

Deterministic and Random Signal Analysis

In this chapter we present the background material needed in the study of the following chapters. The analysis of deterministic and random signals and the study of different methods for their representation are the main topics of this chapter. In addition, we also introduce and study the main properties of some random variables frequently encountered in analysis of communication systems. We continue with a review of random processes, properties of lowpass and bandpass random processes, and series expansion of random processes.

Throughout this chapter, and the book, we assume that the reader is familiar with the properties of the Fourier transform as summarized in Table 2.0–1 and the important Fourier transform pairs given in Table 2.0–2.

In these tables we have used the following signal definitions.

$$\Pi(t) = \begin{cases} 1 & |t| < \frac{1}{2} \\ \frac{1}{2} & t = \pm \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \quad \text{sinc}(t) = \begin{cases} \frac{\sin(\pi t)}{\pi t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

and

$$\text{sgn}(t) = \begin{cases} 1 & t > 0 \\ -1 & t < 0 \\ 0 & t = 0 \end{cases} \quad \Lambda(t) = \Pi(t) \star \Pi(t) = \begin{cases} t+1 & -1 \leq t < 0 \\ -t+1 & 0 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

The unit step signal $u_{-1}(t)$ is defined as

$$u_{-1}(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0 \end{cases}$$

We also assume that the reader is familiar with elements of probability, random variables, and random processes as covered in standard texts such as Papoulis and Pillai (2002), Leon-Garcia (1994), and Stark and Woods (2002).

■ TABLE 2.0-1
Table of Fourier Transform Properties

Property	Signal	Fourier Transform
Linearity	$\alpha x_1(t) + \beta x_2(t)$	$\alpha X_1(f) + \beta X_2(f)$
Duality	$X(t)$	$x(-f)$
Conjugacy	$x^*(t)$	$X^*(-f)$
Time-scaling ($a \neq 0$)	$x(at)$	$\frac{1}{ a } X\left(\frac{f}{a}\right)$
Time-shift	$x(t - t_0)$	$e^{-j2\pi f t_0} X(f)$
Modulation	$e^{j2\pi f_0 t} x(t)$	$X(f - f_0)$
Convolution	$x(t) \star y(t)$	$X(f)Y(f)$
Multiplication	$x(t)y(t)$	$X(f) \star Y(f)$
Differentiation	$\frac{d^n}{dt^n} x(t)$	$(j2\pi f)^n X(f)$
Differentiation in frequency	$t^n x(t)$	$\left(\frac{j}{2\pi}\right)^n \frac{d^n}{df^n} X(f)$
Integration	$\int_{-\infty}^t x(\tau) d\tau$	$\frac{X(f)}{j2\pi f} + \frac{1}{2} X(0)\delta(f)$
Parseval's theorem	$\int_{-\infty}^{\infty} x(t)y^*(t) dt = \int_{-\infty}^{\infty} X(f)Y^*(f) df$	
Rayleigh's theorem	$\int_{-\infty}^{\infty} x(t) ^2 dt = \int_{-\infty}^{\infty} X(f) ^2 df$	

■ 2.1

BANDPASS AND LOWPASS SIGNAL REPRESENTATION

As was discussed in Chap. 1, the process of communication consists of transmission of the output of an information source over a communication channel. In almost all cases, the spectral characteristics of the information sequence do not directly match the spectral characteristics of the communication channel, and hence the information signal cannot be directly transmitted over the channel. In many cases the information signal is a low frequency (baseband) signal, and the available spectrum of the communication channel is at higher frequencies. Therefore, at the transmitter the information signal is translated to a higher frequency signal that matches the properties of the communication channel. This is the modulation process in which the baseband information signal is turned into a bandpass modulated signal. In this section we study the main properties of baseband and bandpass signals.

2.1-1 Bandpass and Lowpass Signals

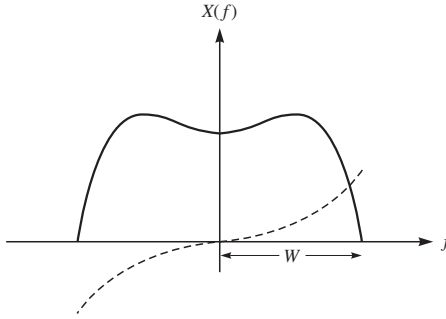
In this section we will show that any real, narrowband, and high frequency signal—called a bandpass signal—can be represented in terms of a complex low frequency

■ TABLE 2.0-2
Table of Fourier Transform Pairs

Time Domain	Frequency Domain
$\delta(t)$	1
1	$\delta(f)$
$\delta(t - t_0)$	$e^{-j2\pi f t_0}$
$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}\delta(f - f_0) + \frac{1}{2}\delta(f + f_0)$
$\sin(2\pi f_0 t)$	$\frac{1}{2j}\delta(f - f_0) - \frac{1}{2j}\delta(f + f_0)$
$\Pi(t)$	$\text{sinc}(f)$
$\text{sinc}(t)$	$\Pi(f)$
$\Lambda(t)$	$\text{sinc}^2(f)$
$\text{sinc}^2(t)$	$\Lambda(f)$
$e^{-\alpha t} u_{-1}(t), \alpha > 0$	$\frac{1}{\alpha + j2\pi f}$
$t e^{-\alpha t} u_{-1}(t), \alpha > 0$	$\frac{1}{(\alpha + j2\pi f)^2}$
$e^{-\alpha t } (\alpha > 0)$	$\frac{2\alpha}{\alpha^2 + (2\pi f)^2}$
$e^{-\pi t^2}$	$e^{-\pi f^2}$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$u_{-1}(t)$	$\frac{1}{2}\delta(f) + \frac{1}{j2\pi f}$
$\frac{1}{2}\delta(t) + j\frac{1}{2\pi t}$	$u_{-1}(f)$
$\delta'(t)$	$j2\pi f$
$\delta^{(n)}(t)$	$(j2\pi f)^n$
$\frac{1}{t}$	$-j\pi \text{sgn}(f)$
$\sum_{n=-\infty}^{\infty} \delta(t - nT_0)$	$\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_0}\right)$

signal, called the lowpass equivalent of the original bandpass signal. This result makes it possible to work with the lowpass equivalents of bandpass signals instead of directly working with them, thus greatly simplifying the handling of bandpass signals. That is so because applying signal processing algorithms to lowpass signals is much easier due to lower required sampling rates which in turn result in lower rates of the sampled data.

The Fourier transform of a signal provides information about the frequency content, or *spectrum*, of the signal. The Fourier transform of a real signal $x(t)$ has *Hermitian symmetry*, i.e., $X(-f) = X^*(f)$, from which we conclude that $|X(-f)| = |X(f)|$ and $\angle X^*(f) = -\angle X(f)$. In other words, for real $x(t)$, the magnitude of $X(f)$ is even and

**FIGURE 2.1-1**

The spectrum of a real-valued lowpass (baseband) signal.

its phase is odd. Because of this symmetry, all information about the signal is in the positive (or negative) frequencies, and in particular $x(t)$ can be perfectly reconstructed by specifying $X(f)$ for $f \geq 0$. Based on this observation, for a real signal $x(t)$, we define the *bandwidth* as the smallest range of *positive frequencies* such that $X(f) = 0$ when $|f|$ is outside this range. It is clear that the bandwidth of a real signal is one-half of its frequency support set.

A *lowpass*, or *baseband*, signal is a signal whose spectrum is located around the zero frequency. For instance, speech, music, and video signals are all lowpass signals, although they have different spectral characteristics and bandwidths. Usually lowpass signals are low frequency signals, which means that in the time domain, they are slowly varying signals with no jumps or sudden variations. The bandwidth of a real lowpass signal is the minimum positive W such that $X(f) = 0$ outside $[-W, +W]$. For these signals the *frequency support*, i.e., the range of frequencies for which $X(f) \neq 0$, is $[-W, +W]$. An example of the spectrum of a real-valued lowpass signal is shown in Fig. 2.1-1. The solid line shows the magnitude spectrum $|X(f)|$, and the dashed line indicates the phase spectrum $\angle X(f)$.

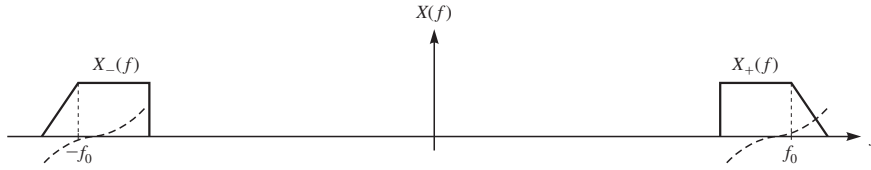
We also define the *positive spectrum* and the *negative spectrum* of a signal $x(t)$ as

$$X_+(f) = \begin{cases} X(f) & f > 0 \\ \frac{1}{2}X(0) & f = 0 \\ 0 & f < 0 \end{cases} \quad X_-(f) = \begin{cases} X(f) & f < 0 \\ \frac{1}{2}X(0) & f = 0 \\ 0 & f > 0 \end{cases} \quad (2.1-1)$$

It is clear that $X_+(f) = X(f)u_{-1}(f)$, $X_-(f) = X(f)u_{-1}(-f)$ and $X(f) = X_+(f) + X_-(f)$. For a real signal $x(t)$, since $X(f)$ is Hermitian, we have $X_-(f) = X_+^*(-f)$.

For a complex signal $x(t)$, the spectrum $X(f)$ is not symmetric; hence, the signal cannot be reconstructed from the information in the positive frequencies only. For complex signals, we define the *bandwidth* as one-half of the entire range of frequencies over which the spectrum is nonzero, i.e., one-half of the *frequency support* of the signal. This definition is for consistency with the definition of bandwidth for real signals. With this definition we can state that in general and for all signals, real or complex, the bandwidth is defined as one-half of the frequency support.

In practice, the spectral characteristics of the message signal and the communication channel do not always match, and it is required that the message signal be *modulated* by one of the many different modulation methods to match its spectral characteristics to

**FIGURE 2.1-2**

The spectrum of a real-valued bandpass signal.

the spectral characteristics of the channel. In this process, the spectrum of the lowpass message signal is translated to higher frequencies. The resulting modulated signal is a bandpass signal.

A *bandpass signal* is a real signal whose frequency content, or spectrum, is located around some frequency $\pm f_0$ which is far from zero. More formally, we define a bandpass signal to be a real signal $x(t)$ for which there exists positive f_0 and W such that the positive spectrum of $X(f)$, i.e., $X_+(f)$, is nonzero only in the interval $[f_0 - W/2, f_0 + W/2]$, where $W/2 < f_0$ (in practice, usually $W \ll f_0$). The frequency f_0 is called the *central frequency*. Obviously, the bandwidth of $x(t)$ is at most equal to W . Bandpass signals are usually high frequency signals which are characterized by rapid variations in the time domain.

An example of the spectrum of a bandpass signal is shown in Figure 2.1-2. Note that since the signal $x(t)$ is real, its magnitude spectrum (solid line) is even, and its phase spectrum (dashed line) is odd. Also, note that the central frequency f_0 is not necessarily the midband frequency of the bandpass signal. Due to the symmetry of the spectrum, $X_+(f)$ has all the information that is necessary to reconstruct $X(f)$. In fact we can write

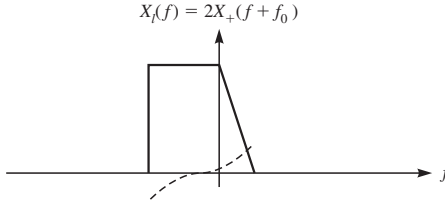
$$X(f) = X_+(f) + X_-(f) = X_+(f) + X_+^*(-f) \quad (2.1-2)$$

which means that knowledge of $X_+(f)$ is sufficient to reconstruct $X(f)$.

2.1-2 Lowpass Equivalent of Bandpass Signals

We start by defining the *analytic signal*, or the *pre-envelope*, corresponding to $x(t)$ as the signal $x_+(t)$ whose Fourier transform is $X_+(f)$. This signal contains only positive frequency components, and its spectrum is not Hermitian. Therefore, in general, $x_+(t)$ is a complex signal. We have

$$\begin{aligned} x_+(t) &= \mathcal{F}^{-1} [X_+(f)] \\ &= \mathcal{F}^{-1} [X(f)u_{-1}(f)] \\ &= x(t) \star \left(\frac{1}{2}\delta(t) + j\frac{1}{2\pi t} \right) \\ &= \frac{1}{2}x(t) + \frac{j}{2}\hat{x}(t) \end{aligned} \quad (2.1-3)$$

**FIGURE 2.1-3**

The spectrum of the lowpass equivalent of the signal shown in Figure 2.1-2.

where $\hat{x}(t) = \frac{1}{\pi t} \star x(t)$ is the *Hilbert transform* of $x(t)$. The Hilbert transform of $x(t)$ is obtained by introducing a phase shift of $-\frac{\pi}{2}$ at positive frequency components of $x(t)$ and $\frac{\pi}{2}$ at negative frequencies. In the frequency domain we have

$$\mathcal{F}[\hat{x}(t)] = -j \operatorname{sgn}(f) X(f) \quad (2.1-4)$$

Some of the properties of the Hilbert transform will be covered in the problems at the end of this chapter.

Now we define $x_l(t)$, the *lowpass equivalent*, or the *complex envelope*, of $x(t)$, as the signal whose spectrum is given by $2X_+(f + f_0)$, i.e.,

$$X_l(f) = 2X_+(f + f_0) = 2X(f + f_0)u_{-1}(f + f_0) \quad (2.1-5)$$

Obviously the spectrum of $x_l(t)$ is located around the zero frequency, and therefore it is in general a complex lowpass signal. This signal is called the *lowpass equivalent* or the *complex envelope* of $x(t)$. The spectrum of the lowpass equivalent of the signal shown in Figure 2.1-2 is shown in Figure 2.1-3.

Applying the modulation theorem of the Fourier transform, we obtain

$$\begin{aligned} x_l(t) &= \mathcal{F}^{-1}[X_l(f)] \\ &= 2x_+(t)e^{-j2\pi f_0 t} \\ &= (x(t) + j\hat{x}(t))e^{-j2\pi f_0 t} \end{aligned} \quad (2.1-6)$$

$$\begin{aligned} &= (x(t) \cos 2\pi f_0 t + \hat{x}(t) \sin 2\pi f_0 t) \\ &\quad + j(\hat{x}(t) \cos 2\pi f_0 t - x(t) \sin 2\pi f_0 t) \end{aligned} \quad (2.1-7)$$

From Equation 2.1-6 we can write

$$x(t) = \operatorname{Re} [x_l(t)e^{j2\pi f_0 t}] \quad (2.1-8)$$

This relation expresses any bandpass signals in terms of its lowpass equivalent. Using Equations 2.1-2 and 2.1-5, we can write

$$X(f) = \frac{1}{2} [X_l(f - f_0) + X_l^*(-f - f_0)] \quad (2.1-9)$$

Equations 2.1-8, 2.1-9, 2.1-5, and 2.1-7 express $x(t)$ and $x_l(t)$ in terms of each other in the time and frequency domains.

The real and imaginary parts of $x_l(t)$ are called the *in-phase component* and the *quadrature component* of $x(t)$, respectively, and are denoted by $x_i(t)$ and $x_q(t)$. Both $x_i(t)$ and $x_q(t)$ are real-valued lowpass signals, and we have

$$x_l(t) = x_i(t) + jx_q(t) \quad (2.1-10)$$

Comparing Equations 2.1–10 and 2.1–7, we conclude that

$$\begin{aligned} x_i(t) &= x(t) \cos 2\pi f_0 t + \hat{x}(t) \sin 2\pi f_0 t \\ x_q(t) &= \hat{x}(t) \cos 2\pi f_0 t - x(t) \sin 2\pi f_0 t \end{aligned} \quad (2.1-11)$$

Solving Equation 2.1–11 for $x(t)$ and $\hat{x}(t)$ gives

$$\begin{aligned} x(t) &= x_i(t) \cos 2\pi f_0 t - x_q(t) \sin 2\pi f_0 t \\ \hat{x}(t) &= x_q(t) \cos 2\pi f_0 t + x_i(t) \sin 2\pi f_0 t \end{aligned} \quad (2.1-12)$$

Equation 2.1–12 shows that any bandpass signal $x(t)$ can be expressed in terms of two lowpass signals, namely, its in-phase and quadrature components.

Equation 2.1–10 expresses $x_i(t)$ in terms of its real and complex parts. We can write a similar relation in polar coordinates expressing $x(t)$ in terms of its magnitude and phase. If we define the *envelope* and *phase* of $x(t)$, denoted by $r_x(t)$ and $\theta_x(t)$, respectively, by

$$r_x(t) = \sqrt{x_i^2(t) + x_q^2(t)} \quad (2.1-13)$$

$$\theta_x(t) = \arctan \frac{x_q(t)}{x_i(t)} \quad (2.1-14)$$

we have

$$x_i(t) = r_x(t) e^{j\theta_x(t)} \quad (2.1-15)$$

Substituting this result into Equation 2.1–8 gives

$$x(t) = \text{Re} [r_x(t) e^{j(2\pi f_0 t + \theta_x(t))}] \quad (2.1-16)$$

resulting in

$$x(t) = r_x(t) \cos (2\pi f_0 t + \theta_x(t)) \quad (2.1-17)$$

A bandpass signal and its envelope are shown in Figure 2.1–4.

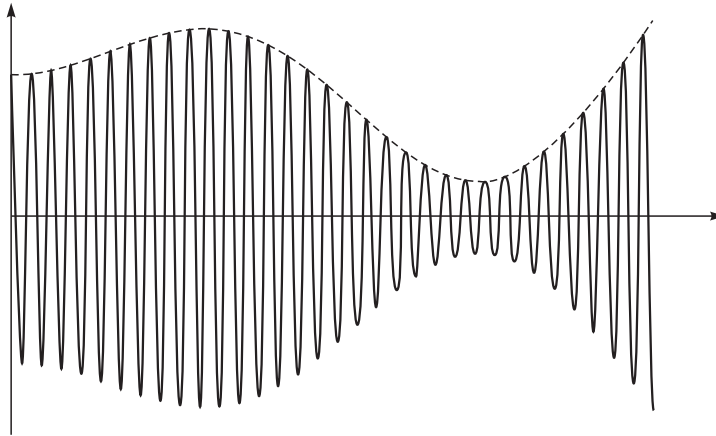


FIGURE 2.1–4

A bandpass signal. The dashed curve denotes the envelope.

It is important to note that $x_l(t)$ —and consequently $x_i(t)$, $x_q(t)$, $r_x(t)$, and $\theta_x(t)$ —depends on the choice of the central frequency f_0 . For a given bandpass signal $x(t)$, different values of f_0 —as long as $X_+(f)$ is nonzero only in the interval $[f_0 - W/2, f_0 + W/2]$, where $W/2 < f_0$ —yield different lowpass signals $x_l(t)$. Therefore, it makes more sense to define the lowpass equivalent of a bandpass signal with respect to a specific f_0 . Since in most cases the choice of f_0 is clear, we usually do not make this distinction.

Equations 2.1–12 and 2.1–17 provide two methods for representing a bandpass signal $x(t)$ in terms of two lowpass signals, one in terms of the in-phase and quadrature components and one in terms of the envelope and the phase. The two relations given in Equations 2.1–8 and 2.1–12 that express the bandpass signal in terms of the lowpass component(s) define the modulation process, i.e., the process of going from lowpass to bandpass. The system that implements this process is called a *modulator*. The structure of a general modulator implementing Equations 2.1–8 and 2.1–12 is shown in Figure 2.1–5(a) and (b). In this figure double lines and double blocks indicate complex values and operations.

Similarly, Equations 2.1–7 and 2.1–11 represent how $x_l(t)$, or $x_i(t)$ and $x_q(t)$, can be obtained from the bandpass signal $x(t)$. This process, i.e., extracting the lowpass signal from the bandpass signal, is called the *demodulation* process and is shown in Figure 2.1–6(a) and (b). In these block diagrams the block denoted by $\mathcal{H}$ represents a Hilbert transform, i.e., an LTI system with impulse response $h(t) = \frac{1}{\pi t}$ and transfer function $H(f) = -j \operatorname{sgn}(f)$.

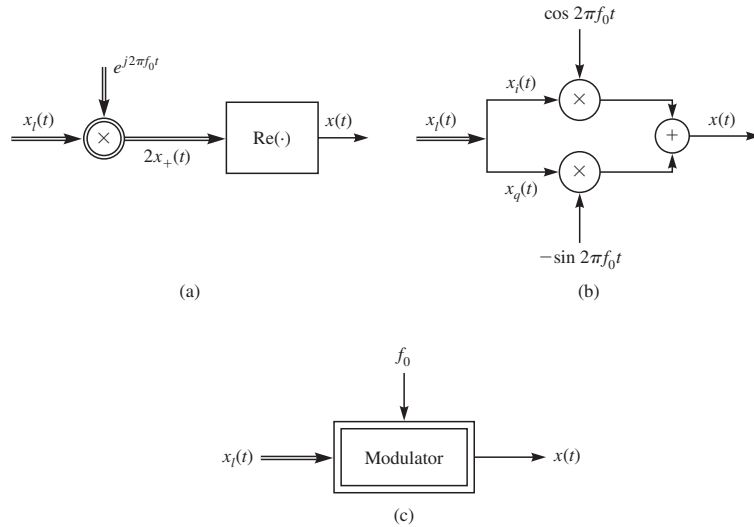
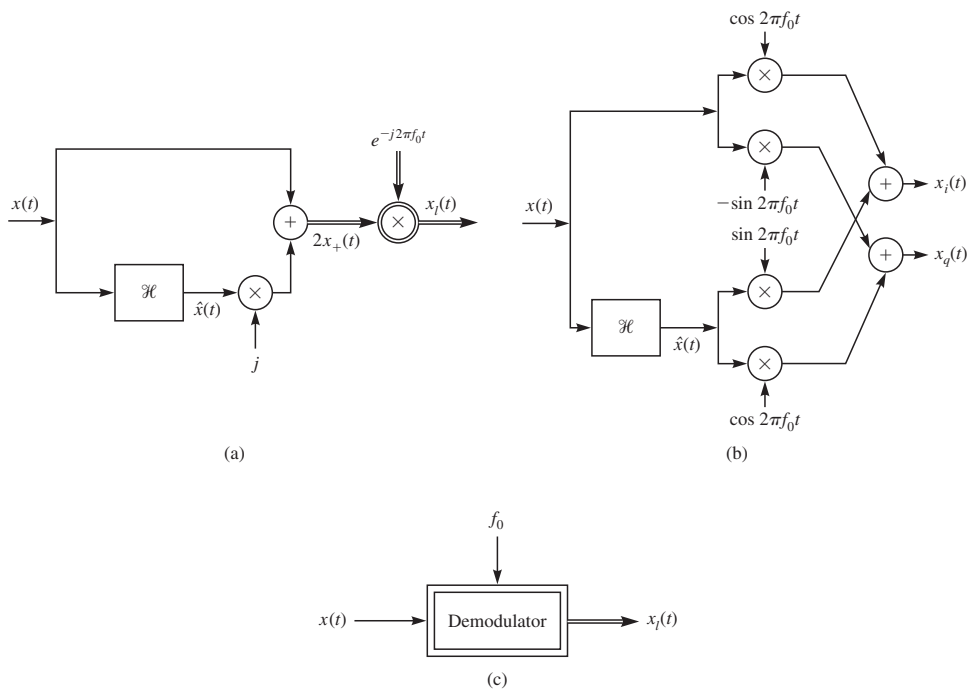


FIGURE 2.1–5

A complex (a) and real (b) modulator. A general representation for a modulator is shown in (c).

**FIGURE 2.1-6**

A complex (a) and real (b) demodulator. A general representation for a demodulator is shown in (c).

2.1-3 Energy Considerations

In this section we study the relation between energy contents of the signals introduced in the preceding pages. The *energy* of a signal $x(t)$ is defined as

$$\mathcal{E}_x = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (2.1-18)$$

and by Rayleigh's relation from Table 2.0-1 we can write

$$\mathcal{E}_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df \quad (2.1-19)$$

Since there is no overlap between $X_+(f)$ and $X_-(f)$, we have $X_+(f)X_-(f) = 0$, and hence

$$\begin{aligned} \mathcal{E}_x &= \int_{-\infty}^{\infty} |X_+(f) + X_-(f)|^2 df \\ &= \int_{-\infty}^{\infty} |X_+(f)|^2 df + \int_{-\infty}^{\infty} |X_-(f)|^2 df \\ &= 2 \int_{-\infty}^{\infty} |X_+(f)|^2 df \\ &= 2\mathcal{E}_{x_+} \end{aligned} \quad (2.1-20)$$

On the other hand,

$$\begin{aligned}
 \mathcal{E}_x &= 2 \int_{-\infty}^{\infty} |X_+(f)|^2 df \\
 &= 2 \int_{-\infty}^{\infty} \left| \frac{X_l(f)}{2} \right|^2 df \\
 &= \frac{1}{2} \mathcal{E}_{x_l}
 \end{aligned} \tag{2.1-21}$$

This shows that the energy in the lowpass equivalent signal is twice the energy in the bandpass signal.

We define the *inner product of two signals* $x(t)$ and $y(t)$ as

$$\langle x(t), y(t) \rangle = \int_{-\infty}^{\infty} x(t)y^*(t) dt = \int_{-\infty}^{\infty} X(f)Y^*(f) df \tag{2.1-22}$$

where we have used Parseval's relation from Table 2.0-1. Obviously

$$\mathcal{E}_x = \langle x(t), x(t) \rangle \tag{2.1-23}$$

In Problem 2.2 we prove that if $x(t)$ and $y(t)$ are two bandpass signals with lowpass equivalents $x_l(t)$ and $y_l(t)$ with respect to the same f_0 , then

$$\langle x(t), y(t) \rangle = \frac{1}{2} \text{Re} [\langle x_l(t), y_l(t) \rangle] \tag{2.1-24}$$

The complex quantity $\rho_{x,y}$, called the *cross-correlation coefficient* of $x(t)$ and $y(t)$, is defined as

$$\rho_{x,y} = \frac{\langle x(t), y(t) \rangle}{\sqrt{\mathcal{E}_x \mathcal{E}_y}} \tag{2.1-25}$$

and represents the normalized inner product between two signals. From $\mathcal{E}_{x_l} = 2\mathcal{E}_x$ and Equation 2.1-24 we can conclude that if $x(t)$ and $y(t)$ are bandpass signals with the same f_0 , then

$$\rho_{x,y} = \text{Re} (\rho_{x_l, y_l}) \tag{2.1-26}$$

Two signals are *orthogonal* if their inner product (and subsequently, their ρ) is zero. Note that if $\rho_{x_l, y_l} = 0$, then using Equation 2.1-26, we have $\rho_{x,y} = 0$; but the converse is not necessarily true. In other words, *orthogonality in the baseband implies orthogonality in the pass band, but not vice versa*.

EXAMPLE 2.1-1. Assume that $m(t)$ is a real baseband signal with bandwidth W , and define two signals $x(t) = m(t) \cos 2\pi f_0 t$ and $y(t) = m(t) \sin 2\pi f_0 t$, where $f_0 > W$. Comparing these relations with Equation 2.1-12, we conclude that

$$\begin{aligned}
 x_i(t) &= m(t) & x_q(t) &= 0 \\
 y_i(t) &= 0 & y_q(t) &= -m(t)
 \end{aligned}$$

or, equivalently,

$$\begin{aligned}x_l(t) &= m(t) \\ y_l(t) &= -jm(t)\end{aligned}$$

Note that here

$$\rho_{x_l, y_l} = j \int_{-\infty}^{\infty} m^2(t) dt = j\mathcal{E}_m$$

Therefore,

$$\rho_{x, y} = \text{Re}(\rho_{x_l, y_l}) = \text{Re}(j\mathcal{E}_m) = 0$$

This means that $x(t)$ and $y(t)$ are orthogonal, but their lowpass equivalents are not orthogonal.

2.1–4 Lowpass Equivalent of a Bandpass System

A bandpass system is a system whose transfer function is located around a frequency f_0 (and its mirror image $-f_0$). More formally, we define a bandpass system as a system whose impulse response $h(t)$ is a bandpass signal. Since $h(t)$ is bandpass, it has a lowpass equivalent denoted by $h_l(t)$ where

$$h(t) = \text{Re}[h_l(t)e^{j2\pi f_0 t}] \quad (2.1-27)$$

If a bandpass signal $x(t)$ passes through a bandpass system with impulse response $h(t)$, then obviously the output will be a bandpass signal $y(t)$. The relation between the spectra of the input and the output is given by

$$Y(f) = X(f)H(f) \quad (2.1-28)$$

Using Equation 2.1–5, we have

$$\begin{aligned}Y_l(f) &= 2Y(f + f_0)u_{-1}(f + f_0) \\ &= 2X(f + f_0)H(f + f_0)u_{-1}(f + f_0) \\ &= \frac{1}{2} [2X(f + f_0)u_{-1}(f + f_0)] [2H(f + f_0)u_{-1}(f + f_0)] \\ &= \frac{1}{2} X_l(f)H_l(f)\end{aligned} \quad (2.1-29)$$

where we have used the fact that for $f > -f_0$, which is the range of frequencies of interest, $u_{-1}^2(f + f_0) = u_{-1}(f + f_0) = 1$. In the time domain we have

$$y_l(t) = \frac{1}{2} x_l(t) \star h_l(t) \quad (2.1-30)$$

Equations 2.1–29 and 2.1–30 show that when a bandpass signal passes through a bandpass system, the input-output relation between the lowpass equivalents is very similar to the relation between the bandpass signals, the only difference being that for the lowpass equivalents a factor of $\frac{1}{2}$ is introduced.

■ 2.2

SIGNAL SPACE REPRESENTATION OF WAVEFORMS

Signal space (or vector) representation of signals is a very effective and useful tool in the analysis of digitally modulated signals. We cover this important approach in this section and show that any set of signals is equivalent to a set of vectors. We show that signals have the same basic properties of vectors. We study methods of determining an equivalent set of vectors for a set of signals and introduce the notion of signal space representation, or *signal constellation*, of a set of waveforms.

2.2–1 Vector Space Concepts

A vector $\mathbf{v}$ in an n -dimensional space is characterized by its n components $v_1 v_2 \cdots v_n$. Let $\mathbf{v}$ denote a column vector, i.e., $\mathbf{v} = [v_1 v_2 \cdots v_n]^t$, where A^t denotes the transpose of matrix A . The *inner product* of two n -dimensional vectors $\mathbf{v}_1 = [v_{11} v_{12} \cdots v_{1n}]^t$ and $\mathbf{v}_2 = [v_{21} v_{22} \cdots v_{2n}]^t$ is defined as

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1 \cdot \mathbf{v}_2 = \sum_{i=1}^n v_{1i} v_{2i}^* = \mathbf{v}_2^H \mathbf{v}_1 \quad (2.2-1)$$

where A^H denotes the *Hermitian transpose* of the matrix A , i.e., the result of first transposing the matrix and then conjugating its elements. From the definition of the inner product of two vectors it follows that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle^* \quad (2.2-2)$$

and therefore,

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = 2 \operatorname{Re} [\langle \mathbf{v}_1, \mathbf{v}_2 \rangle] \quad (2.2-3)$$

A vector may also be represented as a linear combination of orthogonal unit vectors or an *orthonormal basis* $\mathbf{e}_i$, $1 \leq i \leq n$, i.e.,

$$\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i \quad (2.2-4)$$

where, by definition, a unit vector has length unity and v_i is the projection of the vector $\mathbf{v}$ onto the unit vector $\mathbf{e}_i$, i.e., $v_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$. Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are *orthogonal* if $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. More generally, a set of m vectors $\mathbf{v}_k$, $1 \leq k \leq m$, are orthogonal if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $1 \leq i, j \leq m$, and $i \neq j$. The *norm* of a vector $\mathbf{v}$ is denoted by $\|\mathbf{v}\|$ and is defined as

$$\|\mathbf{v}\| = (\langle \mathbf{v}, \mathbf{v} \rangle)^{1/2} = \sqrt{\sum_{i=1}^n |v_i|^2} \quad (2.2-5)$$

which in the n -dimensional space is simply the length of the vector. A set of m vectors is said to be *orthonormal* if the vectors are orthogonal and each vector has a

unit norm. A set of m vectors is said to be *linearly independent* if no one vector can be represented as a linear combination of the remaining vectors. Any two n -dimensional vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ satisfy the *triangle inequality*

$$\|\mathbf{v}_1 + \mathbf{v}_2\| \leq \|\mathbf{v}_1\| + \|\mathbf{v}_2\| \quad (2.2-6)$$

with equality if $\mathbf{v}_1$ and $\mathbf{v}_2$ are in the same direction, i.e., $\mathbf{v}_1 = a\mathbf{v}_2$ where a is a positive real scalar. The *Cauchy-Schwarz inequality* states that

$$|\langle \mathbf{v}_1, \mathbf{v}_2 \rangle| \leq \|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\| \quad (2.2-7)$$

with equality if $\mathbf{v}_1 = a\mathbf{v}_2$ for some complex scalar a . The norm square of the sum of two vectors may be expressed as

$$\|\mathbf{v}_1 + \mathbf{v}_2\|^2 = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 + 2 \operatorname{Re} [\langle \mathbf{v}_1, \mathbf{v}_2 \rangle] \quad (2.2-8)$$

If $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal, then $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$ and, hence,

$$\|\mathbf{v}_1 + \mathbf{v}_2\|^2 = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 \quad (2.2-9)$$

This is the *Pythagorean relation* for two orthogonal n -dimensional vectors. From matrix algebra, we recall that a linear transformation in an n -dimensional vector space is a matrix transformation of the form $\mathbf{v}' = \mathbf{A}\mathbf{v}$, where the matrix $\mathbf{A}$ transforms the vector $\mathbf{v}$ into some vector $\mathbf{v}'$. In the special case where $\mathbf{v}' = \lambda\mathbf{v}$, i.e.,

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

where λ is some scalar, the vector $\mathbf{v}$ is called an *eigenvector* of the transformation and λ is the corresponding *eigenvalue*.

Finally, let us review the *Gram-Schmidt procedure* for constructing a set of orthonormal vectors from a set of n -dimensional vectors $\mathbf{v}_i$, $1 \leq i \leq m$. We begin by arbitrarily selecting a vector from the set, say, $\mathbf{v}_1$. By normalizing its length, we obtain the first vector, say,

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \quad (2.2-10)$$

Next, we may select $\mathbf{v}_2$ and, first, subtract the projection of $\mathbf{v}_2$ onto $\mathbf{u}_1$. Thus, we obtain

$$\mathbf{u}'_2 = \mathbf{v}_2 - (\langle \mathbf{v}_2, \mathbf{u}_1 \rangle) \mathbf{u}_1 \quad (2.2-11)$$

Then we normalize the vector $\mathbf{u}'_2$ to unit length. This yields

$$\mathbf{u}_2 = \frac{\mathbf{u}'_2}{\|\mathbf{u}'_2\|} \quad (2.2-12)$$

The procedure continues by selecting $\mathbf{v}_3$ and subtracting the projections of $\mathbf{v}_3$ into $\mathbf{u}_1$ and $\mathbf{u}_2$. Thus, we have

$$\mathbf{u}'_3 = \mathbf{v}_3 - (\langle \mathbf{v}_3, \mathbf{u}_1 \rangle) \mathbf{u}_1 - (\langle \mathbf{v}_3, \mathbf{u}_2 \rangle) \mathbf{u}_2 \quad (2.2-13)$$

Then the orthonormal vector $\mathbf{u}_3$ is

$$\mathbf{u}_3 = \frac{\mathbf{u}'_3}{\|\mathbf{u}'_3\|} \quad (2.2-14)$$

By continuing this procedure, we construct a set of N orthonormal vectors, where $N \leq \min(m, n)$.

2.2–2 Signal Space Concepts

As in the case of vectors, we may develop a parallel treatment for a set of signals. The *inner product* of two generally complex-valued signals $x_1(t)$ and $x_2(t)$ is denoted by $\langle x_1(t), x_2(t) \rangle$ and defined as

$$\langle x_1(t), x_2(t) \rangle = \int_{-\infty}^{\infty} x_1(t) x_2^*(t) dt \quad (2.2-15)$$

similar to Equation 2.1–22. The signals are *orthogonal* if their inner product is zero. The *norm of a signal* is defined as

$$\|x(t)\| = \left(\int_{-\infty}^{\infty} |x(t)|^2 dt \right)^{1/2} = \sqrt{\mathcal{E}_x} \quad (2.2-16)$$

where $\mathcal{E}_x$ is the energy in $x(t)$. A set of m signals is *orthonormal* if they are orthogonal and their norms are all unity. A set of m signals is *linearly independent* if no signal can be represented as a linear combination of the remaining signals. The *triangle inequality* for two signals is simply

$$\|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\| \quad (2.2-17)$$

and the Cauchy–Schwarz inequality is

$$|\langle x_1(t), x_2(t) \rangle| \leq \|x_1(t)\| \cdot \|x_2(t)\| = \sqrt{\mathcal{E}_{x_1} \mathcal{E}_{x_2}} \quad (2.2-18)$$

or, equivalently,

$$\left| \int_{-\infty}^{\infty} x_1(t) x_2^*(t) dt \right| \leq \left| \int_{-\infty}^{\infty} |x_1(t)|^2 dt \right|^{1/2} \left| \int_{-\infty}^{\infty} |x_2(t)|^2 dt \right|^{1/2} \quad (2.2-19)$$

with equality when $x_2(t) = ax_1(t)$, where a is any complex number.

2.2–3 Orthogonal Expansions of Signals

In this section, we develop a vector representation for signal waveforms, and thus we demonstrate an equivalence between a signal waveform and its vector representation. Suppose that $s(t)$ is a deterministic signal with finite energy

$$\mathcal{E}_s = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (2.2-20)$$

Furthermore, suppose that there exists a set of functions $\{\phi_n(t), n = 1, 2, \dots, K\}$ that are orthonormal in the sense that

$$\langle \phi_n(t), \phi_m(t) \rangle = \int_{-\infty}^{\infty} \phi_n(t) \phi_m^*(t) dt = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \quad (2.2-21)$$

We may approximate the signal $s(t)$ by a weighted linear combination of these functions, i.e.,

$$\hat{s}(t) = \sum_{k=1}^K s_k \phi_k(t) \quad (2.2-22)$$

where $\{s_k, 1 \leq k \leq K\}$ are the coefficients in the approximation of $s(t)$. The approximation error incurred is

$$e(t) = s(t) - \hat{s}(t)$$

Let us select the coefficients $\{s_k\}$ so as to minimize the energy $\mathcal{E}_e$ of the approximation error. Thus,

$$\mathcal{E}_e = \int_{-\infty}^{\infty} |s(t) - \hat{s}(t)|^2 dt \quad (2.2-23)$$

$$= \int_{-\infty}^{\infty} \left| s(t) - \sum_{k=1}^K s_k \phi_k(t) \right|^2 dt \quad (2.2-24)$$

The optimum coefficients in the series expansion of $s(t)$ may be found by differentiating Equation 2.2-23 with respect to each of the coefficients $\{s_k\}$ and setting the first derivatives to zero. Alternatively, we may use a well-known result from estimation theory based on the mean square error criterion, which, simply stated, is that the minimum of $\mathcal{E}_e$ with respect to the $\{s_k\}$ is obtained when the error is orthogonal to each of the functions in the series expansion. Thus,

$$\int_{-\infty}^{\infty} \left[s(t) - \sum_{k=1}^K s_k \phi_k(t) \right] \phi_n^*(t) dt = 0, \quad n = 1, 2, \dots, K \quad (2.2-25)$$

Since the functions $\{\phi_n(t)\}$ are orthonormal, Equation 2.2-25 reduces to

$$s_n = \langle s(t), \phi_n(t) \rangle = \int_{-\infty}^{\infty} s(t) \phi_n^*(t) dt, \quad n = 1, 2, \dots, K \quad (2.2-26)$$

Thus, the coefficients are obtained by projecting the signal $s(t)$ onto each of the functions $\{\phi_n(t)\}$. Consequently, $\hat{s}(t)$ is the projection of $s(t)$ onto the K -dimensional signal space spanned by the functions $\{\phi_n(t)\}$, and therefore it is orthogonal to the error signal $e(t) = s(t) - \hat{s}(t)$, i.e., $\langle e(t), \hat{s}(t) \rangle = 0$. The minimum mean-square approximation error is

$$\mathcal{E}_{\min} = \int_{-\infty}^{\infty} e(t) s^*(t) dt \quad (2.2-27)$$

$$= \int_{-\infty}^{\infty} |s(t)|^2 dt - \int_{-\infty}^{\infty} \sum_{k=1}^K s_k \phi_k(t) s^*(t) dt \quad (2.2-28)$$

$$= \mathcal{E}_s - \sum_{k=1}^K |s_k|^2 \quad (2.2-29)$$

which is nonnegative, by definition. When the minimum mean square approximation error $\mathcal{E}_{\min} = 0$,

$$\mathcal{E}_s = \sum_{k=1}^K |s_k|^2 = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (2.2-30)$$

Under the condition that $\mathcal{E}_{\min} = 0$, we may express $s(t)$ as

$$s(t) = \sum_{k=1}^K s_k \phi_k(t) \quad (2.2-31)$$

where it is understood that equality of $s(t)$ to its series expansion holds in the sense that the approximation error has zero energy.

When every finite energy signal can be represented by a series expansion of the form in Equation 2.2-31 for which $\mathcal{E}_{\min} = 0$, the set of orthonormal functions $\{\phi_n(t)\}$ is said to be *complete*.

EXAMPLE 2.2-1. TRIGONOMETRIC FOURIER SERIES: Consider a finite energy real signal $s(t)$ that is zero everywhere except in the range $0 \leq t \leq T$ and has a finite number of discontinuities in this interval. Its periodic extension can be represented in a Fourier series as

$$s(t) = \sum_{k=0}^{\infty} \left(a_k \cos \frac{2\pi kt}{T} + b_k \sin \frac{2\pi kt}{T} \right) \quad (2.2-32)$$

where the coefficients $\{a_k, b_k\}$ that minimize the mean square error are given by

$$\begin{aligned} a_0 &= \frac{1}{T} \int_0^T s(t) dt \\ a_k &= \frac{2}{T} \int_0^T s(t) \cos \frac{2\pi kt}{T} dt, \quad k = 1, 2, 3, \dots \\ b_k &= \frac{2}{T} \int_0^T s(t) \sin \frac{2\pi kt}{T} dt, \quad k = 1, 2, 3, \dots \end{aligned} \quad (2.2-33)$$

The set of functions $\{1/\sqrt{T}, \sqrt{2/T} \cos 2\pi kt/T, \sqrt{2/T} \sin 2\pi kt/T\}$ is a complete set for the expansion of periodic signals on the interval $[0, T]$, and, hence, the series expansion results in zero mean square error.

EXAMPLE 2.2-2. EXPONENTIAL FOURIER SERIES: Consider a general finite energy signal $s(t)$ (real or complex) that is zero everywhere except in the range $0 \leq t \leq T$ and has a finite number of discontinuities in this interval. Its periodic extension can be represented in an exponential Fourier series as

$$s(t) = \sum_{n=-\infty}^{\infty} x_n e^{j2\pi \frac{n}{T} t} \quad (2.2-34)$$

where the coefficients $\{x_n\}$ that minimize the mean square error are given by

$$x_n = \frac{1}{T} \int_{-\infty}^{\infty} x(t) e^{-j2\pi \frac{n}{T} t} dt \quad (2.2-35)$$

The set of functions $\{\sqrt{1/T}e^{j2\pi \frac{n}{T}t}\}$ is a complete set for expansion of periodic signals on the interval $[0, T]$, and, hence, the series expansion results in zero mean square error.

2.2–4 Gram-Schmidt Procedure

Now suppose that we have a set of finite energy signal waveforms $\{s_m(t), m = 1, 2, \dots, M\}$ and we wish to construct a set of orthonormal waveforms. The *Gram-Schmidt orthogonalization procedure* allows us to construct such a set. This procedure is similar to the one described in Section 2.2–1 for vectors. We begin with the first waveform $s_1(t)$, which is assumed to have energy $\mathcal{E}_1$. The first orthonormal waveform is simply constructed as

$$\phi_1(t) = \frac{s_1(t)}{\sqrt{\mathcal{E}_1}} \quad (2.2-36)$$

Thus, $\phi_1(t)$ is simply $s_1(t)$ normalized to unit energy. The second waveform is constructed from $s_2(t)$ by first computing the projection of $s_2(t)$ onto $\phi_1(t)$, which is

$$c_{21} = \langle s_2(t), \phi_1(t) \rangle = \int_{-\infty}^{\infty} s_2(t) \phi_1^*(t) dt \quad (2.2-37)$$

Then $c_{21}\phi_1(t)$ is subtracted from $s_2(t)$ to yield

$$\gamma_2(t) = s_2(t) - c_{21}\phi_1(t) \quad (2.2-38)$$

This waveform is orthogonal to $\phi_1(t)$, but it does not have unit energy. If $\mathcal{E}_2$ denotes the energy of $\gamma_2(t)$, i.e.,

$$\mathcal{E}_2 = \int_{-\infty}^{\infty} \gamma_2^2(t) dt$$

the normalized waveform that is orthogonal to $\phi_1(t)$ is

$$\phi_2(t) = \frac{\gamma_2(t)}{\sqrt{\mathcal{E}_2}} \quad (2.2-39)$$

In general, the orthogonalization of the k th function leads to

$$\phi_k(t) = \frac{\gamma_k(t)}{\sqrt{\mathcal{E}_k}} \quad (2.2-40)$$

where

$$\gamma_k(t) = s_k(t) - \sum_{i=1}^{k-1} c_{ki}\phi_i(t) \quad (2.2-41)$$

$$c_{ki} = \langle s_k(t), \phi_i(t) \rangle = \int_{-\infty}^{\infty} s_k(t) \phi_i^*(t) dt, \quad i = 1, 2, \dots, k-1 \quad (2.2-42)$$

$$\mathcal{E}_k = \int_{-\infty}^{\infty} \gamma_k^2(t) dt \quad (2.2-43)$$

Thus, the orthogonalization process is continued until all the M signal waveforms $\{s_m(t)\}$ have been exhausted and $N \leq M$ orthonormal waveforms have been constructed. The dimensionality N of the signal space will be equal to M if all the signal waveforms are linearly independent, i.e., none of the signal waveforms is a linear combination of the other signal waveforms.

EXAMPLE 2.2-3. Let us apply the Gram-Schmidt procedure to the set of four waveforms illustrated in Figure 2.2-1. The waveform $s_1(t)$ has energy $\mathcal{E}_1 = 2$, so that

$$\phi_1(t) = \sqrt{\frac{1}{2}} s_1(t)$$

Next we observe that $c_{21} = 0$; hence, $s_2(t)$ and $\phi_1(t)$ are orthogonal. Therefore, $\phi_2(t) = s_2(t)/\sqrt{\mathcal{E}_2} = \sqrt{\frac{1}{2}} s_2(t)$. To obtain $\phi_3(t)$, we compute c_{31} and c_{32} , which are $c_{31} = \sqrt{2}$ and $c_{32} = 0$. Thus,

$$\gamma_3(t) = s_3(t) - \sqrt{2}\phi_1(t) = \begin{cases} -1 & 2 \leq t \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

Since $\gamma_3(t)$ has unit energy, it follows that $\phi_3(t) = \gamma_3(t)$. Determining $\phi_4(t)$, we find that $c_{41} = -\sqrt{2}$, $c_{42} = 0$, and $c_{43} = 1$. Hence,

$$\gamma_4(t) = s_4(t) + \sqrt{2}\phi_1(t) - \phi_3(t) = 0$$

Consequently, $s_4(t)$ is a linear combination of $\phi_1(t)$ and $\phi_3(t)$ and, hence, $\phi_4(t) = 0$. The three orthonormal functions are illustrated in Figure 2.2-1(b).

Once we have constructed the set of orthonormal waveforms $\{\phi_n(t)\}$, we can express the M signals $\{s_m(t)\}$ as linear combinations of the $\{\phi_n(t)\}$. Thus, we may write

$$s_m(t) = \sum_{n=1}^N s_{mn} \phi_n(t), \quad m = 1, 2, \dots, M \quad (2.2-44)$$

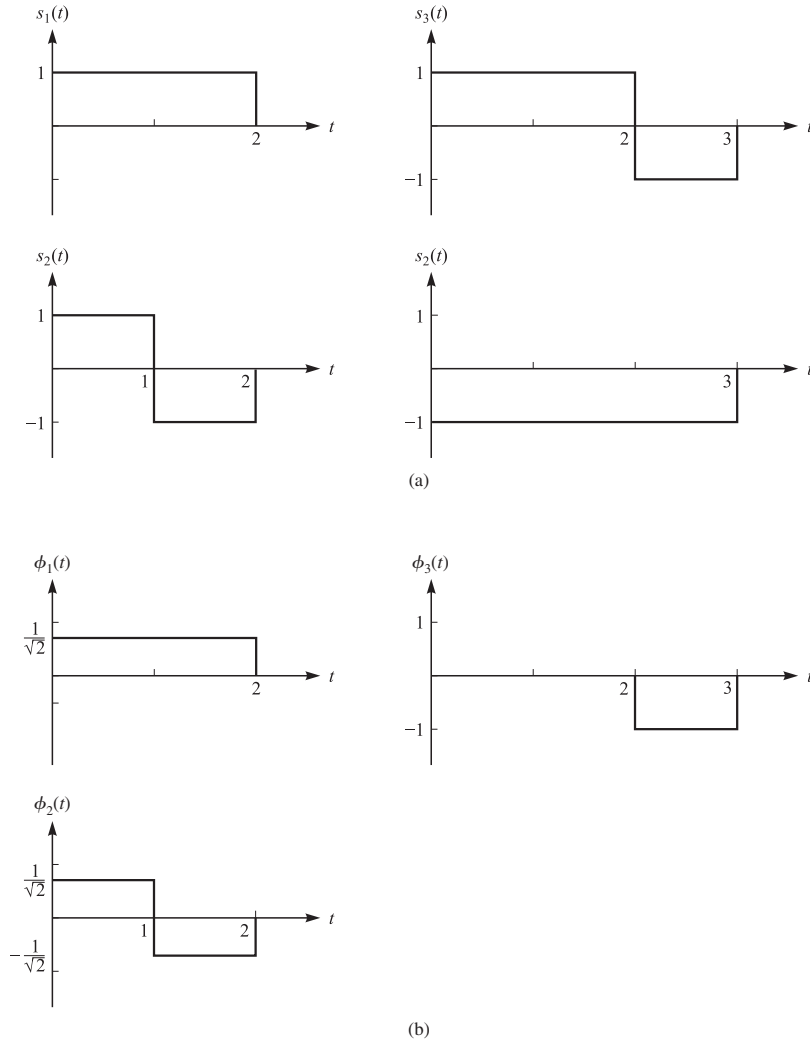
Based on the expression in Equation 2.2-44, each signal may be represented by the vector

$$\mathbf{s}_m = [s_{m1} \ s_{m2} \ \cdots \ s_{mN}]^t \quad (2.2-45)$$

or, equivalently, as a point in the N -dimensional (in general, complex) signal space with coordinates $\{s_{mn}, n = 1, 2, \dots, N\}$. Therefore, a set of M signals $\{s_m(t)\}_{m=1}^M$ can be represented by a set of M vectors $\{\mathbf{s}_m\}_{m=1}^M$ in the N -dimensional space, where $N \leq M$. The corresponding set of vectors is called the *signal space representation*, or *constellation*, of $\{s_m(t)\}_{m=1}^M$. If the original signals are real, then the corresponding vector representations are in $\mathbb{R}^N$; and if the signals are complex, then the vector representations are in $\mathbb{C}^N$. Figure 2.2-2 demonstrates the process of obtaining the vector equivalent from a signal (signal-to-vector mapping) and vice versa (vector-to-signal mapping).

From the orthonormality of the basis $\{\phi_n(t)\}$ it follows that

$$\mathcal{E}_m = \int_{-\infty}^{\infty} |s_m(t)|^2 dt = \sum_{n=1}^N |s_{mn}|^2 = \|\mathbf{s}_m\|^2 \quad (2.2-46)$$

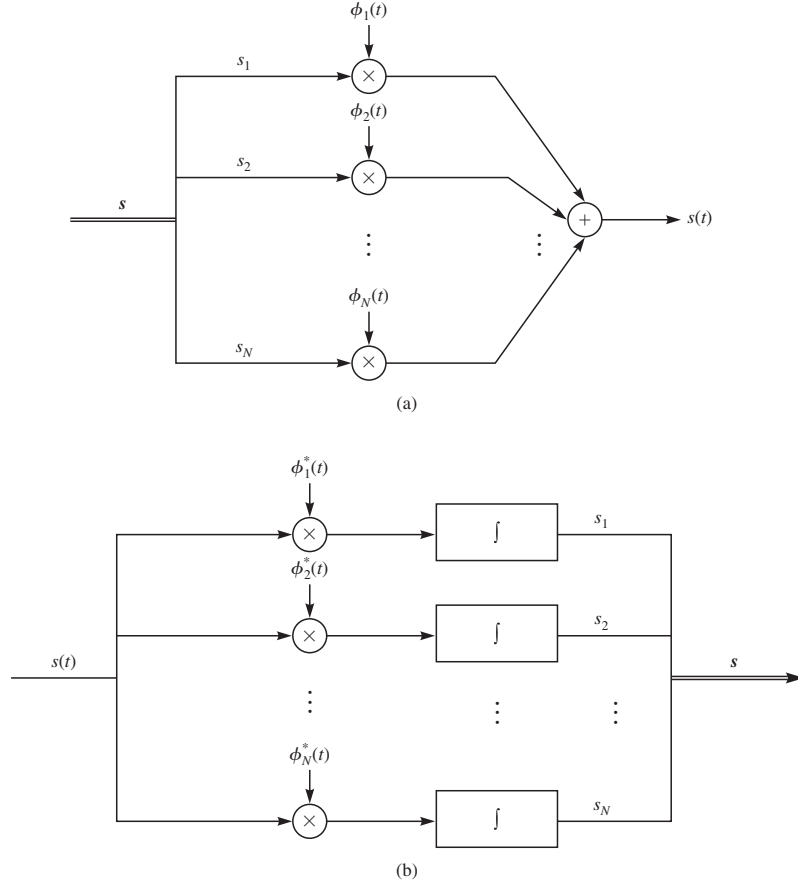
**FIGURE 2.2-1**

Gram-Schmidt orthogonalization of the signal $\{s_m(t), m = 1, 2, 3, 4\}$ and the corresponding orthonormal basis.

The energy in the m th signal is simply the square of the length of the vector or, equivalently, the square of the Euclidean distance from the origin to the point s_m in the N -dimensional space. Thus, any signal can be represented geometrically as a point in the signal space spanned by the orthonormal functions $\{\phi_n(t)\}$. From the orthonormality of the basis it also follows that

$$\langle s_k(t), s_l(t) \rangle = \langle s_k, s_l \rangle \quad (2.2-47)$$

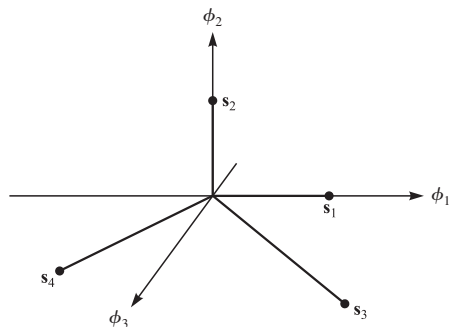
This shows that the inner product of two signals is equal to the inner product of the corresponding vectors.

**FIGURE 2.2-2**

Vector to signal (a), and signal to vector (b) mappings.

EXAMPLE 2.2-4. Let us obtain the vector representation of the four signals shown in Figure 2.2-1(a) by using the orthonormal set of functions in Figure 2.2-1(b). Since the dimensionality of the signal space is $N = 3$, each signal is described by three components. The signal $s_1(t)$ is characterized by the vector $\mathbf{s}_1 = (\sqrt{2}, 0, 0)^t$. Similarly, the signals $s_2(t)$, $s_3(t)$, and $s_4(t)$ are characterized by the vectors $\mathbf{s}_2 = (0, \sqrt{2}, 0)^t$, $\mathbf{s}_3 = (\sqrt{2}, 0, 1)^t$, and $\mathbf{s}_4 = (-\sqrt{2}, 0, 1)^t$, respectively. These vectors are shown in Figure 2.2-3. Their lengths are $\|\mathbf{s}_1\| = \sqrt{2}$, $\|\mathbf{s}_2\| = \sqrt{2}$, $\|\mathbf{s}_3\| = \sqrt{3}$, and $\|\mathbf{s}_4\| = \sqrt{3}$, and the corresponding signal energies are $\mathcal{E}_k = \|\mathbf{s}_k\|^2$, $k = 1, 2, 3, 4$.

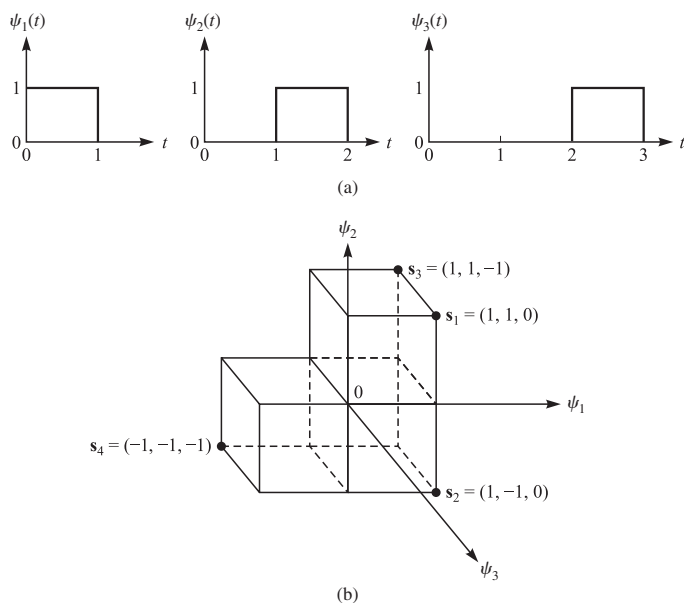
We have demonstrated that a set of M finite energy waveforms $\{s_m(t)\}$ can be represented by a weighted linear combination of orthonormal functions $\{\phi_n(t)\}$ of dimensionality $N \leq M$. The functions $\{\phi_n(t)\}$ are obtained by applying the Gram-Schmidt orthogonalization procedure on $\{s_m(t)\}$. It should be emphasized, however, that the functions $\{\phi_n(t)\}$ obtained from the Gram-Schmidt procedure are not unique. If we

**FIGURE 2.2-3**

The four signal vectors represented as points in three-dimensional space.

alter the order in which the orthogonalization of the signals $\{s_m(t)\}$ is performed, the orthonormal waveforms will be different and the corresponding vector representation of the signals $\{s_m(t)\}$ will depend on the choice of the orthonormal functions $\{\phi_n(t)\}$. Nevertheless, the dimensionality of the signal space N will not change, and the vectors $\{s_m\}$ will retain their geometric configuration; i.e., their lengths and their inner products will be invariant to the choice of the orthonormal functions $\{\phi_n(t)\}$.

EXAMPLE 2.2-5. An alternative set of orthonormal functions for the four signals in Figure 2.2-1(a) is illustrated in Figure 2.2-4(a). By using these functions to expand $\{s_n(t)\}$, we obtain the corresponding vectors $s_1 = (1, 1, 0)^t$, $s_2 = (1, -1, 0)^t$, $s_3 = (1, 1, -1)^t$, and $s_4 = (-1, -1, -1)^t$, which are shown in Figure 2.2-4(b). Note that the vector lengths are identical to those obtained from the orthonormal functions $\{\phi_n(t)\}$.

**FIGURE 2.2-4**

An alternative set of orthonormal functions for the four signals in Figure 2.2-1(a) and the corresponding signal points.

Bandpass and Lowpass Orthonormal Basis

Let us consider the case in which the signal waveforms are bandpass and represented as

$$s_m(t) = \text{Re} [s_{ml}(t)e^{j2\pi f_0 t}], \quad m = 1, 2, \dots, M \quad (2.2-48)$$

where $\{s_{ml}(t)\}$ denotes the lowpass equivalent signals. Recall from Section 2.1–1 that if two lowpass equivalent signals are orthogonal, the corresponding bandpass signals are orthogonal too. Therefore, if $\{\phi_{nl}(t), n = 1, \dots, N\}$ constitutes an orthonormal basis for the set of lowpass signals $\{s_{ml}(t)\}$, then the set $\{\phi_n(t), n = 1, \dots, N\}$ where

$$\phi_n(t) = \sqrt{2} \text{Re} [\phi_{nl}(t)e^{j2\pi f_0 t}] \quad (2.2-49)$$

is a set of orthonormal signals, where $\sqrt{2}$ is a normalization factor to make sure each $\phi_n(t)$ has unit energy. However, this set is not necessarily an orthonormal *basis* for expansion of $\{s_m(t), m = 1, \dots, M\}$. In other words, there is no guarantee that this set is a complete basis for expansion of the set of signals $\{s_m(t), m = 1, \dots, M\}$. Here our goal is to see how an orthonormal basis for representation of bandpass signals can be obtained from an orthonormal basis used for representation of the lowpass equivalents of the bandpass signals.

Since we have

$$s_{ml}(t) = \sum_{n=1}^N s_{mln} \phi_{nl}(t), \quad m = 1, \dots, M \quad (2.2-50)$$

where

$$s_{mln} = \langle s_{ml}(t), \phi_{nl}(t) \rangle, \quad m = 1, \dots, M, \quad n = 1, \dots, N \quad (2.2-51)$$

from Equations 2.2–48 and 2.2–50 we can write

$$s_m(t) = \text{Re} \left[\left(\sum_{n=1}^N s_{mln} \phi_{nl}(t) \right) e^{j2\pi f_0 t} \right], \quad m = 1, \dots, M \quad (2.2-52)$$

or

$$s_m(t) = \text{Re} \left[\sum_{n=1}^N s_{mln} \phi_{nl}(t) \right] \cos 2\pi f_0 t - \text{Im} \left[\sum_{n=1}^N s_{mln} \phi_{nl}(t) \right] \sin 2\pi f_0 t \quad (2.2-53)$$

In Problem 2.6 we will see that when an orthonormal set of signals $\{\phi_{nl}(t), n = 1, \dots, N\}$ constitutes an N -dimensional complex basis for representation of $\{s_{ml}(t), m = 1, \dots, M\}$, then the set $\{\phi_n(t), \tilde{\phi}_n(t), n = 1, \dots, N\}$, where

$$\begin{aligned} \phi_n(t) &= \sqrt{2} \text{Re} [\phi_{nl}(t)e^{j2\pi f_0 t}] = \sqrt{2}\phi_{ni}(t) \cos 2\pi f_0 t - \sqrt{2}\phi_{nq}(t) \sin 2\pi f_0 t \\ \tilde{\phi}_n(t) &= -\sqrt{2} \text{Im} [\phi_{nl}(t)e^{j2\pi f_0 t}] = -\sqrt{2}\phi_{ni}(t) \sin 2\pi f_0 t - \sqrt{2}\phi_{nq}(t) \cos 2\pi f_0 t \end{aligned} \quad (2.2-54)$$

constitutes a $2N$ -dimensional orthonormal basis that is *sufficient* for representation of M bandpass signals

$$s_m(t) = \text{Re} [s_{ml}(t)e^{j2\pi f_0 t}], \quad m = 1, \dots, M \quad (2.2-55)$$

In some cases not all basis functions in the set of basis given by Equation 2.2–54 are necessary, and only a subset of them would be sufficient to expand the bandpass signals. In Problem 2.7 we will further show that

$$\tilde{\phi}(t) = -\hat{\phi}(t) \quad (2.2-56)$$

where $\hat{\phi}(t)$ denotes the Hilbert transform of $\phi(t)$.

From Equation 2.2–52 we have

$$\begin{aligned} s_m(t) &= \text{Re} \left[\left(\sum_{n=1}^N s_{mln} \phi_{nl}(t) \right) e^{j2\pi f_0 t} \right] \\ &= \sum_{n=1}^N \text{Re} \left[(s_{mln} \phi_{nl}(t)) e^{j2\pi f_0 t} \right] \\ &= \sum_{n=1}^N \left[\frac{s_{mln}^{(r)}}{\sqrt{2}} \phi_n(t) + \frac{s_{mln}^{(i)}}{\sqrt{2}} \tilde{\phi}_n(t) \right] \end{aligned} \quad (2.2-57)$$

where we have assumed that $s_{mln} = s_{mln}^{(r)} + js_{mln}^{(i)}$. Equations 2.2–54 and 2.2–57 show how a bandpass signal can be expanded in terms of the basis used for expansion of its lowpass equivalent. In general, lowpass signals can be represented by an N -dimensional complex vector, and the corresponding bandpass signal can be represented by $2N$ -dimensional real vectors. If the complex vector

$$s_{ml} = (s_{ml1}, s_{ml2}, \dots, s_{mlN})^t$$

is a vector representation for the lowpass signal $s_{ml}(t)$ using the lowpass basis $\{\phi_{nl}(t), n = 1, \dots, N\}$, then the vector

$$s_m = \left(\frac{s_{ml1}^{(r)}}{\sqrt{2}}, \frac{s_{ml2}^{(r)}}{\sqrt{2}}, \dots, \frac{s_{mlN}^{(r)}}{\sqrt{2}}, \frac{s_{ml1}^{(i)}}{\sqrt{2}}, \frac{s_{ml2}^{(i)}}{\sqrt{2}}, \dots, \frac{s_{mlN}^{(i)}}{\sqrt{2}} \right)^t \quad (2.2-58)$$

is a vector representation of the bandpass signal

$$s_m(t) = \text{Re} [s_{ml}(t) e^{j2\pi f_0 t}]$$

when the bandpass basis $\{\phi_n(t), \tilde{\phi}_n(t), n = 1, \dots, N\}$ given by Equations 2.2–54 and 2.2–57 is used.

EXAMPLE 2.2–6. Let us assume M bandpass signals are defined by

$$s_m(t) = \text{Re} [A_m g(t) e^{j2\pi f_0 t}] \quad (2.2-59)$$

where A_m 's are arbitrary complex numbers and $g(t)$ is a real lowpass signal with energy $\mathcal{E}_g$. The lowpass equivalent signals are given by

$$s_{ml}(t) = A_m g(t)$$

and therefore the unit-energy signal $\phi(t)$ defined by

$$\phi(t) = \frac{g(t)}{\sqrt{\mathcal{E}_g}}$$

is sufficient to expand all $s_{ml}(t)$'s.

We have

$$s_{ml}(t) = A_m \sqrt{\mathcal{E}_g} \phi(t)$$

thus, corresponding to each $s_{ml}(t)$ we have a single complex scalar $A_m \sqrt{\mathcal{E}_g} = (A_m^{(r)} + jA_m^{(i)}) \sqrt{\mathcal{E}_g}$; i.e., the lowpass signals constitute one complex dimension (or, equivalently, two real dimensions). From Equation 2.2–54 we conclude that

$$\begin{aligned}\phi(t) &= \sqrt{\frac{2}{\mathcal{E}_g}} g(t) \cos 2\pi f_0 t \\ \tilde{\phi}(t) &= -\sqrt{\frac{2}{\mathcal{E}_g}} g(t) \sin 2\pi f_0 t\end{aligned}$$

can be used as a basis for expansion of the bandpass signals.

Using this basis and Equation 2.2–57, we have

$$\begin{aligned}s_m(t) &= A_m^{(r)} \sqrt{\frac{\mathcal{E}_g}{2}} \phi(t) + A_m^{(i)} \sqrt{\frac{\mathcal{E}_g}{2}} \tilde{\phi}(t) \\ &= A_m^{(r)} g(t) \cos 2\pi f_0 t - A_m^{(i)} g(t) \sin 2\pi f_0 t\end{aligned}$$

which agrees with the straightforward expansion of Equation 2.2–59. Note that in the special case where all A_m 's are real, $\phi(t)$ is sufficient to represent the bandpass signals and $\tilde{\phi}(t)$ is not necessary.

2.3

SOME USEFUL RANDOM VARIABLES

In subsequent chapters, we shall encounter several different types of random variables. In this section we list these frequently encountered random variables, their probability density functions (PDFs), their cumulative distribution functions (CDFs), and their moments. Our main emphasis will be on the Gaussian random variable and many random variables that are derived from the Gaussian random variable.

The Bernoulli Random Variable

The Bernoulli random variable is a discrete binary-valued random variable taking values 1 and 0 with probabilities p and $1 - p$, respectively. Therefore the probability mass function (PMF) for this random variable is given by

$$P[X = 1] = p \quad P[X = 0] = 1 - p \quad (2.3-1)$$

The mean and variance of this random variable are given by

$$\begin{aligned}E[X] &= p \\ \text{VAR}[X] &= p(1 - p)\end{aligned} \quad (2.3-2)$$

The Binomial Random Variable

The binomial random variable models the sum of n independent Bernoulli random variables with common parameter p . The PMF of this random variable is given by

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n \quad (2.3-3)$$

For this random variable we have

$$\begin{aligned} E[X] &= np \\ \text{VAR}[X] &= np(1 - p) \end{aligned} \quad (2.3-4)$$

This random variable models, for instance, the number of errors when n bits are transmitted over a communication channel and the probability of error for each bit is p .

The Uniform Random Variable

The uniform random variable is a continuous random variable with PDF

$$p(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \quad (2.3-5)$$

where $b > a$ and the interval $[a, b]$ is the range of the random variable. Here we have

$$E[X] = \frac{b + a}{2} \quad (2.3-6)$$

$$\text{VAR}[X] = \frac{(b - a)^2}{12} \quad (2.3-7)$$

The Gaussian (Normal) Random Variable

The Gaussian random variable is described in terms of two parameters $m \in \mathbb{R}$ and $\sigma > 0$ by the PDF

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (2.3-8)$$

We usually use the shorthand form $\mathcal{N}(m, \sigma^2)$ to denote the PDF of Gaussian random variables and write $X \sim \mathcal{N}(m, \sigma^2)$. For this random variable

$$\begin{aligned} E[X] &= m \\ \text{VAR}[X] &= \sigma^2 \end{aligned} \quad (2.3-9)$$

A Gaussian random variable with $m = 0$ and $\sigma = 1$ is called a *standard normal*. A function closely related to the Gaussian random variable is the Q function defined as

$$Q(x) = P[\mathcal{N}(0, 1) > x] = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt \quad (2.3-10)$$

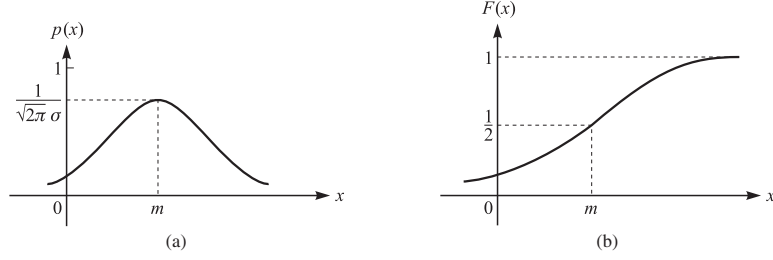


FIGURE 2.3-1
PDF and CDF of a Gaussian random variable.

The CDF of a Gaussian random variable is given by

$$\begin{aligned}
 F(x) &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \\
 &= 1 - \int_x^{\infty} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \\
 &= 1 - \int_{\frac{x-m}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \\
 &= 1 - Q\left(\frac{x-m}{\sigma}\right)
 \end{aligned} \tag{2.3-11}$$

where we have introduced the change of variable $u = (t - m)/\sigma$. The PDF and the CDF of a Gaussian random variable are shown in Figure 2.3-1.

In general if $X \sim \mathcal{N}(m, \sigma^2)$, then

$$\begin{aligned}
 P[X > \alpha] &= Q\left(\frac{\alpha - m}{\sigma}\right) \\
 P[X < \alpha] &= Q\left(\frac{m - \alpha}{\sigma}\right)
 \end{aligned} \tag{2.3-12}$$

Following are some of the important properties of the Q function:

$$Q(0) = \frac{1}{2} \quad Q(\infty) = 0 \tag{2.3-13}$$

$$Q(-\infty) = 1 \quad Q(-x) = 1 - Q(x) \tag{2.3-14}$$

Some useful bounds for the Q function for $x > 0$ are

$$\begin{aligned}
 Q(x) &\leq \frac{1}{2} e^{-\frac{x^2}{2}} \\
 Q(x) &< \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\
 Q(x) &> \frac{x}{(1+x^2)\sqrt{2\pi}} e^{-\frac{x^2}{2}}
 \end{aligned} \tag{2.3-15}$$

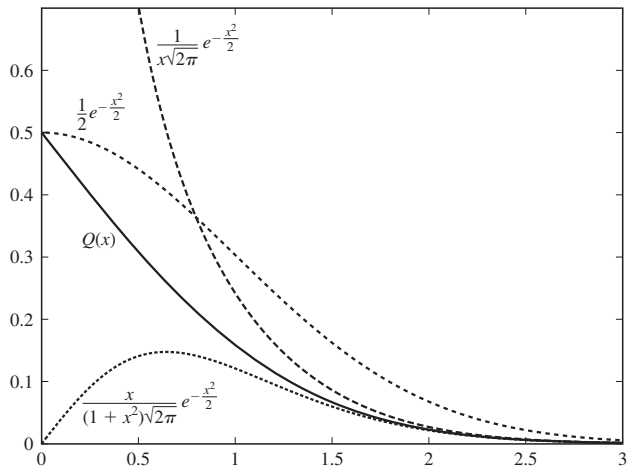


FIGURE 2.3-2
Plot of $Q(x)$ and its upper and lower bounds.

From the last two bounds we conclude that for large x we have

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.3-16)$$

A plot of the Q function bounds is given in Figure 2.3-2. Tables 2.3-1 and 2.3-2 give values of the Q function.

TABLE 2.3-1
Table of Q Function Values

x	$Q(x)$	x	$Q(x)$	x	$Q(x)$	x	$Q(x)$
0	0.500000	1.8	0.035930	3.6	0.000159	5.4	3.3320×10^{-8}
0.1	0.460170	1.9	0.028717	3.7	0.000108	5.5	1.8990×10^{-8}
0.2	0.420740	2	0.022750	3.8	7.2348×10^{-5}	5.6	1.0718×10^{-8}
0.3	0.382090	2.1	0.017864	3.9	4.8096×10^{-5}	5.7	5.9904×10^{-9}
0.4	0.344580	2.2	0.013903	4	3.1671×10^{-5}	5.8	3.3157×10^{-9}
0.5	0.308540	2.3	0.010724	4.1	2.0658×10^{-5}	5.9	1.8175×10^{-9}
0.6	0.274250	2.4	0.008198	4.2	1.3346×10^{-5}	6	9.8659×10^{-10}
0.7	0.241960	2.5	0.006210	4.3	8.5399×10^{-6}	6.1	5.3034×10^{-10}
0.8	0.211860	2.6	0.004661	4.4	5.4125×10^{-6}	6.2	2.8232×10^{-10}
0.9	0.184060	2.7	0.003467	4.5	3.3977×10^{-6}	6.3	1.4882×10^{-10}
1	0.158660	2.8	0.002555	4.6	2.1125×10^{-6}	6.4	7.7689×10^{-11}
1.1	0.135670	2.9	0.001866	4.7	1.3008×10^{-6}	6.5	4.0160×10^{-11}
1.2	0.115070	3	0.001350	4.8	7.9333×10^{-7}	6.6	2.0558×10^{-11}
1.3	0.096800	3.1	0.000968	4.9	4.7918×10^{-7}	6.7	1.0421×10^{-11}
1.4	0.080757	3.2	0.000687	5	2.8665×10^{-7}	6.8	5.2309×10^{-12}
1.5	0.066807	3.3	0.000483	5.1	1.6983×10^{-7}	6.9	2.6001×10^{-12}
1.6	0.054799	3.4	0.000337	5.2	9.9644×10^{-8}	7	1.2799×10^{-12}
1.7	0.044565	3.5	0.000233	5.3	5.7901×10^{-8}	7.1	6.2378×10^{-13}

■ TABLE 2.3-2
Selected Q Function
Values

$Q(x)$	x
10^{-1}	1.2816
10^{-2}	2.3263
10^{-3}	3.0902
10^{-4}	3.7190
10^{-5}	4.2649
10^{-6}	4.7534
10^{-7}	5.1993
0.5×10^{-5}	4.4172
0.25×10^{-5}	4.5648
0.667×10^{-5}	4.3545

Another function closely related to the Q function is the *complementary error function*, defined as

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad (2.3-17)$$

The complementary error function is related to the Q function as follows:

$$\begin{aligned} Q(x) &= \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \\ \operatorname{erfc}(x) &= 2Q(\sqrt{2}x) \end{aligned} \quad (2.3-18)$$

The characteristic function[†] of a Gaussian random variable is given by

$$\Phi_X(\omega) = e^{j\omega m - \frac{1}{2}\omega^2 \sigma^2} \quad (2.3-19)$$

Problem 2.21 shows that for an $\mathcal{N}(m, \sigma^2)$ random variable we have

$$E[(X - m)^n] = \begin{cases} 1 \times 3 \times 5 \times \cdots \times (2k - 1) \sigma^{2k} = \frac{(2k)! \sigma^{2k}}{2^k k!} & \text{for } n = 2k \\ 0 & \text{for } n = 2k + 1 \end{cases} \quad (2.3-20)$$

from which we can obtain moments of the Gaussian random variable.

The sum of n independent Gaussian random variables is a Gaussian random variable whose mean and variance are the sum of the means and the sum of the variances of the random variables, respectively.

[†]Recall that for any random variable X , the *characteristic function* is defined by $\Phi_X(\omega) = E[e^{j\omega X}]$. The *moment generating function* (MGF) is defined by $\Theta_X(t) = E[e^{tX}]$. Obviously, $\Theta(t) = \Phi(-jt)$ and $\Phi(\omega) = \Theta(j\omega)$.

The Chi-Square (χ^2) Random Variable

If $\{X_i, i = 1, \dots, n\}$ are iid (independent and identically distributed) zero-mean Gaussian random variables with common variance σ^2 and we define

$$X = \sum_{i=1}^n X_i^2$$

then X is a χ^2 random variable with n degrees of freedom. The PDF of this random variable is given by

$$p(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(\frac{n}{2})\sigma^n} x^{\frac{n}{2}-1} e^{-\frac{x}{2\sigma^2}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-21)$$

where $\Gamma(x)$ is the *gamma function* defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad (2.3-22)$$

The gamma function has simple poles at $x = 0, -1, -2, -3, \dots$ and satisfies the following properties. The gamma function can be thought of as a generalization of the notion of factorial.

$$\Gamma(x+1) = x\Gamma(x),$$

$$\Gamma(1) = 1$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad (2.3-23)$$

$$\Gamma\left(\frac{n}{2} + 1\right) = \begin{cases} \left(\frac{n}{2}\right)! & n \text{ even and positive} \\ \sqrt{\pi} \frac{n(n-2)(n-4)\dots 3 \times 1}{2^{\frac{n+1}{2}}} & n \text{ odd and positive} \end{cases}$$

When n is even, i.e., $n = 2m$, the CDF of the χ^2 random variable with n degrees of freedom has a closed form given by

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{2\sigma^2}} \sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{x}{2\sigma^2}\right)^k & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-24)$$

The mean and variance of a χ^2 random variable with n degrees of freedom are given by

$$\begin{aligned} E[X] &= n\sigma^2 \\ \text{VAR}[X] &= 2n\sigma^4 \end{aligned} \quad (2.3-25)$$

The characteristic function for a χ^2 random variable with n degrees of freedom is given by

$$\Phi(\omega) = \left(\frac{1}{1 - 2j\omega\sigma^2} \right)^{\frac{n}{2}} \quad (2.3-26)$$

The special case of a χ^2 random variable with two degrees of freedom is of particular interest. In this case the PDF is given by

$$p(x) = \begin{cases} \frac{1}{2\sigma^2} e^{-\frac{x}{2\sigma^2}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-27)$$

This is the PDF of an *exponential random variable* with mean equal to $2\sigma^2$.

The χ^2 random variable is a special case of a *gamma random variable*. A gamma random variable is defined by a PDF of the form

$$p(x) = \begin{cases} \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-28)$$

where $\lambda, \alpha > 0$. A χ^2 random variable is a gamma random variable with $\lambda = \frac{1}{2\sigma^2}$ and $\alpha = \frac{n}{2}$.

Plots of the χ^2 random variable with n degrees of freedom for different values of n are shown in Figure 2.3–3.

The Noncentral Chi-Square (χ^2) Random Variable

The *noncentral χ^2 random variable* with n degrees of freedom is defined similarly to a χ^2 random variable in which X_i 's are independent Gaussians with common variance σ^2 but with different means denoted by m_i . This random variable has a PDF of the form

$$p(x) = \begin{cases} \frac{1}{2\sigma^2} \left(\frac{x}{s^2}\right)^{\frac{n-2}{4}} e^{-\frac{s^2+x}{2\sigma^2}} I_{\frac{n}{2}-1} \left(\frac{s}{\sigma^2} \sqrt{x}\right) & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-29)$$

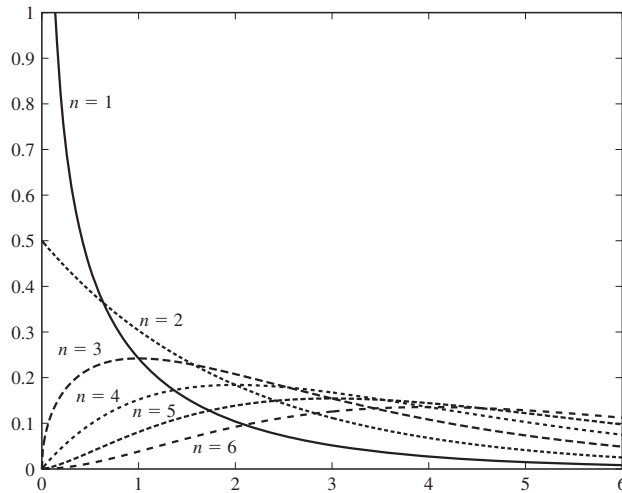


FIGURE 2.3-3

The PDF of the χ^2 random variable for different values of n . All plots are shown for $\sigma = 1$.

where s is defined as

$$s = \sqrt{\sum_{i=1}^n m_i^2} \quad (2.3-30)$$

and $I_\alpha(x)$ is the *modified Bessel function of the first kind and order α* given by

$$I_\alpha(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{\alpha+2k}}{k! \Gamma(\alpha + k + 1)}, \quad x \geq 0 \quad (2.3-31)$$

where $\Gamma(x)$ is the gamma function defined by Equation 2.3-22. The function $I_0(x)$ can be written as

$$I_0(x) = \sum_{k=0}^{\infty} \left(\frac{x^k}{2^k k!} \right)^2 \quad (2.3-32)$$

and for $x > 1$ can be approximated by

$$I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}} \quad (2.3-33)$$

Two other expressions for $I_0(x)$, which are used frequently, are

$$\begin{aligned} I_0(x) &= \frac{1}{\pi} \int_0^\pi e^{\pm x \cos \phi} d\phi \\ I_0(x) &= \frac{1}{2\pi} \int_0^{2\pi} e^{x \cos \phi} d\phi \end{aligned} \quad (2.3-34)$$

The CDF of this random variable, when $n = 2m$, can be written in the form

$$F(x) = \begin{cases} 1 - Q_m\left(\frac{s}{\sigma}, \frac{\sqrt{x}}{\sigma}\right) & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-35)$$

where $Q_m(a, b)$ is the *generalized Marcum Q function* and is defined as

$$\begin{aligned} Q_m(a, b) &= \int_b^\infty x \left(\frac{x}{a} \right)^{m-1} e^{-(x^2+a^2)/2} I_{m-1}(ax) dx \\ &= Q_1(a, b) + e^{-(a^2+b^2)/2} \sum_{k=1}^{m-1} \left(\frac{b}{a} \right)^k I_k(ab) \end{aligned} \quad (2.3-36)$$

In Equation 2.3-36, $Q_1(a, b)$ is the *Marcum Q function* defined as

$$Q_1(a, b) = \int_b^\infty x e^{-\frac{a^2+x^2}{2}} I_0(ax) dx \quad (2.3-37)$$

or

$$Q_1(a, b) = e^{-\frac{a^2+b^2}{2}} \sum_{k=0}^{\infty} \left(\frac{a}{b} \right)^k I_k(ab), \quad b \geq a > 0 \quad (2.3-38)$$

This function satisfies the following properties:

$$\begin{aligned} Q_1(x, 0) &= 1 \\ Q_1(0, x) &= e^{-\frac{x^2}{2}} \\ Q_1(a, b) &\approx Q(b - a) \quad \text{for } b \gg 1 \text{ and } b \gg b - a \end{aligned} \quad (2.3-39)$$

For a noncentral χ^2 random variable, the mean and variance are given by

$$\begin{aligned} E[X] &= n\sigma^2 + s^2 \\ \text{VAR}[X] &= 2n\sigma^4 + 4\sigma^2 s^2 \end{aligned} \quad (2.3-40)$$

and the characteristic function is given by

$$\Phi(\omega) = \left(\frac{1}{1 - 2j\omega\sigma^2} \right)^{\frac{n}{2}} e^{\frac{j\omega s^2}{1 - 2j\omega\sigma^2}} \quad (2.3-41)$$

The Rayleigh Random Variable

If X_1 and X_2 are two iid Gaussian random variables each distributed according to $\mathcal{N}(0, \sigma^2)$, then

$$X = \sqrt{X_1^2 + X_2^2} \quad (2.3-42)$$

is a *Rayleigh random variable*. From our discussion of the χ^2 random variables, it is readily seen that a Rayleigh random variable is the square root of a χ^2 random variable with two degrees of freedom. We can also conclude that the Rayleigh random variable is the square root of an exponential random variable as given by Equation 2.3-27. The PDF of a Rayleigh random variable is given by

$$p(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-43)$$

and its mean and variance are

$$\begin{aligned} E[X] &= \sigma \sqrt{\frac{\pi}{2}} \\ \text{VAR}[X] &= \left(2 - \frac{\pi}{2} \right) \sigma^2 \end{aligned} \quad (2.3-44)$$

In general, the n th moment of a Rayleigh random variable is given by

$$E[X^k] = (2\sigma^2)^{k/2} \Gamma\left(\frac{k}{2} + 1\right) \quad (2.3-45)$$

and its characteristic function is given by

$$\Phi_X(\omega) = {}_1F_1\left(1, \frac{1}{2}; -\frac{1}{2}\omega^2\sigma^2\right) + j\sqrt{\frac{\pi}{2}}\omega\sigma e^{-\frac{\omega^2\sigma^2}{2}} \quad (2.3-46)$$

where ${}_1F_1(a, b; x)$ is the *confluent hypergeometric function* defined by

$${}_1F_1(a, b; x) = \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b)x^k}{\Gamma(a)\Gamma(b+k)k!}, \quad b \neq 0, -1, -2, \dots \quad (2.3-47)$$

The function ${}_1F_1(a, b; x)$ can also be written as the integral

$${}_1F_1(a, b; x) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{xt} t^{a-1} (1-t)^{b-a-1} dt \quad (2.3-48)$$

In Beaulieu (1990), it is shown that

$${}_1F_1\left(1, \frac{1}{2}; -x\right) = -e^{-x} \sum_{k=0}^{\infty} \frac{x^k}{(2k-1)k!} \quad (2.3-49)$$

The CDF of a Rayleigh random variable can be easily found by integrating the PDF. The result is

$$F(x) = \begin{cases} 1 - e^{-\frac{x^2}{2\sigma^2}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-50)$$

The PDF of a Rayleigh random variable is plotted in Figure 2.3-4.

A generalized version of the Rayleigh random variable is obtained when we have n iid zero-mean Gaussian random variables $\{X_i, 1 \leq i \leq n\}$ where each X_i has an $\mathcal{N}(0, \sigma^2)$ distribution. In this case

$$X = \sqrt{\sum_{i=1}^n X_i^2} \quad (2.3-51)$$

has a *generalized Rayleigh distribution*. The PDF for this random variable is given by

$$p(x) = \begin{cases} \frac{x^{n-1}}{2^{\frac{n-2}{2}} \sigma^n \Gamma(\frac{n}{2})} e^{-\frac{x^2}{2\sigma^2}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-52)$$

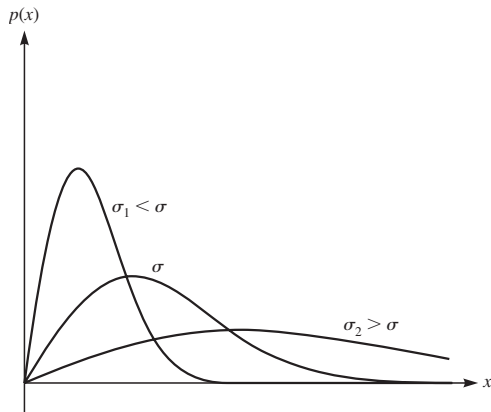


FIGURE 2.3-4

The PDF of the Rayleigh random variable for three different values of σ .

For the generalized Rayleigh, and with $n = 2m$, the CDF is given by

$$F(x) = \begin{cases} 1 - e^{-\frac{x^2}{2\sigma^2}} \sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{x^2}{2\sigma^2} \right)^k & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-53)$$

The k th moment of a generalized Rayleigh for any integer value of n (even or odd) is given by

$$E[X^k] = (2\sigma^2)^{\frac{k}{2}} \frac{\Gamma\left(\frac{n+k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \quad (2.3-54)$$

The Ricean Random Variable

If X_1 and X_2 are two independent Gaussian random variables distributed according to $\mathcal{N}(m_1, \sigma^2)$ and $\mathcal{N}(m_2, \sigma^2)$ (i.e., the variances are equal and the means may be different), then

$$X = \sqrt{X_1^2 + X_2^2} \quad (2.3-55)$$

is a Ricean random variable with PDF

$$p(x) = \begin{cases} \frac{x}{\sigma^2} I_0\left(\frac{sx}{\sigma^2}\right) e^{-\frac{x^2+s^2}{2\sigma^2}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-56)$$

where $s = \sqrt{m_1^2 + m_2^2}$ and $I_0(x)$ is given by Equation 2.3-32. It is clear that a Ricean random variable is the square root of a noncentral χ^2 random variable with two degrees of freedom.

It is readily seen that for $s = 0$, the Ricean random variable reduces to a Rayleigh random variable. For large s the Ricean random variable can be well approximated by a Gaussian random variable.

The CDF of a Ricean random variable can be expressed as

$$F(x) = \begin{cases} 1 - Q_1\left(\frac{s}{\sigma}, \frac{x}{\sigma}\right) & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-57)$$

where $Q_1(a, b)$ is defined by Equations 2.3-37 and 2.3-38.

The first two moments of the Ricean random variable are given by

$$\begin{aligned} E[X] &= \sigma \sqrt{\frac{\pi}{2}} {}_1F_1\left(-\frac{1}{2}, 1, -\frac{s^2}{2\sigma^2}\right) \\ &= \sigma \sqrt{\frac{\pi}{2}} e^{-\frac{K}{2}} \left[(1+K) I_0\left(\frac{K}{2}\right) + K I_1\left(\frac{K}{2}\right) \right] \\ E[X^2] &= 2\sigma^2 + s^2 \end{aligned} \quad (2.3-58)$$

where K is the Rice factor defined in Equation 2.3-60.

In general, the k th moment of this random variable is given by

$$E[X^k] = (2\sigma^2)^{\frac{k}{2}} \Gamma\left(1 + \frac{k}{2}\right) {}_1F_1\left(-\frac{k}{2}, 1; -\frac{s^2}{2\sigma^2}\right) \quad (2.3-59)$$

Another form of the Ricean density function is obtained by defining the *Rice factor* K as

$$K = \frac{s^2}{2\sigma^2} \quad (2.3-60)$$

If we define $A = s^2 + 2\sigma^2$, the Ricean PDF can be written as

$$p(x) = \begin{cases} \frac{2(K+1)}{A} x e^{-\frac{K+1}{A}(x^2 + \frac{AK}{K+1})} I_0\left(2x\sqrt{\frac{K(K+1)}{A}}\right) & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-61)$$

For the normalized case when $A = 1$ (or, equivalently, when $E[X^2] = s^2 + 2\sigma^2 = 1$) this reduces to

$$p(x) = \begin{cases} 2(K+1)x e^{-(K+1)(x^2 + \frac{K}{K+1})} I_0(2x\sqrt{K(K+1)}) & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-62)$$

A plot of the PDF of a Ricean random variable for different values of K is shown in Figure 2.3–5.

Similar to the Rayleigh random variable, a *generalized Ricean random variable* can be defined as

$$X = \sqrt{\sum_{i=1}^n X_i^2} \quad (2.3-63)$$

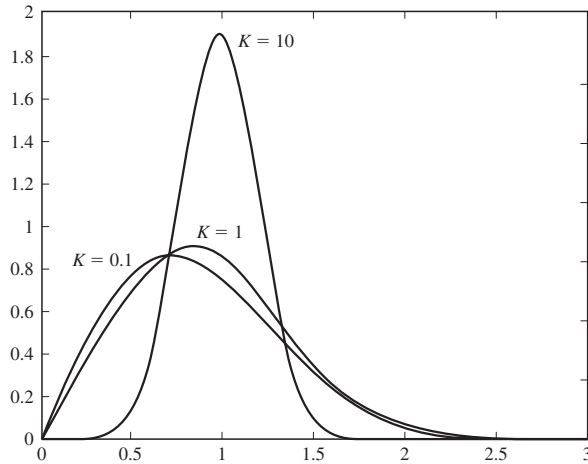


FIGURE 2.3–5

The Ricean PDF for different values of K . For small K this random variable reduces to a Rayleigh random variable, and for large K it is well approximated by a Gaussian random variable.

where X_i 's are independent Gaussians with mean m_i and common variance σ^2 . In this case the PDF is given by

$$p(x) = \begin{cases} \frac{x^{\frac{n}{2}}}{\sigma^2 s^{\frac{n-2}{2}}} e^{-\frac{x^2+s^2}{2\sigma^2}} I_{\frac{n}{2}-1} \left(\frac{xs}{\sigma^2} \right) & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-64)$$

and the CDF is given by

$$F(x) = \begin{cases} 1 - Q_m \left(\frac{s}{\sigma}, \frac{x}{\sigma} \right) & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-65)$$

where

$$s = \sqrt{\sum_{i=1}^n m_i^2}$$

The k th moment of a generalized Ricean is given by

$$E[X^k] = (2\sigma^2)^{\frac{k}{2}} e^{-\frac{s^2}{2\sigma^2}} \frac{\Gamma(\frac{n+k}{2})}{\Gamma(\frac{n}{2})} {}_1F_1 \left(\frac{n+k}{2}, \frac{n}{2}; \frac{s^2}{2\sigma^2} \right) \quad (2.3-66)$$

The Nakagami Random Variable

Both the Rayleigh distribution and the Rice distribution are frequently used to describe the statistical fluctuations of signals received from a multipath fading channel. These channel models are considered in Chapters 13 and 14. Another distribution that is frequently used to characterize the statistics of signals transmitted through multipath fading channels is the Nakagami m distribution. The PDF for this distribution is given by Nakagami (1960) as

$$p(x) = \begin{cases} \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega} \right)^m x^{2m-1} e^{-mx^2/\Omega} & x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-67)$$

where Ω is defined as

$$\Omega = E[X^2] \quad (2.3-68)$$

and the parameter m is defined as the ratio of moments, called the *fading figure*,

$$m = \frac{\Omega^2}{E[(X^2 - \Omega)^2]}, \quad m \geq \frac{1}{2} \quad (2.3-69)$$

A normalized version of Equation 2.3-67 may be obtained by defining another random variable $Y = X/\sqrt{\Omega}$ (see Problem 2.42). The n th moment of X is

$$E[X^n] = \frac{\Gamma(m + \frac{n}{2})}{\Gamma(m)} \left(\frac{\Omega}{m} \right)^{n/2} \quad (2.3-70)$$

The mean and the variance for this random variable are given by

$$\begin{aligned} E[X] &= \frac{\Gamma(m + \frac{1}{2})}{\Gamma(m)} \left(\frac{\Omega}{m}\right)^{1/2} \\ \text{VAR}[X] &= \Omega \left(1 - \frac{1}{m} \left(\frac{\Gamma(m + \frac{1}{2})}{\Gamma(m)}\right)^2\right) \end{aligned} \quad (2.3-71)$$

By setting $m = 1$, we observe that Equation 2.3-67 reduces to a Rayleigh PDF. For values of m in the range $\frac{1}{2} \leq m \leq 1$, we obtain PDFs that have larger tails than a Rayleigh-distributed random variable. For values of $m > 1$, the tail of the PDF decays faster than that of the Rayleigh. Figure 2.3-6 illustrates the Nakagami PDF for different values of m .

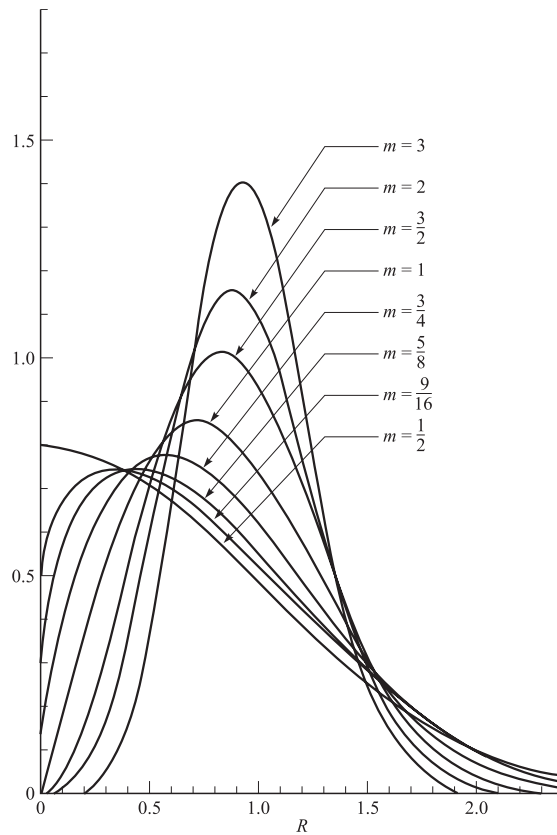


FIGURE 2.3-6

The PDF for the Nakagami m distribution, shown with $\Omega = 1$. m is the fading figure.

The Lognormal Random Variable

Suppose that a random variable Y is normally distributed with mean m and variance σ^2 . Let us define a new random variable X that is related to Y through the transformation $Y = \ln X$ (or $X = e^Y$). Then the PDF of X is

$$p(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2} x} e^{-(\ln x - m)^2 / 2\sigma^2} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.3-72)$$

For this random variable

$$\begin{aligned} E[X] &= e^{m + \frac{\sigma^2}{2}} \\ \text{VAR}[X] &= e^{2m + \sigma^2} (e^{\sigma^2} - 1) \end{aligned} \quad (2.3-73)$$

The lognormal distribution is suitable for modeling the effect of shadowing of the signal due to large obstructions, such as tall buildings, in mobile radio communications. Examples of the lognormal PDF are shown in Figure 2.3–7.

Jointly Gaussian Random Variables

An $n \times 1$ column random vector $\mathbf{X}$ with components $\{X_i, 1 \leq i \leq n\}$ is called a *Gaussian vector*, and its components are called *jointly Gaussian random variables* or

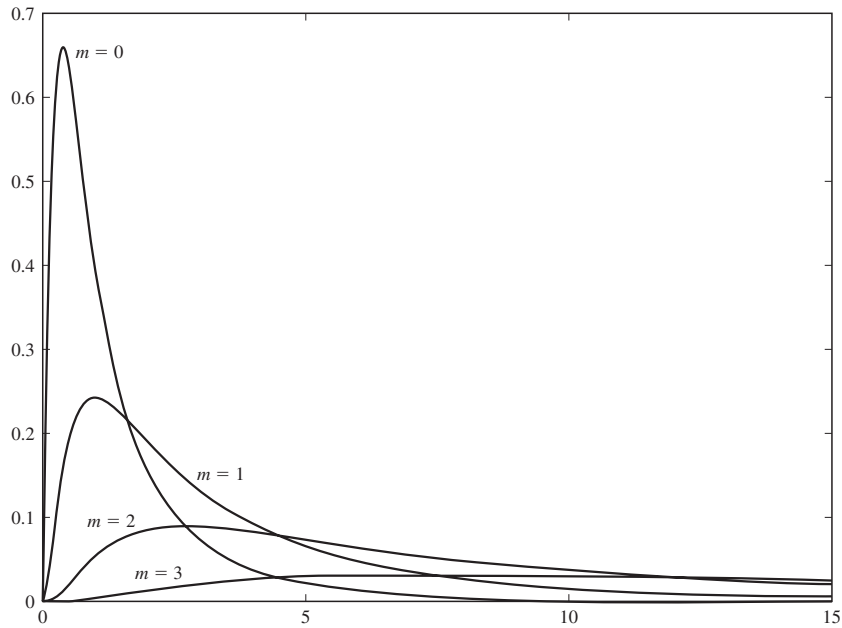


FIGURE 2.3–7

Lognormal PDF with $\sigma = 1$ for different values of m .

multivariate Gaussian random variables if the joint PDF of X_i 's can be written as

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}(\det \mathbf{C})^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^t \mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})} \quad (2.3-74)$$

where $\mathbf{m}$ and $\mathbf{C}$ are the mean vector and covariance matrix, respectively, of $\mathbf{X}$ and are given by

$$\begin{aligned} \mathbf{m} &= \mathbf{E}[\mathbf{X}] \\ \mathbf{C} &= \mathbf{E}[(\mathbf{X} - \mathbf{m})(\mathbf{X} - \mathbf{m})^t] \end{aligned} \quad (2.3-75)$$

From this definition it is clear that

$$C_{ij} = \text{COV}[X_i, X_j] \quad (2.3-76)$$

and therefore $\mathbf{C}$ is a symmetric matrix. From elementary probability it is also well known that $\mathbf{C}$ is nonnegative definite.

In the special case of $n = 2$, we have

$$\begin{aligned} \mathbf{m} &= \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \end{aligned} \quad (2.3-77)$$

where

$$\rho = \frac{\text{COV}[X_1, X_2]}{\sigma_1\sigma_2}$$

is the correlation coefficient of the two random variables. In this case the PDF reduces to

$$p(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{\left(\frac{x_1-m_1}{\sigma_1}\right)^2 + \left(\frac{x_2-m_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-m_1}{\sigma_1}\right)\left(\frac{x_2-m_2}{\sigma_2}\right)}{2(1-\rho^2)}} \quad (2.3-78)$$

where m_1, m_2, σ_1^2 and σ_2^2 are means and variances of the two random variables and ρ is their correlation coefficient. Note that in the special case when $\rho = 0$ (i.e., when the two random variables are uncorrelated), we have

$$p(x_1, x_2) = \mathcal{N}(m_1, \sigma_1^2) \times \mathcal{N}(m_2, \sigma_2^2)$$

This means that the two random variables are independent, and therefore for this case independence and uncorrelatedness are equivalent. This property is true for general jointly Gaussian random variables.

Another important property of jointly Gaussian random variables is that linear combinations of jointly Gaussian random variables are also jointly Gaussian. In other words, if $\mathbf{X}$ is a Gaussian vector, the random vector $\mathbf{Y} = \mathbf{A}\mathbf{X}$, where the invertible matrix $\mathbf{A}$ represents a linear transformation, is also a Gaussian vector whose mean and

covariance matrix are given by

$$\begin{aligned} \mathbf{m}_Y &= \mathbf{A}\mathbf{m}_X \\ \mathbf{C}_Y &= \mathbf{A}\mathbf{C}_X\mathbf{A}^t \end{aligned} \quad (2.3-79)$$

This property is developed in Problem 2.23.

In summary, jointly Gaussian random variables have the following important properties:

1. For jointly Gaussian random variables, *uncorrelated* is equivalent to *independent*.
2. Linear combinations of jointly Gaussian random variables are themselves jointly Gaussian.
3. The random variables in any subset of jointly Gaussian random variables are jointly Gaussian, and any subset of random variables conditioned on random variables in any other subset is also jointly Gaussian (all joint subsets and all conditional subsets are Gaussian).

We also emphasize that any set of independent Gaussian random variables is jointly Gaussian, but this is not necessarily true for a set of dependent Gaussian random variables.

Table 2.3–3 summarizes some of the properties of the most important random variables.

■ 2.4

BOUNDS ON TAIL PROBABILITIES

Performance analysis of communication systems requires computation of error probabilities of these systems. In many cases, as we will observe in the following chapters, the error probability of a communication system is expressed in terms of the probability that a random variable exceeds a certain value, i.e., in the form of $P[X > \alpha]$. Unfortunately, in many cases these probabilities cannot be expressed in closed form. In such cases we are interested in finding upper bounds on these tail probabilities. These upper bounds are of the form $P[X > \alpha] \leq \beta$. In this section we describe different methods for providing and tightening such bounds.

The Markov Inequality

The Markov inequality gives an upper bound on the tail probability of nonnegative random variables. Let us assume that X is a nonnegative random variable, i.e., $p(x) = 0$ for all $x < 0$, and assume $\alpha > 0$ is an arbitrary positive real number. The Markov inequality states that

$$P[X \geq \alpha] \leq \frac{E[X]}{\alpha} \quad (2.4-1)$$

■ TABLE 2.3-3

Properties of Important Random Variables

RV (Parameters)	PDF or PMF	E [X]	VAR [X]	$\Phi_X(\omega) = E \left[e^{j\omega X} \right]$
Bemoulli (p)	$P(X = 1) = 1 - P(X = 0) = p$ $0 \leq p \leq 1$	p	$p(1 - p)$	$pe^{j\omega} + (1 - p)$
Binomial (n, p)	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ $0 \leq k \leq n, 0 \leq p \leq 1$	np	$np(1 - p)$	$(pe^{j\omega} + (1 - p))^n$
Uniform (a, b)	$\frac{1}{b-a}, a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{j\omega b} - e^{j\omega a}}{j\omega(b-a)}$
Exponential (λ)	$\lambda e^{-\lambda x}, \lambda > 0, x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - j\omega}$
Gaussian (m, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$ $\sigma > 0$	m	σ^2	$e^{jom - \frac{\omega^2 \sigma^2}{2}}$
Gamma (λ, α)	$\frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$ $x \geq 0, \lambda, \alpha > 0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - j\omega} \right)^\alpha$
$\chi^2(n, \sigma^2)$	$\frac{1}{2^{n/2} \Gamma(\frac{n}{2}) \sigma^n} x^{\frac{n}{2}-1} e^{-\frac{x}{2\sigma^2}}$ $x, \sigma > 0, n \in \mathbb{N}$	$n\sigma^2$	$2n\sigma^4$	$\left(\frac{1}{1-2j\omega\sigma^2} \right)^{n/2}$
Noncentral $\chi^2(n, s, \sigma^2)$	$\frac{1}{2\sigma^2} \left(\frac{x}{s} \right)^{\frac{n-2}{2}} e^{-\frac{x+s}{2\sigma^2}} I_{\frac{n}{2}-1} \left(\frac{s}{\sigma^2} \sqrt{x} \right)$ $x, s, \sigma > 0, n \in \mathbb{N}$	$n\sigma^2 + s^2$	$2n\sigma^4 + 4\sigma^2 s^2$	$\left(\frac{1}{1-2j\omega\sigma^2} \right)^{n/2} e^{\frac{j\omega s^2}{1-2j\omega\sigma^2}}$
Rayleigh (σ^2)	$\frac{x}{\sigma^2} e^{-x/2\sigma^2}$ $x, \sigma > 0$	$\sigma \sqrt{\frac{\pi}{2}}$	$\left(2 - \frac{\pi}{2} \right) \sigma^2$	${}_1F_1 \left(1, \frac{1}{2}; -\frac{\omega^2 \sigma^2}{2} \right) + j \sqrt{\frac{\pi}{2}} \omega \sigma e^{-\frac{\omega^2 \sigma^2}{2}}$
Ricean (σ^2, s)	$\frac{s}{\sigma^2} I_0 \left(\frac{sx}{\sigma^2} \right) e^{-\frac{x+s}{2\sigma^2}}$ $x, s, \sigma > 0$	$\sigma \sqrt{\frac{\pi}{2}} {}_1F_1 \left(-\frac{1}{2}, 1, -\frac{s^2}{2\sigma^2} \right)$	$2\sigma^2 + s^2 - (E[X])^2$	—
Jointly Gaussian ($\mathbf{m}, \mathbf{C}$)	$\frac{1}{(2\pi)^{n/2} \det(\mathbf{C})} e^{-\frac{1}{2}[(\mathbf{x}-\mathbf{m})^T \mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})]}$ $\mathbf{C}$ symmetric and positive definite	$\mathbf{m}$	$\mathbf{C}$ (cov. matrix)	$e^{jm^T \omega - \frac{1}{2} \omega^T \mathbf{C} \omega}$

To see this, we observe that

$$\begin{aligned}
 E[X] &= \int_0^{\infty} xp(x) dx \\
 &\geq \int_{\alpha}^{\infty} xp(x) dx \\
 &\geq \alpha \int_{\alpha}^{\infty} p(x) dx \\
 &= \alpha P[X \geq \alpha]
 \end{aligned} \tag{2.4-2}$$

Dividing both sides by α gives the desired inequality.

Chernov Bound

The Chernov bound is a very tight and useful bound that is obtained from the Markov inequality. Unlike the Markov inequality that is applicable only to nonnegative random variables, the Chernov bound can be applied to all random variables.

Let X be an arbitrary random variable, and let δ and ν be arbitrary real numbers ($\nu \neq 0$). Define random variable Y by $Y = e^{\nu X}$ and constant α by $\alpha = e^{\nu\delta}$. Obviously, Y is a nonnegative random variable and α is a positive real number. Applying the Markov inequality to Y and α yields

$$P[e^{\nu X} \geq e^{\nu\delta}] \leq \frac{E[e^{\nu X}]}{e^{\nu\delta}} = E[e^{\nu(X-\delta)}] \tag{2.4-3}$$

The event $\{e^{\nu X} \geq e^{\nu\delta}\}$ is equivalent to the event $\{\nu X \geq \nu\delta\}$ which for positive or negative values of ν is equivalent to $\{X \geq \delta\}$ or $\{X \leq \delta\}$, respectively. Therefore we have

$$P[X \geq \delta] \leq E[e^{\nu(X-\delta)}], \quad \text{for all } \nu > 0 \tag{2.4-4}$$

$$P[X \leq \delta] \leq E[e^{\nu(X-\delta)}], \quad \text{for all } \nu < 0 \tag{2.4-5}$$

Since the two inequalities are valid for all positive and negative values of ν , respectively, it makes sense to find the values of ν that give the tightest possible bounds. To this end, we differentiate the right hand of the inequalities with respect to ν and find its root; this is the value of ν that gives the tightest bound. From this point on, we will consider only the first inequality. The extension to the second inequality is straightforward.

Let us define function $g(\nu)$ to denote the right side of the inequalities, i.e.,

$$g(\nu) = E[e^{\nu(X-\delta)}]$$

Differentiating $g(\nu)$, we have

$$g'(\nu) = E[(X - \delta)e^{\nu(X-\delta)}] \tag{2.4-6}$$

The second derivative of $g(\nu)$ is given by

$$g''(\nu) = E[(X - \delta)^2 e^{\nu(X-\delta)}]$$

It is easily seen that for all ν , we have $g''(\nu) > 0$ and hence $g(\nu)$ is convex and $g'(\nu)$ is an increasing function, and therefore can have only one root. In addition, since $g(\nu)$ is convex, this single root minimizes $g(\nu)$ and therefore results in the the tightest bound. Putting $g'(\nu) = 0$, we find the root to be obtained by solving the equation

$$E[Xe^{\nu X}] = \delta E[e^{\nu X}] \quad (2.4-7)$$

Equation 2.4-7 has a single root ν^* that gives the tightest bound. The only thing that remains to be checked is to see whether this ν^* satisfies the $\nu^* > 0$ condition. Since $g'(\nu)$ is an increasing function, its only root is positive if $g'(0) < 0$. From Equation 2.4-6 we have

$$g'(0) = E[X] - \delta$$

therefore $\nu^* > 0$ if and only if $\delta > E[X]$.

Summarizing, from Equations 2.4-4 and 2.4-5 we conclude

$$P[X \geq \delta] \leq e^{-\nu^* \delta} E[e^{\nu^* X}], \quad \text{for } \delta > E[X] \quad (2.4-8)$$

$$P[X \leq \delta] \leq e^{-\nu^* \delta} E[e^{\nu^* X}], \quad \text{for } \delta < E[X] \quad (2.4-9)$$

where ν^* is the solution of Equation 2.4-7. Equations 2.4-8 and 2.4-9 are known as Chernov bounds. Finding optimal ν^* by solving Equation 2.4-7 is sometimes difficult. In such cases a numerical approximation or an educated guess gives a suboptimal bound. The Chernov bound can also be given in terms of the moment generating function (MGF) $\Theta_X(\nu) = E[e^{\nu X}]$ as

$$P[X \geq \delta] \leq e^{-\nu^* \delta} \Theta_X(\nu^*), \quad \text{for } \delta > E[X] \quad (2.4-10)$$

$$P[X \leq \delta] \leq e^{-\nu^* \delta} \Theta_X(\nu^*), \quad \text{for } \delta < E[X] \quad (2.4-11)$$

EXAMPLE 2.4-1. Consider the Laplace PDF given by

$$p(x) = \frac{1}{2} e^{-|x|} \quad (2.4-12)$$

Let us evaluate the upper tail probability $P[X \geq \delta]$ for some $\delta > 0$ from the Chernov bound and compare it with the true tail probability, which is

$$P[X \geq \delta] = \int_{\delta}^{\infty} \frac{1}{2} e^{-x} dx = \frac{1}{2} e^{-\delta} \quad (2.4-13)$$

First note that $E[X] = 0$, and therefore the condition $\delta > E[X]$ needed to use the upper tail probability in the Chernov bound is satisfied. To solve Equation 2.4-7 for ν^* , we must determine $E[Xe^{\nu X}]$ and $E[e^{\nu X}]$. For the PDF in Equation 2.4-12, we find that $E[Xe^{\nu X}]$ and $E[e^{\nu X}]$ converge only if $-1 < \nu < 1$, and for this range of values of ν we have

$$\begin{aligned} E[Xe^{\nu X}] &= \frac{2\nu}{(\nu+1)^2(\nu-1)^2} \\ E[e^{\nu X}] &= \frac{1}{(1+\nu)(1-\nu)} \end{aligned} \quad (2.4-14)$$

Substituting these values into Equation 2.4–7, we obtain the quadratic equation

$$v^2\delta + 2v - \delta = 0$$

which has the solutions

$$v^* = \frac{-1 \pm \sqrt{1 + \delta^2}}{\delta} \quad (2.4-15)$$

Since v^* must be in the $(-1, +1)$ interval for $E[Xe^{vX}]$ and $E[e^{vX}]$ to converge, the only acceptable solution is

$$v^* = \frac{-1 + \sqrt{1 + \delta^2}}{\delta} \quad (2.4-16)$$

Finally, we evaluate the upper bound in Equation 2.4–8 by substituting for v^* from Equation 2.4–16. The result is

$$P[X \geq \delta] \leq \frac{\delta^2}{2(-1 + \sqrt{1 + \delta^2})} e^{1 - \sqrt{1 + \delta^2}} \quad (2.4-17)$$

For $\delta \gg 1$, Equation 2.4–17 reduces to

$$P(X \geq \delta) \leq \frac{\delta}{2} e^{-\delta} \quad (2.4-18)$$

We note that the Chernov bound decreases exponentially as δ increases. Consequently, it approximates closely the exact tail probability given by Equation 2.4–13.

EXAMPLE 2.4-2. In performance analysis of communication systems over fading channels, we encounter random variables of the form

$$X = d^2 R^2 + 2RdN \quad (2.4-19)$$

where d is a constant, R is a Ricean random variable with parameters s and σ representing channel attenuation due to fading, and N is a zero-mean Gaussian random variable with variance $\frac{N_0}{2}$ representing channel noise. It is assumed that R and N are independent random variables. We are interested to apply the Chernov bounding technique to find an upper bound on $P[X < 0]$. From the Chernov bound given in Equation 2.4–5, we have

$$P[X \leq 0] \leq E[e^{vX}], \quad \text{for all } v < 0 \quad (2.4-20)$$

To determine $E[e^{vX}]$, we use the well-known relation

$$E[Y] = E[E[Y|X]] \quad (2.4-21)$$

from elementary probability. We note that conditioned on R , X is a Gaussian random variable with mean $d^2 R^2$ and variance $2R^2 d^2 N_0$. Using the relation for the moment generating function of a Gaussian random variable from Table 2.3–3, we have

$$E[e^{vX}|R] = e^{vd^2 R^2 + v^2 d^2 N_0 R^2} = e^{vd^2(1+N_0 v)R^2} \quad (2.4-22)$$

Now noting that R^2 is a noncentral χ^2 random variable with two degrees of freedom, and using the characteristic function for this random variable from Table 2.3–3, we obtain

$$\begin{aligned} E[e^{\nu X}] &= E[E[e^{\nu X} | R]] \\ &= E[e^{\nu d^2(1+N_0\nu)R^2}] \\ &= \frac{1}{1 - 2\nu d^2(1 + N_0\nu)\sigma^2} e^{\frac{\nu d^2(1+N_0\nu)s^2}{1-2\nu d^2(1+N_0\nu)\sigma^2}} \end{aligned} \quad (2.4-23)$$

where we have used Equation 2.4–21. From Equations 2.4–20 and 2.4–23 we conclude that

$$P[X \leq 0] \leq \min_{\nu < 0} \frac{1}{1 - 2\nu d^2(1 + N_0\nu)\sigma^2} e^{\frac{\nu d^2(1+N_0\nu)s^2}{1-2\nu d^2(1+N_0\nu)\sigma^2}} \quad (2.4-24)$$

It can be easily verified by differentiation that in the range of interest ($\nu < 0$), the right-hand side is an increasing function of $\lambda = \nu d^2(1 + N_0\nu)$, and therefore the minimum is achieved when λ is minimized. By simple differentiation we can verify that λ is minimized for $\nu = -\frac{1}{2N_0}$, resulting in

$$P[X \leq 0] \leq \frac{1}{1 + \frac{d^2}{2N_0}\sigma^2} e^{-\frac{\frac{d^2}{4N_0}s^2}{1 + \frac{d^2}{2N_0}\sigma^2}} \quad (2.4-25)$$

If we use Equation 2.3–61 or 2.3–62 for the Ricean random variable, we obtain the following bounds:

$$P[X \leq 0] \leq \frac{K+1}{K+1 + \frac{A^2 d^2}{4N_0}} e^{-\frac{\frac{A^2 K d^2}{4N_0}}{K+1 + \frac{A^2 d^2}{4N_0}}} \quad (2.4-26)$$

and

$$P[X \leq 0] \leq \frac{K+1}{K+1 + \frac{d^2}{4N_0}} e^{-\frac{\frac{K d^2}{4N_0}}{K+1 + \frac{d^2}{4N_0}}} \quad (2.4-27)$$

For the case of Rayleigh fading channels, in which $s = 0$, these relations reduce to

$$P[X \leq 0] \leq \frac{1}{1 + \frac{d^2}{2N_0}\sigma^2} \quad (2.4-28)$$

Chernov Bound for Sums of Random Variables

Let $\{X_i\}$, $1 \leq i \leq n$, denote a sequence of iid random variables and define

$$Y = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.4-29)$$

We are interested to find a bound on $P[Y > \delta]$, where $\delta > E[X]$. Applying the Chernov bound, we have

$$\begin{aligned} P[Y > \delta] &= P\left[\sum_{i=1}^n X_i > n\delta\right] \\ &\leq E\left[e^{\nu(\sum_{i=1}^n X_i - n\delta)}\right] \\ &= [E[e^{\nu(X-\delta)}]]^n, \quad \nu > 0 \end{aligned} \quad (2.4-30)$$

To find the optimal choice of ν we equate the derivative of the right-hand side to zero

$$\frac{d}{d\nu} [E[e^{\nu(X-\delta)}]]^n = n [E[e^{\nu(X-\delta)}]]^{n-1} E[(X-\delta)e^{\nu(X-\delta)}] = 0 \quad (2.4-31)$$

The single root of this equation is obtained by solving

$$E[Xe^{\nu X}] = \delta E[e^{\nu X}] \quad (2.4-32)$$

which is exactly Equation 2.4-7. Therefore, for the sum of iid random variables we find the ν^* solution of Equation 2.4-7, and then we use

$$P[Y > \delta] \leq [E[e^{\nu^*(X-\delta)}]]^n = e^{-n\nu^*\delta} [E[e^{\nu^* X}]]^n \quad (2.4-33)$$

EXAMPLE 2.4-3. The X_i 's are binary iid random variables with $P[X = 1] = 1 - p$ and $P[X = -1] = p$, where $p < \frac{1}{2}$. We are interested to find a bound on

$$P\left[\sum_{i=1}^n X_i > 0\right]$$

We have $E[X] = p - (1 - p) = 2p - 1 < 0$. Assuming $\delta = 0$, the condition $\delta > E[X]$ is satisfied, and the preceding development can be applied to this case. We have

$$E[Xe^{\nu X}] = pe^{\nu} - (1 - p)e^{-\nu} \quad (2.4-34)$$

and Equation 2.4-7 becomes

$$pe^{\nu} - (1 - p)e^{-\nu} = 0 \quad (2.4-35)$$

which has the unique solution

$$\nu^* = \frac{1}{2} \ln \frac{1 - p}{p} \quad (2.4-36)$$

Using this value, we have

$$E[e^{\nu^* X}] = p\sqrt{\frac{1 - p}{p}} + (1 - p)\sqrt{\frac{p}{1 - p}} = 2\sqrt{p(1 - p)} \quad (2.4-37)$$

Substituting this result into Equation 2.4-33 results in

$$P\left[\sum_{i=1}^n X_i > 0\right] \leq [4p(1 - p)]^{\frac{n}{2}} \quad (2.4-38)$$

Since for $p < \frac{1}{2}$ we have $4p(1-p) < 1$, the bound given in Equation 2.4–38 tends to zero exponentially.

■ 2.5

LIMIT THEOREMS FOR SUMS OF RANDOM VARIABLES

If $\{X_i, i = 1, 2, 3, \dots\}$ represents a sequence of iid random variables, then it is intuitively clear that the running average of this sequence, i.e.,

$$Y_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.5-1)$$

should in some sense converge to the average of the random variables. Two limit theorems, i.e., the *law of large numbers* (LLN) and the *central limit theorem* (CLT), rigorously state how the running average of the random variable behaves as n becomes large.

The (strong) law of large numbers states that if $\{X_i, i = 1, 2, \dots\}$ is a sequence of iid random variables with $E[X_1] < \infty$, then

$$\frac{1}{n} \sum_{i=1}^n X_i \longrightarrow E[X_1] \quad (2.5-2)$$

where the type of convergence is *convergence almost everywhere* (a.e.) or *convergence almost surely* (a.s.), meaning the set of points in the probability space for which the left-hand side does not converge to the right-hand side has zero probability.

The central limit theorem states that if $\{X_i, i = 1, 2, \dots\}$ is a sequence of iid random variables with $m = E[X_1] < \infty$ and $\sigma^2 = \text{VAR}[X_1] < \infty$, then we have

$$\frac{\frac{1}{n} \sum_{i=1}^n X_i - m}{\frac{\sigma}{\sqrt{n}}} \longrightarrow \mathcal{N}(0, 1) \quad (2.5-3)$$

The type of convergence in the CLT is *convergence in distribution*, meaning the CDF of the left-hand side converges to the CDF of $\mathcal{N}(0, 1)$ as n increases.

■ 2.6

COMPLEX RANDOM VARIABLES

A complex random variable $Z = X + jY$ can be considered as a pair of real random variables X and Y . Therefore, we treat a complex random variable as a two-dimensional random vector with components X and Y . The PDF of a complex random variable is defined to be the joint PDF of its real and complex parts. If X and Y are jointly Gaussian random variables, then Z is a complex Gaussian random variable. The PDF of a zero-mean complex Gaussian random variable Z with iid real and imaginary parts

is given by

$$p(z) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.6-1)$$

$$= \frac{1}{2\pi\sigma^2} e^{-\frac{|z|^2}{2\sigma^2}} \quad (2.6-2)$$

For a complex random variable Z , the mean and variance are defined by

$$E[Z] = E[X] + j E[Y] \quad (2.6-3)$$

$$\text{VAR}[Z] = E[|Z|^2] - |E[Z]|^2 = \text{VAR}[X] + \text{VAR}[Y] \quad (2.6-4)$$

2.6-1 Complex Random Vectors

A complex random vector is defined as $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$, where $\mathbf{X}$ and $\mathbf{Y}$ are real-valued random vectors of size n . We define the following real-valued matrices for a complex random vector $\mathbf{Z}$.

$$\mathbf{C}_X = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^t] \quad (2.6-5)$$

$$\mathbf{C}_Y = E[(\mathbf{Y} - E[\mathbf{Y}])(\mathbf{Y} - E[\mathbf{Y}])^t] \quad (2.6-6)$$

$$\mathbf{C}_{XY} = E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{Y} - E[\mathbf{Y}])^t] \quad (2.6-7)$$

$$\mathbf{C}_{YX} = E[(\mathbf{Y} - E[\mathbf{Y}])(\mathbf{X} - E[\mathbf{X}])^t] \quad (2.6-8)$$

Matrices $\mathbf{C}_X$ and $\mathbf{C}_Y$ are the *covariance matrices* of real random vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively, and hence they are symmetric and nonnegative definite. It is clear from above that $\mathbf{C}_{YX} = \mathbf{C}_{XY}^t$.

The PDF of $\mathbf{Z}$ is the joint PDF of its real and imaginary parts. If we define the $2n$ -dimensional real vector

$$\tilde{\mathbf{Z}} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \quad (2.6-9)$$

then the PDF of the complex vector $\mathbf{Z}$ is the PDF of the real vector $\tilde{\mathbf{Z}}$. It is clear that $\mathbf{C}_{\tilde{\mathbf{Z}}}$, the covariance matrix of $\tilde{\mathbf{Z}}$, can be written as

$$\mathbf{C}_{\tilde{\mathbf{Z}}} = \begin{bmatrix} \mathbf{C}_X & \mathbf{C}_{XY} \\ \mathbf{C}_{YX} & \mathbf{C}_Y \end{bmatrix} \quad (2.6-10)$$

We also define the following two, in general complex-valued, matrices

$$\mathbf{C}_Z = E[(\mathbf{Z} - E[\mathbf{Z}])(\mathbf{Z} - E[\mathbf{Z}])^H] \quad (2.6-11)$$

$$\tilde{\mathbf{C}}_Z = E[(\mathbf{Z} - E[\mathbf{Z}])(\mathbf{Z} - E[\mathbf{Z}])^t] \quad (2.6-12)$$

where $\mathbf{A}^t$ denotes the transpose and $\mathbf{A}^H$ denotes the Hermitian transpose of $\mathbf{A}$ ($\mathbf{A}$ is transposed and each element of it is conjugated). $\mathbf{C}_Z$ and $\tilde{\mathbf{C}}_Z$ are called the *covariance* and the *pseudocovariance* of the complex random vector $\mathbf{Z}$, respectively. It is easy to

verify that for any $\mathbf{Z}$, the covariance matrix is Hermitian[†] and nonnegative definite. The pseudocovariance is *skew-Hermitian*.

From these definitions it is easy to verify the following relations.

$$\mathbf{C}_Z = \mathbf{C}_X + \mathbf{C}_Y + j(\mathbf{C}_{YX} - \mathbf{C}_{XY}) \quad (2.6-13)$$

$$\tilde{\mathbf{C}}_Z = \mathbf{C}_X - \mathbf{C}_Y + j(\mathbf{C}_{XY} + \mathbf{C}_{YX}) \quad (2.6-14)$$

$$\mathbf{C}_X = \frac{1}{2} \text{Re}[\mathbf{C}_Z + \tilde{\mathbf{C}}_Z] \quad (2.6-15)$$

$$\mathbf{C}_Y = \frac{1}{2} \text{Re}[\mathbf{C}_Z - \tilde{\mathbf{C}}_Z] \quad (2.6-16)$$

$$\mathbf{C}_{YX} = \frac{1}{2} \text{Im}[\mathbf{C}_Z + \tilde{\mathbf{C}}_Z] \quad (2.6-17)$$

$$\mathbf{C}_{XY} = \frac{1}{2} \text{Im}[\tilde{\mathbf{C}}_Z - \mathbf{C}_Z] \quad (2.6-18)$$

Proper and Circularly Symmetric Random Vectors

A complex random vector $\mathbf{Z}$ is called *proper* if its pseudocovariance is zero, i.e., if $\tilde{\mathbf{C}}_Z = \mathbf{0}$. From Equation 2.6-14 it is clear that for a proper random vector we have

$$\mathbf{C}_X = \mathbf{C}_Y \quad (2.6-19)$$

$$\mathbf{C}_{XY} = -\mathbf{C}_{YX} \quad (2.6-20)$$

Substituting these results into Equations 2.6-13 to 2.6-18 and 2.6-10, we conclude that for proper random vectors

$$\mathbf{C}_Z = 2\mathbf{C}_X + 2j\mathbf{C}_{YX} \quad (2.6-21)$$

$$\mathbf{C}_X = \mathbf{C}_Y = \frac{1}{2} \text{Re}[\mathbf{C}_Z] \quad (2.6-22)$$

$$\mathbf{C}_{YX} = -\mathbf{C}_{XY} = \frac{1}{2} \text{Im}[\mathbf{C}_Z] \quad (2.6-23)$$

$$\mathbf{C}_{\tilde{Z}} = \begin{bmatrix} \mathbf{C}_X & \mathbf{C}_{XY} \\ -\mathbf{C}_{XY} & \mathbf{C}_X \end{bmatrix} \quad (2.6-24)$$

For the special case of $n = 1$, i.e., when we are dealing with a single complex random variable $Z = X + jY$, the conditions for being proper become

$$\text{VAR}[X] = \text{VAR}[Y] \quad (2.6-25)$$

$$\text{COV}[X, Y] = -\text{COV}[Y, X] \quad (2.6-26)$$

which means that Z is proper if X and Y have equal variances and are uncorrelated. In this case $\text{VAR}[Z] = 2 \text{VAR}[X]$. Since in the case of jointly Gaussian random variables uncorrelated is equivalent to independent, we conclude that a complex Gaussian random

[†]Matrix $\mathbf{A}$ is Hermitian if $\mathbf{A} = \mathbf{A}^H$. It is skew-Hermitian if $\mathbf{A}^H = -\mathbf{A}$.

variable Z is proper if and only if its real and complex parts are independent with equal variance. For a zero-mean proper complex Gaussian random variable, the PDF is given by Equation 2.6–2.

If the complex random vector $\mathbf{Z} = \mathbf{X} + j\mathbf{Y}$ is Gaussian, meaning that $\mathbf{X}$ and $\mathbf{Y}$ are jointly Gaussian, then we have

$$p(\mathbf{z}) = p(\tilde{\mathbf{z}}) = \frac{1}{(2\pi)^n (\det \mathbf{C}_{\tilde{\mathbf{Z}}})^{\frac{1}{2}}} e^{-\frac{1}{2}(\tilde{\mathbf{z}} - \tilde{\mathbf{m}})^t \mathbf{C}_{\tilde{\mathbf{Z}}}^{-1} (\tilde{\mathbf{z}} - \tilde{\mathbf{m}})} \quad (2.6-27)$$

where

$$\tilde{\mathbf{m}} = \mathbf{E} [\tilde{\mathbf{Z}}] \quad (2.6-28)$$

It can be shown that in the special case where $\mathbf{Z}$ is a proper n -dimensional complex Gaussian random vector, with mean $\mathbf{m} = \mathbf{E} [\mathbf{Z}]$ and nonsingular covariance matrix $\mathbf{C}_Z$, its PDF can be written as

$$p(\mathbf{z}) = \frac{1}{\pi^n \det \mathbf{C}_Z} e^{-\frac{1}{2}(\mathbf{z} - \mathbf{m})^\dagger \mathbf{C}_Z^{-1} (\mathbf{z} - \mathbf{m})} \quad (2.6-29)$$

A complex random vector $\mathbf{Z}$ is called *circularly symmetric* or *circular* if rotating the vector by any angle does not change its PDF. In other words, a complex random vector $\mathbf{Z}$ is circularly symmetric if $\mathbf{Z}$ and $e^{j\theta} \mathbf{Z}$ have the same PDF for all θ . In Problem 2.34 we will see that if $\mathbf{Z}$ is circular, then it is zero-mean and proper, i.e., $\mathbf{E} [\mathbf{Z}] = \mathbf{0}$ and $\mathbf{E} [\mathbf{Z}\mathbf{Z}^t] = \mathbf{0}$. In Problem 2.35 we show that if $\mathbf{Z}$ is a zero-mean proper Gaussian complex vector, then $\mathbf{Z}$ is circular. In other words, *for complex Gaussian random vectors being zero-mean and proper is equivalent to being circular*.

In Problem 2.36 we show that if $\mathbf{Z}$ is a proper complex vector, then any *affine transformation* of it, i.e., any transform of the form $\mathbf{W} = \mathbf{A}\mathbf{Z} + \mathbf{b}$, is also a proper complex vector. Since we know that if $\mathbf{Z}$ is Gaussian, so is $\mathbf{W}$, we conclude that if $\mathbf{Z}$ is a proper Gaussian vector, so is $\mathbf{W}$. For more details on properties of proper and circular random variables and random vectors, the reader is referred to Neeser and Massey (1993) and Eriksson and Koivunen (2006).

■ 2.7

RANDOM PROCESSES

Random processes, stochastic processes, or random signals are fundamental in the study of communication systems. Modeling information sources and communication channels requires a good understanding of random processes and techniques for analyzing them. We assume that the reader has a knowledge of the basic concepts of random processes including definitions of mean, autocorrelation, cross-correlation, stationarity, and ergodicity as given in standard texts such as Leon-Garcia (1994), Papoulis and Pillai (2002), Stark and Woods (2002). In the following paragraphs we present a brief review of the most important properties of random processes.

The mean $m_X(t)$ and the *autocorrelation function* of a random process $X(t)$ are defined as

$$m_X(t) = E[X(t)] \quad (2.7-1)$$

$$R_X(t_1, t_2) = E[X(t_1)X^*(t_2)] \quad (2.7-2)$$

The *cross-correlation function* of two random processes $X(t)$ and $Y(t)$ is defined by

$$R_{XY}(t_1, t_2) = E[X(t_1)Y^*(t_2)] \quad (2.7-3)$$

Note that $R_X(t_2, t_1) = R_X^*(t_1, t_2)$, i.e., $R_X(t_1, t_2)$ is Hermitian. For the cross-correlation we have $R_{YX}(t_2, t_1) = R_{XY}^*(t_1, t_2)$.

2.7-1 Wide-Sense Stationary Random Processes

Random process $X(t)$ is *wide-sense stationary* (WSS) if its mean is constant and $R_X(t_1, t_2) = R_X(\tau)$ where $\tau = t_1 - t_2$. For WSS processes $R_X(-\tau) = R_X^*(\tau)$. Two processes $X(t)$ and $Y(t)$ are *jointly wide-sense stationary* if both $X(t)$ and $Y(t)$ are WSS and $R_{XY}(t_1, t_2) = R_{XY}(\tau)$. For jointly WSS processes $R_{YX}(-\tau) = R_{XY}^*(\tau)$. A complex process is WSS if its real and imaginary parts are jointly WSS.

The *power spectral density* (PSD) or *power spectrum* of a WSS random process $X(t)$ is a function $\mathcal{S}_X(f)$ describing the distribution of power as a function of frequency. The unit for power spectral density is watts per hertz. The *Wiener-Khinchin theorem* states that for a WSS process, the power spectrum is the Fourier transform of the autocorrelation function $R_X(\tau)$, i.e.,

$$\mathcal{S}_X(f) = \mathcal{F}[R_X(\tau)] \quad (2.7-4)$$

Similarly, the *cross spectral density* (CSD) of two jointly WSS processes is defined as the Fourier transform of their cross-correlation function.

$$\mathcal{S}_{XY}(f) = \mathcal{F}[R_{XY}(\tau)] \quad (2.7-5)$$

The cross spectral density satisfies the following symmetry property:

$$\mathcal{S}_{XY}(f) = \mathcal{S}_{YX}^*(f) \quad (2.7-6)$$

From properties of the autocorrelation function it is easy to verify that the power spectral density of any real WSS process $X(t)$ is a real, nonnegative, and even function of f . For complex processes, power spectrum is real and nonnegative, but not necessarily even. The cross spectral density can be a complex function, even when both $X(t)$ and $Y(t)$ are real processes.

If $X(t)$ and $Y(t)$ are jointly WSS random processes, then $Z(t) = aX(t) + bY(t)$ is a WSS random process with autocorrelation and power spectral density given by

$$R_Z(\tau) = |a|^2 R_X(\tau) + |b|^2 R_Y(\tau) + ab^* R_{XY}(\tau) + ba^* R_{YX}(\tau) \quad (2.7-7)$$

$$\mathcal{S}_Z(f) = |a|^2 \mathcal{S}_X(f) + |b|^2 \mathcal{S}_Y(f) + 2 \operatorname{Re}[ab^* \mathcal{S}_{XY}(f)] \quad (2.7-8)$$

In the special case where $a = b = 1$, we have $Z(t) = X(t) + Y(t)$, which results in

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + R_{XY}(\tau) + R_{YX}(\tau) \quad (2.7-9)$$

$$\mathcal{S}_Z(f) = \mathcal{S}_X(f) + \mathcal{S}_Y(f) + 2 \operatorname{Re} [\mathcal{S}_{XY}(f)] \quad (2.7-10)$$

and when $a = 1$ and $b = j$, we have $Z(t) = X(t) + jY(t)$ and

$$R_Z(\tau) = R_X(\tau) + R_Y(\tau) + j(R_{YX}(\tau) + R_{XY}(\tau)) \quad (2.7-11)$$

$$\mathcal{S}_Z(f) = \mathcal{S}_X(f) + \mathcal{S}_Y(f) + 2 \operatorname{Im} [\mathcal{S}_{XY}(f)] \quad (2.7-12)$$

When a WSS process $X(t)$ passes through an LTI system with impulse response $h(t)$ and transfer function $H(f) = \mathcal{F}[h(t)]$, the output process $Y(t)$ and $X(t)$ are jointly WSS and the following relations hold:

$$m_Y = m_X \int_{-\infty}^{\infty} h(t) dt \quad (2.7-13)$$

$$R_{XY}(\tau) = R_X(\tau) \star h^*(-\tau) \quad (2.7-14)$$

$$R_Y(\tau) = R_X(\tau) \star h(\tau) \star h^*(-\tau) \quad (2.7-15)$$

$$m_Y = m_X H(0) \quad (2.7-16)$$

$$\mathcal{S}_{XY}(f) = \mathcal{S}_X(f) H^*(f) \quad (2.7-17)$$

$$\mathcal{S}_Y(f) = \mathcal{S}_X(f) |H(f)|^2 \quad (2.7-18)$$

The power in a WSS process $X(t)$ is the sum of the powers at all frequencies, and therefore it is the integral of the power spectrum over all frequencies. We can write

$$P_X = E[|X(t)|^2] = R_X(0) = \int_{-\infty}^{\infty} \mathcal{S}_X(f) df \quad (2.7-19)$$

Gaussian Random Processes

A real random process $X(t)$ is Gaussian if for all positive integers n and for all $(t_1, t_2, \dots, t_n)$, the random vector $(X(t_1), X(t_2), \dots, X(t_n))^t$ is a Gaussian random vector; i.e., random variables $\{X(t_i)\}_{i=1}^n$ are jointly Gaussian random variables. Similar to jointly Gaussian random variables, linear filtering of Gaussian random processes results in a Gaussian random process, even when the filtering is time-varying.

Two real random processes $X(t)$ and $Y(t)$ are jointly Gaussian if for all positive integers n, m and all $(t_1, t_2, \dots, t_n)$, and $(t'_1, t'_2, \dots, t'_m)$, the random vector

$$(X(t_1), X(t_2), \dots, X(t_n), Y(t'_1), Y(t'_2), \dots, Y(t'_m))^t$$

is a Gaussian vector. For two jointly Gaussian random processes $X(t)$ and $Y(t)$, being uncorrelated, i.e., having

$$R_{XY}(t + \tau, t) = E[X(t + \tau)] E[Y(t)] \quad \text{for all } t \text{ and } \tau \quad (2.7-20)$$

is equivalent to being independent.

A complex process $Z(t) = X(t) + jY(t)$ is Gaussian if $X(t)$ and $Y(t)$ are jointly Gaussian processes.

White Processes

A process is called a *white process* if its power spectral density is constant for all frequencies; this constant value is usually denoted by $\frac{N_0}{2}$.

$$\mathcal{S}_X(f) = \frac{N_0}{2} \quad (2.7-21)$$

Using Equation 2.7–19, we see that the power in a white process is infinite, indicating that white processes cannot exist as a physical process. Although white processes are not physically realizable processes, they are very useful, closely modeling some important physical phenomenon including the *thermal noise*.

Thermal noise is the noise generated in electric devices by thermal agitation of electrons. Thermal noise can be closely modeled by a random process $N(t)$ having the following properties:

1. $N(t)$ is a stationary process.
2. $N(t)$ is a zero-mean process.
3. $N(t)$ is a Gaussian process.
4. $N(t)$ is a white process whose power spectral density is given by

$$\mathcal{S}_N(f) = \frac{N_0}{2} = \frac{kT}{2} \quad (2.7-22)$$

where T is the ambient temperature in kelvins and k is *Boltzmann's constant*, equal to 38×10^{-23} J/K.

Discrete-Time Random Processes

Discrete-time random processes have similar properties to continuous time processes. In particular the PSD of a WSS discrete-time random process is defined as the discrete-time Fourier transform of its autocorrelation function

$$\mathcal{S}_X(f) = \sum_{m=-\infty}^{\infty} R_X(m) e^{-j2\pi f m} \quad (2.7-23)$$

and the autocorrelation function can be obtained as the inverse Fourier transform of the power spectral density as

$$R_X(m) = \int_{-1/2}^{1/2} \mathcal{S}_X(f) e^{j2\pi f m} df \quad (2.7-24)$$

The power in a discrete-time random process is given by

$$P = E[|X(n)|^2] = R_X(0) = \int_{-1/2}^{1/2} \mathcal{S}_X(f) df \quad (2.7-25)$$

2.7-2 Cyclostationary Random Processes

A random process $X(t)$ is *cyclostationary* if its mean and autocorrelation function are periodic functions with the same period T_0 . For a cyclostationary process we have

$$m_X(t + T_0) = m_X(t) \quad (2.7-26)$$

$$R_X(t_1 + T_0, t_2 + T_0) = R_X(t_1, t_2) \quad (2.7-27)$$

Cyclostationary processes are encountered frequently in the study of communication systems because many modulated processes can be modeled as cyclostationary processes. For a cyclostationary process, the average autocorrelation function is defined as the average of the autocorrelation function over one period

$$\overline{R_X(\tau)} = \frac{1}{T_0} \int_0^{T_0} R_X(t + \tau, t) dt \quad (2.7-28)$$

The (average) power spectral density for a cyclostationary process is defined as the Fourier transform of the average autocorrelation function, i.e.,

$$\mathcal{S}_X(f) = \mathcal{F}[\overline{R_X(\tau)}] \quad (2.7-29)$$

EXAMPLE 2.7-1. Let $\{a_n\}$ denote a discrete-time WSS random process with mean $m_a(n) = E[a_n] = m_a$ and autocorrelation function $R_a(m) = E[a_{n+m}a_n^*]$. Define the random process

$$X(t) = \sum_{n=-\infty}^{\infty} a_n g(t - nT) \quad (2.7-30)$$

for an arbitrary deterministic function $g(t)$. We have

$$m_X(t) = E[X(t)] = m_a \sum_{n=-\infty}^{\infty} g(t - nT) \quad (2.7-31)$$

This function is obviously periodic with period T . For the autocorrelation function we have

$$R_X(t + \tau, t) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} E[a_n a_m^*] g(t + \tau - nT) g^*(t - mT) \quad (2.7-32)$$

$$= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} R_a(n - m) g(t + \tau - nT) g^*(t - mT) \quad (2.7-33)$$

It can readily be verified that

$$R_X(t + \tau + T, t + T) = R_X(t + \tau, t) \quad (2.7-34)$$

Equations 2.7-31 and 2.7-34 show that $X(t)$ is a cyclostationary process.

2.7–3 Proper and Circular Random Processes

For a complex random process $Z(t) = X(t) + jY(t)$, we define the covariance and the pseudocovariance, similar to the case of complex random vectors, as

$$C_Z(t + \tau, t) = E[Z(t + \tau)Z^*(t)] \quad (2.7-35)$$

$$\tilde{C}_Z(t + \tau, t) = E[Z(t + \tau)Z(t)] \quad (2.7-36)$$

It is easy to verify that similar to Equations 2.6–13 and 2.6–14, we have

$$C_Z(t + \tau, t) = C_X(t + \tau, t) + C_Y(t + \tau, t) + j(C_{YX}(t + \tau, t) - C_{XY}(t + \tau, t)) \quad (2.7-37)$$

$$\tilde{C}_Z(t + \tau, t) = C_X(t + \tau, t) - C_Y(t + \tau, t) + j(C_{YX}(t + \tau, t) + C_{XY}(t + \tau, t)) \quad (2.7-38)$$

A complex random process $Z(t)$ is *proper* if its pseudocovariance is zero, i.e., $\tilde{C}_Z(t + \tau, t) = 0$. For a proper random process we have

$$C_X(t + \tau, t) = C_Y(t + \tau, t) \quad (2.7-39)$$

$$C_{YX}(t + \tau, t) = -C_{XY}(t + \tau, t) \quad (2.7-40)$$

and

$$C_Z(t + \tau, t) = 2C_X(t + \tau, t) + j2C_{YX}(t + \tau, t) \quad (2.7-41)$$

If $Z(t)$ is a zero-mean process, then all covariances in Equations 2.7–35 to 2.7–41 are substituted with auto- or cross-correlations. When $Z(t)$ is WSS, all auto- and cross-correlations are functions of τ only. A proper Gaussian random process is a random process for which, for all n and all $(t_1, t_2, \dots, t_n)$, the complex random vector $(Z(t_1), Z(t_2), \dots, Z(t_n))^t$ is a proper Gaussian vector.

A complex random process $Z(t)$ is circular if for all θ , $Z(t)$ and $e^{j\theta}Z(t)$ have the same statistical properties. Similar to the case of complex vectors, it can be shown that if $Z(t)$ is circular, then it is both proper and zero-mean. For the case of Gaussian processes, being proper and zero-mean is equivalent to being circular. Also similar to the case of complex vectors, passing a circular Gaussian process through a linear (not necessarily time-invariant) system results in a circular Gaussian process at the output.

2.7–4 Markov Chains

Markov chains are discrete-time, discrete-valued random processes in which the current value depends on the entire past values only through the most recent values. In a j th-order Markov chain, the current value depends on the past values only through the most recent j values, i.e.,

$$\begin{aligned} P[X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots] \\ = P[X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_{n-j} = x_{n-j}] \end{aligned} \quad (2.7-42)$$

It is convenient to consider the set of the most recent j values as the *state* of the Markov chain. With this definition the current state of the Markov chain, i.e., $S_n = (X_n, X_{n-1}, \dots, X_{n-j+1})$, depends only on the most recent state $S_{n-1} = (X_{n-1}, X_{n-2}, \dots, X_{n-j})$. That is,

$$P[S_n = s_n | S_{n-1} = s_{n-1}, S_{n-2} = s_{n-2}, \dots] = P[S_n = s_n | S_{n-1} = s_{n-1}] \quad (2.7-43)$$

which represents a first-order Markov chain in terms of the state variable S_n . Note that with this notation, X_n is a deterministic function of state S_n . We can generalize this notion to the case where the state evolves according to Equation 2.7-43 but the output—or the value of the random process X_n —depends on state S_n through a conditional probability mass function

$$P[X_n = x_n | S_n = s_n] \quad (2.7-44)$$

With this background, we define a Markov chain[†] as a finite-state machine with state at time n , denoted by S_n , taking values in the set $\{1, 2, \dots, S\}$ such that Equation 2.7-43 holds and the value of the random process at time n , denoted by X_n and taking values in a discrete set, depends statistically on the state through the conditional PMF $P[X_n = x_n | S_n = s_n]$.

The internal development of the process depends on the set of states and the probabilistic law that governs the transitions between the states. If $P[S_n | S_{n-1}]$ is independent of n (time), the Markov chain is called *homogeneous*. In this case the probability of transition from state i to state j , $1 \leq i, j \leq S$, is independent of n and is denoted by P_{ij}

$$P_{ij} = P[S_n = j | S_{n-1} = i] \quad (2.7-45)$$

In a homogeneous Markov chain, we define the *state transition matrix*, or *one-step transition matrix*, $\mathbf{P}$ as a matrix with elements P_{ij} . The element at row i and column j denotes the probability of a direct transition from state i to state j . $\mathbf{P}$ is a matrix with nonnegative elements, and the sum of each row of it is equal to 1. The n -step transition matrix gives the probabilities of moving from i to j in n steps. For discrete-time homogeneous Markov chains, the n -step transition matrix is equal to $\mathbf{P}^n$. All Markov chains studied here are assumed to be homogeneous.

The row vector $\mathbf{p}(n) = [p_1(n) \ p_2(n) \ \dots \ p_S(n)]$, where $p_i(n)$ denotes the probability of being in state i at time n , is the *state probability vector* of the Markov chain at time n . From this definition it is clear that

$$\mathbf{p}(n) = \mathbf{p}(n-1)\mathbf{P} \quad (2.7-46)$$

and

$$\mathbf{p}(n) = \mathbf{p}(0)\mathbf{P}^n \quad (2.7-47)$$

[†]Strictly speaking, this is the definition of a finite-state Markov chain (FSMC), which is the only class of Markov chains studied in this book.

If $\lim_{n \rightarrow \infty} \mathbf{P}^n$ exists and all its rows are equal, we denote each row of the limit by $\mathbf{p}$, i.e.,

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{bmatrix} \mathbf{p} \\ \mathbf{p} \\ \vdots \\ \mathbf{p} \end{bmatrix} \quad (2.7-48)$$

In this case

$$\lim_{n \rightarrow \infty} \mathbf{p}(n) = \lim_{n \rightarrow \infty} \mathbf{p}(0) \mathbf{P}^n = \mathbf{p}(0) \begin{bmatrix} \mathbf{p} \\ \mathbf{p} \\ \vdots \\ \mathbf{p} \end{bmatrix} = \mathbf{p} \quad (2.7-49)$$

This means that starting from *any* initial probability vector $\mathbf{p}(0)$, the Markov chain stabilizes at the state probability vector given by $\mathbf{p}$, which is called the *steady-state*, *equilibrium*, or *stationary* state probability distribution of the Markov chain. Since after reaching the steady-state probability distribution these probabilities do not change, $\mathbf{p}$ can be obtained as the solution of the equation

$$\mathbf{p} \mathbf{P} = \mathbf{p} \quad (2.7-50)$$

that satisfies the conditions $p_i \geq 0$ and $\sum_i p_i = 1$ (i.e., it is a probability vector). If a Markov chain starts from state $\mathbf{p}$, then it will always remain in this state, because $\mathbf{p} \mathbf{P} = \mathbf{p}$. Some basic questions are the following: Does $\mathbf{p} \mathbf{P} = \mathbf{p}$ always have a solution that is a probability vector? If yes, under what conditions is this solution unique? Under what conditions does $\lim_{n \rightarrow \infty} \mathbf{P}^n$ exist? If the limit exists, does the limit have equal rows?

If it is possible to move from any state of a Markov chain to any other state in a finite number of steps, the Markov chain is called *irreducible*. The *period* of state i of a Markov chain is the greatest common divisor (GCD) of all n such that $P_{ii}(n) > 0$. State i is *aperiodic* if its period is equal to 1. A finite-state Markov chain is called *ergodic* if it is irreducible and all its states are aperiodic.

It can be shown that in an ergodic Markov chain $\lim_{n \rightarrow \infty} \mathbf{P}^n$ always exists and all rows of the limit are equal, i.e., Equation 2.7-48 holds. In this case a unique stationary (steady-state) state probability distribution exists and starting from any initial state probability vector, the Markov chain ends up in the steady-state state probability vector $\mathbf{p}$.

EXAMPLE 2.7-2. A Markov chain with four states is described by the finite-state diagram shown in Figure 2.7-1. For this Markov chain we have

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{6} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{5}{6} & 0 & \frac{1}{6} & 0 \end{bmatrix} \quad (2.7-51)$$

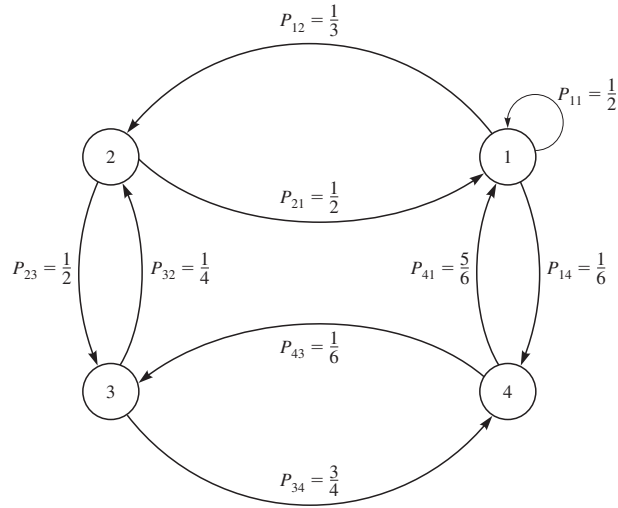


FIGURE 2.7-1
State transition diagram for a FSMC.

It is easily verified that this Markov chain is irreducible and aperiodic, and thus ergodic. To find the steady-state probability distribution, we can either find the limit of $\mathbf{P}^n$ as $n \rightarrow \infty$ or solve Equation 2.7-50. The result is

$$\mathbf{p} \approx [0.49541 \quad 0.19725 \quad 0.12844 \quad 0.17889] \quad (2.7-52)$$

2.8

SERIES EXPANSION OF RANDOM PROCESSES

Series expansion of random processes results in expressing the random processes in terms of a sequence of random variables as coefficients of orthogonal or orthonormal basis functions. This type of expansion reduces working with random processes to working with random variables, which in many cases are easier to handle. In the following we describe two types of series expansions for random processes. First we describe the sampling theorem for band-limited random processes, and then we continue with the Karhunen-Loeve expansion of random processes, which is a more general expansion.

2.8-1 Sampling Theorem for Band-Limited Random Processes

A deterministic real signal $x(t)$ with Fourier transform $X(f)$ is called band-limited if $X(f) = 0$ for $|f| > W$, where W is the highest frequency contained in $x(t)$. Such a signal is uniquely represented by samples of $x(t)$ taken at a rate of $f_s \geq 2W$ samples/s. The minimum rate $f_N = 2W$ samples/s is called the *Nyquist rate*. For complex-valued signals W is one-half of the frequency support of the signal; i.e., if W_1 and W_2

are the lowest and the highest frequency components of the signal, respectively, then $2W = W_2 - W_1$. The signal can be perfectly reconstructed from its sampled values if the sampling rate is at least equal to $2W$. The difference, however, is that the sampled values are complex in this case, and for specifying each sample, two real numbers are required. This means that a real signal can be perfectly described in terms of $2W$ real numbers per second, or it has $2W$ *degrees of freedom* or *real dimensions* per second. For a complex signal the number of degrees of freedom is $4W$ per second, which is equivalent to $2W$ *complex dimensions* or $4W$ real dimensions per second.

Sampling below the Nyquist rate results in frequency aliasing. The band-limited signal sampled at the Nyquist rate can be reconstructed from its samples by use of the interpolation formula

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2W}\right) \text{sinc}\left[2W\left(t - \frac{n}{2W}\right)\right] \quad (2.8-1)$$

where $\{x(n/2W)\}$ are the samples of $x(t)$ taken at $t = n/2W$, $n = 0, \pm 1, \pm 2, \dots$. Equivalently, $x(t)$ can be reconstructed by passing the sampled signal through an ideal lowpass filter with impulse response $h(t) = \text{sinc}(2Wt)$. Figure 2.8-1 illustrates the signal reconstruction process based on ideal interpolation. Note that the expansion of $x(t)$ as given by Equation 2.8-1 is an orthogonal expansion and not an orthonormal expansion since

$$\int_{-\infty}^{\infty} \text{sinc}\left[2W\left(t - \frac{n}{2W}\right)\right] \text{sinc}\left[2W\left(t - \frac{m}{2W}\right)\right] dt = \begin{cases} \frac{1}{2W} & n = m \\ 0 & n \neq m \end{cases} \quad (2.8-2)$$

A stationary stochastic process $X(t)$ is said to be band-limited if its power spectral density $\mathcal{S}_X(f) = 0$ for $|f| > W$. Since $\mathcal{S}_X(f)$ is the Fourier transform of the autocorrelation function $R_X(\tau)$, it follows that $R_X(\tau)$ can be represented as

$$R_X(\tau) = \sum_{n=-\infty}^{\infty} R_X\left(\frac{n}{2W}\right) \text{sinc}\left[2W\left(\tau - \frac{n}{2W}\right)\right] \quad (2.8-3)$$

where $\{R_X(n/2W)\}$ are samples of $R_X(\tau)$ taken at $\tau = n/2W$, $n = 0, \pm 1, \pm 2, \dots$. Now, if $X(t)$ is a band-limited stationary stochastic process, then $X(t)$ can be represented as

$$X(t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2W}\right) \text{sinc}\left[2W\left(t - \frac{n}{2W}\right)\right] \quad (2.8-4)$$

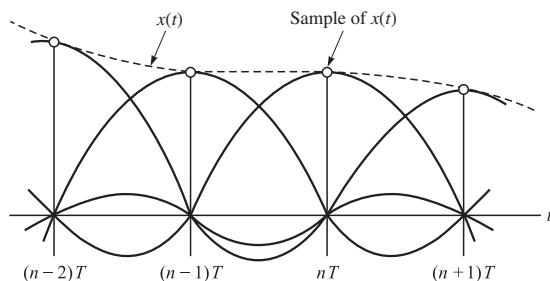


FIGURE 2.8-1
Sampling and reconstruction from samples.

where $\{X(n/2W)\}$ are samples of $X(t)$ taken at $t = n/2W, n = 0, \pm 1, \pm 2, \dots$. This is the sampling representation for a stationary stochastic process. The samples are random variables that are described statistically by appropriate joint probability density functions. If $X(t)$ is a WSS process, then random variables $\{X(n/2W)\}$ represent a WSS discrete-time random process. The autocorrelation of the sample random variables is given by

$$\begin{aligned} E \left[X \left(\frac{n}{2W} \right) X^* \left(\frac{m}{2W} \right) \right] &= R_X \left(\frac{n-m}{2W} \right) \\ &= \int_{-W}^W \mathcal{S}_X(f) e^{j2\pi f \frac{n-m}{2W}} df \end{aligned} \quad (2.8-5)$$

If the process $X(t)$ is filtered white Gaussian noise, then it is zero-mean and its power spectrum is flat in the $[-W, W]$ interval. In this case the samples are uncorrelated, and since they are Gaussian, they are independent as well.

The signal representation in Equation 2.8-4 is easily established by showing that (Problem 2.44)

$$E \left[\left| X(t) - \sum_{n=-\infty}^{\infty} X \left(\frac{n}{2W} \right) \text{sinc} \left[2W \left(t - \frac{n}{2W} \right) \right] \right|^2 \right] = 0 \quad (2.8-6)$$

Hence, equality between the sampling representation and the stochastic process $X(t)$ holds in the sense that the mean square error is zero.

2.8-2 The Karhunen-Loève Expansion

The sampling theorem presented above gives a straightforward method for orthogonal expansion of band-limited processes. In this section we present the Karhunen-Loève expansion, an orthonormal expansion that applies to a large class of random processes and results in uncorrelated random variables as expansion coefficients. We present only the results of the Karhunen-Loève expansion. The reader is referred to Van Trees (1968) or Loève (1955) for details.

There are many ways in which a random process can be expanded in terms of a sequence of random variables $\{X_n\}$ and an orthonormal basis $\{\phi_n(t)\}$. However, if we require the additional condition that the random variables X_n be mutually uncorrelated, then the orthonormal bases have to be the solutions of an eigenfunction problem given by an integral equation whose kernel is the autocovariance function of the random process. Solving this integral equation results in the orthonormal basis $\{\phi_n(t)\}$, and projecting the random process on this basis results in the sequence of uncorrelated random variables $\{X_n\}$.

The Karhunen-Loève expansion states that under mild conditions, a random process $X(t)$ with autocovariance function

$$C_X(t_1, t_2) = R_X(t_1, t_2) - m_X(t_1)m_X^*(t_2) \quad (2.8-7)$$

can be expanded over an interval of interest $[a, b]$ in terms of an orthonormal basis $\{\phi_n(t)\}_{n=1}^{\infty}$ such that the coefficients of expansion are uncorrelated. The $\phi_n(t)$'s are solutions (eigenfunctions) of the integral equation

$$\int_a^b C_X(t_1, t_2) \phi_n(t_2) dt_2 = \lambda_n \phi_n(t_1), \quad a < t_1 < b \quad (2.8-8)$$

with appropriate normalization such that

$$\int_a^b |\phi_n(t)|^2 dt = 1 \quad (2.8-9)$$

The Karhunen-Loève expansion is given by

$$\hat{X}(t) = \sum_{n=1}^{\infty} X_n \phi_n(t), \quad a < t < b$$

with the following properties:

1. Random variables X_n denoting the coefficients of the expansion are projections of the random process $X(t)$ on the basis functions, i.e.,

$$X_n = \langle X(t), \phi_n(t) \rangle = \int_a^b X(t) \phi_n^*(t) dt \quad (2.8-10)$$

2. Random variables X_n are mutually uncorrelated. Moreover, the variance of X_n is λ_n .

$$\text{COV}[X_n, X_m] = \begin{cases} \lambda_n & n = m \\ 0 & n \neq m \end{cases} \quad (2.8-11)$$

3. We have

$$E[\hat{X}(t)] = E[X(t)] = m_X(t), \quad a < t < b \quad (2.8-12)$$

4. $\hat{X}(t)$ is equal to $X(t)$ in the mean square sense

$$E[|X(t) - \hat{X}(t)|^2] = 0, \quad a < t < b \quad (2.8-13)$$

5. The covariance $C_X(t_1, t_2)$ can be expanded in terms of the bases and the eigenvalues as given in Equation 2.8-14. This result is known as *Mercer's theorem*.

$$C_X(t_1, t_2) = \sum_{n=1}^{\infty} \lambda_n \phi_n(t_1) \phi_n(t_2), \quad a < t_1, t_2 < b \quad (2.8-14)$$

6. The eigenfunctions $\{\phi_n(t)\}_{n=1}^{\infty}$ form a complete basis for expansion of all signals $g(t)$ which have finite energy in the interval $[a, b]$. In other words, if $g(t)$ is such that

$$\int_a^b |g(t)|^2 dt < \infty$$

then we can expand it in terms of $\{\phi_n(t)\}$ as

$$g(t) = \sum_{n=1}^{\infty} g_n \phi_n(t), \quad a < t < b \quad (2.8-15)$$

where

$$g_n = \langle g(t), \phi_n(t) \rangle = \int_a^b g(t) \phi_n^*(t) dt \quad (2.8-16)$$

Equation 2.8-13, which states the Karhunen-Loève expansion, is usually written in the form

$$X(t) = \sum_{n=1}^{\infty} X_n \phi_n(t), \quad a < t < b \quad (2.8-17)$$

where it is understood that the equality is in the mean square sense. The $\{\phi_n(t)\}$ are obtained by solving Equation 2.8-8 and normalizing the solutions, and the coefficients $\{X_n\}$ are obtained by using Equation 2.8-10.

It is worthwhile noting that the Karhunen-Loève expansion applies to both WSS and nonstationary processes. In the special case where the process is zero-mean, the autocovariance function $C_X(t_1, t_2)$ is substituted with the autocorrelation function $R_X(t_1, t_2)$. If the process $X(t)$ is a Gaussian process, $\{X_n\}$ are independent Gaussian random variables.

EXAMPLE 2.8-1. Let $X(t)$ be a zero-mean white process with power spectral density $\frac{N_0}{2}$. To derive the Karhunen-Loève expansion for this process over an arbitrary interval $[a, b]$, we have to solve the integral equation

$$\int_a^b \frac{N_0}{2} \delta(t_1 - t_2) \phi_n(t_2) dt_2 = \lambda_n \phi_n(t_1), \quad a < t_1 < b \quad (2.8-18)$$

where $\frac{N_0}{2} \delta(t_1 - t_2)$ is the autocorrelation function of the white process. Using the sifting property of the impulse function, we have

$$\frac{N_0}{2} \phi_n(t_1) = \lambda_n \phi_n(t_1), \quad a < t_1 < b \quad (2.8-19)$$

From this equation we see that $\phi_n(t)$ can be any arbitrary function. Therefore, any orthonormal basis can be used for expansion of white processes, and all coefficients of the expansion X_n will have the same variance of $\frac{N_0}{2}$.

2.9

BANDPASS AND LOWPASS RANDOM PROCESSES

In general, bandpass and lowpass random processes can be defined as WSS processes $X(t)$ for which the autocorrelation function $R_X(\tau)$ is either a bandpass or a lowpass signal. Recall that the autocorrelation function is an ordinary deterministic function with a Fourier transform which represents the power spectral density of the random

process $X(t)$. Therefore, for a bandpass process the power spectral density is located around frequencies $\pm f_0$, and for lowpass processes the power spectral density is located around zero frequency.

To be more specific, we define a bandpass (or narrowband) process as a real, zero-mean, and WSS random process whose autocorrelation function is a bandpass signal.

Inspired by Equations 2.1–11, we define the *in-phase* and *quadrature* components of a bandpass random process $X(t)$ as

$$\begin{aligned} X_i(t) &= X(t) \cos 2\pi f_0 t + \hat{X}(t) \sin 2\pi f_0 t \\ X_q(t) &= \hat{X}(t) \cos 2\pi f_0 t - X(t) \sin 2\pi f_0 t \end{aligned} \quad (2.9-1)$$

We will now show that

1. $X_i(t)$ and $X_q(t)$ are jointly WSS zero-mean random processes.
2. $X_i(t)$ and $X_q(t)$ have the same power spectral density.
3. $X_i(t)$ and $X_q(t)$ are both lowpass processes; i.e., their power spectral density is located around $f = 0$.

We also define the *lowpass equivalent process* $X_l(t)$ as

$$X_l(t) = X_i(t) + j X_q(t) \quad (2.9-2)$$

and we will derive an expression for its autocorrelation function and power spectral density. In addition we will see that $X_l(t)$ is a proper random process.

Since $X(t)$ by assumption is zero-mean, so is $\hat{X}(t)$, its Hilbert transform. This is obvious since the Hilbert transform is just a filtering operation. From this observation, it is clear that $X_i(t)$ and $X_q(t)$ are both zero-mean processes.

To derive the autocorrelation function of $X_l(t)$, we have

$$\begin{aligned} R_{X_l}(t + \tau, t) &= E[X_l(t + \tau)X_l(t)] \\ &= E[(X(t + \tau) \cos 2\pi f_0(t + \tau) + \hat{X}(t + \tau) \sin 2\pi f_0(t + \tau)) \\ &\quad \times (X(t) \cos 2\pi f_0 t + \hat{X}(t) \sin 2\pi f_0 t)] \end{aligned} \quad (2.9-3)$$

Expanding this relation, we have

$$\begin{aligned} R_{X_l}(t + \tau, t) &= R_X(\tau) \cos 2\pi f_0(t + \tau) \cos 2\pi f_0 t \\ &\quad + R_{X\hat{X}}(t + \tau, t) \cos 2\pi f_0(t + \tau) \sin 2\pi f_0 t \\ &\quad + R_{\hat{X}X}(t + \tau, t) \sin 2\pi f_0(t + \tau) \cos 2\pi f_0 t \\ &\quad + R_{\hat{X}\hat{X}}(t + \tau, t) \sin 2\pi f_0(t + \tau) \sin 2\pi f_0 t \end{aligned} \quad (2.9-4)$$

Since the Hilbert transform is the result of passing the process through an LTI system, we conclude that $X(t)$ and $\hat{X}(t)$ are jointly WSS and therefore all the auto- and cross-correlations in Equation 2.9–4 are functions of τ only. Using Equations 2.7–17

and 2.7–18, we can easily show that (see Problem 2.56)

$$\begin{aligned} R_{X\hat{X}}(\tau) &= -\hat{R}_X(\tau) \\ R_{\hat{X}X}(\tau) &= \hat{R}_X(\tau) \\ R_{\hat{X}\hat{X}}(\tau) &= R_X(\tau) \end{aligned} \quad (2.9-5)$$

Substituting these results into Equation 2.9–4 and using standard trigonometric identities yield

$$R_{X_i}(\tau) = R_X(\tau) \cos(2\pi f_0 \tau) + \hat{R}_X(\tau) \sin(2\pi f_0 \tau) \quad (2.9-6)$$

Similarly, we can show that

$$R_{X_q}(\tau) = R_{X_i}(\tau) = R_X(\tau) \cos(2\pi f_0 \tau) + \hat{R}_X(\tau) \sin(2\pi f_0 \tau) \quad (2.9-7)$$

$$R_{X_i X_q}(\tau) = -R_{X_q X_i}(\tau) = R_X(\tau) \sin(2\pi f_0 \tau) - \hat{R}_X(\tau) \cos(2\pi f_0 \tau) \quad (2.9-8)$$

These relations show that $X_i(t)$ and $X_q(t)$ are zero-mean jointly WSS processes with equal autocorrelation functions (and thus equal power spectral densities).

To derive the common power spectral density of $X_i(t)$ and $X_q(t)$ and their cross spectral density, we derive the Fourier transforms of Equations 2.9–7 and 2.9–8. We need to use the modulation property of the Fourier transform and the fact that the Fourier transform of $\hat{R}_X(\tau)$ is equal to $-j \operatorname{sgn}(f) \mathcal{S}_X(f)$. Given these facts, it is straightforward to derive

$$\begin{aligned} \mathcal{S}_{X_i}(f) = \mathcal{S}_{X_q}(f) &= \begin{cases} \mathcal{S}_X(f + f_0) + \mathcal{S}_X(f - f_0) & |f| < f_0 \\ 0 & \text{otherwise} \end{cases} \quad (2.9-9) \\ \mathcal{S}_{X_i X_q}(f) = -\mathcal{S}_{X_q X_i}(f) &= \begin{cases} j[\mathcal{S}_X(f + f_0) - \mathcal{S}_X(f - f_0)] & |f| < f_0 \\ 0 & \text{otherwise} \end{cases} \quad (2.9-10) \end{aligned}$$

Equation 2.9–9 states that the common power spectral density of the in-phase and quadrature components of $X(t)$ is obtained by shifting the power spectral density of $X(t)$ to left and right by f_0 and adding the results and then removing all components outside $[-f_0, f_0]$. This result also shows that both $X_i(t)$ and $X_q(t)$ are lowpass processes. From Equation 2.9–10 we see that if $\mathcal{S}_X(f + f_0) = \mathcal{S}_X(f - f_0)$ for $|f| < f_0$, then $\mathcal{S}_{X_i X_q}(f) = 0$ and consequently, $R_{X_i X_q}(\tau) = 0$. Since $X_i(t)$ and $X_q(t)$ are zero-mean processes, from $R_{X_i X_q}(\tau) = 0$ we conclude that under this condition $X_i(t)$ and $X_q(t)$ are uncorrelated. One of the cases where we have $\mathcal{S}_X(f + f_0) = \mathcal{S}_X(f - f_0)$ for $|f| < f_0$ occurs when $\mathcal{S}_X(f)$ is symmetric around f_0 , in which case the in-phase and quadrature components will be uncorrelated processes.

We define the complex process $X_I(t) = X_i(t) + j X_q(t)$ as the lowpass equivalent of $X(t)$. Since $X_i(t)$ and $X_q(t)$ are both lowpass processes, we conclude that $X_I(t)$ is also a lowpass process. Comparing Equations 2.9–7 and 2.9–8 with Equations 2.7–39 and 2.7–40, we can conclude that $X_I(t)$ is a proper random process, and therefore, from

Equation 2.7–41, we have

$$R_{X_I}(\tau) = 2R_{X_i}(\tau) + 2jR_{X_q X_i}(\tau) \quad (2.9-11)$$

$$= 2[R_X(\tau) + j\widehat{R}_X(\tau)]e^{-j2\pi f_0 \tau} \quad (2.9-12)$$

where we have used Equations 2.9–7 and 2.9–8. Comparing Equations 2.9–12 and 2.1–6, we observe that $R_{X_I}(\tau)$ is twice the lowpass equivalent of $R_X(\tau)$. In other words, *the autocorrelation function of the lowpass equivalent process $X_I(t)$ is twice the lowpass equivalent of the autocorrelation function of the bandpass process $X(t)$.*

Taking Fourier transform of both sides of Equation 2.9–12, we obtain

$$\mathcal{S}_{X_I}(f) = \begin{cases} 4\mathcal{S}_X(f + f_0) & |f| < f_0 \\ 0 & \text{otherwise} \end{cases} \quad (2.9-13)$$

and consequently,

$$\mathcal{S}_X(f) = \frac{1}{4}[\mathcal{S}_{X_I}(f - f_0) + \mathcal{S}_{X_I}(-f - f_0)] \quad (2.9-14)$$

We also observe that if $X(t)$ is a Gaussian process, then $X_i(t)$, $X_q(t)$, and $X_I(t)$ will be jointly Gaussian processes; and since $X_I(t)$ is Gaussian, zero-mean, and proper, we conclude that $X_I(t)$ is a circular process as well. In this case if $\mathcal{S}_X(f + f_0) = \mathcal{S}_X(f - f_0)$ for $|f| < f_0$, then $X_i(t)$ and $X_q(t)$ will be independent processes.

EXAMPLE 2.9–1. White Gaussian noise with power spectral density of $\frac{N_0}{2}$ passes through an ideal bandpass filter with transfer function

$$H(f) = \begin{cases} 1 & |f - f_0| < W \\ 0 & \text{otherwise} \end{cases}$$

where $W < f_0$. The output, called *filtered white noise*, is denoted by $X(t)$. This process has a power spectral density of

$$\mathcal{S}_X(f) = \begin{cases} \frac{N_0}{2} & |f - f_0| < W \\ 0 & \text{otherwise} \end{cases}$$

Since $\mathcal{S}_X(f + f_0) = \mathcal{S}_X(f - f_0)$ for $|f| < f_0$, and the process is Gaussian, $X_i(f)$ and $X_q(f)$ are independent lowpass processes. Using Equation 2.9–9, we conclude that

$$\mathcal{S}_{X_i}(f) = \mathcal{S}_{X_q}(f) = \begin{cases} N_0 & |f| < W \\ 0 & \text{otherwise} \end{cases}$$

and from Equation 2.9–13, we conclude that

$$\mathcal{S}_{X_I}(f) = \begin{cases} 2N_0 & |f| < W \\ 0 & \text{otherwise} \end{cases}$$

■ 2.10

BIBLIOGRAPHICAL NOTES AND REFERENCES

In this chapter we have provided a review of basic concepts and definitions in signal analysis, the theory of probability, and stochastic processes. An advanced book on signal analysis that covers most of the material presented here in detail is the book by Franks (1969). The texts by Davenport and Root (1958), Davenport (1970), Papoulis and Pillai (2002), Peebles (1987), Helstrom (1991), Stark and Woods (2002), and Leon-Garcia (1994) provide engineering-oriented treatments of probability and stochastic processes. A more mathematical treatment of probability theory may be found in the text by Loève (1955). Finally, we cite the book by Miller (1964), which treats multidimensional Gaussian distributions.



PROBLEMS

2.1 Prove the following properties of Hilbert transforms:

- a. If $x(t) = x(-t)$, then $\hat{x}(t) = -\hat{x}(-t)$.
- b. If $x(t) = -x(-t)$, then $\hat{x}(t) = \hat{x}(-t)$.
- c. If $x(t) = \cos \omega_0 t$, then $\hat{x}(t) = \sin \omega_0 t$.
- d. If $x(t) = \sin \omega_0 t$, then $\hat{x}(t) = -\cos \omega_0 t$.
- e. $\hat{\hat{x}}(t) = -x(t)$

$$f. \int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} \hat{x}^2(t) dt$$

$$g. \int_{-\infty}^{\infty} x(t)\hat{x}(t) dt = 0$$

2.2 Let $x(t)$ and $y(t)$ denote two bandpass signals, and let $x_l(t)$ and $y_l(t)$ denote their lowpass equivalents with respect to some frequency f_0 . We know that in general $x_l(t)$ and $y_l(t)$ are complex signals.

1. Show that

$$\int_{-\infty}^{\infty} x(t)y(t) dt = \frac{1}{2} \operatorname{Re} \left[\int_{-\infty}^{\infty} x_l(t)y_l^*(t) dt \right]$$

2. From this conclude that $\mathcal{E}_x = \frac{1}{2} \mathcal{E}_{x_l}$, i.e., the energy in a bandpass signal is one-half the energy in its lowpass equivalent.

2.3 Suppose that $s(t)$ is either a real- or complex-valued signal that is represented as a linear combination of orthonormal functions $\{f_n(t)\}$, i.e.,

$$\hat{s}(t) = \sum_{k=1}^K s_k f_k(t)$$

where

$$\int_{-\infty}^{\infty} f_n(t)f_m^*(t) dt = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Determine the expressions for the coefficients $\{s_k\}$ in the expansion $\hat{s}_i(t)$ that minimize the energy

$$\mathcal{E}_e = \int_{-\infty}^{\infty} |s(t) - \hat{s}(t)|^2 dt$$

and the corresponding residual error $\mathcal{E}_e$.

- 2.4** Suppose that a set of M signal waveforms $\{s_{lm}(t)\}$ is complex-valued. Derive the equations for the Gram-Schmidt procedure that will result in a set of $N \leq M$ orthonormal signal waveforms.
- 2.5** Carry out the Gram-Schmidt orthogonalization of the signals in Figure 2.2–1(a) in the order $s_4(t)$, $s_3(t)$, $s_1(t)$, and thus obtain a set of orthonormal functions $\{f_m(t)\}$. Then determine the vector representation of the signals $\{s_n(t)\}$ by using the orthonormal functions $\{f_m(t)\}$. Also determine the signal energies.
- 2.6** Assuming that the set of signals $\{\phi_{nl}(t), n = 1, \dots, N\}$ is an orthonormal basis for representation of $\{s_{ml}(t), m = 1, \dots, M\}$, show that the set of functions given by Equation 2.2–54 constitutes a $2N$ orthonormal basis that is *sufficient* for representation of M bandpass signals given in Equation 2.2–55.
- 2.7** Show that
- $$\widetilde{\widehat{\phi}}(t) = -\widehat{\widetilde{\phi}}(t)$$
- where $\widehat{\phi}(t)$ denotes the Hilbert transform and ϕ and $\widetilde{\phi}$ are given by Equation 2.2–54.
- 2.8** Determine the correlation coefficients ρ_{km} among the four signal waveforms $\{s_i(t)\}$ shown in Figure 2.2–1 and their corresponding Euclidean distances.
- 2.9** Prove that $s_l(t)$ is generally a complex-valued signal, and give the condition under which it is real. Assume that $s(t)$ is a real-valued bandpass signal.
- 2.10** Consider the three waveforms $f_n(t)$ shown in Figure P2.10.

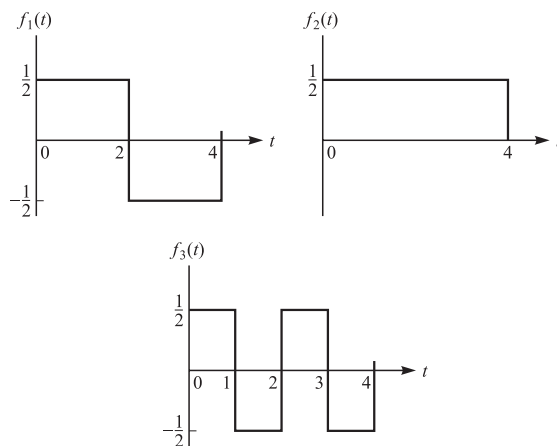


FIGURE P2.10

- a. Show that these waveforms are orthonormal.
 b. Express the waveform $x(t)$ as a linear combination of $f_n(t)$, $n = 1, 2, 3$, if

$$x(t) = \begin{cases} -1 & 0 \leq t < 1 \\ 1 & 1 \leq t < 3 \\ -1 & 3 \leq t < 4 \end{cases}$$

and determine the weighting coefficients.

2.11 Consider the four waveforms shown in Figure P2.11.

- a. Determine the dimensionality of the waveforms and a set of basis functions.
 b. Use the basis functions to represent the four waveforms by vectors s_1 , s_2 , s_3 , and s_4 .
 c. Determine the minimum distance between any pair of vectors.

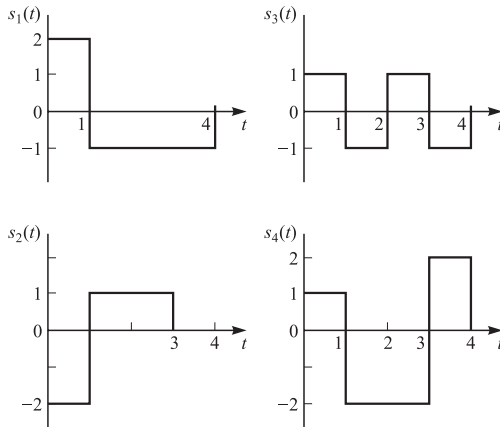


FIGURE P2.11

2.12 Determine a set of orthonormal functions for the four signals shown in Figure P2.12.

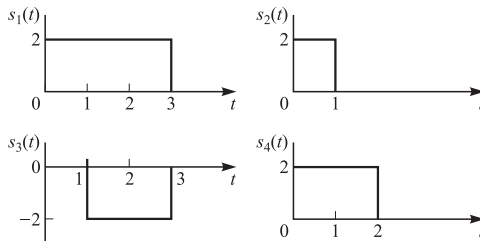


FIGURE P2.12

2.13 A random experiment consists of drawing a ball from an urn that contains 4 red balls numbered 1, 2, 3, 4 and three black balls numbered 1, 2, 3. The following events are defined.

1. E_1 = The number on the ball is even.
2. E_2 = The color of the ball is red, and its number is greater than 1.
3. E_3 = The number on the ball is less than 3.
4. $E_4 = E_1 \cup E_3$
5. $E_5 = E_1 \cup (E_2 \cap E_3)$

Answer the following questions.

1. What is $P(E_2)$?
2. What is $P(E_3|E_2)$?
3. What is $P(E_2|E_4E_3)$?
4. Are E_3 and E_5 independent?

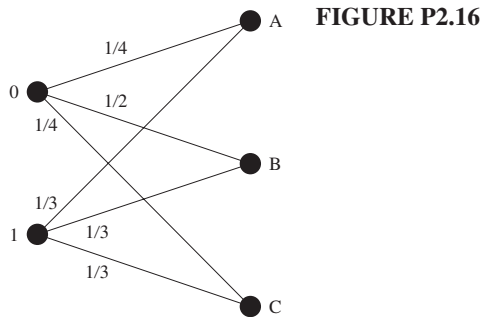
2.14 In a certain city three car brands A, B, C have 20%, 30% and 50% of the market share, respectively. The probability that a car needs major repair during its first year of purchase for the three brands is 5%, 10%, and 15%, respectively.

1. What is the probability that a car in this city needs major repair during its first year of purchase?
2. If a car in this city needs major repair during its first year of purchase, what is the probability that it is made by manufacturer A?

2.15 The random variables X_i , $i = 1, 2, \dots, n$, have joint PDF $p(x_1, x_2, \dots, x_n)$. Prove that

$$p(x_1, x_2, x_3, \dots, x_n) = p(x_n|x_{n-1}, \dots, x_1)p(x_{n-1}|x_{n-2}, \dots, x_1) \cdots p(x_3|x_2, x_1)p(x_2|x_1)p(x_1)$$

2.16 A communication channel with binary input and ternary output alphabets is shown in Figure P2.16. The probability of the input being 0 is 0.4. The transition probabilities are shown on the figure.



1. If the channel output is A, what is the best decision on channel input that minimizes the error probability? Repeat for the cases where channel output is B and C.
2. If a 0 is transmitted and an optimal decision scheme (the one derived in part 1) is used at the receiver, what is the probability of error?
3. What is the overall error probability for this channel if the optimal decision scheme is used at the receiver.

2.17 The PDF of a random variable X is $p(x)$. A random variable Y is defined as

$$Y = aX + b$$

where $a < 0$. Determine the PDF of Y in terms of the PDF of X .

2.18 Suppose that X is a Gaussian random variable with zero mean and unit variance. Let

$$Y = aX^3 + b, \quad a > 0$$

Determine and plot the PDF of Y .

2.19 The noise voltage in an electric circuit can be modeled as a Gaussian random variable with mean equal to zero and variance equal to 10^{-8} .

1. What is the probability that the value of the noise exceeds 10^{-4} ? What is the probability that it exceeds 4×10^{-4} ? What is the probability that the noise value is between -2×10^{-4} and 10^{-4} ?
2. Given that the value of the noise is positive, what is the probability that it exceeds 10^{-4} ?

2.20 X is a $\mathcal{N}(0, \sigma^2)$ random variable. This random variable is passed through a system whose input-output relation is given by $y = g(x)$. Find the PDF or the PMF of the output random variable Y in each of the following cases.

1. Square-law device, $g(x) = ax^2$.
2. Limiter,

$$g(x) = \begin{cases} -b & x \leq -b \\ b & x \geq b \\ x & |x| < b \end{cases}$$

3. Hard limiter,

$$g(x) = \begin{cases} a & x > 0 \\ 0 & x = 0 \\ b & x < 0 \end{cases}$$

4. Quantizer, $g(x) = x_n$ for $a_n \leq x < a_{n+1}$, $1 \leq n \leq N$, where x_n lies in the interval $[a_n, a_{n+1}]$ and the sequence $\{a_1, a_2, \dots, a_{N+1}\}$ satisfies the conditions $a_1 = -\infty$, $a_{N+1} = \infty$ and for $i > j$ we have $a_i > a_j$.

2.21 Shows that for an $\mathcal{N}(m, \sigma^2)$ random variable we have

$$E[(X - m)^n] = \begin{cases} 1 \times 3 \times 5 \times \dots \times (2k - 1) \sigma^{2k} = \frac{(2k)! \sigma^{2k}}{2^k k!} & \text{for } n = 2k \\ 0 & \text{for } n = 2k + 1 \end{cases}$$

2.22 *a.* Let X_r and X_i be statistically independent zero-mean Gaussian random variables with identical variance. Show that a (rotational) transformation of the form

$$Y_r + jY_i = (X_r + jX_i)e^{j\phi}$$

results in another pair (Y_r, Y_i) of Gaussian random variables that have the same joint PDF as the pair (X_r, X_i) .

- b.* Note that

$$\begin{bmatrix} Y_r \\ Y_i \end{bmatrix} = \mathbf{A} \begin{bmatrix} X_r \\ X_i \end{bmatrix}$$

where $\mathbf{A}$ is a 2×2 matrix. As a generalization of the two-dimensional transformation of the Gaussian random variables considered in (a), what property must the linear transformation $\mathbf{A}$ satisfy if the PDFs for $\mathbf{X}$ and $\mathbf{Y}$, where $\mathbf{Y} = \mathbf{A}\mathbf{X}$, $\mathbf{X} = (X_1 X_2 \dots X_n)$, and $\mathbf{Y} = (Y_1 Y_2 \dots Y_n)$ are identical?

2.23 Show that if $\mathbf{X}$ is a Gaussian vector, the random vector $\mathbf{Y} = \mathbf{A}\mathbf{X}$, where the invertible matrix $\mathbf{A}$ represents a linear transformation, is also a Gaussian vector whose mean and

covariance matrix are given by

$$\begin{aligned}\mathbf{m}_Y &= \mathbf{A}\mathbf{m}_X \\ \mathbf{C}_Y &= \mathbf{A}\mathbf{C}_X\mathbf{A}^t\end{aligned}$$

2.24 The random variable Y is defined as

$$Y = \sum_{i=1}^n X_i$$

where the X_i , $i = 1, 2, \dots, n$, are statistically independent random variables with

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases}$$

- Determine the characteristic function of Y .
- From the characteristic function, determine the moments $E(Y)$ and $E(Y^2)$.

2.25 This problem provides some useful bounds on $Q(x)$.

- By integrating $e^{-\frac{u^2+v^2}{2}}$ on the region $u > x$ and $v > x$ in $\mathbb{R}^2$, where $x > 0$, then changing to polar coordinates and upper bounding the integration region by the region $r > \sqrt{2}x$ in the first quadrant, show that $Q(x) \leq \frac{1}{2}e^{-\frac{x^2}{2}}$ for all $x \geq 0$.
- Apply integration by parts to

$$\int_x^\infty e^{-\frac{y^2}{2}} \frac{dy}{y^2}$$

and show that

$$\frac{x}{\sqrt{2\pi}(1+x^2)}e^{-\frac{x^2}{2}} < Q(x) < \frac{1}{\sqrt{2\pi}x}e^{-\frac{x^2}{2}}$$

for all $x > 0$.

- Based on the result of part 2 show that, for large x ,

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

2.26 Let $X_1, X_2, X_3, \dots$ denote iid random variables each uniformly distributed on $[0, A]$, where $A > 0$. Let $Y_n = \min\{X_1, X_2, \dots, X_n\}$.

- What is the PDF of Y_n ?
- Show that if both A and n go to infinity such that $\frac{n}{A} = \lambda$, where $\lambda > 0$ is a constant, the density function of Y_n tends to an exponential density function. Specify this density function.

2.27 The four random variables X_1, X_2, X_3, X_4 are zero-mean jointly Gaussian random variables with covariance $C_{ij} = E(X_i X_j)$ and characteristic function $\Phi_X(\omega_1, \omega_2, \omega_3, \omega_4)$. Show that

$$E(X_1 X_2 X_3 X_4) = C_{12}C_{34} + C_{13}C_{24} + C_{14}C_{23}$$

2.28 Let

$$\Theta_X(t) = E[e^{tX}]$$

denote the moment generating function of random variable X .

1. Using the Chernov bound, show that

$$\ln P[X \geq \alpha] \leq -\max_{t \geq 0}(\alpha t - \ln \Theta_X(t))$$

2. Define

$$I(\alpha) = \max_{t \geq 0}(\alpha t - \ln \Theta_X(t))$$

as the *large-deviation rate function* of the random variable X , and let $X_1, X_2, \dots, X_n$ be iid. Define $S_n = (X_1 + X_2 + \dots + X_n)/n$. Show that for $\alpha \geq E[X]$

$$\frac{1}{n} \ln P[S_n \geq \alpha] \leq -I(\alpha)$$

or equivalently

$$P[S_n \geq \alpha] \leq e^{-nI(\alpha)}$$

Note: It can be shown that for $\alpha \geq E[X]$, we have $P[S_n \geq \alpha] = e^{-nI(\alpha) + o(n)}$, where $o(n) \rightarrow 0$ as $n \rightarrow \infty$. This result is known as the *large-deviation theorem*.

3. Now assume the X_i 's are exponential, i.e.,

$$p_X(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Using the large-deviation result, show that

$$P[S_n \geq \alpha] = \alpha^n e^{-n(\alpha-1) + o(n)}$$

for $\alpha \geq 1$.

2.29 From the characteristic functions for the central chi-square and noncentral chi-square random variables given in Table 2.3–3, determine their corresponding first and second moments.

2.30 The PDF of a Cauchy distributed random variable X is

$$p(x) = \frac{a/\pi}{x^2 + a^2}, \quad -\infty < x < \infty$$

a. Determine the mean and variance of X .

b. Determine the characteristic function of X .

2.31 Let R_0 denote a Rayleigh random variable with PDF

$$f_{R_0}(r_0) = \begin{cases} \frac{r_0}{\sigma^2} e^{-\frac{r_0^2}{2\sigma^2}} & r_0 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

and R_1 be Ricean with PDF

$$f_{R_1}(r_1) = \begin{cases} \frac{r_1}{\sigma^2} I_0\left(\frac{\mu r_1}{\sigma^2}\right) e^{-\frac{r_1^2 + \mu^2}{2\sigma^2}} & r_1 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Furthermore, assume that R_0 and R_1 are *independent*. Show that

$$P(R_0 > R_1) = \frac{1}{2} e^{-\frac{\mu^2}{4\sigma^2}}$$

2.32 Suppose that we have a complex-valued Gaussian random variable $Z = X + jY$, where (X, Y) are statistically independent variables with zero mean and variance $E[X^2] = E[Y^2] = \sigma^2$. Let $R = Z + m$, where $m = m_r + jm_i$ and define R as $R = A + jB$. Clearly, $A = X + m_r$ and $B = Y + m_i$. Determine the following probability density functions:

1. $p_{A,B}(a, b)$
2. $p_{U,\Phi}(u, \phi)$, where $U = \sqrt{A^2 + B^2}$ and $\Phi = \tan^{-1} B/A$
3. $p_U(u)$

Note: In part 2 it is convenient to define $\theta = \tan^{-1}(m_i/m_r)$ so that

$$m_r = \sqrt{m_r^2 + m_i^2} \cos \theta, \quad m_i = \sqrt{m_r^2 + m_i^2} \sin \theta$$

Furthermore, you must use Equation 2.3–34, defining $I_0(\cdot)$ as the modified Bessel function of order zero.

2.33 The random variable Y is defined as

$$Y = \frac{1}{n} \sum_{i=1}^n X_i$$

where X_i , $i = 1, 2, \dots, n$, are statistically independent and identically distributed random variables each of which has the Cauchy PDF given in Problem 2.30.

- a. Determine the characteristic function of Y .
- b. Determine the PDF of Y .
- c. Consider the PDF of Y in the limit as $n \rightarrow \infty$. Does the central limit theorem hold? Explain your answer.

2.34 Show that if $\mathbf{Z}$ is circular, then it is zero-mean and proper, i.e., $E[\mathbf{Z}] = \mathbf{0}$ and $E[\mathbf{Z}\mathbf{Z}'] = \mathbf{0}$.

2.35 Show that if $\mathbf{Z}$ is a zero-mean proper Gaussian complex vector, then $\mathbf{Z}$ is circular.

2.36 Show that if $\mathbf{Z}$ is a proper complex vector, then any transform of the form $\mathbf{W} = \mathbf{A}\mathbf{Z} + \mathbf{b}$ is also a proper complex vector.

2.37 Assume that random processes $X(t)$ and $Y(t)$ are individually and jointly stationary.

- a. Determine the autocorrelation function of $Z(t) = X(t) + Y(t)$.
- b. Determine the autocorrelation function of $Z(t)$ when $X(t)$ and $Y(t)$ are uncorrelated.
- c. Determine the autocorrelation function of $Z(t)$ when $X(t)$ and $Y(t)$ are uncorrelated and have zero means.

2.38 The autocorrelation function of a stochastic process $X(t)$ is

$$R_X(\tau) = \frac{1}{2} N_0 \delta(\tau)$$

Such a process is called *white noise*. Suppose $x(t)$ is the input to an ideal bandpass filter having the frequency response characteristic shown in Figure P2.38. Determine the total noise power at the output of the filter.

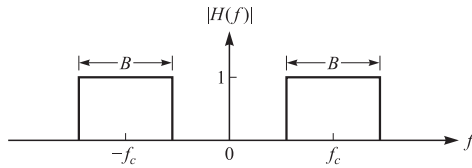


FIGURE P2.38

2.39 A lowpass Gaussian stochastic process $X(t)$ has a power spectral density

$$S(f) = \begin{cases} N_0 & |f| < B \\ 0 & \text{otherwise} \end{cases}$$

Determine the power spectral density and the autocorrelation function of $Y(t) = X^2(t)$.

2.40 The covariance matrix of three random variables X_1 , X_2 , and X_3 is

$$\begin{bmatrix} C_{11} & 0 & C_{13} \\ 0 & C_{22} & 0 \\ C_{31} & 0 & C_{33} \end{bmatrix}$$

The linear transformation $\mathbf{Y} = \mathbf{A}\mathbf{X}$ is made where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Determine the covariance matrix of $\mathbf{Y}$.

2.41 Let $X(t)$ be a stationary real normal process with zero mean. Let a new process $Y(t)$ be defined by

$$Y(t) = X^2(t)$$

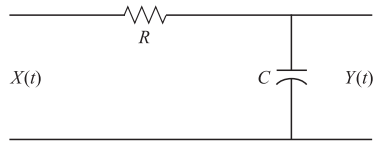
Determine the autocorrelation function of $Y(t)$ in terms of the autocorrelation function of $X(t)$. *Hint:* Use the result on Gaussian variables derived in Problem 2.27.

2.42 For the Nakagami PDF, given by Equation 2.3–67, define the normalized random variable $X = R/\sqrt{\Omega}$. Determine the PDF of X .

2.43 The input $X(t)$ in the circuit shown in Figure P2.43 is a stochastic process with $E[X(t)] = 0$ and $R_X(\tau) = \sigma^2 \delta(\tau)$; i.e., $X(t)$ is a white noise process.

a. Determine the spectral density $S_Y(f)$.

b. Determine $R_Y(\tau)$ and $E[Y^2(t)]$.

**FIGURE P2.43**

2.44 Demonstrate the validity of Equation 2.8–6.

2.45 Use the Chernoff bound to show that $Q(x) \leq e^{-x^2/2}$.

2.46 Determine the mean, the autocorrelation sequence, and the power density spectrum of the output of a system with unit sample response

$$h(n) = \begin{cases} 1 & n = 0 \\ -2 & n = 1 \\ 1 & n = 2 \\ 0 & \text{otherwise} \end{cases}$$

when the input $x(n)$ is a white noise process with variance σ_x^2 .

2.47 The autocorrelation sequence of a discrete-time stochastic process is $R(k) = \left(\frac{1}{2}\right)^{|k|}$. Determine its power density spectrum.

2.48 A discrete-time stochastic process $X(n) \equiv X(nT)$ is obtained by periodic sampling of a continuous-time zero-mean stationary process $X(t)$, where T is the sampling interval; i.e., $f_s = 1/T$ is the sampling rate.

- Determine the relationship between the autocorrelation function of $X(t)$ and the autocorrelation sequence of $X(n)$.
- Express the power density spectrum of $X(n)$ in terms of the power density spectrum of the process $X(t)$.
- Determine the conditions under which the power density spectrum of $X(n)$ is equal to the power density spectrum of $X(t)$.

2.49 The random process $V(t)$ is defined as

$$V(t) = X \cos 2\pi f_c t - Y \sin 2\pi f_c t$$

where X and Y are random variables. Show that $V(t)$ is wide-sense stationary if and only if $E(X) = E(Y) = 0$, $E(X^2) = E(Y^2)$, and $E(XY) = 0$.

2.50 Consider a band-limited zero-mean stationary stochastic process $X(t)$ with power density spectrum

$$\mathcal{S}_X(f) = \begin{cases} 1 & |f| \leq W \\ 0 & \text{otherwise} \end{cases}$$

$X(t)$ is sampled at a rate $f_s = 1/T$ to yield a discrete-time process $X(n) \equiv X(nT)$.

- Determine the expression for the autocorrelation sequence of $X(n)$.
- Determine the minimum value of T that results in a white (spectrally flat) sequence.

c. Repeat (b) if the power density spectrum of $X(t)$ is

$$\mathcal{S}_X(f) = \begin{cases} 1 - |f|/W & |f| \leq W \\ 0 & \text{otherwise} \end{cases}$$

2.51 Show that the functions

$$f_k(t) = \text{sinc} \left[2W \left(t - \frac{k}{2W} \right) \right], \quad k = 0, \pm 1, \pm 2, \dots$$

are orthogonal over the real line, i.e.,

$$\int_{-\infty}^{\infty} f_k(t) f_j(t) dt = \begin{cases} 1/2W & k = j \\ 0 & \text{otherwise} \end{cases}$$

Therefore, the sampling theorem reconstruction formula may be viewed as a series expansion of the band-limited signal $s(t)$, where the weights are samples of $s(t)$ and the $\{f_k(t)\}$ are the set of orthogonal functions used in the series expansion.

2.52 The noise equivalent bandwidth of a system is defined as

$$B_{\text{eq}} = \frac{1}{G} \int_0^{\infty} |H(f)|^2 df$$

where $G = \max |H(f)|^2$. Using this definition, determine the noise equivalent bandwidth of the ideal bandpass filter shown in Figure P2.38 and the low-pass system shown in Figure P2.43.

2.53 Suppose that $N(t)$ is a zero-mean stationary narrowband process. The autocorrelation function of the equivalent lowpass process $Z(t) = X(t) + jY(t)$ is defined as

$$R_Z(\tau) = E [Z^*(t)Z(t + \tau)]$$

a. Show that

$$E [Z(t)Z(t + \tau)] = 0$$

b. Suppose $R_z(\tau) = N_0\delta(\tau)$, and let

$$V = \int_0^T Z(t) dt$$

Determine $E [V^2]$ and $E [|V|^2]$.

2.54 Determine the autocorrelation function of the stochastic process

$$X(t) = A \sin(2\pi f_c t + \Theta)$$

where f_c is a constant and Θ is a uniformly distributed phase, i.e.,

$$p(\theta) = \frac{1}{2\pi}, \quad 0 \leq \theta \leq 2\pi$$

- 2.55** Let $Z(t) = X(t) + jY(t)$ be a complex random process, where $X(t)$ and $Y(t)$ are real-valued, independent, zero-mean, and jointly stationary Gaussian random processes. We assume that $X(t)$ and $Y(t)$ are both band-limited processes with a bandwidth of W and a flat spectral density within their bandwidth, i.e.,

$$S_X(f) = S_Y(f) = \begin{cases} N_0 & |f| \leq W \\ 0 & \text{otherwise} \end{cases}$$

1. Find $E[Z(t)]$ and $R_Z(t + \tau, t)$, and show that $Z(t)$ is WSS.
2. Find the power spectral density of $Z(t)$.
3. Assume $\phi_1(t), \phi_2(t), \dots, \phi_n(t)$ are orthonormal, i.e.,

$$\int_{-\infty}^{\infty} \phi_j(t) \phi_k^*(t) dt = \begin{cases} 1 & j = k \\ 0 & \text{otherwise} \end{cases}$$

and all $\phi_j(t)$'s are band-limited to $[-W, W]$. Define random variables Z_j as the projections of $Z(t)$ on the $\phi_j(t)$'s, i.e.,

$$Z_j = \int_{-\infty}^{\infty} Z(t) \phi_j^*(t) dt, \quad j = 1, 2, \dots, n$$

Determine $E[Z_j]$ and $E[Z_j Z_k^*]$ and conclude that the Z_j 's are iid zero-mean Gaussian random variables. Find their common variance.

4. Let $Z_j = Z_{jr} + jZ_{ji}$, where Z_{jr} and Z_{ji} denote the real and imaginary parts, respectively, of Z_j . Comment on the joint probability distribution of the $2n$ random variables

$$(Z_{1r}, Z_{1i}, Z_{2r}, Z_{2i}, \dots, Z_{nr}, Z_{ni})$$

5. Let us define

$$\hat{Z}(t) = Z(t) - \sum_{j=1}^n Z_j \phi_j(t)$$

to be the error in expansion of $Z(t)$ as a linear combination of $\phi_j(t)$'s. Show that $E[\hat{Z}(t) Z_k^*] = 0$ for all $k = 1, 2, \dots, n$. In other words, show that the error $\hat{Z}(t)$ and all the Z_k 's are uncorrelated. Can you say $\hat{Z}(t)$ and the Z_k 's are independent?

- 2.56** Let $X(t)$ denote a (real, zero-mean, WSS) bandpass process with autocorrelation function $R_X(\tau)$ and power spectral density $S_X(f)$, where $S_X(0) = 0$, and let $\hat{X}(t)$ denote the Hilbert transform of $X(t)$. Then $\hat{X}(t)$ can be viewed as the output of a filter, with impulse response $\frac{1}{\pi t}$ and transfer function $-j \text{sgn}(f)$, whose input is $X(t)$. Recall that when $X(t)$ passes through a system with transfer function $H(f)$ and the output is $Y(t)$, we have $S_Y(f) = S_X(f) |H(f)|^2$ and $S_{XY}(f) = S_X(f) H^*(f)$.

1. Prove that $R_{\hat{X}}(\tau) = R_X(\tau)$.
2. Prove that $R_{X\hat{X}}(\tau) = -\hat{R}_X(\tau)$.
3. If $Z(t) = X(t) + j\hat{X}(t)$, determine $S_Z(f)$.
4. Define $X_I(t) = Z(t)e^{-j2\pi f_0 t}$. Show that $X_I(t)$ is a lowpass WSS random process, and determine $S_{X_I}(f)$. From the expression for $S_{X_I}(f)$, derive an expression for $R_{X_I}(\tau)$.

2.57 A noise process has a power spectral density given by

$$S_n(f) = \begin{cases} 10^{-8} \left(1 - \frac{|f|}{10^8}\right) & |f| < 10^8 \\ 0 & |f| > 10^8 \end{cases}$$

This noise is passed through an ideal bandpass filter with a bandwidth of 2 MHz centered at 50 MHz.

1. Find the power content of the output process.
2. Write the output process in terms of the in-phase and quadrature components, and find the power in each component. Assume $f_0 = 50$ MHz.
3. Find the power spectral density of the in-phase and quadrature components.
4. Now assume that the filter is not an ideal filter and is described by

$$|H(f)|^2 = \begin{cases} \frac{|f|}{10^6} - 49 & 49 \text{ MHz} < |f| < 51 \text{ MHz} \\ 0 & \text{otherwise} \end{cases}$$

Repeat parts 1, 2, and 3 with this assumption.