Chapter Ten

Noise and Nonlinear Distortion

The effect of noise is critical to the performance of most RF and microwave communications, radar, and remote sensing systems because noise ultimately determines the threshold for the minimum signal that can be reliably detected by a receiver. Noise power in a receiver will be introduced from the external environment through the receiving antenna, as well as generated internally by the receiver circuitry. Here we will study the sources of noise in RF and microwave systems, and the characterization of components in terms of noise temperature and noise figure, including the effect of impedance mismatch. The additional noise-related topics of transistor amplifier noise figure, oscillator phase noise, and antenna noise temperature will be discussed in later chapters.

We will also discuss the related topics of compression, harmonic distortion, intermodulation distortion, and dynamic range. These can have important limiting effects when large signal levels are present in mixers, amplifiers, and other components that use nonlinear devices such as diodes and transistors.

10.1 NOISE IN MICROWAVE CIRCUITS

Noise power is a result of random processes such as the flow of charges or holes in an electron tube or solid-state device, propagation through the ionosphere or other ionized gas, or, most basic of all, the thermal vibrations in any component at a temperature above absolute zero. Noise can be passed into a microwave system from external sources, or generated within the system itself. In either case the noise level of a system sets the lower limit on the strength of a signal that can be detected in the presence of the noise. Thus, it is generally desired to minimize the residual noise level of a radar or communications receiver to achieve the best performance. In some cases, such as radiometers or radio astronomy systems, the desired signal is actually the noise power received by an antenna, and it is necessary to distinguish between the received noise power and the undesired noise generated by the receiver system itself.

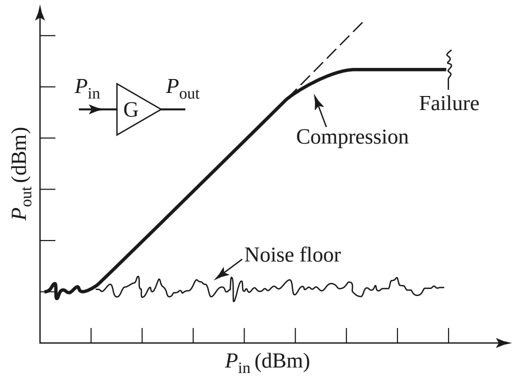


FIGURE 10.1 Illustrating the dynamic range of a realistic amplifier.

Dynamic Range and Sources of Noise

In previous chapters we have implicitly assumed that all components were *linear* (meaning that the output signal level is directly proportional to the input signal level), and *deterministic* (meaning that the output signal is predictable from the input signal). In reality no component can perform in this way over an unlimited range of input/output signal levels. In practice, however, there is usually a range of signal levels over which such assumptions are approximately valid; this range is called the *dynamic range* of the component.

As an example, consider a realistic microwave transistor amplifier having a power gain G, as shown in Figure 10.1. If the amplifier were ideal, the output power would be related to the input power as $P_{out} = GP_{in}$, and this relation would hold true for any value of P_{in}. Thus, if $P_{in} = 0$, we would have $P_{out} = 0$, and if $P_{in} = 10^6$ W and $G = 10$ dB, we would have $P_{out} = 10^7$ W. Neither of these results would actually occur in practice, however. Because of noise generated by the amplifier itself, some nonzero noise power will always be delivered by the amplifier, even when the input power is zero. At the other extreme, very high input power will cause the amplifier to fail. Thus, the actual relation between the output and input power will be as shown in Figure 10.1. At very low input power levels, the output will be dominated by the noise generated by the amplifier. This level is often called the *noise floor* of the component or system; typical values may range from -80 to -140 dBm over the bandwidth of the system, with the lowest values being obtained with thermally cooled components. Above the noise floor, the amplifier will have a range of input power for which $P_{out} = GP_{in}$ is closely approximated. This is the usable *dynamic range* of the component. At the upper end of this range, the output will begin to saturate, meaning that the output power no longer increases linearly as the input power increases. Excessive input power will lead to failure of the amplifier.

Noise that is generated internally in a device or component is usually caused by random motions of charges or charge carriers in devices and materials. Such motions may be due to any of several mechanisms, leading to various types of noise:

- *Thermal noise* is the most basic type of noise, being caused by thermal vibration of bound charges. It is also known as *Johnson* or *Nyquist* noise.
- *Shot noise* is due to random fluctuations of charge carriers in an electron tube or solid-state device.
- *Flicker noise* occurs in solid-state components and vacuum tubes. Flicker noise power varies inversely with frequency, and so is often called $1/f$-noise.

- *Plasma noise* is caused by random motion of charges in an ionized gas, such as a plasma, the ionosphere, or sparking electrical contacts.
- *Quantum noise* results from the quantized nature of charge carriers and photons; it is often insignificant relative to other noise sources.

External noise may be introduced into a system either by a receiving antenna or by electromagnetic coupling. Some sources of external RF noise include the following:

- Thermal noise from the ground
- Cosmic background noise from the sky
- Noise from stars (including the sun)
- Lightning
- Gas discharge lamps
- Radio, TV, and cellular stations
- Wireless devices
- Microwave ovens
- Deliberate jamming devices

The characterization of noise effects in RF and microwave systems in terms of noise temperature and noise figure will apply to all types of noise, regardless of the source, as long as the spectrum of the noise is relatively flat over the bandwidth of the system. Noise with a flat frequency spectrum is called *white noise*.

Noise Power and Equivalent Noise Temperature

Consider a resistor at a physical temperature of T degrees kelvin (K), as depicted in Figure 10.2. The electrons in the resistor are in random motion, with a kinetic energy that is proportional to the temperature. These random motions produce small, random voltage fluctuations at the resistor terminals, as illustrated in Figure 10.2. This voltage has a zero average value but a nonzero root mean square (rms) value given by Planck's blackbody radiation law,

$$V_n = \sqrt{\frac{4hf\ BR}{e^{hf/kT} - 1}}, \tag{10.1}$$

where

$h = 6.626 \times 10^{-34}$ J-sec is Planck's constant.
$k = 1.380 \times 10^{-23}$ J/K is Boltzmann's constant.
$T =$ the temperature in degrees kelvin (K).
$B =$ the bandwidth of the system in Hz.
$f =$ the center frequency of the bandwidth in Hz.
$R =$ the resistance in Ω.

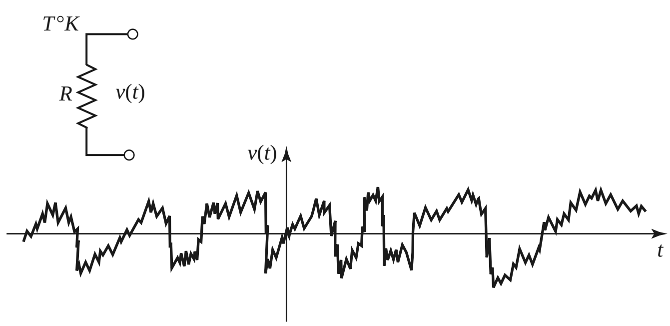


FIGURE 10.2 A random voltage generated by a noisy resistor.

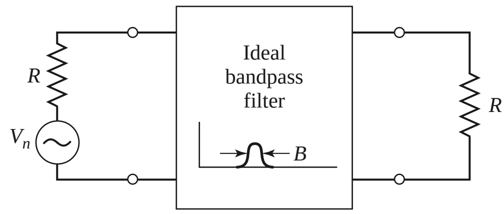


FIGURE 10.3 Equivalent circuit of a noisy resistor delivering maximum power to a load resistor through an ideal bandpass filter.

This result comes from quantum mechanical considerations, and is valid for any frequency f. At microwave frequencies the above result can be simplified by making use of the fact that $hf \ll kT$. (As a worst-case example, let $f = 100$ GHz and $T = 100$ K. Then $hf = 6.6 \times 10^{-23} \ll kT = 1.4 \times 10^{-21}$.) Using the first two terms of a Taylor series expansion for the exponential in (10.1) gives

$$e^{hf/kt} - 1 \simeq \frac{hf}{kT},$$

so that (10.1) reduces to

$$V_n = \sqrt{4kTBR}. \tag{10.2}$$

This is the *Rayleigh–Jeans approximation,* and is the result that is most commonly used in microwave work [1]. For very high frequencies or very low temperatures, however, this approximation may be invalid, in which case (10.1) should be used.

The noisy resistor of Figure 10.2 can be replaced with a Thevenin equivalent circuit consisting of a noiseless resistor and a generator with a voltage given by (10.2), as shown in Figure 10.3. Connecting a load resistor R results in maximum power transfer from the noisy resistor, with the result that power delivered to the load in a bandwidth B is

$$P_n = \left(\frac{V_n}{2R}\right)^2 R = \frac{V_n^2}{4R} = kTB, \tag{10.3}$$

since V_n is an rms voltage. This important result gives the maximum available noise power from the noisy resistor at temperature T. Note that this noise power is independent of frequency; such a noise source has a power spectral density that is constant with frequency, and is an example of a white noise source. The noise power is directly proportional to the bandwidth, which in practice is usually limited by the passband of the RF or microwave system. Independent white noise sources can be treated as Gaussian-distributed random variables, so the noise powers (variances) of independent noise sources are additive.

The following trends can be observed from (10.3):

- As $B \to 0$, $P_n \to 0$. This means that systems with smaller bandwidths collect less noise power.
- As $T \to 0$, $P_n \to 0$. This means that cooler devices and components generate less noise power.
- As $B \to \infty$, $P_n \to \infty$. This is the so-called *ultraviolet catastrophe,* which does not occur in reality because (10.2)–(10.3) are not valid as f (or B) $\to \infty$; (10.1) must be used in this case.

If an arbitrary source of noise (thermal or nonthermal) is "white," so that the noise power is not a strong function of frequency over the bandwidth of interest, it can be modeled as an equivalent thermal noise source, and characterized with an *equivalent noise*

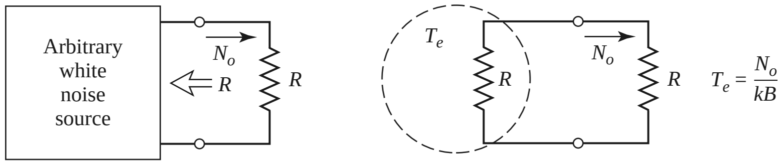


FIGURE 10.4 The equivalent noise temperature, T_e, of an arbitrary white noise source.

temperature. Thus, consider the arbitrary white noise source of Figure 10.4, which has a driving-point impedance of R and delivers a noise power N_o to a load resistor R. This noise source can be replaced by a noisy resistor of value R at temperature T_e, where T_e is an equivalent temperature selected so that the same noise power is delivered to the load. That is,

$$T_e = \frac{N_o}{kB}. \tag{10.4}$$

Components and systems can then be characterized by saying that they have an equivalent noise temperature T_e; this implies some fixed bandwidth B, which is generally the operational bandwidth of the component or system.

For example, consider a noisy amplifier with a bandwidth B and gain G. Let the amplifier be matched to noiseless source and load resistors, as shown in Figure 10.5. If the source resistor is at a (hypothetical) temperature of $T_s = 0\,K$, then the input power to the amplifier will be $N_i = 0$, and the output noise power N_o will be due only to the noise generated by the amplifier itself. We can obtain the same load noise power by driving an ideal noiseless amplifier with a resistor at the temperature

$$T_e = \frac{N_o}{GkB}, \tag{10.5}$$

so that the output power in both cases is $N_o = GkT_eB$. Then T_e is the equivalent noise temperature of the amplifier.

It is sometimes useful for measurement purposes to have a calibrated noise source. A passive noise source may simply consist of a resistor held at a constant temperature, either

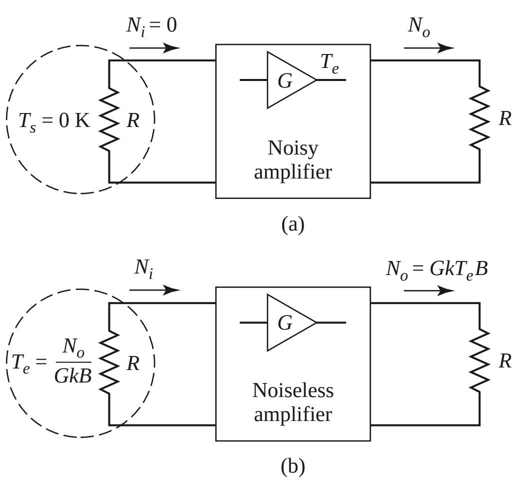


FIGURE 10.5 Defining the equivalent noise temperature of a noisy amplifier. (a) Noisy amplifier. (b) Noiseless amplifier.

in a temperature-controlled oven, or in a cryogenic flask. Active noise sources may use a diode, transistor, or tube to provide a calibrated noise power output. Noise generators can be characterized by an equivalent noise temperature, but a more common measure of noise power for such components is the *excess noise ratio* (ENR), defined as

$$\text{ENR (dB)} = 10 \log \frac{N_g - N_o}{N_o} = 10 \log \frac{T_g - T_0}{T_0}, \tag{10.6}$$

where N_g and T_g are the noise power and equivalent noise temperature of the generator, and N_o and T_0 are the noise power and temperature associated with a room-temperature ($T_0 = 290$ K) passive source (a matched load). Solid-state noise generators typically have ENRs ranging from 20 to 40 dB.

Measurement of Noise Temperature

In principle, the equivalent noise temperature of a component can be determined by measuring the output power when a matched load at 0 K is connected at the input of the component. In practice, of course, a 0 K source temperature cannot be obtained, so a different method must be used. If two matched loads at significantly different temperatures are available, then the *Y-factor method* can be applied.

This technique is illustrated in Figure 10.6, where the amplifier (or other component) under test is connected to one of two matched loads at different temperatures, and the output power is measured for each case. Let T_1 be the temperature of the hot load and T_2 the temperature of the cold load ($T_1 > T_2$), and let P_1 and P_2 be the respective powers measured at the amplifier output. The output noise power consists of noise power generated by the amplifier as well as noise power from the source resistor. Thus we have

$$N_1 = GkT_1B + GkT_eB, \tag{10.7a}$$

$$N_2 = GkT_2B + GkT_eB, \tag{10.7b}$$

which are two equations for the two unknowns, T_e and GB (the gain–bandwidth product of the amplifier). Define the Y-factor as

$$Y = \frac{N_1}{N_2} = \frac{T_1 + T_e}{T_2 + T_e} > 1, \tag{10.8}$$

which is determined as the ratio of the output power measurements. Then (10.7) can be solved for the equivalent noise temperature of the device under test as

$$T_e = \frac{T_1 - YT_2}{Y - 1}, \tag{10.9}$$

in terms of the load temperatures and the Y-factor.

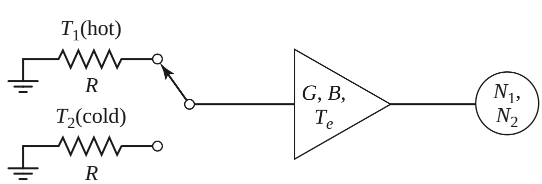


FIGURE 10.6 The Y-factor method for measuring the equivalent noise temperature of an amplifier.

Note that to obtain accurate results from this method, the two source temperatures must not be too close together. If they are, N_1 will be close to N_2, Y will be close to unity, and the evaluation of (10.9) will involve the subtractions of numbers close to each other, resulting in a loss of accuracy. In practice, one noise source is usually a load resistor at room temperature ($T_0 = 290$ K), while the other noise source is either "hotter" or "colder," depending on whether T_e is greater or less than T_0. An active noise generator can be used as a "hot" source, while a "cold" source can be obtained by immersing a load resistor in liquid nitrogen ($T = 77$ K) or liquid helium ($T = 4$ K).

EXAMPLE 10.1 NOISE TEMPERATURE MEASUREMENT

An X-band amplifier has a gain of 20 dB and a 1 GHz bandwidth. Its equivalent noise temperature is to be measured via the Y-factor method. The following data are obtained:

$$\text{For } T_1 = 290 \text{ K}, \qquad N_1 = -62.0 \text{ dBm}.$$
$$\text{For } T_2 = 77 \text{ K}, \qquad N_2 = -64.7 \text{ dBm}.$$

Determine the equivalent noise temperature of the amplifier. If the amplifier is used with a source having an equivalent noise temperature of $T_s = 450$ K, what is the output noise power from the amplifier, in dBm?

Solution

From (10.8), the Y-factor in dB is

$$Y = (N_1 - N_2) \text{ dB} = (-62.0) - (-64.7) = 2.7 \text{ dB},$$

which is a numeric value of $Y = 1.86$. Using (10.9) gives the equivalent noise temperature as

$$T_e = \frac{T_1 - YT_2}{Y - 1} = \frac{290 - (1.86)(77)}{1.86 - 1} = 170 \text{ K}.$$

If a source with an equivalent noise temperature of $T_s = 450$ K drives the amplifier, the noise power into the amplifier will be kT_sB. The total noise power out of the amplifier will be

$$N_o = GkT_sB + GkT_eB = 100(1.38 \times 10^{-23})(10^9)(450 + 170)$$
$$= 8.56 \times 10^{-10} \text{ W} = -60.7 \text{ dBm}.$$

■

10.2 NOISE FIGURE

Definitio of Noise Figure

We have seen that a noisy microwave component can be characterized by an equivalent noise temperature. An alternative characterization is the *noise figure* of the component, which is a measure of the degradation in the signal-to-noise ratio between the input and output of the component. The *signal-to-noise ratio* is the ratio of desired signal power to undesired noise power, and so is dependent on the signal power. When noise and a desired signal are applied to the input of a noiseless network, both noise and signal will be attenuated or amplified by the same factor, so that the signal-to-noise ratio will be unchanged. However, if the network is noisy, the output noise power will be increased more than the

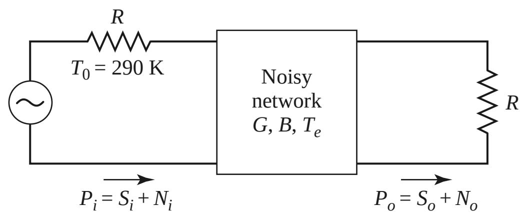


FIGURE 10.7 Determining the noise figure of a noisy network.

output signal power, so that the output signal-to-noise ratio will be reduced. The noise figure, F, is a measure of this reduction in signal-to-noise ratio, and is defined as

$$F = \frac{S_i/N_i}{S_o/N_o} \geq 1, \tag{10.10}$$

where S_i, N_i are the input signal and noise powers, and S_o, N_o are the output signal and noise powers. By definition, the input noise power is assumed to be the noise power resulting from a matched resistor at $T_0 = 290$ K; that is, $N_i = kT_0B$.

Consider Figure 10.7, which shows noise power N_i and signal power S_i being fed into a noisy two-port network. The network is characterized by a gain, G, a bandwidth, B, and an equivalent noise temperature, T_e. The input noise power is $N_i = kT_0B$, and the output noise power is a sum of the amplified input noise and the internally generated noise: $N_o = kGB(T_0 + T_e)$. The output signal power is $S_o = GS_i$. Using these results in (10.10) gives the noise figure as

$$F = \frac{S_i}{kT_0B}\frac{kGB(T_0 + T_e)}{GS_i} = 1 + \frac{T_e}{T_0} \geq 1. \tag{10.11}$$

In dB, $F = 10\log(1 + T_e/T_0)$ dB ≥ 0. If the network were noiseless, T_e would be zero, giving $F = 1$, or 0 dB. Solving (10.11) for T_e gives

$$T_e = (F - 1)T_0. \tag{10.12}$$

It is important to keep in mind two things concerning the definition of noise figure: noise figure is defined for a matched input source, and for a noise source equivalent to a matched load at temperature $T_0 = 290$ K. Noise figure and equivalent noise temperatures are interchangeable characterizations of the noise properties of a component.

An important special case occurs in practice for a two-port network consisting of a passive, lossy component, such as an attenuator or lossy transmission line, held at a physical temperature T. Consider such a network with a matched source resistor that is also at temperature T, as shown in Figure 10.8. The power gain, G, of a lossy network is less than unity; the loss factor, L, can be defined as $L = 1/G > 1$. Because the entire system is in thermal equilibrium at the temperature T, and has a driving point impedance of R, the output noise power must be $N_o = kTB$. However, we can also think of this power as coming from the source resistor (attenuated by the lossy line), and from the noise generated by the line itself. Thus we also have that

$$N_o = kTB = GkTB + GN_{\text{added}}, \tag{10.13}$$

where N_{added} is the noise generated by the line, as if it appeared at the input terminals of the line. Solving (10.13) for this power gives

$$N_{\text{added}} = \frac{1 - G}{G}kTB = (L - 1)kTB. \tag{10.14}$$

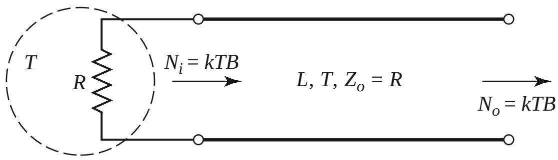


FIGURE 10.8 Determining the noise figure of a lossy line or attenuator with loss L and temperature T.

Then (10.4) shows that the lossy line has an equivalent noise temperature (referred to the input) given by

$$T_e = \frac{1-G}{G}T = (L-1)T. \tag{10.15}$$

From (10.11) the noise figure is

$$F = 1 + (L-1)\frac{T}{T_0}. \tag{10.16}$$

If the line is at temperature T_0, then $F = L$. For instance, a 6 dB attenuator at room temperature has a noise figure of $F = 6$ dB.

Noise Figure of a Cascaded System

In a typical microwave system the input signal travels through a cascade of many different components, each of which may degrade the signal-to-noise ratio to some degree. If we know the noise figure (or noise temperature) of the individual stages, we can determine the noise figure (or noise temperature) of the cascade connection of stages. We will see that the noise performance of the first stage is usually the most critical, an interesting result that is very important in practice.

Consider the cascade of two components, having gains G_1, G_2, noise figures F_1, F_2, and equivalent noise temperatures T_{e1}, T_{e2}, as shown in Figure 10.9. We wish to find the overall noise figure and equivalent noise temperature of the cascade, as if it were a single component. The overall gain of the cascade is G_1G_2.

Using noise temperatures, we can write the noise power at the output of the first stage as

$$N_1 = G_1kT_0B + G_1kT_{e1}B, \tag{10.17}$$

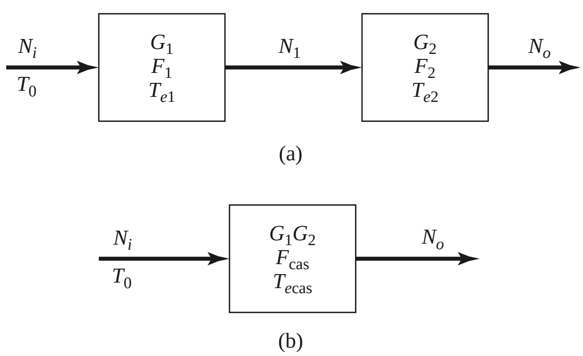


FIGURE 10.9 Noise figure and equivalent noise temperature of a cascaded system. (a) Two cascaded networks. (b) Equivalent network.

since $N_i = kT_0B$ for noise figure calculations. The noise power at the output of the second stage is

$$N_o = G_2N_1 + G_2kT_{e2}B$$
$$= G_1G_2kB\left(T_0 + T_{e1} + \frac{1}{G_1}T_{e2}\right). \tag{10.18}$$

For the equivalent system we have

$$N_o = G_1G_2kB(T_{\text{cas}} + T_0), \tag{10.19}$$

so comparison with (10.18) gives the noise temperature of the cascade system as

$$T_{\text{cas}} = T_{e1} + \frac{1}{G_1}T_{e2}. \tag{10.20}$$

Using (10.12) to convert the temperatures in (10.20) to noise figures yields the noise figure of the cascade system as

$$F_{\text{cas}} = F_1 + \frac{1}{G_1}(F_2 - 1). \tag{10.21}$$

Equations (10.20) and (10.21) show that the noise characteristics of a cascaded system are dominated by the characteristics of the first stage since the effect of the second stage is reduced by the gain of the first (assuming $G_1 > 1$). Thus, for the best overall system noise performance, the first stage should have a low noise figure and at least moderate gain. Expense and effort should be devoted primarily to the first stage, as opposed to later stages, since later stages have a diminished impact on the overall noise performance.

Equations (10.20) and (10.21) can be generalized to an arbitrary number of stages, as follows:

$$T_{\text{cas}} = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1G_2} + \cdots, \tag{10.22}$$

$$F_{\text{cas}} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \cdots. \tag{10.23}$$

EXAMPLE 10.2 NOISE ANALYSIS OF A WIRELESS RECEIVER

The block diagram of a wireless receiver front-end is shown in Figure 10.10. Compute the overall noise figure of this subsystem. If the input noise power from a feeding antenna is $N_i = kT_AB$, where $T_A = 150$ K, find the output noise power in dBm. If we require a minimum signal-to-noise ratio (SNR) of 20 dB at the output of the receiver, what is the minimum signal voltage that should be applied

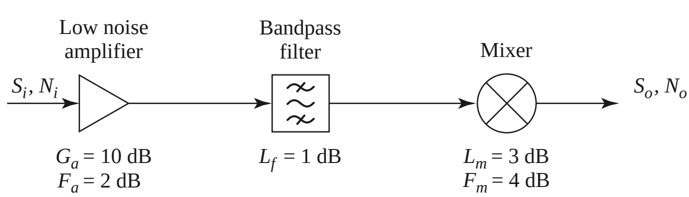


FIGURE 10.10 Block diagram of a wireless receiver front-end for Example 10.2.

at the receiver input? Assume the system is at temperature T_0, with a characteristic impedance of 50 Ω, and an IF bandwidth of 10 MHz.

Solution
We first perform the required conversions from dB to numerical values:

$$G_a = 10 \text{ dB} = 10 \qquad G_f = -1.0 \text{ dB} = 0.79 \qquad G_m = -3.0 \text{ dB} = 0.5$$
$$F_a = 2 \text{ dB} = 1.58 \qquad F_f = 1 \text{ dB} = 1.26 \qquad F_m = 4 \text{ dB} = 2.51$$

Next, use (10.23) to find the overall noise figure of the system:

$$F = F_a + \frac{F_f - 1}{G_a} + \frac{F_m - 1}{G_a G_f} = 1.58 + \frac{(1.26 - 1)}{10} + \frac{(2.51 - 1)}{(10)(0.79)}$$
$$= 1.80 = 2.55 \text{ dB}.$$

The best way to compute the output noise power is to use noise temperatures. From (10.12), the equivalent noise temperature of the overall system is

$$T_e = (F - 1)T_0 = (1.80 - 1)(290) = 232 \text{ K}.$$

The overall gain of the system is $G = (10)(0.79)(0.5) = 3.95$. Then we can find the output noise power as

$$N_o = k(T_A + T_e)BG = (1.38 \times 10^{-23})(150 + 232)(10 \times 10^6)(3.95)$$
$$= 2.08 \times 10^{-13} \text{ W} = -96.8 \text{ dBm}.$$

For an output SNR of 20 dB = 100, the input signal power must be

$$S_i = \frac{S_o}{G} = \frac{S_o}{N_o}\frac{N_o}{G} = 100\frac{2.08 \times 10^{-13}}{3.95} = 5.27 \times 10^{-12} \text{ W} = -82.8 \text{ dBm}.$$

For a 50 Ω system impedance, this corresponds to an input signal voltage of

$$V_i = \sqrt{Z_o S_i} = \sqrt{(50)(5.27 \times 10^{-12})} = 1.62 \times 10^{-5} \text{ V} = 16.2 \ \mu\text{V (rms)}.$$

Note: It may be tempting to compute the output noise power from the definition of the noise figure, as

$$N_o = N_i F\left(\frac{S_o}{S_i}\right) = N_i F G = kT_A BFG$$
$$= (1.38 \times 10^{-23})(150)(10 \times 10^6)(1.8)(3.95) = 1.47 \times 10^{-13} \text{ W}.$$

This is an *incorrect* result! The reason for the disparity with the earlier result is that the definition of noise figure assumes an input noise level of kT_0B, while this problem involves an input noise of kT_AB, with $T_A = 150 \text{ K} \neq T_0$. This is a common error, and suggests that when computing absolute noise power it is often safer to use noise temperatures to avoid this confusion. ■

Noise Figure of a Passive Two-Port Network

We previously derived the noise figure for a matched lossy line or attenuator by using a thermodynamic argument. Here we generalize that technique to evaluate the noise figure of general passive networks (networks that do not contain active devices such as diodes or transistors, which generate nonthermal noise). In addition, this method will account for the change in noise figure that occurs when a component is impedance mismatched at either its

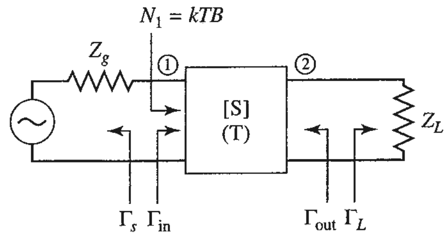


FIGURE 10.11 A passive two-port network with impedance mismatches. The network is at physical temperature T.

input or output port. Generally it is easier and more accurate to find the noise characteristics of an active device, such as a diode or transistor, by direct measurement than by calculation from first principles.

Figure 10.11 shows an arbitrary passive two-port network, with a generator at port 1 and a load at port 2. The network is characterized by its scattering matrix, $[S]$. In the general case, impedance mismatches may exist at each port, and we define these mismatches in terms of the following reflection coefficients:

$$\begin{aligned}
\Gamma_s &= \text{reflection coefficient looking toward generator,} \\
\Gamma_{in} &= \text{reflection coefficient looking toward port 1 of network,} \\
\Gamma_{out} &= \text{reflection coefficient looking toward port 2 of network,} \\
\Gamma_L &= \text{reflection coefficient looking toward load.}
\end{aligned}$$

If we assume the network is at temperature T, and that an available input noise power of $N_1 = kTB$ is applied to the input of the network, we can write the available output noise power at port 2 as

$$N_2 = G_{21}kTB + G_{21}N_{added}, \tag{10.24}$$

where N_{added} is the noise power generated internally by the network (referenced to port 1), and G_{21} is the *available power gain* of the network from port 1 to port 2. The available power gain can be expressed in terms of the scattering parameters of the network and the port mismatches as (also see Section 12.1),

$$G_{21} = \frac{\text{power available from network}}{\text{power available from source}} = \frac{|S_{21}|^2(1-|\Gamma_S|^2)}{|1-S_{11}\Gamma_S|^2(1-|\Gamma_{out}|^2)}. \tag{10.25}$$

As derived in Example 4.7, the output port mismatch is given by

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1-S_{11}\Gamma_S}. \tag{10.26}$$

Observe that when the network is matched to its external circuitry, so that $\Gamma_s = 0$ and $S_{22} = 0$, we have $\Gamma_{out} = 0$ and $G_{21} = |S_{21}|^2$, which is the gain of the network when it is matched. Also observe that the available gain of the network does not depend on the load mismatch, Γ_L. This is because available gain is defined in terms of the maximum power that is available from the network, which occurs when the load impedance is conjugately matched to the output impedance of the network.

Since the input noise power is kTB, and the network is passive and at temperature T, the network is in thermodynamic equilibrium, and so the available output noise power must be $N_2 = kTB$. Then we can solve for N_{added} from (10.24) to give

$$N_{\text{added}} = \frac{1 - G_{21}}{G_{21}} kTB. \tag{10.27}$$

Then the equivalent noise temperature of the network is

$$T_e = \frac{N_{\text{added}}}{kB} = \frac{1 - G_{21}}{G_{21}} T, \tag{10.28}$$

and the noise figure of the network is

$$F = 1 + \frac{T_e}{T_0} = 1 + \frac{1 - G_{21}}{G_{21}} \frac{T}{T_0}. \tag{10.29}$$

Note the similarity of (10.27)–(10.29) to the results in (10.14)–(10.16) for the lossy line—the essential difference is that here we are using the available gain of the network, which accounts for impedance mismatches between the network and the external circuit. We can illustrate the use of this result with some applications to problems of practical interest.

Noise Figure of a Mismatched Lossy Line

Earlier we found the noise figure of a lossy transmission line under the assumption that it was matched to its input and output circuits. Now we consider the case where the line is mismatched to its input circuit. Figure 10.12 shows a transmission line of length ℓ at temperature T, with a power loss factor $L = 1/G$, and an impedance mismatch between the line and the generator. Thus, $Z_g \neq Z_0$, and the reflection coefficient looking toward the generator is

$$\Gamma_s = \frac{Z_g - Z_0}{Z_g + Z_0} \neq 0.$$

The scattering matrix of the lossy line of characteristic impedance Z_0 can be written as

$$[S] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{e^{-j\beta\ell}}{\sqrt{L}}, \tag{10.30}$$

where β is the propagation constant of the line. Using the elements of (10.30) in (10.26) gives the reflection coefficient looking into port 2 of the line as

$$\Gamma_{\text{out}} = S_{22} + \frac{S_{12}S_{21}\Gamma_s}{1 - S_{11}\Gamma_s} = \frac{\Gamma_s}{L} e^{-2j\beta\ell}. \tag{10.31}$$

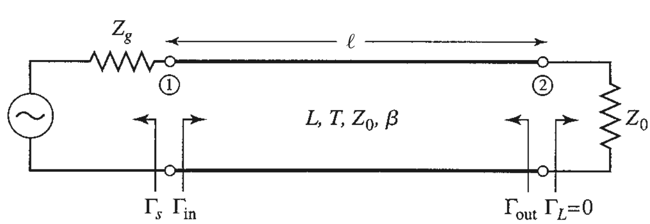


FIGURE 10.12 A lossy transmission line at temperature T with an impedance mismatch at its input port.

Then the available gain, from (10.25), is

$$G_{21} = \frac{\frac{1}{L}(1 - |\Gamma_s|^2)}{1 - |\Gamma_{\text{out}}|^2} = \frac{L(1 - |\Gamma_s|^2)}{L^2 - |\Gamma_s|^2}. \tag{10.32}$$

We can verify two limiting cases of (10.32): when $L = 1$ we have $G_{21} = 1$, and when $\Gamma_s = 0$ we have $G_{21} = 1/L$. Using (10.32) in (10.28) gives the equivalent noise temperature of the mismatched lossy line as

$$T_e = \frac{1 - G_{21}}{G_{21}} T = \frac{(L-1)(L + |\Gamma_s|^2)}{L(1 - |\Gamma_s|^2)} T. \tag{10.33}$$

The corresponding noise figure can then be evaluated using (10.11). Observe that when the line is matched, $\Gamma_s = 0$, and (10.33) reduces to $T_e = (L-1)T$, in agreement with the result for the matched lossy line given by (10.15). If the line is lossless, then $L = 1$, and (10.33) reduces to $T_e = 0$ regardless of mismatch, as expected. However, when the line is lossy and mismatched, so that $L > 1$ and $|\Gamma_s| > 0$, then the noise temperature given by (10.33) is greater than $T_e = (L-1)T$, the noise temperature of the matched lossy line. The reason for this increase is that the lossy line actually delivers noise power out of both its ports, but when the input port is mismatched some of the available noise power at port 1 is reflected from the source back into port 1 and appears at port 2. When the generator is matched to port 1, none of the available power from port 1 is reflected back into the line, so the noise power available at port 2 is a minimum. This result implies that impedance matching is important in minimizing noise temperature and noise figure.

EXAMPLE 10.3 APPLICATION TO A WILKINSON POWER DIVIDER

Find the noise figure of a Wilkinson power divider when one of the output ports is terminated in a matched load. Assume an insertion loss factor of L from the input to either output port.

Solution

From Chapter 7 the scattering matrix of a Wilkinson divider is given as

$$[S] = \frac{-j}{\sqrt{2L}} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

where the factor $L \geq 1$ accounts for the dissipative loss from port 1 to port 2 or 3 (note that dissipative loss is distinct from the -3 dB power division ratio). To evaluate the noise figure of the Wilkinson divider, we first terminate port 3 with a matched load; this converts the three-port device into a two-port device. If we assume a matched source at port 1, we have $\Gamma_s = 0$. Equation (10.26) then gives $\Gamma_{\text{out}} = S_{22} = 0$, and so the available gain can be calculated from (10.25) as

$$G_{21} = |S_{21}|^2 = \frac{1}{2L}.$$

The equivalent noise temperature of the Wilkinson divider is, from (10.28),

$$T_e = \frac{1 - G_{21}}{G_{21}} T = (2L - 1)T,$$

where T is the physical temperature of the divider. Using (10.11) gives the noise figure as

$$F = 1 + \frac{T_e}{T_0} = 1 + (2L - 1)\frac{T}{T_0}.$$

Observe that if the divider is at room temperature, then $T = T_0$ and the above reduces to $F = 2L$. If the divider is at room temperature and lossless, this reduces to $F = 2 = 3$ dB. In this case the source of the noise power is the isolation resistor contained in the Wilkinson divider circuit.

Because the network is matched at its input and output, it is easy to obtain these same results using the thermodynamic argument directly. Thus, if we apply an input noise power of kTB to port 1 of the matched divider at temperature T, the system will be in thermal equilibrium and the output noise power must be kTB. We can also express the output noise power as the sum of the input power times the gain of the divider, and N_{added}, the noise power added by the divider itself (referenced to the input to the divider):

$$kTB = \frac{kTB}{2L} + \frac{N_{added}}{2L}.$$

Solving for N_{added} gives $N_{added} = kTB(2L - 1)$, so the equivalent noise temperature is

$$T_e = \frac{N_{added}}{kB} = (2L - 1)T,$$

in agreement with the above. ■

Noise Figure of a Mismatched Amplifie

Finally, consider the effect of an input impedance mismatch on the noise figure of an amplifier. As shown in Figure 10.13, the amplifier, when matched, has a gain G, a noise figure F, and a bandwidth B. The amplifier output is matched, but there is an impedance mismatch at the input represented by the reflection coefficient, Γ. Our previous results involving the effect of mismatch on noise figure made use of (10.29), but that was derived for a passive network and so cannot be directly used in this case. Instead we will use noise temperatures.

Since we are dealing with noise figure, let the input noise power to the amplifier be $N_i = kT_0B$. Then the output noise power from the amplifier (referenced to the input) is given by

$$N_o = kT_0GB\left(1 - |\Gamma|^2\right) + kT_0\,(F - 1)\,GB \tag{10.34}$$

where the first term is due to the input noise power, decreased by the reflection at the input, and the second term is the noise power due to the amplifier itself, based on the equivalent noise temperature as given by (10.12). For an applied signal power S_i, the output signal

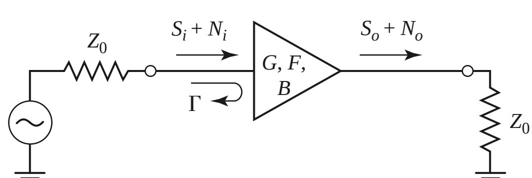


FIGURE 10.13 A noisy amplifier with an impedance mismatch at its input.

power is

$$S_o = G\left(1 - |\Gamma|^2\right)S_i. \tag{10.35}$$

The overall noise figure, F_m, of the mismatched amplifier can be found from (10.10) as

$$F_m = \frac{S_i N_o}{S_o N_i} = 1 + \frac{F - 1}{1 - |\Gamma|^2}. \tag{10.36}$$

Observe from (10.36) the limiting case that $F_m = F$ when $|\Gamma| = 0$ (no mismatch), and that this is the minimum noise figure that can be achieved since F_m increases as the mismatch increases. This result demonstrates that good noise figure requires good impedance matching. This problem would be more complicated if a mismatch also existed at the output of the amplifier, particularly if the amplifier is not unilateral.

10.3 NONLINEAR DISTORTION

We have seen that thermal noise is generated by any lossy component. Since all realistic components have at least a small loss, the ideal linear component does not exist in practice because all realistic devices are nonlinear at very low signal levels due to noise effects. In addition, practical components may also become nonlinear at high signal levels. In the case of active devices, such as diodes and transistors, this may be due to effects such as gain compression or the generation of spurious frequency components due to device nonlinearities, but all devices ultimately fail at very high power levels. In either case, these effects set a minimum and maximum realistic power range, or *dynamic range*, over which a given component or network will operate as desired. In this section we will study the response of nonlinear devices in general, and two definitions of dynamic range. These results will be useful for our later discussions of amplifiers (Chapter 12), mixers (Chapter 13), and wireless receivers (Chapter 14).

Devices such as diodes and transistors have nonlinear characteristics, and it is this nonlinearity that is of great utility for desirable functions such as amplification, detection, and frequency conversion [2]. Nonlinear device characteristics, however, can also lead to undesirable effects such as gain compression and the generation of spurious frequency components. These effects may lead to increased losses, signal distortion, and possible interference with other radio channels or services. Some of the many possible effects of nonlinearity in RF and microwave circuits are listed below [3]:

- Harmonic generation (multiples of a fundamental signal)
- Saturation (gain reduction in an amplifier)
- Intermodulation distortion (products of a two-tone input signal)
- Cross-modulation (modulation transfer from one signal to another)
- AM-PM conversion (amplitude variation causes phase shift)
- Spectral regrowth (intermodulation with many closely spaced signals)

Figure 10.14 shows a general nonlinear network, having an input voltage v_i and an output voltage v_o. In the most general sense, the output response of a nonlinear circuit can

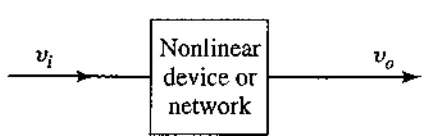


FIGURE 10.14 A general nonlinear device or network.

be modeled as a Taylor series in terms of the input signal voltage:

$$v_o = a_0 + a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + \cdots, \tag{10.37}$$

where the Taylor coefficients are defined as

$$a_0 = v_o(0) \qquad \text{(DC output)} \tag{10.38a}$$

$$a_1 = \left.\frac{dv_o}{dv_i}\right|_{v_i=0} \qquad \text{(linear output)} \tag{10.38b}$$

$$a_2 = \left.\frac{d^2 v_o}{dv_i^2}\right|_{v_i=0} \qquad \text{(squared output)} \tag{10.38c}$$

and higher order terms. Different functions can be obtained from the nonlinear network depending on the dominance of particular terms in the expansion. The constant term, with coefficient a_0, in (10.37) leads to rectification, converting an AC input signal to DC. The linear term, with coefficient a_1, models a linear attenuator ($a_1 < 1$) or amplifier ($a_1 > 1$). The second-order term, with coefficient a_2, can be used for mixing and other frequency conversion functions. Practical nonlinear devices usually have a series expansion containing many nonzero terms, and a combination of several of the above effects will occur. We will consider some important special cases below.

Gain Compression

First consider the case where a single-frequency sinusoid is applied to the input of a general nonlinear network, such as an amplifier:

$$v_i = V_0 \cos \omega_0 t. \tag{10.39}$$

Equation (10.37) gives the output voltage as

$$\begin{aligned} v_o &= a_0 + a_1 V_0 \cos \omega_0 t + a_2 V_0^2 \cos^2 \omega_0 t + a_3 V_0^3 \cos^3 \omega_0 t + \cdots \\ &= \left(a_0 + \frac{1}{2} a_2 V_0^2\right) + \left(a_1 V_0 + \frac{3}{4} a_3 V_0^3\right) \cos \omega_0 t + \frac{1}{2} a_2 V_0^2 \cos 2\omega_0 t \\ &\quad + \frac{1}{4} a_3 V_0^3 \cos 3\omega_0 t + \cdots. \end{aligned} \tag{10.40}$$

This result leads to the voltage gain of the signal component at frequency ω_0:

$$G_v = \frac{v_o^{(\omega_0)}}{v_i^{(\omega_0)}} = \frac{a_1 V_0 + \frac{3}{4} a_3 V_0^3}{V_0} = a_1 + \frac{3}{4} a_3 V_0^2, \tag{10.41}$$

where we have retained only terms through the third order.

The result of (10.41) shows that the voltage gain is equal to a_1, the coefficient of the linear term, as expected, but with an additional term proportional to the square of the input voltage amplitude. In most practical amplifiers a_3 typically has the opposite sign of a_1, so that the output of the amplifier tends to be reduced from the expected linear dependence for large values of V_0. This effect is called *gain compression*, or *saturation*. Physically, this is usually due to the fact that the instantaneous output voltage of an amplifier is limited by the power supply voltage used to bias the active device.

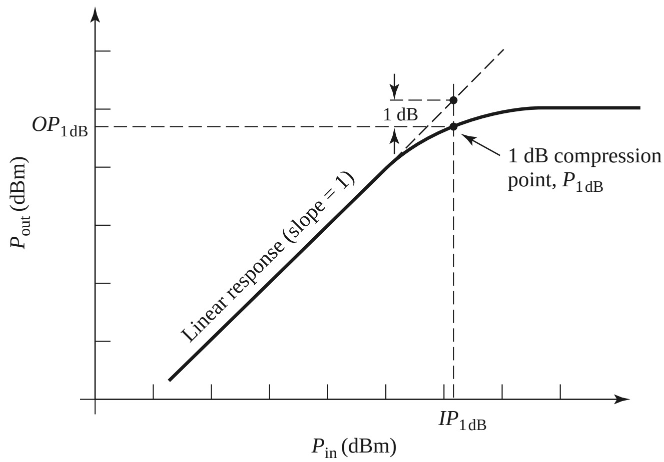


FIGURE 10.15 Definition of the 1 dB compression point for a nonlinear amplifier.

A typical amplifier response is shown in Figure 10.15. For an ideal linear amplifier a plot of the output power versus input power would be a straight line with a slope of unity, and the power gain of the amplifier given by the ratio of the output power to the input power. The amplifier response of Figure 10.15 tracks the ideal response over a limited range, then begins to saturate, resulting in reduced gain. To quantify the linear operating range of the amplifier, we define the *1 dB compression point* as the power level for which the output power has decreased by 1 dB from the ideal linear characteristic. This power level is usually denoted by $P_{1\text{dB}}$, and can be stated in terms of either input power ($IP_{1\text{dB}}$) or output power ($OP_{1\text{dB}}$). The 1 dB compression point is typically given as the larger of these two options, so for amplifiers $P_{1\text{dB}}$ is usually specified as an output power, while for mixers $P_{1\text{dB}}$ is usually specified in terms of input power. The relation between a compression point referenced at the input versus the output is given as, in dB, $OP_{1\text{dB}} = IP_{1\text{dB}} + G - 1$ dB [4, 5].

Harmonic and Intermodulation Distortion

Observe from the expansion of (10.40) that a portion of the input signal at frequency ω_0 is converted to other frequency components. For example, the first term of (10.40) represents a DC voltage, which would be a useful response in a rectifier application. The voltage components at frequencies $2\omega_0$ or $3\omega_0$ can be useful for frequency multiplier circuits. In amplifiers, however, the presence of other frequency components will lead to signal distortion if those components are in the passband of the amplifier.

For a single input frequency, or *tone*, ω_0, the output will in general consist of harmonics of the input frequency of the form $n\omega_0$, for $n = 0, 1, 2, \ldots$. Often these harmonics lie outside the passband of the amplifier and so do not interfere with the desired signal at frequency ω_0. The situation is different, however, when the input signal consists of two closely spaced frequencies.

Consider a *two-tone* input voltage, consisting of two closely spaced frequencies ω_1 and ω_2:

$$v_i = V_0(\cos \omega_1 t + \cos \omega_2 t). \tag{10.42}$$

From (10.37) the output is

$$
\begin{aligned}
v_o &= a_0 + a_1 V_0(\cos\omega_1 t + \cos\omega_2 t) + a_2 V_0^2(\cos\omega_1 t + \cos\omega_2 t)^2 \\
&\quad + a_3 V_0^3(\cos\omega_1 t + \cos\omega_2 t)^3 + \cdots \\
&= a_0 + a_1 V_0 \cos\omega_1 t + a_1 V_0 \cos\omega_2 t + \frac{1}{2} a_2 V_0^2 (1 + \cos 2\omega_1 t) + \frac{1}{2} a_2 V_0^2 (1 + \cos 2\omega_2 t) \\
&\quad + a_2 V_0^2 \cos(\omega_1 - \omega_2)t + a_2 V_0^2 \cos(\omega_1 + \omega_2)t \\
&\quad + a_3 V_0^3 \left(\frac{3}{4}\cos\omega_1 t + \frac{1}{4}\cos 3\omega_1 t\right) + a_3 V_0^3 \left(\frac{3}{4}\cos\omega_2 t + \frac{1}{4}\cos 3\omega_2 t\right) \\
&\quad + a_3 V_0^3 \left[\frac{3}{2}\cos\omega_2 t + \frac{3}{4}\cos(2\omega_1 - \omega_2)t + \frac{3}{4}\cos(2\omega_1 + \omega_2)t\right] \\
&\quad + a_3 V_0^3 \left[\frac{3}{2}\cos\omega_1 t + \frac{3}{4}\cos(2\omega_2 - \omega_1)t + \frac{3}{4}\cos(2\omega_2 + \omega_1)t\right] + \cdots. \qquad (10.43)
\end{aligned}
$$

where standard trigonometric identities have been used to expand the initial expression. We see that the output spectrum consists of harmonics of the form

$$m\omega_1 + n\omega_2, \qquad (10.44)$$

with $m, n = 0, \pm 1, \pm 2, \pm 3, \ldots$. These combinations of the two input frequencies are called *intermodulation products*, and the *order* of a given product is defined as $|m| + |n|$. For example, the squared term of (10.43) gives rise to the following four intermodulation products of second order:

$2\omega_1$	(second harmonic of ω_1)	$m = 2$	$n = 0$	order $= 2$,
$2\omega_2$	(second harmonic of ω_2)	$m = 0$	$n = 2$	order $= 2$,
$\omega_1 - \omega_2$	(difference frequency)	$m = 1$	$n = -1$	order $= 2$,
$\omega_1 + \omega_2$	(sum frequency)	$m = 1$	$n = 1$	order $= 2$.

All of these second-order products are undesired in an amplifier, but in a mixer the sum or difference frequencies form the desired outputs. In either case, if ω_1 and ω_2 are close, all of the second-order products will be far from ω_1 or ω_2 and can easily be filtered (either passed or rejected) from the output of the component. Note from (10.43) that the ratio of the amplitude of the second-order intermodulation product $\omega_1 - \omega_2$ (or $\omega_1 + \omega_2$) to the amplitude of a second harmonic $2\omega_1$ (or $2\omega_2$) is 2.0, so the second-order harmonic power will be 6 dB less than the power in the second-order sum or difference terms.

The cubed term of (10.43) leads to six third-order intermodulation products: $3\omega_1$, $3\omega_2$, $2\omega_1 + \omega_2$, $2\omega_2 + \omega_1$, $2\omega_1 - \omega_2$, and $2\omega_2 - \omega_1$. The first four of these will again be located far from ω_1 or ω_2, and will typically be outside the passband of the component. However, the two difference terms produce products located near the original input signals at ω_1 and ω_2, and so cannot be easily filtered from the passband of an amplifier. Figure 10.16 shows a typical spectrum of the second- and third-order two-tone intermodulation products. For an arbitrary input signal consisting of many frequencies of varying amplitude and phase, the resulting in-band intermodulation products will cause distortion of the output signal. This effect is called *third-order intermodulation distortion*.

It can be seen from (10.43) that the ratio of the amplitude of the third-order intermodulation product $2\omega_1 - \omega_2$ (or $2\omega_2 - \omega_1$) to the amplitude of the third harmonic $3\omega_1$ (or $3\omega_2$) is 3.0, so the third-order harmonic power will be 9.54 dB less than the power in the third-order intermodulation terms.

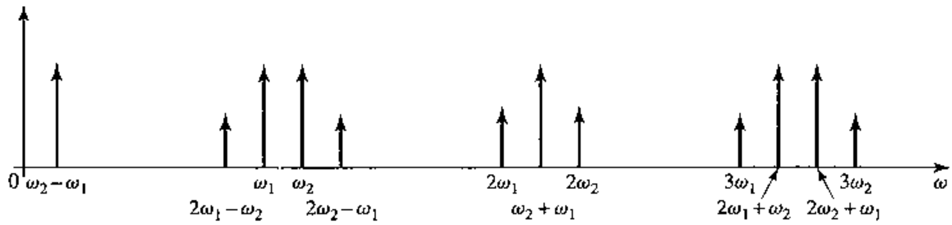


FIGURE 10.16 Output spectrum of second- and third-order two-tone intermodulation products, assuming $\omega_1 < \omega_2$.

Third-Order Intercept Point

Equation (10.43) shows that as the input voltage V_0 increases, the voltage associated with the third-order products increases as V_0^3. Since power is proportional to the square of voltage, we can also say that the output power of third-order products must increase as the cube of the input power. So for small input powers the third-order intermodulation products will be very small, but will increase quickly as input power increases. We can view this effect graphically by plotting the output power for the first- and third-order products versus input power on log-log scales (or in dB), as shown in Figure 10.17.

The output power of the first-order, or linear, product is proportional to the input power, and so the line describing this response has a slope of unity (before the onset of compression). The line describing the response of the third-order products has a slope of 3. (The second-order products would have a slope of 2, but since these products are generally not in the passband of the component, we have not plotted their response in Figure 10.17.) Both the linear and third-order responses will exhibit compression at high input powers, so we show the extension of their idealized responses with dotted lines. Since these two lines have different slopes, they will intersect, typically at a point above the onset of compression, as shown in the figure. This hypothetical intersection point where the first-order and third-order powers would be equal is called the *third-order intercept point*, denoted as IP_3;

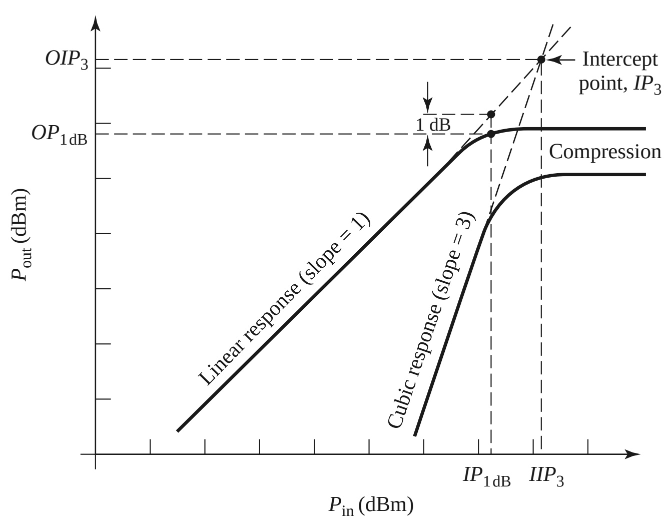


FIGURE 10.17 Third-order intercept diagram for a nonlinear component.

it may be specified as either an input power level (IIP_3), or an output power level (OIP_3). The relation between an intercept point referenced at the input versus the output is simply $OIP_3 = G\,(IIP_3)$. As with the 1 dB compression point, the reference for IP_3 is typically chosen to result in the largest value, so IP_3 is usually referenced at the output for amplifiers and at the input for mixers. As depicted in Figure 10.17, IP_3 generally occurs at a higher power level than $P_{1\text{dB}}$, the 1 dB compression point. Many practical components follow the approximate rule that IP_3 is 10–15 dB greater than $P_{1\text{dB}}$, assuming these powers are referenced at the same point.

We can express IP_3 in terms of the Taylor coefficients of the expansion of (10.43) as follows. Define P_{ω_1} as the output power of the desired signal at frequency ω_1. Then from (10.43) we have

$$P_{\omega_1} = \frac{1}{2} a_1^2 V_0^2. \tag{10.45}$$

Similarly, define $P_{2\omega_1-\omega_2}$ as the output power of the intermodulation product of frequency $2\omega_1 - \omega_2$. Then from (10.43) we have

$$P_{2\omega_1-\omega_2} = \frac{1}{2}\left(\frac{3}{4}a_3 V_0^3\right)^2 = \frac{9}{32} a_3^2 V_0^6. \tag{10.46}$$

By definition, these two powers are equal at the third-order intercept point. If we define the input signal voltage at the intercept point as V_{IP}, then equating (10.45) and (10.46) gives

$$\frac{1}{2} a_1^2 V_{IP}^2 = \frac{9}{32} a_3^2 V_{IP}^6.$$

Solving for V_{IP} yields

$$V_{IP} = \sqrt{\frac{4a_1}{3a_3}}. \tag{10.47}$$

Since OIP_3 is equal to the linear response of P_{ω_1} at the intercept point, we have from (10.45) and (10.47) that

$$OIP_3 = \left. P_{\omega_1}\right|_{V_0=V_{IP}} = \frac{1}{2} a_1^2 V_{IP}^2 = \frac{2a_1^3}{3a_3}, \tag{10.48}$$

where IP_3 in this case is referred to the output port. These expressions will be useful in the following sections.

Intercept Point of a Cascaded System

As in the case of noise figure, a cascade connection of components usually has the effect of degrading (lowering) the third-order intercept point. Unlike noise powers, however, intermodulation products in a cascaded system are deterministic and may be in phase coherence, in which case we cannot simply add powers but must deal with voltages [5]. We will first consider the coherent (in-phase) cascade case, then the noncoherent case.

With reference to Figure 10.18, G_1 and OIP_3' are the power gain and third-order intercept point for the first stage, and G_2 and OIP_3'' are the corresponding values for the second stage. Let P_{ω_1}' be the first-stage output power of the desired signal at frequency ω_1, and let $P_{2\omega_1-\omega_2}'$ be the first-stage output power at the third-order intermodulation product. From (10.46), $P_{2\omega_1-\omega_2}'$ can be rewritten in terms of P_{ω_1}' and OIP_3', using (10.45) and (10.48), as

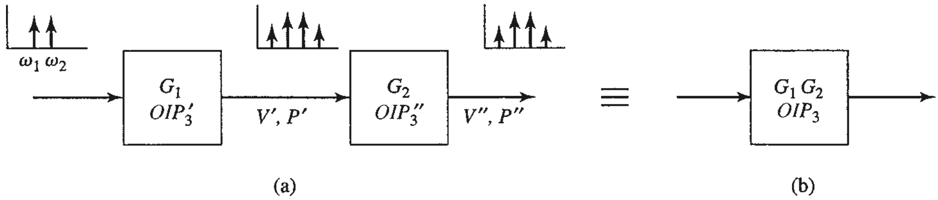


FIGURE 10.18 Third-order intercept point for a cascaded system. (a) Two cascaded networks. (b) Equivalent network.

follows:

$$P'_{2\omega_1-\omega_2} = \frac{9a_3^2 V_0^6}{32} = \frac{\frac{1}{8}a_1^6 V_0^6}{\frac{4a_1^6}{9a_3^2}} = \frac{(P'_{\omega_1})^3}{(OIP'_3)^2}. \tag{10.49}$$

The first-stage output voltage associated with this power is

$$V'_{2\omega_1-\omega_2} = \sqrt{P'_{2\omega_1-\omega_2} Z_0} = \frac{\sqrt{\left(P'_{\omega_1}\right)^3 Z_0}}{OIP'_3}, \tag{10.50}$$

where Z_0 is the system impedance.

For coherent intermodulation products, the total third-order distortion voltage at the output of the second stage is the sum of the above voltage times the voltage gain of the second stage, and the distortion voltage generated by the second stage. This is because these voltages are deterministic and phase related, unlike uncorrelated noise powers that arise in cascaded components. Adding these voltages gives the worst-case result for the overall distortion level because there may be phase delays within the stages that could cause partial cancellation. Thus we can write the worst-case total distortion voltage at the output of the second stage as

$$V''_{2\omega_1-\omega_2} = \frac{\sqrt{G_2\left(P'_{\omega_1}\right)^3 Z_0}}{OIP'_3} + \frac{\sqrt{\left(P''_{\omega_1}\right)^3 Z_0}}{OIP''_3}.$$

Since $P''_{\omega_1} = G_2 P'_{\omega_1}$, we have

$$V''_{2\omega_1-\omega_2} = \left(\frac{1}{G_2\left(OIP'_3\right)} + \frac{1}{OIP''_3}\right)\sqrt{\left(P''_{\omega_1}\right)^3 Z_0}. \tag{10.51}$$

The total output distortion power is

$$P''_{2\omega_1-\omega_2} = \frac{\left(V''_{2\omega_1-\omega_2}\right)^2}{Z_0} = \left(\frac{1}{G_2\left(OIP'_3\right)} + \frac{1}{OIP''_3}\right)^2 \left(P''_{\omega_1}\right)^3 = \frac{\left(P''_{\omega_1}\right)^3}{(OIP_3)^2}. \tag{10.52}$$

Thus the third-order intercept point of the cascaded system with coherent products is

$$OIP_3 = \left(\frac{1}{G_2\left(OIP'_3\right)} + \frac{1}{OIP''_3}\right)^{-1}. \tag{10.53}$$

Note that $OIP_3 = G_2\left(OIP'_3\right)$ for $OIP''_3 \to \infty$, which is the limiting case when the second stage has no third-order distortion.

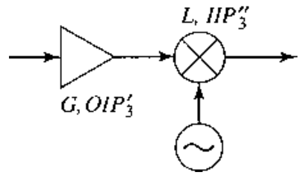


FIGURE 10.19 System for Example 10.4.

If the intermodulation products from each stage have relatively random phases, which may occur when the intermodulation products are not very close to the fundamental signals, it may be proper to treat the individual contributions as incoherent, allowing us to add powers. It is straightforward to show that the overall intercept point in this case is given by

$$OIP_3 = \left(\frac{1}{G_2^2 (OIP_3')^2} + \frac{1}{(OIP_3'')^2} \right)^{-1/2}. \tag{10.54}$$

EXAMPLE 10.4 CALCULATION OF CASCADE INTERCEPT POINT

A low-noise amplifier and mixer are shown in Figure 10.19. The amplifier has a gain of 20 dB and a third-order intercept point of 22 dBm (referenced at output), and the mixer has a conversion loss of 6 dB and a third-order intercept point of 13 dBm (referenced at input). Find the intercept points of the cascade network for both a phase coherence assumption and a random-phase (noncoherence) assumption.

Solution

First we transfer the reference of IP_3 for the mixer from its input to its output:

$$OIP_3'' = (IIP_3'')\, G_2 = 13 \text{ dBm} - 6 \text{ dB} = 7 \text{ dBm}.$$

Converting the necessary dB values to numerical values yields:

$$\begin{aligned} OIP_3' &= 22 \text{ dBm} = 158 \text{ mW} && \text{(for amplifier)}, \\ OIP_3'' &= 7 \text{ dBm} = 5 \text{ mW} && \text{(for mixer)}, \\ G_2 &= -6 \text{ dB} = 0.25 && \text{(for mixer)}. \end{aligned}$$

Assuming coherence, equation (10.53) gives the intercept point of the cascade as

$$OIP_3 = \left(\frac{1}{G_2\,(OIP_3')} + \frac{1}{OIP_3''} \right)^{-1} = \left(\frac{1}{(0.25)(158)} + \frac{1}{5} \right)^{-1}$$

$$= 4.4 \text{ mW} = 6.4 \text{ dBm},$$

which is seen to be lower than the minimum IP_3 of the individual components. Equation (10.54) gives the results for the noncoherent case as

$$OIP_3 = \left(\frac{1}{G_2^2\,(OIP_3')^2} + \frac{1}{(OIP_3'')^2} \right)^{-1/2} = \left(\frac{1}{(0.25)^2\,(158)^2} + \frac{1}{(5)^2} \right)^{-1/2}$$

$$= 4.96 \text{ mW} = 6.9 \text{ dBm}.$$

As expected, the noncoherent case results in a slightly higher intercept point. ■

Passive Intermodulation

The above discussion of intermodulation distortion was in the context of active circuits involving diodes and transistors, but it is also possible for intermodulation products to be generated by passive nonlinear effects in connectors, cables, antennas, or almost any component where there is a metal-to-metal contact. This effect is called *passive intermodulation* (PIM) and, as in the case of intermodulation in amplifiers and mixers, it occurs when signals at two or more closely spaced frequencies mix to produce spurious products.

Passive intermodulation can be caused by a number of factors, such as poor mechanical contact, oxidation of junctions between ferrous-based metals, contamination of conducting surfaces at RF junctions, or the use of nonlinear materials such as carbon fiber composites or ferromagnetic materials. In addition, when high powers are involved, thermal effects may contribute to the overall nonlinearity of a junction. It is very difficult to predict PIM levels from first principles, so measurement techniques must usually be used.

Because of the third-power dependence of the third-order intermodulation products with input power, passive intermodulation is usually only significant when input signal powers are relatively large. This is frequently the case in cellular telephone base station transmitters, which may operate with powers of 30–40 dBm, with many closely spaced RF channels. It is often desired to maintain the PIM level below −125 dBm, with two 40 dBm transmit signals. This is a very wide dynamic range, and requires careful selection of components used in the high-power portions of the transmitter, including cables, connectors, and antenna components. Because these components are often exposed to the weather, deterioration due to oxidation, vibration, and sunlight must be offset by a careful maintenance program. Communications satellites often face similar problems with passive intermodulation. Passive intermodulation is generally not a problem in receiver systems due to the much lower power levels.

10.4 DYNAMIC RANGE

Linear and Spurious Free Dynamic Range

We can define *dynamic range* in a general sense as the operating range for which a component or system has desirable characteristics. For a power amplifier this may be the power range that is limited at the low end by noise and at the high end by the compression point. This is essentially the linear operating range for the amplifier, and is called the *linear dynamic range* (LDR). For low-noise amplifiers or mixers, operation may be limited by noise at the low end and the maximum power level for which intermodulation distortion becomes unacceptable. This is effectively the operating range for which spurious responses are minimal, and it is called the *spurious-free dynamic range* (SFDR).

We can find the linear dynamic range LDR as the ratio of $P_{1\text{dB}}$, the 1 dB compression point, to the noise level of the component, as shown in Figure 10.20. In dB, this can be written in terms of output powers as

$$\text{LDR (dB)} = OP_{1\text{dB}} - N_o, \tag{10.55}$$

for $OP_{1\text{dB}}$ and N_o expressed in dBm. Note that some authors prefer to define the linear dynamic range in terms of a minimum detectable power level. This definition is more appropriate for a receiver system rather than an individual component, as it depends on factors external to the component itself, such as the type of modulation used, the recommended system SNR, effects of error-correcting coding, and related factors.

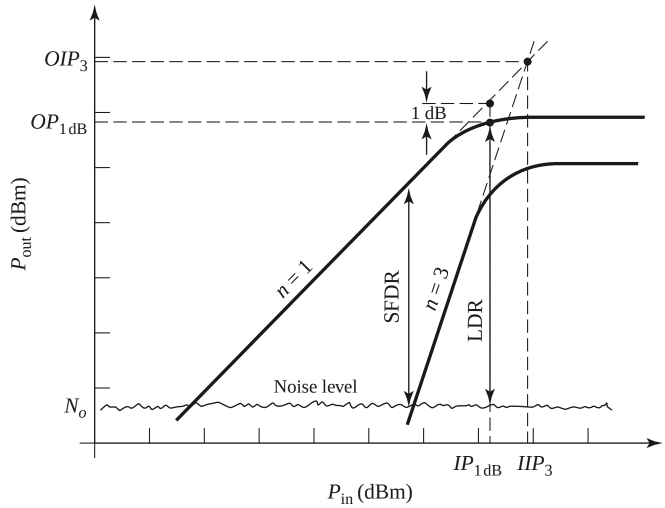


FIGURE 10.20 Illustrating linear dynamic range (LDR) and spurious free dynamic range (SFDR).

The spurious free dynamic range is defined as the maximum output signal power for which the power of the third-order intermodulation product is equal to the noise level of the component, divided by the output noise level. This situation is shown in Figure 10.20. If P_{ω_1} is the output power of the desired signal at frequency ω_1, and $P_{2\omega_1-\omega_2}$ is the output power of the third-order intermodulation product, then the spurious free dynamic range can be expressed as

$$\text{SFDR} = \frac{P_{\omega_1}}{P_{2\omega_1-\omega_2}}, \tag{10.56}$$

with $P_{2\omega_1-\omega_2}$ taken equal to the noise level of the component. As in (10.49), $P_{2\omega_1-\omega_2}$ can be written in terms of OIP_3 and P_{ω_1} as

$$P_{2\omega_1-\omega_2} = \frac{(P_{\omega_1})^3}{(OIP_3)^2}. \tag{10.57}$$

Observe that this result clearly shows that the third-order intermodulation power increases as the cube of the input signal power. Solving (10.57) for P_{ω_1} and applying the result to (10.56) gives the spurious free dynamic range in terms of OIP_3 and N_o, the output noise power of the component:

$$\text{SFDR} = \left.\frac{P_{\omega_1}}{P_{2\omega_1-\omega_2}}\right|_{P_{2\omega_1-\omega_2}=N_o} = \left(\frac{OIP_3}{N_o}\right)^{2/3}. \tag{10.58}$$

This result can be written in terms of dB as

$$\text{SFDR (dB)} = \frac{2}{3}(OIP_3 - N_o), \tag{10.59}$$

for OIP_3 and N_o expressed in dBm. Although this result was derived for the $2\omega_1 - \omega_2$ product, the same result applies for the $2\omega_2 - \omega_1$ product.

In a receiver it may be required to have a minimum detectable signal level, or minimum SNR, in order to achieve a specified performance level. This requires an increase in the input signal level, resulting in a corresponding decrease in dynamic range, since the spurious power level is still equal to the noise power. In this case, the spurious free dynamic range of (10.59) would be modified as [5, 6]:

$$\text{SFDR (dB)} = \frac{2}{3}(OIP_3 - N_o) - \text{SNR}. \tag{10.60}$$

EXAMPLE 10.5 DYNAMIC RANGES

A receiver has a noise figure of 7 dB, a 1 dB compression point of 25 dBm (referenced to output), a gain of 40 dB, and a third-order intercept point of 35 dBm (referenced to output). If the receiver is fed with an antenna having a noise temperature of $T_A = 150$ K, and the desired output SNR is 10 dB, find the linear and spurious free dynamic ranges. Assume a receiver bandwidth of 100 MHz.

Solution

The noise power at the receiver output can be calculated using noise temperatures as

$$N_o = GkB[T_A + (F-1)T_0] = 10^4(1.38 \times 10^{-23})(10^8)[150 + (4.01)(290)]$$
$$= 1.8 \times 10^{-8} \text{ W} = -47.4 \text{ dBm}.$$

The linear dynamic range is, from (10.55), in dB,

$$\text{LDR} = OP_{1\text{dB}} - N_o = 25 \text{ dBm} + 47.4 \text{ dBm} = 72.4 \text{ dB}.$$

Equation (10.60) gives the spurious free dynamic range as

$$\text{SFDR} = \frac{2}{3}(OIP_3 - N_o) - \text{SNR} = \frac{2}{3}(35 + 47.4) - 10 = 44.9 \text{ dB}.$$

Observe that SFDR $\ll$ LDR. ■

REFERENCES

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Volume I, Microwave Remote Sensing, Fundamentals and Radiometry*. Addison-Wesley, Reading, Mass., 1981.

[2] M. E. Hines, “The Virtues of Nonlinearity—Detection, Frequency Conversion, Parametric Amplification and Harmonic Generation,” *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-32, pp. 1097–1104, September 1984.

[3] S. A. Maas, *Nonlinear Microwave and RF Circuits*, 2nd ed., Artech House, Norwood, Mass., 2003.

[4] K. Chang, *RF and Microwave Wireless Systems*, John Wiley & Sons, New York, 2000.

[5] W. Egan, *Practical RF System Design*, John Wiley & Sons, Hoboken, N.J., 2003.

[6] M. Steer, *Microwave and RF Design: A Systems Approach*, SciTech, Raleigh, N.C., 2010.

PROBLEMS

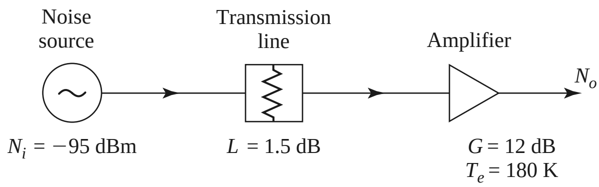
10.1 The noise figure of a microwave receiver front-end is measured using the Y-factor method. A noise source having an ENR of 22 dB, and a liquid nitrogen cold load (77 K) are used, resulting in a measured Y-factor ratio of 15.83 dB. What is the noise figure of the receiver?

10.2 Assume that measurement error introduces an uncertainty of ΔY into the measurement of Y in a Y-factor measurement. Derive an expression for the normalized error, $\Delta T_e/T_e$, of the equivalent noise temperature in terms of $\Delta Y/Y$ and the temperatures T_1, T_2, and T_e. Minimize this result with respect to T_e to obtain an expression for T_e in terms of T_1 and T_2 that will result in minimum error.

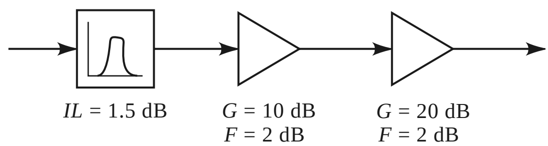
10.3 A lossy transmission line has a noise figure of F_0 at temperature $T_0 = 290$ K. Calculate and plot the noise figure of this line as its physical temperature ranges from $T = 0$ K to 1000 K, for $F_0 = 1$ dB and for $F_0 = 3$ dB.

10.4 An amplifier with a gain of 12 dB, a bandwidth of 150 MHz, and a noise figure of 4 dB feeds a receiver with a noise temperature of 900 K. Find the noise figure of the overall system.

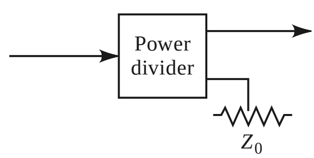
10.5 A cellular telephone receiver front-end circuit is shown below. The operating frequency is 1805–1880 MHz, and the physical temperature of the system is 300 K. A noise source with $N_i = -95$ dBm is applied to the receiver input. (a) What is the equivalent noise temperature of the source over the operating bandwidth? (b) What is the noise figure (in dB) of the amplifier? (c) What is the noise figure (in dB) of the cascaded transmission line and amplifier? (d) What is the total noise power output (in dBm) of the receiver over the operating bandwidth?



10.6 Consider the wireless local area network (WLAN) receiver front-end shown below, where the bandwidth of the bandpass filter is 100 MHz centered at 2.4 GHz. If the system is at room temperature, find the noise figure of the overall system. What is the resulting signal-to-noise ratio at the output if the input signal power level is −90 dBm? Can the components be rearranged to give a better noise figure?



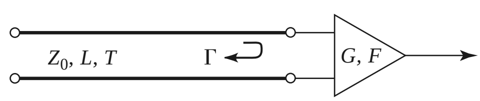
10.7 A two-way power divider has one output port terminated in a matched load, as shown below. Find the noise figure of the resulting two-port network if the divider is (a) an equal-split two-way resistive divider, (b) a two-way Wilkinson divider, and (c) a 3 dB quadrature hybrid. Assume the divider in each case is matched, and at room temperature.



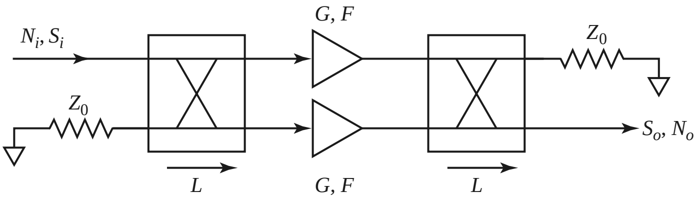
10.8 Show that, for fixed loss $L > 1$, the equivalent noise temperature of a mismatched lossy line given in (10.33) is minimized when $|\Gamma_S| = 0$.

10.9 Consider the mismatched amplifier of Figure 10.13, having a noise figure F when matched at its input. Calculate and plot the resulting noise figure as the input reflection coefficient magnitude, $|\Gamma|$, varies from 0 to 0.9 for $F = 1$, 3, and 10 dB.

10.10 A lossy line at temperature T feeds an amplifier with noise figure F, as shown below. If an impedance mismatch Γ is present at the input of the amplifier, find the overall noise figure of the system.



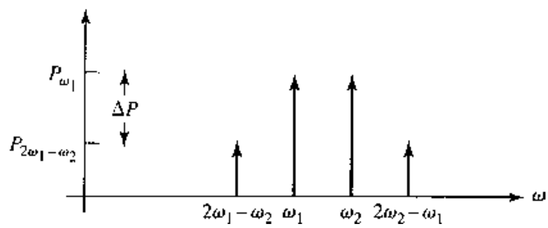
10.11 A balanced amplifier circuit is shown below. The two amplifiers are identical, each with power gain G and noise figure F. The two quadrature hybrids are also identical, with an insertion loss from the input to either output of $L > 1$ (not including the 3 dB power division factor). Derive an expression for the overall noise figure of the balanced amplifier. What does this result reduce to when the hybrids are lossless?



10.12 Show that the following relations involving the third-order intercept point of a two-port nonlinear network are valid. $P^i_{\omega_1}$ and $P^o_{\omega_1}$ are the input and output power levels of an applied two-tone signal, and $P^i_{2\omega_1-\omega_2}$ and $P^o_{2\omega_1-\omega_2}$ are the power levels of the third-order products referenced to the input and output.

$$\frac{OIP_3 - P^o_{\omega_1}}{IIP_3 - P^i_{\omega_1}} = 1, \qquad \frac{OIP_3 - P^o_{2\omega_1-\omega_2}}{IIP_3 - P^i_{2\omega_1-\omega_2}} = 3.$$

10.13 In practice, the third-order intercept point is extrapolated from measured data taken at input power levels well below IP_3. For the spectrum analyzer display shown below, where ΔP is the difference in power between P_{ω_1} and $P_{2\omega_1-\omega_2}$, show that the third-order intercept point is given by $OIP_3 = P_{\omega_1} + (1/2)\Delta P$. Calculate the input and output third-order intercept points for the following data: $P_{\omega_1} = 5$ dBm, $P_{2\omega_1-\omega_2} = -27$ dBm, $P_{\text{in}} = -4$ dBm.



10.14 A two-tone input with a 6 dB difference in the two signal levels is applied to a nonlinear component. What is the relative power ratio of the resulting two third-order intermodulation products $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$, if ω_1 and ω_2 are close together?

10.15 Find the third-order intercept points for the problem of Example 10.4 when the positions of the amplifier and mixer are reversed.

10.16 It is possible to approximately relate the 1 dB compression point to the third-order intercept point. For a single-tone input, use (10.40) to find the amplitudes of the fundamental and third harmonic terms, and assume that a_3 is of opposite sign to a_1. Let V_0 be the voltage where the third-order term reduces the first-order power by 1 dB, and solve for $|a_3/a_1|$. For a two-tone input, use (10.43) to find the amplitude of the third-order intermodulation product, then use (10.44) to relate $OP_{1\text{dB}}$ to OIP_3.

10.17 An amplifier with a bandwidth of 1 GHz has a gain of 15 dB and a noise temperature of 250 K. If the 1 dB compression point occurs for an output power level of 5 dBm, what is the linear dynamic range of the amplifier?

10.18 A receiver subsystem has a noise figure of 6 dB, a 1 dB compression point of 21 dBm (referenced to output), a gain of 30 dB, and a third-order intercept point of 33 dBm (referenced to output). If the subsystem is fed with a noise source with $N_i = -105$ dBm and the desired output SNR is 8 dB, find the linear and spurious free dynamic ranges of the subsystem. Assume a system bandwidth of 20 MHz.