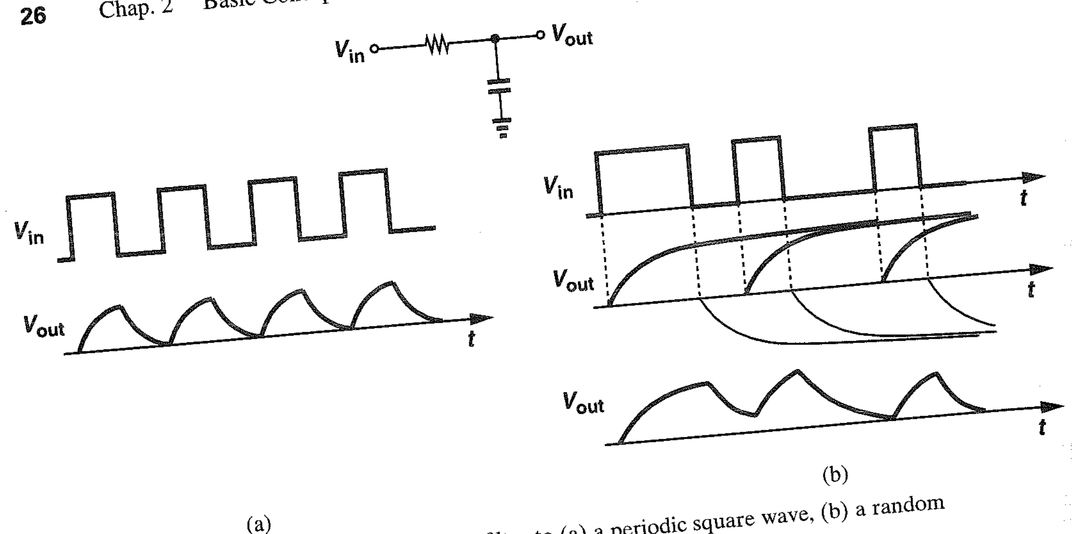

Figure 2.13 Response of a low-pass filter to (a) a periodic square wave, (b) a random sequence of ONEs and ZEROs.

Now suppose the output of a digital system consists of a random sequence of ONEs and ZEROs, each ONE represented by an ideal rectangular pulse and each ZERO by the absence of such a pulse. If this sequence is applied to a low-pass filter, the output can be obtained as the superposition of the responses to each input bit [Fig. 2.13(b)]. We note that each bit level is corrupted by decaying tails created by previous bits. Called "intersymbol interference" (ISI), this phenomenon leads to higher error rate in the detection of random waveforms that are transmitted through band-limited channels.

The problem of ISI is particularly troublesome in wireless communications because the bandwidth allocated to each channel is fairly narrow. Methods of reducing ISI include pulse shaping ("Nyquist signaling") in the transmitter and "equalization" in the receiver. We briefly describe Nyquist signaling here and refer the reader to the extensive literature on equalization for ISI mitigation [3, 4].

In order to reduce ISI, the pulse shape can be chosen such that it is less susceptible to interference with its shifted replicas. In Nyquist signaling, each pulse is allowed to overlap with past and future pulses, but the shape is selected such that ISI is zero at certain points in time. Illustrated in Fig. 2.14, the idea is that all other pulses go through zero at the point when the present pulse reaches its peak. Thus, if the bit stream is sampled at $t = kT_S$, no ISI exists.

A simple calculation leads to a basic condition for Nyquist signals. For a pulse shape, $p(t)$, to introduce zero ISI, we have

$$p(kT_S) = 1 \quad \text{if } k = 1 \qquad (2.48)$$

$$= 0 \quad \text{if } k \neq 0 \qquad (2.49)$$

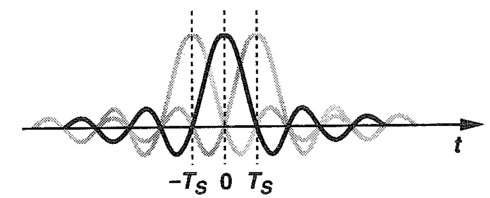

Figure 2.14 Pulse shape with no ISI.

Using a train of impulses to sample this pulse, we obtain

$$p(t) \cdot \sum \delta(t - kT_S) = \delta(t). \qquad (2.50)$$

Taking the Fourier transform of both sides, we have

$$P(f) * \frac{1}{T_S} \sum \delta\left(f - \frac{k}{T_S}\right) = 1. \qquad (2.51)$$

That is,

$$\frac{1}{T_S} \sum P\left(f - \frac{k}{T_S}\right) = 1. \qquad (2.52)$$

Originally proposed by Nyquist and shown in Fig. 2.15, this result indicates that the shifted replicas of $P(f)$ must add up to a flat spectrum. For example, a sinc waveform satisfies this condition because its Fourier transform is a rectangular box.

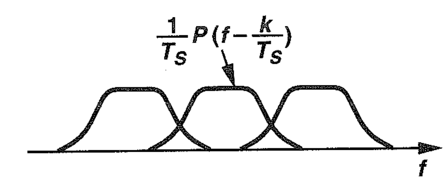

Figure 2.15 Nyquist's condition for the spectrum of a pulse shape that gives no ISI.

A sinc pulse shape, however, introduces difficulties in the design of the system. The filter required to produce the rectangular spectrum becomes quite complex if a sharp cutoff is necessary. Furthermore, the substantial signal energy near the edge of the spectrum complicates the filtering requirements in both the transmit and receive paths. In addition, the sinc waveform decays slowly with time, introducing considerable ISI in the presence of timing errors in the sampling command.

A pulse shape often employed in Nyquist signaling is related to a "raised cosine" spectrum. Shown in Fig. 2.16, the time- and frequency-domain expressions of this function are, respectively,

$$p(t) = \frac{\sin(\pi t/T_S)}{\pi t/T_S} \frac{\cos(\pi \alpha t/T_S)}{1 - 4\alpha^2 t^2/T_S^2} \qquad (2.53)$$

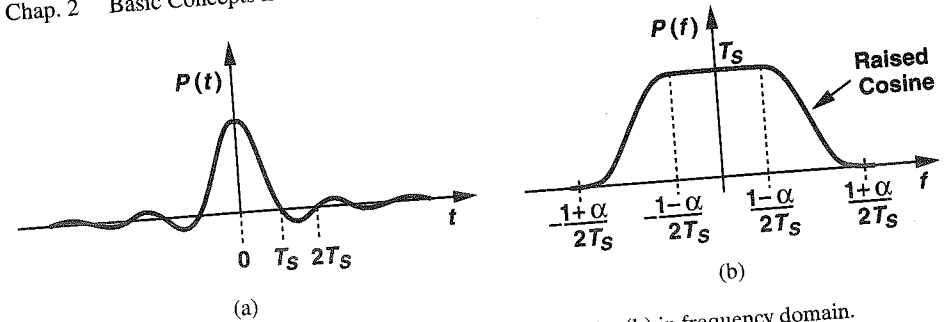

Figure 2.16 Raised-cosine pulse: (a) in time domain, (b) in frequency domain.

and

$$
\begin{aligned}
P(f) &= T_S \quad 0 < |f| < \frac{1-\alpha}{2T_S} \\
&= \frac{T_S}{2}\left[1 + \cos\frac{\pi T_S}{\alpha}\left(|f| - \frac{1-\alpha}{2T_S}\right)\right] \quad \frac{1-\alpha}{2T_S} < |f| < \frac{1+\alpha}{2T_S} \\
&= 0 \quad |f| > \frac{1+\alpha}{2T_S},
\end{aligned}
\tag{2.54}
$$

where $0 < \alpha < 1$ is the "rolloff" factor. It is interesting to note that (1) $p(t)$ decays faster than a sinc function, (2) for $\alpha = 0$, $p(t)$ reduces to a sinc function, and (3) $P(f)$ is similar to a box spectrum but with smooth edges.

The trade-off in the choice of α is between the decay rate in the time domain and the excess bandwidth (with respect to a box spectrum) in the frequency domain. Typical values of α are between 0.3 and 0.5.

Raised-cosine signaling can also be visualized as shown in Fig. 2.17. Here, each bit is represented by an impulse, and the data stream is applied to a filter whose transfer function is given by (2.54). From this point of view, the operation is called "raised-cosine filtering."

As explained in Chapter 3, many applications incorporate a filter whose transfer function is equal to the *square root* of that in (2.54).

2.3 RANDOM PROCESSES AND NOISE

Random processes are an integral part of communications, used to represent both signals and noise. In this section, we provide a brief review of random processes and noise to the extent required for and in a language suited to RF design. The primary goal is to develop an intuitive understanding of these phenomena and the relationships governing their behavior. The reader is assumed to be familiar with concepts such as random variables and probability density functions (PDFs).

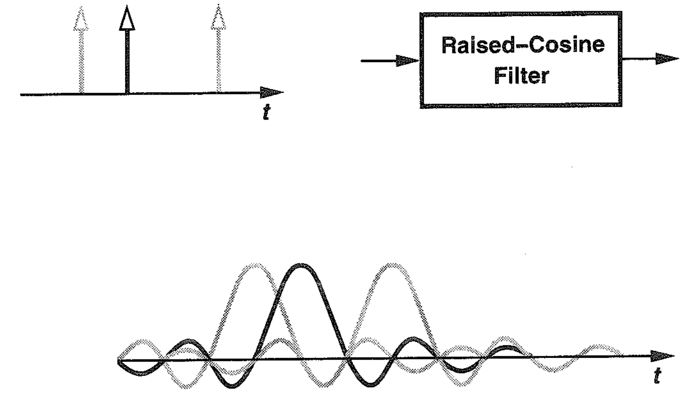

Figure 2.17 Raised-cosine filtering.

2.3.1 Random Processes

The trouble with random processes is that they are random! Engineers who are used to dealing with well-defined, deterministic, "hard facts" often find the concept of randomness difficult to grasp, especially if it must be incorporated mathematically. To overcome this fear of randomness, we approach the problem from an intuitive angle.

We consider a phenomenon random because we do not know or simply do not *need* to know everything about it. We characterize the process with only a few parameters and functions and solve most problems without any other information about the process. Experience shows that this approach is feasible and adequate in many applications, including RF design. In other words, we are fortunate that most random processes encountered in RF design lend themselves to relatively simple modeling.

For our purposes, a random (actually a "stochastic") process can be defined as "a family of time functions." If we measure the noise voltage across a resistor as a function of time today, the waveform is different from that measured tomorrow[2] (Fig. 2.18). To know everything about the noise voltage, we would need to perform an infinite number of measurements, each one for an infinite length of time. Since a single waveform measurement in general does not provide adequate knowledge of the process, even simplest random processes extend in two dimensions; i.e., they require a *collection* of measurements, hence the phrase "family of time functions." This is the principal difference between random and deterministic signals—and the primary source of confusion. In using an ordinary signal generator, we always consider the output a single predictable and well-defined waveform (except perhaps for the phase at the power-up time, which is usually unimportant). With a random signal, e.g., the

[2] This should not be confused with time variance in a system.

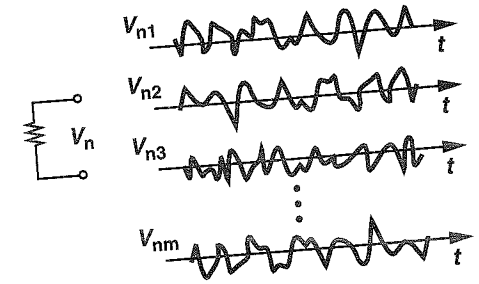

Figure 2.18 Noise viewed as a family of time functions.

voice going through a phone line, we do not have this luxury and must resort to statistics obtained from multiple measurements.

We must emphasize that the reason why the theory of random signals is useful and practical is that in most cases of interest such signals can be modeled with simple statistical functions that indicate, among other things, *how much* and *how fast* the amplitude varies with time. Furthermore, the statistical models can be used in conjunction with the familiar theory of deterministic signals and systems, often allowing us to momentarily forget about randomness and utilize more intuitive analysis techniques.

How is a random process characterized? What aspects of its statistics are important? How are these aspects incorporated in system analysis? We answer these questions by first making some simplifying assumptions.

Statistical Ensembles As mentioned above, full characterization of a (continuous-time) random signal, e.g., the noise voltage across a resistor, requires a "doubly infinite" set: an infinite number of measurements, each for an infinite length of time (Fig. 2.18). Now suppose, rather than one resistor, we consider a very large number of identical resistors and measure their noise voltages simultaneously (still for a very long time). We would expect these two experiments to yield the same statistical results. The large set of resistor noise voltages is called an "ensemble," and each of the waveforms is called a "sample function."

How do we measure the average value of the noise voltage of a resistor? Our familiar approach is to measure the noise, $n(t)$, for a long time, T, and calculate the average (or dc component) as

$$< n(t) >= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} n(t)dt. \tag{2.55}$$

This notion of dc component of a random signal is called the "time average."

Another definition of the average value is based on simultaneous sampling of all the waveforms in an ensemble (Fig. 2.19). Here, we compute the average by adding the sampled values and normalizing the sum to the number

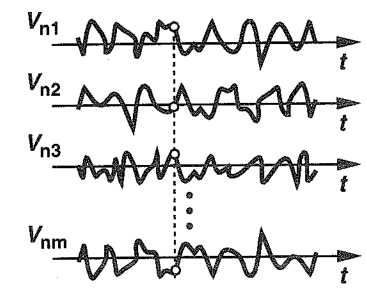

Figure 2.19 Averaging over sample functions.

of waveforms. Called the "ensemble average" or "statistical average," this type is formally defined as

$$\overline{n(t)} = \int_{-\infty}^{+\infty} n(t) P_n(n) dn, \tag{2.56}$$

where $P_n(n)$ is the probability density function of the process.

From the above definitions arise two questions. First, is the time average measured today equal to that measured tomorrow? Not necessarily. A process whose statistical properties are invariant to a time shift is called "stationary" (more accurately "strict-sense stationary"). Thus, the concept of time average is useful for stationary processes, e.g., noise voltage of a resistor held at a constant temperature. Fortunately, most of the random phenomena in RF systems can be considered stationary.

The second question is: Is the time average of a stationary process equal to the ensemble average? Not always, but for most random processes of interest in this book, we can assume so, thus avoiding ensemble averages.

The time and ensemble averages defined above are of first order. Higher-order averages can also be defined. Of particular interest are second-order averages, for they represent the power of signals. In the time domain,

$$< n^2(t) >= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} n^2(t)dt, \tag{2.57}$$

which is also called the "mean square" power (with respect to a 1-Ω resistor) if $n(t)$ is a voltage quantity. The second-order ensemble average is

$$\overline{n^2(t)} = \int_{-\infty}^{+\infty} n^2(t) P_n(n) dn. \tag{2.58}$$

For our purposes, $< n^2(t) > = \overline{n^2(t)}$.

Probability Density Function When considering a random signal in the time domain, we usually need to know how often its amplitude is between certain limits. For example, if a binary data sequence is corrupted by additive noise (Fig. 2.20), it is important to find the probability that a logical ONE is interpreted as a ZERO and vice versa, that is, how often the noise amplitude exceeds half of the signal amplitude. The amplitude statistics of a random signal $x(t)$ is characterized by the probability density function, $P_x(x)$, defined as

$$P_x(x)dx = \text{probability of } x < X < x + dx, \tag{2.59}$$

where X is the measured value of $x(t)$ at some point in time. To estimate the PDF, we sample $x(t)$ at many points (for many functions in the ensemble), construct bins of small width, choose the bin height equal to the number of samples whose value falls between the two edges of the bin, and normalize the bin heights to the total number of samples. Note that the PDF provides no information as to *how fast* the random signal varies in the time domain.

Figure 2.20 Binary signal corrupted by noise.

An important example of PDFs is the Gaussian (or normal) distribution. The central limit theorem states that if many independent random processes with arbitrary PDFs are added, the PDF of the sum approaches a Gaussian distribution. It is therefore not surprising that many natural phenomena exhibit Gaussian statistics. For example, since the noise of a resistor results from random "walk" of a very large number of electrons, each having relatively independent statistics, the overall amplitude follows a Gaussian PDF.

The Gaussian PDF is defined as

$$P_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\frac{-(x-m)^2}{2\sigma^2}, \tag{2.60}$$

where σ and m are the standard deviation and mean of the distribution, respectively.

From the PDF of the amplitude of a random signal, we can also answer the following question: If a large number of samples are taken, what percentage will fall between x_1 and x_2? This is given by the area under $P_x(x)$ from x_1 to x_2, and for a Gaussian PDF:

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} \exp\frac{-(x-m)^2}{2\sigma^2} dx. \tag{2.61}$$

For finite x_1 and x_2, the integral on the right-hand side must be calculated numerically. A simpler version of this integral, called the error function, is tabulated in many references:

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\frac{-u^2}{2} du. \tag{2.62}$$

It is useful to remember that for a Gaussian distribution approximately 68% of the sampled values fall between $m - \sigma$ and $m + \sigma$ and 99% between $m - 3\sigma$ and $m + 3\sigma$.

Power Spectral Density Since our knowledge of random signals in the time domain is usually quite limited, it is often necessary to characterize such signals in the frequency domain as well. In fact, as we will see throughout this book, the frequency-domain behavior of random signals and noise proves much more useful in RF design than do their time-domain characteristics.

For a deterministic signal $x(t)$, the frequency information is embodied in the Fourier transform:

$$X(f) = \int_{-\infty}^{+\infty} x(t)\exp(-j2\pi ft)dt. \tag{2.63}$$

While it may seem natural to use the same definition for random signals, we must note that the Fourier transform exists only for signals with finite energy,[3]

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 dt < \infty, \tag{2.64}$$

i.e., only if $|x(t)|^2$ drops rapidly enough as $t \to \infty$. As shown in Fig. 2.21, this condition is violated by two classes of signals: periodic waveforms and random signals. In most cases, however, these waveforms have a finite power:

$$P = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 dt < \infty. \tag{2.65}$$

For periodic signals with $P < \infty$, the Fourier transform can still be defined by representing each component of the Fourier series with an impulse in the frequency domain. For random signals, on the other hand, this is generally not possible because a frequency impulse indicates the existence of a deterministic sinusoidal component. Another practical problem is that even if we somehow define a Fourier transform for a random (stationary or nonstationary) process, the result itself is also a random process [5].

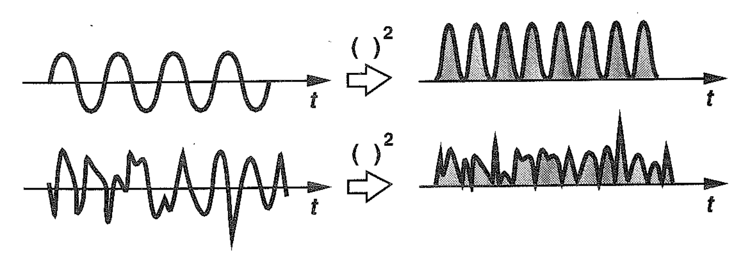

Figure 2.21 Signals with infinite energy.

[3] The definition of energy can be visualized if $x(t)$ is a voltage applied across a 1-Ω resistor.

From the above discussion we infer that frequency-domain characteristics of random signals are embodied in a function different from a direct Fourier transform. The power spectral density (PSD) (also called the "spectral density" or simply the "spectrum") is such a function. Before giving a formal definition of PSD, we describe its meaning from an intuitive point of view [6]. The spectral density, $S_x(f)$, of a random signal $x(t)$ shows how much power the signal carries in a unit bandwidth around frequency f. As illustrated in Fig. 2.22, if we apply the signal to a bandpass filter with a 1-Hz bandwidth centered at f and measure the average output power over a sufficiently long time (on the order of 1 s), we obtain an estimate of $S_x(f)$. If this measurement is performed for each value of f, the overall spectrum of the signal is obtained. This is in fact the principle of operation of spectrum analyzers.[4]

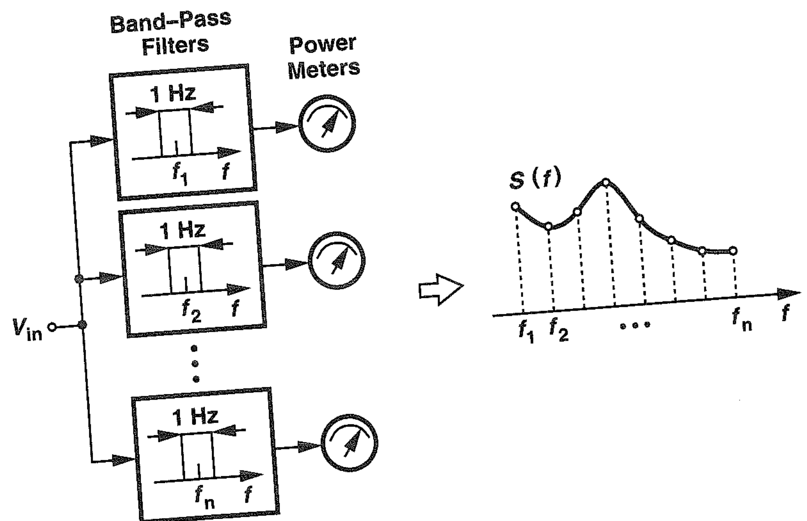

Figure 2.22 Measurement of spectrum.

The formal definition of the PSD is as follows [3]:

$$S_x(f) = \lim_{T\to\infty} \frac{\overline{|X_T(f)|^2}}{T}, \tag{2.66}$$

where

$$X_T(f) = \int_0^T x(t)\exp(-j2\pi ft)dt. \tag{2.67}$$

[4] Building a low-loss BPF with 1-Hz bandwidth and a center frequency of, say, 1 GHz is impractical. Thus, actual spectrum analyzers both translate the spectrum to a lower center frequency and measure the power in a band wider than 1 Hz.

The definition can be understood with the aid of a corresponding computational algorithm (Fig. 2.23): (1) truncate $x(t)$ to a relatively long interval $[0, T]$, (2) calculate the Fourier transform of the result and hence $|X_T(f)|^2$, (3) repeat steps 1 and 2 for many sample functions of $x(t)$ (e.g., for many noise voltage waveforms measured across a resistor), and (4) take the average of all $|X_T(f)|^2$ functions to arrive at $\overline{|X_T(f)|^2}$ and normalize the result to T. This algorithm proves useful in time-domain simulations incorporating random noise waveforms.

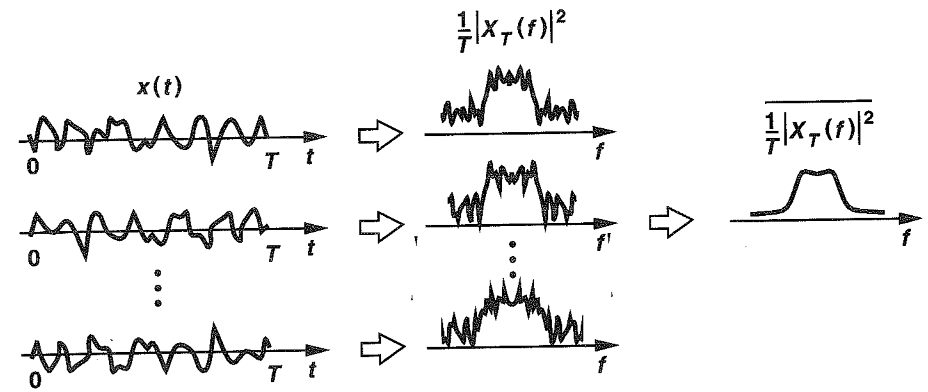

Figure 2.23 Algorithm for PSD estimation.

Since $S_x(f)$ is an even function of f for real $x(t)$ [3], as depicted in Fig. 2.24(a) the total power carried by $x(t)$ in the frequency range $[f_1\ f_2]$ is equal to

$$\int_{-f_2}^{-f_1} S_x(f)df + \int_{f_1}^{f_2} S_x(f)df = \int_{f_1}^{f_2} 2S_x(f)df. \tag{2.68}$$

In fact, the right-hand side integral is the quantity measured by a spectrum analyzer; i.e., the negative-frequency part of the spectrum is folded around the vertical axis and is added to the positive-frequency part [Fig. 2.24(b)]. We call the representation of Fig. 2.24(a) the "two-sided" spectrum and that of Fig. 2.24(b) the "one-sided" spectrum.

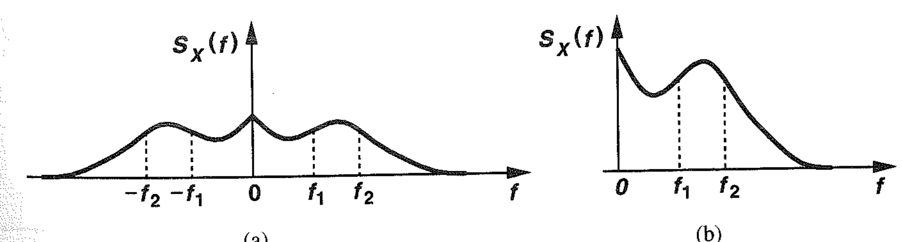

Figure 2.24 (a) Two-sided and (b) one-sided spectra.

In graphical analysis of frequency-domain operations, it is generally simpler to use a two-sided spectrum, whereas actual noise calculations are more easily carried out with a one-sided spectrum. Nevertheless, these two representations bear no fundamental difference—though they can cause confusion.

As an example of $S_x(f)$, we consider the thermal noise voltage across a resistor of value R. The two-sided PSD is

$$S_x(f) = 2kTR, \tag{2.69}$$

where k is the Boltzmann constant and equal to 1.38×10^{-23} J/K and T is the absolute temperature. Such a flat spectrum is called "white" because it contains the same level of power at all frequencies.

Equation (2.69) raises two interesting questions. First, is the total noise power of a resistor [the area under $S_x(f)$] infinite? In reality, $S_x(f)$ is flat for only $|f| < 100$ GHz, dropping beyond this frequency such that the total power remains finite [3]. Second, is the dimension of $2kTR$ power per unit bandwidth (W/Hz)? No, the actual dimension is mean square voltage per unit bandwidth. We tacitly assume that this voltage is applied across a 1-Ω resistor to generate a *power* of $2kTR$ in a 1-Hz bandwidth. In circuit noise calculations, we often write

$$\overline{V_n^2} = 4kTR \cdot \Delta f, \tag{2.70}$$

where $\overline{V_n^2}$ is the mean square noise voltage generated by resistor R in a bandwidth Δf. Called the "spot noise" for $\Delta f = 1$ Hz, $\overline{V_n^2}$ is measured in V^2/Hz.

To summarize the concepts of PDF and PSD, we note that the former is a statistical indication of how often the amplitude of a random process falls in a given range of values while the latter shows how much power the signal is expected to contain in a small frequency interval. In general, the PDF and PSD bear no relationship: thermal noise has a Gaussian PDF and a white PSD, whereas flicker ($1/f$) noise has the same type of PDF but a PSD proportional to $1/f$.

Random Signals in Linear Systems The principal reason for defining the power spectral density function is that it allows many of the frequency-domain operations used with deterministic signals to be applied to random processes as well. It can be shown that if a signal with spectral density $S_x(f)$ is applied to a linear time-invariant system with transfer function $H(s)$ (Fig. 2.25), then the output spectrum is

$$S_y(f) = S_x(f)|H(f)|^2, \tag{2.71}$$

where $H(f) = H(s = j2\pi f)$ [3]. This agrees with our intuition that the spectrum of the signal is shaped by the transfer function of the system. It can also be shown that if $x(t)$ is Gaussian, so is $y(t)$ [3].

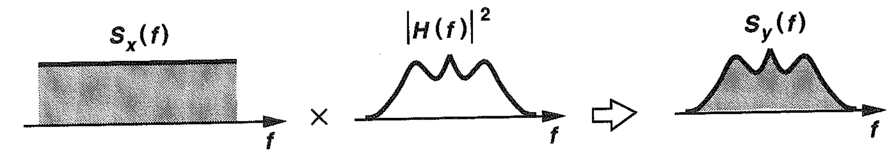

Figure 2.25 Noise shaping in a linear system.

2.3.2 Noise

Noise can be loosely defined as any random interference unrelated to the signal of interest.[5] This definition distinguishes between noise and deterministic phenomena such as harmonic distortion and intermodulation. As other random processes, noise is characterized by a PDF and a PSD.

Present in all circuits is thermal noise, generated by resistors, base and emitter resistance of bipolar devices, and channel resistance of MOSFETs [Fig. 2.26(a)]. The thermal noise of MOS devices is modeled as a current source connected between the drain and source with a PSD,

$$\overline{I_n^2} = 4kT\left(\frac{2}{3}g_m\right), \tag{2.72}$$

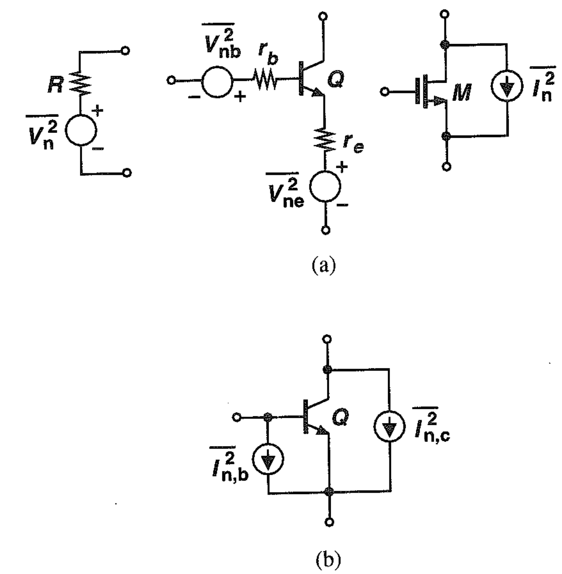

Figure 2.26 (a) Thermal and (b) shot noise in devices.

[5] It is often said that if there were no noise, there would be no analog designers.

where g_m is the transconductance of the transistor. Derived for long-channel devices [7], the factor 2/3 may need to be replaced with higher values for channel lengths below 1 μm [8]. Note that the distributed gate resistance of MOSFETs also contributes thermal noise [9], but the effect can be minimized through careful layout.

In addition to thermal noise, active devices may exhibit shot and flicker noise as well. Shot noise is a Gaussian white process associated with the transfer of charge across an energy barrier (e.g., a pn junction) having a PSD

$$\overline{I_n^2} = 2qI, \tag{2.73}$$

where q is the charge of an electron and I the average current. For a bipolar transistor, the collector and base current shot noise is modeled as a current source connected between the collector and emitter and another between the base and the emitter [Fig. 2.26(b)].

Flicker noise arises from random trapping of charge at the oxide-silicon interface of MOSFETs. Represented as a voltage source in series with the gate, the noise density is given by

$$\overline{V_n^2} = \frac{K}{WLCox}\frac{1}{f}, \tag{2.74}$$

where K is a process-dependent constant. While the effect of flicker noise may seem negligible at high frequencies, we must note that nonlinearity or time variance in circuits such as mixers or oscillators can translate the $1/f$-shaped spectrum to the RF range (Chapters 6 and 7).

Input-Referred Noise The noise of a two-port system can be modeled by two input noise generators: a series voltage source and a parallel current source (Fig. 2.27) [10]. In general, the correlation between the two sources must be taken into account. We use an example to illustrate the idea. Consider the circuit shown in Fig. 2.28(a), where we assume proper biasing ensures that M_1 is in saturation and carries a drain current of I_D. This circuit has only one dominant source of thermal noise: that due to the channel and represented by $\overline{I_{nD}^2}$. For the model of Fig. 2.28(b), we calculate $\overline{V_n^2}$ by shorting the input port and $\overline{I_n^2}$ by leaving it open. Since the circuits of Figs. 2.28(a) and (b) must produce

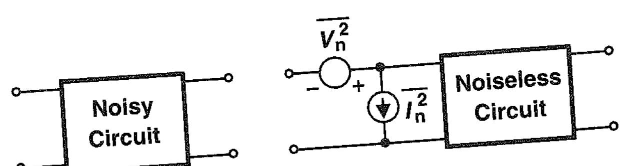

Figure 2.27 Representation of noise by input noise generators.

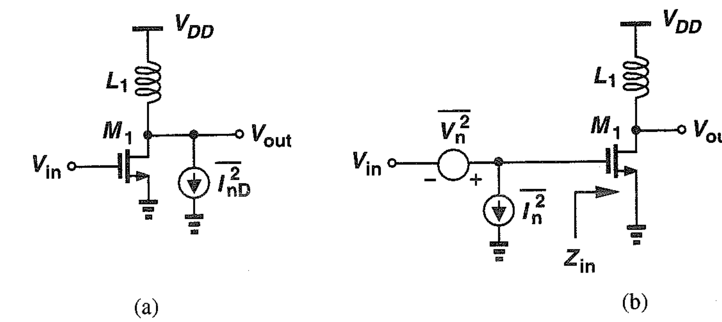

Figure 2.28 (a) MOS amplifier, (b) equivalent input noise generators.

the same output noise in both cases, we have $g_m^2\overline{V_n^2} = \overline{I_{nD}^2}$ and $g_m^2\overline{I_n^2}|Z_{in}|^2 = \overline{I_{nD}^2}$, where Z_{in} denotes the input impedance of the circuit. Thus, for $\overline{I_{nD}^2} = 4kT(2g_m/3)$, we obtain $\overline{V_n^2} = 8kT/(3g_m)$ and $\overline{I_n^2} = 8kT/(3g_m|Z_{in}|^2)$. Since V_n and I_n represent the same noise mechanism, they are correlated.

We note that if $|Z_{in}| \to \infty$, $\overline{I_n^2} \to 0$, and $\overline{V_n^2}$ is sufficient to represent the noise. At radio frequencies, however, $|Z_{in}|$ is relatively low (in some cases around 50 Ω by design), thereby necessitating the use of both $\overline{V_n^2}$ and $\overline{I_n^2}$.

The key point in the above example is that even though the actual circuit may have no physical input noise current, the representation using input-referred sources must include $\overline{I_n^2}$.

Noise Figure In many analog circuits, the signal-to-noise ratio (SNR), defined as the ratio of the signal power to the total noise power, is an important parameter. In RF design, on the other hand, even though the ultimate goal is to maximize the SNR for the received and detected signal, most of the front-end receiver blocks are characterized in terms of their "noise figure" rather than the input-referred noise. This is partly for computational convenience and partly from tradition.

Noise figure has been defined in a number of different ways. The most commonly accepted definition is

$$\text{noise figure} = \frac{SNR_{in}}{SNR_{out}}, \tag{2.75}$$

where SNR_{in} and SNR_{out} are the signal-to-noise ratios measured at the input and output, respectively. Note that the above ratio is called the "noise factor" in most textbooks, with the term noise figure applied to $10\log_{10}$(noise factor). We do not make this distinction here.

It is important to understand the physical meaning of (2.75). Noise figure is a measure of how much the SNR degrades as the the signal passes through a system. If a system has no noise, then $SNR_{out} = SNR_{in}$, regardless of the gain. This is because both the input signal and the input noise are amplified (or attenuated) by the same factor and no additional noise is introduced. Therefore, the noise figure of a noiseless system is equal to unity. In reality, the finite noise of a system degrades the SNR, yielding $NF > 1$.

Compared to input-referred noise, the definition of NF in (2.75) may seem rather complicated: it depends on not only the noise of the circuit under consideration but the SNR of the *preceding* stage. In fact, if the input signal contains no noise, $SNR_{in} = \infty$ and $NF = \infty$, even though the circuit may have only a finite internal noise. For such a case, NF is not a meaningful parameter. In RF design, on the other hand, this does not occur because even the signal in the first stage of a receiver is corrupted by the noise due to the radiation resistance of the antenna.

Calculation of the noise figure is generally simpler than (2.75) may suggest. As depicted in Fig. 2.29, we assume SNR_{in} is the ratio of the input signal power to the noise generated by the source resistance, R_S, and modeled by $\overline{V_{RS}^2}$. If the voltage gain from V_{in} to the input port of the circuit (node P) is equal to α, the SNR measured at this node is

$$SNR_{in} = \frac{\alpha^2 V_{in}^2}{\alpha^2 \overline{V_{RS}^2}}. \tag{2.76}$$

For a voltage gain of A_v from P to V_{out}, the SNR measured at the output is equal to

$$SNR_{out} = \frac{\alpha^2 A_v^2 V_{in}^2}{[\overline{V_{RS}^2} + \overline{(V_n + I_n R_S)^2}]\alpha^2 A_v^2} \tag{2.77}$$

$$= \frac{V_{in}^2}{[\overline{V_{RS}^2} + \overline{(V_n + I_n R_S)^2}]}, \tag{2.78}$$

where V_n and $I_n R_S$ are added before squaring to account for their correlation. It follows that,

$$NF = \frac{\overline{V_{RS}^2} + \overline{(V_n + I_n R_S)^2}}{\overline{V_{RS}^2}} \tag{2.79}$$

$$= 1 + \frac{\overline{(V_n + I_n R_S)^2}}{\overline{V_{RS}^2}}. \tag{2.80}$$

The NF is usually specified for a 1-Hz bandwidth at a given frequency. Called the "spot" noise figure to emphasize the very small bandwidth, this quantity can be obtained from (2.80) as

$$NF = 1 + \frac{\overline{(V_n + I_n R_S)^2}}{4kTR_S}, \tag{2.81}$$

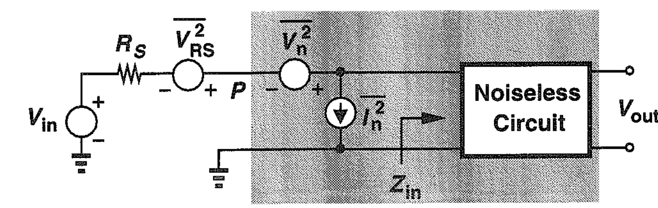

Figure 2.29 Calculation of noise figure.

where V_n and I_n are also measured in unity bandwidth and at the same frequency.

Equation (2.81) indicates that the noise figure is a function of the source impedance, R_S. In general, knowledge of the NF for a given source impedance is not sufficient to calculate the NF for a different source impedance because R_S appears in both the numerator and the denominator of the fraction in (2.81). In traditional RF systems, however, most building blocks are designed so as to exhibit 50-Ω input and output resistance[6] (with negligible reactance), thereby avoiding ambiguity in NF calculation. As we will see in Chapter 6, this issue still requires attention in certain cases.

For simulation purposes, it is beneficial to write (2.81) as

$$NF = \frac{4kTR_S + \overline{(V_n + I_n R_S)^2}}{4kTR_S} \tag{2.82}$$

$$= \frac{A^2[4kTR_S + \overline{(V_n + I_n R_S)^2}]}{A^2} \frac{1}{4kTR_S} \tag{2.83}$$

$$= \frac{\overline{V_{n,out}^2}}{A^2} \frac{1}{4kTR_S}, \tag{2.84}$$

where $A = \alpha A_v$ and $\overline{V_{n,out}^2}$ represents the total noise at the output. Thus, to calculate NF in Fig. 2.29, we divide the total output noise power by the square of the voltage gain from V_{in} to V_{out} and normalize the result to the noise of R_S.

As an example of noise figure calculation, consider the single resistor, R_P, shown in Fig. 2.30(a). What is the noise figure of this circuit with respect to a source resistance R_S? From Fig. 2.30(b), the total output noise voltage is given by

$$\overline{V_{n,out}^2} = 4kT(R_S||R_P), \tag{2.85}$$

and the gain is

$$A_v = \frac{R_P}{R_S + R_P}. \tag{2.86}$$

[6] In TV systems, the characteristic and termination impedances are 75 Ω.