$NF_{\min}$ and base resistance of BJT

skimu@me.com

January 7, 2019

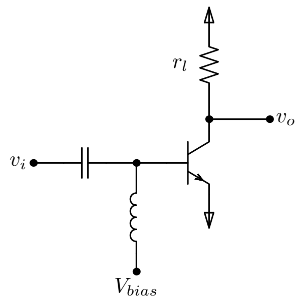
Abstract

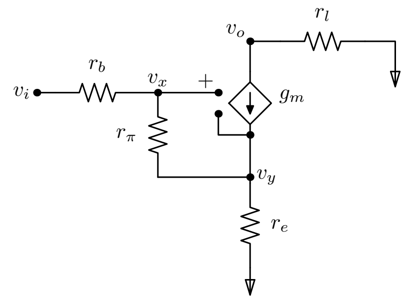
A detailed derivation of $NF_{\min}$ formula for emitter common amplifier, which is widely used as RF preamp, is demonstrated. A lot of text books on RF design explains $NF_{\min}$ formula by adjusting input network or bias current. They are adequate for those who use discrete devices or does not have access to IC design. In this design note, we will derive $NF_{\min}$ formula by finding optimum size (either emitter area or number of parallel devices) for a given DC bias current density. In this way, we can keep the gain and the band-width of the amplifier constant. As the result, $NF_{\min}$ is a sole function of $g_m r_b$ and f_T.

1 $NF_{\min}$ and base resistance

1.1 Emitter common amplifier and its small signal model

The circuit we will analyze today and its small signal model is shown below.





V_{bias} maintains DC collector current density (collector current divided by emitter area) and it is usually come from replica circuit. r_l is the load which is usually the input impedance of the next stage through matching network.

The small signal model for the BJT consists of g_m, r_b, r_π and r_e. The transconductance g_m is equal to I_c/v_t where I_c is the DC bias current and v_t is the thermal voltage (kT/q). Emitter resistance r_e is usually metallic (PTAT) and it is usually small and negligible. However when you use the device at very high current density, the effect would become visible. We include this to see how it affects the performance. For example, $g_m r_e$ is the ratio of the voltage drop across r_e and the thermal voltage (v_t) and we will see effect of r_e when this number gets something like 0.1 or 0.2. Smaller devices in finer process technology tend to have relatively large $g_m r_e$. We have embedded c_{be} into r_π, i.e., r_π is not pure resistance but has reactance component, and we call resulting beta as complex beta:

$$g_m r_\pi = \beta(s) = \frac{\beta_0}{1 + s\tau_\beta}, \quad \tau_T = c_{be}/g_m, \quad \tau_\beta = \beta_0\, \tau_T,$$

where τ_T is the inverse of transition frequency $\tau_T = 1/\omega_T = 1/2\pi f_T$ and β_0 is the DC current gain (the ratio of collector current and base current). The base-collector capacitance (c_μ) is small compared to c_{be} and we take the effect of it as slightly larger c_{be} (Miller approximation). Base resistance is come from metal and contacts as well as silicon and its characteristic is mixture of these. There is r_o in between v_o and v_y. $g_m r_o$ is equal to V_A/v_t, where V_A is the early voltage of the device. We assume $g_m r_o$ is very big compared $g_m r_l$ and ignore it.

Usually, β_0 is in the order of 100, $g_m r_b$ can be anywhere 0.1 to 10, depending on process technology and DC bias current density, it gets larger at higher current density. $g_m r_e$ is can be anywhere negligible to somewhere around one. $g_m r_o$ is usually well more than a thousand.

1.2 Transfer function and input impedance

Equation for node v_o, v_x and v_y is respectively,

$$g_m(v_x - v_y) + \frac{v_o}{r_l} = 0, \qquad \frac{v_x - v_i}{r_b} + \frac{v_x - v_y}{r_\pi} = 0,$$
$$\frac{v_y - v_x}{r_\pi} - g_m(v_x - v_y) + \frac{v_y}{r_e} = 0.$$

Following substitution makes it easier to solve this equation:

$$v_X = v_x, \quad v_Y = v_x - v_y.$$

Therefore

$$\begin{aligned} v_o &= -g_m r_l \, v_Y, \\ v_Y &= \frac{1}{1 + g_m r_e + (g_m r_b + g_m r_e)/\beta(s)}. \end{aligned}$$

Transfer function $A_V = v_o/v_i$ is calculated as follows:

$$\begin{aligned} A_V = v_o/v_i &= -\frac{g_m r_l}{1 + g_m r_e + g_m(r_e + r_b)/\beta(s)}, \\ &= -\frac{g_m r_l}{1 + g_m r_e + g_m(r_e + r_b)/\beta_0} \cdot \frac{1}{1 + s\frac{\tau_\beta}{\beta_0} \cdot \frac{1}{1 + g_m r_e + g_m(r_b + r_e)/\beta_0}}. \end{aligned}$$

From here, we assume that $g_m r_e$ is very small, i.e., the transistor is operated at relatively low DC bias current density. Let's insert $g_m r_e = 0$:

$$A_V = -\frac{g_m r_l}{1 + g_m r_b/\beta(s)} = -\frac{g_m r_l}{1 + g_m r_b/\beta_0} \cdot \frac{1}{1 + s\frac{\tau_\beta}{\beta_0} \cdot \frac{1}{1 + g_m r_b/\beta_0}}.$$

Recalling that $g_m r_b$ is much smaller than β_0:

$$A_V \sim -g_m r_l \cdot \frac{1}{1 + s\,\tau_T}. \qquad (g_m r_b \ll \beta_0)$$

Note that $g_m r_l$ is equal to $I_c r_l / v_t$, which is DC voltage across r_l in the unit of v_t, if r_l is a resistor.

Input impedance (z_i) is

$$\begin{aligned} z_i &= r_b + r_\pi = (g_m r_b + \beta(s))/g_m, \\ &= \frac{g_m r_b + \beta_0}{g_m} \cdot \frac{1 + s\, g_m r_b \tau_\beta / (\beta_0 + g_m r_b)}{1 + s\,\tau_\beta}. \end{aligned}$$

This is about β_0/g_m at low frequencies and r_b at high frequencies. Recalling that $\tau_T = \tau_\beta/\beta_0$ and that $\tau_T = c_{be}/g_m$ and noting that β_0 is likely much larger than $g_m r_b$:

$$z_i \sim \frac{g_m r_b + \beta_0}{g_m} \cdot \frac{1 + s\, g_m r_b \tau_T}{1 + s\,\tau_\beta}.$$

1.3 Noise sources

Each resistance or conductance has at least one noise source, thermal noise. There may be another noise source. That is flicker noise. However we will ignore it because we know that flicker noise is not significant in most cases when we use this amplifier as RF preamp. Each noise source is modeled as a current source in parallel with the resistor associated with it. But we do not know the amount. The only thing we know is power spectral density. We denote power spectral density of noise source i_N as $\left\langle i_N^2 \right\rangle_f$ or $\left\langle i_N^2 \right\rangle_\omega$. The former is with ordinary frequency [A^2/Hz] and the latter is with angular frequency [A^2/(rad/s)]. Although we do not use angular frequency in the following discussion, angular frequency is sometimes more convenient than ordinary frequency, and the relation between these two is $\left\langle i_N^2 \right\rangle_\omega = \left\langle i_N^2 \right\rangle_f /2\pi$.

To find the contribution of a noise source i_N, we first find transfer function $Z(s)$ from the source (i_N) to the output node (v_o), with all other sources set to zero.

$$Z(s) = v_o/i_N.$$

The contribution of the noise source i_N is calculated as follows.

$$\left\langle v_o^2 \right\rangle_f = |Z(s)|^2 \left\langle i_N^2 \right\rangle_f$$

Since noise from different resistance or conductance are independent each other, we can simply sum up to find total noise spectral density.

There are four noise sources in BJT. The first is collector current noise i_c, which is in parallel with g_m. Its spectral density is proportional to the DC bias current:

$$\left\langle i_c^2 \right\rangle_f = 2qI_c = \frac{4kTg_m}{2}.$$

Next will be base current noise i_π, which is in parallel with r_π. Its spectral density is proportional to the DC bias current as well.

$$\left\langle i_\pi^2 \right\rangle_f = 2qI_b = \frac{4kTg_m}{2\beta_0}$$

The third is base resistance noise, which is in parallel with r_b. Its spectral density is

$$\left\langle i_{r_b}^2 \right\rangle_f = \frac{4kT}{r_b} = \frac{4kTg_m}{g_m r_b}.$$

The last is emitter resistance noise, which is in parallel with r_e. Its spectral density is

$$\left\langle i_{r_e}^2 \right\rangle_f = \frac{4kT}{r_e} = \frac{4kTg_m}{g_m r_e}.$$

There must be noise source associated with r_o but since r_o is very large compared to other resistances we can ignore it.

$4kTg_m$ is baseline noise amount and $1/2$, $1/2\beta_0$, $g_m r_b$ and $g_m r_b$ are called the noise factor of each noise source.

1.4 Input referred noise

Set v_i to the ground and calculate contributions for each noise source (i_N) to the output (v_o) and sum everything up. The input referred noise is noise spectral density at the output referred back to the input, i.e., divide by $|A_V|^2$. Here we set $r_e = 0$.

Base resistance Place current source i_N in parallel with r_b. Spectral density of base resistance noise is $\langle i_N^2 \rangle_f = 4kT/r_b$. Equation for node v_x and v_o is respectively,

$$i_N + \frac{v_x}{r_b} + \frac{v_x}{r_\pi} = 0, \qquad g_m v_x + \frac{v_o}{r_l} = 0.$$

Therefore

$$v_o = r_b \; i_N \; A_V.$$

Spectral density is

$$\langle v_o^2 \rangle_f = r_b^2 \; \langle i_N^2 \rangle_f \; |A_V|^2 = \frac{4kT}{g_m} \cdot g_m r_b \; |A_V|^2 .$$

Base current Place current source i_N in parallel with r_π. It is the same place as in r_b, since we set v_i to the ground. Therefore the transfer function is the same as above. Spectral density of base current noise is $\langle i_N^2 \rangle_f = 2qI_B = 2kTg_m/\beta_0$. Therefore

$$\langle v_o^2 \rangle_f = r_b^2 \; \langle i_N^2 \rangle_f \; |A_V|^2 = \frac{4kT}{g_m} \cdot \frac{(g_m r_b)^2}{2\beta_0} \, |A_V|^2 .$$

This is very small compared to base resistance noise. ($g_m r_b \ll \beta_0$)

Collector current Place current source i_N between collector and emitter. Spectral density of collector current noise is $\langle i_N^2 \rangle_f = 2qI_c$. Equation for node v_o:

$$\frac{v_o}{r_l} + i_N = 0.$$

Therefore

$$v_o = -i_N \; r_l.$$

Spectral density:

$$\langle v_o^2 \rangle_f = \langle i_N^2 \rangle_f \; r_l^2 = 2qI_c \, r_l^2 = \frac{4kT}{g_m} \cdot \frac{(g_m r_l)^2}{2}.$$

Load resistor Since load resistor is input impedance from the next stage, noise from it belongs to the next stage.

Total noise In the end, total noise spectral density on the output is sum of base resistance noise and collector current noise.

$$\left\langle v_{\text{total}}^2 \right\rangle_f \sim \frac{4kT}{g_m} \left(g_m r_b \left| A_V \right|^2 + \frac{(g_m r_l)^2}{2} \right).$$

Input referred:

$$\left\langle v_n^2 \right\rangle_f = \frac{\left\langle v_{\text{total}}^2 \right\rangle_f}{\left| A_V \right|^2} = \frac{4kT}{g_m} \left(g_m r_b + \frac{(g_m r_l)^2}{2 \left| A_V \right|^2} \right).$$

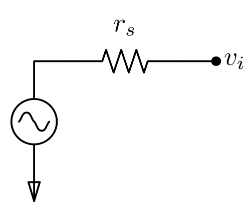
Recalling that $A_V = -g_m r_l / \left(1 + g_m r_b / \beta(s)\right)$,

$$\begin{aligned}
\left\langle v_n^2 \right\rangle_f &= \frac{4kT}{g_m} \left(g_m r_b + \frac{(g_m r_l)^2}{2} \cdot \left| \frac{1 + g_m r_b / \beta}{g_m r_l} \right|^2 \right), \\
&= \frac{4kT}{g_m} \left(g_m r_b + \frac{1}{2} \left(1 + g_m r_b \left(\frac{1}{\beta} + \frac{1}{\beta^*} \right) + \frac{(g_m r_b)^2}{|\beta|^2} \right) \right) \\
&= \frac{4kT}{g_m} \left(g_m r_b \left(1 + \frac{1}{\beta_0} \right) + \frac{1}{2} + \frac{(g_m r_b)^2}{2 \left| \beta \right|^2} \right), \\
&\sim \frac{4kT}{g_m} \left(\frac{1}{2} + g_m r_b + \frac{(g_m r_b)^2}{2 \left| \beta \right|^2} \right).
\end{aligned}$$

We have used the fact that $\beta_0 \gg 1$ in the last formula. The first term is from collector current and you have to have this to get transconductance. Note that input referred noise does not depend on r_l or the gain of the amplifier. This can also be written as follows.

$$\left\langle v_n^2 \right\rangle_f = 4kT \left(r_b + \frac{1}{2 g_m} + \frac{g_m r_b \cdot r_b}{2 \left| \beta \right|^2} \right).$$

1.5 Noise figure



To take source noise into account we can put equivalent noise resistance in series, i.e., substitute $r_s + r_b$ for r_b. r_s is often the same as source impedance.

$$\left\langle v_n^2 \right\rangle_f = 4kT \left((r_s + r_b) + \frac{1}{2 g_m} + \frac{g_m \left(r_s + r_b \right)^2}{2 \left| \beta \right|^2} \right).$$

Noise figure F is the ratio of input referred noise and source noise:

$$F = \frac{\left\langle v_n^2 \right\rangle_f}{4kT r_s} = 1 + \frac{r_b}{r_s} + \frac{1}{2\, g_m r_s} + \frac{g_m \left(r_s + r_b \right)^2}{2\, r_s \left| \beta \right|^2}.$$

We see that there is an optimum r_s for a given operating condition (g_m and r_b) or an optimum transistor size for a given r_s and collector bias current density. Here, we would like to find optimum transistor size[1] N for a given r_s while keeping collector current per unit size J_c constant so that we can keep f_T (bandwidth of the amplifier) constant. g_m and r_b can be written as a function of N as follows:

$$g_m = N \cdot J_c/v_t, \quad r_b = \rho/N,$$

where ρ is unit base resistance at collector bias current density J_c. Note that $g_m r_b = \rho\, J_c/v_t$ stays constant. Now F can be written as a function of N:

$$F = 1 + \frac{g_m r_b}{|\beta|^2} + \left(\rho + \frac{v_t}{2J_c} + \frac{\rho\, g_m r_b}{2\,|\beta|^2}\right)\frac{1}{r_s N} + \frac{J_c}{2\,|\beta|^2\, v_t} \cdot r_s N.$$

Therefore, minimum of noise figure $F_{\min}$ and corresponding $N_{\min}$ is respectively[2],

$$\begin{aligned}
F_{\min} &= 1 + \frac{g_m r_b}{|\beta|^2} + \frac{1}{|\beta|}\sqrt{\frac{1}{4} + \frac{g_m r_b}{2} + \frac{(g_m r_b)^2}{4\,|\beta|^2}}, \\
N_{\min} &= \frac{\rho|\beta|}{r_s}\sqrt{\frac{1}{(g_m r_b)^2} + \frac{2}{g_m r_b} + \frac{1}{|\beta|^2}}.
\end{aligned}$$

And collector bias current $I_{\min}$ which gives $F_{\min}$:

$$I_{\min} = J_c \cdot N_{\min} = \frac{v_t\,|\beta|}{r_s}\sqrt{1 + 2g_m r_b + \frac{(g_m r_b)^2}{|\beta|^2}},$$

or the transconductance which gives $F_{\min}$ is

$$g_m \text{ that gives } F_{\min} \; = \frac{I_{\min}}{v_t} = \frac{|\beta|}{r_s}\sqrt{1 + 2g_m r_b + \frac{(g_m r_b)^2}{|\beta|^2}}.$$

Note that $F_{\min}$ does not depend on the input source impedance (r_s), and that $|\beta|$ is roughly ω_T/ω at $\omega \gg \omega_\beta$ where most of our signals are at, i.e., $|\beta|$ does not even depend on β_0 at frequencies of our interest.

[1] This can be either number of parallel device or emitter area.

[2] See Appendix.

Appendix: Useful formula

BJT Model

$$g_m = I_c/v_t, \qquad g_m r_o = V_A/v_t, \qquad g_m r_\pi = \beta(s) = \frac{\beta_0}{1+s\,\tau_\beta}$$

$$\frac{1}{1+\beta(s)} = \frac{1}{1+\beta_0}\cdot\frac{1+s\,\tau_\beta}{1+s\,\tau_\beta/(1+\beta_0)} \sim \frac{1}{1+\beta_0}\cdot\frac{1+s\,\tau_\beta}{1+s\,\tau_T}$$

$$\frac{\beta(s)}{1+\beta(s)} = \frac{\beta_0}{1+\beta_0}\cdot\frac{1}{1+s\,\tau_\beta/(1+\beta_0)} \sim \frac{\beta_0}{1+\beta_0}\cdot\frac{1}{1+s\,\tau_T}$$

Minimum value

$$\min\left(\frac{A}{x}+Bx\right) = \sqrt{AB}, \qquad x_{\min} = \sqrt{\frac{A}{B}}.$$

Integral

$$\int_0^\infty \frac{d\omega/2\pi}{1+(\omega/\omega_0)^2} = \frac{\omega_0}{2\pi}\tan^{-1}\left(\frac{\omega}{\omega_0}\right)\Big|_0^\infty = \frac{1}{4}\omega_0$$

Two pole transfer function and its step response

$$\frac{1}{1+s\,b+s^2\,a} = \frac{1}{(1+s\,\tau_\oplus)(1+s\,\tau_\ominus)} \qquad (a>0,\ b>0)$$

$$1/\tau_{\oplus,\ominus} = \frac{b\pm\sqrt{b^2-4a}}{2a} = \frac{b}{2a}\left(1\pm\sqrt{1-4a/b^2}\right)$$

Discriminant $4a/b^2$:

$$\begin{aligned} 4a/b^2 < 1 \rightarrow &\quad \text{Exponential settling} \\ = 1 \rightarrow &\quad \text{Critical damping} \\ > 1 \rightarrow &\quad \text{Ringing} \end{aligned}$$

If $4a/b^2 \ll 1$,

$$1/\tau_\oplus = b/a - 1/b, \qquad 1/\tau_\ominus = 1/b$$

Canonical form of two pole amplifier

$$A(s) = \frac{N}{Q+s\,B+s^2\,A} = \frac{A_0}{(1+s\,\tau_A A_0)(1+s\,\tau_\oplus)}$$

If $4AQ/B^2 \ll 1$:

$$A_0 = N/Q, \quad \tau_A = B/N, \quad 1/\tau_\oplus = B/A - Q/B$$

Laplace transform

$$\mathcal{L}\{\delta(t)\} = 1, \quad \mathcal{L}\{1\} = \frac{1}{s}, \quad \mathcal{L}\{e^{-t/\tau_1}\} = \frac{\tau_1}{1+s\,\tau_1}, \quad \mathcal{L}\{t/\tau_1\, e^{-t/\tau_1}\} = \frac{\tau_1}{(1+s\,\tau_1)^2}.$$

$$\begin{aligned}
\frac{1}{(1+s\,\tau_1)(1+s\,\tau_2)} &= \frac{1}{\tau_1-\tau_2}\left(\frac{\tau_1}{1+s\,\tau_1} - \frac{\tau_2}{1+s\,\tau_2}\right) \\
\frac{s}{(1+s\,\tau_1)(1+s\,\tau_2)} &= -\frac{1}{\tau_1-\tau_2}\left(\frac{1}{\tau_1}\cdot\frac{\tau_1}{1+s\,\tau_1} - \frac{1}{\tau_2}\cdot\frac{\tau_2}{1+s\,\tau_2}\right) \\
\frac{1+s\,\tau_3}{(1+s\,\tau_1)(1+s\,\tau_2)} &= \frac{1}{\tau_1-\tau_2}\left(\frac{\tau_1-\tau_3}{\tau_1}\cdot\frac{\tau_1}{1+s\,\tau_1} - \frac{\tau_2-\tau_3}{\tau_2}\cdot\frac{\tau_2}{1+s\,\tau_2}\right) \\
\frac{s}{(1+s\,\tau_1)^2} &= \frac{1}{\tau_1^2}\left(\frac{\tau_1}{1+s\,\tau_1} - \frac{\tau_1}{(1+s\,\tau_1)^2}\right)
\end{aligned}$$

$$\begin{aligned}
\frac{1}{s\,(1+s\,\tau_1)} &= \frac{1}{s} - \frac{\tau_1}{1+s\,\tau_1} \\
\frac{1}{s\,(1+s\,\tau_1)(1+s\,\tau_2)} &= \frac{1}{s} - \frac{\tau_1}{\tau_1-\tau_2}\cdot\frac{\tau_1}{1+s\,\tau_1} + \frac{\tau_2}{\tau_1-\tau_2}\cdot\frac{\tau_2}{1+s\,\tau_2} \\
\frac{1}{s\,(1+s\,\tau_1)^2} &= \frac{1}{s} - \frac{\tau_1}{1+s\,\tau_1} - \frac{\tau_1}{(1+s\,\tau_1)^2} \\
\frac{1+s\,\tau_3}{s\,(1+s\,\tau_1)(1+s\,\tau_2)} &= \frac{1}{s} - \frac{\tau_1-\tau_3}{\tau_1-\tau_2}\cdot\frac{\tau_1}{1+s\,\tau_1} + \frac{\tau_2-\tau_3}{\tau_1-\tau_2}\cdot\frac{\tau_2}{1+s\,\tau_2}
\end{aligned}$$

Approximation If $\tau_1 \gg \tau_2$,

$$\begin{aligned}
\frac{s}{(1+s\,\tau_1)(1+s\,\tau_2)} &\sim \frac{1}{\tau_1\tau_2}\left(\frac{\tau_2}{1+s\,\tau_2} - \frac{\tau_2}{\tau_1}\cdot\frac{\tau_1}{1+s\,\tau_1}\right) \\
\frac{1}{s\,(1+s\,\tau_1)(1+s\,\tau_2)} &\sim \frac{1}{s} - \frac{\tau_1}{1+s\,\tau_1} + \frac{\tau_2}{\tau_1}\cdot\frac{\tau_2}{1+s\,\tau_2}
\end{aligned}$$