 = R_2 \frac{\tau_1 - \tau_2}{\tau_2 - \tau_3}\left(1 + \frac{\tau_4}{\tau_3}\right) = 92055\,\Omega \tag{36}$$

Then recalculating C_1 from (18) gives:

$$C_1 = 34.578\,\text{nF} \tag{37}$$

and recalculating R_3 as in (33) gives:

$$R_3 = 424.3\,\Omega \tag{38}$$

As is shown in Fig.8, with these results the response deviates by less than 0.01 dB.

An identical result for R_3 could be obtained also from equation (17) if we consider that at 50 kHz the absolute value of the transfer function (still with $R_4 = 0$) must be $|j + 1| = \sqrt{2}$. This means that we can equate:

$$j+1=\frac{R_1+R_2}{R_3}\,\frac{s\tau_2+1}{(s\tau_1+1)(s\tau_3+1)}+\frac{R_4}{R_3}+1$$

So 1 cancels, and after multiplying all by R_3 and taking the absolute value we have:

$$R_3=(R_1+R_2)\left|\frac{s\tau_2+1}{(s\tau_1+1)(s\tau_3+1)}\right|_{s=j2\pi 50000}$$

Hence, because s must be evaluated at $j2\pi 50000$, the inclusion of $\tau_4=1/2\pi 50000$ in equation (36) is hereby justified.

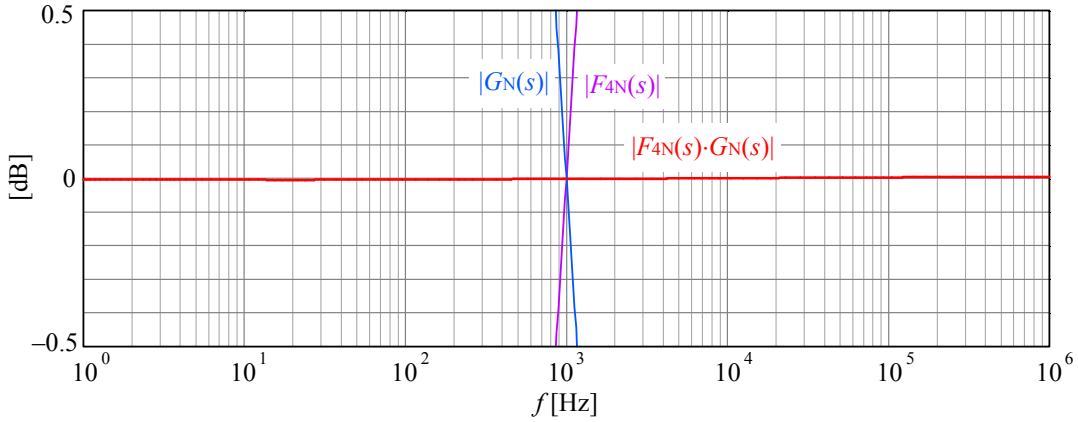


Fig.8: Acceptable response flatness is achieved by accounting for τ_4 in equation (36).

Now that R_3 is known it is easy to calculate the required system gain. Because we kept $R_4=0$, the calculated value of R_3 must actually be redistributed to R_4+R_3. But the final gain will still depend mostly on R_3. To determine the gain from the circuit input and output requirements, we must make a short digression into history.

Regarding the input requirements, the nominal record modulation velocity is 5cm/s at 1kHz (another RIAA agreement), and phonograph cartridges of the moving magnet and variable reluctance type are usually designed to deliver 3–5 $\mathrm{mV_{rms}}$ at such modulation (as given by the sensitivity parameter in manufacturer's specifications). Moving coil types usually deliver 0.1–0.5 $\mathrm{mV_{rms}}$. But there are exceptions to this rule.

As for the output, there is a recommended standard value for the 'line input' in audio equipment, and it is often being referred to as -10 dBm. How much is that?

The unit B ('bel', established at Bell Labs in the 1930s, and named in honour to *Alexander Graham Bell*) is defined as the decade logarithm of the network power transfer ratio, $\log_{10}(P_{\mathrm{out}}/P_{\mathrm{in}})$. The dB is the 'decibel', or $1/10$ of a bel, thus the same power ratio expressed in dB equals $10\log_{10}(P_{\mathrm{out}}/P_{\mathrm{in}})$. The suffix 'm' in dBm stands for the input reference power of 1 mW applied to a load of 600 Ω (both values inherited from the telephone line development era), and this power represents the 0 dBm reference. Strictly speaking, the dB unit should be used **only for expressing power ratios, and only when the input and output load impedance is the same**. Generally, in amplifiers this seldom the case. So over time it has become customary to express voltage ratios (amplifier gain) over different input and output loading also in dB, which is sometimes more correctly indicated as dBV (but this often means that the

reference signal level is 1V_{rms}). Since the power is proportional to voltage squared, the dB value for voltages is calculated as $10\log_{10}(V_{\text{out}}^2/V_{\text{in}}^2)$, or $20\log_{10}(V_{\text{out}}/V_{\text{in}})$.

The electrical power is the product of voltage and current, $P = VI$, and because the current is impedance dependent, $I = V/Z$, the power is proportional to the voltage squared, $P = V^2/Z$. Given the power and the impedance, the voltage required is equal to the square root of the power-impedance product, $V = \sqrt{P\,Z}$. For a 0 dBm level this means $\sqrt{0.001 \cdot 600} = 0.7746\,\text{V}_{\text{rms}}$. Then, -10dBm will be $\sqrt{0.0001 \cdot 600}$, or lower by a factor of $10^{(-10/20)}$, or $0.3162 \cdot 0.7746 = 0.245\,\text{V}_{\text{rms}}$.

So by assuming a 5 mV input and a 245 mV output, the minimum required gain at 1 kHz is $245/5 = 49\times$. For a 3mV signal, the gain should be 82×. And for the 0.5 mV of a moving coil cartridge we need a gain of 490×.

With this knowledge we return to the circuit. The feedback branch impedance expression was given in (10). Excluding R_4 for a moment, the impedance is:

$$Z_{\text{f}}(j\omega) = \frac{R_1}{j\omega C_1 R_1 + 1} + \frac{R_2}{j\omega C_2 R_2 + 1} \tag{39}$$

and its absolute value is:

$$|Z_{\text{f}}(j\omega)| = \sqrt{Z_{\text{f}}(j\omega) \cdot Z_{\text{f}}(-j\omega)} \tag{40}$$

or explicitly:

$$\begin{aligned}
|Z_{\text{f}}(j\omega)| &= \sqrt{\left(\frac{R_1}{j\omega C_1 R_1 + 1} + \frac{R_2}{j\omega C_2 R_2 + 1}\right)\left(\frac{R_1}{-j\omega C_1 R_1 + 1} + \frac{R_2}{-j\omega C_2 R_2 + 1}\right)} \\
&= \sqrt{\frac{R_1(j\omega C_2 R_2 + 1) + R_2(j\omega C_1 R_1 + 1)}{(j\omega C_1 R_1 + 1)(j\omega C_2 R_2 + 1)} \cdot \frac{R_1(-j\omega C_2 R_2 + 1) + R_2(-j\omega C_1 R_1 + 1)}{(-j\omega C_1 R_1 + 1)(-j\omega C_2 R_2 + 1)}} \\
&= \sqrt{\frac{[R_1(j\omega C_2 R_2 + 1) + R_2(j\omega C_1 R_1 + 1)][R_1(-j\omega C_2 R_2 + 1) + R_2(-j\omega C_1 R_1 + 1)]}{(\omega^2 C_1^2 R_1^2 + 1)(\omega^2 C_2^2 R_2^2 + 1)}} \\
&= \sqrt{\frac{R_1^2(\omega^2 C_2^2 R_2^2 + 1) + 2R_1 R_2(\omega^2 C_1 R_1 C_2 R_2 + 1) + R_2^2(\omega^2 C_1^2 R_1^2 + 1)}{(\omega^2 C_1^2 R_1^2 + 1)(\omega^2 C_2^2 R_2^2 + 1)}}
\end{aligned} \tag{41}$$

With the values chosen in (21), (22), (36) and (37), the absolute value of the feedback impedance at $f = 1$ kHz (with $\omega = 2\pi f$) is:

$$R_{\text{f1k}} = |Z_{\text{f}}(j2\pi 1000)| = 9825\,\Omega \tag{42}$$

Then the amplifier's gain $A_{1\text{k}}$ at 1 kHz is given by:

$$A_{1\text{k}} = \frac{R_{\text{f1k}} + R_4}{R_3} + 1 \tag{43}$$

Without R_4, and for $A_{1\text{k}} = 49\times$, the value of R_3 should be 48× smaller than R_{f1k}, or about $205\,\Omega$. But we have shown that to achieve an acceptable response flatness, a value of $R_3 + R_4 = 424\,\Omega$ is necessary. Note however that there are two other reasons for having a higher resistance value in the feedback to ground path.

One reason is that the amplifier must be capable of driving that impedance at high frequencies, where the feedback impedance is the same or lower, owed to the capacitive path. And not just at the nominal output level of 245 mV_{rms}, but also at some 20 to 30 dB higher level, 2.45–7.75 V_{rms}, or almost 11 V peak! Such a headroom is necessary for clean processing of high level transients in high quality recordings. With only 200 ohm, this requires 55 mA peak at 20 kHz, not including any additional loading by the following circuitry. At those levels the amplifier's output overload protection must not be activated if low distortion is expected, and most IC amplifiers will distort already at 10–15 mA.

The other reason is amplifier's stability. Namely, many amplifiers with high bandwidth are being offered as 'decompensated' (typical examples include the popular OP37 and NE5534; this last one may be compensated externally, but others give no such option). This means that their internal dominant pole is set higher than the usual ~100 Hz, with the consequence of the secondary pole of the output stage compromising the loop phase margin, therefore a minimal gain of 4× or 5× is required (at unity gain, or with a capacitive load of an ordinary 1m long coaxial cable, such amplifiers will oscillate). Of course, it would be best and simplest to employ a unity gain stable amplifier, but until recently a stable audio grade amplifier with low noise, high output current, high bandwidth, high slew rate, and low distortion was either unavailable or was a compromise in one way or another.

On the other hand, a high effective feedback resistance would also have a high thermal noise, comparable to that of the cartridge, so lowering the value of R_3 in favor to R_4 gives us just what is required. In Table 1 we give several values of gain, R_3, and R_4 for various cartridge sensitivities, calculated after equation (43).

Table 1

Sensitivity [mv @ 1kHz, 5cm/s]	Gain	Gain [dB]	R_3 [Ω]	R_4 [Ω]
5.0	49×	33.8	206	218
3.5	70×	36.9	146	278
3.0	81.67×	38.24	127	299
0.5	490×	53.8	21	403

Regarding standard component values (E24 set for resistors, and E12 for capacitors), here we give some parallel combinations which result in values close enough to those required. Of course, if easily available, the E48 or E96 set may be preferred, because tolerances add up for combinations of components (either serial or parallel). This undesirable property may be reduced slightly by selecting one high and one low value in the hope that the tolerance of the high value will not change the combined value by much, so even if a 1% tolerance of each component result in a 2% total, the actual individual variation will be less than 1%, and will therefore spoil the value by a lower amount. Whilst resistors are easily found in 0.5% or even 0.1% ranges, capacitors are usually in the 20%, 10%, and 5% ranges, but in the 5% range only the most common decade values are ordinarily available. So hand-picking the correct values from a large bunch will be unavoidable in practice.

The values for R_2 (7500 Ω) and C_2 (10 nF) have been already chosen from the standard E24 set, so we only need to measure and select those which will be within 0.1%. For the remaining values here are a few possible parallel combinations which

will be within 0.5% or so, assuming exact values (of course, it is always possible to find slightly different values which will give better results):

$$R_3 = 206.3\,\Omega \rightarrow \frac{1}{\frac{1}{240} + \frac{1}{1500}} = 206.9\,\Omega \rightarrow +0.3\%$$

$$R_4 = 218.3\,\Omega \rightarrow \frac{1}{\frac{1}{240} + \frac{1}{2400}} = 2018.2\,\Omega \rightarrow -0.05\%$$

$$R_1 = 92055\,\Omega \rightarrow \frac{1}{\frac{1}{160\text{k}} + \frac{1}{220\text{k}}} = 92.6\,\text{k}\Omega \rightarrow +0.6\%$$

$$C_1 = 34.58\,\text{nF} \rightarrow 33\text{ nF} + 1.5\,\text{nF} = 34.5\,\text{nF} \rightarrow -0.2\%$$

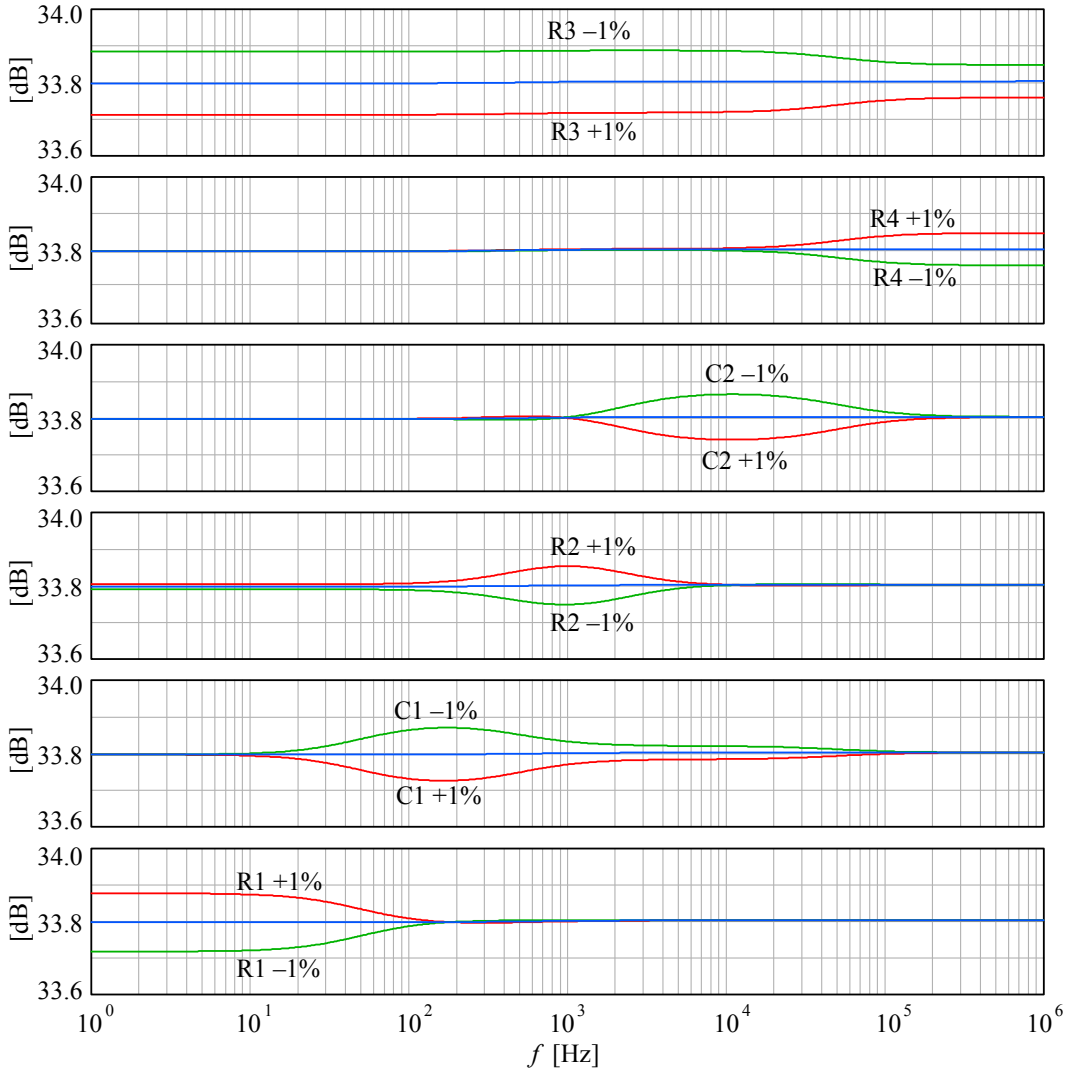


Fig.9: Influence of components' ±1% tolerance on various parts of the spectrum.

Another possibility for C_1 would be a parallel combination of 22 nF and 12 nF nominal, each with a slightly larger value, or maybe a pair of 68 nF in series. In most cases the actual value of a component is on the low side of the tolerance range, and this can be used to advantage when selecting combined values manually to obtain a value closer to the optimal.

In Fig.9 we report the influence of each individual component change by $\pm 1\%$ on the system's flatness deviation. We can see that such changes result in less than 0.1 dB deviation. Note however, that **combined variations of two or more components can have a greater combined effect**. This is especially true for C_1, the influence of which extends into the spectral range influenced by both C_2 and R_2. Likewise, R_2 will also alter the spectral range under dominant influence of C_1 and R_1. A slight inverse variations from both C_2 and R_1 is seen between 200 Hz and 1 kHz.

R_3 and R_4 are less problematic, since R_4 affects only the ultrasonic range, and R_3 affects the whole audible range equally.

The Cartridge and its Load

We now want include the effect of the cartridge's internal impedance and its load into the equation. Since at high frequencies the cartridge impedance is dominantly inductive, and its load is dominantly capacitive (because of the connecting cable and the amplifier's input capacitance), there will be a pronounced resonance at or above 15 kHz, and this must be taken into account.

In the following analysis we shall assume that the previously calculated circuit response $G_3(s)$ will correctly equalize the encoding function $F_4(s)$, so the product $F_4(s) \cdot G_3(s)$ is flat, and can be neglected for cartridge response optimization.

In Fig.4 the generator v_{pu} represents the cartridge nominal output (usually between 3–5 mV for the 5 cm/s groove modulation velocity at 1 kHz). The resistor R_{pu} represents the coil wire resistance and L_{pu} represents the coil and magnetic core inductance. The signal v_{p} is the output signal when the cartridge is loaded by its nominal load R_{b} and the cable and amplifier's input capacitance C_{b}. But L_{pu} and R_{pu} can vary widely between manufacturers, and also between models from the same manufacturer! Some popular examples: Sonus Blue and Gold series have 150 mH and 300 Ω, Ortofon 2M Black has 630 mH and 1200 Ω; but notable exceptions include Grado G1+ with only 55 mH and 700 Ω, and Stanton 681EEE MkIII with 930 mH and 13 kΩ. And we are not even considering low output moving coil types here!

We shall illustrate the influence of the cartridge impedance components by using the following three examples to cover the most common range of values:

Table 2

	Type		
Impedance	Z_{A}	Z_{B}	Z_{C}
L_{pu} [mH]	150	300	600
R_{pu} [Ω]	300	600	1200

Here we are interested only in the response variation with frequency that is owed to the electrical impedance of the pickup and its load. The mechanical resonance

of the needle, the cantilever, and the magnet moving mass, with the cantilever support compliance (elasticity modulus) are given by the type of cartridge, and are not influenced by the circuitry. Mechanical cartridge-tonearm resonances usually affect the deep bass and sub-audio frequencies, whilst the cantilever resonance and other mechanical properties affect the upper audio range. To compensate for this some manufacturers specify R_b and C_b such to intentionally enhance high frequencies, so checking the actual response will be necessary when using a different load.

Low frequency resonances can be dealt with either by increasing mass, or compliance, or both. However high compliance cartirdges are not suitable for many tonearms, and a high mass at the end of the tonearm increases the system's inertia, which in turn increases record wear. Likewise, a high turntable mass causes high axis pressure, increasing mechanical wear and in time developing rumble. A judicious compromise is necessary. Making the turntable support mass high, and adding soft suspension to decouple all from other supporting masses, thus preventing the bass transmission through the material, is often the best way to go.

Returning to the cartridge equivalent circuit, the ratio $v_\mathrm{p}/v_\mathrm{pu}$ can be found from voltages on their associated impedances:

$$\frac{v_\mathrm{pu} - v_\mathrm{p}}{sL_\mathrm{pu} + R_\mathrm{pu}} = \frac{v_\mathrm{p}}{\dfrac{1}{sC_\mathrm{b} + \dfrac{1}{R_\mathrm{b}}}} \tag{44}$$

With a little reordering we obtain the transfer function:

$$H(s) = \frac{v_\mathrm{p}}{v_\mathrm{pu}} = \frac{R_\mathrm{b}}{R_\mathrm{g} + R_\mathrm{b}} \cdot \frac{\dfrac{1}{L_\mathrm{pu}C_\mathrm{b}}\left(\dfrac{R_\mathrm{pu}}{R_\mathrm{b}} + 1\right)}{s^2 + s\left(\dfrac{1}{C_\mathrm{b}R_\mathrm{b}} + \dfrac{R_\mathrm{pu}}{L_\mathrm{pu}}\right) + \dfrac{1}{L_\mathrm{pu}C_\mathrm{b}}\left(\dfrac{R_\mathrm{pu}}{R_\mathrm{b}} + 1\right)} \tag{45}$$

Fig.10 shows a plot of (45) against frequency for the three representative cartridge impedances given in Table 2.

As can be seen from the plot, low impedance cartridges tend to peak in the ultrasonic part of the spectrum, at 20–30 kHz (this was once considered desirable for quadrophonic reproduction), whilst high impedance versions tend to attenuate above 10 kHz. Both effects can be reduced by lowering C_b, but this will be ultimately limited by the cable capacitance.

The peaking can also be reduced by lowering R_b, but this attenuates the signal because of the frequency independent term in (45):

$$a = \frac{R_\mathrm{b}}{R_\mathrm{pu} + R_\mathrm{b}} \tag{46}$$

However pay attention to the cartridge mechanical response, because some cartridge manufacturers compensate its influence by deliberately specifying a lighter electrical loading, enhancing high frequencies. Always check the response using a good test record, either with continuous tone at various frequencies or by a 1 kHz squarewave.

As will be shown later, making the amplifier's input impedance adaptable by a small DIL switch array can save us a lot of soldering.

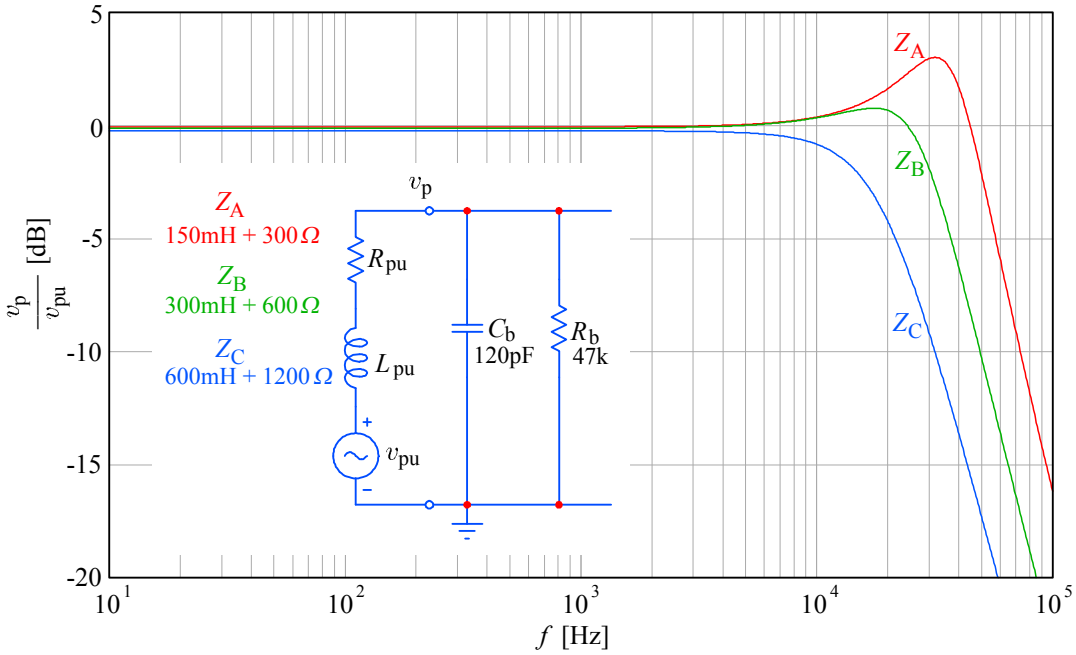


Fig.10: Influence of cartridge impedance on the frequency response. Lower impedance types are generally preferable for lower noise, but they require lower cable and input capacitance C_b, and/or lower input resistance R_b. See Fig.11 and Fig.12.

The resonant frequency is given by the last term of the denominator of (45):

$$f_r = \frac{\omega_r}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{1}{L_{pu}C_b}\left(\frac{R_{pu}}{R_b}+1\right)} \tag{47}$$

The damping factor $\xi = 1/2Q$ is obtained from the denominator's middle term:

$$2\xi\omega_r = \frac{1}{C_b R_b} + \frac{R_{pu}}{L_{pu}} \tag{48}$$

$$\xi = \frac{1}{2}\sqrt{\frac{L_{pu}}{C_b R_b (R_{pu}+R_b)} + \frac{2R_{pu}}{R_{pu}+R_b} + \frac{C_b R_b R_{pu}^2}{L_{pu}(R_{pu}+R_b)}} \tag{49}$$

The influence of C_b is shown in Fig.11. Lower value results in lower peaking and larger bandwidth. Since C_b includes the capacitance of the cable, which usually has about 100 pF/m, shortening the cable may be beneficial, if possible.

The load resistance R_b influences the system damping, as Fig.12 shows. Lower value results in lower peaking. What is required is not a 'critical damping', ($\xi = 0.5$), which insures no overshoot on transients, but rather a Bessel-like damping, with $\xi = 1/\sqrt{3}$, which gives a slightly higher bandwidth and an envelope time delay equal for all frequencies up to the upper bandwidth limit. Time delay is the phase derivative of frequency, $\tau_e = d\varphi/d\omega$, so if τ_e is flat, the phase must vary linearly with frequency within the bandwidth of interest, therefore such systems are often being referred to as

'linear phase'. What is more important is that such systems also exhibit the fastest rise time, Fig.13, with minimal overshoot (~0.5% of the step amplitude). This means no ringing on transients.

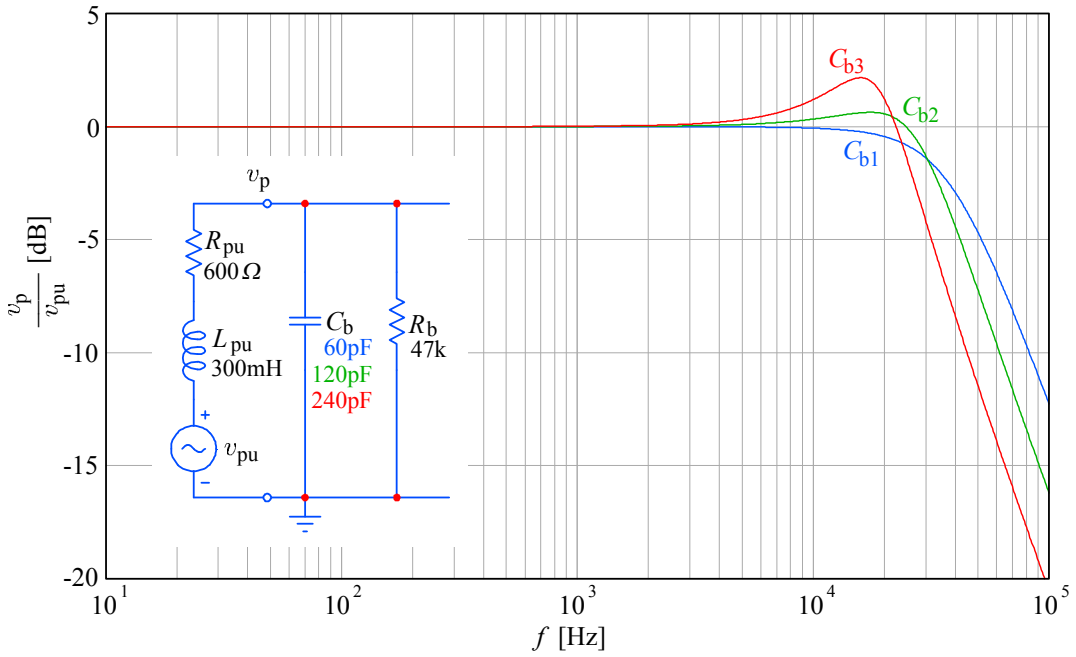


Fig.11: Influence of capacitive loading on the frequency response. Lower C_b values are preferable.

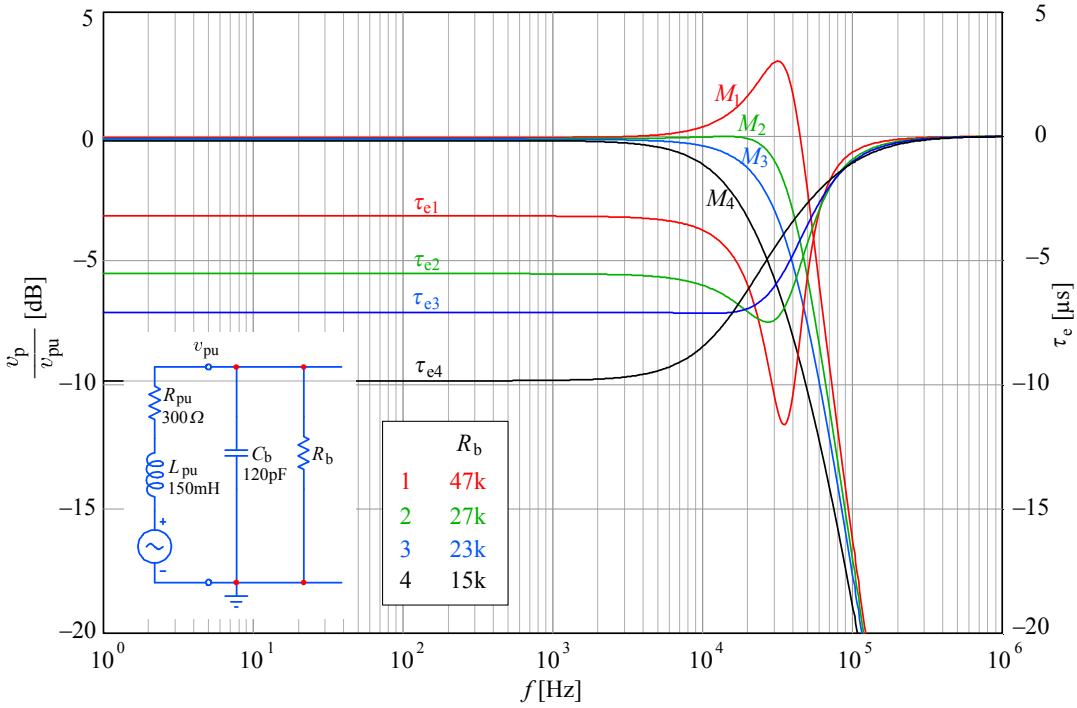


Fig.12: Influence of resistive loading on the frequency response magnitude (M) and the envelope time delay (τ_e). Lower values reduce the high frequency peaking and increase the pass-band time delay, but also decrease the time delay at resonance. Bessel-like response (maximally flat time delay, τ_3) is obtained with $R_b = 23\,k\Omega$.

Generally speaking, because both ω_r and ξ are affected by C_b and R_b, it is important to have the load adjusted to the system by observing the rise time overshoot

on an oscilloscope using some low frequency square wave as the driving signal. Note also that the total system thermal noise is proportional to the square root of the real part of the total input impedance and the bandwidth. The total noise of the system will be examined in detail a little later.

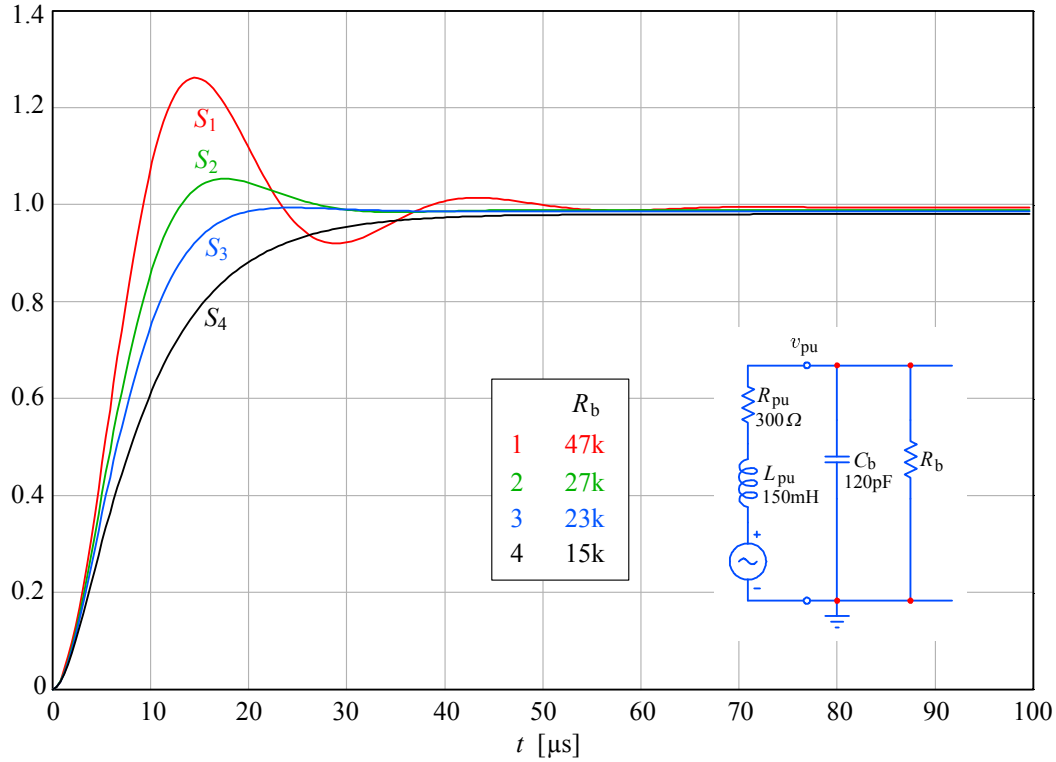


Fig.13: Time domain step response for the same conditions as in Fig.12.

The System

The complete system transfer function is obtained by multiplying $F_4(s)$ (4) with $H(s)$ (44) and $G_3(s)$ (17):

$$E(s) = F_4(s) \cdot H(s) \cdot G_3(s) \tag{50}$$

Regarding the cartridge response $H(\mathrm{s})$, here we are using $R_{\mathrm{pu}} = 600\,\Omega$ and $L_{\mathrm{pu}} = 300\,\mathrm{mH}$ because those values are fairly close to the average of what is found on the market. However, the reader is encouraged to insert the values of his own cartridge and loading in order to check his own system, either on a circuit simulator or with the aid of a mathematical program. Also we shall use $C_{\mathrm{b}} = 120\,\mathrm{pF}$ because the majority of people will probably be reluctant to cut in half the 1 m long cable from the turntable (which usually has 100 pF/m), and the amplifier input will have an additional 20 pF at least, and some will offer the possibility of adjusting C_{b} (to higher values).

For $G(s)$ we shall use the values already calculated, and the reader can change the values of R_3 and R_4 to obtain the correct gain for his own cartridge sensitivity. Fig.14 shows the resulting responses. Note that we are using the 1 kHz normalization, that is $F_{4\mathrm{N}}(s)$ and $G_{3\mathrm{N}}(s)$ responses.

Note that the cartridge response peaking at 15 kHz (or above) is lowered by the single pole at 50 kHz, which needs to be added after the RIAA equalization circuit.

Since cartridges have very different impedances, any remaining peaking will need to be adjusted by R_b and C_b. Fig.15 shows the complete schematic. The gain settings for different cartridge sensitivity must be made so that $R_3 + R_4$ is constant!

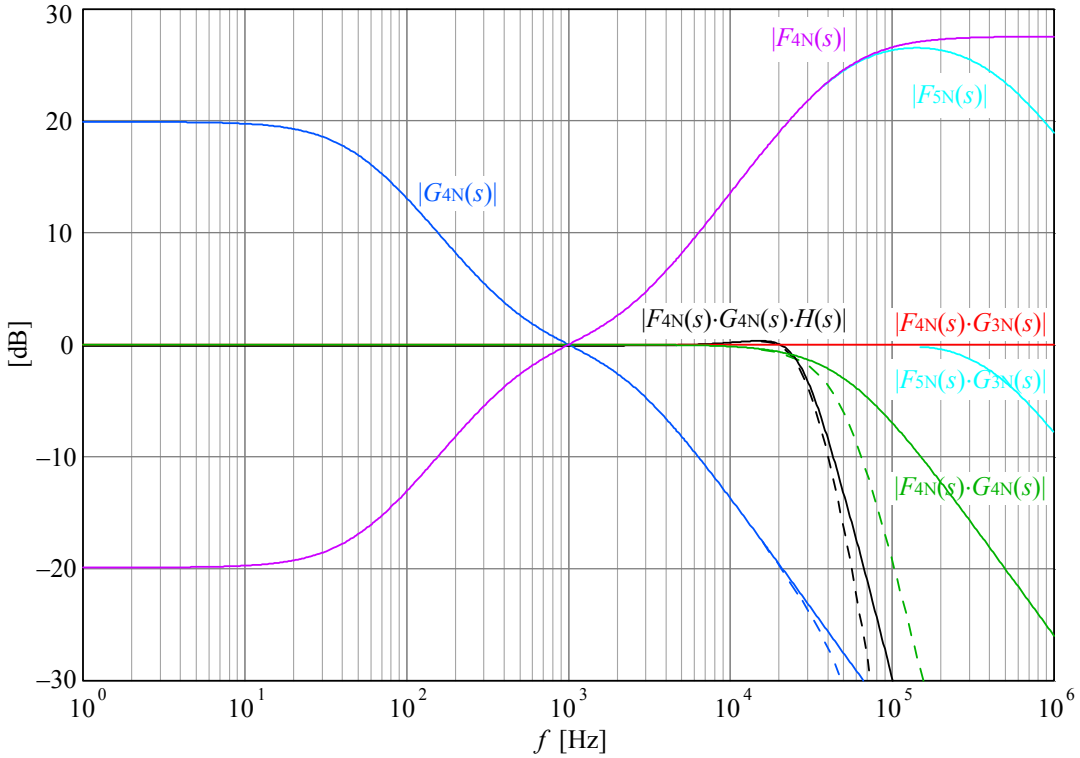


Fig.14: System flatness check by multiplying $F_4(s)$ with $G_3(s)$ is the flat red curve. An additional pole at 50 kHz causes the $G_4(s)$ response to continue dropping at -6 dB/octave beyond 20 kHz, with the corresponding deviation from the flat response indicated by the green line. The black line shows the cartridge response $H(s)$ modified by the system. The dashed lines show the influence of a 2nd-order Butterworth filter at 50 kHz in the record cutter system; it affects the responses only beyond 20 kHz.

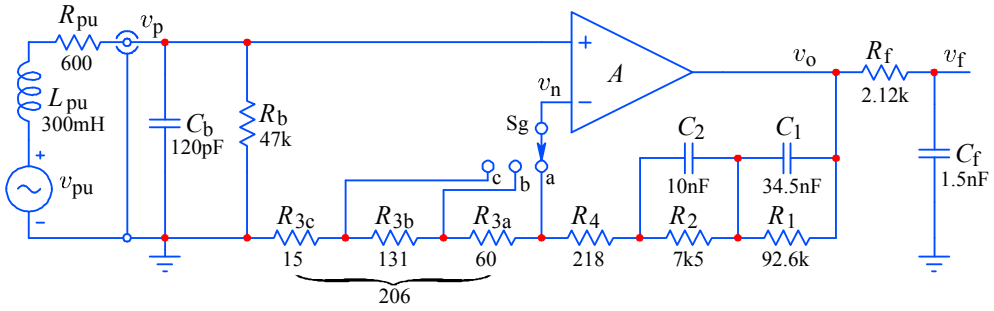


Fig.15: Complete circuit schematic with values; a switch sets the gain for different cartridge sensitivities: a) normal (5mV), b) low (3.5mV), c) moving coil (0.3mV).

The transfer function for the complete circuit in Fig.15 is again a simple product of $H(s)$ (45), $G(s)$ (17), and the transfer function of the $C_f R_f$ low pass filter; this is shown by the blue curve in Fig.14. For the evaluation of the response flatness we must multiply the circuit transfer function by $F_4(s)$.

$$\frac{v_f}{v_{pu}} = F_4(s) \cdot H(s) \cdot G_3(s) \cdot J_1(s) = F_4(s) \cdot H(s) \cdot G_3(s) \cdot \frac{1}{sC_f R_f + 1} \qquad (51)$$

In Fig.16 we show how the influence of the low pass filter J_1 can be take into account when setting up the load of the cartridge. We compare the manufacturer's recommended load settings with optimized settings for six popular cartridge types, both in frequency domain and time domain. We assume the RIAA equalization is done correctly, and take into account only the 400 kHz encoding bandwidth limit together with the 50 kHz limit. It is clear that the manufacturer's settings seldom result in adequate time domain response. So the optimization has been performed for the time domain response minimizing the overshoot of the step response waveform and maximizing the bandwidth, and the resulting frequency response is shown along.

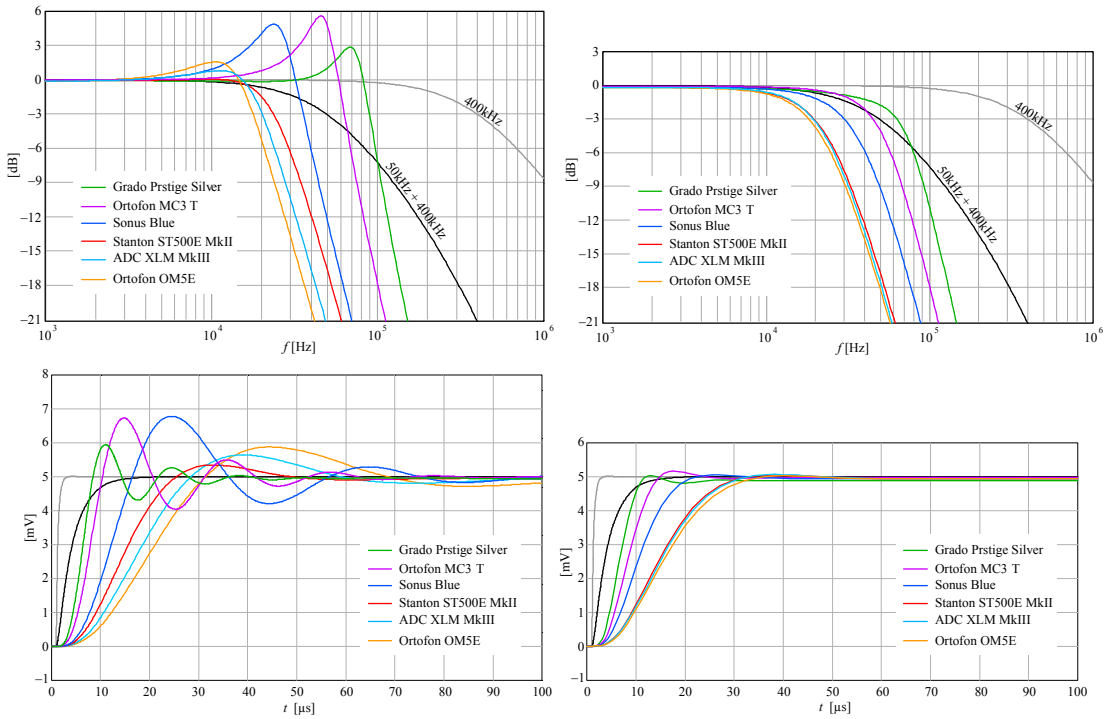


Fig.16: Modifying cartridge load (C_b and R_b) to obtain the lowest overshoot for a maximal bandwidth. Most manufacturers prefer a high degree of high frequency peaking (left hand figures), as this gives a sense of 'brightness' and 'openness'. A proper adjustment can be achieved for most cartridges (right hand figures), even with very different internal impedance.

The DC Offset

An additional problem to solve in the circuit of Fig.15 is the DC offset. This has been traditionally dealt with by adding a large capacitor in series with R_3. Such low frequency filtering helps to reduce the amplification of the rumble generated by the turntable shaft, bearing, and driving mechanism, and preventing possible acoustic feedback below 20 Hz if the loudspeaker system with a large subwoofer is used.

In 1972 the IEC proposed an additional break point at 20 Hz to be part of the reproduction standard (but not the recording!). So a low frequency roll off is desirable for several reasons. But church organ music fans will probably make a grim face at this, since the C_0 bass pedal note is at 16.203 Hz. Of course, it is always possible to make the cut off below 16 Hz, but then the mentioned problems may resurface.

The relation for the low frequency cut off, which multiplies (51), is simply:

$$F_0(s) = \frac{s}{s + \dfrac{1}{C_0 R_3}} \tag{52}$$

where C_0 is the capacitor added between R_3 and ground. But because of the very low value of R_3, cutting the response at 20 Hz or below requires a large capacitance value, about 50 µF for $R_3 = 206\,\Omega$ and 16 Hz. And because electrolytic capacitors are the only type offering such large values, two back to back electrolytics of a double value, in parallel with a 0.1–1 µF polystyrene or polypropylene capacitor are required in order to avoid asymmetrical impedance and bypass the large serial inductance of the electrolytic capacitors.

Readers not willing to renounce to the full bass depth of the Bach's *Toccata and Fugue in D minor*, or the Richard Strauss' *Also sprach Zarathustra*, will have to use 2× 150 µF (for each channel) for a −3 dB limit at ~5.5 Hz, thus keeping those precious 16 Hz within the desirable pass band.

Shown in Fig.17 is a modified feedback configuration proposed recently by Kendall Castor-Perry in *Linear Audio* Vol.1, pp.133–137.

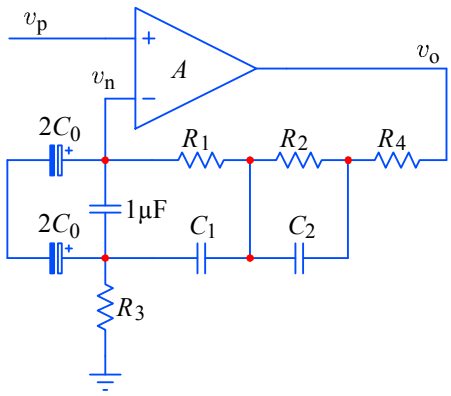


Fig.17: One version of the DC blocking principle proposed by Kendall Castor-Perry.

This arrangement allows most of the feedback current to pass via C_1 and C_2 to R_3, and only at low frequencies a very small current via R_1 and R_2 will go through those large nonlinear electrolytics, so their influence is minimized.

But those large capacitors are now floating, so they represent a large stray capacitance to surrounding fields from transformers, RF sources (mobile phones), etc. Fig.18 shows an easy way of avoiding large capacitors by using a simple and cheap positive integrator servo loop with only a single capacitor of moderate value.

By itself, the DC offset is not a problem in audio. But if too large, it can reduce the signal headroom on one polarity, since the gain at DC is about 500×. For example, an OP37 with its ~10 µV input offset in place of the main amplifier A will exhibit a maximum output offset of 5 mV or so, which is tolerable. But the 4 mV input offset of an NE5534 will be amplified to 2V.

The integrator in Fig.18 with an OP277 (an improved version of the popular OP07) can easily reduce this below 20 µV at the output node v_o. Also, any system with a large DC offset will generate a large step at the output at power on. This is prevented if the power supply rises slowly, say < 2V/s.

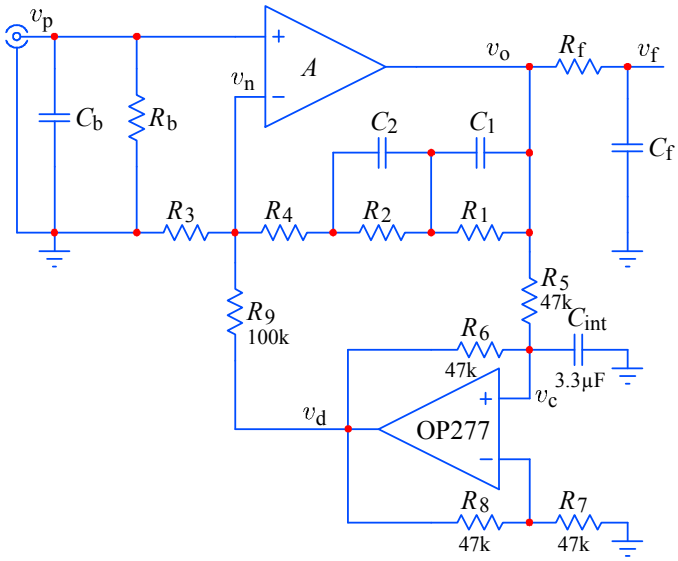


Fig.18: An integrator servo loop using an OP277 can reduce the DC offset to less than 20 µV.

We start the analysis of the loop by finding the current i_c through C_{int}:

$$i_c = \frac{v_c}{\frac{1}{sC_{int}}} \tag{53}$$

Here v_c is the voltage across the integration capacitor C_{int}. The OP277 is configured for a gain of 2× by two equal resistors in its negative feedback loop, $R_7 = R_8$, and as long as the circuit is in the linear range, the voltage at the inverting input will be equal to v_c. The integration condition is achieved if $R_6/R_5 = R_8/R_7$. Therefore:

$$\frac{v_d - v_c}{R_8} = \frac{v_c}{R_7} \tag{54}$$

We can rearrange this as:

$$v_d = v_c\left(\frac{R_8}{R_7} + 1\right) = 2v_c \tag{55}$$

So we can express the sum of all the currents at the v_c node:

$$\frac{v_o - v_c}{R_5} = i_c + \frac{v_c - v_d}{R_6} \tag{56}$$

By replacing i_c with (53), and v_d with (55), and solving for v_c gives:

$$v_c = v_o \frac{1}{sC_{int}R_5 - \frac{R_5}{R_6} + 1} \tag{57}$$

Thus by making $R_5 = R_6$ we have only:

$$v_c = v_o \frac{1}{sC_{int}R_5} = \frac{v_o}{R_5} \cdot \frac{1}{sC_{int}} \tag{58}$$

Equation (58) is the response of an integrator, integrating the current v_o/R_5 on C_{int}. We know this because of the form $v_o/s\tau$; in contrast, an ordinary low pass filter would have $v_o/(s\tau+1)$ in Laplace space. The driving impedance seen by C_{int} is very high, limited by the tolerance of R_5 and R_6 and the input impedance of the OP277.

Now $v_d = 2v_c$ (55) is applied via R_9 to the inverting input v_n of A. At DC, and with $v_p = 0$, the voltage at the inverting input v_n of the main amplifier A is the input offset voltage of A. The following equation holds for the DC servo loop:

$$\frac{v_d - v_n}{R_9} = \frac{v_n}{R_3} + \frac{v_n - v_o}{R_{DC}} \tag{59}$$

where $R_{DC} = R_1 + R_2 + R_4$. This we reorder as:

$$v_d = v_n\left(1 + \frac{R_9}{R_3} + \frac{R_9}{R_{DC}}\right) - v_o\frac{R_9}{R_{DC}} \tag{60}$$

We are interested in the value of v_a, so:

$$v_o = v_n\left(1 + \frac{R_9}{R_3} + \frac{R_9}{R_{DC}}\right)\frac{R_{DC}}{R_9} - v_d\frac{R_{DC}}{R_9} \tag{61}$$

We replace v_d with $2v_c$ from (55), and then replace v_c from (58):

$$v_o = v_n\left(\frac{R_{DC}}{R_3} + 1 + \frac{R_{DC}}{R_9}\right) - 2\frac{v_o}{R_5}\cdot\frac{1}{sC_{int}}\frac{R_{DC}}{R_9} \tag{62}$$

Effectively (62) defines a transfer function with a single zero at $1/C_{int}R_5$, or $v_o/v_n = K_1 s/(s + K_2/C_{int}R_5)$. We have intentionally left v_o/R_5 on the right hand side, since it is the DC error current, proportional to v_o, integrated to correct back v_o. The equation (62) means that the input offset v_n is amplified by slightly more than the closed loop DC gain of A, and then reduced (note the negative sign!) by the integration of the error current on C_{int}, itself amplified by $2(R_1 + R_2 + R_4)/R_9$.

Instead of the expression in the Laplace space, we could have written the relationship between v_c and i_c in the usual time domain integral form:

$$v_c = \frac{1}{C_{int}}\int_0^\tau i_c dt \tag{63}$$

with the integration time $\tau \gg C_{int}R_5$, so that the integration loop has enough time to settle close to the final offset value of v_c, and consequently also v_o.

With $R_5 = 47\,\text{k}\Omega$, and $C_{int} = 3.3\,\mu\text{F}$, the low frequency limit ($-3\,\text{dB}$ below pass band) is about 1 Hz. This is enough for the settling speed of the loop, but with a large sub woofer the very low frequency rumble and acoustic feedback may still be amplified enough to pose problems. What is needed is an additional high pass filter of high order to let pass all above 16 Hz unaffected, but cut sharply everything below.

The High Pass Filter

The advantage of a separate HPF is that it can be tailored to user's needs adding capacitors in parallel by a 3-way rotary switch, or bypassed completely. For this

task we use an active third order high pass filter as shown in Fig.19, a configuration known in filter theory as the *Sallen–Key* type, with a cut off frequency of, say, 10 Hz, and an attenuation slope of $18\,\mathrm{dB}/2f$ below the cut off.

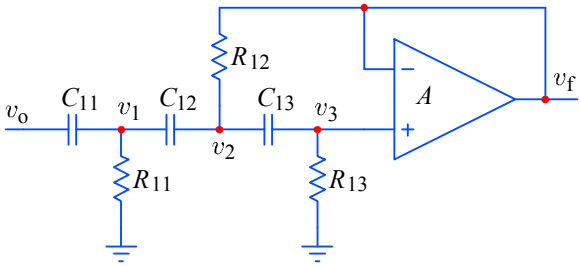


Fig.19: A third order high pass filter of the *Sallen–Key* type.

The filter transfer function is obtained from the following node equations:

at v_1:
$$\frac{v_o - v_1}{\frac{1}{sC_{11}}} = \frac{v_1 - v_2}{\frac{1}{sC_{12}}} + \frac{v_1}{R_{11}} \tag{64}$$

at v_2:
$$\frac{v_1 - v_2}{\frac{1}{sC_{12}}} = \frac{v_2 - v_3}{\frac{1}{sC_{13}}} + \frac{v_2 - v_f}{R_{12}} \tag{65}$$

at v_3:
$$\frac{v_2 - v_3}{\frac{1}{sC_{13}}} = \frac{v_3}{R_{13}} \tag{66}$$

For a unity gain non-inverting amplifier $v_3 \approx v_f$. First we substitute v_3 with v_f in (66) and (65), then express v_2 from (66), and replace v_2 with its expression in (65), then find v_1 from (65), and replace v_1 in (64) with that. In that order we obtain:

$$v_2 = v_f\left(\frac{1}{sC_{13}R_{13}} + 1\right) \tag{67}$$

$$v_1 = v_f\left[\frac{1}{s^2C_{12}R_{12}C_{13}R_{13}} + \frac{1}{sC_{13}R_{13}}\left(\frac{C_{13}}{C_{12}} + 1\right) + 1\right] \tag{68}$$

$$v_o = v_f\left[\frac{1}{s^3C_{11}R_{11}C_{12}R_{12}C_{13}R_{13}} + \frac{1}{s^2C_{12}R_{12}C_{13}R_{13}}\left(\frac{C_{12}}{C_{11}} + 1\right)\right.$$
$$\left. + \frac{1}{s^2C_{11}R_{11}C_{13}R_{13}}\left(\frac{C_{13}}{C_{12}} + 1\right) + \frac{1}{sC_{13}R_{13}}\left(\frac{C_{13}}{C_{11}} + \frac{C_{13}}{C_{12}} + 1\right) + \frac{1}{sC_{11}R_{11}} + 1\right] \tag{69}$$

From (69) we can express the transfer function v_f/v_o. In addition, from filter theory, specifically the Sallen–Key configuration such as in Fig.19, we know that the sensitivity to component variations is minimized if the serial elements are made equal, so $C_{11} = C_{12} = C_{12} = C$. With this we rewrite (69) as:

$$\frac{v_\mathrm{f}}{v_\mathrm{o}} = \frac{1}{\frac{1}{s^3C^3R_{11}R_{12}R_{13}} + \frac{2}{s^2C^2R_{13}}\left(\frac{1}{R_{12}} + \frac{1}{R_{11}}\right) + \frac{1}{sC}\left(\frac{3}{R_{13}} + \frac{1}{R_{11}}\right) + 1} \tag{70}$$

Finally, by multiplying both the numerator and the denominator by s^3, and with a little regrouping we obtain:

$$\frac{v_\mathrm{f}}{v_\mathrm{o}} = \frac{s^3}{s^3 + s^2\frac{1}{C}\left(\frac{3}{R_{13}} + \frac{1}{R_{11}}\right) + s\frac{2}{C^2R_{13}}\left(\frac{1}{R_{12}} + \frac{1}{R_{11}}\right) + \frac{1}{C^3R_{11}R_{12}R_{13}}} \tag{71}$$

At high frequencies the term s^3/s^3 dominates, so the filter has unity gain. At low frequencies the resonant frequency is determined by the last term in the denominator:

$$\omega_0 = \frac{1}{C}\sqrt[3]{\frac{1}{R_{11}R_{12}R_{13}}} \tag{72}$$

It is of course possible to set also the three resistor values equal, for simplicity, and obtain the so called critically damped response. However the transition from pass band to the attenuation slope towards low frequencies is very gradual in that case, and we would rather have a steeper transition, with greater attenuation of very low frequencies. As most turntables have their mechanical resonance and the dominant rumble spectrum within 2–5 Hz, we want to suppers 5 Hz considerably.

The sharpest transition is offered by the Chebyshev family of filters. However, a sharp cutoff also means a large phase jump at ω_0, and therefore a large derivative $\tau_\mathrm{e} = d\varphi/d\omega$, the envelope time delay. And if the frequencies near the resonance pass through the system with a delay larger than for the rest of the spectrum, long oscillations occur at the resonance at every fast signal transition. The Bessel family of filters offers the best compromise in this respect, but at the expense of a very gradual pass-band to stop-band transition. The Butterworth filter family has a maximally flat magnitude and a steeper stop-band transition, but also a longer ringing in the time domain. Note however that ringing at subsonic frequency (near 10 Hz) will be less problematic than at the high end of the spectrum, mainly because of the cutoff inherently present in any loudspeaker system, even those with very large subwoofers. But also the rest of the amplification chain will have additional AC coupling at various stages to help. So let us calculate the system response with Bessel and Butterworth poles, and then decide which to implement in the system.

The values of Bessel poles for a third order system, normalized to the angular frequency $\omega_\mathrm{n} = 1\,\mathrm{rad}/s$ (see Appendix 4 for details) are:

$$\begin{aligned} s_3 &= -1.0474 - j0.9993 \\ s_2 &= -1.0474 + j0.9993 \\ s_1 &= -1.3227 \end{aligned} \tag{73}$$

However those values are given for a low-pass system. For a high-pass system we must first invert the pole values, $1/s_\mathrm{i}$. Then in order to obtain a 10 Hz cut off we must multiply the inverted poles by $\omega_0 = 2\pi 10\,\mathrm{rad/s}$. We then set C to some convenient value, and solve a third order equation for the values of $R_{11,12,13}$ (see Appendix 4). However, a large C requires low values of R, lowering the input

impedance of the filter, which will attenuate the signal because of the presence of R_5 after the RIAA amplifier. $C = 330\,\text{nF}$ and $R_{11} = 47\,\text{k}\Omega$ represent a good choice.

For the 3rd-order Bessel high pass filter system, the resistors are related as $R_{11} : R_{12} : R_{13} = 1 : 0.7 : 3.9$. By selecting $C = 330\,\text{nF}$, the resistor values are $R_{11} = 47\,\text{k}\Omega$, $R_{12} = 33\,\text{k}\Omega$, $R_{13} = 180\,\text{k}\Omega$. If necessary, R_{12} can be decreased and R_{13} increased in proportion, or the other way round, in order to trim the response just above the cut off frequency. This may or may not be necessary, depending on the Q–factor of the integrator loop and its peaking at very low frequencies.

For the Butterworth response the pole values are:

$$\begin{aligned} s_3 &= -0.5000 - j0.8660 \\ s_2 &= -0.5000 + j0.8660 \\ s_1 &= -1.0000 \end{aligned} \tag{74}$$

Because Butterworth poles are on the unit circle, the inverted values for the high-pass response are the same. Of course the multiplication by $\omega_0 = 2\pi 10\,\text{rad/s}$ is necessary to obtain the 10 Hz cut off. The calculation of component values then follows the same procedure as for the Bessel poles. The values obtained are optimized to $C = 220\,\text{nF}$ and are: $R_{11} = 56\,\text{k}\Omega$, $R_{12} = 22\,\text{k}\Omega$, $R_{13} = 390\,\text{k}\Omega$. See Appendix 4 for details.

In Fig.20 we compare the frequency response of the Bessel and Butterworth 3rd-order high pass filters having a 10 Hz cut off, together with the 1st-order filter resulting from the integrator DC loop correction, which cuts off at about 1 Hz.

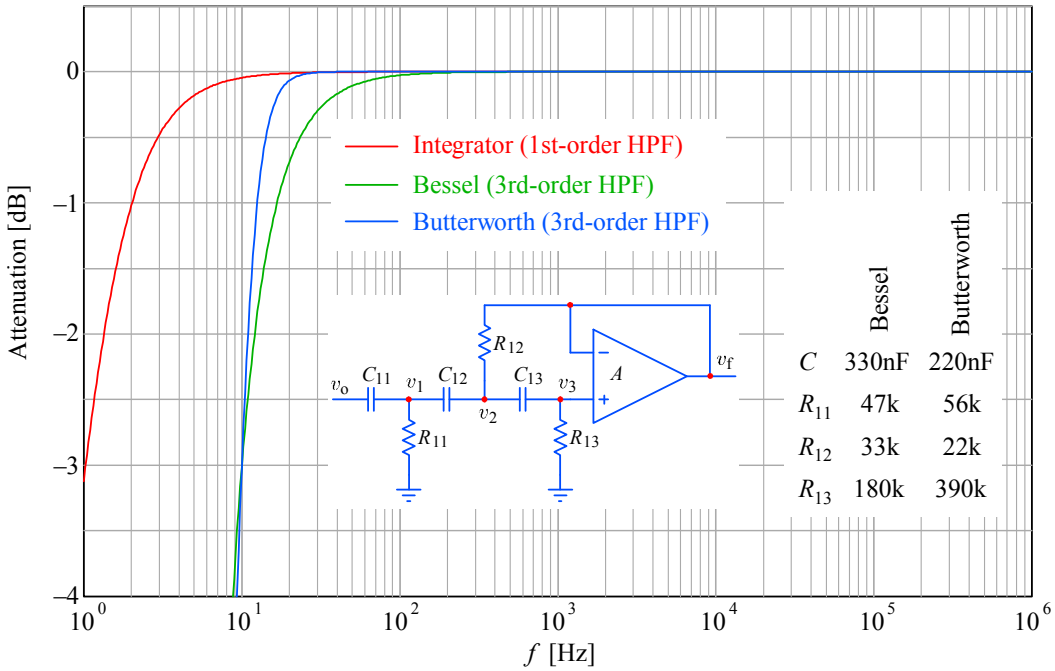


Fig.20: Comparison of the frequency responses of 3rd-order Bessel and Butterworth high pass filters, together with the effective 1 Hz cut off owed to the integrator DC correction loop. The Bessel filter attenuates 16 Hz by about −1 dB, the Butterworth by only −0.2 dB.

Note that the filter input impedance in the pass band is only slightly less than R_{11} alone, because of feedback bootstrapping of $R_{12,13}$. Therefore the loading of the $R_f C_f$ low pass filter (51) remains low and does not alter the total pass band gain.

We are now able to evaluate the flatness of the response of the whole system by simply multiplying all the relevant transfer functions, as in (40), but with the integrator and the high pass filter included. In Fig.14 we have plotted the response at the RIAA equalization amplifier, multiplied by the inverse RIAA function, the response of the cartridge network, and the low pass filter output. Here in Fig.21 we plot the same total response with the influence of the integrator for DC offset removal, and the 3rd-order Butterworth high pass filter at 10 Hz.

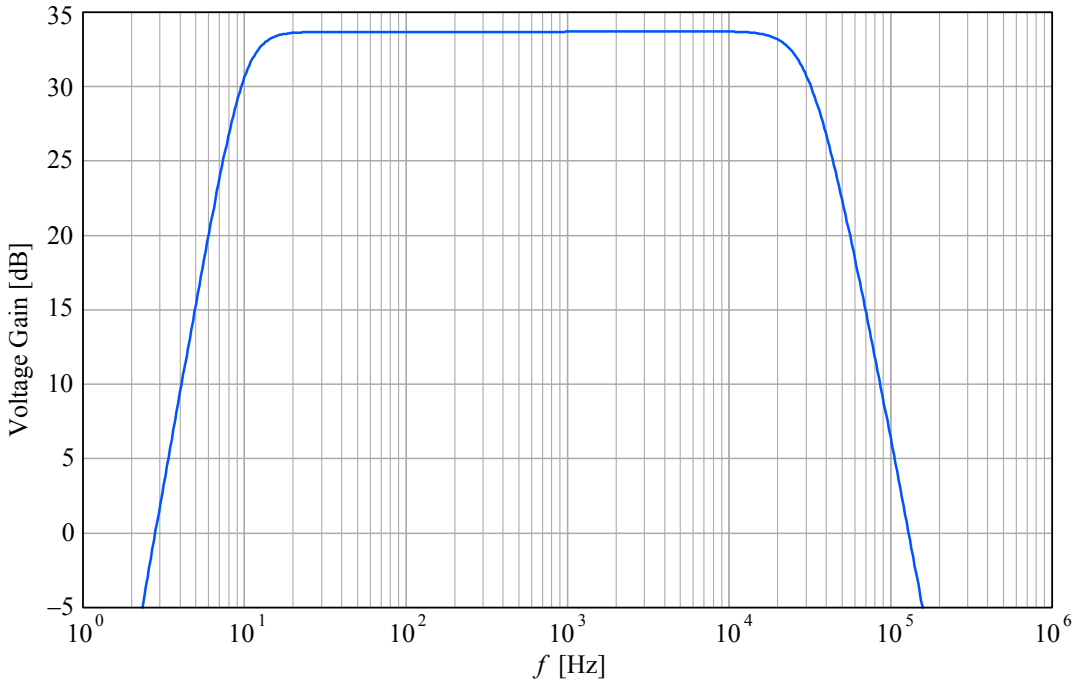


Fig.21: The complete system response, including the DC integration loop, a Butterworth 3rd-order high-pass filter, and the 1st-order low-pass filter at 50 kHz. The high frequency response is dominated by the cartridge's internal electrical impedance and its loading, which is a 2nd-order low-pass system, resulting in a 3rd-order response at high frequencies.

And in Fig.22 we use a vertical scale of 4 dB in total in order to expose the pass-band details. The response is flat within 0.1 dB from 16 Hz to 16 kHz.

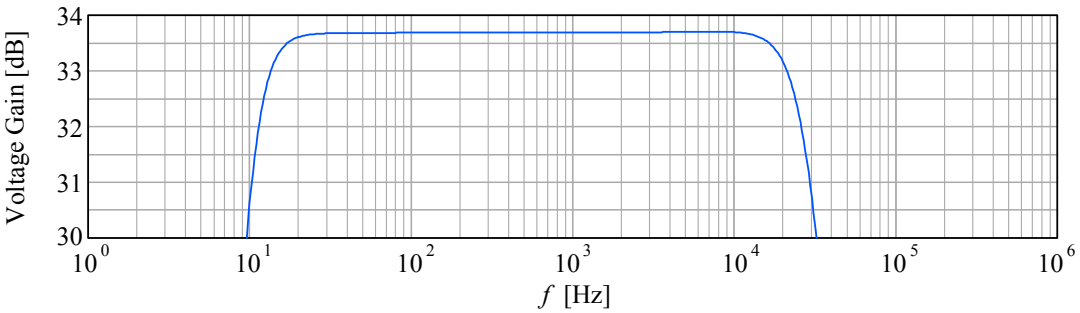


Fig.22: Same as in Fig.21, but with only 4 dB vertical scale to expose the passband details. The response is within 0.2 dB from 16 Hz to 16 kHz, with −3 dB at 10 Hz and 29 kHz.

The theoretically derived response has been checked numerically by using Matlab (by The MathWorks), see the code in Appendix 5. Of course MathCad (by

PTC), or Mathematica (by Wolfram Research), or other tools can be used equally well. It is a good practice to also check the calculations independently by using one of the many circuit simulation computer programmes, some are available freely online. Most of those programmes offer the possibility to vary the components within the expected tolerance range and use the Monte Carlo statistical analysis to evaluate the variations in response.

Readers who want to build the circuit for their own use, or in limited quantities, can measure and select the components to 0.1% tolerance from a bunch of 1% parts using a good RLC-meter. Note however that using a simple hand held $3\frac{1}{2}$ digit multimeter for such measurement will not guarantee a precision better than 0.5% ± 1 digit. Use of at least $4\frac{1}{2}$ digit measuring equipment is recommended.

Noise Analysis

Noise is a random quantum phenomenon and fundamentally unpredictable. However, during the last 150 years a monumental amount of theoretical and experimental work has been accomplished, allowing some clever thermodynamic and statistical approximations to be used, resulting in relatively simple relations by which the circuit noise behaviour can be assessed easily.

In the old days of tubes/valves and later discrete transistor amplifiers the circuit thermal noise had to be optimized by carefully setting the input devices' working point and bias current, because of the different dependence of the noise current and noise voltage with bias. In modern low noise integrated circuits the current noise component is low and the noise voltage dominates. There is not much we can do about the circuit noise besides keeping all the impedances reasonably low.

The choice of the cartridge will probably be made on its mechanical and electrical characteristics, but for the system noise performance we must consider also the sensitivity and the electrical impedance of the cartridge. Thermal noise voltage is proportional to the square root of the real part of the cartridge impedance, its load, and the bandwidth, so low impedance cartridges are preferred. However, the cartridge sensitivity is also a function of its impedance (the number of turns of the coil and the coil resistance), as well as the magnetic field strength. Thus a low impedance cartridge with a good magnet can be as sensitive as a higher impedance one with a poor magnet.

The noise evaluation of the system starts from the well known thermal noise relation of the signal source, the noise voltage being proportional to the square root of the cartridge impedance and bandwidth:

$$v_{\mathrm{nc}} = \sqrt{4k_{\mathrm{B}}TR_{\theta}\Delta f} \tag{75}$$

where:

$k_{\mathrm{B}} = 1.3807{\times}10^{-23}\,\mathrm{VAs/K}$ ..... Boltzmann thermal energy constant
$T = 293\,\mathrm{K}$ ..... nominal ambient absolute temperature
$R_{\theta} = \Re\{Z_{\mathrm{c}}\}$ ..... thermal real part of the impedance
$\Delta f = f_{\mathrm{H}} - f_{\mathrm{L}}$ ..... bandwidth of interest (20 kHz in audio)

But (75) is valid only if R_{θ} is independent of frequency. Because only the real part of the impedance generates thermal noise, and because the impedance varies with

frequency, we must first calculate the equivalent impedance seen by the amplifier's input, and then evaluate its real part within the audio spectrum. This gives the noise voltage spectral density, which is a function of frequency.

The total impedance seen by the amplifier's input is the cartridge internal impedance in parallel with its load impedance:

$$Z_{\mathrm{c}}(j\omega) = \frac{1}{\dfrac{1}{R_{\mathrm{pu}} + j\omega L_{\mathrm{pu}}} + j\omega C_{\mathrm{b}} + \dfrac{1}{R_{\mathrm{b}}}} \tag{76}$$

We want to express this as the sum of its real and imaginary part. We first get rid of the double fractions:

$$Z_{\mathrm{c}}(j\omega) = \frac{R_{\mathrm{b}}(R_{\mathrm{pu}} + j\omega L_{\mathrm{pu}})}{R_{\mathrm{b}} + (R_{\mathrm{pu}} + j\omega L_{\mathrm{pu}})(j\omega C_{\mathrm{b}} R_{\mathrm{b}} + 1)} \tag{77}$$

Multiply the terms in the denominator:

$$Z_{\mathrm{c}}(j\omega) = \frac{R_{\mathrm{b}}(R_{\mathrm{pu}} + j\omega L_{\mathrm{pu}})}{R_{\mathrm{b}} + R_{\mathrm{pu}} - \omega^2 C_{\mathrm{b}} R_{\mathrm{b}} L_{\mathrm{pu}} + j\omega(C_{\mathrm{b}} R_{\mathrm{b}} R_{\mathrm{pu}} + L_{\mathrm{pu}})} \tag{78}$$

Rationalize the denominator by multiplying both the numerator and the denominator by the complex conjugate of the denominator:

$$Z_{\mathrm{c}}(j\omega) = \frac{R_{\mathrm{b}}(R_{\mathrm{pu}} + j\omega L_{\mathrm{pu}})[R_{\mathrm{b}} + R_{\mathrm{pu}} - \omega^2 C_{\mathrm{b}} R_{\mathrm{b}} L_{\mathrm{pu}} - j\omega(C_{\mathrm{b}} R_{\mathrm{b}} R_{\mathrm{pu}} + L_{\mathrm{pu}})]}{(R_{\mathrm{b}} + R_{\mathrm{pu}} - \omega^2 C_{\mathrm{b}} R_{\mathrm{b}} L_{\mathrm{pu}})^2 + \omega^2 (C_{\mathrm{b}} R_{\mathrm{b}} R_{\mathrm{pu}} + L_{\mathrm{pu}})^2} \tag{79}$$

We perform the multiplications in the numerator, cancel the terms with opposite sign, and with some reordering and normalization we group the real and imaginary parts:

$$Z_{\mathrm{c}}(j\omega) = R_{\mathrm{pu}} \frac{1 - \omega^2 C_{\mathrm{b}} L_{\mathrm{pu}} + \frac{R_{\mathrm{pu}}}{R_{\mathrm{b}}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{pu}}}\left(\omega C_{\mathrm{b}} R_{\mathrm{pu}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{b}}}\right) + j\left[\frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{pu}}}(1 - \omega^2 C_{\mathrm{b}} L_{\mathrm{pu}}) - \omega C_{\mathrm{b}} R_{\mathrm{pu}}\right]}{\left(1 - \omega^2 C_{\mathrm{b}} L_{\mathrm{pu}} + \frac{R_{\mathrm{pu}}}{R_{\mathrm{b}}}\right)^2 + \left(\omega C_{\mathrm{b}} R_{\mathrm{pu}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{b}}}\right)^2} \tag{80}$$

The real part of the impedance is then:

$$\Re\{Z_{\mathrm{c}}(j\omega)\} = R_{\mathrm{pu}} \frac{1 - \omega^2 C_{\mathrm{b}} L_{\mathrm{pu}} + \frac{R_{\mathrm{pu}}}{R_{\mathrm{b}}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{pu}}}\left(\omega C_{\mathrm{b}} R_{\mathrm{pu}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{b}}}\right)}{\left(1 - \omega^2 C_{\mathrm{b}} L_{\mathrm{pu}} + \frac{R_{\mathrm{pu}}}{R_{\mathrm{b}}}\right)^2 + \left(\omega C_{\mathrm{b}} R_{\mathrm{pu}} + \frac{\omega L_{\mathrm{pu}}}{R_{\mathrm{b}}}\right)^2} \tag{81}$$

Fig.23 shows the real part of the input impedance as the function of frequency, after (81) in comparison with the absolute value of (80). At low frequencies the value is approximately R_{pu} (600 Ω), then rises owed to the inductance L_{pu} to nearly R_{b} (47 kΩ) at the resonant frequency $1/2\pi\sqrt{C_{\mathrm{b}} L_{\mathrm{pu}}}$, and finally falls back owed to C_{b}.

Because of this, the voltage noise spectral density follows a similar function:

$$e_{\mathrm{n}Z} = \frac{v_{\mathrm{nc}}}{\sqrt{\Delta f}} = \sqrt{4 k_{\mathrm{B}} T \Re\{Z_{\mathrm{c}}\}} \tag{82}$$

and with the assumed values for an average cartridge and its nominal load the spectrum is shown in Fig.24. It goes from about 3 to about 27 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at resonance.

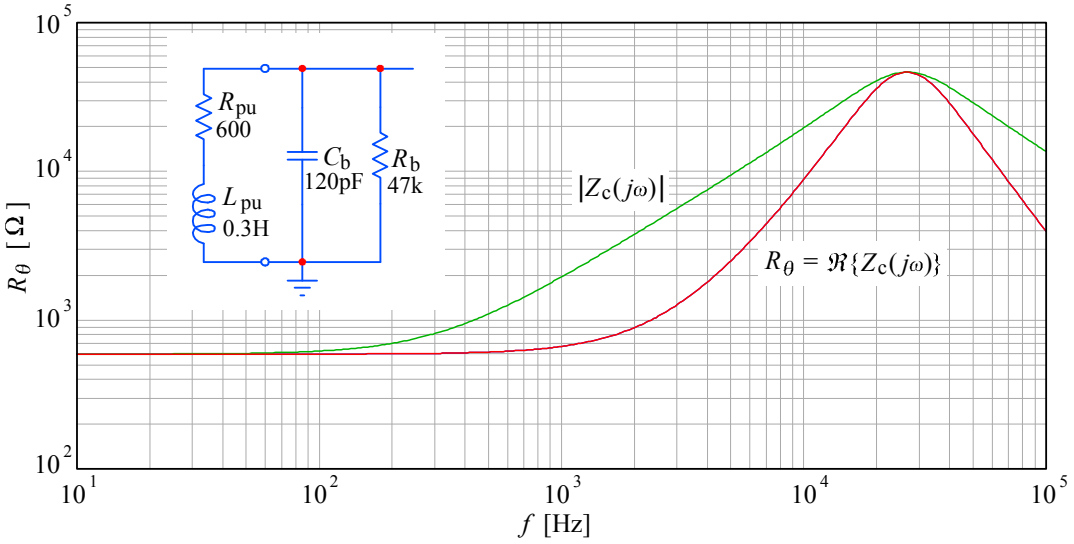


Fig.23: The real part of the input impedance compared to the absolute value.

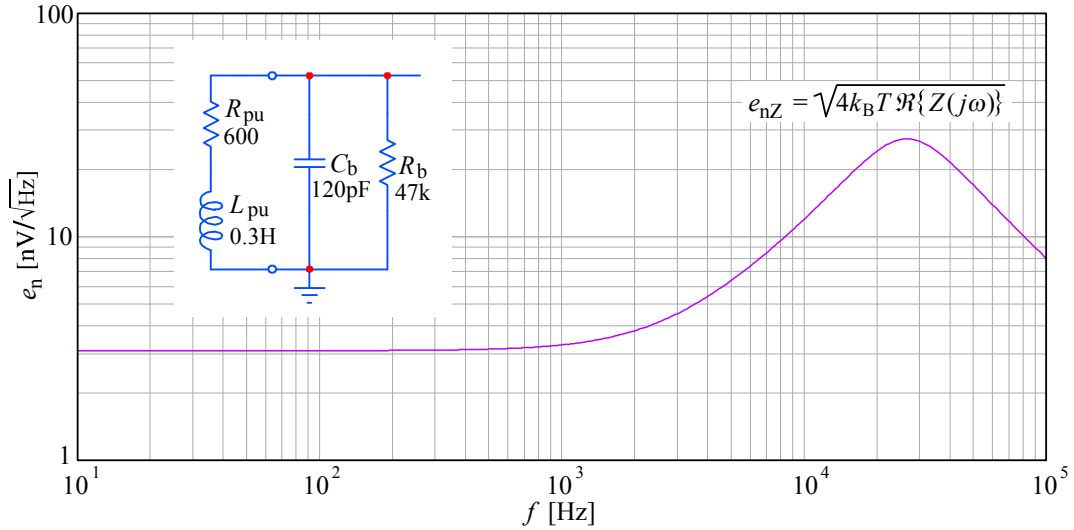


Fig.24: Thermal noise voltage spectral density is proportional to $\sqrt{\Re\{Z_c(j\omega)\}}$, (82).

The thermal noise of the RIAA feedback network must also be taken into account, however the feedback thermal impedance is equal to the real part of (39) with R_4 added, and all in parallel with R_3. Because of the low value of R_3 the thermal noise of the feedback network may be approximated by R_3 alone within the audio band (it falls eventually to $R_3||R_4$ beyond 50 kHz):

$$e_{nR} = \sqrt{4\,k_B T R_3} = 1.8\,\text{nV}/\sqrt{\text{Hz}} \tag{83}$$

In addition to the thermal noise of the signal source impedance and its load, the amplifier's input voltage noise and current must be taken into account.

The amplifier's input noise current must be multiplied by the absolute value of the input impedance (not juts the real part!) in order to obtain the equivalent noise voltage, therefore the absolute value of equation (80) must be used for $|Z_c|$:

$$e_{niZ} = i_n\,|Z_c(j\omega)| \tag{84}$$

Likewise, a similar input noise current flows through the feedback network, so it should be multiplied by R_3 (simplified!) to obtain the equivalent noise voltage:

$$e_{\mathrm{n}iR} = i_{\mathrm{n}}\, R_3 \tag{85}$$

However, modern integrated circuit amplifiers have very low input noise current, so even if Z_{c} increases with frequency the equivalent voltage noise is relatively low. As an example, the good old NE5534 has a typical input noise current 2.5 $\mathrm{pA}/\sqrt{\mathrm{Hz}}$ at 30 Hz, and 0.6 $\mathrm{pA}/\sqrt{\mathrm{Hz}}$ at 1 kHz and above. On the other hand its input noise voltage is about 7 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 30 Hz, and 4 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 1 kHz. To develop a comparable voltage noise, the input noise current should flow through a resistance of 8 kΩ. Therefore the input noise current will be significant only above some 4 kHz. Similar, maybe slightly better results will be obtained with the OP37.

But check the OPA627, OPA637, or the latest OPA827, the voltage noise is similar, $e_{\mathrm{n}A} = 4\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$, but the current noise is 500× lower, about 2.2 $\mathrm{fA}/\sqrt{\mathrm{Hz}}$. This makes the influence of the current noise insignificant, more so because the different noise components are not correlated, and must be added by power:

$$e_{\mathrm{nTi}} = \sqrt{e_{\mathrm{n}Z}^2 + e_{\mathrm{n}R}^2 + e_{\mathrm{n}i}^2 + e_{\mathrm{n}A}^2} \tag{86}$$

Note that both the amplifier input noise current and input noise voltage increase towards very low frequencies (owed to the so called $1/f$ noise component) and towards high frequencies (because of the excess noise component), and both influences are outside the audio band. But because RIAA equalization enhances low frequencies, and because some amplifiers have a relatively high corner of the $1/f$ spectrum, it might be important to model the amplifier voltage noise more accurately. With f_{c} being the corner frequency below which the noise starts to increase from the level at 1 kHz, the spectrum is modeled by:

$$e_{\mathrm{n}A} = e_{\mathrm{n}A}(1\,\mathrm{kHz})\frac{f_{\mathrm{c}} + f}{f} \tag{87}$$

The corner frequency must be found by adjustment from amplifier's data, since most manufacturers specify only the effective noise at 1 kHz and 30 Hz. Usually f_{c} will be between 10 and 100 Hz; here we use $f_{\mathrm{c}} = 50$ Hz. In addition, excess noise increases the amplifier noise in proportion to $\sqrt{f}$, but its corner frequency is often at 10 kHz or above, thus outside of the audio range.

For the output noise spectrum, we multiply the spectrum of the total input noise e_{nTi} by system gain $G_3(s)$ (17) and low pass (51) and high pass (71) filters:

$$e_{\mathrm{nTo}} = e_{\mathrm{nTi}}\, G_3(s)\, J_1(s)\, J_{3\mathrm{B}}(s) \tag{88}$$

Since RIAA equalization decreases with increasing frequency, the output noise will be substantially lower. The output spectral noise density is shown in Fig.25, along with the individual input components and the total input noise spectral density.

To obtain the effective (root-mean-square, rms) noise voltage form the spectral density, the output spectral power (voltage squared!) density function must be integrated in frequency within the audio band, from $f_{\mathrm{L}} = 16$ Hz to $f_{\mathrm{H}} = 20$ kHz, and then the square root is taken to obtain the effective noise voltage:

$$v_{no}^2 = \int_{f_L}^{f_H} e_{no}^2 \, df \tag{89}$$

$$\overline{v_{no}} = \sqrt{v_{no}^2} \tag{90}$$

Using the OPA827 data, the effective output noise voltage is about 71 μV_{rms}. Considering the nominal input signal level of 5 mV with the gain of about 50× gives 250 mV, so the signal to noise ratio is about 71 dB. By using a better amplifier with the input voltage noise density of just $2\,nV/\sqrt{Hz}$, about 74 dB can be achieved, and further amplifier noise reduction will have no effect. This is because the input noise is dominated by the cartridge thermal noise. To achieve low noise a low impedance and high sensitivity cartridge is always preferred.

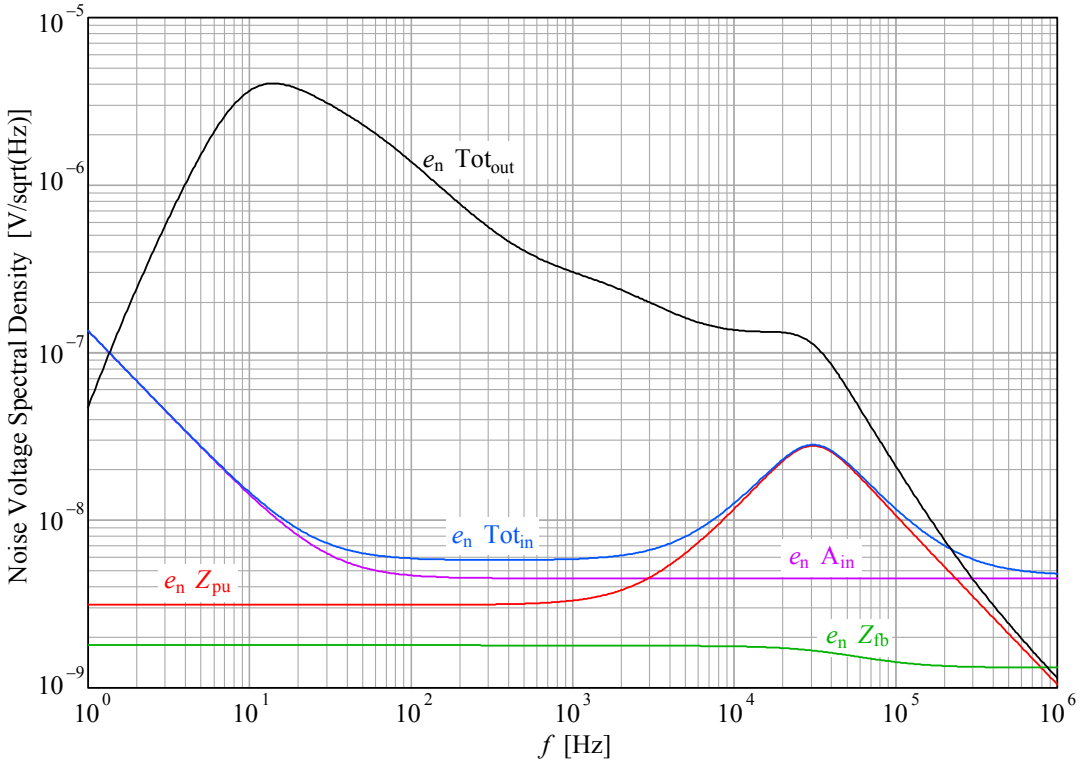


Fig.25: Noise spectral density: e_{nZpu} is the thermal noise of the real part of the impedance of the cartridge and its load; e_{nAin} is the amplifier input voltage noise. The amplifier current noise is not shown because it is far too low. e_{nR_3} is the feedback network thermal impedance, simplified to just R_3. The total input equivalent noise density is denoted by e_{nTi}. The output noise density e_{nTout} is equal to e_{nTi} multiplied by the gain $G_3(s)$, the low-pass filter $J_1(s)$ and the high-pass filter $J_3(s)$.

For example, the Grado Prestige Silver or Gold series can take the noise down by about 8 dB, mainly owed to its very low inductance ($L_{pu} = 40$ mH) and high sensitivity (5 mV at 5 cm/s at 1 kHz). With such a cartridge it makes sense using amplifiers with less than $2\,nV/\sqrt{Hz}$, achieving $v_{no} = 26\,\mu$Vrms, and a S/N ratio of almost 80 dB (see the example in Fig.27 and Fig.28).

Note also that the actual dynamic range will often be greater than the S/N ratio by at least 10 dB with high quality recordings. But note also that the vinyl noise will dominate, even with a poor cartridge and amplifier. Sadly, in too many cases with high

quality vinyl the analog tape recorder hiss will clearly be heard well above the vinyl noise! Therefore for a really low noise consider a CD instead of the LP.

See Appendix 5 for the numerical integration of the noise spectrum in Matlab code.

A Couple of Interesting Variations

The conventional circuit configurations (Fig.3,4,5) are not ideal, but are good compromises. In the past, with the amplifier's open loop gain and bandwidth low, the input impedance of the amplifier would present an additional load to the signal source, in particular with circuits of Fig.4,5 where the feedback varies with frequency. Modern amplifiers have adequate gain and bandwidth, so this became a non-issue.

Nevertheless, non-inverting circuit configurations may suffer from input impedance asymmetry problems, since the feedback impedance is usually low and nearly constant ($\sim R_3$), whilst the signal input impedance varies widely. The amplifier differential input actually senses a very small differential signal floating on a relatively large common mode signal, and under such conditions the impedance asymmetry may result in a certain degree of nonlinearity. In this application the input signal causing the common mode error is small, about $25\,\mathrm{mV_{rms}}$ nominal at $10\,\mathrm{kHz}$, but the purists among us may have their objections anyway. However, the impedance asymmetry allows another distortion 'mechanism' to appear, and it gets worse in phonograph preamplifiers because the signal increases with frequency. Since the cartridge impedance also increases with frequency, the input transistor internal capacitance (either from collector to base in bipolar junction transistors or from drain to gate in field effect transistors) changes nonlinearly with the signal, and the Miller effect (see Appendix 2) further amplifies this by the gain of the input stage, resulting in rather high intermodulation distortion (see Appendix 3 for details).

Such errors can be reduced by using a very high impedance current source to provide bias for the input differential stage, and the cascode configuration to get rid of the Miller effect at the input, but this does not eliminate the distortion completely.

The only way to get rid of common mode and dynamic nonlinearity errors is to use an inverting amplifier configuration, but ordinary inverting amplifiers have noise problems, because the input impedance is large ($47\,\mathrm{k\Omega}$) throughout the spectrum, not just at the cartridge resonance, and to provide enough inverting gain the feedback impedance must be suitably larger, and large impedance means large thermal noise and large amplifier current noise components.

But if the signal source can be made floating, Fig.26 shows an attractive inverting amplifier configuration having the advantages of both the non-inverting and inverting configuration, and disadvantages of none. The only problem with the floating source is that cartridges usually have their grounding and shielding tied to the left channel ground. This means that a rewiring of both the cartridge and the tone arm is necessary. Not many serious audiophiles will be prepared to perform such a delicate operation on their expensive equipment (with a possible exception of a few complete lunatics, like myself;o).

Otherwise, the circuit components are the same in both cases, so there will be no change in the design philosophy. If you are apt using a sharp scalpel, and have a fine soldering iron, and other precision tools, a steady hand and good eye sight, you

might give it a try. In my case it worked perfectly, but to be honest I did not notice any dramatic changes in sound quality, most probably because the distortion of the basic circuit was already well below the distortion on the vinyl.

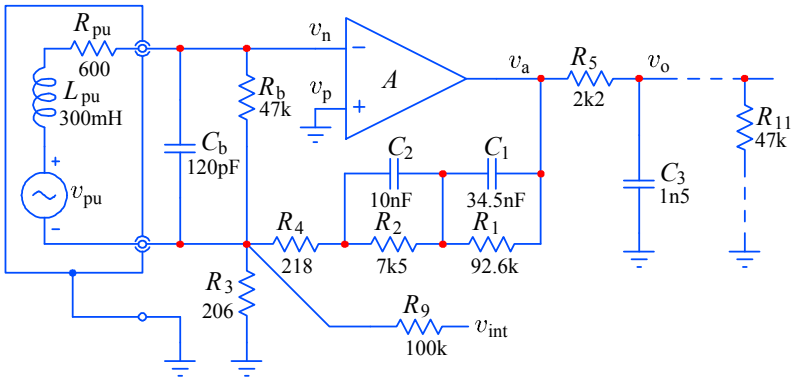


Fig.26: Inverting configuration employing a floating signal source reduces input stage distortion.

For highest quality cartridges with low internal impedance and high sensitivity there is another circuit version is worth considering. It uses discrete jFETs for the front-end, and ICs for the rest of the circuitry, as shown below.

Today's best IC audio amplifiers in terms of both noise and distortion are probably the LME series (developed by National Semiconductor, now owned by Texas Instruments). In particular, the LME49990 offers $0.9\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ input voltage noise, $1.8\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ input current noise, along with a distortion figure with an impressive number of zeros before the first significant digit. If your favourite cartridge has high electrical impedance, the current noise component (as well as the input bias current) may still be too high for you, so the already mentioned OPA827, or the OPA1641, or the LME49880, all with a jFET input, and consequently negligible current noise, will be the natural choice, even if their voltage noise is slightly higher, about 4–5 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ between 100 Hz and 10 kHz.

Fig.27 shows a way around this dilemma if we consider using discrete or integrated pairs of jFETs for the amplifier input stage.

The LSK170 is a single, and LSK389 a dual low noise jFET (both produced by Linear Integrated Systems; the devices are production process improved versions of the well known 2SK170 and 2SK389), with the input voltage noise of $0.9\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 1 kHz, and $2.5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 10 Hz, and the input current noise down into the $\mathrm{fA}/\sqrt{\mathrm{Hz}}$ level. A differential pair has a total noise voltage greater by $\sqrt{2}$.

With these jFETs used to form the input differential amplifier stage, their output can be connected to an ordinary NE5534 in place of its own input stage, itself being disabled by connecting the input pins 2 and 3 to the negative supply rail. The input to the second stage is available at pins 1 and 8 of the NE5534, which are ordinarily used for offset correction.

So we have a composite amplifier with extremely low input noise, very low distortion, a 600 Ω load drive capability, and the circuit is easy to build. The connection, shown in Fig.27, appeared in a data sheet catalogue of Siliconix jFETs in the late 1970s, using a 2N5911 dual jFET at that time.

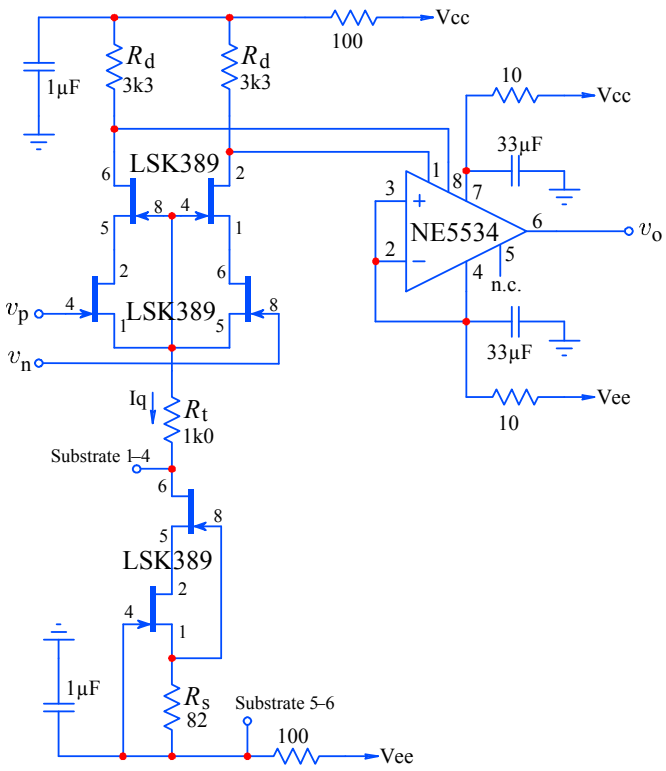


Fig.27: Composite jFET and bipolar amplifier.

Fig.28 shows the noise performance of the circuit in Fig.27, using the Grado Prestige Silver cartridge. Compare that to Fig.25, where the noise of a OPA827 opamp with our typical average cartridge is shown.

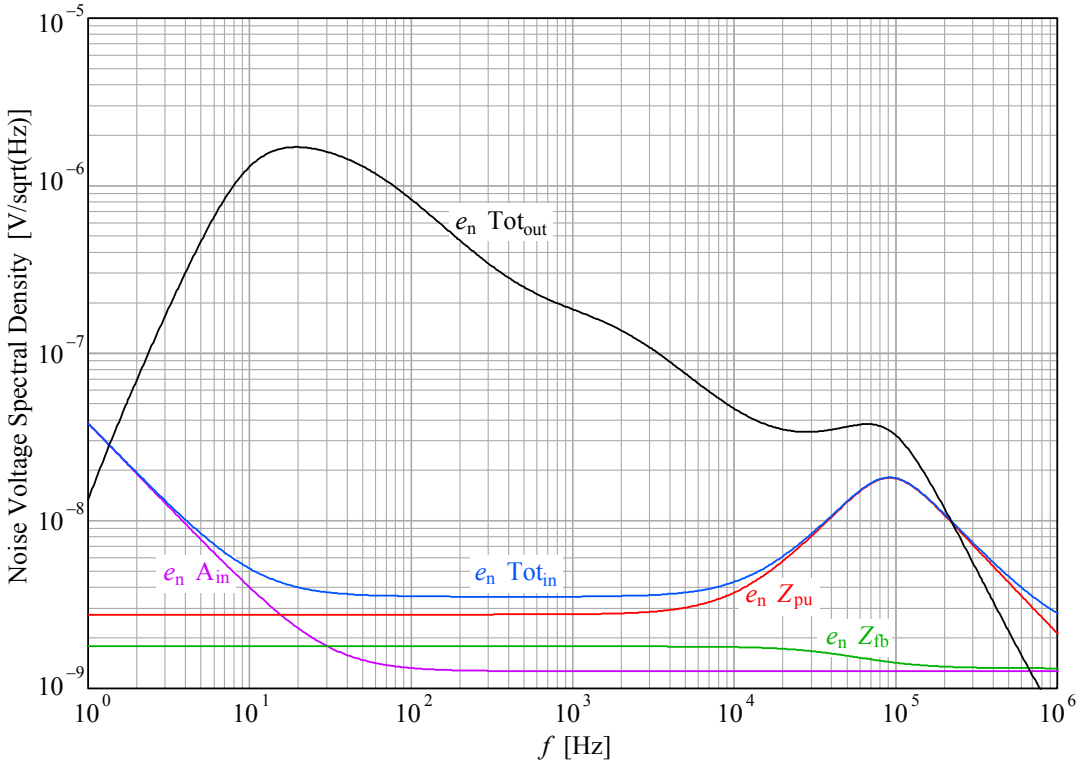


Fig.28: Noise is much lower using a Grado Prestige Silver with a composite amplifier with jFETs.

Note that jFETs, in contrast to bipolar transistors, exhibit their optimal noise performance at a relatively high drain current, about 2 mA according to LSK389 specifications. This means that drain resistors R_d of relatively low value must be added externally, since the internal resistors are 12 kΩ (as found in NE5534 specifications). The differential amplifier bias current I_q is set by the constant current source, itself realized by another pair of cascode connected jFETs with a source resistor R_s which can be adjusted for optimal bias current. And instead of a simple differential pair, a differential cascode can be used to further reduce distortion and increase the input impedance and bandwidth. See Appendix 3 for details.

The LSK389 pair is relatively well matched, so the input differential stage will have a good common mode rejection. Nevertheless, as with any jFET differential stage, the input offset can be as large as 10 mV, so a proper DC reduction as described before is indispensable. Likewise, proper power supply decoupling is obligatory, given the low signal levels. Also in the case of using LSK389, there are two substrate pins on the chip, which must be connected to a suitable negative point in the circuit. Note that the LSK389 have relatively low V_{gsoff}, typically in the ~300 mV range, so a rather low resistor value is needed for R_s to obtain a 1 mA current per FET.

As shown in Fig.28, with such an amplifier very low input voltage noise can be achieved, lower than the thermal noise of the ~200 Ohm of R_3. Such an amplifier is ideal for low impedance cartridges, like the Grado Prestige Silver ($L_{pu} = 40$ mH, $R_{pu} = 475\,\Omega$). Low inductance requires lower C_b for proper compensation and a low R_b for resonance damping. So with $C_b = 120$ pF (which includes a 100 pF cable) and $R_b = 27$ kΩ the impedance peaking is low and is shifted up away from the critical 2–5 kHz region where the human ear is most sensitive. The resulting output noise voltage is around 22 μV_{rms}, giving a S/N ratio of 81 dB.

Fig.29 shows a further advantage of the low signal source impedance: the very high bandwidth extension up to 60 kHz.

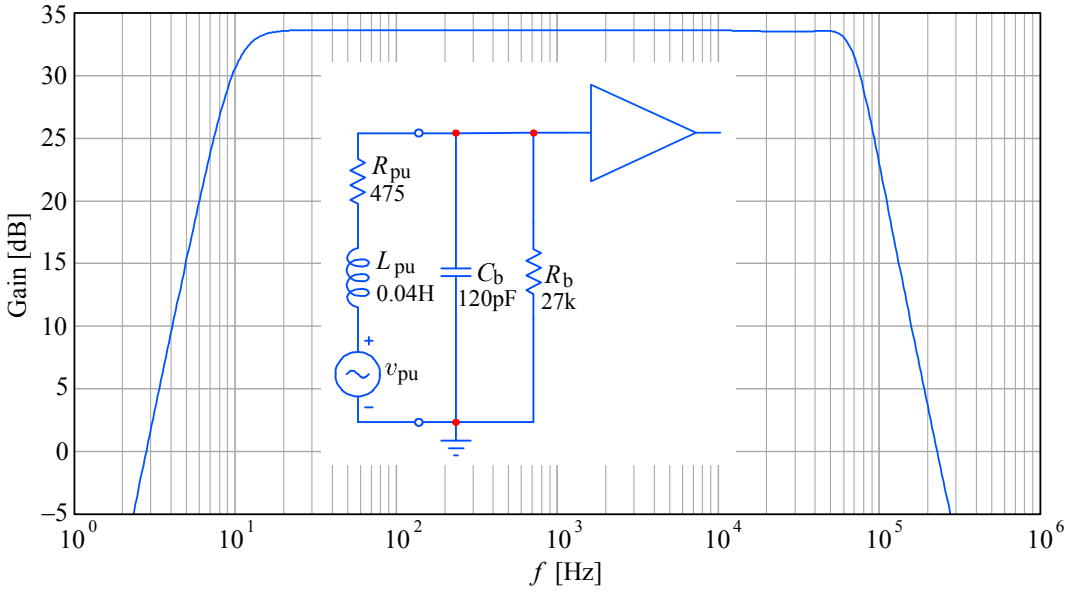


Fig.29: A cartridge like Grado Prestige Silver with low inductance extends the bandwidth to 60 kHz.

To conclude, readers who like to experiment, or want to evaluate different cartridges and system setups, might want to build this adaptable circuit, shown in

Fig.30, which uses a triple switch to configure the circuit for accepting either a floating or a grounded signal source, as well as vary the input impedance to suit different cartridges.

The input impedance components C_b and R_b are broken into a fixed (C_{b0}, R_{b0}) and variable part (C_{bv}, R_{bv}), the latter preferably adjustable in steps by jumpers or DIL binary switches. Note that the cable capacitance will be slightly higher in the grounded configuration, compared to the floating configuration, because of the cable shield, but with a low capacitance cable of 100 pF/m, and a 50-60 cm length, it will be manageable.

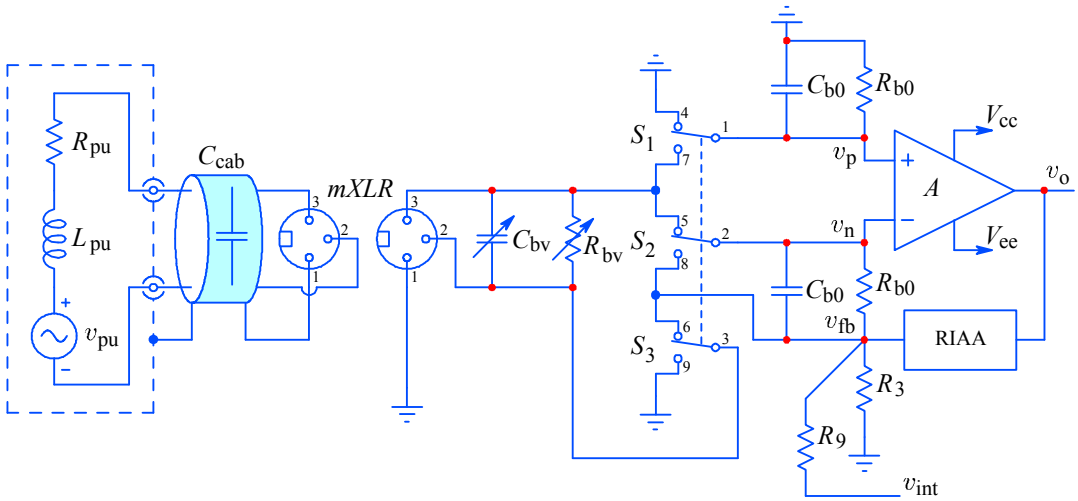


Fig.30: A triple switch allows using a floating or a grounded signal source.

A differential output amplifier may be in interesting addition for professional audio use. Differential signaling is used to avoid ground loops and cancel any supply currents flowing between various ground points in form of a common mode signal. Here is a simple circuit which converts a single-ended input from the high pass filter into a differential signal. The circuit uses either an LME49725 or an OPA1632 fully differential amplifier, both capable of driving a 600 Ω line. The only difference between the two is the connection of the chip thermal pad, see their data sheet. The LME may be preferred because its quiescent current is lower, only 5 mA. But other similar amplifiers can be used as well. All that is required at the receiving end is a differential to single ended converter in form of a usual opamp differential amplifier.

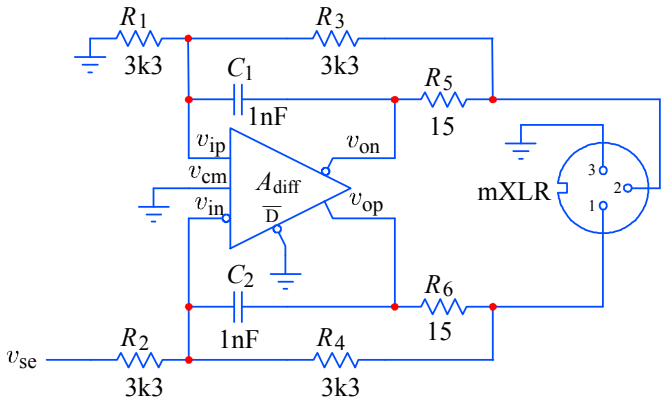


Fig.31: Differential output amplifier; can be LME49725, OPA1632, or similar.

Appendix 1: A Zero-Pole Network for the Encoding Circuit Model

In circuit theory the transfer function of a network is expressed in a form of a ratio of two polynomials, with the complex frequency s as the independent variable. The transfer function zero is the particular value of the independent variable at which the value of the numerator polynomial is zero. Similarly, the transfer function pole is that particular value of the independent variable for which the value of the denominator polynomial is zero (and hence the transfer function value is $\pm\infty$). In general, zeros are associated with a phase lead, and poles with a phase lag. Fig.A1.1 shows a simple network suitable for exemplary analysis.

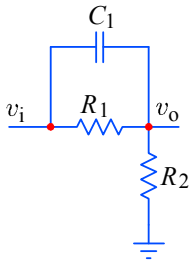


Fig.A1.1: Zero-Pole Network

If the voltage at the input v_i is non-zero, there will be a current through the network. In accordance with the Kirchhoff law the sum of the currents in the node v_o will be zero, so we can equate the incoming and outgoing current:

$$\frac{v_\mathrm{i} - v_\mathrm{o}}{\dfrac{1}{sC_1 + \dfrac{1}{R_1}}} = \frac{v_\mathrm{o}}{R_2} \tag{A1.1}$$

We want to solve this for the ratio $v_\mathrm{o}/v_\mathrm{i}$, which is the network transfer function. So we first try to get rid of the multiple fractions:

$$\frac{v_\mathrm{i} - v_\mathrm{o}}{\dfrac{R_1}{sC_1R_1 + 1}} = \frac{v_\mathrm{o}}{R_2} \tag{A1.2}$$

$$(v_\mathrm{i} - v_\mathrm{o})(sC_1R_1 + 1) = v_\mathrm{o}\frac{R_1}{R_2} \tag{A1.3}$$

We separate the voltages:

$$v_\mathrm{o}\left(sC_1R_1 + 1 + \frac{R_1}{R_2}\right) = v_\mathrm{i}(sC_1R_1 + 1) \tag{A1.4}$$

and we obtain the transfer function:

$$\frac{v_\mathrm{o}}{v_\mathrm{i}} = \frac{sC_1R_1 + 1}{sC_1R_1 + 1 + \dfrac{R_1}{R_2}} \tag{A1.5}$$

The transfer function is a ratio of two polynomials of the first order in s. But from (A1.5) it is difficult to get an insight, we want an expression like the general form:

$$F(s) = A\frac{s-z}{s-p}$$

Here A is a frequency independent gain or attenuation, z is the zero, and p is the pole. We want to arrange the expression (A1.5) in a similar way. So we divide both the numerator and the denominator by C_1R_1:

$$\frac{v_\mathrm{o}}{v_\mathrm{i}} = \frac{s + \dfrac{1}{C_1R_1}}{s + \dfrac{1}{C_1R_1}\left(1 + \dfrac{R_1}{R_2}\right)} \tag{A1.6}$$

Then the zero is:

$$z = -\frac{1}{C_1R_1} \tag{A1.7}$$

and the pole is:

$$p = -\frac{1}{C_1R_1}\left(1 + \frac{R_1}{R_2}\right) \tag{A1.8}$$

Alternatively, the time constant corresponding to the zero is:

$$\tau_\mathrm{z} = C_1R_1 \tag{A1.9}$$

and the time constant corresponding to the pole is:

$$\tau_\mathrm{p} = C_1R_1\frac{R_2}{R_1+R_2} \tag{A1.10}$$

We first choose the value of the capacitor because capacitors are available in fewer values and tolerance. We select:

$$C_1 = 33\,\mathrm{nF} \tag{A1.11}$$

From (A1.9) we obtain R_1:

$$R_1 = 96455\,\Omega \tag{A1.12}$$

Then from the (A1.10) we obtain R_2:

$$R_2 = \frac{\tau_\mathrm{p}R_1}{C_1R_1 - \tau_\mathrm{p}} = R_1\frac{\tau_\mathrm{p}}{\tau_\mathrm{z} - \tau_\mathrm{p}} \tag{A1.13}$$

The value of R_2 is:

$$R_2 = 10717\,\Omega \tag{A1.14}$$

Note that the resistance ratio is exactly:

$$R_1/R_2 = 9.0 \tag{A1.15}$$

because from (A1.9) and (A1.10) $\tau_\mathrm{z}/\tau_\mathrm{p} = 3183\,\mu\mathrm{s}/318.3\,\mu\mathrm{s} = (R_1+R_2)/R_2$.

In a similar manner we calculate the second network values for the $\tau_z = 75\,\mu s$ and $\tau_p = 3.18\,\mu s$ time constants. By selecting C_2:

$$C_2 = 10\,nF \tag{A1.16}$$

the values of resistors obtained are:

$$R_3 = 7500\,\Omega \tag{A1.17}$$

$$R_4 = 332.418\,\Omega \tag{A1.18}$$

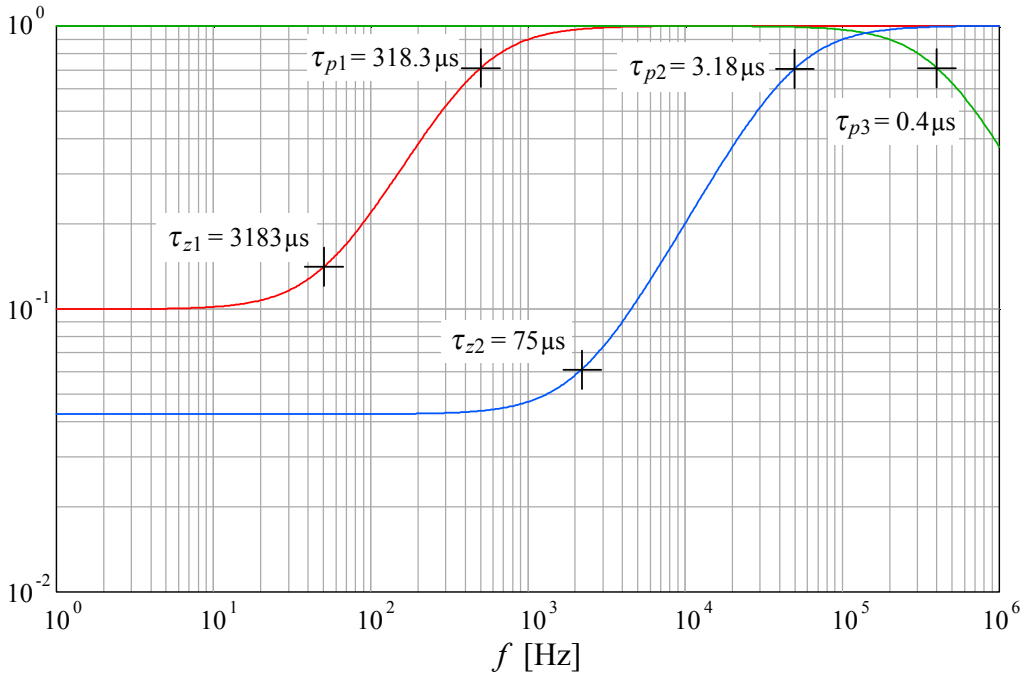


Fig.A1.2: The responses of the zero-pole network of Fig.A1.1 for the two characteristic time constant pairs (red and blue), with the (optional) high frequency pole of the cutter driving amplifier (green). The combined responses give the realistic RIAA encoding curve.

In accordance with the relations above, we can model the RIAA encoding standard to serve as the reference circuit for checking the accuracy of the equalization network. In a circuit simulator programme we can simply use two networks as in Fig.A1.1 and separate them by dependent ViVo generators (voltage in, voltage out). Fig.A1.3 shows the configuration (with optional R_5 and C_3). Note that the gain of the E_1 dependent generator is set equal to the second stage attenuation at DC, $A = (1 + R_3/R_4) = 23.562$, bringing the system gain at 1kHz very close to unity.

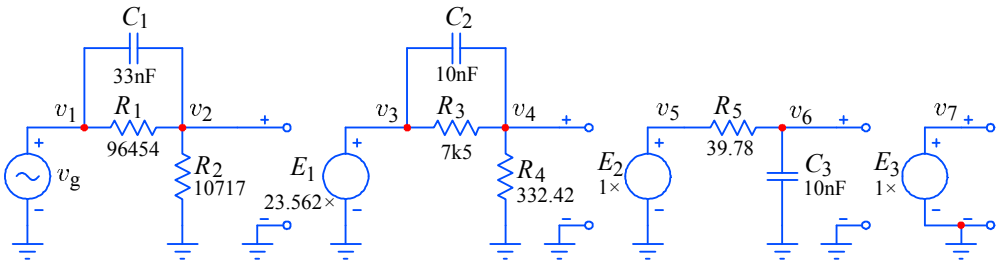


Fig.A1.3: Model of the RIAA encode network for circuit simulation testing.

With the circuit in Fig.A1.3 an input signal with a 5 mV amplitude simulates the standard 5 cm/s groove modulation reference level. In Fig.A1.4 the combined frequency response of all the sections is shown:

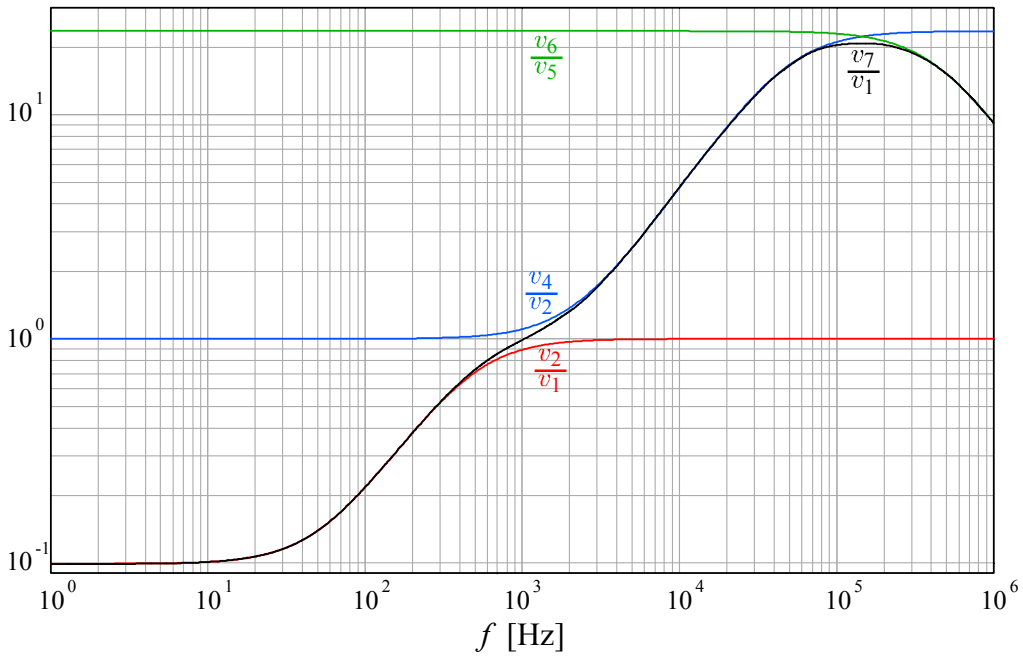


Fig.A1.4: Frequency response of the model of the RIAA encode network of Fig.A1.3.

However, to test an actual RIAA amplifier in hardware we need a completely passive realization. Fig.A1.5 shows the configuration which provides a response identical to the response above. The circuit attenuates the signal v_g by about 200× at 1kHz, so 1 V (0.5 V at v_1) gives 5 mV at v_4. The circuit has a low output impedance at v_3, going from 100 Ω at l.f. to 76 Ω at h.f., so thermal noise will be low (<0.26 μV_{rms} for 50 kHz bandwidth), and the cartridge thermal noise will be dominant. A suitable cartridge impedance (R_{pu} and L_{pu}) can be added for a realistic test. Subtract 76 Ω from R_{pu} for correct high frequency damping. L_{pu} must be wound on a ferrite pot core, use a pot with a small air gap (~0.1 mm) for good h.f. response.

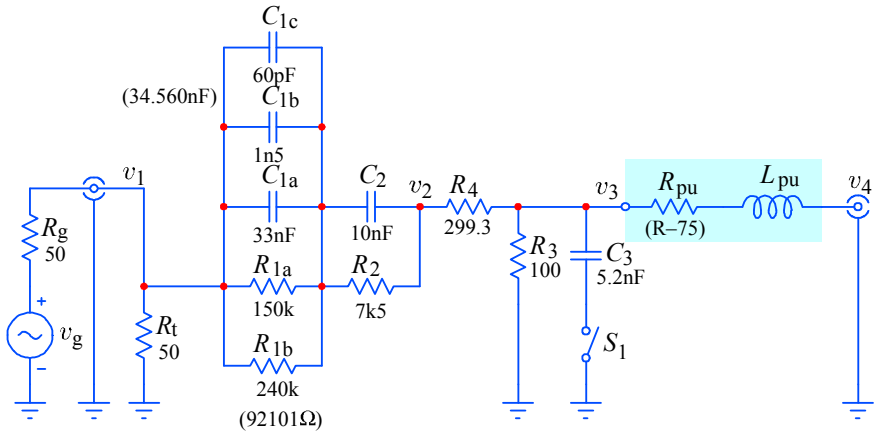


Fig.A1.5: Model of the passive RIAA encode network for testing an actual equalizer. $R_g||R_t = 25\,\Omega$ was subtracted from R_4 to get a correct response at high frequencies. C_3 optionally limits the bandwidth at 400 kHz. Attenuation at 1 kHz is about 200×. Optionally, R_{pu} and L_{pu} model the cartridge impedance.

Appendix 2: Amplifier Gain and Bandwidth Analysis

We want to justify certain simplifications used in the circuit analysis regarding the difference between the real and ideal amplifier performance. Fig.A2.1 shows an amplifier with a differential input and some negative feedback (part of the output voltage fed to the inverting input):

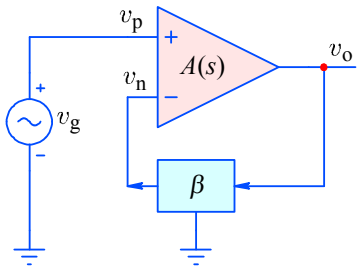


Fig.A2.1: A general differential amplifier model with feedback.

We express the output voltage v_o of a differential amplifier as a function of the input voltage difference $v_p - v_n$, and the gain A, itself a function of the complex frequency s:

$$v_o = (v_p - v_n)A(s) \tag{A2.1}$$

The gain $A(s)$ at DC (or at very low frequencies) has a value A_0:

$$A(0) = A_0 \tag{A2.2}$$

At higher frequencies there is a cut off at a frequency where the output power is only $1/2$ of the power at very low frequencies. Since power is proportional to voltage squared, for the output voltage this represents a $1/\sqrt{2}$ ratio (this is often being referred to as the -3 dB point, which is allowed to be expressed in dB only if the load impedance is independent of frequency, which is seldom the case).

The frequency at which this cut off occurs is determined by the dominant pole s_d. For simple amplifiers the pole is real and negative:

$$s_d = -\omega_d = -2\pi f_d \tag{A2.3}$$

This is because we associate the negative half of the complex plane with passive components (energy loss), in contrast with the positive half, which represents active components (energy generation). In actual amplifiers the dominant pole is often determined internally by the effective input resistance of the second amplifying stage, and the total capacitance at the same node:

$$\omega_d = \frac{1}{RC} \tag{A2.4}$$

The capacitance is either a sum of stray capacitances of both interconnected stages or is intentionally increased for stability reasons; more on stability later.

If all other poles are well above the frequency at which the amplifier ceases to amplify, they will not influence the amplifier passband behaviour significantly. As is

shown in Fig.A2.2, internally the amplifier can be modeled by three stages, with the first stage being a differential transconductance (current output) stage, the second is a voltage gain stage and the third stage is a unity gain power buffer.

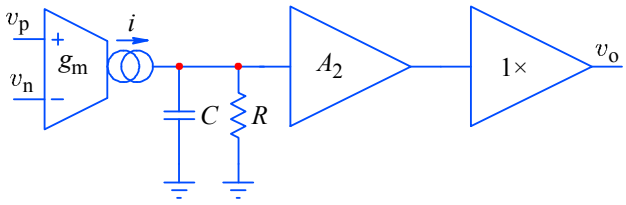


Fig.A2.2: Internal amplifier model. The differential transconductance (current output) stage drives the voltage gain stage with its input impedance modeled by the parallel RC network, and the output is buffered by a unity gain power stage.

In accordance with this model, the total gain at DC must be a product of the transconductance g_m by R, and the voltage gain of the second stage:

$$A_0 = g_m R \cdot A_2 \tag{A2.5}$$

whilst the frequency dependence is owed to the current $i = g_m(v_p - v_n)$ driving the parallel RC network:

$$v_{RC} = i\frac{1}{sC + \frac{1}{R}} = i\frac{\frac{1}{C}}{s + \frac{1}{CR}} = iR\frac{\frac{1}{CR}}{s + \frac{1}{CR}} \tag{A2.6}$$

Then the frequency dependence of the gain may be modeled mathematically by a Cauchy polynomial of the first order:

$$F(s) = \frac{1}{s - s_d} \tag{A2.7}$$

where $s_d = -\omega_d = -2\pi f_d = -1/RC$ is the dominant pole (there are other poles at much higher frequencies, often above the transition frequency f_T, which is the frequency where the amplifier gain falls below unity, $|A(\omega_T)| = 1$).

But the DC gain of expression (A2.7) varies with s_d, and we want the DC gain to be set by A_0 alone, whilst the frequency dependent part is determined by s_d. We therefore normalize this expression by first evaluating it at DC ($s = 0$):

$$F(0) = \frac{1}{0 - s_d} = \frac{1}{-s_d} \tag{A2.8}$$

and then take the ratio $F(s)/F(0)$:

$$\frac{F(s)}{F(0)} = \frac{\frac{1}{s - s_d}}{\frac{1}{-s_d}} = \frac{-s_d}{s - s_d} \tag{A2.9}$$

As a result we have the same expression as in equation (A2.6).

Now we can express the gain $A(s)$ by the frequency independent part A_0 and the frequency dependent polynomial in s:

$$A(s) = A_0 \frac{-s_\mathrm{d}}{s - s_\mathrm{d}} \tag{A2.10}$$

Therefore the output voltage can be written as:

$$v_\mathrm{o} = (v_\mathrm{p} - v_\mathrm{n}) A_0 \frac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}} \tag{A2.11}$$

Note that this relation holds regardless of whether we are dealing with an amplifier with no feedback (open loop) or with some negative feedback (closed loop). The amplifier's closed loop gain depends also on the feedback attenuation factor β, by which the output voltage is attenuated and fed to the inverting input:

$$v_\mathrm{n} = \beta\, v_\mathrm{o} \tag{A2.12}$$

By inserting this back into (A2.11) we have:

$$v_\mathrm{o} = (v_\mathrm{p} - \beta\, v_\mathrm{o}) A_0 \frac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}} \tag{A2.13}$$

By separating the voltage variables:

$$v_\mathrm{o} \left(1 + \beta\, A_0 \frac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}} \right) = v_\mathrm{p} A_0 \frac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}} \tag{A2.14}$$

we can define the closed loop transfer function:

$$\frac{v_\mathrm{o}}{v_\mathrm{p}} = \frac{A_0 \dfrac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}}}{1 + \beta\, A_0 \dfrac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}}} \tag{A2.15}$$

We get rid of the double fractions:

$$\frac{v_\mathrm{o}}{v_\mathrm{p}} = A_0 \frac{\omega_\mathrm{d}}{s + \omega_\mathrm{d}(1 + \beta\, A_0)} \tag{A2.16}$$

and make the coefficients in the numerator and the denominator equal:

$$\frac{v_\mathrm{o}}{v_\mathrm{p}} = \frac{A_0}{1 + \beta\, A_0} \cdot \frac{\omega_\mathrm{d}(1 + \beta\, A_0)}{s + \omega_\mathrm{d}(1 + \beta\, A_0)} \tag{A2.17}$$

and we have the definition for the system DC gain and its frequency dependence. By dividing the DC gain by A_0 we have the following form, from which it will be easier to recognize how the system response is affected by A_0:

$$\frac{v_\mathrm{o}}{v_\mathrm{p}} = \frac{1}{\dfrac{1}{A_0} + \beta} \cdot \frac{\omega_\mathrm{d}(1 + \beta\, A_0)}{s + \omega_\mathrm{d}(1 + \beta\, A_0)} \tag{A2.18}$$

With A_0 very high, say, 10^5 or higher, the system cut off frequency is also very high:

$$\omega_\mathrm{h} = \omega_\mathrm{d}(1 + \beta\, A_0) \tag{A2.19}$$

and the system closed loop gain factor becomes:

$$A_c = \frac{1}{\frac{1}{A_0} + \beta} \approx \frac{1}{\beta} \tag{A2.20}$$

Fig.A2.3 shows the relations between the open and closed loop gain responses and the loop feedback factor.

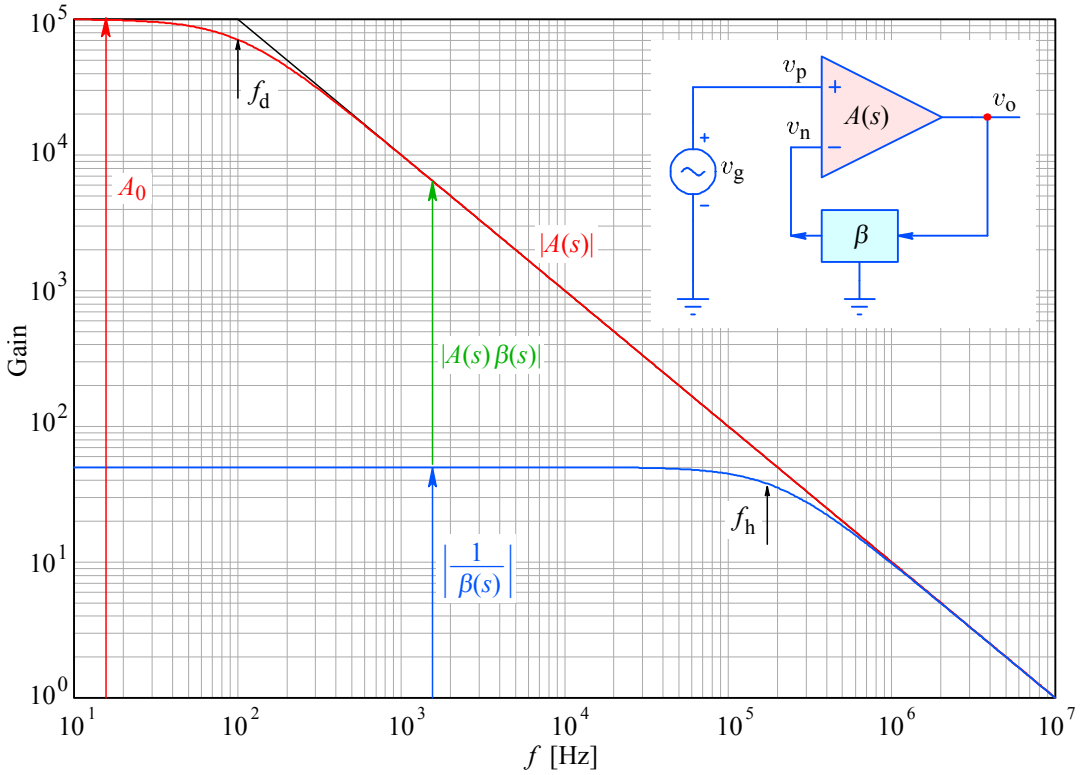


Fig.A2.3: Open loop and closed loop ($\beta = 1/50$) gain and frequency response comparisons.

Indeed, equations (A2.19) and (A2.20) tell us that the closed loop system gain will be closely approximated by the factor $1/\beta$ over a very wide frequency range if the open loop gain is high enough. Alternatively, the same is true if the dominant pole can be made high, however, this is a technological issue with many limiting factors.

There are of course secondary effects, and one of the most important is the slew rate limiting, which was the most severe distortion mechanism in early integrated circuits, and even in discrete power amplifiers. This mechanism is not immediately obvious from the internal amplifier configuration of Fig.A2.2, yet it is fundamentally limited by that configuration. Fig.A2.4 shows a conventional input differential stage employing bipolar transistors. What is not explicitly shown in Fig.A2.2 is that the maximum current available from the g_m stage is limited, $|i_1 - i_2| \leq i_0$. The reason for this is the configuration of the input differential amplifier and its requirements. To understand this we must take a close look at Fig.A2.4.

The transistors Q_1 and Q_2 form the differential amplifier. Its quiescent current bias is provided by the constant current source made by Q_3 and Q_4. The transistors Q_5 and Q_6 form a current mirror, which provides symmetrical output current to the second stage, which consists of the Q_7 and a current source i_2.

Q_4 provides strong feedback around Q_3, thus keeping its collector current constant, and independent of the common mode signal $(v_p + v_n)/2$. Because of this, the impedance seen by the emitters of Q_1 and Q_2 is very high, allowing the emitters to share the bias current seamlessly, thus keeping low the error of the difference of the collector currents, $i_{c1} - i_{c2}$. This difference drives the second stage amplifier Q_7.

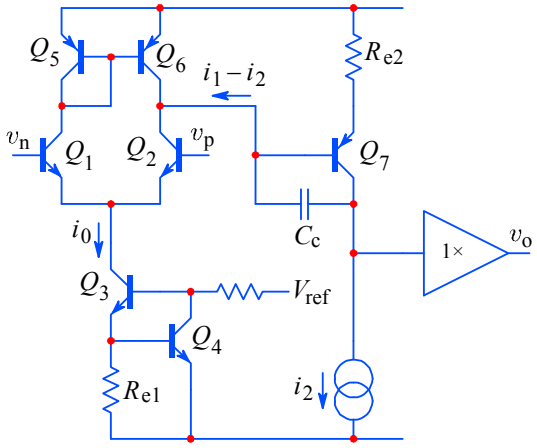


Fig.A2.4: A simplified schematic diagramme of a conventional differential amplifier.

Note that C_c is in the feedback of Q_7, which inverts the signal. For every mV of accumulated charge on the base side there are some -200 mV on the collector side of the capacitor C_c, and all this charge must be supplied by Δi_c. Therefore the **Miller effect** will increase the effective value of the loading capacitance by the Q_7 gain (often between 200× and 500×).

It is obvious that constant current is necessary for low differential error. But it is also obvious that the capacitance C_c slows the response. If the input voltage difference changes too quickly, the output voltage will not follow as quickly, and the result is momentary loss of feedback and saturation of the input differential amplifier. In such conditions one transistor of the $Q_{1,2}$ pair will be fully open and the other closed, thus the current difference becomes equal to i_0 and as a result the C_c will cause the voltage at the base of Q_7 to increase linearly with time, instead of following the input signal variations.

The frequency at which the output cannot follow the input depends on the input signal amplitude, as we are going to show. But the most important question is: can an RIAA amplifier (gain decreasing with increasing frequency!) suffer from slew rate problems?

The input resistance at the base of Q_7 is equal to the emitter resistance R_{e2} multiplied by the Q_7 gain $(A_2 + 1)$ (A_2 is equal to the Q_7 current gain):

$$R_b = R_{e2}(A_2 + 1) \tag{A2.21}$$

The Q_7 collector voltage is equal to v_o, and if the stage gain is A_2, the Q_7 base voltage must be (Q_7 inverts the signal!):

$$v_b = \frac{-v_o}{A_2} \tag{A2.22}$$

For the difference current of frequency s we can write:

$$\Delta i_{\mathrm{c}}(s) = \frac{v_{\mathrm{b}}}{R_{\mathrm{b}}} + \frac{v_{\mathrm{b}} - v_{\mathrm{o}}}{\frac{1}{sC_{\mathrm{c}}}} \tag{A2.23}$$

If we assume $R_{\mathrm{e2}} = 100\,\Omega$ and $A_2 = 200$, the effective base resistance R_{b} will be about 20 kΩ. With a power supply of 15 V, the maximum output voltage v_{o} will be close to some 13 V, and with $A_2 = 200$ the base voltage v_{b} would change by only 0.065 V. The base current will then be $i_{\mathrm{b}} = v_{\mathrm{b}}/R_{\mathrm{b}} \approx 3\,\mu\mathrm{A}$. Now the input differential stage will have minimum noise if biased by i_0 of about 200 µA, and $\Delta i_{\mathrm{cmax}} = i_0$. This means that the majority of input current will flow through the compensating capacitor C_{c}, and the influence of R_{b} can be neglected. So we can approximate (A2.23) with:

$$i_0 = \frac{v_{\mathrm{b}} + v_{\mathrm{b}} A_2}{\frac{1}{sC_{\mathrm{c}}}} \tag{A2.24}$$

First we equate the voltages:

$$\frac{i_0}{sC_{\mathrm{c}}} = v_{\mathrm{b}}(1 + A_2) \tag{A2.25}$$

and we can express the input admittance:

$$\frac{i_0}{v_{\mathrm{b}}} = sC_{\mathrm{c}}(1 + A_2) \tag{A2.26}$$

Obviously, the effective input admittance is being increased by the system gain factor. This increase was first explained by *John M. Miller* (*J.M. Miller*, Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit, Scientific Papers of the Bureau of Standards, 15(351):367–385, USA, 1920), so the effect and the equivalent capacitance C_{M} are named in his honour:

$$C_{\mathrm{M}} = C_{\mathrm{c}}(1 + A_2) \tag{A2.27}$$

Since the frequency operator is equivalent to a time differential, $s = j(d/dt)$, we can also write this as:

$$\frac{i_0}{C_{\mathrm{c}}(1 + A_2)} = \frac{dv_{\mathrm{b}}}{dt} \tag{A2.28}$$

The derivative dv_{b}/dt is the slew rate at the base of Q_7. Of course, the output slew rate dv_{o}/dt is greater A_2 times:

$$\frac{dv_{\mathrm{o}}}{dt} = -A_2 \frac{v_{\mathrm{b}}}{dt} \tag{A2.29}$$

By assuming a sine wave of amplitude V and frequency ω, the output signal is:

$$v_{\mathrm{o}} = V \sin\omega t \tag{A2.30}$$

The maximal slope of such a signal is obtained by a time differentiation:

$$\frac{dv_{\mathrm{o}}}{dt} = V\omega \cos\omega t \tag{A2.31}$$

and since the cosine function has a maximum of 1 at $t = 0$, the slew rate is both amplitude and frequency dependent:

$$\left|\frac{dv_o}{dt}\right|_{max} = V\omega \qquad \text{(A2.32)}$$

If we assume the total gain $A_0 = 10^5$ distributed as $A_1 = 500\times$ and $A_2 = 200\times$, the value of the compensation capacitance $C_c = 22\,\text{pF}$, and the available maximum differential current $i_0 = 200\,\mu\text{A}$, the output slew rate limit will be a generous $9\,\text{V}/\mu\text{s}$. In contrast, an ordinary RIAA amplifier does not have to put out more than $3.5\,V_{peak}$ (+20 dBm) at 20 kHz, which is equivalent to $0.44\,\text{V}/\mu\text{s}$.

Clearly, in normal operation an RIAA amplifier will never get near the limit, but for the slowest of classical operational amplifiers, like the µA741 ($0.5\,\text{V}/\mu\text{s}$).

Also, in contrast to amplifiers with purely resistive feedback, the RIAA equalization amplifier of Fig.4 will have a feedback factor high and nearly constant with frequency, as illustrated in Fig.A2.5, up to the highest frequencies of interest, because of the gain falling with frequency.

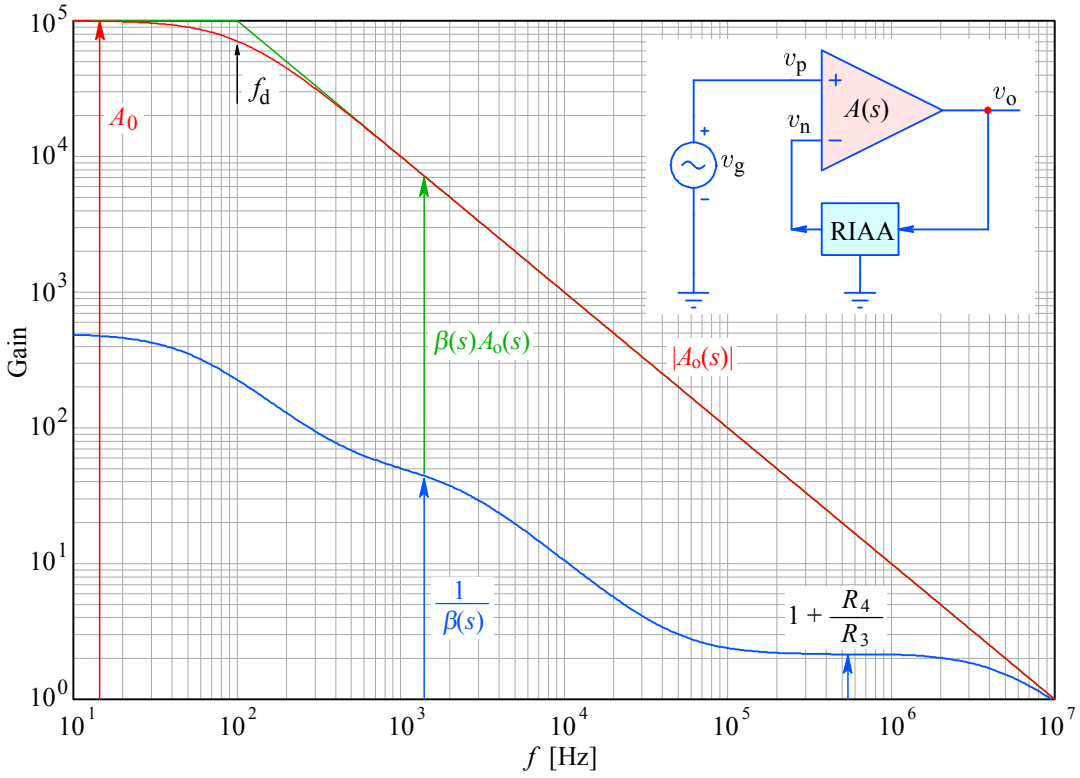


Fig.A2.5: Because of the frequency dependent feedback the RIAA equalization amplifier has a feedback factor $A\beta$ high and changing very little with frequency (in contrast to an amplifier with purely resistive feedback, Fig.A2.3, where the feedback factor is falling at a rate of $-20\,\text{dB}/\text{decade}$, becoming only $\sqrt{2}$ at the system cut off frequency f_h).

Because the system linearity is governed by the amount of feedback, having more feedback at higher frequencies means lower distortion. Of course, other distortion mechanisms may (in suboptimal designs) spoil the performance, but having enough feedback is always beneficial.

Audiophiles still often engage in endless discussions on why and how does feedback reduce distortion. Why is feedback more effective for low frequencies, why it seems to be more effective with certain distortion mechanisms and less with others, what are its limitations, and how does it sometimes make the problem worse.

The short answer is that an amplifier with a closed loop feedback amplifies any externally generated signals, but reduces any internally generated signal. Signals within the feedback loop, be it intentionally introduced signals, or a consequence of transfer function nonlinearity, or noise, are being all reduced by the feedback factor, that is what remains of the open loop gain when we take away the closed loop gain. But this is only a general rule of thumb, and to see how exactly the system behaves in some particular cases we need to do some modeling and model analysis.

A simplified amplifier distortion mechanism model is shown in Fig.A2.6. An ordinary amplifier is usually composed of an input differential voltage gain stage A, followed by a unity gain power buffer stage, with the negative feedback loop closed by the resistive divider having an attenuation of β. Most of the distortion is usually generated by the output buffer in various forms: crossover distortion, current gain nonlinearity, gain difference between the output transistor pair, etc. All these mechanisms can be included in the model by adding a summing stage Σ, where some arbitrary distortion signal v_d is summed with the signal v_a from a distortionless voltage amplifier, and that sum appears as v_o from a distortionless output stage.

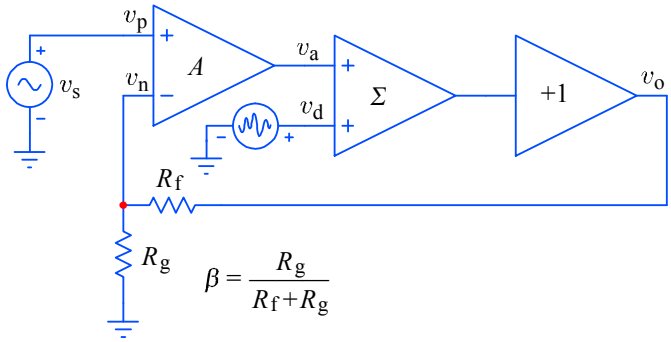


Fig.A2.6: Simplified amplifier distortion model: a distortionless differential input stage with the gain A drives a distortionless summing stage Σ, followed by a distortionless unity gain buffer $(+1)$. A resistive divider closes the feedback loop, dividing the output signal by β. Any internal nonlinearity and noise is modeled by a generator producing the distortion signal v_d.

In this way the distortion as part of the output signal returns to the input by the feedback loop attenuated by β, and then amplified by $-A$ (phase inverted!), where it becomes part of the amplified signal v_a, which is again summed by the distortion signal v_d, and so forth. In this way the distortion signal is being cumulatively reduced by the feedback loop factor $(1 + A\beta)$. Following the input $(v_p = v_s)$ signal path in Fig.A2.6 we can write for the differential voltage gain stage:

$$v_a = A(v_s - v_n) \tag{A2.33}$$

and for the output voltage:

$$v_o = v_a + v_d \tag{A2.34}$$

The feedback attenuates the output voltage to v_n which closes the loop:

$$v_\mathrm{n} = \beta\, v_\mathrm{o} \tag{A2.35}$$

By inserting (A2.35) into (A2.33) we have:

$$v_\mathrm{a} = A(v_\mathrm{s} - \beta\, v_\mathrm{o}) \tag{A2.36}$$

and taking this result into (A2.34) we have:

$$v_\mathrm{o} = A(v_\mathrm{s} - \beta\, v_\mathrm{o}) + v_\mathrm{d} \tag{A2.37}$$

This we solve for v_o, first by multiplying each part in the parentheses by A:

$$v_\mathrm{o} = Av_\mathrm{s} - A\beta\, v_\mathrm{o} + v_\mathrm{d} \tag{A2.38}$$

then transporting the factors containing v_o on the left hand side:

$$v_\mathrm{o} + A\beta\, v_\mathrm{o} = Av_\mathrm{s} + v_\mathrm{d} \tag{A2.39}$$

and extracting the common v_o factor:

$$v_\mathrm{o}(1 + A\beta\,) = Av_\mathrm{s} + v_\mathrm{d} \tag{A2.40}$$

and finally by dividing all by the feedback factor $(1 + A\beta)$:

$$v_\mathrm{o} = \frac{A}{1 + A\beta} v_\mathrm{s} + \frac{1}{1 + A\beta} v_\mathrm{d} \tag{A2.41}$$

We can now rewrite the term multiplying the input signal v_s in the following way:

$$\frac{A}{1 + A\beta} = \frac{A\beta}{1 + A\beta} \cdot \frac{1}{\beta} \tag{A2.42}$$

By making A very high the term $A\beta/(1 + A\beta) \to 1$. This can be thought of as a small error term $(1 - \varepsilon)$ of the closed loop gain $1/\beta$. So in the end we have the signal v_s amplified by the closed loop gain $1/\beta$ (set by the resistor ratio), and in addition the distortion signal which is reduced by a factor $(1 + A\beta)$:

$$v_\mathrm{o} = \frac{1}{\beta} v_\mathrm{s} + \frac{1}{1 + A\beta} v_\mathrm{d} \tag{A2.43}$$

or explicitly:

$$v_\mathrm{o} = \frac{R_\mathrm{f} + R_\mathrm{g}}{R_\mathrm{g}} v_\mathrm{s} + \frac{1}{1 + A\dfrac{R_\mathrm{g}}{R_\mathrm{f} + R_\mathrm{g}}} v_\mathrm{d} \tag{A2.44}$$

For example, by making $R_\mathrm{f} = 9R_\mathrm{g}$ and $A = 10^5$, the closed loop signal gain will be very close to $10\times$, and the distortion will be reduced by as much as 10^{-4}.

Of course, the trouble is that A is not constant, but decreases with frequency, as we have seen before. So instead of just maximizing A it is often a better idea to increase the frequency of the dominant pole from the usual 10–100 Hz to 1 kHz, or higher if possible; this then provides more loop gain at high frequencies and allows the amplifier to react faster to any error. Of course, in RIAA correction the gain decreases with frequency too, leaving more loop gain for correction. As seen in Fig.A2.5, at 1 kH the loop gain can be about 500× and at 10 kHz still about 100×.

Appendix 3: A Discussion on Input jFETs

The book "Designing with Field effect Transistors" is a classical reference, published in 1981, written by the application staff of Siliconix, Inc., and edited by one of the corporation founders and the application department leader Arthur D. Evans, [9]. The following analysis is not a substitute for it, or any other book; I only present some of the important points in a didactic way for a quick reader's insight.

As a general rule, at the amplifier's input a bipolar junction transistor (BJT) differential pair will be preferred for low impedance signal sources, and a junction field effect transistor (jFET) differential pair will be preferred for high impedance signal sources. The dominantly inductive (within the audio band) nature of a phonograph cartridge means that there is a low impedance at low frequencies and a high impedance at high frequencies. Since this impedance starts to increase from about 1 kHz, and the noise power also increases with frequency, and the human ear is most sensitive in the range between 2–5 kHz, the system noise should be made low within this region.

Another reason for choosing jFETs at the input is that their performance with respect to noise becomes better at higher drain current, which also makes it easier to obtain a higher gain, higher bandwidth, and higher slew rate limit, thus also a lower intermodulation distortion. In addition their larger V_{gs} (compared to $k_B T/q \approx 26\,\text{mV}$ of a BJT) means larger signal handling, better linearity, and a more gradual saturation. In contrast, BJTs have usually their optimal noise performance at a relatively small collector current, compromising mostly the slew rate.

On the negative side, matching of a JFET pair is usually worse than for BJTs, meaning a much higher DC offset. And because of their high quiescent current (desirable for noise) a lower drain resistance is necessary for adapting to the following amplifier stage, which means a somewhat lower system gain than with BJTs.

Another point of concern might be the influence of the parasitic gate-drain capacitance C_{gd}, mainly at frequencies in the upper audio band, because of the high source impedance. Also, C_{gd} varies nonlinearly with the signal, causing larger intermodulation distortion at higher frequencies. This can be most problematic when the total system gain is inadequately low, since it falls with frequency anyway.

But let us look at the equations relating the most important jFET parameters. We start from the 'static' transfer function, the drain current as a function of the gate-source voltage, $I_d(V_{gs})$, shown in Fig.A3.1. An N-channel jFET is biased in a way similar to a thermionic triode, a positive drain voltage V_{dd}, and a negative gate voltage V_{gs} (obviously, for P-channel jFETs the voltage and current polarities change).

The gate is in the form of a reverse biased diode with the channel, and the resulting field modulates the effective channel width, and consequently the current through the channel. Like in a thermionic triode, the current through the device, I_d, has a maximal value, I_{dss}, when $V_{gs} = 0$, and $I_d = 0$ when $V_{gs} = V_p$, the so called pinch-off voltage (the gate voltage which causes a field strong enough to deplete the channel of charge carriers, so there can be no current; in manufacturing datasheet V_p is ordinarily labeled as V_{gsoff}; in literature a V_T label can often be found, meaning a conducting threshold voltage).

However, unlike a triode, the modulation of the channel width also changes the parasitic gate-drain capacitance C_{gd} and this change is nonlinear, so that will have some unwanted effects, which will be discussed a little later.

The similarity with a thermionic triode extends also to the transfer function law: The drain current varies in an exponential manner with the gate-source voltage, which can be written as:

$$I_d = I_{dss}\left(1 - \frac{V_{gs}}{V_p}\right)^n \qquad \text{(A3.1)}$$

The exponent can be experimentally found to be $n \approx 2$, so the law is approximately quadratic. In our analysis we shall assume for simplicity an exact quadratic function, but the results would not be very different with an actual value.

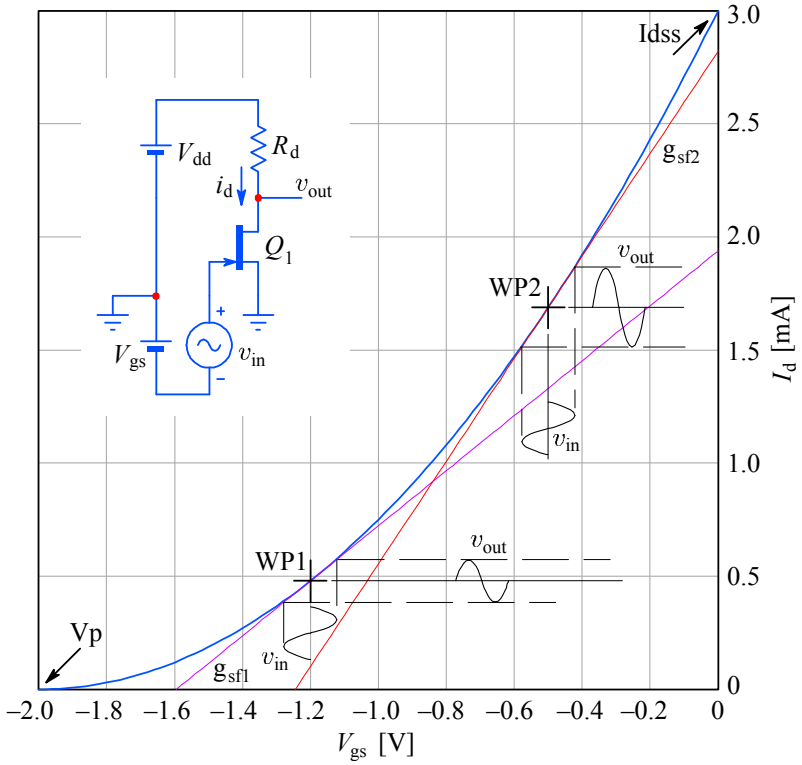


Fig.A3.1: The jFET drain current as a function of gate-source voltage is approximately quadratic. The dynamic small transconductance is bias dependent and is greater at lower gate voltage and higher drain current.

We are interested in the small signal gain of a jFET, and therefore we need a derivative of the function (A3.1) at some particular DC bias setting ('working point'). The common source forward transconductance g_{sf} is found by first choosing some DC V_{gs} resulting in some DC I_d, and we introduce a small AC signal about that DC level, with the amplitude $v_{in} = \Delta V_{gs}$, and consequently $v_{out} = R_d \Delta I_d$. To maximize the dynamic range of this simple amplifier we set the value of R_d such that $R_d I_d = V_{dd}/2$.

$$\frac{\frac{v_{out}}{R_d}}{v_{in}} = g_{sf} = \left.\frac{\Delta I_d}{\Delta V_{gs}}\right|_{V_{gs},I_d} \qquad \text{(A3.2)}$$

We need to separate the DC and AC variables. To make this easier we shall rewrite equation (A3.1) as:

$$I_\mathrm{d} = \frac{I_\mathrm{dss}}{V_\mathrm{p}^2}(V_\mathrm{p} - V_\mathrm{gs})^2 = \frac{I_\mathrm{dss}}{V_\mathrm{p}^2}\left(V_\mathrm{p}^2 - 2V_\mathrm{p}V_\mathrm{gs} + V_\mathrm{gs}^2\right) \tag{A3.3}$$

and further:

$$I_\mathrm{d} = I_\mathrm{dss} - 2\frac{I_\mathrm{dss}}{V_\mathrm{p}}V_\mathrm{gs} + \frac{I_\mathrm{dss}}{V_\mathrm{p}^2}V_\mathrm{gs}^2 \tag{A3.4}$$

As we make the current difference and the voltage difference infinitesimally small, we are accustomed to write the derivative as d instead of the (difference) Δ. With differentiation we eliminate the DC term, with a constant and a linear term remaining:

$$g_\mathrm{sf} = \frac{dI_\mathrm{d}}{dV_\mathrm{gs}} = 0 - 2\frac{I_\mathrm{dss}}{V_\mathrm{p}} + 2\frac{I_\mathrm{dss}}{V_\mathrm{p}^2}V_\mathrm{gs} \tag{A3.5}$$

Note that when $V_\mathrm{gs} = 0$, then $I_\mathrm{d} = I_\mathrm{dss}$, and $\mathrm{g_{sf}} = \mathrm{g_{sf0}}$, which means that the maximal transconductance achievable for a jFET is given by that constant term in (A3.5):

$$g_\mathrm{sf0} = -2\frac{I_\mathrm{dss}}{V_\mathrm{p}} \tag{A3.6}$$

and this figure is often stated by the manufacturer in the device's data. FETs tend to have large tolerances for V_p, say a nominal $-2\,\mathrm{V}$ value can vary from $-0.5\,\mathrm{V}$ to $-6\,\mathrm{V}$ in an actual device, with corresponding variations of I_dss, so g_sf0 usually gives a more reliable performance estimation. Note also that the remaining V_gs term in (A3.5) is the DC bias value, and since the term has a sign opposite to the g_sf0, it shows how the g_sf is reduced by making V_gs larger. The negative value of g_sf means that the output signal is inverted with respect to the input signal.

By combining (A3.5) and (A3.6) we can write:

$$g_\mathrm{sf} = g_\mathrm{sf0}\left(1 - \frac{V_\mathrm{gs}}{V_\mathrm{p}}\right) \tag{A3.7}$$

or alternatively:

$$g_\mathrm{sf} = g_\mathrm{sf0}\sqrt{\frac{I_\mathrm{d}}{I_\mathrm{dss}}} \tag{A3.8}$$

An often neglected aspect of jFET performance is the output conductance:

$$g_\mathrm{os} = \frac{\Delta I_\mathrm{d}}{\Delta V_\mathrm{ds}} \tag{A3.9}$$

Because $V_\mathrm{ds} = V_\mathrm{dd} - R_\mathrm{d}I_\mathrm{d}$, it is obvious that the signal gain also depends on g_os and consequently on the supply voltage and the drain resistance:

$$A = \frac{v_\mathrm{o}}{v_\mathrm{g}} = g_\mathrm{sf}\frac{R_\mathrm{d}}{1 + g_\mathrm{os}R_\mathrm{d}} \tag{A3.10}$$

so only if $g_\mathrm{os} \ll 1/R_\mathrm{d}$ can the signal gain be approximated as $A \approx g_\mathrm{sf}R_\mathrm{d}$.

This condition requires a suitably high supply voltage V_{dd}, so that the output signal variation $v_o = \Delta I_d R_d$ can be relatively small in comparison with V_{dd}.

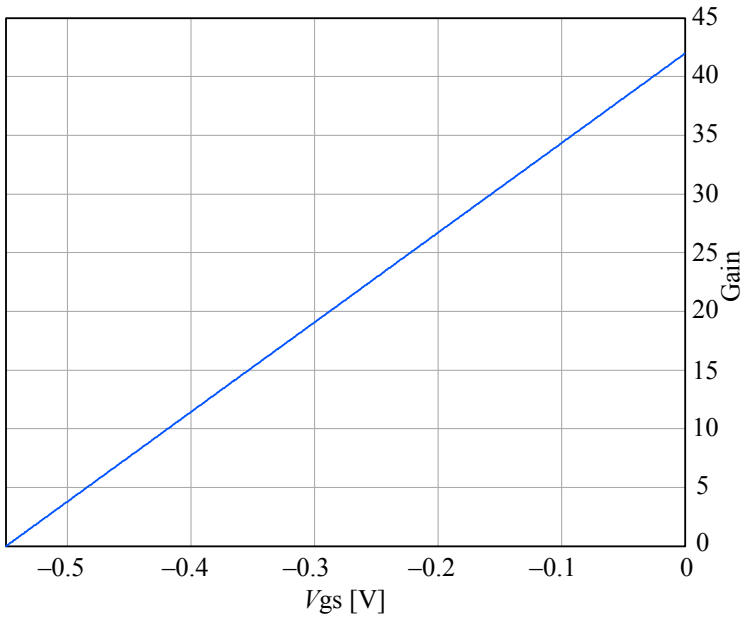


Fig.A3.2: Gain A as a function of V_{gs}, with $V_{dd} = 12\,\text{V}$ and $R_d = 3.3\,\text{k}\Omega$.

We must now discuss some dynamic effects. At high frequencies and large amplitudes the parasitic gate-drain capacitance C_{gd} becomes particularly important. The capacitance per unit area can be expressed as:

$$C = \left(\frac{K}{V_{bi} - V_g} \right)^k \tag{A3.11}$$

Here V_{bi} represents the 'built in' space charge potential, usually about 0.6 V, and K is a constant determined by the channel geometry. For an abrupt junction the exponent $k = 1/2$, and for a linearly graded junction $k = 1/3$. A typical FET has a grading coefficient between these values. An acceptable approximation is $k = 0.4$.

A similar relation is valid for the gate-source capacitance C_{gs}. Because of the lower voltage across it (only V_{gs}), C_{gs} is typically much larger than C_{gd}, and it also varies more with the gate voltage. However, the variation of C_{gd} will often be dominant because of the Miller effect, which was already discussed in Appendix 2.

To make the analysis a bit easier for the purpose of acquiring a feeling of what is going on, let us assume that the signal generator impedance is purely resistive, say $R_G = 40\,\text{k}\Omega$. Then the gate input capacitance will form a low pass filter for the signal, with a transfer function magnitude:

$$\left| \frac{v_g}{v_G} \right| = \frac{1}{\sqrt{1 + (2\pi f\, C_{in} R_G)^2}} \tag{A3.12}$$

where v_G is the signal generator voltage, v_g is the jFET gate voltage, and C_{in} is:

$$C_{in} = C_{gs} + C_{gd}(1 + A) \tag{A3.13}$$

and here we have accounted for the Miller effect caused by the voltage gain A. C_{gs} and C_{gd} are functions of V_{gs} and V_{gd}, so we need to determine the bias condition and gain A. In general a junction capacitance C_j varies with the junction voltage V_j:

$$C_j = C_0 \frac{1}{1 + \left(\frac{V_j}{V_{bi}} \right)^k} \tag{A3.14}$$

Here C_0 represents a measured nominal capacitance for a specific bias, as given in the manufacturer's data sheet. With C_{j1} known for a certain bias, we can determine C_{j2} for a different bias as:

$$C_{j2} = C_{j1} \left(\frac{V_{bi} - V_{j1}}{V_{bi} - V_{j2}} \right)^k \tag{A3.15}$$

We can determine the required V_{gs} from (A3.1), here assuming $n = 2$:

$$V_{gs} = V_p \left(1 - \sqrt{\frac{I_d}{I_{dss}}} \right) \tag{A3.16}$$

Then A is determined by (A3.10), and we can see that $V_j = V_{gd}$:

$$-V_{gd} = V_{dd} - R_d I_d - V_g \tag{A3.17}$$

where V_g is the gate to ground potential, which in the case of a grounded source is equal to V_{gs}.

For example, as shown in Fig.A3.3, the LSK389 data specify the input capacitance as $C_{iss} = 25\,\text{pF}$, and the reverse transfer capacitance $C_{rss} = 5\,\text{pF}$. Note that those values are somewhat larger than C_{gs} and C_{gd}, respectively, but the ratio is similar. However, at a bias of $V_{gs} = 150\,\text{mV}$ and $I_d = 1.8\,\text{mA}$ the gain will be about $A = 30\times$, and the effective input capacitance owed to C_{gd} becomes $(A + 1)C_{gd}$, or about 35 pF. Of course, the variation of the capacitance with the signal is amplified by the same gain factor.

Let us see what impact will that have on a sine wave of 5 kHz. The input signal amplitude (at nominal 5 cm/s modulation) will be $17\,\text{mV}_{peak}$, so the V_{gs} bias of -150 mV is being varied by ±11%, which in turn modulates the total input capacitance by a few %. With the signal source impedance of our typical average cartridge, this will introduce a distortion of a few %, mostly of 2nd harmonic. Fig.A3.4 shows the input and output signal and their normalized difference. The signal gain is about 30×.

Note however that this analysis is valid for a single jFET amplifier. But based on this some amplifier designers conclude that a similar distortion mechanism is present in differential amplifiers as well. Of course, the differential signal level is reduced by feedback by s few orders of magnitude, but in noninverting amplifiers the full signal amplitude is present in the form of a common mode signal, with equally devastating effects. The question is: how much truth is there is those claims? And if that is not so, why?

First thing to note is that in a differential jFET pair the source of one device is loaded by the source of the other, acting in a complementary manner for the

differential signal. This means that the nonlinearity effect of C_{gs} is greatly reduced, even in case of a mismatch between the two devices.

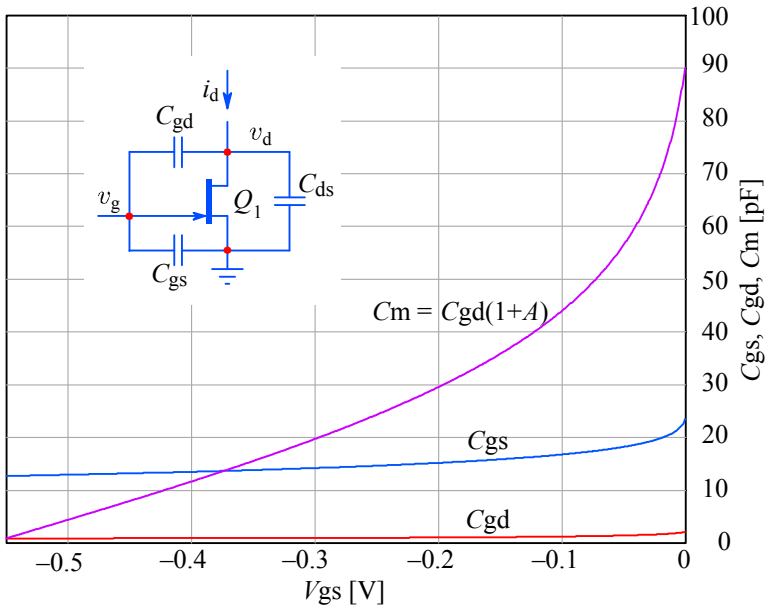


Fig.A3.3: Input capacitance of a typical low-capacitance jFET is the sum of the gate-source capacitance C_{gs} and the Miller capacitance $C_{gd}(1 + A)$. The capacitances vary nonlinearly with the signal. The Miller effect makes the largest capacitance variation.

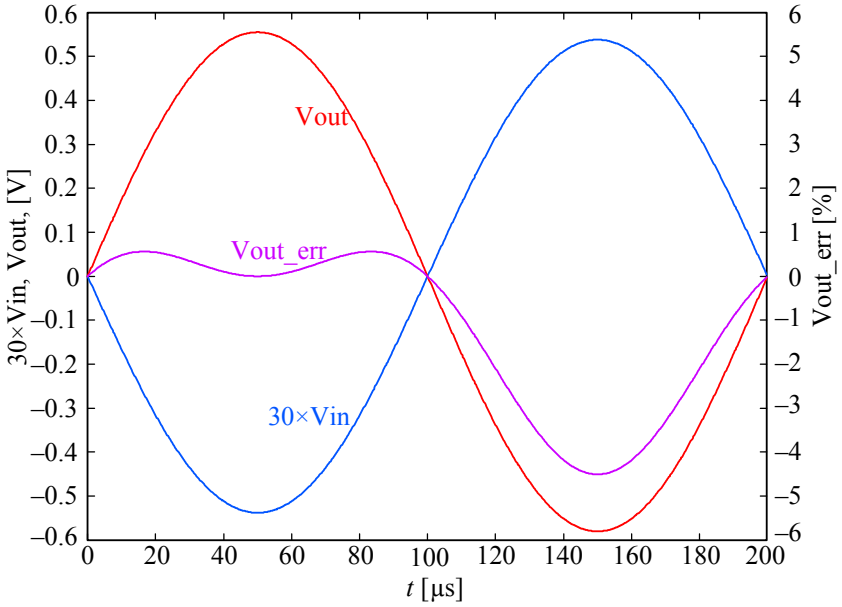


Fig.A3.4: Distortion of a 5 kHz signal for a single jFET amplifier; see text for details.

Next, C_{gd} is also being modulated in a complementary manner, but the feedback impedance in an RIAA equalizer is always very low, whilst the input impedance varies widely. However, the cable connecting the cartridge to the amplifier will usually have 100 pF/m, with some additional capacitance in parallel, necessary for the correct inductance compensation, and both appear in parallel with the amplifier input capacitance, so the variations of the input capacitance represent a much smaller fraction of the total.

As the differential signal is reduced by feedback, so is the nonlinearity of the gain transfer function and of the effective input capacitance. This means that more feedback in a loop is a desirable factor, as long as the bandwidth is correctly taken care of, and any secondary poles of the amplifier are kept below the unity gain level, thus the amplifier stability is not compromised.

What about the common mode signal? It is being reduced by the symmetry of the differential pair, but this can be also affected by the impedance of the current source which supplies the bias current to the pair. If the impedance of the current source is not high enough, the sources of the differential pair do not share the same current, some of the current is lost mainly through the drain capacitance and the real conductance of the current source. With spoiled symmetry, part of the common mode signal becomes a differential mode signal, causing trouble.

A simple way of greatly increasing the impedance of the current source is to use the 'cascode' configuration. The name stems from the valve/tube epoch circuits in which the anode of the first device was connected to the cathode of the second, forming a 'cascade to cathode'. The configuration using semiconducting devices is similar, but different types require different bias method. The configuration using jFETs is particularly simple, Fig.A3.5, because of the similarity of the gate voltage polarity with thermionic devices.

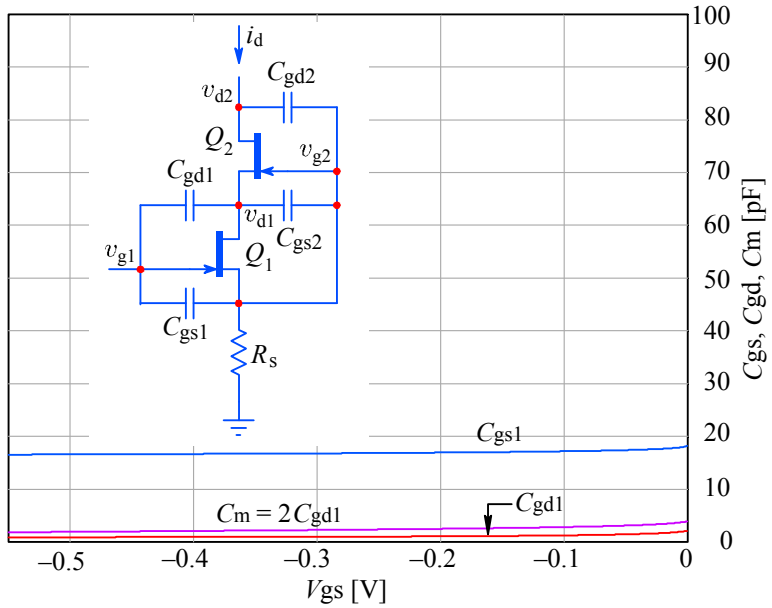


Fig.A3.5: A jFET cascode connection and its capacitances. Because the voltage on Q_1 is equal to V_{gs2} and changes very little, the effective input capacitances also changes very little. The Miller effect is greatly reduced by the low impedance of Q_2 source loading the drain of Q_1, making the Q_1 gain ≈ 1, so $C_m \approx 2C_{gd1}$.

The cascode connection has high linearity in the saturation region, also for very large signals. The Miller effect is greatly reduced because the drain of Q_1 is loaded by the source of Q_2, making its gain $A \approx 1$, so the effective Miller capacitance is only $C_m = 2C_{gd}$, and linearity is improved because the drain voltage of Q_1 varies little. Therefore bandwidth and linearity are extended to high frequencies and high signal levels. Also the leakage current to the gate of Q_1 is greatly reduced, because its V_{dg} is small, lower than the V_{gs} of Q_2.

We are not going to deal with the cascode amplifier in full length, but only show how the gain depends on the gain of individual stages. The gain of Q_1 (which acts as a common source stage) can be written as:

$$A_{v1} = \frac{v_{\mathrm{d1}}}{v_{\mathrm{g1}}} = -g_{\mathrm{m1}} \frac{r_{\mathrm{d1}} r_{s2}}{r_{\mathrm{d1}} + r_{s2}} \tag{A3.18}$$

and the gain of Q_2 (which acts as a common gate stage) can be approximated as:

$$A_{v2} = \frac{v_{\mathrm{d2}}}{v_{\mathrm{s2}}} \approx -g_{\mathrm{m2}} \frac{r_{\mathrm{d2}} R_{\mathrm{d}}}{r_{\mathrm{d2}} + R_{\mathrm{d}}} \tag{A3.19}$$

where R_{d} is the drain resistance of Q_2. Two things become important here: first, Q_2 reduces the loading of the drain of Q_1 by $g_{\mathrm{m2}} r_{\mathrm{d2}}$:

$$r_{\mathrm{s2}} = \frac{r_{\mathrm{d2}} + R_{\mathrm{d}}}{1 + g_{\mathrm{m2}} r_{\mathrm{d2}}} \tag{A3.20}$$

and second, the system gain is the product of the gains of each stage:

$$A_v = \frac{v_{\mathrm{d2}}}{v_{\mathrm{g1}}} = A_{v_1} A_{v2} = g_{\mathrm{m1}} g_{\mathrm{m2}} \frac{r_{\mathrm{d1}} r_{s2}}{r_{\mathrm{d1}} + r_{s2}} \frac{r_{\mathrm{d2}} R_{\mathrm{d}}}{r_{\mathrm{d2}} + R_{\mathrm{d}}} \tag{A3.21}$$

which by inserting () into () becomes:

$$A_v = g_{\mathrm{m1}} g_{\mathrm{m2}} R_{\mathrm{d}} \frac{r_{\mathrm{d1}} r_{\mathrm{d2}}}{r_{\mathrm{d1}}(1 + g_{\mathrm{m2}} r_{\mathrm{d2}}) + r_{\mathrm{d2}} + R_{\mathrm{d}}} \tag{A3.22}$$

Note that because the denominator is much greater than the numerator, the effective gain of a cascode stage is similar to that of a single common source stage, but its dependence on internal resistance variability is greatly reduced. Therefore it makes sense to employ the cascode also for the differential pair itself, as in Fig.27, since such a simple circuit addition solves many problems.

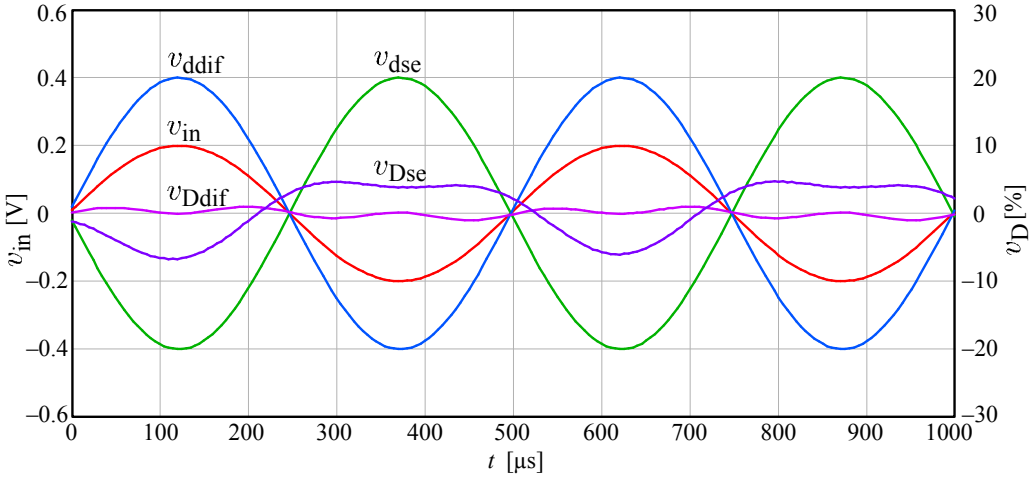


Fig.A3.6: A computer simulation, using model parameters of the SK170 jFET, compares the distortion of a jFET single-ended and differential cascode circuit. Note that the signal voltage v_{in} is not drawn to scale. The single-ended cascode has some 5% distortion, dominantly 2nd harmonic, whilst the differential cascode has only 0.5%, but dominantly 3rd harmonic, both at 0.4 V peak output at 2 kHz. The loop gain of an ordinaary opamp will reduce this by at least 500×.

Appendix 4: The Third Order 10 Hz High Pass Filter

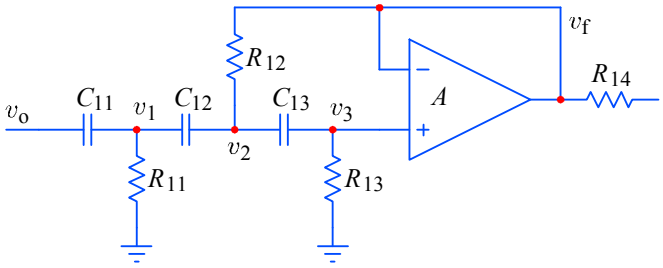


Fig.A4.1: Third order high pass filter of the Sallen–Key configuration

From the theory of active filters, in particular the Sallen–Key circuit topology we know that the component values spread and the system's response sensitivity to component variations is minimized by making all the serial components (capacitors in this case) equal: $C_{11} = C_{12} = C_{13} = C$.

We have already derived the transfer function from (64) to (71):

$$\frac{v_f}{v_o} = \frac{s^3}{s^3 + s^2\frac{1}{C}\left(\frac{3}{R_{13}} + \frac{1}{R_{11}}\right) + s\frac{2}{C^2}\left(\frac{1}{R_{12}R_{13}} + \frac{1}{R_{11}R_{13}}\right) + \frac{1}{C^3 R_{11}R_{12}R_{13}}}$$

If we compare this with the general third order high pass form :

$$F(s) = A_0\frac{s^3}{(s-s_1)(s-s_2)(s-s_3)} =$$

$$= A_0\frac{s^3}{s^3 - s^2(s_1+s_2+s_3) + s(s_1s_2 + s_1s_3 + s_2s_3) - s_1s_2s_3} \qquad \text{(A4.1)}$$

we note that the gain $A_0 = 1$, whilst the system cut off frequency is obtained from $\omega_0^3 = -s_1s_2s_3 = 1/C^3R_{11}R_{12}R_{13}$. Still, we have four component values to define, but only three polynomial coefficients. But we can normalize the values by making $C = 1$, obtain the resistor ratios from the polynomial coefficients, and then denormalize the CR products for the required cut off frequency.

Anyway, the pole values are also given in normalized form to $\omega_0 = 1\,\text{rad/s}$, either by tables for systems of different order, or by numerical algorithms to compute the values. The algorithm for Bessel poles is quite complicated, but the algorithm for polynomial coefficients is relatively simple. Bessel polynomials follow these rules:

$$\begin{aligned} B_0(s) &= 1 \\ B_1(s) &= s + 1 \\ B_n(s) &= (2n-1)\,B_{n-1}(s) + s^2 B_{n-2}(s) \end{aligned} \qquad \text{(A4.2)}$$

The coefficients c_k of the resulting polynomial of order n can be calculated as:

$$c_k = \left.\frac{(2n-k)!}{2^{(n-k)}\,k!\,(n-k)!}\right|_{k=0,1,2,\ldots,n-1,n} \qquad \text{(A4.3)}$$

For the third order ($n = 3$) Bessel system the polynomial coefficients are:

$$k = 3 \Rightarrow \qquad c_3 = \frac{(2{\cdot}3-3)!}{2^{(3-3)}\,3!\,(3-3)!} = \frac{3!}{2^0\,3!\,0!} = \frac{3{\cdot}2{\cdot}1}{1{\cdot}3{\cdot}2{\cdot}1} = \frac{6}{6} = 1 \tag{A4.4}$$

$$k = 2 \Rightarrow \qquad c_2 = \frac{(2{\cdot}3-2)!}{2^{(3-2)}\,2!\,(3-2)!} = \frac{4!}{2^1\,2!\,1!} = \frac{4{\cdot}3{\cdot}2{\cdot}1}{2{\cdot}2{\cdot}1} = \frac{24}{4} = 6 \tag{A4.5}$$

$$k = 1 \Rightarrow \qquad c_1 = \frac{(2{\cdot}3-1)!}{2^{(3-1)}\,1!\,(3-1)!} = \frac{5!}{2^2\,1!\,2!} = \frac{120}{4{\cdot}1{\cdot}2} = \frac{120}{8} = 15 \tag{A4.6}$$

$$k = 0 \Rightarrow \qquad c_0 = \frac{(2{\cdot}3-0)!}{2^{(3-0)}\,0!\,(3-0)!} = \frac{6!}{2^3\,0!\,3!} = \frac{6{\cdot}5{\cdot}4{\cdot}3{\cdot}2{\cdot}1}{8{\cdot}1{\cdot}6} = \frac{720}{48} = 15 \tag{A4.7}$$

The denominator of the transfer function will then be the following polynomial:

$$\sum_{k=3}^{0} c_k s^k = c_3 s^3 + c_2 s^2 + c_1 s^1 + c_0 s^0 = s^3 + 6s^2 + 15s + 15 \tag{A4.8}$$

The poles for this polynomial are:

$$s_3 = -1.0474 - j0.9993$$
$$s_2 = -1.0474 + j0.9993$$
$$s_1 = -1.3227$$

As stated in the main text, these are the poles for a low pass system, and for a high pass system we need to invert the poles, $1/s_{\mathrm{i}}$. So the values to be used are:

$$s_3 = \frac{1}{-1.0474-j0.9993} = -0.4998 + j9.4768$$
$$s_2 = \frac{1}{-1.0474+j0.9993} = -0.4998 - j9.4768$$
$$s_1 = \frac{1}{-1.3227} = -0.7560 \tag{A4.9}$$

Now we can solve the following set of equations:

$$-s_1 - s_2 - s_3 = \frac{3}{R_{11}} + \frac{1}{R_{13}} \tag{A4.10}$$

$$s_1 s_2 + s_1 s_3 + s_2 s_3 = 2\left(\frac{1}{R_{11}R_{12}} + \frac{1}{R_{11}R_{13}}\right) \tag{A4.11}$$

$$-s_1 s_2 s_3 = \frac{1}{R_{11}R_{12}R_{13}} \tag{A4.12}$$

To shorten the expressions and reduce the possibility of transcription errors we set:

$$x = -s_1 - s_2 - s_3 \tag{A4.13}$$

$$y = s_1 s_2 + s_1 s_3 + s_2 s_3 \tag{A4.14}$$

$$z = -s_1 s_2 s_3 \tag{A4.15}$$

And for simplicity we also reassign:

$$R_1 = R_{11} \tag{A4.16}$$

$$R_2 = R_{12} \tag{A4.17}$$

$$R_3 = R_{13} \tag{A4.18}$$

Then, from (A4.11) and (A4.14) we have:

$$y = \frac{2}{R_1}\left(\frac{1}{R_2} + \frac{1}{R_3}\right) \tag{A4.19}$$

$$R_1 = \frac{2}{y}\left(\frac{1}{R_2} + \frac{1}{R_3}\right) \tag{C.20}$$

From (A4.12) and (A4.15):

$$z = \frac{1}{\frac{2}{y}\left(\frac{1}{R_2} + \frac{1}{R_3}\right)R_2R_3} \tag{A4.21}$$

$$z = \frac{y}{2(R_2 + R_3)} \tag{A4.22}$$

$$R_2 + R_3 = \frac{y}{2z} \tag{A4.23}$$

$$R_2 = \frac{y}{2z} - R_3 \tag{A4.24}$$

From (A4.10) and (A4.13):

$$x = \frac{3}{\frac{2}{y}\left(\frac{1}{R_2} + \frac{1}{R_3}\right)} + \frac{1}{R_3} \tag{A4.25}$$

$$x = \frac{3}{\frac{2}{y}\left(\frac{1}{\frac{y}{2z} - R_3} + \frac{1}{R_3}\right)} + \frac{1}{R_3} \tag{A4.26}$$

$$x = \frac{3y}{2\left(\frac{2z}{y - 2zR_3} + \frac{1}{R_3}\right)} + \frac{1}{R_3} \tag{C.27}$$

$$x = \frac{3y}{2\left[\frac{2zR_3 + y - 2zR_3}{(y - 2zR_3)R_3}\right]} + \frac{1}{R_3} \tag{A4.28}$$

$$x = \frac{3y}{2\frac{y}{(y - 2zR_3)R_3}} + \frac{1}{R_3} \tag{A4.29}$$

$$x = \frac{3(y - 2zR_3)R_3}{2} + \frac{1}{R_3} \tag{A4.30}$$

$$x = \frac{3}{2}yR_3 - 3zR_3^2 + \frac{1}{R_3} \tag{A4.31}$$

$$xR_3 = \frac{3}{2}yR_3^2 - 3zR_3^3 + 1 \tag{A4.32}$$

$$3zR_3^3 - \frac{3}{2}yR_3^2 + xR_3 - 1 = 0 \tag{A4.33}$$

$$R_3^3 - \frac{y}{2z}R_3^2 + \frac{x}{3z}R_3 - \frac{1}{3z} = 0 \tag{A4.34}$$

We have now a third order equation for R_3. To simplify it further, we can substitute:

$$p = -\frac{y}{2z} \tag{A4.35}$$

$$q = \frac{x}{3z} \tag{A4.36}$$

$$r = -\frac{1}{3z} \tag{A4.37}$$

The equation (A4.34) becomes:

$$R_3^3 + R_3^2 p + R_3 q + r = 0 \tag{A4.38}$$

We solve this for R_3 by using a general form solution for the third order equation:

$$R_3 = -\frac{2}{3}\sqrt{p^2 - 3q}\,\sin\left[\frac{1}{3}\arctan\frac{\sqrt{3}\,(p^2 - 3q)^{3/2}(3q - p^2)^{-3/2}(2p^3 - 9pq + 27r)}{9\sqrt{4rp^3 - p^2q^2 - 18pqr + 4q^3 + 27r^2}}\right] - \frac{p}{3} \tag{A4.39}$$

Note that there are also two complex-conjugate solutions for R_3, but only the real value solution is needed. From a known R_3 we can then find:

$$R_2 = \frac{s_1 s_2 + s_1 s_3 + s_2 s_3}{2 s_1 s_2 s_3} - R_3 \tag{A4.40}$$

$$R_1 = \frac{1}{R_2 R_3 s_1 s_2 s_3} \tag{A4.41}$$

We can now find C from a given -3 dB bandwidth limit $\omega_0 = \sqrt[3]{-s_1 s_2 s_3}$:

$$\omega_0 = \frac{1}{\sqrt[3]{C^3 R_1 R_2 R_3}} \tag{A4.42}$$

$$C = \frac{1}{\omega_0 \sqrt[3]{R_1 R_2 R_3}} \tag{A4.43}$$

For Butterworth poles the procedure is identical, except that we start from the values for a Butterworth 3rd-order system:

$$s_3 = -0.5000 - j0.8660$$
$$s_2 = -0.5000 + j0.8660$$
$$s_1 = -1.0000 \qquad \text{(A4.44)}$$

and we do not need to invert those, because those poles are on the unit circle (as for the whole Butterworth family), so the inverted values are the same.

Appendix 5: Matlab Code Used for the RIAA Amplifier Design

```
% RIAA optimization
% NOTE: the figure numbers are not the same as in the text!

% frequency vectors (100 samples/decade)
f=logspace(0,6,601);
s=j*2*pi*f;
% discrete frequencies
fp=[1,4,10,20,50,500,1000,2122,20000,50000,400000];
sp=j*2*pi*fp;
% 1kHz
f1k=1000;
s1k=j*2*pi*f1k; % == s(301)
% 50kHz
f50k=50000;
s50k=j*2*pi*f50k;

% time constants definitions:
T1=3183e-6;  % [s] .... f1=50Hz
T2=318.3e-6; % [s] .... f2=500Hz
T3=75e-6;    % [s] .... f3=2122Hz
T4=3.183e-6; % [s] .... f4=50kHz
T5=0.4e-6;   % [s] .... f5=400kHz

% Inverse functions
% nominal DC attenuation for all inverse functions:
ax=T2*T4/(T1*T3);
% 3 time constants:
Fi3=ax*(s*T1-1).*(s*T3-1)./(s*T2-1);
% 4 time constants:
Fi4=ax*(s*T1-1).*(s*T3-1)./((s*T2-1).*(s*T4-1));
% 5 time constants:
Fi5=ax*(s*T1-1).*(s*T3-1)./((s*T2-1).*(s*T4-1).*(s*T5-1));
% Magnitudes:
Mi3=20*log10(abs(Fi3));
Mi4=20*log10(abs(Fi4));
Mi5=20*log10(abs(Fi5));

% reference value at 1kHz for normalization:
% Fi41k=(s1k*T1-1).*(s1k*T3-1)./((s1k*T2-1).*(s1k*T4-1));
% Mi41k=20*log10(abs(Fi41k));
figure(1)
semilogx(f,Mi4,'-r',...
         f,Mi5,'-m')
title('Input - inverting RIAA')
xlabel('f [Hz]')
ylabel('Attenuation [dB]')
grid

% RIAA equalizer reference values:
% for T3=75e-6;
R2=7500;
C2=1e-8;
% Theoretical (no T4):
% R1=R2*(T1-T2)/(T2-T3); --> result too low!
% artificial correction with T4:
R1=R2*(1+T4/T3)*(T1-T2)/(T2-T3);
C1=T1/R1;
% feedback impedance
Zfb=1 ./(1/R1+s*C1) + 1 ./(1/R2+s*C2);
% at 1kHz:
Zfb1k=1 ./(1/R1+s1k*C1) + 1 ./(1/R2+s1k*C2);
% at 50kHz:
Zfb50k=1 ./(1/R1+s50k*C1) + 1 ./(1/R2+s50k*C2);
```

```
% nominal attenuation at DC:
a_dc=T1*T3/(T2*T4);
% calculate total R34=R4+R3:
R34=(R1+R2)/(a_dc-1);
disp(['--> R34 = ', num2str(R34), ' Ohm'])
% nominal gain requirement
% output level: -10dBm (0.1mW at 600 Ohm)
% input level: 5mV rms @ 5cm/s modulation velocity @ 1kHz
Ar1k=sqrt(1e-4*600)/5e-3;
R3=abs(Zfb1k+R34)/Ar1k;
R4=R34-R3;
disp('------------------------------')
disp(['--> R2 = ', num2str(R2*1e-3), ' kOhm'])
disp(['--> C2 = ', num2str(C2*1e+9), ' nF'])
disp(['--> R1 = ', num2str(R1*1e-3), ' kOhm'])
disp(['--> C1 = ', num2str(C1*1e+9), ' nF'])
disp(['--> R4 = ', num2str(R4), ' Ohm'])
disp(['--> R3 = ', num2str(R3), ' Ohm'])
disp('------------------------------')

% 50kHz output low pass filter 1st-order:
Rf1=2120;
Cf1=1.5e-9;
Tf1=Cf1*Rf1;
% complex frequency response:
J1=1 ./(s*Tf1+1);
% magnitude:
Mj1=20*log10(abs(J1));

% 50kHz Butterworth low pass filter 2nd-order:
% normalized poles:
p2=(sqrt(2)/2)*[-1-j;-1+j];
% actual poles for 50kHz bandwidth:
p2=p2*2*pi*5e+4;
% complex frequency response:
J2=prod(-p2) ./((s-p2(1)).*(s-p2(2)));
% magnitude:
Mj2=20*log10(abs(J2));

% combined effect of J1 and J2:
J3=J1.*J2;
Mj3=20*log10(abs(J3));

% equalizer gain
Gr1 = 1 + R4/R3 + Zfb/R3;
Mgr1=20*log10(abs(Gr1));
% with LPF J1:
Gr2 = Gr1 .* J1;
Mgr2=20*log10(abs(Gr2));
% equalizer driving with Fi4:
% normalization to 1kHz:
Fi4n=Fi4/Fi4(301);
Mfi4n=20*log10(abs(Fi4n));
FGr1=Fi4n.*Gr1;
FGr2=Fi4n.*Gr2;
Mfgr1=20*log10(abs(FGr1));
Mfgr2=20*log10(abs(FGr2));

figure(2)
semilogx(f,Mgr2,            '--b',...
         f,Mgr1,            '-b',...
         f,Mfi4n+Mgr1(301), '-m',...
         f,Mfgr1,           '-r',...
         f,Mfgr2,           '-c',...
         f,Mj1+Mgr1(301),   '-k',...
         f,Mj2+Mgr1(301),   '-g',...
         f,Mj3+Mgr1(301),   '-g',...
```

```
          f,Mj3+Mgr1(301),    '--k')
axis([min(f), max(f), 0, 65]);
title('Amplifier - RIAA Equalizer')
xlabel('f [Hz]')
ylabel('20*log10(Vo/Vi) [dB]')
grid

Fc1=(Fi4/Fi4(301)).*Gr1; % Fi re 1kHz
Fc2=(Fi5/Fi5(301)).*Gr1; % Fi re 1kHz
Fc3=(Fi4/Fi4(301)).*Gr2; % Fi re 1kHz
Fc4=(Fi5/Fi5(301)).*Gr2; % Fi re 1kHz
Mc1=20*log10(abs(Fc1));
Mc2=20*log10(abs(Fc2));
Mc3=20*log10(abs(Fc3));
Mc4=20*log10(abs(Fc4));

figure(3)
semilogx(f,Mc1, '-r',...
          f,Mc2, '-m',...
          f,Mc3, '-g',...
          f,Mc4, '-b')
title('Input, Compensation, and Resulting Output')
xlabel('f [Hz]')
ylabel('20*log10(Vo/Vi) [dB]')
grid

put=4;
% Pickup Z
% ----------------------
if (put==1)
     % Stanton ST500E MkII
     Lp=0.400;      % [H]
     Rp=635;        % [Ohm]
     Rbn=47e+3;     % [Ohm]
     Cbn=100e-12;   % [F]
     Rbo=47e+3;     % [Ohm]
     Cbo=250e-12;   % [F]
elseif(put==2)
     % Sonus Blue Label
     Lp=0.150;      % [H]
     Rp=300;        % [Ohm]
     Rbn=47e+3;     % [Ohm]
     Cbn=200e-12;   % [F]
     Rbo=47e+3;     % [Ohm]
     Cbo=39e-12;    % [F]
elseif(put==3)
     % General Average
     Lp=0.300;      % [H]
     Rp=600;        % [Ohm]
     Rbn=47e+3;     % [Ohm]
     Cbn=120e-12;   % [F]
     Rbo=47e+3;     % [Ohm]
     Cbo=90e-12;    % [F]
else
     % Grado Prestige Silver
     Lp=0.045;      % [H]
     Rp=475;        % [Ohm]
     Rbn=47e+3;     % [Ohm]
     Cbn=250e-12;   % [F]
     Rbo=27e+3;     % [Ohm]
     Cbo=120e-12;   % [F]
end

% amp input impedance - nominal cartrdige load
Zin=1 ./(1/Rbn + s*Cbn);
% with optimized load
Zio=1 ./(1/Rbo + s*Cbo);
```

```
% cartridge + load transfer function
Hpun=Zin./(Rp+s*Lp+Zin);
Hpuo=Zio./(Rp+s*Lp+Zio);
Mpun=20*log10(abs(Hpun));
Mpuo=20*log10(abs(Hpuo));

figure(4)
semilogx(f,Mpun,'-g',...
         f,Mpuo,'-b')
axis([min(f), max(f),-4,1]);
title('Pickup with Loading')
xlabel('f [Hz]')
ylabel('Vpu [dB]')
text(30,-0.8,[' Lp = ', num2str(Lp*1000),' mH'])
text(30,-1.3,[' Rp = ', num2str(Rp),' Ohm'])
text(30,-1.8,['Rbn = ', num2str(Rbn/1000),' kOhm'])
text(30,-2.3,['Cbn = ', num2str(Cbn*1e+12),' pF'])
text(30,-2.8,['Rbo = ', num2str(Rbo/1000),' kOhm'])
text(30,-3.3,['Cbo = ', num2str(Cbo*1e+12),' pF'])
grid

Fcxn=Fc4.*Hpun;
Fcxo=Fc4.*Hpuo;
Mcxn=20*log10(abs(Fcxn));
Mcxo=20*log10(abs(Fcxo));

figure(5)
semilogx(f,Mc4,             '-m',...
         f,Mpun+Mgr1(301), '-g',...
         f,Mpuo+Mgr1(301), '-m',...
         f,Mpuo+Mgr1(301), '--g',...
         f,Mcxn,            '-k',...
         f,Mcxo,            '-m',...
         f,Mcxo,            '--k')
axis([min(f), max(f), 30, 35]);
title('Output, pickup, and combined response')
xlabel('f [Hz]')
ylabel('Vout, Vpu*Vout [dB]')
grid

% DC integrator, HPF 1Hz
Rint=47000;
Cint=3.3e-6;
wint=1/(Cint*Rint);
K1=s ./(s-wint);
Mk1=20*log10(abs(K1));
% HighPass Filter 3rd order
% Bessel Low-Pass, normalized to 1Hz
[z,p3]=bestap(3,'n');
% inversion for High-Pass, 1Hz
p3=1 ./p3;
% for 10Hz:
p3=p3*2*pi*10;
% High-Pass Response:
K3=s.^3 ./((s-p3(1)).*(s-p3(2)).*(s-p3(3)));
Mk3=20*log10(abs(K3));
% Butterworth, 10Hz
[z,q3]=buttapx(3);
q3=q3*2*pi*10;
K3u=s.^3 ./((s-q3(1)).*(s-q3(2)).*(s-q3(3)));
Mk3u=20*log10(abs(K3u));

figure(6)
semilogx(f, Mk1,  '-r',...
         f, Mk3u, '-g',...
         f, Mk3,  '-b')
axis([min(f), max(f), -4, 1]);
```

```
grid
xlabel('f [Hz]')
ylabel('Attenuation [dB]')
title('Compare Bessel and Butterworth HPFs')

% all together now:
X=Fi4n.*Hpuo.*Gr1.*J1.*K1.*K3u;
Mx=20*log10(abs(X));

figure(7)
semilogx(f,Mx,'-b')
xlabel('f [Hz]')
ylabel('Gain [dB]')
title('full system simulation')
axis([min(f), max(f), -10, +40]);
grid;

% Noise Analysis - equivalent input noise:
% absolute Temperature:
Ta=293; % [K], ==20°C
% Boltzmann constant:
kB = 1.3807e-23; % [VAs/K]

% Pickup impedance
Zpun = 1 ./(1 ./(Rp + s*Lp) + 1/Rbo + s*Cbo);
% Pickup thermal niose voltage spectral density:
enZpu=sqrt(4*kB*Ta*real(Zpun));

% amplifier input noise voltage s.d., manufacturer data:

% [V/sqrt(Hz)] (opamps: NE5534, OP37, OPA656, OPA657, OPA857, LM49880)
% enAmp=4e-9;

% [V/sqrt(Hz)] LSK398 dual jFET dif. pair
enAmp=sqrt(2)*0.9e-9;

% 1/f corner frequency
% fcn=300; % [Hz] - opamps general
% fcn=20; % [Hz] - OPA827
fcn=50; % [Hz] - LSK398
% amp input noise + 1/f noise
Fnamp=(fcn+f)./f;
enAmp=enAmp.*Fnamp;

% amplifier input noise current s.d.
inAmp=2.2e-15; % [A/sqrt(Hz)]
inAmp=inAmp.*Fnamp; % 1/f noise
einAmp=sqrt((inAmp.*abs(Zpun)).^2 + (inAmp*R3).^2);

% feedback network impedance simplified to R3:
Rfbn=sqrt(4*kB*Ta*R3)*ones(size(f));

% total input noise:
% Fntot=sqrt(enZpu.^2 + enAmp.^2 + einAmp.^2 + Rfbn.^2);
% current noise einAmp too smal, neglected
Fntot=sqrt(enZpu.^2 + enAmp.^2 + Rfbn.^2);

% output noise = input noise * RIAA equ gain
Fnout=Fntot.*abs(Gr2.*K3u);

figure(8)
loglog(f,enZpu, '-r',...
       f,enAmp, '-m',...
       f,Rfbn,  '-g',...
       f,Fntot, '-b',...
       f,Fnout, '-k')
%       f,einAmp,'-g') % neglected
```

```
title('Noise Source Voltages S.D. and Total Input Noise S.D.')
xlabel('f [Hz]')
ylabel('V/sqrt(Hz)')
grid

% output rms noise:
% make the df vector:
df=diff(f);
df=[df(1), df];
% square the voltage and integrate noise power from 20Hz to 20kHz,
% then take the square root:
Vnrms = sqrt(sum(Fnout.^2 .* df));
disp('------------')
disp(['Vnrms = ', num2str(Vnrms), ' Vrms'])
snr=0.25/Vnrms;
disp(['S/N ratio = 20*log10(0.25/Vnrms) = ', num2str(20*log10(snr)), '
dB'])
```

References:

[1] Switched On Bach, Original Recording, Format: Vinyl
Johann Sebastian Bach (Composer),
Walter Carlos (Performer)
<https://www.amazon.com/Bach-Switched-Johann-Sebastian/dp/B000L1H8ZS/ref=sr_1_1?ie=UTF8&qid=1473693687&sr=8-1&keywords=switched+on+bach+vinyl>
See also:
<https://en.wikipedia.org/wiki/Johann_Sebastian_Bach>
<https://en.wikipedia.org/wiki/The_Well-Tempered_Clavier>
One example of the very many WTC recordings:
<https://www.amazon.com/Johann-Sebastian-Bach-Gustav-Leonhardt/dp/B016OQ5NSW/ref=sr_1_8?ie=UTF8&qid=1474108355&sr=8-8&keywords=wohltemperierte+klavier+LP>

[2] Peter J. Baxandall
<https://en.wikipedia.org/wiki/Peter_Baxandall>
<http://www.keith-snook.info/wireless-world-magazine/Wireless-World-1968/Noise%20in%20Transistor%20Circuits%20-%20P%20J%20Baxandall.pdf>

[3] John Linsley Hood
<https://en.wikipedia.org/wiki/John_Linsley_Hood>
<http://www.keith-snook.info/wireless-world-magazine/Wireless-World-1969/Modular%20Pre-amplifier%20Design-DCD.pdf>

[4] On RIAA Equalization Networks
Stanley P. Lipshitz
Journal of the AES, Vol. 27, No. 6, June 1979, pp. 458–481

Comments on "On RIAA Equalization Networks"
Peter J. Baxandall
Journal of the Audio Engineering Society, 29(1), Jan 1981.

Author's Reply to "Comments on 'On RIAA Equalization Networks'"
Stanley P. Lipshitz
Journal of the Audio Engineering Society, 29(1), Jan 1981.

A High Accuracy Inverse RIAA Network
Stanley Lipshitz and *Walt Jung*
The Audio Amateur, Issue 1/1980, pp. 22–24

[5] AES Standard Playback Curve,
Audio Engineering, Vol. 35, No. 1, January 1951, pp. 22 and 45
<https://www.neumann.com/?lang=en&id=about_us_history_part_4>
<https://en.wikipedia.org/wiki/RIAA_equalization>

[6] A Collection of Wireless World Audio Articles
<http://www.keith-snook.info/wireless-world-magazine/wireless-world-articles.html>

[7] <http://www.keith-snook.info/wireless-world-magazine/Wireless-World-1956/Disc%20Recording%20Characteristics%20RIAA%20-%20W%20H%20Livy.pdf>

[8] Hendrik Wade Bode
<https://en.wikipedia.org/wiki/Hendrik_Wade_Bode>
<https://en.wikipedia.org/wiki/Bode_plot>

[9] Designing with Field Effect Transistors
Evans, A.D., (editor),
Siliconix, Inc., 1981, McGraw-Hill