Communication circuits

Chapter 5

For EE312

by

A. DAHIMENE

Institut d'Electricité et d'Electronique

Université M'hamed Bougara
Boumerdes

B.

Content

Chapter 5

Receiver Circuits

5.1 Basic definitions

Frequency mixing

In this part, we introduce an important technique used in receiver and transmitter design: Frequency mixing. The mixer is a device capable of changing a carrier frequency for any type of modulation. It is based on the frequency translation theorem of Fourier theory. We start first with real signal mixing.

Consider a general bandpass signal $x_{rf}(t) = r(t)\cos\left[\omega_{rf}t + \phi(t)\right]$. If we multiply this signal by a sinewave $x_{lo}(t) = B\cos\omega_{lo}t$, the result is:

$$z(t) = x_{rf}(t)\times x_{lo}(t) = Br(t)\cos\left[\omega_{rf}t + \phi(t)\right]\cos\omega_{lo}t$$

Using trigonometric identities, we see that this signal is the sum of two bandpass signals:

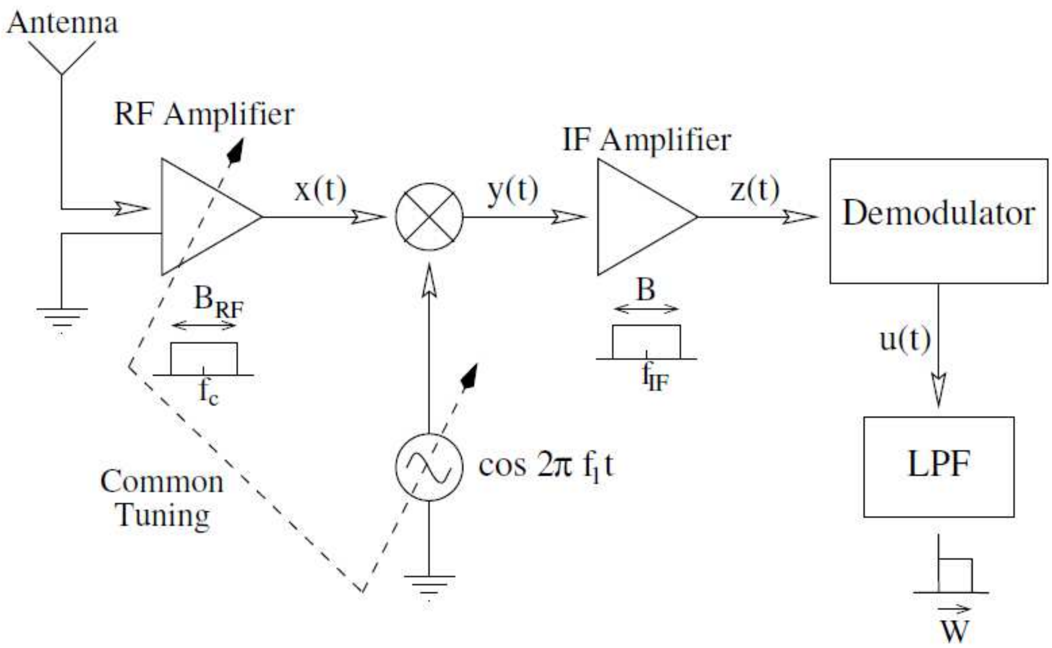
$$z(t) = \frac{B}{2}r(t)\cos\left[\left(\omega_{rf} + \omega_{lo}\right)t + \phi(t)\right] + \frac{B}{2}r(t)\cos\left[\left(\omega_{rf} - \omega_{lo}\right)t + \phi(t)\right].$$

Using a bandpass filter tuned at either the sum or the difference frequency, we obtain a bandpass signal having the same complex envelop (i.e. the same information) but a different carrier frequency. The new frequency is usually called "*intermediate frequency*" f_{if}.

If $f_{if} = f_{rf} + f_{lo}$, we say that we are doing "*up mixing*". On the other hand, if $f_{if} = f_{rf} - f_{lo}$ or $f_{if} = f_{lo} - f_{rf}$, we say that we are performing down mixing.

The "*mixer*" is an important electronic subsystem in any communication receiver or transmitter. It is the basic building block of the "*superheterodyne*" receiver. This concept of receiver was introduced in order to solve the very complex problem of amplifying and selecting one radio station among a large number of stations transmitting at different frequencies.

The first solution that comes to mind is to use a "*tunable*" bandpass filter. However, the construction of a very selective tunable bandpass filter is very complex. Furthermore, due to component aging, such system is prone to random changes and mistuning after a while. It is much easier to build a fixed frequency very selective filter. So, instead of translating the center frequency of a tunable filter before the different signals, it is much easier to translate the frequency of the signals before the center frequency of a fixed bandpass filter. This is the concept of the super heterodyne receiver. The superheterodyne receiver is composed of a tunable local oscillator ganged with a wide band tunable RF amplifier, a mixer and a fixed frequency IF amplifier. It is built using the block diagram shown below.



Let us assume we are using down mixing. The RF amplifier pre-selects a band of frequencies containing a small number of stations around the station at frequency f_c. The bandwidth B_{RF} is large compared to the bandwidth B required by the modulation used (FM, AM, any linear one) but smaller than $2f_{IF}$, the intermediate frequency. Using down mixing, we must have:

$f_{IF} = f_c - f_1$ giving $f_c = f_1 + f_{IF}$ or $f_{IF} = f_1 - f_c$ giving $f_c = f_1 - f_{IF}$.

From the above two relations, we see that if the RF filter does not exist, then we can receive two different stations if we simply use the local oscillator for tuning. These two stations are separated by $2f_{IF}$. These two frequencies are called "*image frequencies*". The job of the tunable rf amplifier is to eliminate one of them so that it will not interfere with the station that we want to receive. Intermediate frequency for broadcast receivers has been standardized to the values of 455 kHz for AM and 10.7 MHz for FM.

Another technique is used to avoid this problem of image frequencies. It is based on frequency translation using complex phasors. Consider the high frequency signal:

$x_{rf}(t) = r(t)\cos\left[\omega_{rf}t + \phi(t)\right]$. The associated analytic signal is: $x_{rf+}(t) = x_{rf}(t) + j\hat{x}_{rf}(t) = r(t)\exp\left(j\phi(t)\right)\exp\left(j\omega_{rf}t\right)$. If we multiply this signal by the phasor: $\exp\left(-j\omega_{lo}t\right)$, we obtain the signal: $z(t) = r(t)\exp\left(j\phi(t)\right)\exp\left(j\left(\omega_{rf} - \omega_{lo}\right)t\right)$. The real part is:

$x_{if}(t) = \text{Re}\left[z(t)\right] = r(t)\cos\left[\left(\omega_{rf} - \omega_{lo}\right)t + \phi(t)\right]$. This is the correct translated signal. So, the process of performing the above operation is:

$$x_{if}(t) = \text{Re}\left[\left(x_{rf}(t) + j\hat{x}(t)\right)\left(\cos\omega_{lo}t - j\sin\omega_{lo}t\right)\right]$$ giving:

$$x_{if}(t) = x_{rf}(t)\cos\omega_{lo}t + \hat{x}_{rf}(t)\sin\omega_{lo}t$$

This leads to the following block diagram:

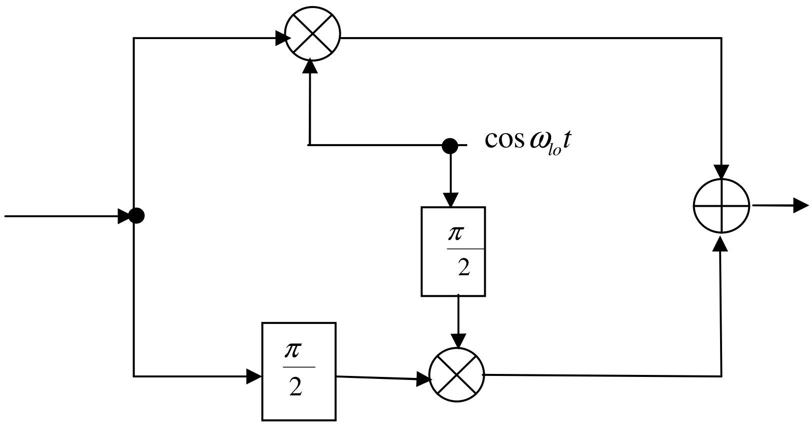


Figure 1 Imageless mixer

The above circuit can be used without any image rejection filter before.

Basic problems in receiver design

So, the front end of a receiver is usually composed of an RF (radio frequency) amplifier that can be tunable but with quite wideband filters, a mixer and a fixed frequency narrowband IF amplifier.

Each of the above stages has specific problems to solve.

a) The RF amplifier is used to select a signal that has very small amplitude and eliminate image frequencies. There are basically two problems that the electronic engineer has to solve in RF design:

1. If the signal that we have to select is very small, the RF amplifier should not add its own noise. So, the design should be oriented toward low noise strategy.

2. If the signal that we have to select is in a crowded environment, the amplifier should be as linear as possible in order to avoid problems of cross modulation and inter modulation.

b) The mixer is a three port element. It should provide the best isolation between the ports. At the same time, it should also simplify the filtering of spurious signals.

c) The IF amplifier is usually a high gain fixed frequency narrowband amplifier. Care should be taken to stability because high gain amplifiers have a tendency to unwanted oscillations.

5.2 RF Amplifiers

In this part of the course, we are going to analyze noise and non linearities in RF design.

Noise in electronic systems

We call noise any waveform that has some random properties and that affects the wanted signal. If the waveform is known, we call it interference. So, in theory, we should use probability theory in order to study noise. However, in our course, we are going to characterize noise as if it were a small deterministic power signal. A noisy signal is characterized by an rms amplitude and band of frequency occupied by the signal. In a receiver, some noisy signals are generated inside the receiver while other are generated in the transmission link and amplified by the receiver.

Thermal noise

Thermal noise is a random signal generated inside pure metals. It is due to the fact that in metals, a crystal is formed by covalent bonds between atoms, however, most of the electron of outer shell are free and form a cloud around positive charged ions. When temperature rises, the ions are going to have larger and larger vibrations around their equilibrium point. A charge in movement amounts to a current and we can model the crystal as a sum of a very large number of small random voltages. So, a random signal will appear across the crystal. This signal can be observed by putting a resistor at the input of an oscilloscope and use a very large gain. So, a metallic resistor can be modelled as a noiseless resistance in series with a noise source.

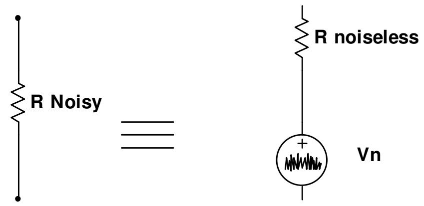


Figure 2 Noisy Resistor

The resistor is observed with a circuit that has a bandwidth Δ*f*. The rms value of the noise voltage (in probability theory, it is called the standard deviation) is given by:

$$v_n^2 = 4kTR\Delta f \tag{1}$$

In the above equation, *R* is the resistance of the resistor in Ω's, *T* is the absolute temperature in °*K*, $k = 1.38\ 10^{-23}$ *J/°K* is Boltzmann's constant and Δ*f* is the band of frequencies in *Hz*. The quantity 4*kTR* is a power density (measured in V^2/Hz) that is independent on frequency. This type of noise is called white because we find that it has the same density at all frequencies just as white light is a sum of lightwaves at all visible wavelengths. If we have two resistors in series with values R_1 and R_2 at the same temperature, the rms voltage across the series combination of the two resistors is:

$$v_{R_1+R_2}^2 = 4kT\left(R_1 + R_2\right)\Delta f = v_{R_1}^2 + v_{R_2}^2$$

We remark that the total rms voltage squared is the sum of the squares of the rms voltage contributions of each resistor. We obtain the same type of result for a parallel configuration.

$$v_{R_1 // R_2}^2 = 4kT\left(\frac{R_1 R_2}{R_1 + R_2}\right)\Delta f$$

If we have devices that are connected in parallel, it is sometimes easier to use Norton's theorem and model the noisy resistor as a noiseless conductance in parallel with a noise current source.

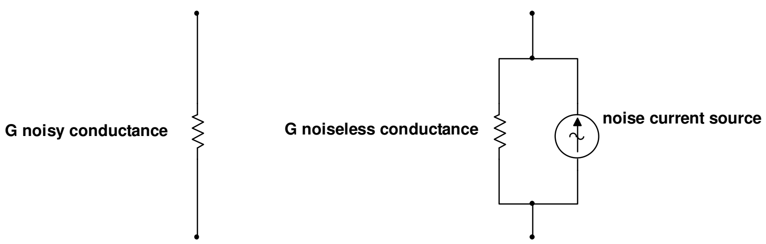


Figure 3 Noisy conductance

Here, the noise current rms value is given by:

$$i_n^2 = 4kTG\Delta f \tag{2}$$

Example: Consider a metal film resistor having $R = 100\ \text{k}\Omega$ in a circuit having a bandwidth $B = 1$ MHz at room temperature (20°C), the rms noise voltage across the resistor will be:

$$v_n = \sqrt{4 \times 1.38\,10^{-23} \times (273+20) \times 100\,10^{3} \times 10^{6}} = 40.6\mu\text{V}$$

This voltage can be observed using the highest sensitivity of an oscilloscope.

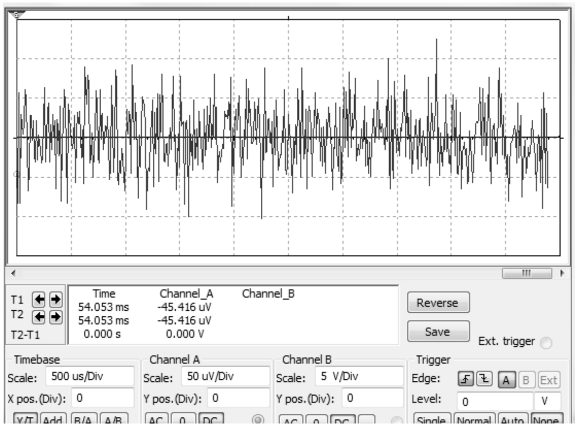


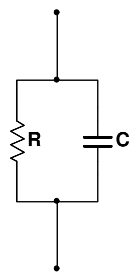
Figure 4 Type of waveform observable using high gain oscilloscope

The above figure shows the type of waveform that can be produced by the above resistor. We can observe the random nature of the signal.

When we have a complex impedance: $Z(f) = R(f) + jX(f)$, the noise voltage will be produced by the resistive part of the impedance. The rms value of the noise voltage that appears across the two terminals used to measure *Z*(*f*) is given by:

$$v_n^2 = \int_0^{+\infty} 4kTR(f)df \tag{3}$$

Example: Consider the parallel connection of a resistance *R* and a capacitance *C*.



The impedance of the above circuit is:

$$Z(f) = \frac{1}{Y(f)} = \frac{1}{\frac{1}{R} + jC2\pi f} = \frac{R}{1+(RC2\pi f)^2} - j\frac{R^2C2\pi f}{1+(RC2\pi f)^2}$$

So, the mean square (square of the rms value) of the voltage is:

$$v_n^2 = 4kTR\int_0^{+\infty} \frac{1}{1+\left(2\pi RC\right)^2} df$$

Making the change of variable $x = 2\pi RCf$, we obtain

$$v_n^2 = \frac{2kT}{\pi C}\int_0^{+\infty} \frac{1}{1+\left(x\right)^2} dx = \frac{2kT}{\pi C}\left[\tan^{-1} x\right]_0^{+\infty} = \frac{kT}{C}$$

We can remark that the noise voltage does not depend on the resistor (which is the device that produces noise) but on the capacitor. This is due to the filtering effect of the capacitor.

Excess noise

When we observe physical resistors, we remark that they produce more noise than the one predicted by equation(1). This noise is called "excess noise" and in general, it is frequency dependent. It decreases with frequency. Since it contains more low frequency than high frequency, it is called "pink noise" by analogy to pink light that contains more red (low frequency) than blue (high frequency) light. When we want to design low noise amplifiers, we should use metal film resistors that produce only thermal noise and avoid carbon types (film or agglomerated) that produce a lot of excess noise.

Shot noise

We have seen that the current through a junction is due to a discrete number of carriers flowing through the junction. The DC current is just the average value and there exists a random fluctuation around this DC value. This fluctuation is called "shot noise". It is also a white noise. Its rms value squared is given by:

$$i_n^2 = 2qI_{DC}\Delta f \tag{4}$$

I_{DC} is the DC current flowing through the junction, $q = 1.6\ 10^{-19}\ C$ is the electron charge and Δf is the bandwidth of the circuit used to observe the signal. This noise is the one that will be produced by the junctions in a biased bipolar junction transistor. In this case also, there exists excess noise that will appear at low frequency.

Field effect transistors also produce noise. We cannot characterize all noise sources. You will see this characterization in more advanced courses.

Antenna temperature

One important source of noise at the input of any radio receiver is the noise picked by the receiving antenna. The antenna has always some bandwidth and it can receive signals that are within that band of frequencies. The antenna is directed toward some region and any signal generated from this direction will be picked by the antenna. This can be electromagnetic energy produced by the sun, the stars, the cosmic noise, man-made noise, etc. We characterize this noise by the antenna temperature. The antenna is characterized by its "radiation resistance". The antenna radiation resistance characterizes the power transfer of the antenna and it is not the "ohmic" resistance of the material used to build the antenna. Let us call this resistance R. The antenna temperature is the temperature that we should apply to a thermal noise resistance of value R to obtain the same noise voltage as the one measured across the antenna, with the same bandwidth Δf. So, if we measure a noise voltage v_n, the antenna temperature will be:

$$T_A = \frac{v_n^2}{4kR\Delta f}$$

It is clear that the antenna temperature is a fictitious temperature. A very directive antenna with a narrow bandwidth will have a small temperature and we refer to it as a cold antenna while an antenna that picks a lot of noise will be qualified as a hot one.

Example: consider an antenna with a radiation resistance of 50Ω picking a noise voltage of 0.1 μV and having a bandwidth of 10 kHz. Its temperature is:

$$T_A = \frac{v_n^2}{4kRB} = 362°K$$

Available noise power

Instead of characterizing noise sources by their voltages or their currents, we can characterize them by the available noise power that they can deliver to a load. First, we have to define this notion of available power. If we consider a voltage source with internal impedance $Z_S = R_S + jX_S$ connected to a load $Z_L = R_L + jX_L$, the source delivers maximum power to the load when we have matching, i.e. when $R_S = R_L$ and $X_S = -X_L$.

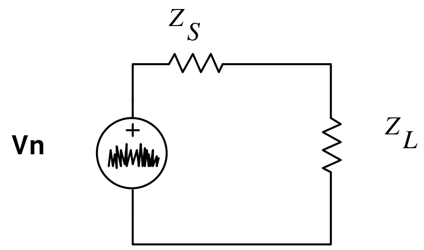


Figure 5 noise source and load

The available power is the maximum power that the source can deliver to a load. It is clear that this power is the power dissipated by the load under matching conditions. Under these conditions, the reactances compensate each other and the load resistance is equal to the source resistance. The rms voltage across the load is equal to half of the rms voltage of the source. So, the available power is:

$$P = \left(\frac{v_n^2}{2}\right)\frac{1}{R_S} = \frac{v_n^2}{4R_s} \tag{5}$$

The available power provided by a metallic resistor generating thermal noise is:

$$P = \frac{4kTR\Delta f}{4R} = kT\Delta f \tag{6}$$

The noise power density is:

$$p = \frac{P}{\Delta f} = kT \qquad \text{in} \quad \text{W/Hz} \tag{7}$$

When we consider electronic systems, we usually characterize them by their power gain. The power gain that we consider here is the available power gain. This gain is the ratio of the available power at the output of the amplifier and the available power at the input of the amplifier.

Signal to noise ratio

At any point in an electronic circuit, we can find some signal power P_S and some noise power P_n. The ratio of the two powers is called the signal to noise ratio (*SNR*).

$$SNR = \frac{P_S}{P_n} \tag{8}$$

Or in dB: $SNR_{dB} = 10\log SNR$

It is clear that a large value of *SNR* is what we should aim to in communication circuits. Depending on the application, there is always a minimum value of *SNR* in order to be able to demodulate signals. We need *SNR* of about 10 *dB* in order to be able to demodulate AM, about 12 *dB* for FM demodulation and even higher values for more complex modulations.

Noise factor, Noise figure and sensitivity

An amplifier is characterized by its gain (we consider the available gain) and the internal noise it produces. The total noise at the output of the amplifier is equal to the amplified noise present at the input added to the internal noise produced by the amplifier.

Noise factor

The noise factor is a quantity that characterizes the noisiness of an amplifier. The amplifier has an available power gain *G*. The noise factor is the following ratio:

$$F = \frac{available\ output\ noise\ power}{available\ output\ noise\ power\ due\ to\ the\ source} \tag{9}$$

Let us call P_{no} the available output noise power, P_{ni} the available input noise power

$$F = \frac{P_{no}}{GP_{ni}} \tag{10}$$

By multiplying and dividing (10) by the available signal power at the input and the output, P_{si} and P_{so}, and using the definition of the signal to noise ratio:

$$SNR_i = \frac{P_{si}}{P_{ni}} = \frac{P_{si}}{kT_0 \Delta f}$$

And $$SNR_o = \frac{P_{so}}{P_{no}}$$

We obtain an alternate definition of the noise factor:

$$F = \frac{SNR_i}{SNR_o} \tag{11}$$

We can write: $P_{no} = GP_{ni} + P_{ao}$, P_{ao} is the noise generated inside the amplifier and appearing at the output. So, we can express the noise factor as:

$$F = \frac{GP_{ni} + P_{ao}}{GP_{ni}} = 1 + \frac{P_{ao}}{GP_{ni}} \tag{12}$$

We introduce now a noise power P_{ai} which is the power we should add to P_{ni} to produce the output noise power P_{no} using a noiseless amplifier. Equation(12) becomes:

$$F = 1 + \frac{P_{ai}}{P_{ni}} \tag{13}$$

In equation(13), we can consider total noise power and the noise factor is called average noise factor or we can consider power densities (measured in *W/Hz*) and we call it "spot noise factor". The input noise power density is by definition taken as the power density of a thermal noise source at a reference temperature $T_0 = 290\ °K$. The noise factor becomes:

$$F = 1 + \frac{P_{ai}}{kT_0} \tag{14}$$

Noise figure

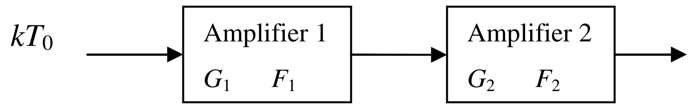
The noise figure is just the noise factor expressed in *dB*.

$$NF = 10 \log F \tag{15}$$

An ideal noiseless amplifier does not add any power to the reference input one. Its noise factor is $F = 1$ or a noise figure $NF = 0\ dB$.

Noise factor of cascaded amplifiers

Consider the following diagram:



We have a cascade of two amplifiers. The input of the first amplifier is a thermal noise at the reference temperature. The added noise power density at the input of the first amplifier is:

$$P_{ai1} = kT_0 \left(F_1 - 1 \right)$$

The added noise power density at the input of the second amplifier is:

$P_{ai2} = kT_0 \left(F_2 - 1 \right)$ so if we consider the cascade of the two amplifiers as noiseless, we can consider this noise as a noise at the input amplified by the first amplifier. This noise is:

$P_{ai2,input} = \frac{kT_0}{G_1} \left(F_2 - 1 \right)$. So, the total added noise at the input of the cascade of the two amplifiers is the sum of the two powers. Finally, the total noise factor is:

$$F = 1 + \frac{P_{ai1} + P_{ai2,input}}{kT_0} = 1 + \frac{kT_0 \left(F_1 - 1 \right) + \frac{kT_0}{G_1} \left(F_2 - 1 \right)}{kT_0}$$

And finally:

$$F = F_1 + \frac{F_2 - 1}{G_1} \tag{16}$$

If we have *N* cascaded amplifiers, we obtain Friis' formula:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_N - 1}{G_1 G_2 \cdots G_{N-1}} \tag{17}$$

Friis' formula shows that it is principally the front end of the receiver that should be optimized for noise, especially if the gain of the preamplifiers is very large.

Noise factor of an attenuator

An attenuator is characterized by a loss *L*, which is the opposite of a gain:

$$L = \frac{P_{in}}{P_{out}}$$

It can be shown that the noise factor of an attenuator is equal to the loss.

$$F = L \tag{18}$$

Example:

Consider a cable having a total attenuation of 10*dB* and a preamplifier having a gain of 10*dB* and a noise figure of 2*dB*. If we connect the cable to the source directly and we amplify the signal with the preamplifier, we have:

$NF_1 = L_{dB} = 10\ dB$. $G_1 = 1/L$ so $G_{1dB} = -10\ dB$. The gains and noise factors are:

$F_1 = 10$, $G_1 = 0.1$ while $F_2 = 1.58$ and $G_2 = 10$.

So, the overall noise factor is $F = 15.8$.

If we consider the total noise referred to the input $(kT_0 + P_{ai})\Delta f$, if the bandwidth is 1 MHz, the result is:

$P_{in} = FkT\Delta f = 6.34\ 10^{-14}\ W$. This is also the noise that will appear at the output because the product of the two gains is one.

If we reverse the cable and the preamplifier, the overall noise factor becomes:

F = 2.48 and the noise power becomes: $P_{in} = FkT\Delta f = 9.94\ 10^{-15}\ W$, which is much smaller. This is why a low noise preamplifier is always located at the antenna location in satellite TV receivers.

Noise temperature

When we consider low noise amplifier, the range of values of the noise factor is very small. If we consider equation(14), we can replace the power density P_{ai} by a thermal noise of a resistance at a temperature T_e: $P_{ai} = kT_e$, and we obtain: $F = 1 + \frac{kT_e}{kT_0}$ giving

$$T_e = T_0\left(F - 1\right) \tag{19}$$

Example:

In the previous example, we have found $F = 2.48$ giving $T_e = 429\ °K$.

If $F = 1.2$ then $T_e = 58\ °K$.

Sensitivity

The sensitivity of a receiver is defined as the amount of signal (power or voltage) required to achieve a given output signal to noise ratio. Using definition(11), we can write:

$SNR_i = F \times SNR_o$, so, the required power at the input is: $P_{si} = F \times P_{ni} \times SNR_o$. Finally:

$$P_{si} = FkT_0\Delta f\left(SNR_o\right) \tag{20}$$

Example:

Consider a system that has an input impedance of 50 Ω, a signal to noise ratio after the amplifier of 0 *dB*, a noise figure of 8 dB and a bandwidth of 10 kHz. What is the *minimum detectable signal*?

From the given data, we have: $F = 6.3$, $kT_0\Delta f = 4\ 10^{-16}$ and $SNR_o = 1$

So, $P_{si} = 6.3\times 4\ 10^{-16} = 2.5\ 10^{-15}\ W$. Since the source resistance is 50 Ω, the minimum detectable voltage is: $v_{in} = \sqrt{4R_sP_{si}} = 7.1\,10^{-7}V = 0.71\mu V$

The above result does not take into account the noise picked by the antenna. This noise is represented by the antenna temperature. If we want to take it into account, we have to modify the value of the input SNR and add to the reference thermal noise the noise $kT_a\Delta f$ added by the antenna. We obtain the following formula for the evaluation of the output signal to noise ratio:

$$SNR_o = \frac{P_{si}}{kT_a\Delta f + FkT_0\Delta f} \qquad (21)$$

So, the minimum detectable signal is:

$$P_{si} = \left(kT_a + FkT_0\right)\Delta f \times SNR_o \qquad (22)$$

Intermodulation, cross-modulation in RF amplifiers

When the amplifier possesses some nonlinearity, and many signals are present, we can encounter problems of intermodulation and cross modulation. To have a better understanding of these problems, let us consider the following input, output characteristic:

$$y = k_1x + k_2x^2 + k_3x^3 + \ldots$$

We assume that we can stop the above development to the third degree. For most small signal preamplifiers, the coefficient k_2 is zero. The voltage gain of the above amplifier is the derivative of the above characteristic.

$$G_V = \frac{dy}{dx} = k_1 + 3k_3x^2$$

In preamplifiers, the power gain G_{PdB} decreases with increasing amplitude. This power gain is equal to $20\log\left|G_V\right|$ plus a constant. So, we must have $k_1 \times k_3 < 0$. In the following, we will assume that the preamplifier is non-inverting. So, $k_1 > 0$ and $k_3 < 0$. The nonlinearity of the preamplifier is quite small, so $\left|\frac{k_3}{k_1}\right| << 1$. If the input is the sum of two sinewaves, we obtain the following result for the output:

Let $x(t) = A_1\cos\omega_1 t + A_2\cos\omega_2 t$, then:

$$
\begin{aligned}
y(t) &= k_1\left(A_1\cos\omega_1 t + A_2\cos\omega_2 t\right)\\
&+k_2\left[A_1^2\frac{1+\cos 2\omega_1 t}{2}+A_2^2\frac{1+\cos 2\omega_2 t}{2}+\frac{A_1A_2}{2}\left(\cos\left(\omega_1+\omega_2\right)t+\cos\left(\omega_1+\omega_2\right)t\right)\right]\\
&+k_3\left\{\left[A_1^3\left(\frac{3\cos\omega_1 t}{4}+\frac{\cos 3\omega_1 t}{4}\right)+A_2^3\left(\frac{3\cos\omega_2 t}{4}+\frac{\cos 3\omega_2 t}{4}\right)\right]\right.\\
&\quad+A_1^2A_2\left[\frac{3}{2}\cos\omega_2 t+\frac{3}{4}\cos\left(2\omega_1+\omega_2\right)t+\frac{3}{4}\cos\left(2\omega_1-\omega_2\right)t\right]\\
&\quad\left.+A_2^2A_1\left[\frac{3}{2}\cos\omega_1 t+\frac{3}{4}\cos\left(2\omega_2+\omega_1\right)t+\frac{3}{4}\cos\left(2\omega_2-\omega_1\right)t\right]\right\}
\end{aligned}
$$

There exist many frequencies at the output. We can list them along with their amplitudes in the following table.

Frequency	Amplitude
ω_1	$k_1A_1+k_3\left(\frac{3A_1^3}{4}+\frac{3A_1A_2^2}{2}\right)$
ω_2	$k_1A_2+k_3\left(\frac{3A_2^3}{4}+\frac{3A_2A_1^2}{2}\right)$
$\omega_1-\omega_2$	$k_2\frac{A_1A_2}{2}$
$\omega_1+\omega_2$	$k_2\frac{A_1A_2}{2}$
$2\omega_1$	$k_2\frac{A_1^2}{2}$
$2\omega_2$	$k_2\frac{A_2^2}{2}$
$2\omega_1-\omega_2$	$k_3\frac{3A_1^2A_2}{4}$
$2\omega_2-\omega_1$	$k_3\frac{3A_2^2A_1}{4}$
$2\omega_1+\omega_2$	$k_3\frac{3A_1^2A_2}{4}$
$2\omega_2+\omega_1$	$k_3\frac{3A_2^2A_1}{4}$
$3\omega_1$	$k_3\frac{A_1^3}{4}$
$3\omega_2$	$k_3\frac{A_2^3}{4}$

Table 1 Frequencies generated by third order system

Depending on the needed signal and on the amplitude of the unwanted ones, we have several different cases.

Single tone gain compression

If the amplitude A_2 is zero, the amplitude of the output term at the frequency ω_1 is:

$$A_{1out} = k_1 A_1 + k_3 \frac{3A_1^3}{4} = k_1 A_1 \left(1 + \frac{3k_3 A_1^2}{4k_1}\right)$$

So, we have the following relation:

$$20\log A_{1out} = 20\log A_1 + 20\log k_1 + 20\log\left(1 + \frac{3k_3 A_1^2}{4k_1}\right)$$

If the input and output impedances are equal, we have the following relation in *dBm*:

$$P_o = P_i + G_P + 20\log\left(1 + \frac{3k_3 A_1^2}{4k_1}\right) \tag{23}$$

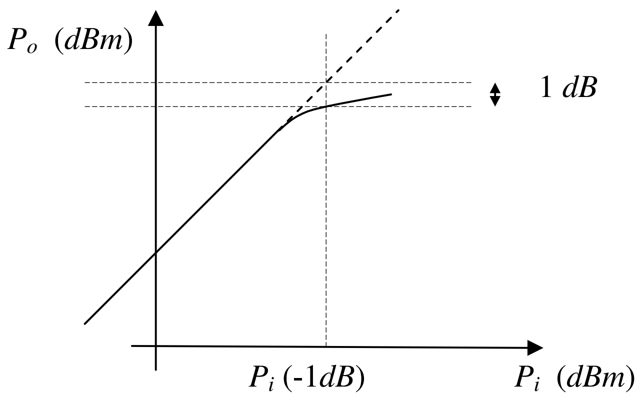


Figure 6 Input Output relation

In Figure 6, the slope of the curve is one for any input power smaller than the point indicated by P_i (−1*dB*). This point is called "1 *dB* compression point". This point indicates the departure from the linear characteristic indicated by the dotted line.

Desensitization

If we are tuned to ω_1 and the RF amplifier has a bandwidth that lets a large signal at ω_2 pass, we can encounter this problem of "desensitization". From Table 1, the amplitude at ω_1 is given by

$$A_{1out} = k_1 A_1 + k_3\left(\frac{3A_1^3}{4} + \frac{3A_1 A_2^2}{2}\right) = k_1 A_1\left(1 + \frac{k_3}{k_1}\left(\frac{3A_1^2}{4} + \frac{3A_2^2}{4}\right)\right) \tag{24}$$

We can see that if A_2 is large, we are going to have a reduction of the output amplitude that will depend on A_2.

Cross-modulation

Let us consider the previous case, but now, the amplitude A_2 is modulated. So, we have $A_2 = A(1+ms(t))$, replacing in equation(24) , we obtain:

$$A_{1out} = k_1A_1 + k_3\frac{3A_1^2}{4} + k_3\frac{2A_1}{2}A^2\left(1+ms(t)\right)^2 = k_1A_1 + k_3\frac{3A_1^2}{4} + k_3\frac{2A_1}{2}A^2\left(1+2ms(t)+m^2s^2(t)\right)$$

It is clear from the above expression that the modulation has passed from the carrier at ω_2 to the carrier at ω_1.

Third order intermodulation distortion

If we are tuned to a frequency ω_0 and the two frequencies ω_1 and ω_2 are quite close to it, they can generate components at $2\omega_1 - \omega_2$ or $2\omega_2 - \omega_1$ that are even closer to ω_0. The amplitudes of these interfering signals are:

$k_3\frac{3A_1^2A_2}{4}$ at $2\omega_1 - \omega_2$ or $k_3\frac{3A_2^2A_1}{4}$ at $2\omega_2 - \omega_1$. This distortion is measured when the two amplitudes A_1 and A_2 are equal to the wanted signal amplitude. So, the amplitude of this third order intermodulation distortion becomes:

$$A_{3third} = k_3\frac{3A_1^3}{4} \quad (25)$$

In terms of power in *dBm*, the third order interfering power is given by:

$$P_{3third} = 20\log k_3\frac{3A_1^3}{4} = 3P_i + 20\log\frac{3k_3}{4}\ dBm \quad (26)$$

If we plot this power in the same graph as the input output one shown in Figure 6, we obtain a line of slope three.

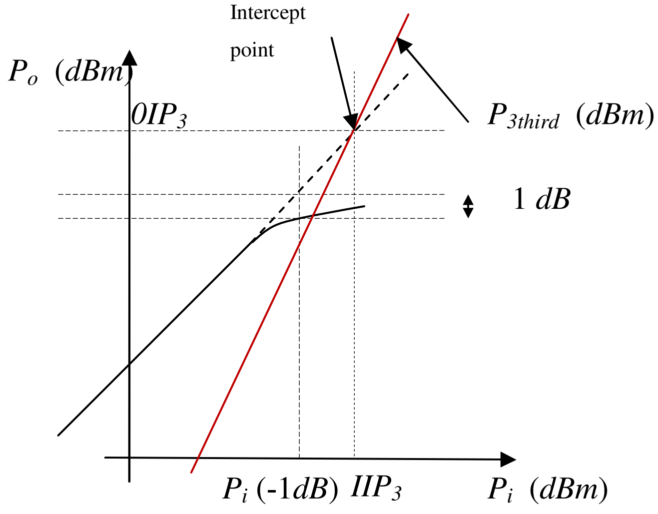


Figure 7 Third order intercept point

Since the (ideal) gain curve has a slope of one and the third order one has a slope of three, they must intersect at some point. This point is called the "third order intercept point" and it is defined by the input intercept point IIP_3 (input power in dBm at the intersection) and the output intercept point OIP_3 (output power in dBm at the intersection). It is evident that these values cannot be measured directly. However, using sensitive spectrum analyzers, we can plot the linear part of the curves shown in Figure 7 and extrapolate.

A low noise RF amplifier is limited by two factors: The minimum detectable signal, which defines its sensitivity and the amount of interference that it can tolerate. The minimum detectable signal should be as small as possible while the third order intercept point should be as high as possible. A high value of the intercept point implies also that 1 dB compression point is also high. In fact, we can show that $IIP3 \approx P_i(-1dB) + 9.6dB$. In many design, these two objectives (low noise and low distortion) cannot be achieved simultaneously.

5.3 Mixers

In mixer design, the objective is to multiply two carriers. We can achieve this goal either by using nonlinear elements (a quadratic Taylor series development will produce the desired result; see Table 1: terms corresponding to k_2) or by switching functions (we turn on and off a switch). To develop the basic definitions, we are going to start the analysis of passive diode mixers.

Diode mixers

The mixers that we are going to study in this part use ideal diodes. The local oscillator voltage is assumed to have very large amplitude and the signal voltage has very small amplitude. In what follows, the intermediate frequency is the difference frequency: $f_{if} = f_s - f_{lo}$.

Unbalanced diode mixer

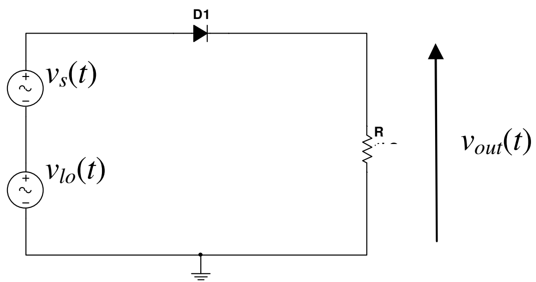


Figure 8 Single diode mixer

In the above figure, the output voltage $v_{out}(t)$ is:

$$v_{out}(t) = \begin{cases} v_s(t) + v_{lo}(t) & v_s(t) + v_{lo}(t) \geq 0 \\ 0 & v_s(t) + v_{lo}(t) < 0 \end{cases} \tag{27}$$

Where $v_s(t) = V_s \cos \omega_s t$ and $v_{lo}(t) = V_{lo} \cos \omega_{lo} t$ along with $V_{lo} >> V_s$.

Since V_{lo} >> V_s, we can modify (27) and write:

$$v_{out}(t) = \begin{cases} v_s(t) + v_{lo}(t) & \cos \omega_{lo} t \geq 0 \\ 0 & \cos \omega_{lo} t < 0 \end{cases} \tag{28}$$

Let us introduce the following function:

$S_W(t) = u(\cos \omega_{lo} t)$. This is a square wave that is equal to one when $\cos \omega_{lo} t$ is positive and to zero when $\cos \omega_{lo} t$ is negative. This function is periodic, its fundamental frequency is $f_{lo} = \dfrac{\omega_{lo}}{2\pi}$ and its Fourier series development is:

$$S_W(t) = \frac{1}{2} + \sum_{k=1}^{+\infty} \frac{2(-1)^{(k-1)}}{\pi(2k-1)} \cos(2k-1)\omega_{lo} t \tag{29}$$

So, the voltage at the output can be written as:

$$\begin{aligned} v_{out}(t) &= (V_s \cos \omega_s t + V_{lo} \cos \omega_{lo} t) S_W(t) \\ &= (V_s \cos \omega_s t + V_{lo} \cos \omega_{lo} t)\left[\frac{1}{2} + \sum_{k=1}^{+\infty} \frac{2(-1)^{(k-1)}}{\pi(2k-1)} \cos(2k-1)\omega_{lo} t\right] \end{aligned}$$

We are just going to look at the first terms of the Fourier series:

$v_{out}(t) = (V_s \cos \omega_s t + V_{lo} \cos_{lo} t)\left(\dfrac{1}{2} + \dfrac{2}{\pi}\cos \omega_{lo} t - \dfrac{2}{3\pi}\cos 3\omega_{lo} t + \cdots\right)$.This is going to result in:

$$v_{out}(t) = \frac{V_s}{2}\cos \omega_s t + \frac{V_{lo}}{2}\cos \omega_{lo} t + \frac{V_s}{\pi}\cos(\omega_s + \omega_{lo})t + \frac{V_s}{\pi}\cos(\omega_s - \omega_{lo})t + \cdots \tag{30}$$

We find that the output signal contains the RF signal $v_s(t)$, the local oscillator voltage $v_{lo}(t)$, a component at the sum of frequencies and the desired component at the difference plus a large number of other components. The following figure shows the output voltage measured using a spectrum analyzer with the following data: $V_s = 0.1\ V$, $f_s = 1$ MHz, $V_{lo} = 10\ V$ and $f_{lo} = 1.5$ MHz.

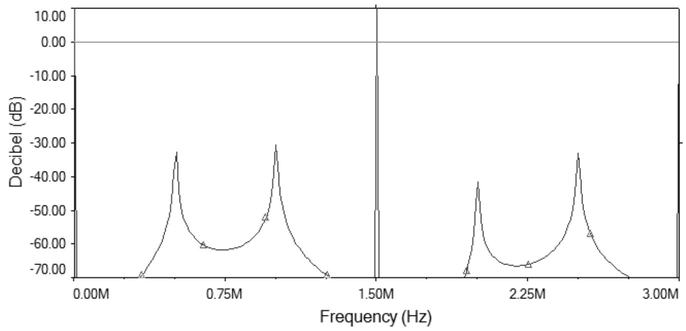


Figure 9 Single diode output spectrum

We see in the above figure that there is a component at the intermediate frequency (500 kHz), a very large component at the local oscillator frequency (1.5 MHz) and a component at the RF frequency (1 MHz). We can also observe the sum frequency (2.5 MHz), a component at 2 MHz (due to a square law nonlinearity in the diode model), and a large one at 3 MHz (coming from the product of vlo (t) with the fundamental of SW (t). The other signals at frequencies above 3 MHz are not shown. The signal is very hard to filter.

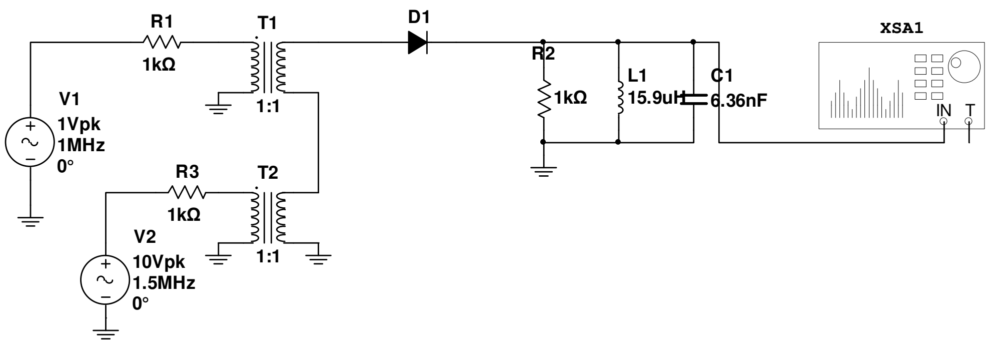


Figure 10 Single diode mixer with filtering

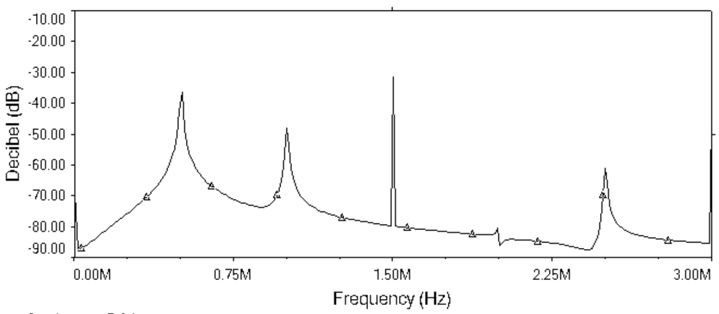


Figure 11 output spectrum from single diode mixer with filtering

Figure 10 and Figure 11 show the result of an implementation with a filter implemented using a tank circuit tuned at the intermediate frequency. We see the presence of spurious signals. Furthermore, the large local oscillator component can lead to distortion by desensitization as seen in the previous section.

Balanced diode mixer

We have seen in the previous part that the most troublesome signal is the large local oscillator signal. The balanced mixer uses a combination of two diode mixers connected in such way to eliminate the local oscillator voltage at the output.

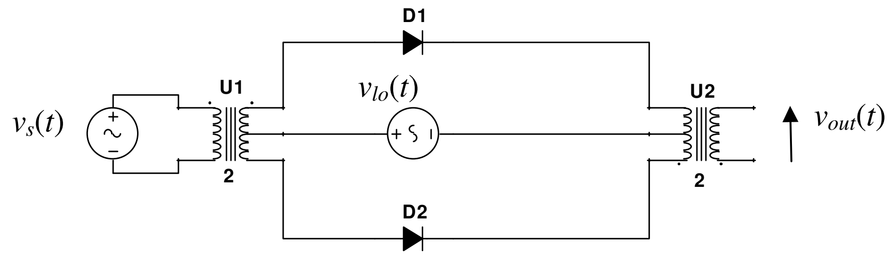


Figure 12 Balanced diode mixer

In the above figure, we see clearly that we have two mixers. The input of the top mixer (using the diode D1) is $v_s(t) + v_{lo}(t)$ while the input of the bottom one (using D2) is $v_{lo}(t) - v_s(t)$. The currents at the outputs from the mixers are opposite in the second transformer. So, the output voltage will be proportional to the difference of the outputs of the mixers.

$$v_{out}(t) = \left(v_s(t) - v_{lo}(t)\right) S_W(t) - \left(v_{lo}(t) - v_s(t)\right) S_W(t) = 2v_s(t) S_W(t) \tag{31}$$

The first components are:

$$v_{out}(t) = V_s \cos \omega_s t + \frac{2V_s}{\pi} \cos\left(\omega_s + \omega_{lo}\right) t + \frac{2V_s}{\pi} \cos\left(\omega_s - \omega_{lo}\right) t + \cdots \tag{32}$$

We can see that the local oscillator voltage has disappeared. However, any slight variation in the symmetry of the circuit will be the cause of a leakage of the local oscillator voltage to the output.

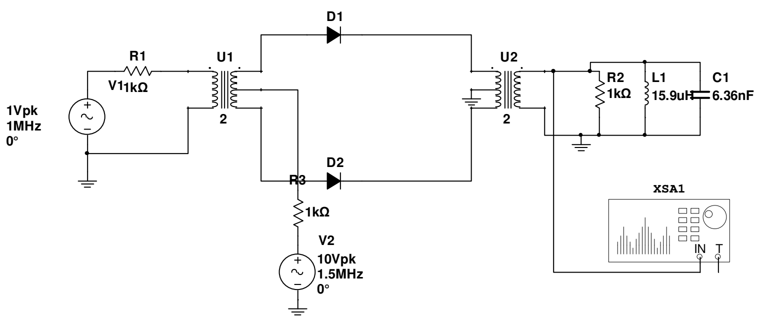


Figure 13 Balanced diode mixer with filtering

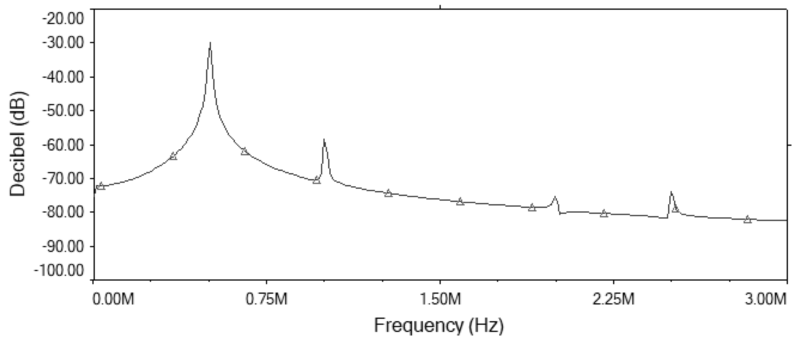


Figure 14 Output spectrum from balanced mixer with filtering

In Figure 13 and Figure 14, we see an implementation of a balanced mixer with a filter implemented using a tank circuit. The output spectrum does not contain any component at the local oscillator frequency or its multiples. In the following example, we are going to see the effect of a difference in the two diodes.

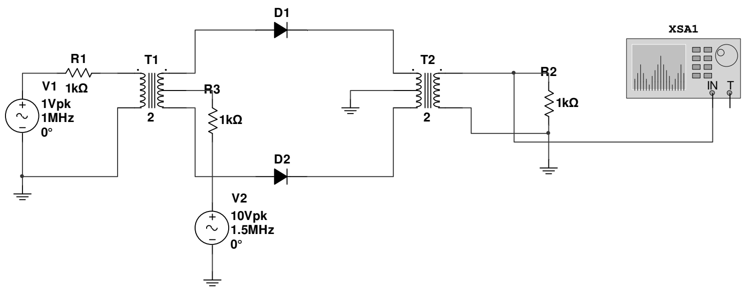


Figure 15 Balanced diode mixer without filtering

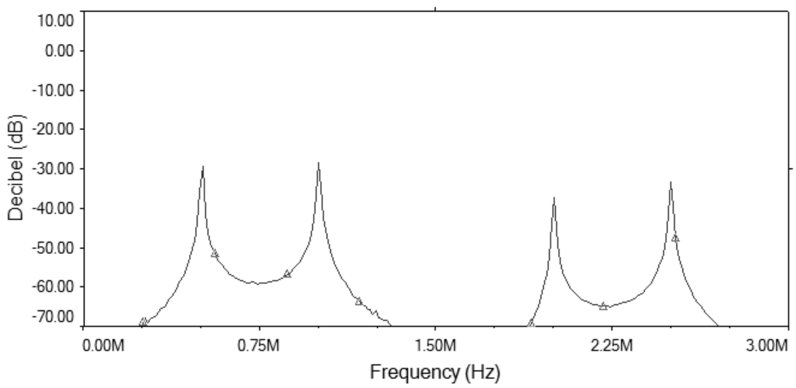


Figure 16 Output spectrum for exact balance

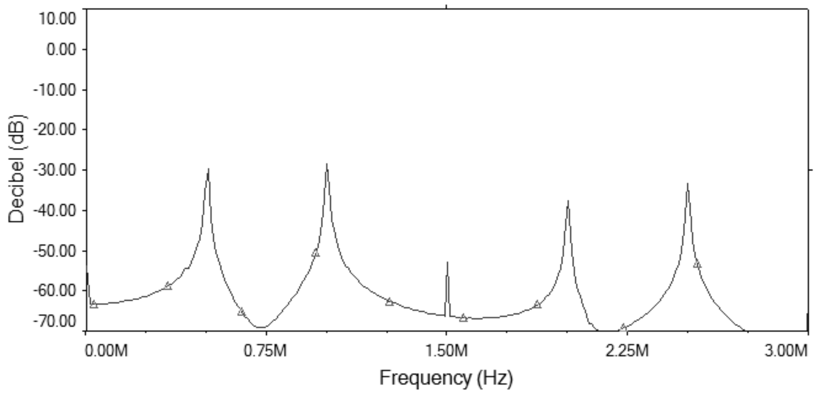


Figure 17 Output spectrum for unbalanced diodes

In Figure 16, the output spectrum does not contain any component at the local oscillator frequency due to the perfect symetry of the circuit. In Figure 17, we modified the saturation current of the diodes. D1 has $I_s = 10^{-14}\,A$ while D2 has $I_s = 2\times10^{-14}\,A$. We remark a small amount of power at the local oscillator frequency (1.5 MHz).

Double balanced mixer

In the balanced mixer, we were able to eliminate the local oscillator component from the output signal. However, the signal frequency still appears at the output. This is due to the DC term in the Fourier series development of the switching function $S_W(t)$. If we use a balanced mixer with the diodes that are on when the local oscillator voltage is positive and another one with diodes that are on when the local oscillator voltage is negative, we can solve the problem.

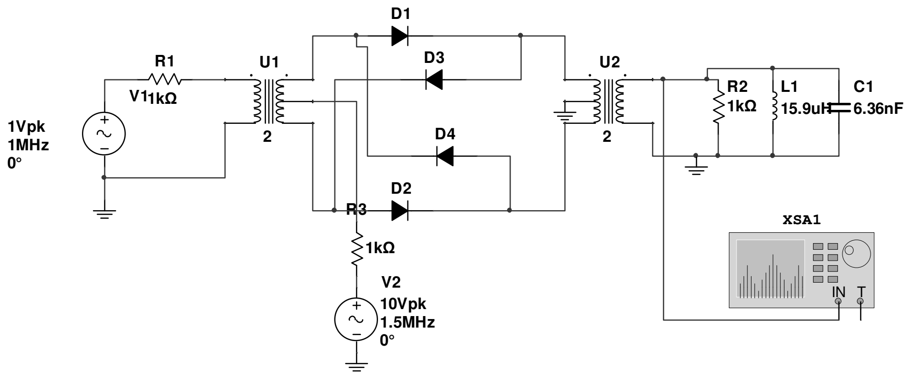


Figure 18 Double balanced mixer with filtering

In Figure 18, we remark that we have two balanced mixer that are implemented. The first one uses diodes D1 and D2. Those two diodes are on when the local oscillator voltage is positive. Its contribution to the output is given by equation(31):

$$v_1(t) = 2v_s(t)S_W(t)$$

The other balanced mixer uses diodes D3 and D4. these diodes are connected in reverse. So, its contribution to the output is:

$$v_2(t) = 2v_s(t)\overline{S}_W(t)$$

The switching function $\overline{S}_W(t)$ is a function that is equal to one when $S_W(t)$ is zero and to zero when $S_W(t)$ is one. So, $\overline{S}_W(t) = 1 - S_W(t)$. The diodes are also crossed with D1 and D2. So, the total output is proportional to the difference between v_1 and v_2.

$$v_{out}(t) = 2v_s(t)S_W(t) - 2v_s(t)\left(1 - S_W(t)\right) = 2v_s(t)\left(2S_W(t) - 1\right) \tag{33}$$

The function $(2S_W(t) - 1)$ is a square wave with no DC term and it is equal to 1 for half period and -1 for the other half. So, its Fourier series is:

$$S_W(t) = \sum_{k=1}^{+\infty} \frac{4(-1)^{(k-1)}}{\pi(2k-1)} \cos(2k-1)\,\omega_{lo} t \tag{34}$$

The first components are:

$$v_{out}(t) = \frac{4V_s}{\pi}\cos(\omega_s + \omega_{lo})t + \frac{4V_s}{\pi}\cos(\omega_s - \omega_{lo})t + \ldots \tag{35}$$

The output spectrum of the circuit shown in Figure 18 is shown below:

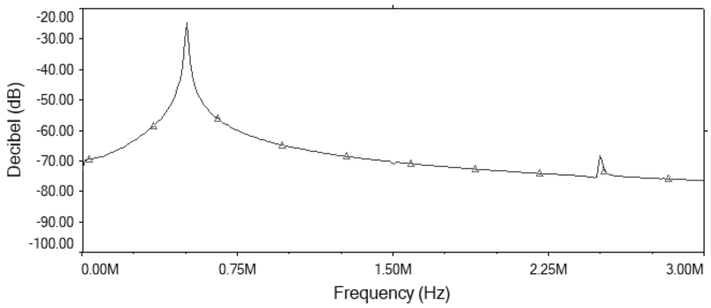


Figure 19 Output spectrum of double balanced mixer with filtering

We see that the signal at the output is only the desired IF frequency signal with a very small signal at the sum frequency (2.5 MHz).

This terminology: unbalanced, balanced and double balanced is used for other type of mixers. We say that a mixer is unbalanced if both the signal (RF) and the local oscillator frequencies appear at the output. If only the signal frequency appears at the output, we say that the mixer is balanced and if both signals are eliminated, the mixer is called double balanced.

Active Mixers

The diode mixers are passive mixers. The RF power entering the device is larger than the power of the signal at the IF port. They are thus characterized by a "Conversion loss". The mixers that we are going to consider now have a conversion gain.

Field Effect Transistor Mixers

Field effect transistors have a half square law characteristics. So, with proper biasing, they can be used with advantage in implementing mixers. Since both JFET and MOSFET transistors have the same type of transfer, we are going to analyze the JFET case. MOSFET transistors have similar results (with a difference in biasing). We have seen that a JFET is characterized by a pinch-off voltage V_p and a drain saturation current I_{DSS}. Its transfer function is:

$$i_D = \begin{cases} I_{DSS}\left(1-\frac{v_{GS}}{V_p}\right)^2 & V_p \le v_{GS} \le 0 \\ 0 & v_{GS} < V_p \\ \text{forbidden} & v_{GS} > 0 \end{cases} \tag{36}$$

Let the gate to source voltage be: $v_{GS} = v_s(t) + v_{lo}(t) + V_{DC}$. We also use $V_x = V_p - V_{DC}$ and we assume that v_{GS} remains always inside the square law region. The drain current becomes:

$$i_D(t) = \frac{I_{DSS}}{V_p^2}\left[V_x^2 + v_s^2(t) + v_{lo}^2(t) - 2V_x v_s(t) - 2V_x v_{lo}(t) + 2v_s(t)v_{lo}(t)\right] \tag{37}$$

In (37), there is only one term of interest: it is the one with the product of v_s and v_{lo}. When we replace $v_s(t) = V_s\cos\omega_s t$ and $v_{lo}(t) = V_{lo}\cos\omega_{lo}t$, the last term provides the following component at the IF frequency:

$$i_{\omega_s-\omega_{lo}}(t) = \frac{I_{DSS}}{V_p^2}V_{lo}V_s\cos\left(\omega_s - \omega_{lo}\right)t \tag{38}$$

In this case, we can define a "conversion transconductance". It is the ratio of the output peak current as the IF frequency and the input peak voltage at the RF frequency. In this case, the conversion transconductance is:

$$G_c = \frac{I_{DSS}}{V_p^2}V_{lo} \tag{39}$$

We see that the conversion transconductance is proportional to the amplitude of the local oscillator. In general, we have $V_{lo} >> V_s$. So, the largest possible value for the amplitude of the local oscillator while remaining inside the square law region is $V_{lo} = \frac{|V_p|}{2}$ along with a DC biasing voltage in the middle of the square law region: $V_{DC} = \frac{V_p}{2}$. So, the largest conversion transconductance while remaining in the square law region is:

$$G_{c,\max} = -\frac{I_{DSS}}{2V_p} \tag{40}$$

The next figure (Figure 20) shows an implementation with a JFET having Vp = -2 V. We can observe the output spectrum in Figure 21. We remark components at the IF (sum and difference) frequencies: 500 kHz and 2.5 MHz, the RF frequency: 1 MHz, the local oscillator frequency: 1.5 MHz and at their double: 2 MHz and 3 MHz.

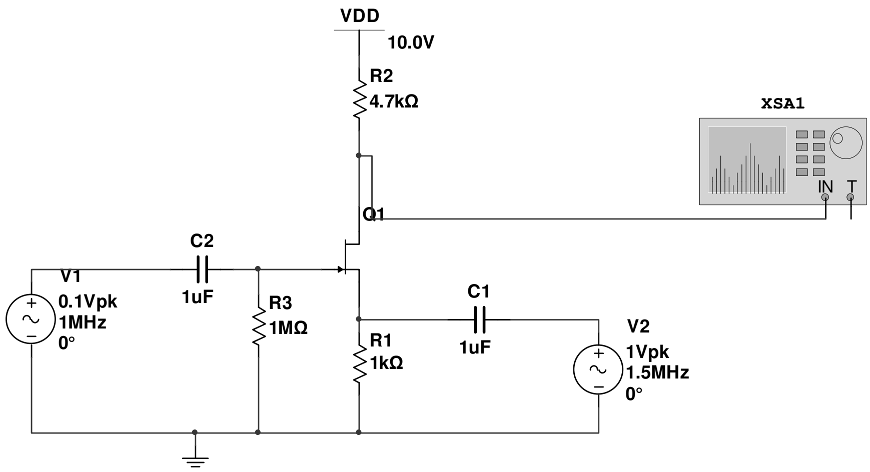


Figure 20 JFET mixer without filtering

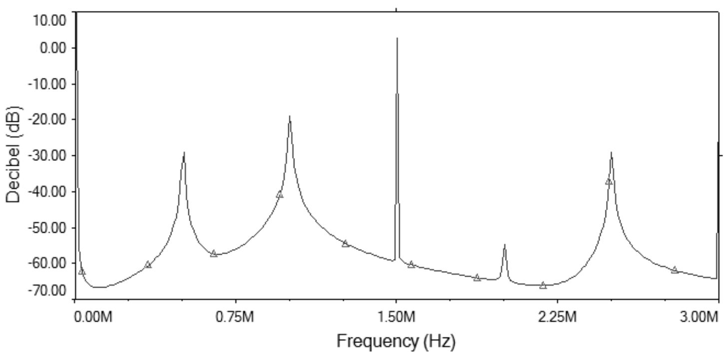


Figure 21 Drain voltage spectrum for unfiltered mixer

Figure 22, on the other hand, shows a possible practical implementation with filtering at all three ports. We remark that the unwanted signals have a much smaller output.

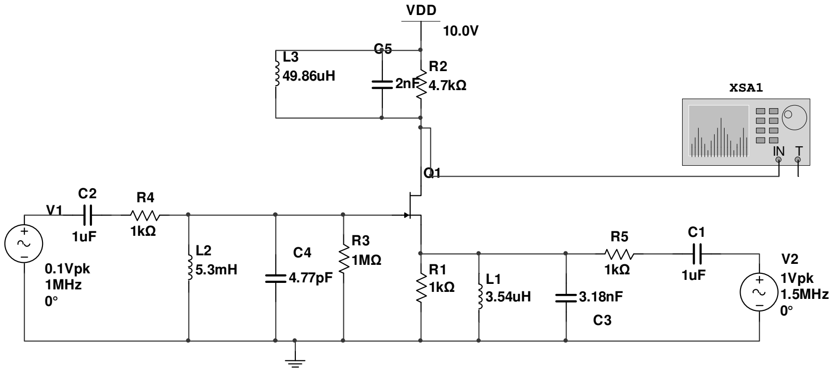


Figure 22 JFET mixer with filtering

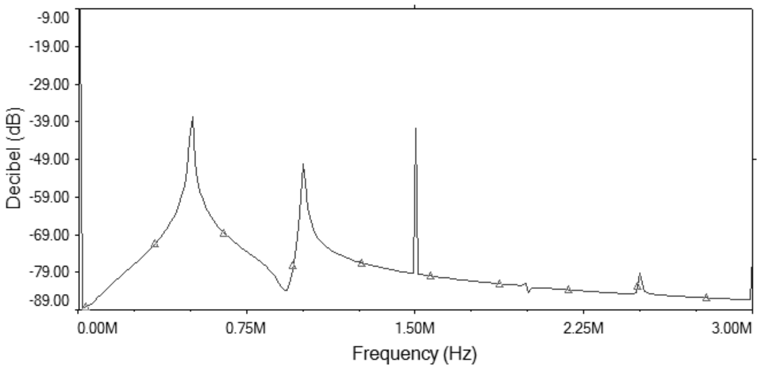


Figure 23 Drain voltage spectrum for filtered JFET mixer

Bipolar Junction Transistor Mixers

BJT have an exponential nonlinearity which can be used to implement mixing. For BJT mixers, we add the RF and local oscillator signal with a DC voltage to apply it between the base and the emitter of the transistor.

$$v_{BE}(t) = V_{DC} + v_s(t) + v_{lo}(t)$$

The emitter current is given by:

$$i_E(t) = I_{ES}\exp\left(\frac{qv_{BE}}{kT}\right) = I_{ES}\exp\left(\frac{qV_{DC}}{kT}\right)\exp\left(\frac{qv_s(t)}{kT}\right)\exp\left(\frac{qv_{lo}(t)}{kT}\right)$$

Replacing $v_s(t) = V_s\cos\omega_s t$ and $v_{lo}(t) = V_{lo}\cos\omega_{lo} t$ gives:

$$i_E(t) = I_{ES}\exp\left(\frac{qV_{DC}}{kT}\right)\exp\left(\frac{qV_{lo}}{kT}\cos\omega_{lo}t\right)\exp\left(\frac{qV_s}{kT}\cos\omega_s t\right)$$

$$= I_{ES}\exp\left(\frac{qV_{DC}}{kT}\right)\exp\left(x\cos\omega_{lo}t\right)\exp\left(y\cos\omega_s t\right)$$

Where $x = \dfrac{qV_{lo}}{kT}$ and $y = \dfrac{qV_s}{kT}$.

So, we have a product of the exponentials of the two voltages and not the product of the two voltages. By using the Fourier series development of the exponentials, we arrive to:

$$i_E(t) = I_{ES}\exp\left(\frac{qV_{DC}}{kT}\right)I_0(x)I_0(y)\left(1+\sum_{n=1}^{\infty}\frac{2I_n(x)}{I_0(x)}\cos n\omega_{lo}t\right)\left(1+\sum_{m=1}^{\infty}\frac{2I_m(y)}{I_0(y)}\cos m\omega_s t\right)$$

So, the emitter current is composed of:

DC current: $I_{DC} = I_{ES}\exp\left(\frac{qV_{DC}}{kT}\right)I_0(x)I_0(y)$

Components at harmonics of the local oscillator frequency: $I_{DC}\sum_{n=1}^{\infty}\frac{2I_n(x)}{I_0(x)}\cos\left(n\omega_{lo}t\right)$

Components at harmonics of the RF frequency: $I_{DC}\sum_{m=1}^{\infty}\frac{2I_m(y)}{I_0(y)}\cos\left(m\omega_{lo}t\right)$

Mixing between all these components: $4I_{DC}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{I_n(x)I_m(y)}{I_0(x)I_0(y)}\cos n\omega_{lo}t\cos m\omega_s t$

There is only one term of interest in all of the above ones. It is:

$$i_{IF}(t) = \frac{2I_{DC}I_1(x)I_1(y)}{I_0(x)I_0(y)}\cos\left(\omega_s-\omega_{lo}\right)t$$

If the RF input signal is small, then $y \ll 1$ and we have:

$\frac{2I_1(y)}{I_0(y)} \approx y = \frac{qV_s}{kT}$ and $i_{IF}(t) = \frac{I_1(x)}{I_0(x)}\frac{qI_{DC}}{kT}V_s\cos\left(\omega_s-\omega_{lo}\right)t$

Finally, the conversion transconductance is:

$$G_c = \frac{I_1(x)}{I_0(x)}g_{in} \tag{41}$$

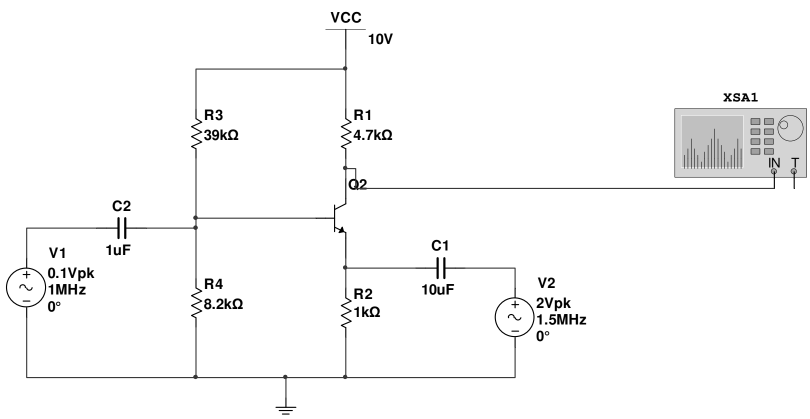


Figure 24 BJT mixer without filtering

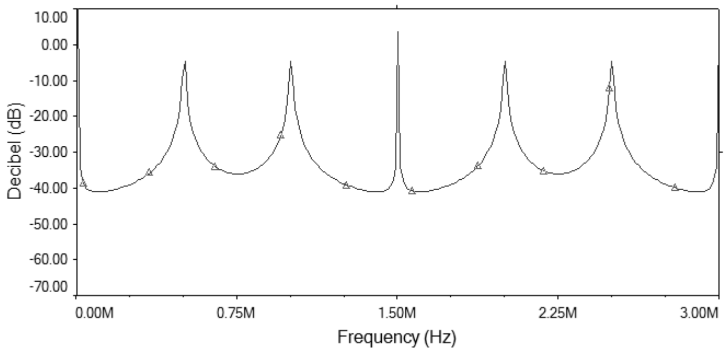


Figure 25 Output spectrum BJT mixer without filtering

Figure 24 shows a possible implementation of a BJT mixer. In Figure 25, we observe the larger output at the IF frequency (500 kHz). However, we remark the large amount of undesired signals. The following two figures concern BJT mixer with output filtering. We remark that the unwanted signals are attenuated, but they are still present. The output spectrum of the JFET mixer is much cleaner.

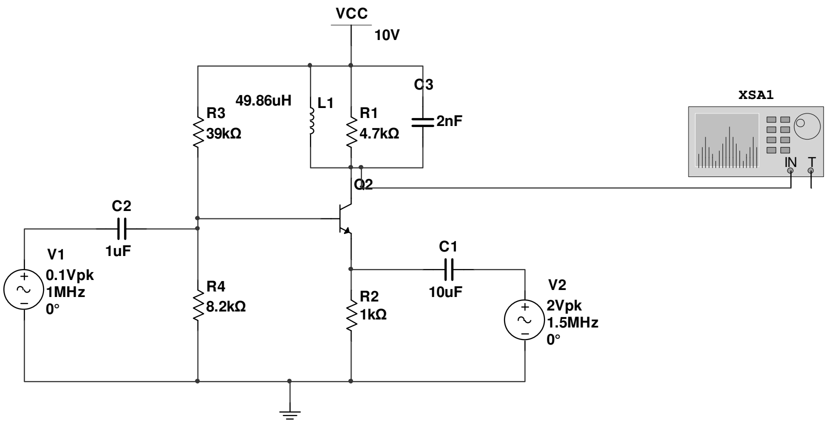


Figure 26 BJT mixer with filtering

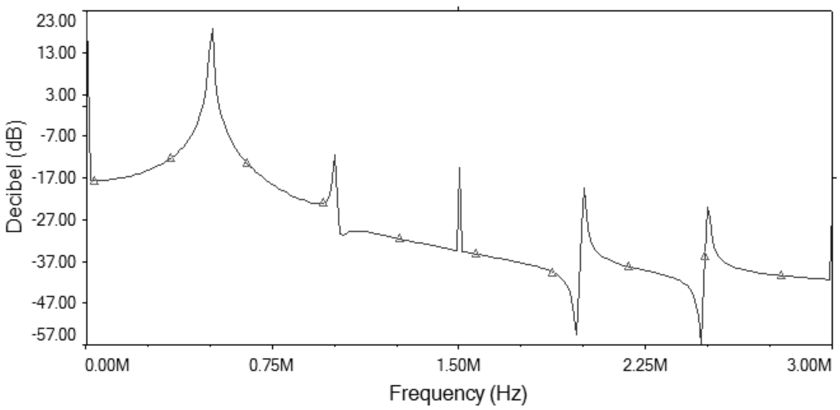


Figure 27 Output spectrum BJT mixer with filtering

Differential pair mixer

Consider the following circuit:

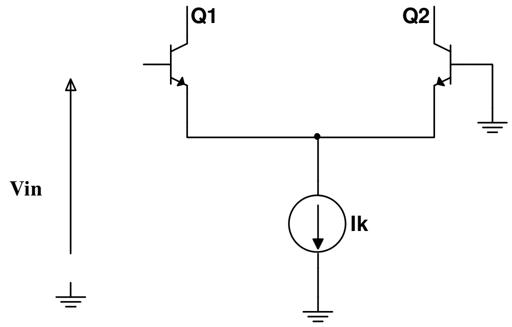


Figure 28 Differential pair

If we assume that α = 1 for both transistors, then we know that the current at the collector of Q1 is: $i_{C1} = \frac{I_k}{2}\left(1 + \tanh\left(\frac{z}{2}\right)\right)$ and the current at the collector of Q2 is $i_{C2} = \frac{I_k}{2}\left(1 - \tanh\left(\frac{z}{2}\right)\right)$, where $z = \frac{qVin}{kT}$. Let $Vin = v_{lo}(t) = V_{lo}\cos\omega_{lo}t$ be the local oscillator voltage and $I_k = I_{DC} + bv_s(t)$, $v_s(t)$ being the rf voltage.

The collector current of Q1 becomes:

$$i_{C1}(t) = \frac{\left(I_{DC} + bv_s(t)\right)}{2}\left(1 + \tanh\frac{x}{2}\cos\omega_{lo}t\right) = \frac{\left(I_{DC} + bv_s(t)\right)}{2} + \left(I_{DC} + bv_s(t)\right)\sum_{n=1}^{\infty} a_{2n-1}(x)\cos\left(2n-1\right)\omega_{lo}t$$

where $x = \frac{qV_{lo}}{kT}$.

We see clearly that the above circuit can be used as a mixer. The conversion transconductance is

$$G_c = \frac{ba_1(x)}{2} \tag{42}$$

The following circuit can be used as a controlled current source for the above differential pair.

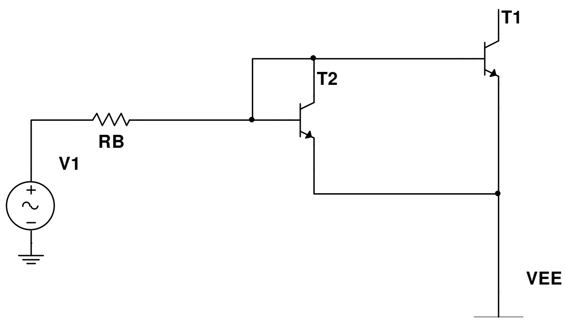


Figure 29 variable current mirror

Here again, if we assume that $\alpha = 1$, the current provided by the source is: $I_k = \frac{V_{EE} - v_1(t) - 0.75}{R_B}$, so the DC current is $I_{DC} = \frac{V_{EE} - 0.75}{RB}$ and $b = \frac{-1}{R_B}$.

Another way of implementing a mixer is to apply a small RF signal between the bases of the two transistors and make the current source depend on the local oscillator voltage. So, looking at Figure 29, the volt age source becomes $v_1(t) = V_{lo}\cos\omega_{lo}t$ and the current driving the differential pair becomes :

$$I_k = \frac{V_{EE} - V_{lo}\cos\omega_{lo}t - 0.75}{R_B} = I_{k0} - \frac{V_{lo}}{R_B}\cos\omega_{lo}t \qquad (43)$$

Where $I_{k0} = \frac{V_{EE} - 0.75}{R_B}$.

If we use a differential load, the output voltage will be proportional to the difference between the collector currents. So:

$$i_{out}(t) = i_{C1}(t) - i_{C2}(t) = I_k \tanh\left(\frac{z}{2}\right)$$

Now, the RF voltage is applied between the bases, so:

$$z = \frac{qv_s(t)}{kT} = \frac{qV_s}{kT}\cos\omega_s t = x\cos\omega_s t$$

If x is small, we can write $\tanh\left(\frac{x}{2}\cos\omega_s t\right) \approx \frac{x}{2}\cos\omega_s t$. After replacement in $i_{out}(t)$, we obtain:

$$i_{out} = \left(I_{k0} - \frac{V_{lo}}{R_B}\cos\omega_{lo}t\right)\frac{qV_s}{2kT}\cos\omega_s t = g_{in}V_s\cos\omega_s t - \frac{g_{in}V_{lo}}{R_B I_{k0}}V_s\cos_{lo}t\cos\omega_s t$$

In this case, the conversion transconductance become:

$$G_c = \frac{g_m V_{lo}}{R_B V_s} \qquad (44)$$

Where $g_{in} = \frac{qI_{k0}}{2kT}$ and $g_m = \frac{g_{in}}{2}$.

The Gilbert Cell

The Gilbert cell consists on the interconnection of two differential pair mixers in such way to build a double balanced mixer. We have seen above that a differential pair implements a balanced mixer. Consider the following circuit:

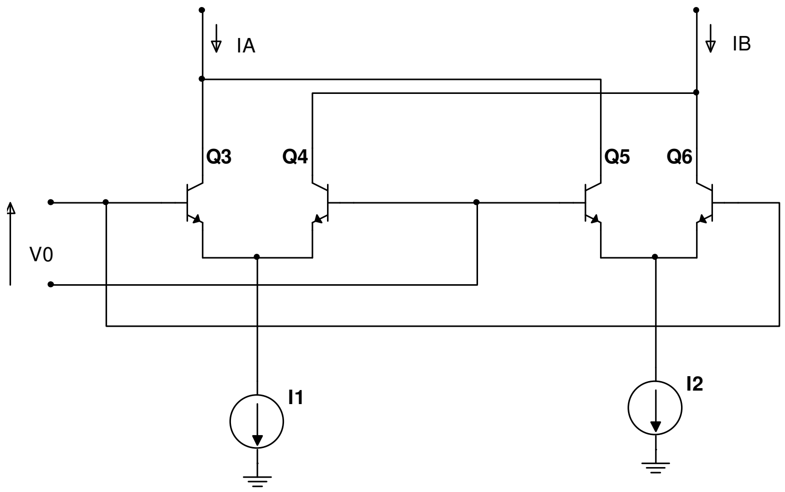


Figure 30 Interconnection of two differential pairs

The different collector currents are:

$$i_{C3} = \frac{I_1}{2}\left(1+\tanh\left(\frac{z}{2}\right)\right) \; i_{C4} = \frac{I_1}{2}\left(1-\tanh\left(\frac{z}{2}\right)\right) \; i_{C5} = \frac{I_2}{2}\left(1-\tanh\left(\frac{z}{2}\right)\right) \; i_{C6} = \frac{I_2}{2}\left(1+\tanh\left(\frac{z}{2}\right)\right)$$

Where $z = \frac{qv_0}{kT}$

So the currents IA and IB are:

$$\text{IA} = i_{C3} + i_{C5} = \frac{I_1 + I_2}{2} + \frac{I_1 - I_2}{2}\tanh\left(\frac{z}{2}\right) \tag{45}$$

$$\text{IB} = i_{C4} + i_{C6} = \frac{I_1 + I_2}{2} - \frac{I_1 - I_2}{2}\tanh\left(\frac{z}{2}\right) \tag{46}$$

The currents I1 and I2 are collector currents from a differential pair. However, we are going to use resistance in series with the emitters in order to have linear amplification. So, let us consider the following circuit:

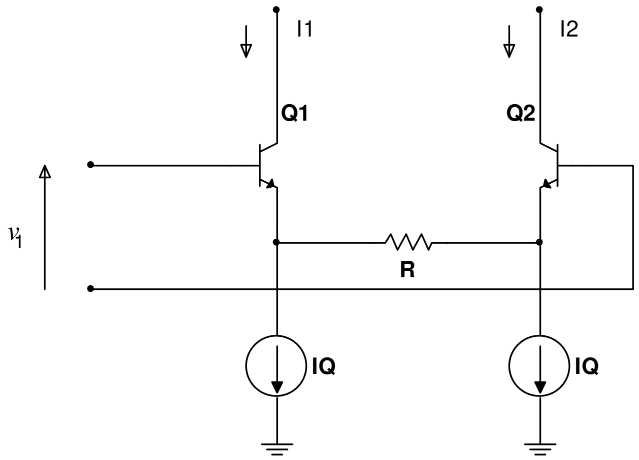


Figure 31 Differential pair with emitter degeneracy

The next circuit is equivalent to the circuit shown in Figure 31, except for a DC voltage drop across the emitter resistances.

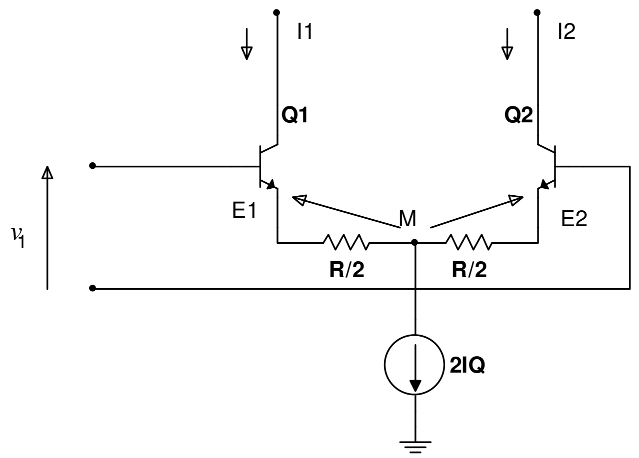


Figure 32 Equivalent differential pair with degeneracy

From Figure 32, and using the result of chapter 3 ($g_{in}R$ >> 1), we can write:

$$v_1 = V_{E1M} + V_{ME2} = \left(r_{in1} + \frac{R}{2} \right) i_{E1} - \left(r_{in2} + \frac{R}{2} \right) i_{E2}$$

We have also $i_{E1} + i_{E2} = 2I_Q$. We can write:

$i_{E1} = I_Q + i$ and $i_{E2} = I_Q - i$, i being the variation of the emitter current from the DC average value I_Q flowing in each transistor. Furthermore, $r_{in1} = r_{in2} = r_{in} = \frac{kT}{qI_Q}$. So, we can write: $i = \frac{v_1}{R + 2r_{in}}$. If we assume also that α = 1, then the two collector currents are equal to the emitter currents: $I_1 = i_{E1} = I_Q + i$ and $I_2 = i_{E2} = I_Q - i$. Finally, equations(45) and (46) can be rephrased as:

$$\text{IA} = I_Q + i\tanh\left(\frac{z}{2}\right) = I_Q + \frac{v_1}{R+2r_{in}}\tanh\left(\frac{z}{2}\right) \tag{47}$$

$$\text{IB} = I_Q - i\tanh\left(\frac{z}{2}\right) = I_Q - \frac{v_1}{R+2r_{in}}\tanh\left(\frac{z}{2}\right) \tag{48}$$

Let $v_1(t) = v_s(t) = V_s \cos\omega_s t$ and $v_0(t) = v_{lo}(t) = V_{lo}\cos\omega_{lo}t$. The output currents become:

$$\text{IA} = I_Q + \frac{V_s \cos_s t}{R+2r_{in}}\tanh\left(\frac{qV_{lo}}{2kT}\cos\omega_{lo}t\right) = I_Q + \frac{V_s \cos_s t}{R+2r_{in}}\tanh\left(\frac{x}{2}\cos\omega_{lo}t\right)$$

$$\text{IB} = I_Q - \frac{V_s \cos_s t}{R+2r_{in}}\tanh\left(\frac{qV_{lo}}{2kT}\cos\omega_{lo}t\right) = I_Q - \frac{V_s \cos_s t}{R+2r_{in}}\tanh\left(\frac{x}{2}\cos\omega_{lo}t\right)$$

And using the Fourier development of $\tanh\left(\frac{x}{2}\cos_{lo} t\right) = 2\sum_{n=1}^{\infty} a_{2n-1}(x)\cos\left(2n-1\right)\omega_{lo}t$, we obtain:

$$\text{IA} = I_Q + 2\frac{V_s \cos_s t}{R+2r_{in}}\sum_{n=1}^{\infty} a_{2n-1}(x)\cos\left(2n-1\right)\omega_{lo}t$$

$$\text{IB} = I_Q - 2\frac{V_s \cos_s t}{R+2r_{in}}\sum_{n=1}^{\infty} a_{2n-1}(x)\cos\left(2n-1\right)\omega_{lo}t$$

Finally, the conversion transconductance of the cell is:

$$G_c = \frac{a_1(x)}{R+2r_{in}} \tag{49}$$

We see that the output currents do not contain any component at the RF frequency nor at the local oscillator frequency or their harmonics. So, it is a double balanced mixer. The Gilbert cell can also be implemented using MOS transistors. It is quite common in all communication circuits (PLL, Mixers, Modulators, etc.).

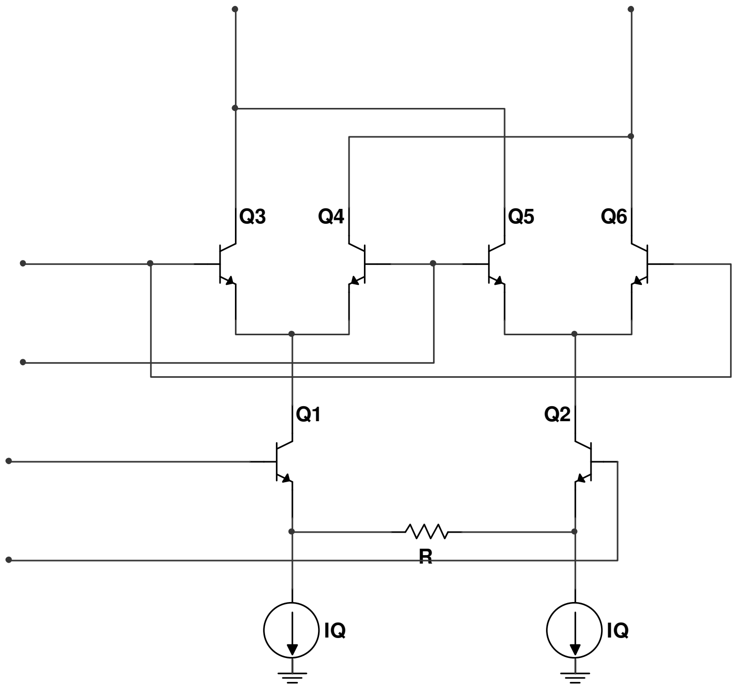


Figure 33 Complete BJT Gilbert Cell

5.4 Small signal amplifiers

The analysis that follows applies to both RF and IF amplifiers. In the design of these amplifiers, we take into account that the signal level is very small. So, the amplifiers are going to use small signals varying around the Q point. In this type of design, we can use measured two port parameters. In the HF and VHF range, the set of 2-port parameters that is most commonly used is the admittance parameters. We have already defined them. We repeat this definition below:

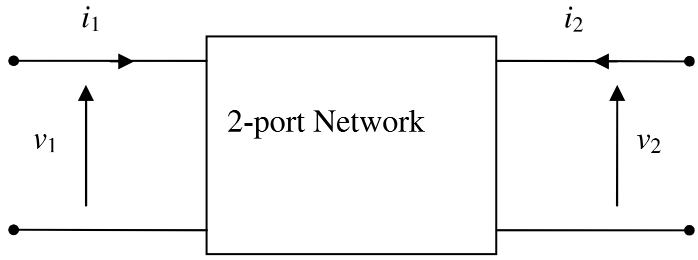


Figure 34 2-port network

In the representation by the admittance parameters, the currents i_1 and i_2 are expressed as a function of the voltages v_1 and v_2.

$$\begin{aligned} i_1 &= y_{11}v_1 + y_{12}v_2 \\ i_2 &= y_{21}v_1 + y_{22}v_2 \end{aligned} \tag{50}$$

In matrix form, we have:

$$\begin{pmatrix} i_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{51}$$

The admittance parameters are also called short circuit admittance parameters. If we look back at equations(50), we can measure the different parameters by applying the definition:

$$\begin{aligned} y_{11} &= \left.\frac{i_1}{v_1}\right|_{v_2=0} & y_{12} &= \left.\frac{i_1}{v_2}\right|_{v_1=0} \\ y_{21} &= \left.\frac{i_2}{v_1}\right|_{v_2=0} & y_{11} &= \left.\frac{i_2}{v_2}\right|_{v_1=0} \end{aligned} \tag{52}$$

The fact that we must use a short circuit (in AC) to measure the parameter eliminates the effect of parasitic elements. This is why we find that these parameters are used at high frequency. However, at the microwave level, we must consider the currents as waves and it is better to use another set: the scattering parameters. The range of frequencies where the use of the admittance parameters is advisable is the HF and VHF bands (from 3 to 300 MHz). The admittance parameters are small signal parameters. They are measured at a given Q point. They are also frequency dependent. So, they are provided in data sheets either at some fixed frequency or as a set of curves giving y_{ij} as a function of frequency.

Hybrid pi model

There exist also frequency independent models for both the BJT and the FET transistor that can be used up to quite high frequency.

The Giacoletto model:

This model is a small signal model of the BJT using some physically dependent resistances and capacitances.

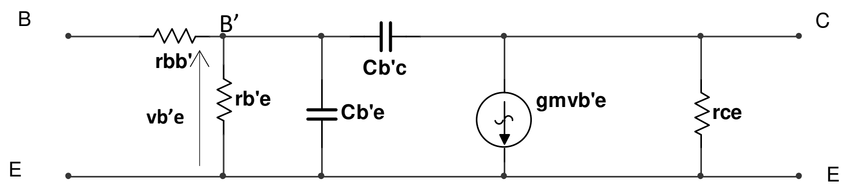


Figure 35 Hybrid Pi model of BJT

The elements of the above circuit can be derived from the transistor data sheet. In order to compute them, we must introduce BJT cut off frequencies.

The β cut off frequency f_β: Due to the different capacitances, the current gain drops when the frequency increases. The common emitter current gain is frequency dependent and we can write:

$$|\beta| = \frac{\beta_0}{\sqrt{1+\left(\frac{f}{f_\beta}\right)^2}} \tag{53}$$

β_0 is the low frequency common emitter current gain.

When we connect the BJT in common base, we introduce the α cut off frequency f_α.

$$|\alpha| = \frac{\alpha_0}{\sqrt{1+\left(\frac{f}{f_\alpha}\right)^2}} \tag{54}$$

α_0 is the low frequency common base current gain.

The two cut off frequencies are related:

$$f_\alpha = \beta_0 f_\beta \tag{55}$$

We remark that f_α is equal to the gain bandwidth product. In data sheets, it is denoted f_T: total bandwidth.

Characteristic	Symbol	Min	Max	Unit
Small-Signal Characteristics				
Current-gain—Bandwidth Product (I_C = 10 mA dc, V_{CE} = 20 V, f = 100 MHz)	f_T			MHz
2N3903		250	—	
2N3904		300	—	

Figure 36 Specification of f_T in 2N3903, 2N3904 data sheet.

The resistance $r_{bb'}$ (called also r_x in some references) is a resistance due to the path between the base contact and the point in the middle of the base where the emission of carriers takes place. Its value is a few ohms (2 to 20 Ω). It is usually neglected in high frequency calculations. The transconductance g_m is the small signal transconductance.

$$g_m = \frac{qI_{CQ}}{kT} \tag{56}$$

The resistance $r_{b'e}$ (called also r_π) is the dynamic resistance of the base emitter junction seen from the base.

$$r_{b'e} = \beta_0 r_e = \frac{\beta_0}{g_m} \tag{57}$$

The capacitance $C_{b'c}$ is the capacitance of the reverse biased collector base junction. It decreases with increasing V_{CC}. It is provided in transistor data sheets but it is called C_{ob} (output capacitance of a common base amplifier).

The output resistance r_{ce} (also called r_o) is the finite output resistance of a common emitter connected transistor due to the "Early" effect. It is inversely proportional to the collector biasing current I_{CQ}. An estimate of this resistance is:

$$r_{ce} \approx \frac{V_A}{I_{CQ}} \tag{58}$$

Where V_A is the Early voltage which is around 20 to 150 V.

The capacitance $C_{b'e}$ is the "diffusion" capacitance of the forward biased base emitter junction.

$$C_{b'e} = \frac{g_m}{2\pi f_T} - C_{b'c} \approx \frac{g_m}{2\pi f_T} \tag{59}$$

The FET Hybrid Pi Model

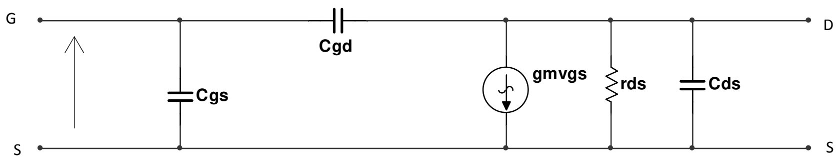


Figure 37 Hybrid Pi model of FET

We remark that the equivalent model is essentially capacitive. The input impedance of any FET is capacitive (true capacitive for the MOSFET and reverse biased junction for the JFET). The different capacitances are provided in data sheets. However, they are not directly given. Instead, manufacturers provide short circuit capacitances that are related.

The input capacitance with drain-source short circuited:

$$C_{gss} = C_{iss} = C_{gs} + C_{gd} \tag{60}$$

The output capacitance with gate-source short circuited:

$$C_{oss} = C_{gd} + C_{ds} \tag{61}$$

The reverse capacitance with gate-source and drain-source short circuited:

$$C_{rss} = C_{gd} \tag{62}$$

The output resistance r_{ds} is proportional to the inverse of the biasing drain current. We can show that:

$$r_{ds} = r_o \approx \frac{1}{\lambda I_{DQ}} \tag{63}$$

The constant λ is called the channel length modulation parameter. Its inverse is the Early voltage (same as the BJT transistor). Typical values for λ are between 0.01 V^{-1} to 0.03 V^{-1}.

Finally, the small signal transconductance is g_m which depends on the drain bias current.

For a JFET, the characteristic is:

$$i_D = I_{DSS}\left(1 - \frac{v_{GS}}{V_p}\right)^2$$

The small signal transconductance is:

$$g_m = \frac{-2}{V_p}\sqrt{I_{DSS} I_{DQ}} \tag{64}$$

It varies between 0 and $g_{m0} = \frac{-2I_{DSS}}{V_p}$.

For a MOSFET, the characteristic is:

$$i_D = \beta\left(v_{GS} - V_{th}\right)^2$$

The small signal transconductance is:

$$g_m = 2\sqrt{\beta I_{DQ}} \tag{65}$$

These models are useful for wide band design of high frequency amplifiers. However, they are not very accurate. It is preferable to use measured parameters such as the admittance parameters.

Design using admittance parameters[1]

Consider the following two port network with a source having an admittance Y_S and a load having an admittance Y_L.

[1] This section is just an introduction to this type of design. The interested reader should consult: Carson, R.,S., *High Frequency Amplifiers*, John Wiley & sons, 1976.

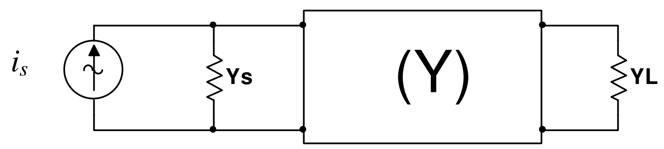


Figure 38 Two port network with source and load

The basic equations(50) are repeated below:

$$i_1 = y_{11}v_1 + y_{12}v_2$$
$$i_2 = y_{21}v_1 + y_{22}v_2$$

Along with:

$$i_1 = i_s - v_1 Y_S \text{ and } i_2 = -v_2 Y_L \tag{66}$$

Designing such amplifier means finding the values of Y_S and Y_L that optimize some criterion: gain, noise along with constraints of stability and bandwidth.

The input and output admittances are given by:

$$Y_{in} = \frac{i_1}{v_1} = y_{11} - \frac{y_{12}y_{21}}{y_{22} + Y_L} \tag{67}$$

$$Y_{out} = \left.\frac{i_2}{v_2}\right|_{i_s=0} = y_{22} - \frac{y_{12}y_{21}}{y_{11} + Y_S} \tag{68}$$

The output admittance is the Thevenin admittance of the device and the source. We remark that, if $y_{12} \neq 0$, the input admittance depends on the load and the output admittance depends on the source. If $y_{12} = 0$, the device is unilateral and this interaction disappears (we have uncoupling). The reverse trans-admittance y_{12} indicates that there is a feedback in the two port device. We have seen in chapter 4 that under some conditions of gain and phase, the feedback can lead to oscillations. These oscillations can occur if the input or the output conductance becomes negative for some combination of load and source impedances.

Let $y_{ik} = g_{ik} + jb_{ik}$ and $Y_S = G_S + jB_S$ along with $Y_L = G_L + jB_L$ and $Y_{in} = G_{in} + jB_{in}$, $Y_{out} = G_{out} + jB_{out}$.

The stability of two port devices has been thoroughly studied and we are going to cite the most important results.

The Linvill stability criterion

The following factor should be computed:

$$C = \frac{|y_{21}y_{12}|}{2g_{11}g_{22} - \mathrm{Re}[y_{21}y_{12}]} \tag{69}$$

If $0 \le C < 1$, the device is unconditionally stable. It means that we can select any values for the source and load impedance to satisfy any optimization criterion. Consider the following parameters of the 2N4957 measured at $f = 200$ MHz, $V_{CEQ} = 10$ V and $I_{CQ} = 2$ mA:

$y_{11} = (2.7 + j6.8)\ \text{m}\Omega^{-1}$; $y_{12} = (0 - j0.5)\ \text{m}\Omega^{-1}$; $y_{21} = (53 - j22)\ \text{m}\Omega^{-1}$; $y_{22} = (0.1 + j1.5)\ \text{m}\Omega^{-1}$

Using the above data, the Linvill factor is $C = 2.486$. So, the transistor is potentially unstable.

The Stern stability criterion

If $C \ge 1$, Stern has developed another stability criterion which depends on the load and the source impedances. The Stern factor is:

$$K = \frac{2(g_{11} + G_S)(g_{22} + G_L)}{|y_{21}y_{12}| + \text{Re}[y_{21}y_{12}]} \tag{70}$$

If K is greater than 1, the circuit is stable. If K is less than 1, the circuit is potentially unstable.

Looking at (69) and(70), we can see that that if $y_{12} = 0$ (unilateral device), $C = 0$ and $K = \infty$. The circuit is unconditionally stable and there is total freedom in the choice of source and load impedance. Furthermore, there is no interaction between the input and output impedance. This implies that every circuit (input and output) can be tuned independently. If $C < 1$, we still have freedom in our choice of source and load impedance, however, there will be a problem in alignability because of the influence of the output on the input and vice versa.

In general, a good design should minimize the noise factor and maximize the power gain of the amplifier. We define below this notion of power gain.

Operating power gain

This gain is the ratio of the power delivered to the output load over the power delivered to the input port.

$$G_p = \frac{\text{power delivered to the load}}{\text{power into the input port}} = \frac{|V_2|^2 G_L}{|V_1|^2 G_{in}}$$

We finally obtain:

$$G_p = \frac{|y_{21}|^2 G_L}{|Y_L + y_{22}|^2 G_{in}} \tag{71}$$

We remark that the operating power gain does not depend on the source admittance.

Available Gain

This gain has already been defined previously and it is the ratio of the available power at the output over the available power at the input. It represents the maximum power gain that we can obtain from the active device.

$$G_A = \frac{\text{available power at the output port}}{\text{available power at the input port}}$$

$$G_A = \frac{|y_{21}|^2 G_S}{\mathrm{Re}\left[\left(y_{11}y_{22} - y_{12}y_{21} + y_{22}Y_S\right)\left(y_{11} + Y_S\right)^*\right]} \tag{72}$$

The available gain does not depend on the output load admittance.

Transducer Gain

The transducer gain is defined as the ratio of the power delivered at the load over the available power delivered at the input.

$$G_T = \frac{\text{power delivered to the load}}{\text{power available from the source}}$$

$$G_T = \frac{4G_S G_L |y_{21}|^2}{\left|\left(y_{11} + Y_S\right)\left(y_{22} + Y_L\right) - y_{12}y_{21}\right|^2} \tag{73}$$

We see that G_T depends on both the source and the load impedance.

Maximum Available Gain

There exists also a figure of merit which the maximum available gain. The maximum available gain is the available gain of the device if it were unilateral ($y_{12} = 0$).

$$MAG = \frac{|y_{21}|^2}{4g_{11}g_{22}} \tag{74}$$

It is evident that this gain cannot be achieved with simple BJT or FET based amplifiers. If we have matching at both ports, then we can have: $G_{T\max} = G_A$.

Design using unilateral device

If $y_{12} = 0$, the device is unilateral and we have uncoupling between the output and input. If we consider equations (67) and(68), the input and output impedances are: $Y_{in} = y_{11}$ and $Y_{out} = y_{22}$. If we want to maximize the gain, we must select:

$$Y_S = Y_{in}^* \text{ and } Y_L = Y_{out}^* \tag{75}$$

In this case, the power gain is $G_{T\max} = G_A = MAG$.

If we want to optimize the noise behavior, we must select the source resistance according to a minimum noise factor. So, in this case $R_S = \frac{1}{G_S}$ is fixed (given in data sheet of the active device). So, we are going to have a smaller gain. In this case we have:

$$B_S = -b_{11},\ G_L = g_{22},\ B_L = -b_{22} \tag{76}$$

The transducer gain becomes:

$$G_T = \frac{G_S |y_{21}|^2}{(g_{11}+G_S)^2 g_{22}} \tag{77}$$

Design for an unconditionally stable device

If $y_{12} \neq 0$, we must compute the Linvill stability factor C. If $0 \le C < 1$, the device is unconditionally stable and we are free to select Y_S and Y_L according to any criterion. The problem that we encounter here is the interdependence between the input impedance and the output impedance. Y_{in} depends on Y_L and Y_{out} depends on Y_S. If we want to optimize the gain, we can compute the partial derivatives of G_T with respect to G_S, B_S, G_L and B_L and set them to zero. We obtain:

$$G_S = \frac{\sqrt{\left[2g_{11}g_{22} - \mathrm{Re}(y_{12}y_{21})\right]^2 - |y_{12}y_{21}|^2}}{2g_{22}} \tag{78}$$

$$B_S = -b_{11} + \frac{\mathrm{Im}(y_{12}y_{21})}{2g_{11}} \tag{79}$$

$$G_L = \frac{\sqrt{\left[2g_{11}g_{22} - \mathrm{Re}(y_{12}y_{21})\right]^2 - |y_{12}y_{21}|^2}}{2g_{11}} = \frac{G_S g_{22}}{g_{11}} \tag{80}$$

$$B_L = -b_{22} + \frac{\mathrm{Im}(y_{12}y_{21})}{2g_{11}} \tag{81}$$

Replacing in the expression of G_T, we obtain:

$$G_T = \frac{|y_{21}|^2}{2g_{11}g_{22} - \mathrm{Re}(y_{12}y_{21}) + \sqrt{\left[2g_{11}g_{22} - \mathrm{Re}(y_{12}y_{21})\right]^2 - |y_{12}y_{21}|^2}} \tag{82}$$

Example[2]: Consider a transistor having $y_{11} = (8 + j6.8)\ \mathrm{m}\Omega^{-1}$, $y_{22} = (0.4 + j1.5)\ \mathrm{m}\Omega^{-1}$, $y_{21} = (53 + j22)\ \mathrm{m}\Omega^{-1}$ and $y_{12} = (0 - j0.1)\ \mathrm{m}\Omega^{-1}$. The Linvill stability factor is: C = 0.667. The

[2] From Krauss, H. L., Bostian, C. W. and Raab, F. H., *Solid State Radio Engineering*, John Wiley & Sons, 1980.

device is unconditionally stable. Using equations(78),(79), (80) and (81), the optimum load and source admittances are: $Y_S = (8.0 - j13.43)\ \text{m}\Omega^{-1}$ and $Y_L = (0.4 - j1.83)\ \text{m}\Omega^{-1}$. The transducer gain is $G_T = 219$ which is smaller than $MAG = 257$.

Design using a potentially unstable device

If the Linvill stability factor is larger than one, one solution is to specify a value for the Stern stability factor and solve for Y_S and Y_L according to some optimizing criterion. Usual values for the Stern factor are : $4 \le K \le 10$. If we want to design the amplifier for a minimum noise factor, the value of $R_S = 1/G_S$ is given in the manufacturer data sheet. We can deduce the load conductance from equation(70).

$$G_L = \frac{K\left[\left|y_{12}y_{21}\right| + \text{Re}\left(y_{12}y_{21}\right)\right]}{2\left(g_{11} + G_S\right)} - g_{22} \tag{83}$$

The susceptances B_S and B_L are harder to compute because of the interdependence between the input and output impedances. We can use the following iterative method:

1. We start with $B_L = -b_{22}$.
2. We compute Y_{in} using(67). This provides a value for B_{in}.
3. We set $B_S = -B_{in}$. Using(68), we compute Y_{out}. This provides a value for B_{out}.
4. We replace $B_L = -B_{out}$.
5. We repeat the procedure from step 2.

We usually obtain convergence after two or three iterations.

Example [3]: We have already considered the data for the 2N4957 measured at $f = 200$ MHz, $V_{CEQ} = 10$ V and $I_{CQ} = 2$ mA:

$y_{11} = (2.7 + j6.8)\ \text{m}\Omega^{-1}$; $y_{12} = (0 - \text{j}0.5)\ \text{m}\Omega^{-1}$; $y_{21} = (53 - j22)\ \text{m}\Omega^{-1}$; $y_{22} = (0.1 + j1.5)\ \text{m}\Omega^{-1}$

The Linvill stability factor has already been computed. It is $C = 2.486$. The transistor is potentially unstable. For a DC collector current of 2 mA, the optimum source resistance is $R_S = 200\Omega$. So, $G_S = 1/R_S = 5\ \text{m}\Omega^{-1}$. Using (83) and a value of the Stern factor $K = 4$, we obtain: $G_L = 4.395\ \text{m}\Omega^{-1}$. The iterative process is used to compute B_S and B_L, giving: $B_S = -11.91\ \text{m}\Omega^{-1}$ and $B_L = -4.55\ \text{m}\Omega^{-1}$.

If we use these values, we obtain a transducer gain $G_T = 216.5$

[3] From Krauss, H. L., Bostian, C. W. and Raab, F. H., *Solid State Radio Engineering*, John Wiley & Sons, 1980.

If the design criterion is the maximum transducer gain, we can compute the partial derivatives of G_T with respect to G_S, B_S, G_L and B_L and set them to zero. We obtain the following results:

$$G_S = \sqrt{\frac{K\left[\left|y_{12}y_{21}\right| + \mathrm{Re}\left(y_{12}y_{21}\right)\right]g_{11}}{2g_{22}}} - g_{11} \tag{84}$$

$$G_L = \sqrt{\frac{K\left[\left|y_{12}y_{21}\right| + \mathrm{Re}\left(y_{12}y_{21}\right)\right]g_{22}}{2g_{11}}} - g_{22} \tag{85}$$

The expressions for B_S and B_L are quite involved. The iterative procedure given above is much easier to implement and it provides the same result. We can also remark that the following ratios are equal:

$$\frac{G_L}{g_{22}} = \frac{G_S}{g_{11}} = R \tag{86}$$

If we use (86) in the expression of the Stern factor (equation(70)), we obtain the following result:

$$\left(1+R\right)^2 = K\left[\frac{\left|y_{12}y_{21}\right| + \mathrm{Re}\left(y_{12}y_{21}\right)}{2g_{11}g_{22}}\right] \tag{87}$$

If we use the above relations with the 2N4957 transistor used in the previous example, we obtain (for a Stern factor $K = 4$):

$G_S = 28.21$ mΩ^{-1} and $G_L = 1.045$ mΩ^{-1}. The susceptances are computed using the iterative method and we obtain: $B_S = -27.3$ mΩ^{-1} and $B_L = -2.26$ mΩ^{-1}. The transducer gain is $G_T = 283.9$. However, using $R_S = 200\ \Omega$ (optimum noise), the noise figure is $NF = 1.8$ dB and the gain is $G_T = 216.5$, while with this design, the noise figure is $NF = 3.5$ dB.

Overall design

An amplifier is always connected between a source with internal impedance not always equal to the inverse of the optimum Y_S computed in the previous parts and load impedance different from the inverse of the optimum Y_L. It needs matching networks at the input and the output.

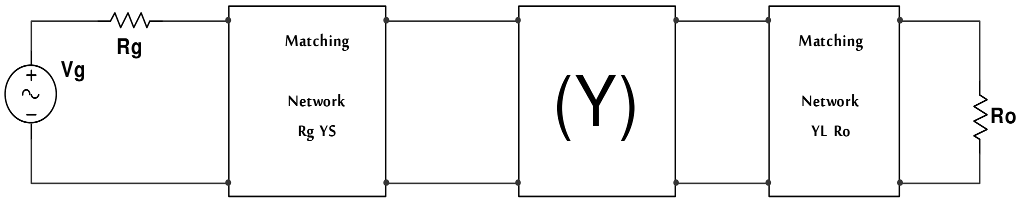


Figure 39 Overall Design

The above figure describes typical design of an amplifier. We first select a device that can be operated under the specified conditions. Using the manufacturer data, we select a Q point. This will determine the small signal parameters of the active device. We then compute the optimal values of the source and load admittances. Finally, we use matching networks to transform the actual source resistance R_g to Y_S and the actual load R_o should appear as Y_L.

As an example, let consider a BJT biased at $I_{CQ} = 1$ mA and $V_{CEQ} = 10$ V. Its admittance parameters are (at $f_0 = 5$ MHz):

$y_{11} = (3.6 + j1.8)\ \text{m}\Omega^{-1}$, $y_{12} = (-0.04 - j0.6)\ \text{m}\Omega^{-1}$, $y_{21} = (29 - j10)\ \text{m}\Omega^{-1}$ and $y_{22} = (0.0029 + j0.6)\ \text{m}\Omega^{-1}$.

The Linvill stability factor is $C = 3.79$. The amplifier is potentially unstable. For optimum noise figure, the source resistance must be $R_S = 400\ \Omega$ giving $G_S = 2.5\ \text{m}\Omega^{-1}$. We select a Stern stability factor $K = 4$. The value of G_L is $G_L = 3.6975\ \text{m}\Omega^{-1}$. We compute the values of the susceptances using the iterative method shown previously. We obtain $B_S = -22.07\ \text{m}\Omega^{-1}$ and $B_L = -3.07\ \text{m}\Omega^{-1}$. The input and output admittances of the device are computed according to (67) and(68). Their values are $Y_{in} = (2.82 + j22.07)\ \text{m}\Omega^{-1}$ and $Y_{out} = (0.4714 + j3.07)\ \text{m}\Omega^{-1}$. We remark that the susceptances are compensated but the conductances are not optimum for maximum gain.

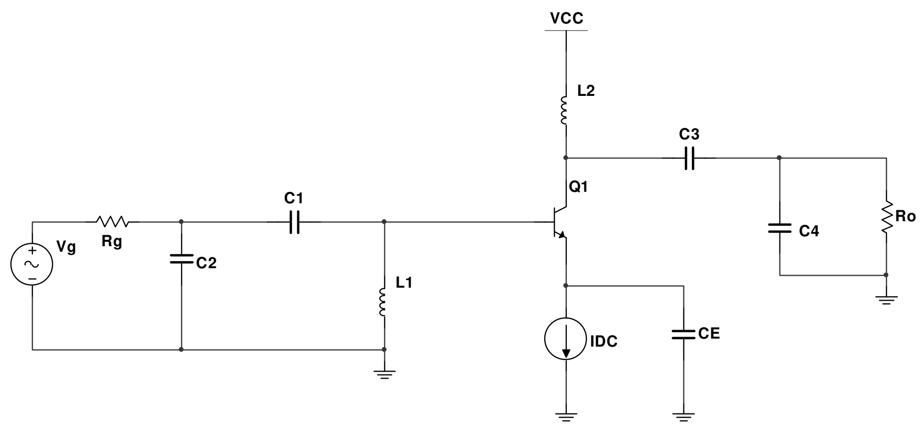


Figure 40 Amplifier at 5 MHz

We select the circuit shown in Figure 40. The matching networks are the common split capacitor matching networks as seen in lab#2. The resistances R_g and R_o are equal and their value is 50Ω. The bandwidths of the input and the output circuits are equal and fixed at 0.25 MHz. The biasing is satisfied with a DC current mirror of 1 mA and a power supply $V_{CC} = 9.25$ V. A value of $C_E = 100$ nF has an impedance of 0.32 Ω at 5 MHz so it provides a good bypass.

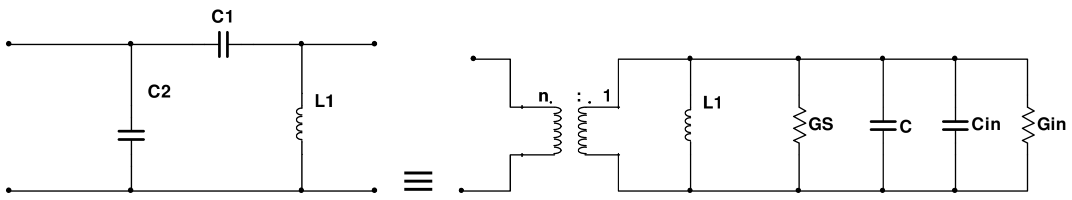


Figure 41 Input network

We start by the design of the input network composed of L1, C1 and C2. It is used to transform the resistance R_g to $Z_S = \frac{1}{Y_S}$, tune at 5 MHz and set the bandwidth of the input at 0.25 MHz. The total conductance that appears at the input of the transistor is $G = G_{in} + G_S = 5.32\ \text{m}\Omega^{-1}$. The total Q of the input network is:

$$Q_T = \frac{f_0}{B} = \frac{5}{0.25} = 20$$

Using $Q_T = RC_T\omega_0$ and $R = \frac{1}{G}$, we can compute the total capacitance C_T of the input tank circuit. This gives $C_T = \frac{GQ_T}{2\pi f_0} = 3387\text{pF}$. This capacitance includes the input capacitance of the active device. Since $B_{in} = C_{in}\omega_0$, we have $C_{in} = \frac{B_{in}}{2\pi f_0} = 702.5\text{pF}$. So, the capacitance that must be implemented by C1 and C2 is: $C = C_T - C_{in} = 2684.5$ pF. The inductance tunes C_T to $f_0 = 5$ MHz. So, $\text{L1} = \frac{1}{C_T\omega_0^2} = 0.3\mu\text{H}$. C is realized by the series combination of C1 and C2 and at the same time, these two capacitances form the turn ratio $n = \frac{\text{C1}}{\text{C1+C2}}$. This turn ratio is also given by $n = \sqrt{\frac{R_g}{R_S}} = \sqrt{\frac{50}{400}} = 0.3535$. $\text{C2} = \frac{C}{n} = 7594$ pF and $\text{C1} = \frac{C}{1-n} = 4152$ pF. As seen in the course (Chapter 2, p.54), the assumptions of narrowband and high Q are justified.

For the output matching network composed of L2, C3 and C4, we repeat the design done in lab#2. Here, the conductance that appears at the primary of the tank circuit is $G = G_{out} + G_L = 4.1689\ \text{m}\Omega^{-1}$. The total Q is also equal to 20. So, the total capacitance appearing at the primary of the ideal transformer is: $C_T = \frac{GQ_T}{2\pi f_0} = 2.654\ \text{nF}$. This capacitance includes the output capacitance of the transistor. So, the capacitance used to perform the impedance transformation is: $C = C_T - C_{out}$. C_{out} is computed from B_{out}. We have

C_{out} = 97.72 pF. So, C = 2.556 nF. The turn ratio is given by $n = \sqrt{\frac{R_o}{R_L}} = \sqrt{R_o G_L} = 0.4299$. The inductance value is $L2 = \frac{1}{C_T \omega_0^2} = 0.381\,\mu\text{H}$. We obtain C3 = 4.484 nF and C4 = 5.945 nF. We verify the assumptions: $Q_{T'}$ = 21.75, Q_E = 16.37 so $nQ_{T'}Q_E$ = 153. The values can be accepted. If we don't have bandwidth constraints, we can use "ell" circuits for matching (it can be the case if we use ceramic or SAW filters.

When we implement the above design, the most difficult part will be tuning of the input and output networks. We have seen that the input admittance depends on the load and that the output admittance depends on the source. So, if we modify the output capacitance to tune the output circuit, this will modify the tuning of the input circuit and vice versa. One way to avoid this problem is to use a unilateral device or to transform the actual device into a unilateral one.

Neutralization

If we connect two devices in parallel, their Y matrices will add.

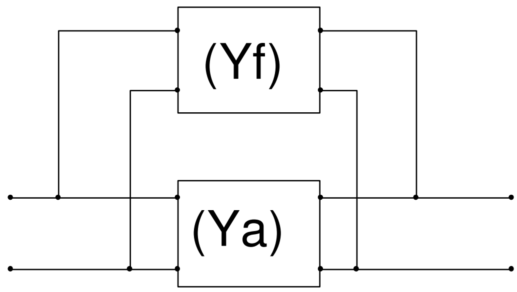


Figure 42 parallel combination

In the above configuration, the total admittance matrix is given by:

$$\mathbf{Yt = Ya + Yf} \tag{88}$$

$$\begin{pmatrix} y_{11t} & y_{12t} \\ y_{21t} & y_{22t} \end{pmatrix} = \begin{pmatrix} y_{11f} & y_{12f} \\ y_{21f} & y_{22f} \end{pmatrix} + \begin{pmatrix} y_{11a} & y_{12a} \\ y_{21a} & y_{22a} \end{pmatrix} \tag{89}$$

If $y_{12f} = -\,y_{12a}$, the composite device becomes unilateral.

Consider the following network consisting of a single admittance connected between the input and the output.

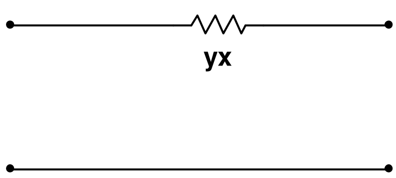


Figure 43 Simple feedback

Its admittance matrix is:

$$\begin{pmatrix} y_x & -y_x \\ -y_x & y_x \end{pmatrix} \tag{90}$$

The admittance matrix of the composite network becomes:

$$\begin{pmatrix} y_{11a} + y_x & y_{12a} - y_x \\ y_{21a} - y_x & y_{22a} + y_x \end{pmatrix} \tag{91}$$

In general, the internal feedback in active devices is capacitive. So, an inductive feedback will compensate for this y_{12}. This can be achieved by the network shown in the following figure.

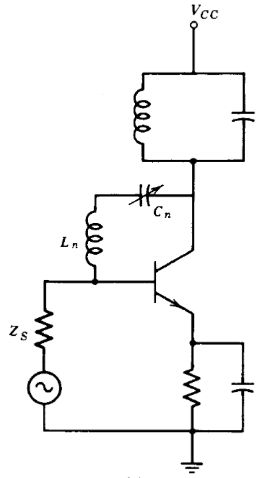


Figure 44 Neutralization

If we use the above circuit for neutralization, the admittance y_x is:

$$y_x = \frac{jC_n\omega}{1 - L_n C_n \omega^2} \tag{92}$$

In order to compensate for y_{12}, y_x must be inductive. Since it is very difficult to implement very small inductances, this reactance is realized with a series combination of an inductance and a capacitance. In the above circuit, the total reactance must be inductive. We can adjust y_x using an adjustable capacitor.

Cascode circuit

The cascode circuit is a cascade of a common emitter (CE) circuit and a common base (CB) circuit or a common source (CS) circuit and a common gate (CG) circuit. This circuit presents a very small reverse transadmittance y_{12}. In this section, we are going to derive the admittance parameters of the cascode circuit from the hybrid pi network seen in page 38. The derivation will be done for BJT. The one for FET is left as an exercise.

First, we provide the admittance parameters as a function of the Giacoletto model. However, we are going to neglect $r_{bb'}$ in order to simplify the expressions.

Let $g_{b'e} = \frac{1}{r_{b'e}}$ and $g_{ce} = \frac{1}{r_{ce}}$. The admittance parameters for the CE BJT are:

$$y_{11e} = g_{b'e} + j\left(C_{b'e} + C_{b'c}\right)\omega \tag{93}$$

$$y_{12e} = -jC_{b'c}\omega \tag{94}$$

$$y_{21e} = g_m - jC_{b'c}\omega \tag{95}$$

$$y_{22e} = g_{ce} + jC_{b'c}\omega \tag{96}$$

We can derive the common base admittance parameters from the common emitter ones (or the common gate parameters from the common source ones). We find also these formulas tabulated in many references).

The common base (gate) parameters as a function of the common emitter (source) admittance are:

$$y_{11b} = y_{11e} + y_{12e} + y_{21e} + y_{22e} \tag{97}$$

$$y_{12b} = -\left(y_{12e} + y_{22e}\right) \tag{98}$$

$$y_{21b} = -\left(y_{21e} + y_{22e}\right) \tag{99}$$

$$y_{22b} = y_{22e} \tag{100}$$

Replacing in the equations (97) to(100), we obtain:

$$y_{11b} = g_{b'e} + g_{ce} + g_m + jC_{b'e}\omega \tag{101}$$

$$y_{12b} = -g_{ce} \tag{102}$$

$$y_{21b} = -g_m - g_{ce} \tag{103}$$

$$y_{22b} = g_{ce} + jC_{b'c}\omega \tag{104}$$

In order to derive the admittance parameters of the cascode, we have to derive the admittance parameters of a cascade of two devices as shown below.

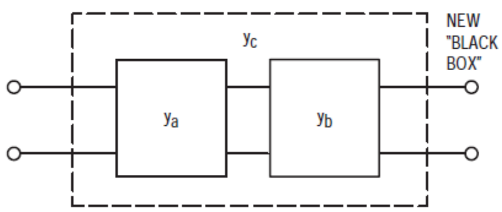


Figure 45 Cascade connection of two devices

The composite admittance parameters are given by:

$$y_{11c} = y_{11a} - \frac{y_{12a} y_{21a}}{y_{22a} + y_{11b}} \tag{105}$$

$$y_{12c} = -\frac{y_{12a} y_{12b}}{y_{22a} + y_{11b}} \tag{106}$$

$$y_{21c} = -\frac{y_{21a} y_{21b}}{y_{22a} + y_{11b}} \tag{107}$$

$$y_{22c} = y_{22b} - \frac{y_{12b} y_{21b}}{y_{22a} + y_{11b}} \tag{108}$$

If we use the above relations directly, we are going to obtain quite complex expressions. In order to be able to draw any significant conclusion, we must simplify them. First, looking at the expressions (101) to(104), we can write: $\left(y_{22a} + y_{11b}\right) \simeq y_{11b}$. It is also evident that $\left|y_{11a}\right| \gg \left|y_{12b}\right|$ and finally, we can always make $y_{11b} \simeq -y_{21b}$.

$$y_{11c} \simeq y_{11a} \tag{109}$$

$$y_{12c} \simeq \frac{y_{12a} y_{12b}}{y_{21b}} \tag{110}$$

$$y_{21c} \simeq y_{21a} \tag{111}$$

$$y_{22c} \simeq y_{22b} \tag{112}$$

So, the cascode circuit has the input admittance of the common emitter, the transadmittance of the common emitter, the output admittance of the common base and a very small reverse transadmittance. If we replace in the above expressions the values derived as a function of the hybrid pi model, we obtain:

$$y_{11c} \simeq g_{b'e1} + j\left(C_{b'e1} + C_{b'c}\right)\omega \tag{113}$$

$$y_{12c} \simeq -j\frac{C_{b'c1} g_{ce2} \omega}{g_{m2} + g_{ce2}} \tag{114}$$

$$y_{21c} \simeq g_{m1} - jC_{b'c1}\omega \tag{115}$$

$$y_{22c} \simeq g_{ce2} + jC_{b'c2}\omega \tag{116}$$

The most important result above is(114). The output impedance of a transistor is very large, so g_{ce} is very small. We have also g_m at the denominator which is quite large. So, we can conclude that the cascode circuit is practically unilateral.

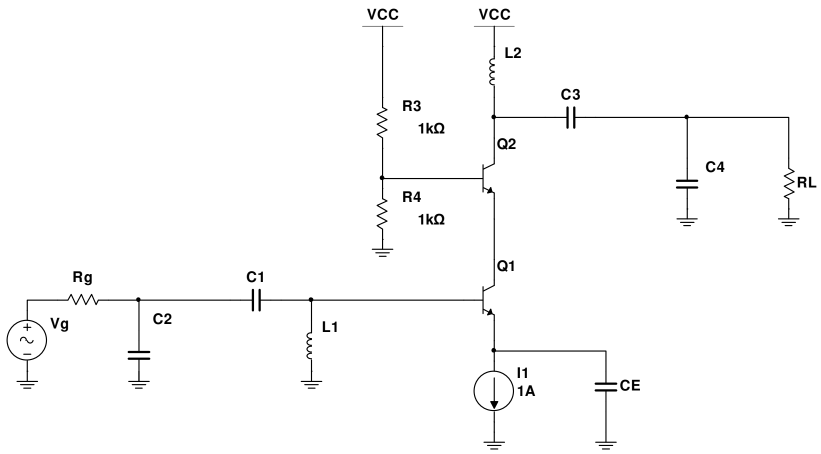


Figure 46 BJT Cascode Circuit

The design of the above circuit is straightforward. The compound device parameters can be calculated from equations (105) to (108) from the transistor datasheet. Those equations are valid for FET devices. A device which has a built in cascode property is the dual gate MOSFET. It can be considered as a cascade of a common source amplifier with a common gate amplifier.

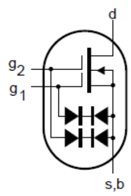


Figure 47 BF996S Dual gate MOSFET

DYNAMIC CHARACTERISTICS

Measuring conditions (common source): I_D = 10 mA; V_{DS} = 15 V; $V_{G2\text{-}S}$ = 4 V; T_{amb} = 25 °C.

SYMBOL	PARAMETER	CONDITIONS	MIN.	TYP.	MAX.	UNIT
$\|Y_{fs}\|$	transfer admittance	f = 1 kHz	15	18	–	mS
$C_{ig1\text{-}s}$	input capacitance at gate 1	f = 1 MHz	–	2.3	2.6	pF
$C_{ig2\text{-}s}$	input capacitance at gate 2	f = 1 MHz	–	1.2	–	pF
C_{rs}	feedback capacitance	f = 1 MHz	–	25	–	fF
C_{os}	output capacitance	f = 1 MHz	–	0.8	–	pF
F	noise figure	f = 200 MHz; G_S = 2 mS; $B_S = B_{Sopt}$	–	1	–	dB
		f = 800 MHz; G_S = 3.3 mS; $B_S = B_{Sopt}$	–	1.8	–	dB
G_P	power gain	f = 200 MHz; G_S = 2 mS; $B_S = B_{Sopt}$; G_L = 0.5 mS; $B_L = B_{Lopt}$	–	25	–	dB
		f = 800 MHz; G_S = 3.3 mS; $B_S = B_{Sopt}$; G_L = 1 mS; $B_L = B_{Lopt}$	–	18	–	dB

The table shown above is from the BF996S dual gate MOSFET. We can remark that the feedback capacitance is practically zero.

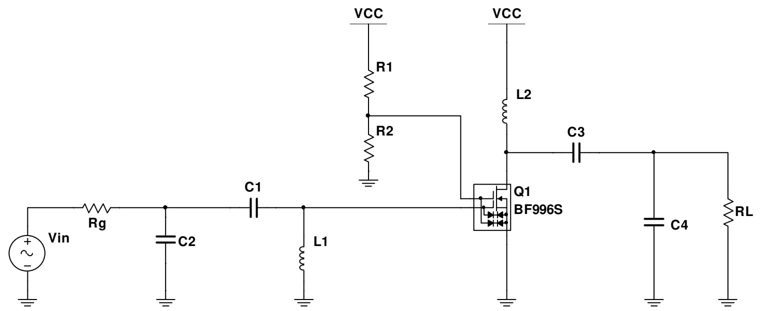


Figure 48 Dual Gate MOSFET Amplifier

The above figure shows a dual gate MOSFET used with matching networks at the input and output.