

Transimpedance Amplifier Analysis

Erik Margan

System Description

In **Fig.1** the transimpedance first order system is shown. It consists of an inverting amplifier accepting the input signal in form of a current from a high impedance signal source, such as a photodiode or a semiconductor based detector for radiation particles, and converts it into an output voltage.

The transimpedance at DC and low frequencies is $v_o/i_i = R_f$. However, the high impedance signal source inevitably has a stray capacitance C_i , which deprives the amplifier from the feedback at high frequencies. Therefore the amplifier's feedback loop must be stabilized by a suitably chosen phase margin compensation capacitance C_f . Owing to the presence of these capacitances, and because of the amplifier's own limitations, the system response at high frequencies will be reduced accordingly.

The system analysis follows from the standard circuit theory in Laplace space. Upon the derived equations the system's response can be optimized by a suitable selection of component values.

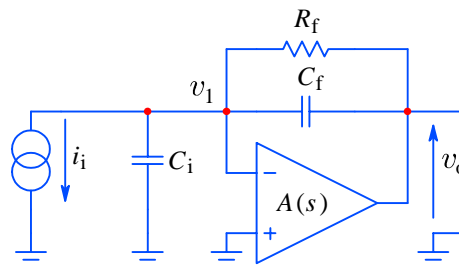


Fig.1: Generalized transimpedance system schematic diagramme

Amplifier Description

The amplifier's inverting open loop voltage gain is modeled as:

$$\frac{v_o}{v_i} = -A(s) = -A_0 \frac{-s_0}{s - s_0} = -A_0 \frac{\omega_0}{s + \omega_0} \quad (1)$$

where:

- s is the complex frequency variable;
- A_0 is the amplifier's open loop DC gain;
- s_0 is the amplifier's real dominant pole, so that:
- $-s_0 = \omega_0 = 2\pi f_0$, and f_0 is the open loop cutoff frequency.

Note: the expression $-s_0/(s - s_0)$ comes from the gain normalization of the function $F(s) = 1/(s - s_0)$, so that $F_N(s) = F(0)/F(s)$. This makes the low frequency gain of $F_N(s)$ unity and independent of s_0 , only its cutoff frequency depends on s_0 .

System Transfer Function

The sum of currents at the v_i node is:

$$i_i = \frac{v_i}{\frac{1}{sC_i}} + \frac{v_i - v_o}{\frac{1}{R_f} + sC_f} \quad (2)$$

From (1) and (2) we have:

$$i_i = \frac{v_o}{-A} sC_i + \left(\frac{v_o}{-A} - v_o \right) \frac{1 + sC_f R_f}{R_f} \quad (3)$$

Reordernig (3) gives:

$$i_i R_f = -v_o \frac{1}{A} [sC_i R_f + (1 + A)(1 + sC_f R_f)] \quad (4)$$

The normalized transfer function is obtained by dividing v_o by $i_i R_f$:

$$\frac{v_o}{i_i R_f} = -A \frac{1}{sC_i R_f + (1 + A)(1 + sC_f R_f)} \quad (5)$$

Since A is a function (1) of s :

$$\frac{v_o}{i_i R_f} = -A_0 \frac{\omega_0}{s + \omega_0} \cdot \frac{1}{sC_i R_f + \left(1 + A_0 \frac{\omega_0}{s + \omega_0}\right)(1 + sC_f R_f)} \quad (6)$$

We multiply the last term in the denominator:

$$\frac{v_o}{i_i R_f} = -A_0 \frac{\omega_0}{s + \omega_0} \cdot \frac{1}{sC_i R_f + 1 + sC_f R_f + A_0 \frac{\omega_0}{s + \omega_0} (1 + sC_f R_f)} \quad (7)$$

and multiply the numerator and the denominator by $(s + \omega_0)$:

$$\frac{v_o}{i_i R_f} = \frac{-A_0 \omega_0}{s^2 (C_i + C_f) R_f + s(1 + \omega_0 C_i R_f + \omega_0 (1 + A_0) C_f R_f) + \omega_0 (1 + A_0)} \quad (8)$$

We divide all the terms by the coefficient of the highest power of s , which is $(C_i + C_f) R_f$, to obtain the canonical form of the normalized transfer function:

$$\frac{v_o}{i_i R_f} = \frac{-\frac{A_0}{1 + A_0} \cdot \frac{\omega_0 (1 + A_0)}{(C_i + C_f) R_f}}{s^2 + s \frac{1 + \omega_0 [C_i + (1 + A_0) C_f] R_f}{(C_i + C_f) R_f} + \frac{\omega_0 (1 + A_0)}{(C_i + C_f) R_f}} \quad (9)$$

The term $A_0/(1 + A_0)$ is the system's DC gain, and it is slightly lower than unity. The error is caused by the finite open loop gain. With A_0 being usually about 10^5 , the error is approximately 10^{-5} , and is independent of frequency, so it can be neglected. The term $\omega_0(1 + A_0)$ represents the transition frequency ω_T (in rad/s) of the open loop amplifier, at which the amplifier has unit gain, $A(\omega_T) = 1$. For high

speed amplifiers, the open loop cutoff frequency f_0 is often between 1 and 10 kHz, so the transition frequency f_T is usually between 100 MHz and 1 GHz, or slightly above.

To design the system, a set of limitations must be accounted for.

In the majority of cases we are given a signal source producing current in response to irradiation (consisting of either photons or particles). The source has a conversion sensitivity S defined as a ratio of the produced instantaneous current i by the irradiation power P_r , or $S = i/P_r$ (in A/W). We would like to have some standard voltage value for a standard amount of input power, say $v_o/P_r = 1 \text{ V}/\mu\text{W}$, or similar, so we need to select a suitable value of the feedback resistance R_f to satisfy the relation $v_o/P_r = R_f S$, so that $v_o = R_f i$ (for constant input).

The source also has a stray capacitance C_i (proportional to the detector's active area, and the dielectric constant of the detector material, and inversely proportional to its thickness). This capacitance will cause a reduction of the feedback signal at high frequencies, so the feedback loop must be phase compensated by a suitably chosen feedback capacitance C_f . The amplifier is chosen on the basis of its noise performance and with enough bandwidth to cover the frequency range of interest, so once the amplifier has been selected, the only element by which we can optimize the system will be the feedback capacitance C_f .

However, even C_f cannot be chosen at will. With a value too large the system will respond slowly, and with a value too small the response may exhibit a large overshoot and long ringing, or even sustained oscillations. A system with a lowest settling time has the poles in conform with a Bessel system. The Bessel system family is optimized for a maximally flat envelope delay up to the system cutoff frequency, therefore all the relevant frequencies will pass through the system with equal delay, and the response will exhibit the fastest possible risetime with minimal overshoot. The optimal component values can be calculated from the system poles, which are then compared to the normalized Bessel poles and scaled accordingly by the system cutoff frequency. By comparing the system transfer function (9) with the general canonical normalized form (10) we obtain two equations from which the poles can be calculated. The general form of the transfer function with only poles is:

$$F(s) = G_0 \frac{(-s_1)(-s_2)}{(s - s_1)(s - s_2)} = G_0 \frac{s_1 s_2}{s^2 + s(-s_1 - s_2) + s_1 s_2} \quad (10)$$

So we can find the system poles from the following two equations:

$$-s_1 - s_2 = \frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{(C_i + C_f)R_f} \quad (11)$$

$$s_1 s_2 = \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f} \quad (12)$$

For the second order Bessel system, normalized to the unity group delay, the values of the poles are:

$$s_{1,2} = -\frac{3}{2} \pm j\frac{\sqrt{3}}{2} = -\frac{3}{2} \left(1 \pm j\frac{\sqrt{3}}{3} \right) \quad (13)$$

To tune the system for the desired response we need to find the system poles from (11) and (12), and with (13) as the guide for achieving the necessary imaginary to real ratio of the pole components.

We start by expressing s_1 from (12):

$$s_1 = \frac{\omega_0(1 + A_0)}{s_2(C_i + C_f)R_f} \quad (14)$$

With this we return to (11):

$$-\frac{\omega_0(1 + A_0)}{s_2(C_i + C_f)R_f} - s_2 = \frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{(C_i + C_f)R_f} \quad (15)$$

We multiply all by $-s_2$ and put everything on the left had side of the equation:

$$s_2^2 + s_2 \frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f} = 0 \quad (16)$$

This is a second order polynomial, and it is solved using the standard textbook expression of a general form:

$$ax^2 + bx + c = 0$$

which has the roots:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = -\frac{b}{2a} \left(1 \pm j \sqrt{\frac{4ac}{b^2} - 1} \right)$$

In our case (16) the coefficient $a = 1$, so we have:

$$\begin{aligned} s_{1,2} &= -\frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{2(C_i + C_f)R_f} \left(1 \pm j \sqrt{\frac{4 \frac{\omega_0(1+A_0)}{(C_i+C_f)R_f}}{\left[\frac{1+\omega_0[C_i+(1+A_0)C_f]R_f}{(C_i+C_f)R_f}\right]^2} - 1} \right) \\ &= -\frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{2(C_i + C_f)R_f} \left(1 \pm j \sqrt{\frac{4\omega_0(1 + A_0)(C_i + C_f)R_f}{\{1 + \omega_0[C_i + (1 + A_0)C_f]R_f\}^2} - 1} \right) \end{aligned} \quad (17)$$

It is difficult to obtain the required value of C_f from this relation, since there is no simple way to solve it analytically. But we can always solve it numerically by entering the known component values and varying C_f until the correct value (13) of the determinand (under the square root) of (17) is obtained:

$$\frac{4\omega_0(1 + A_0)(C_i + C_f)R_f}{\{1 + \omega_0[C_i + (1 + A_0)C_f]R_f\}^2} - 1 = \frac{1}{3} \quad (18)$$

Another possible way would be to calculate the envelope delay of the system for a range of frequencies and varying C_f until the envelope delay is essentially flat up to almost the system's cutoff frequency. We will see the relation for the envelope delay a little later.

In **Fig.2** we have plotted the poles $s_{1,2}$ (17) as a function of changing C_f from 1 pF to 0.1 pF in steps of 0.01 pF. Initially both poles are real and the one closer to the coordinate system's origin is the dominant pole. As C_f decreases the poles move towards each other until they meet, at which point the system is critically damped. By further decreasing C_f the poles form a complex conjugate pair and travel along a circle centered at the origin with a radius equal to the critically damped system values. **Fig.3** shows the transfer function magnitude above the complex plane (Bessel poles).

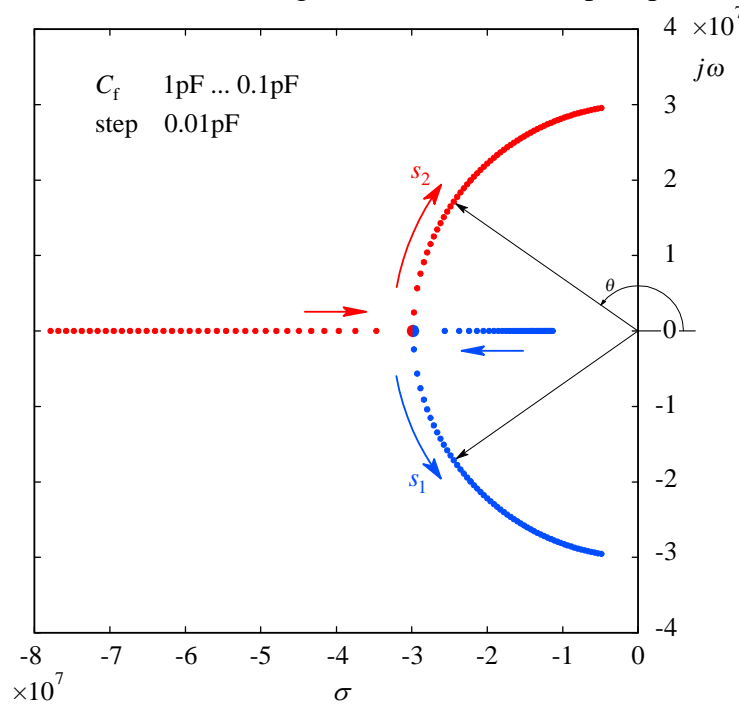


Fig.2: The position of the poles changes with decreasing C_f

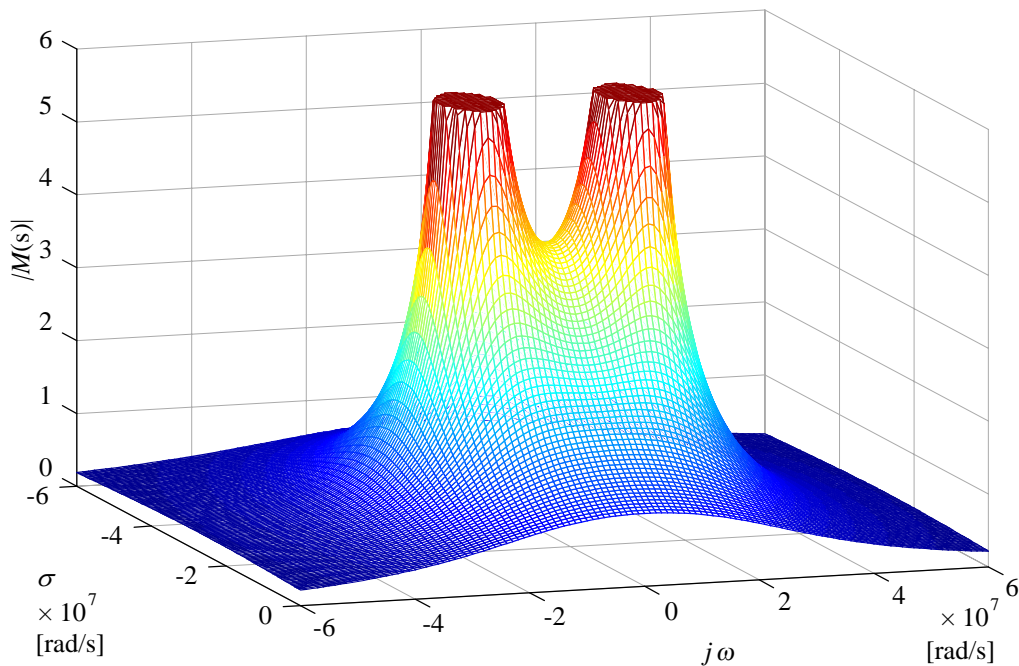


Fig.3: The transfer function magnitude (absolute value) over the complex plane, shown here for the Bessel poles case. Above the poles the magnitude is infinitely high. The shape of the surface cut along the $j\omega$ axis is the system's frequency response (in linear scale).

In most cases the system's high frequency cutoff $f_h = \omega_h/2\pi$, at which the system gain drops by -3 dB relative to its value at low frequencies, can be calculated from the expression for the last characteristic polynomial term (9):

$$\omega_h^2 = \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f} \quad (19)$$

so the cutoff frequency in Hz will be:

$$f_h = \frac{\omega_h}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (20)$$

Note that this holds for only a second order system optimized to have a maximally flat amplitude, or the Butterworth system, where the normalized poles are $s_{1,2} = \sqrt{2}/2 \pm j\sqrt{2}/2 = (\sqrt{2}/2)(1 \pm j)$, and the angle $\theta = \pm 135^\circ$ (as measured from the positive real axis). Because the poles of a Bessel system have a different layout (the angle $\theta = \pm 150^\circ$), the cutoff frequency will be lower by $\sqrt{3}/3$, see (13).

$$f_{h(\text{Bes}2)} = \frac{\sqrt{3}}{3} f_{h(\text{But}2)} \quad (21)$$

Transfer Function Magnitude ('Frequency Response')

The usual meaning of the term 'frequency response' is the magnitude (absolute value) of the complex frequency transfer function. The magnitude can be calculated as the square root of a product of the transfer function with its own complex conjugate:

$$|F(s)| = \sqrt{F(s) \cdot F^*(s)} \quad (22)$$

Ordinarily we are not interested in the shape of the transfer function magnitude over the entire Laplace complex plane (as in **Fig.3**), but only in the shape of that surface cut along the imaginary axis $s = j\omega$ (with $\sigma = 0$):

$$|F(j\omega)| = \sqrt{F(j\omega) \cdot F(-j\omega)} \quad (23)$$

We thus rewrite (9) in response to $j\omega$:

$$\frac{v_o}{i_i R_f} = \frac{-\frac{A_0}{1 + A_0} \cdot \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}}{(j\omega)^2 + j\omega \frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (24)$$

We need first to separate the real and imaginary terms and rationalize the denominator, then multiply the relation by its own complex conjugate, as in (22), (23):

$$\frac{v_o}{i_i R_f} = \frac{-\frac{A_0}{1 + A_0} \cdot \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}}{(j\omega)^2 + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f} + j\omega \frac{1 + \omega_0[C_i + (1 + A_0)C_f]R_f}{(C_i + C_f)R_f}} \quad (25)$$

To keep the equations short we shall use the general form (10), so (25) becomes:

$$\frac{v_o}{i_i R_f} = \frac{-G_0 \cdot s_1 s_2}{(j\omega)^2 + s_1 s_2 + j\omega(-s_1 - s_2)} \quad (26)$$

The denominator is rationalized by multiplying both the numerator and the denominator by the denominator's complex conjugate:

$$\frac{v_o}{i_i R_f} = \frac{-G_0 \cdot s_1 s_2 \cdot [(-j\omega)^2 + s_1 s_2 - j\omega(-s_1 - s_2)]}{[(j\omega)^2 + s_1 s_2 + j\omega(-s_1 - s_2)][(-j\omega)^2 + s_1 s_2 - j\omega(-s_1 - s_2)]} \quad (27)$$

The following terms are present in the denominator after the multiplication:

$$\begin{aligned} & -\omega^4 \\ & + s_1^2 s_2^2 \\ & + \omega^2(-s_1 - s_2)^2 \\ & - \omega^2 s_1 s_2 + \omega^2 s_1 s_2 \\ & - j\omega^3(-s_1 - s_2) + j\omega^3(-s_1 - s_2) \\ & - j\omega(-s_1 - s_2)s_1 s_2 + j\omega(-s_1 - s_2)s_1 s_2 \end{aligned} \quad (28)$$

The last three lines contain terms with alternating signs and they cancel. So the rationalized form is:

$$\frac{v_o}{i_i R_f} = \frac{-G_0 \cdot s_1 s_2 \cdot [\omega^2 + s_1 s_2 - j\omega(-s_1 - s_2)]}{s_1^2 s_2^2 + \omega^2(-s_1 - s_2)^2 - \omega^4} \quad (29)$$

The imaginary unit is contained only in the last term in the brackets of the numerator, so the complex conjugate will only occur there. The common (squared) terms can be extracted from the square root. Then the absolute value is:

$$\left| \frac{v_o}{i_i R_f} \right| = \frac{G_0 \cdot s_1 s_2 \sqrt{[\omega^2 + s_1 s_2 - j\omega(-s_1 - s_2)][\omega^2 + s_1 s_2 + j\omega(-s_1 - s_2)]}}{s_1^2 s_2^2 + \omega^2(-s_1 - s_2)^2 - \omega^4} \quad (30)$$

After multiplying of the brackets under the root the imaginary terms will have alternate signs and cancel, so the system's transfer function magnitude is:

$$\left| \frac{v_o}{i_i R_f} \right| = \frac{G_0 \cdot s_1 s_2 \sqrt{(\omega^2 + s_1 s_2)^2 + \omega^2(-s_1 - s_2)^2}}{s_1^2 s_2^2 + \omega^2(-s_1 - s_2)^2 - \omega^4} \quad (31)$$

Phase Angle

We calculate the phase angle φ of the transfer function of the n^{th} order system as the arctangent of the ratio of the imaginary to real part of the transfer function, which is equivalent to finding the individual phase shift $\varphi_k(\omega)$ of each pole $s_k = \sigma_k \pm j\omega_k$ and then summing them:

$$\varphi(\omega) = \arctan \frac{\Im\{F(s)\}}{\Re\{F(s)\}} = \sum_{k=1}^n \varphi_k(\omega) = \sum_{k=1}^n \arctan \frac{\omega \mp \omega_k}{\sigma_k} \quad (32)$$

Our frequency response function (32) has two complex conjugate poles, therefore the phase response is:

$$\varphi(\omega) = \arctan \frac{\omega - \omega_1}{\sigma_1} + \arctan \frac{\omega + \omega_1}{\sigma_1} \quad (33)$$

Envelope Delay

We obtain the envelope delay as the phase derivative against frequency:

$$\tau_e = \frac{d\varphi(\omega)}{d\omega} \quad (34)$$

Because the phase response (33) is a sum of individual phase shifts for each pole, the same is true for the envelope delay. Each pole contributes a delay:

$$\frac{d\varphi(\omega)}{d\omega} = \frac{d}{d\omega} \left[\arctan \frac{\omega \mp \omega_i}{\sigma_i} \right] = \frac{\sigma_i}{\sigma_i^2 + (\omega \mp \omega_i)^2} \quad (35)$$

and the total envelope delay is the sum of the contributions of each pole.

For the 2-pole case we have:

$$\tau_e = \frac{\sigma_1}{\sigma_1^2 + (\omega - \omega_1)^2} + \frac{\sigma_1}{\sigma_1^2 + (\omega + \omega_1)^2} \quad (36)$$

It is important to note that because the poles of stable systems are on the left side of the complex plane, their real part σ must be negative. In (36) the denominators are sums of squares and thus positive. So the envelope delay is a negative function, the negative sign indicating a time delay. Since this function is a ‘delay’, we might have neglected the negative sign. But there is a deeper meaning in this sign: it reflects the sense of rotation of the phase angle with frequency, and for real stable systems the phase always decreases with frequency. Whenever we see the phase increasing we should watch for the possible source of instability within the system.

In **Fig.4** the system’s transfer function magnitude is plotted, along with the phase angle and the envelope delay (phase derivative against frequency). For this plot the system components have been chosen to conform with a Bessel second order response (constant group delay almost up to the cutoff frequency).

The component values resulting in the Bessel system response are:

Amplifier	Source	Feedback
$A_0 = 10^5$ $\omega_0 = 2\pi \cdot 10^4 \text{ rad/s}$	$C_i = 70 \text{ pF}$	$C_f = 0.58 \text{ pF}$ $R_f = 100 \text{ k}\Omega$

With these components the poles (17) have the following values:

$$s_1 = -2.576 \times 10^7 - j 1.5041 \times 10^7 \text{ rad/s}$$

$$s_2 = -2.576 \times 10^7 + j 1.5041 \times 10^7 \text{ rad/s}$$

The imaginary to real part ratio is $1.5041/2.576 = 0.5839$, which is well within the component tolerances from the ideal value of $\sqrt{3}/3 \approx 0.5774$.

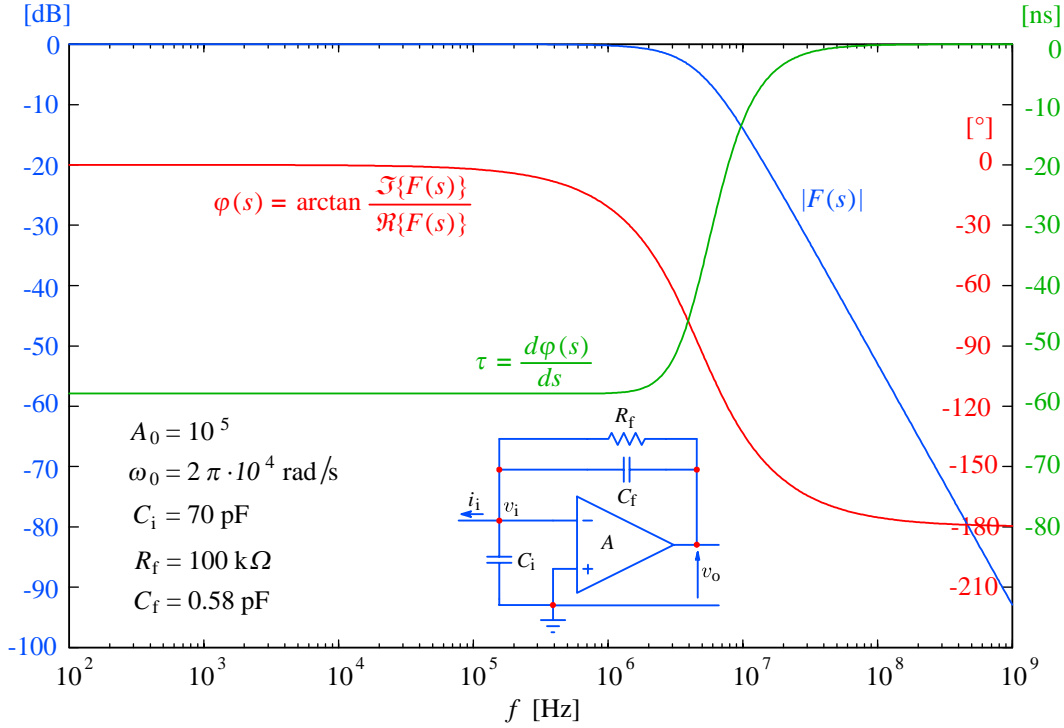


Fig.4: System transfer function (absolute value in dB), phase (in $^\circ$) and group delay (in ns)

Input Impedance Analysis

The amplifier's differential input resistance is assumed to be very high (modern amplifiers having jFET or MOSFET input transistors have their input resistance within the range $10^{12} - 10^{14} \Omega$), and its input capacitance (< 2 pF) can be considered as being a small part of C_i . Then the system's input impedance is:

$$Z_i = \frac{v_i}{i_i} = \frac{v_o}{-A \cdot i_i} = \frac{v_o}{-A_0 \frac{\omega_0}{s + \omega_0} i_i} \quad (37)$$

We need again the relation between the input current i_i and the output voltage v_o (4):

$$i_i R_f = -v_o \frac{1}{A} [sC_i R_f + (1 + A)(1 + sC_f R_f)] \quad (4)$$

which we divide by R_f :

$$i_i = -v_o \frac{1}{AR_f} [sC_i R_f + (1 + A)(1 + sC_f R_f)] \quad (38)$$

By inserting (38) into (37) we have:

$$Z_i = \frac{v_o}{-A \cdot \frac{-v_o}{AR_f} [sC_i R_f + (1 + A)(1 + sC_f R_f)]} \quad (39)$$

We cancel the common terms in the numerator and the denominator:

$$Z_i = \frac{R_f}{sC_i R_f + (1 + A)(1 + sC_f R_f)} \quad (40)$$

Replace A with its full expression from (1):

$$Z_i = \frac{R_f}{sC_i R_f + \left(1 + A_0 \frac{\omega_0}{s + \omega_0}\right) (1 + sC_f R_f)} \quad (41)$$

We multiply the numerator and the denominator by the $s + \omega_0$ term and regroup the coefficients having the same power of s :

$$Z_i = \frac{R_f(s + \omega_0)}{s^2(C_i + C_f)R_f + s[1 + \omega_0 C_i R_f + \omega_0 C_f R_f(1 + A_0)] + \omega_0(1 + A_0)} \quad (42)$$

Divide by the coefficient at the highest power of s :

$$Z_i = \frac{R_f(s + \omega_0) \frac{1}{(C_i + C_f)R_f}}{s^2 + s \frac{1 + \omega_0 C_i R_f + \omega_0 C_f R_f(1 + A_0)}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (43)$$

Make the numerator's frequency dependent term same as the last denominator term:

$$Z_i = \frac{\frac{R_f}{(1 + A_0)}(s + \omega_0) \frac{(1 + A_0)}{(C_i + C_f)R_f}}{s^2 + s \frac{1 + \omega_0 C_i R_f + \omega_0 C_f R_f(1 + A_0)}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (44)$$

From (44) we can extract three terms of the input impedance. The first one is a frequency independent term, which multiplies the two frequency dependent terms:

$$Z_1 = \frac{R_f}{(1 + A_0)} \quad (45)$$

The two frequency dependent terms are: the unity gain normalized band-pass term:

$$Z_2 = \frac{s \frac{(1 + A_0)}{(C_i + C_f)R_f}}{s^2 + s \frac{1 + \omega_0 C_i R_f + \omega_0 C_f R_f(1 + A_0)}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (46)$$

and the unity gain normalized low-pass term:

$$Z_3 = \frac{\frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}}{s^2 + s \frac{1 + \omega_0 C_i R_f + \omega_0 C_f R_f(1 + A_0)}{(C_i + C_f)R_f} + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (47)$$

So the total input impedance (44) is:

$$Z_i = Z_1(Z_2 + Z_3) \quad (48)$$

It is clear from (45) that at low frequencies the input impedance must be very low. Likewise, Z_3 is unity when $s \ll \omega_h$:

$$s \ll \sqrt{\frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (49)$$

and then falls with frequency, owed exclusively to C_i (because C_f is in series with the amplifier's output impedance, which increases at high frequencies, but which we have neglected in this discussion). In contrast, Z_2 increases with frequency and reaches a maximum when $s \approx \omega_h$ and this maximum is proportional to the capacitance ratio:

$$Z_{2max} \approx Z_1 \left(\frac{C_i}{C_f} + 1 \right) \quad (50)$$

after which Z_2 falls off with frequency.

The absolute values (in Ω) of the input impedance and its components are plotted in **Fig.5**.

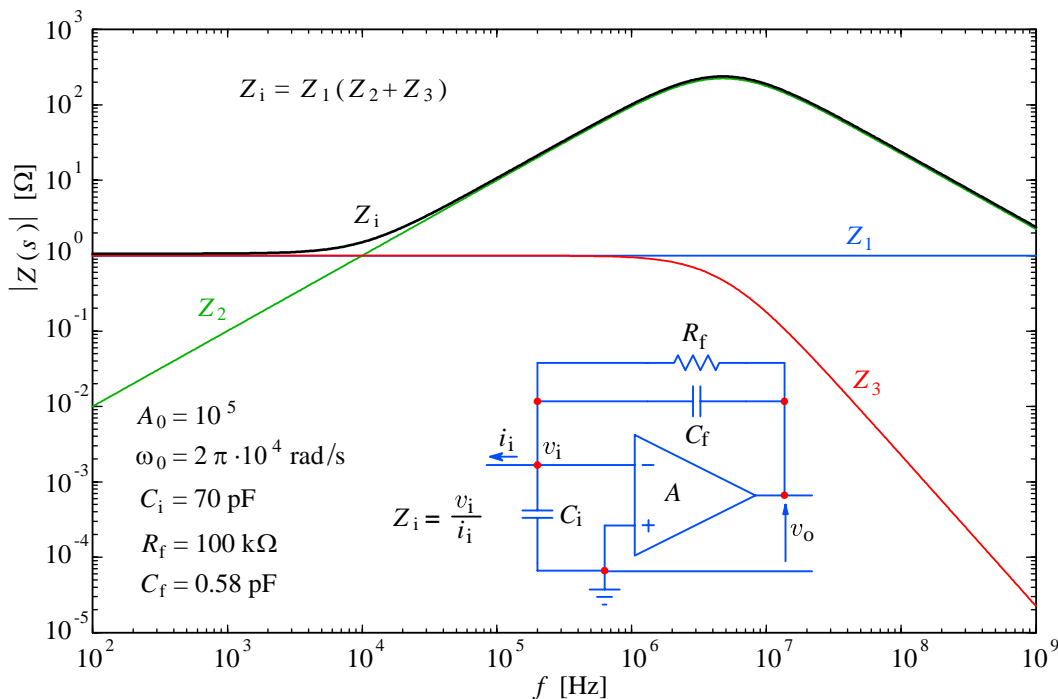


Fig.5: Input impedance with its components (absolute values in Ω)

Time Domain Calculation of the Impulse Response

The system's impulse response can be calculated from the complex transfer function by using the Laplace transform inversion via the Cauchy residue theory. The procedure has been introduced into circuit theory by Oliver Heaviside, who developed the method independently of the existing mathematical knowledge.

The residue of a pole is found from the transfer function by canceling that pole and perform a limiting process in which s approaches the value of that same pole; by repeating the process for all the poles we obtain all the residues. The time domain response is the sum of all the residues.

A general n^{th} order system has n poles. For a system with simple poles (non-repeating), with no zeros in the numerator, and if the impulse response is required, the generalized expression for the residue of the k^{th} pole s_k is:

$$r_k = \lim_{s \rightarrow s_k} (s - s_k) \frac{\prod_{i=1}^n (-s_i)}{\prod_{i=1}^n (s - s_i)} e^{s_k t} \quad (51)$$

So our two-pole system with the pole values as in (13) or (17) will have the residues:

$$r_1 = \lim_{s \rightarrow s_1} (s - s_1) \frac{(-s_1)(-s_2)}{(s - s_1)(s - s_2)} e^{s_1 t} = \frac{s_1 s_2}{(s_1 - s_2)} e^{s_1 t} \quad (52)$$

$$r_2 = \lim_{s \rightarrow s_2} (s - s_2) \frac{(-s_1)(-s_2)}{(s - s_1)(s - s_2)} e^{s_2 t} = \frac{s_1 s_2}{(s_2 - s_1)} e^{s_2 t} \quad (53)$$

In these equations we first cancel the corresponding $(s - s_1)$ and $(s - s_2)$ terms, and then let s assume the value of the particular pole. The system's impulse response is then equal to the sum of the residues:

$$y(t) = r_1 + r_2 = \frac{s_1 s_2}{(s_1 - s_2)} e^{s_1 t} + \frac{s_1 s_2}{(s_2 - s_1)} e^{s_2 t} \quad (54)$$

We can extract the common term:

$$y(t) = \frac{s_1 s_2}{(s_1 - s_2)} (e^{s_1 t} - e^{s_2 t}) \quad (55)$$

We can now write the poles in terms of their real and imaginary part, $s_{1,2} = \sigma_1 \pm j\omega_1$, thus the time domain response (55) can be written as:

$$y(t) = \frac{(\sigma_1 + j\omega_1)(\sigma_1 - j\omega_1)}{(\sigma_1 + j\omega_1 - \sigma_1 + j\omega_1)} [e^{(\sigma_1 + j\omega_1)t} - e^{(\sigma_1 - j\omega_1)t}] \quad (56)$$

We factor out $e^{\sigma_1 t}$ from both exponentials, rearrange the denominator and multiply the numerator to obtain:

$$y(t) = \frac{\sigma_1^2 + \omega_1^2}{2j\omega_1} e^{\sigma_1 t} (e^{j\omega_1 t} - e^{-j\omega_1 t}) \quad (57)$$

Since from Euler's expressions of trigonometric functions follows:

$$\frac{e^{j\omega_1 t} - e^{-j\omega_1 t}}{2j} = \sin \omega_1 t$$

we will have:

$$y(t) = \frac{\sigma_1^2 + \omega_1^2}{\omega_1} e^{\sigma_1 t} \sin \omega_1 t \quad (58)$$

Note that the time domain response of any realizable function is always completely real (the imaginary components cancel)!

Step Response Calculation

The system's step response can be calculated in two ways:

- 1) by the convolution integration of the product of the impulse response (59) and the input unit step; this procedure is easy to execute numerically on a computer, but can be very difficult and often impossible to do it analytically;
- 2) by multiplying the system's transfer function with the Laplace transform of the unit step operator and performing the Laplace transform inversion via residue theory; this process may sometimes be lengthy but is always easily manageable.

We are going to follow the second procedure. The Laplace transform of the unit step function is $1/s$. If we multiply the system's transfer function by $1/s$ we obtain a three pole function, with the new pole at the complex plane origin ($0 + j0$).

$$G(s) = \frac{1}{s}F(s) = \frac{(-s_1)(-s_2)}{s(s-s_1)(s-s_2)} \quad (59)$$

This function has three poles and therefore three residues. We find the residues of the complex conjugate pole pair in the same way as we did for the impulse response:

$$r_1 = \lim_{s \rightarrow s_1} (s - s_1) \frac{(-s_1)(-s_2)}{s(s-s_1)(s-s_2)} e^{s_1 t} = \frac{s_1 s_2}{s_1(s_1 - s_2)} e^{s_1 t} = \frac{s_2}{s_1 - s_2} e^{s_1 t} \quad (60)$$

$$r_2 = \lim_{s \rightarrow s_2} (s - s_2) \frac{(-s_1)(-s_2)}{s(s-s_1)(s-s_2)} e^{s_2 t} = \frac{s_1 s_2}{s_2(s_2 - s_1)} e^{s_2 t} = \frac{s_1}{s_2 - s_1} e^{s_2 t} \quad (61)$$

The residue for the third pole at $s = 0$ will be:

$$r_3 = \lim_{s \rightarrow 0} (s - 0) \frac{(-s_1)(-s_2)}{s(s-s_1)(s-s_2)} e^{st} = \frac{s_1 s_2}{(0 - s_1)(0 - s_2)} e^{0t} = \frac{s_1 s_2}{s_1 s_2} = 1 \quad (62)$$

So our step response will be the sum of these residues:

$$g(t) = r_3 + r_2 + r_1 = 1 + \frac{s_1}{s_2 - s_1} e^{s_2 t} + \frac{s_2}{s_1 - s_2} e^{s_1 t} \quad (63)$$

As before, we can extract the common term from the last two terms:

$$g(t) = 1 + \frac{1}{s_1 - s_2} (s_2 e^{s_1 t} - s_1 e^{s_2 t}) \quad (64)$$

and by writing the poles by their real and imaginary components:

$$g(t) = 1 + \frac{1}{\sigma_1 + j\omega_1 - \sigma + j\omega_1} [(\sigma_1 - j\omega_1) e^{(\sigma_1 + j\omega_1)t} - (\sigma_1 + j\omega_1) e^{(\sigma_1 - j\omega_1)t}] \quad (65)$$

Again we cancel the terms with alternate signs and reorder the expression to obtain:

$$g(t) = 1 + \frac{e^{\sigma_1 t}}{2j\omega_1} [\sigma_1 e^{j\omega_1 t} - j\omega_1 e^{j\omega_1 t} - \sigma_1 e^{-j\omega_1 t} - j\omega_1 e^{-j\omega_1 t}] \quad (66)$$

We regroup the real and imaginary part within the brackets:

$$g(t) = 1 + \frac{e^{\sigma_1 t}}{2j\omega_1} [\sigma_1 (e^{j\omega_1 t} - e^{-j\omega_1 t}) - j\omega_1 (e^{j\omega_1 t} + e^{-j\omega_1 t})] \quad (67)$$

We multiply the brackets by the external exponential term:

$$g(t) = 1 + \frac{e^{\sigma_1 t}}{2j\omega_1} \sigma_1 (e^{j\omega_1 t} - e^{-j\omega_1 t}) - \frac{e^{\sigma_1 t}}{2j\omega_1} j\omega_1 (e^{j\omega_1 t} + e^{-j\omega_1 t}) \quad (68)$$

By moving the denominators to the exponentials with imaginary exponents we get:

$$g(t) = 1 + \frac{\sigma_1}{\omega_1} e^{\sigma_1 t} \frac{e^{j\omega_1 t} - e^{-j\omega_1 t}}{2j} - e^{\sigma_1 t} \frac{e^{j\omega_1 t} + e^{-j\omega_1 t}}{2} \quad (69)$$

By again employing the Euler's trigonometric identities we obtain:

$$g(t) = 1 + \frac{\sigma_1}{\omega_1} e^{\sigma_1 t} \sin \omega_1 t - e^{\sigma_1 t} \cos \omega_1 t \quad (70)$$

And again the resulting step response (70) is a completely real function.

The normalized impulse and step responses are plotted in **Fig.6**.

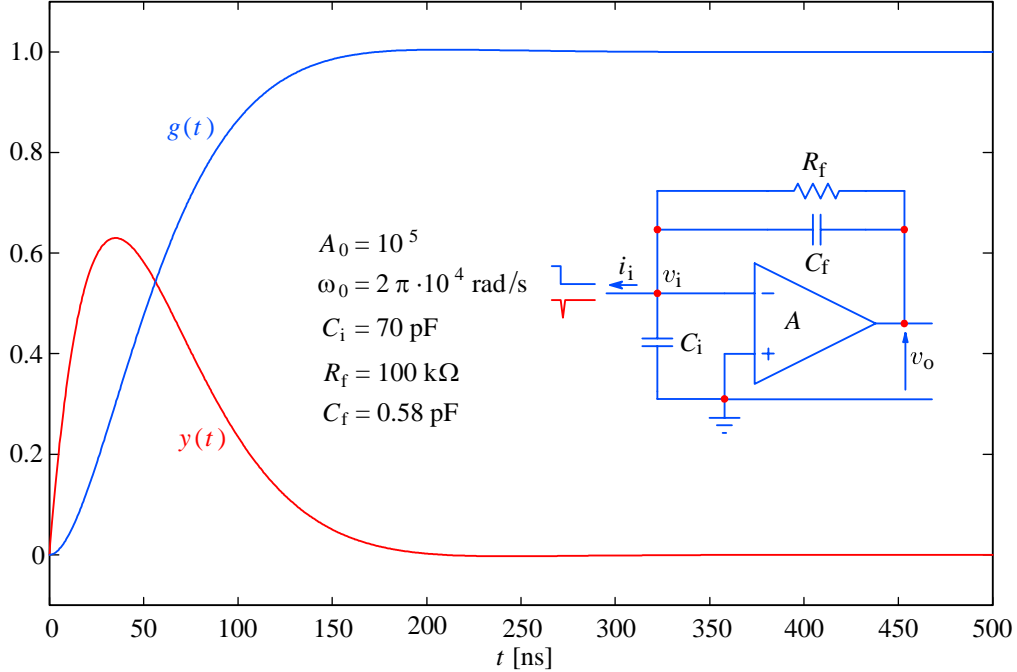


Fig.6: Impulse response $y(t)$ and step response $g(t)$. The ideal second order impulse response rises abruptly from zero; in reality, the presence of a distant non-dominant pole in the amplifier (beyond $A_0\omega_0$) will round up and delay the initial impulse response rising. Because of this non-dominant pole, the step response will also exhibit a slightly increased delay.

Noise Sources and Noise Gain Analysis

The thermal noise sources of the circuit are modeled in **Fig.7**.

The amplifier has two non-coherent noise sources, a differential voltage noise source v_n and the input current noise source i_n . The values of those noise sources are provided by the amplifier's manufacturer in the data sheets. The resistor has its own thermal noise voltage source v_{nR} . All the values are given in terms of noise density

functions (per 1 Hz bandwidth), and to know the actual rms noise we have to account for the system's bandwidth.

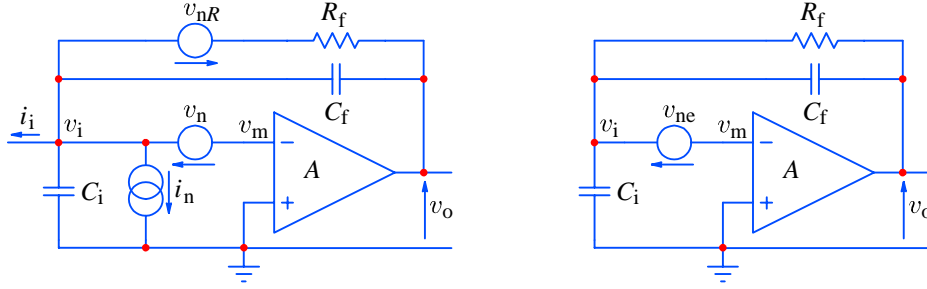


Fig.7: Thermal noise sources and the equivalent total noise source

As shown in **Fig.7**, all the three noise sources are inside the amplifier's feedback loop, therefore we can approximate all the noise sources by a single equivalent source v_{ne} in place of v_n , with a value of:

$$v_{ne} = \sqrt{v_n^2 + (i_n R_f)^2 + v_{nR}^2} \quad (71)$$

Because the noise sources are caused by independent random processes, they are non-coherent, non-correlated, so it is appropriate to sum their powers (voltage or current squared); otherwise the components could be summed directly.

Amplifiers with MOSFET or jFET input transistors usually have their input current noise very low, so in cases when:

$$i_n R_f \leq \frac{v_n}{3}$$

the input current noise can be neglected.

The feedback resistor's thermal noise depends on the resistor's value, its temperature, and the circuit bandwidth $\Delta f = f_H - f_L$. Often $f_L \ll f_H$, therefore the bandwidth can usually be approximated by the upper cutoff frequency of the circuit, $\Delta f \approx f_H = f_h$.

$$v_{nR} = \sqrt{4k_B T R_f \Delta f} \quad (72)$$

where:

k_B is the Boltzmann thermodynamic constant, $k_B = 1.38 \times 10^{-23}$ J/K;

T is the absolute temperature in K; in low power circuits $T \approx 300$ K.

In addition to these noise sources the signal source itself can have its own noise components, i.e., the dark current white noise of a photodiode, which increases with reverse bias and temperature, and also a $1/f$ noise that becomes important in cases where both the amplifier's bandwidth and the signal source bandwidth extend down to DC. However, the signal source noise cannot be distinguished from the signal and is processed by the system in the same way. In contrast, the system's equivalent noise source v_{ne} is inside the feedback loop, and is being processed by the circuit's noise gain.

It is very important to note that the **noise gain is not equal to the signal gain**, in fact it can often be much higher!

The system's noise gain is found by analysing the system's response to the noise source v_{ne} . We start again from the amplifier's inverting input:

$$v_m = -\frac{v_o}{A(s)} = -\frac{v_o}{A_0 \frac{\omega_0}{s + \omega_0}} \quad (73)$$

The voltage at the v_i node is then:

$$v_i = v_m + v_{ne} \quad (74)$$

The current summing at the node v_i is:

$$\frac{v_i}{sC_i} = \frac{v_o - v_i}{\frac{1}{R_f} + sC_f} \quad (75)$$

We regroup the coefficients of the voltage variables:

$$v_i[sC_iR_f + (1 + sC_fR_f)] = v_o(1 + sC_fR_f) \quad (76)$$

By inserting (74) into (76) we obtain:

$$(v_m + v_{ne})[1 + s(C_i + C_f)R_f] = v_o(1 + sC_fR_f) \quad (77)$$

and by replacing v_m with (73):

$$\left(-v_o \frac{s + \omega_0}{A_0\omega_0} + v_{ne}\right)[1 + s(C_i + C_f)R_f] = v_o(1 + sC_fR_f) \quad (78)$$

We again regroup the voltage variables:

$$v_{ne}[1 + s(C_i + C_f)R_f] = v_o \left\{ 1 + sC_fR_f + \frac{s + \omega_0}{A_0\omega_0} [1 + s(C_i + C_f)R_f] \right\} \quad (79)$$

From (79) we obtain the noise gain expression:

$$\frac{v_o}{v_{ne}} = \frac{1 + s(C_i + C_f)R_f}{1 + sC_fR_f + \frac{s + \omega_0}{A_0\omega_0} [1 + s(C_i + C_f)R_f]} \quad (80)$$

By some rearranging we get the final expression:

$$G_n = \frac{v_o}{v_{ne}} = \frac{\frac{A_0}{(1 + A_0)} \left[s\omega_0(1 + A_0) + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f} \right]}{s^2 + s \left[\frac{1}{(C_i + C_f)R_f} + \omega_0 \left(1 + A_0 \frac{C_f}{C_i + C_f} \right) \right] + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (81)$$

The first thing we note is that the noise gain is a non-inverting function, because the sign is positive, which means that the phase of the source at low frequencies is the same as the phase of the output voltage. Further, the term $A_0/(1 + A_0)$ is the frequency independent gain error owed to the finite amplifier's gain; since $A_0 \gg 1$, this term can be neglected.

In a similar way as the input impedance, the noise gain is a sum of two components, one is a band-pass component:

$$G_{nBP} = \frac{s\omega_0(1 + A_0)}{s^2 + s \left[\frac{1}{(C_i + C_f)R_f} + \omega_0 \left(1 + A_0 \frac{C_f}{C_i + C_f} \right) \right] + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (82)$$

and the other is a low-pass component:

$$G_{nLP} = \frac{\frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}}{s^2 + s \left[\frac{1}{(C_i + C_f)R_f} + \omega_0 \left(1 + A_0 \frac{C_f}{C_i + C_f} \right) \right] + \frac{\omega_0(1 + A_0)}{(C_i + C_f)R_f}} \quad (83)$$

In **Fig.8** we have plotted the noise gain (81) and its two components (82), (83).

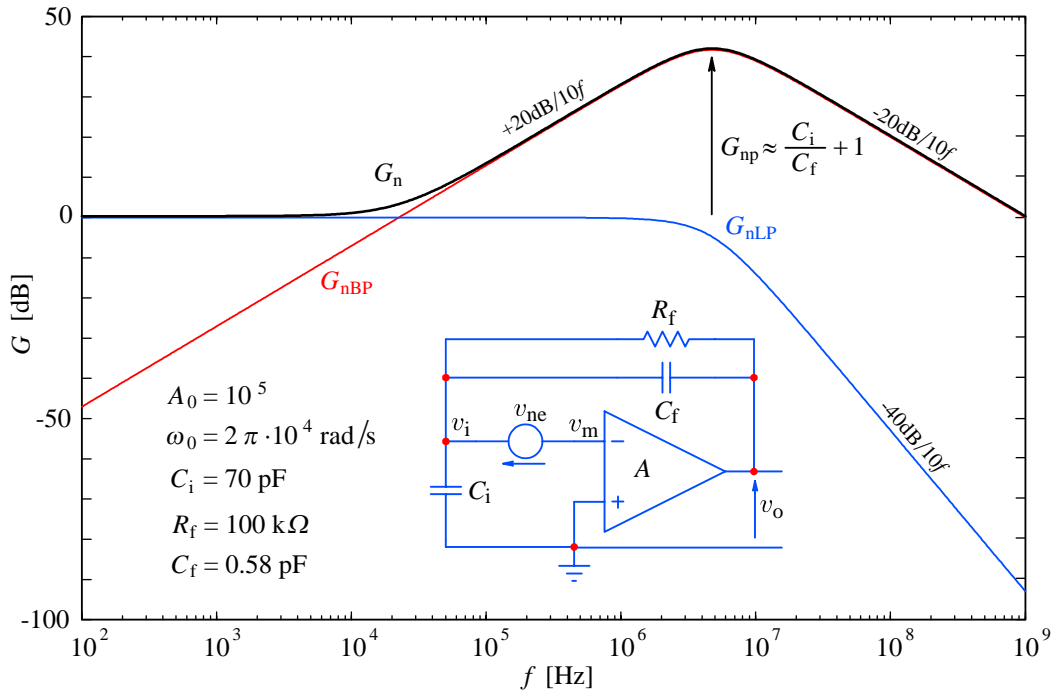


Fig.8: Noise gain G_n and its components

Having determined the noise gain, we want to see its effect on the system's noise spectrum. It is often assumed that the amplifier's shot noise and the resistor's thermal noise are essentially 'white', which (analogous to the white light) means that all the noise frequencies are of equal power. In reality, certain resistor types, like carbon film resistors, have an additional low frequency component noise ('red' or 'excess' noise), inversely proportional to frequency, $\sim 1/f$. Likewise, all amplifiers have the $\sim 1/f$ noise (below about 300 Hz), but certain amplifier types also have a $\sim f$ ('blue') noise component, above 100 kHz. If the system's low frequency cutoff (using an additional high-pass RC filter after the amplifier) is above 1 kHz, we do not need to worry about the low frequency noise. However, the high frequency noise spectrum will be within the system's bandwidth in most cases, so its part below the upper cutoff frequency cannot be neglected.

From the system's noise optimization view it is important to distinguish between the amplifier's shot noise (short for Schottky, or quantum noise) and the resistor's thermal noise (Johnson noise). The shot noise is the consequence of current flow and structural imperfections of the conductor, but in metal conductors it is very low because of the large number of free charge carriers; in semiconductors the number of free charge carriers is much lower and because of the dopants there are relatively many structural imperfections. The shot noise voltage is inversely proportional to the current flow; in FETs it is also independent of temperature, but in bipolar junction transistors it is proportional to temperature because of the base-emitter equivalent resistance $r_e = k_B T / q_e I_e$. In contrast, the thermal noise is present even if there is no current flow in the resistor, and as the name implies it is a function of temperature, actually \sqrt{T} , as will be seen soon. The excess $1/f$ noise in carbon film resistors is proportional to current flow. The $1/f$ noise in amplifiers is mostly proportional to junction leakage because of the reverse bias voltage (say, the collector-base voltage or the drain-gate voltage in jFETs).

If the noise spectrum is not constant with frequency, we must take into account that we will be plotting it in the logarithmic frequency scale, which may influence the shape of the plot. It must be understood that the white noise power spectrum (equal power per 1 Hz bandwidth) can be drawn by a flat line only in a plot with a linear frequency scale. In the usual $\log(f)$ scale every frequency decade is of equal size, so a decade between 1–10 kHz has $10\times$ more 1 Hz bands than the decade between 100 Hz and 1 kHz. Likewise, in the octave between 1–2 kHz there are twice as many 1 Hz bands as between 500 Hz and 1 kHz. This means that a white noise power plotted in the $\log(f)$ scale will be apparently rising in proportion with \sqrt{f} .

However, in amplifiers we are mostly interested in the signal to noise voltage ratio. Since voltage is proportional to the square root of power, the $\log(f)$ scale plot of the white noise voltage will be again constant with frequency. But other types of noise have different spectral distribution, so when plotting those in the $\log(f)$ scale those differences must be taken into account.

Amplifier manufacturers specify the amplifier's current and voltage noise already as a noise density function per 1 Hz bandwidth.

Obviously, in order to have the noise density for the resistor's thermal noise, we must eliminate the Δf from (72). Our R_f of 100 k Ω will thus have the noise density:

$$e_{nR} = \sqrt{4k_B T R_f} = \sqrt{4 \cdot 1.38 \cdot 10^{-23} \cdot 300 \cdot 10^5} = 40.7 \text{ nV}/\sqrt{\text{Hz}} \quad (84)$$

Assume that our amplifier has a voltage noise density of $e_n = 7 \text{ nV}/\sqrt{\text{Hz}}$, and a current noise density of $i_n = 15 \text{ fA}/\sqrt{\text{Hz}}$, thus $i_n R_f = 1.5 \text{ nV}/\sqrt{\text{Hz}}$. Our total equivalent voltage noise density will be:

$$e_{ne} = \sqrt{e_n^2 + (i_n R_f)^2 + e_{nR}^2} = \sqrt{7^2 + 1.5^2 + 40.7^2} \approx 41 \text{ nV}/\sqrt{\text{Hz}} \quad (85)$$

If the amplifier's noise voltage contains a blue component above some frequency f_b , e_n should be multiplied by $\sqrt{1 + f/f_b}$. The noise voltage will start to decrease beyond amplifier's transition frequency owed to the presence of secondary poles (not accounted for in our simplified model).

So our dominant noise source is the resistor's thermal noise. This will be amplified by the system's noise gain:

$$e_{ns} = G_n \cdot e_{ne} \tag{86}$$

Because this function is not constant with frequency, it is not appropriate to simply multiply it by $\sqrt{\Delta f}$. To obtain the actual noise voltage, the expression (86) must be integrated in frequency. Analytical integration will be in most cases very complicated, so it is often done numerically. Alternatively, because the noise sources are not correlated, we may integrate the individual components and then take the root of the sum of squared values. The spectrum of the equivalent noise voltage and its components (some suitably multiplied by the noise gain) is shown in **Fig.9**.

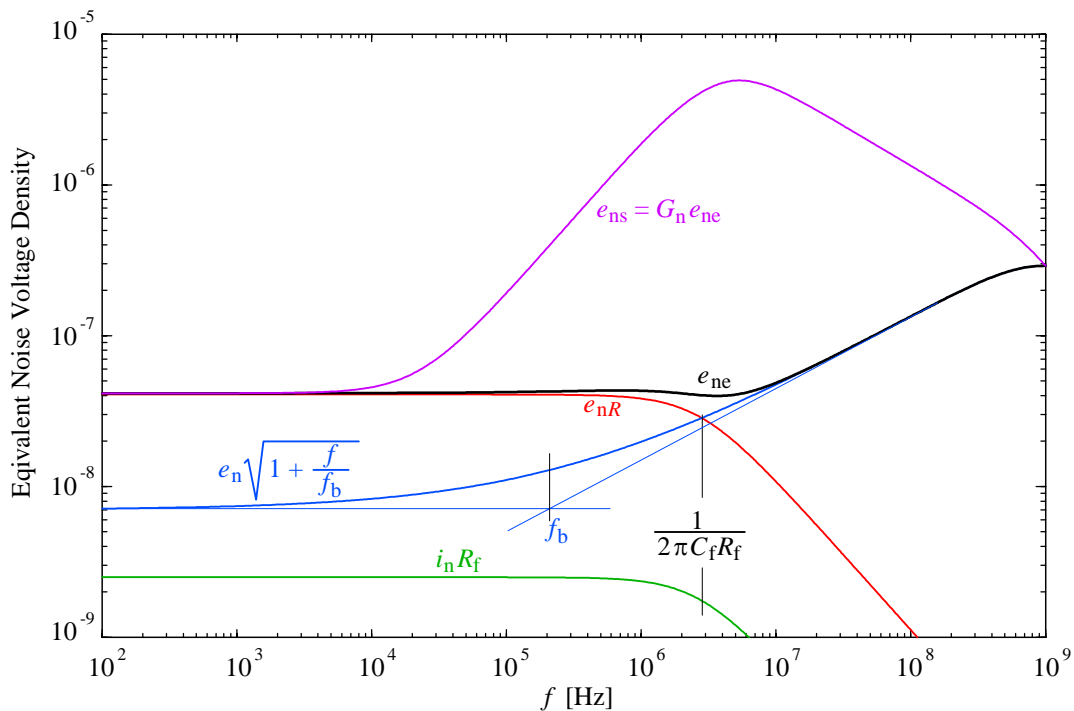


Fig.9: The spectrum of the equivalent noise voltage and its components.

The noise spectral density shows that the dominant noise will be in the frequency range between 5×10^5 and 5×10^7 Hz, where the noise gain has its peak, with values between 2 and $4 \mu\text{V}/\sqrt{\text{Hz}}$. Because of this pronounced peak the noise will appear to be not completely random, but rather having an oscillating component at or near the frequency of the noise gain maximum. This is characteristic for all amplifiers having a pronounced noise gain.