A
RANDOM SIGNALS

In this appendix, we collect and summarize a number of results and establish the notation relating to the representation of stochastic signals. We make no attempt here to provide a detailed discussion of the difficult and subtle mathematical issues of the theory of random processes. Although our approach is not rigorous, we have summarized the important results and the mathematical assumptions implicit in their derivation.

A.1 DISCRETE-TIME RANDOM PROCESSES

The fundamental concept in the mathematical representation of stochastic signals is that of a *random process*. In our discussion of random processes as models for discrete-time signals, we assume that the reader is familiar with basic concepts of the theory of probability, such as random variables, probability distributions, and averages.

In utilizing the random-process model in practical signal-processing applications, we consider a particular sequence to be one of an ensemble of sample sequences. Given a discrete-time signal, the structure, i.e., the underlying probability law, of the corresponding random process is generally not known and must somehow be inferred. It may be possible to make reasonable assumptions about the structure of the process, or it may be possible to estimate the properties of a random-process representation from a finite segment of a typical sample sequence.

Formally, a random process is an indexed family of random variables $\{\mathbf{x}_n\}$ characterized by a set of probability distribution functions that, in general, may be a function of the index n. In using the concept of a random process as a model for discrete-time signals, the index n is associated with the time index. In other words, each sample value $x[n]$ of a random signal is assumed to have resulted from a mechanism that is governed

by a probability law. An individual random variable $\mathbf{x}_n$ is described by the probability distribution function

$$P_{\mathbf{x}_n}(x_n, n) = \text{Probability}\,[\mathbf{x}_n \leq x_n], \tag{A.1}$$

where $\mathbf{x}_n$ denotes the random variable and x_n is a particular value of $\mathbf{x}_n$.[1] If $\mathbf{x}_n$ takes on a continuous range of values, it is equivalently specified by the *probability density function*

$$p_{\mathbf{x}_n}(x_n, n) = \frac{\partial P_{\mathbf{x}_n}(x_n, n)}{\partial x_n}, \tag{A.2}$$

or

$$P_{\mathbf{x}_n}(x_n, n) = \int_{-\infty}^{x_n} p_{\mathbf{x}_n}(x, n)dx. \tag{A.3}$$

The interdependence of two random variables $\mathbf{x}_n$ and $\mathbf{x}_m$ of a random process is described by the joint probability distribution function

$$P_{\mathbf{x}_n,\mathbf{x}_m}(x_n, n, x_m, m) = \text{Probability}\,[\mathbf{x}_n \leq x_n \text{ and } \mathbf{x}_m \leq x_m] \tag{A.4}$$

and by the joint probability density

$$p_{\mathbf{x}_n,\mathbf{x}_m}(x_n, n, x_m, m) = \frac{\partial^2 P_{\mathbf{x}_n,\mathbf{x}_m}(x_n, n, x_m, m)}{\partial x_n \partial x_m}. \tag{A.5}$$

Two random variables are *statistically independent* if knowledge of the value of one does not affect the probability density of the other. If all the random variables of a collection of random variables, $\{\mathbf{x}_n\}$, are statistically independent, then

$$P_{\mathbf{x}_n,\mathbf{x}_m}(x_n, n, x_m, m) = P_{\mathbf{x}_n}(x_n, n) \cdot P_{\mathbf{x}_m}(x_m, m) \qquad m \neq n. \tag{A.6}$$

A complete characterization of a random process requires the specification of all possible joint probability distributions. As we have indicated, these probability distributions may be a function of the time indices m and n. In the case where all the probability distributions are independent of a shift of time origin, the random process is said to be *stationary*. For example, the second-order distribution of a stationary process satisfies

$$P_{\mathbf{x}_{n+k},\mathbf{x}_{m+k}}(x_{n+k}, n+k, x_{m+k}, m+k) = P_{\mathbf{x}_n\mathbf{x}_m}(x_n, n, x_m, m) \quad \text{for all } k. \tag{A.7}$$

In many of the applications of discrete-time signal processing, random processes serve as models for signals in the sense that a particular signal can be considered a sample sequence of a random process. Although the details of such signals are unpredictable—making a deterministic approach to signal representation inappropriate—certain average properties of the ensemble can be determined, given the probability law of the process. These average properties often serve as a useful, although incomplete, characterization of such signals.

[1] In this appendix, boldface type is used to denote the random variables and regular type denotes dummy variables of probability functions.

A.2 AVERAGES

It is often useful to characterize a random variable by averages such as the mean and variance. Since a random process is an indexed set of random variables, we may likewise characterize the process by statistical averages of the random variables making up the random process. Such averages are called *ensemble averages*. We begin the discussion of averages with some definitions.

A.2.1 Definitions

The average, or mean, of a random process is defined as

$$m_{\mathbf{x}_n} = \mathcal{E}\{\mathbf{x}_n\} = \int_{-\infty}^{\infty} x p_{\mathbf{x}_n}(x, n)dx, \tag{A.8}$$

where $\mathcal{E}$ denotes an operator called *mathematical expectation*. In general, the mean (expected value) may depend on n. In addition, if $g(\cdot)$ is a single-valued function, then $g(\mathbf{x}_n)$ is a random variable, and the set of random variables $\{g(\mathbf{x}_n)\}$ defines a new random process. To compute averages of this new process, we can derive probability distributions of the new random variables, or it can be shown that

$$\mathcal{E}\{g(\mathbf{x}_n)\} = \int_{-\infty}^{\infty} g(x) p_{\mathbf{x}_n}(x, n)dx. \tag{A.9}$$

If the random variables are discrete—i.e., if they have quantized values—the integrals become summations over all possible values of the random variable. In that case $\mathcal{E}\{g(x)\}$ has the form

$$\mathcal{E}\{g(\mathbf{x}_n)\} = \sum_x g(x)\hat{p}_{\mathbf{x}_n}(x, n). \tag{A.10}$$

In cases where we are interested in the relationship between two (or more) random processes, we must be concerned with two sets of random variables $\{\mathbf{x}_n\}$ and $\{\mathbf{y}_m\}$. For example, the expected value of a function of two random variables is defined as

$$\mathcal{E}\{g(\mathbf{x}_n, \mathbf{y}_m)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) p_{\mathbf{x}_n,\mathbf{y}_m}(x, n, y, m)dx\,dy, \tag{A.11}$$

where $p_{\mathbf{x}_n,\mathbf{y}_m}(x_m, n, y_m, m)$ is the joint probability density of the random variables $\mathbf{x}_n$ and $\mathbf{y}_m$.

The mathematical expectation operator is a linear operator; that is, it can be shown that

1. $\mathcal{E}\{\mathbf{x}_n + \mathbf{y}_m\} = \mathcal{E}\{\mathbf{x}_n\} + \mathcal{E}\{\mathbf{y}_m\}$; i.e., the average of a sum is the sum of the averages.
2. $\mathcal{E}\{a\mathbf{x}_n\} = a\mathcal{E}\{\mathbf{x}_n\}$; i.e., the average of a constant times $\mathbf{x}_n$ is equal to the constant times the average of $\mathbf{x}_n$.

In general, the average of a product of two random variables is not equal to the product of the averages. When this property holds, however, the two random variables are said to be *linearly independent* or *uncorrelated*. That is, $\mathbf{x}_n$ and $\mathbf{y}_m$ are linearly independent or uncorrelated if

$$\mathcal{E}\{\mathbf{x}_n\mathbf{y}_m\} = \mathcal{E}\{\mathbf{x}_m\} \cdot \mathcal{E}\{\mathbf{y}_m\}. \tag{A.12}$$

It is easy to see from Eqs. (A.11) and (A.12) that a sufficient condition for linear independence is

$$p_{\mathbf{x}_n,\mathbf{y}_m}(x_n, n, y_m, m) = p_{\mathbf{x}_n}(x_n, n) \cdot p_{\mathbf{y}_m}(y_m, m). \tag{A.13}$$

However, it can be shown that Eq. (A.13) is a stronger statement of independence than Eq. (A.12). As previously stated, random variables satisfying Eq. (A.13) are said to be *statistically independent.* If Eq. (A.13) holds for all values of n and m, the random processes $\{\mathbf{x}_n\}$ and $\{\mathbf{y}_m\}$ are said to be statistically independent. Statistically independent random processes are also linearly independent; but the converse is not true: Linear independence does not imply statistical independence.

It can be seen from Eqs. (A.9)–(A.11) that averages generally are functions of the time index. In the case of stationary processes, this is not true: For stationary processes, the mean is the same for all the random variables that constitute the process; i.e., the mean of a stationary process is a constant, which we denote simply m_x.

In addition to the mean of a random process, as defined in Eq. (A.8), a number of other averages are particularly important within the context of signal processing. These are defined next. For notational convenience, we assume that the probability distributions are continuous. Corresponding definitions for discrete random processes can be obtained by applying Eq. (A.10).

The *mean-square* value of $\mathbf{x}_n$ is the average of $|\mathbf{x}_n|^2$; i.e.,

$$\mathcal{E}\{|\mathbf{x}_n|^2\} = \text{mean square} = \int_{-\infty}^{\infty} |x|^2 p_{\mathbf{x}_n}(x, n)dx. \tag{A.14}$$

The mean-square value is sometimes referred to as the *average power.*

The *variance* of $\mathbf{x}_n$ is the mean-square value of $[\mathbf{x}_n - m_{x_n}]$; i.e.,

$$\text{var}[\mathbf{x}_n] = \mathcal{E}\{|(\mathbf{x}_n - m_{x_n})|^2\} = \sigma^2_{\mathbf{x}_n}. \tag{A.15}$$

Since the average of a sum is the sum of the averages, it can easily be shown that Eq. (A.15) can be written as

$$\text{var}[\mathbf{x}_n] = \mathbf{E}\{|\mathbf{x}_n|^2\} - |m_{x_n}|^2. \tag{A.16}$$

In general, the mean-square value and the variance are functions of time; however, they are constant for stationary processes.

The mean, mean square, and variance are simple averages that provide only a small amount of information about a process. A more useful average is the *autocorrelation sequence,* which is defined as

$$\begin{aligned} \phi_{xx}[n, m] &= \mathcal{E}\{\mathbf{x}_n \mathbf{x}_m^*\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_n x_m^* p_{\mathbf{x}_n,\mathbf{x}_m}(x_n, n, x_m, m)dx_n\, dx_m, \end{aligned} \tag{A.17}$$

where * denotes complex conjugation. The autocovariance sequence of a random process is defined as

$$\gamma_{xx}[n, m] = \mathcal{E}\{(\mathbf{x}_n - m_{x_n})(\mathbf{x}_m - m_{x_m})^*\}, \tag{A.18}$$

which can be written as

$$\gamma_{xx}[n, m] = \phi_{xx}[n, m] - m_{x_n} m_{x_m}^*. \tag{A.19}$$

Note that, in general, both the autocorrelation and autocovariance are two-dimensional sequences, i.e., functions of two variables.

The autocorrelation sequence is a measure of the dependence between values of the random processes at different times. In this sense, it partially describes the time variation of a random signal. A measure of the dependence between two different random signals is obtained from the cross-correlation sequence. If $\{\mathbf{x}_n\}$ and $\{\mathbf{y}_m\}$ are two random processes, their cross-correlation is

$$\begin{aligned}\phi_{xx}[n, m] &= \mathcal{E}\{\mathbf{x}_n \mathbf{y}_m^*\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy^* p_{\mathbf{x}_n, \mathbf{y}_m}(x, n, y, m) dx\, dy,\end{aligned} \tag{A.20}$$

where $p_{\mathbf{x}_n, \mathbf{y}_m}(x, n, y, m)$ is the joint probability density of $\mathbf{x}_n$ and $\mathbf{y}_m$. The cross-covariance function is defined as

$$\begin{aligned}\gamma_{xx}[n, m] &= \mathcal{E}\{(\mathbf{x}_n - m_{x_n})(\mathbf{y}_m - m_{y_m})^*\} \\ &= \phi_{xy}[n, m] - m_{x_n} m_{y_m}^*.\end{aligned} \tag{A.21}$$

As we have pointed out, the statistical properties of a random process generally vary with time. However, a stationary random process is characterized by an equilibrium condition in which the statistical properties are invariant to a shift of time origin. This means that the first-order probability distribution is independent of time. Similarly, all the joint probability functions are also invariant to a shift of time origin; i.e., the second-order joint probability distributions depend only on the time difference $(m - n)$. First-order averages such as the mean and variance are independent of time; second-order averages, such as the autocorrelation $\phi_{xx}[n, m]$, are dependent on the time difference $(m - n)$. Thus, for a stationary process, we can write

$$m_x = \mathcal{E}\{\mathbf{x}_n\}, \tag{A.22}$$

$$\sigma_x^2 = \mathcal{E}\{|(\mathbf{x}_n - m_x)|^2\}, \tag{A.23}$$

both independent of n, and now if we denote the time difference by m, we have

$$\phi_{xx}[n + m, n] = \phi_{xx}[m] = \mathcal{E}\{\mathbf{x}_{n+m} \mathbf{x}_n^*\}. \tag{A.24}$$

That is, the autocorrelation sequence of a stationary random process is a one-dimensional sequence, a function of the time difference m.

In many instances, we encounter random processes that are not stationary in the *strict sense*—i.e., their probability distributions are not time invariant—but Eqs. (A.22)–(A.24) still hold. Such random processes are said to be *wide-sense stationary*.

A.2.2 Time Averages

In a signal-processing context, the notion of an ensemble of signals is a convenient mathematical concept that allows us to use the theory of probability to represent the signals. However, in a practical situation, we always have available at most a finite number of finite-length sequences rather than an infinite ensemble of sequences. For example, we might wish to infer the probability law or certain averages of the random-process representation from measurements on a single member of the ensemble. When

the probability distributions are independent of time, it seems intuitively that the amplitude distribution (histogram) of a long segment of an individual sequence of samples should be approximately equal to the single probability density that describes each of the random variables of the random-process model. Similarly, the arithmetic average of a large number of samples of a single sequence should be very close to the mean of the process. To formalize these intuitive notions, we define the time average of a random process as

$$\langle \mathbf{x}_n \rangle = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} \mathbf{x}_n. \tag{A.25}$$

Similarly, the time autocorrelation sequence is defined as

$$\langle \mathbf{x}_{n+m}\mathbf{x}_n^* \rangle = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} \mathbf{x}_{n+m}\mathbf{x}_n^*. \tag{A.26}$$

It can be shown that the preceding limits exist if $\{\mathbf{x}_n\}$ is a stationary process with finite mean. As defined in Eqs. (A.25) and (A.26), these time averages are functions of an infinite set of random variables and thus are properly viewed as random variables themselves. However, under the condition known as *ergodicity*, the time averages in Eqs. (A.25) and (A.26) are equal to constants in the sense that the time averages of almost all possible sample sequences are equal to the same constant. Furthermore, they are equal to the corresponding ensemble average.[2] That is, for any single sample sequence $\{x[n]\}$ for $-\infty < n < \infty$,

$$\langle x[n] \rangle = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x[n] = \mathcal{E}\{\mathbf{x}_n\} = m_x \tag{A.27}$$

and

$$\langle x[n+m]x^*[n] \rangle = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} x[n+m]x^*[n] = \mathcal{E}\{\mathbf{x}_{n+m}\mathbf{x}_n^*\} = \phi_{xx}[m]. \tag{A.28}$$

The time-average operator $\langle \cdot \rangle$ has the same properties as the ensemble-average operator $\mathcal{E}\{\cdot\}$. Thus, we generally do not distinguish between the random variable $\mathbf{x}_n$ and its value in a sample sequence, $x[n]$. For example, the expression $\mathcal{E}\{x[n]\}$ should be interpreted as $\mathcal{E}\{\mathbf{x}_n\} = \langle x[n] \rangle$. In general, a random process for which time averages equal ensemble averages is called an *ergodic process*.

In practice, it is common to assume that a given sequence is a sample sequence of an ergodic random process so that averages can be computed from a single sequence. Of course, we generally cannot compute with the limits in Eqs. (A.27) and (A.28), but instead the quantities

$$\hat{m}_x = \frac{1}{L} \sum_{n=0}^{L-1} x[n], \tag{A.29}$$

$$\hat{\sigma}_x^2 = \frac{1}{L} \sum_{n=0}^{L-1} |x[n] - \hat{m}_x|^2, \tag{A.30}$$

[2] A more precise statement is that the random variables $\langle \mathbf{x}_n \rangle$ and $\langle \mathbf{x}_{n+m}\mathbf{x}_n^* \rangle$ have means equal to m_x and $\phi_{xx}[m]$, respectively, and their variances are zero.

and

$$\langle x[n+m]x^*[n]\rangle_L = \frac{1}{L}\sum_{n=0}^{L-1} x[n+m]x^*[n] \tag{A.31}$$

or similar quantities are often computed as *estimates* of the mean, variance, and autocorrelation. $\hat{m}_x$ and $\hat{\sigma}_x^2$ are referred to as the sample mean and sample variance, respectively. The estimation of averages of a random process from a finite segment of data is a problem of statistics, which we touched on briefly in Chapter 10.

A.3 PROPERTIES OF CORRELATION AND COVARIANCE SEQUENCES

Several useful properties of correlation and covariance functions follow in a simple way from the definitions. These properties are given in this section.

Consider two real stationary random processes $\{\mathbf{x}_n\}$ and $\{\mathbf{y}_n\}$ with autocorrelation, autocovariance, cross-correlation, and cross-covariance being given, respectively, by

$$\phi_{xx}[m] = \mathcal{E}\{\mathbf{x}_{n+m}\mathbf{x}_n^*\}, \tag{A.32}$$

$$\gamma_{xx}[m] = \mathcal{E}\{(\mathbf{x}_{n+m} - m_x)(\mathbf{x}_n - m_x)^*\}, \tag{A.33}$$

$$\phi_{xy}[m] = \mathcal{E}\{\mathbf{x}_{n+m}\mathbf{y}_n^*\}, \tag{A.34}$$

$$\gamma_{xy}[m] = \mathcal{E}\{(\mathbf{x}_{n+m} - m_x)(\mathbf{y}_n - m_y)^*\}, \tag{A.35}$$

where m_x and m_y are the means of the two processes. The following properties are easily derived by simple manipulations of the definitions:

Property 1

$$\gamma_{xx}[m] = \phi_{xx}[m] - |m_x|^2, \tag{A.36a}$$

$$\gamma_{xy}[m] = \phi_{xy}[m] - m_x m_y^*. \tag{A.36b}$$

These results follow directly from Eqs. (A.19) and (A.21), and they indicate that the correlation and covariance sequences are identical for zero-mean processes.

Property 2

$$\phi_{xx}[0] = E[|\mathbf{x}_n|^2] = \text{Mean-square value}, \tag{A.37a}$$

$$\gamma_{xx}[0] = \sigma_x^2 = \text{Variance}. \tag{A.37b}$$

Property 3

$$\phi_{xx}[-m] = \phi_{xx}^*[m], \tag{A.38a}$$

$$\gamma_{xx}[-m] = \gamma_{xx}^*[m], \tag{A.38b}$$

$$\phi_{xy}[-m] = \phi_{yx}^*[m], \tag{A.38c}$$

$$\gamma_{xy}[-m] = \gamma_{yx}^*[m]. \tag{A.38d}$$

Property 4

$$|\phi_{xy}[m]|^2 \leq \phi_{xx}[0]\phi_{yy}[0], \tag{A.39a}$$

$$|\gamma_{xy}[m]|^2 \leq \gamma_{xx}[0]\gamma_{yy}[0]. \tag{A.39b}$$

In particular,

$$|\phi_{xx}[m]| \leq \phi_{xx}[0], \tag{A.40a}$$

$$|\gamma_{xx}[m]| \leq \gamma_{xx}[0]. \tag{A.40b}$$

Property 5. If $\mathbf{y}_n = \mathbf{x}_{n-n_0}$, then

$$\phi_{yy}[m] = \phi_{xx}[m], \tag{A.41a}$$

$$\gamma_{yy}[m] = \gamma_{xx}[m]. \tag{A.41b}$$

Property 6. For many random processes, the random variables become uncorrelated as they become more separated in time. If this is true,

$$\lim_{m\to\infty} \gamma_{xx}[m] = 0, \tag{A.42a}$$

$$\lim_{m\to\infty} \phi_{xx}[m] = |m_x|^2, \tag{A.42b}$$

$$\lim_{m\to\infty} \gamma_{xy}[m] = 0, \tag{A.42c}$$

$$\lim_{m\to\infty} \phi_{xy}[m] = m_x m_y^*. \tag{A.42d}$$

The essence of these results is that the correlation and covariance are finite-energy sequences that tend to die out for large values of m. Thus, it is often possible to represent these sequences in terms of their Fourier transforms or z-transforms.

A.4 FOURIER TRANSFORM REPRESENTATION OF RANDOM SIGNALS

Although the Fourier transform of a random signal does not exist, the autocovariance and autocorrelation sequences of such a signal are aperiodic sequences for which the transform does exist. The spectral representation of the correlation functions plays an important role in describing the input–output relations for a linear time-invariant system when the input is a stochastic signal. Therefore, it is of interest to consider the properties of correlation and covariance sequences and their corresponding Fourier and z-transforms.

Let us define $\Phi_{xx}(e^{j\omega})$, $\Gamma_{xx}(e^{j\omega})$, $\Phi_{xy}(e^{j\omega})$, and $\Gamma_{xy}(e^{j\omega})$ as the Fourier transforms of $\phi_{xx}[m]$, $\gamma_{xx}[m]$, $\phi_{xy}[m]$, and $\gamma_{xy}[m]$, respectively. Since these functions are all discrete-time Fourier transforms of sequences, they must be periodic with period 2π. From Eqs. (A.36a) and (A.36b), it follows that, over one period $|\omega| \leq \pi$,

$$\Phi_{xx}(e^{j\omega}) = \Gamma_{xx}(e^{j\omega}) + 2\pi|m_x|^2\delta(\omega), \qquad |\omega| < \pi, \tag{A.43a}$$

and

$$\Phi_{xy}(e^{j\omega}) = \Gamma_{xy}(e^{j\omega}) + 2\pi m_x m_y^* \delta(\omega), \qquad |\omega| < \pi. \tag{A.43b}$$

In the case of zero-mean processes ($m_x = 0$ and $m_y = 0$), the correlation and covariance functions are identical so that $\Phi_{xx}(e^{j\omega}) = \Gamma_{xx}(e^{j\omega})$ and $\Phi_{xy}(e^{j\omega}) = \Gamma_{xy}(e^{j\omega})$.

From the inverse Fourier transform equation, it follows that

$$\gamma_{xx}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma_{xx}(e^{j\omega}) e^{j\omega m} d\omega, \tag{A.44a}$$

$$\phi_{xx}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{xx}(e^{j\omega}) e^{j\omega m} d\omega, \tag{A.44b}$$

and, consequently,

$$\mathcal{E}\{|x[n]|^2\} = \phi_{xx}[0] = \sigma_x^2 + |m_x|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{xx}(e^{j\omega}) d\omega, \tag{A.45a}$$

$$\sigma_x^2 = \gamma_{xx}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma_{xx}(e^{j\omega}) d\omega. \tag{A.45b}$$

Sometimes it is notationally convenient to define the quantity

$$P_{xx}(\omega) = \Phi_{xx}(e^{j\omega}), \tag{A.46}$$

in which case Eqs.(A.45a) and (A.45b) are expressed as

$$\mathcal{E}\{|x[n]|^2\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\omega) d\omega, \tag{A.47a}$$

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\omega) d\omega - |m_x|^2. \tag{A.47b}$$

Thus, the area under $P_{xx}(\omega)$ for $-\pi \leq \omega \leq \pi$ is proportional to the average power in the signal. In fact, as we discussed in Section 2.10, the integral of $P_{xx}(\omega)$ over a band of frequencies is proportional to the power in the signal in that band. For this reason, the function $P_{xx}(\omega)$ is called the *power density spectrum,* or simply, the *power spectrum.* When $P_{xx}(\omega)$ is a constant independent of ω, the random process is referred to as a white-noise process, or simply, white noise. When $P_{xx}(\omega)$ is constant over a band and zero otherwise, we refer to it as bandlimited white noise.

From Eq. (A.38b), it follows that $P_{xx}(\omega)$ is always real valued, and for real random processes, $\phi_{xx}[m] = \phi_{xx}[-m]$, so in the real case, $P_{xx}(\omega)$ is both real and even; i.e.,

$$P_{xx}(\omega) = P_{xx}(-\omega). \tag{A.48}$$

An additional important property is that the power density spectrum is nonnegative. This point is discussed in Section 2.10.

The *cross power density spectrum* is defined as

$$P_{xy}(\omega) = \Phi_{xy}(e^{j\omega}). \tag{A.49}$$

This function is generally complex, and from Eq. (A.38c), it follows that

$$P_{xy}(\omega) = P_{yx}^*(\omega). \tag{A.50}$$

Finally, as shown in Section 2.10, if $x[n]$ is a random signal input to a linear time-invariant discrete-time system with frequency response $H(e^{j\omega})$, and if $y[n]$ is the corresponding output, then

$$\Phi_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2 \Phi_{xx}(e^{j\omega}) \tag{A.51}$$

and

$$\Phi_{xy}(e^{j\omega}) = H(e^{j\omega})\Phi_{xy}(e^{j\omega}). \tag{A.52}$$

Example A.1 Noise Power Output of Ideal Lowpass Filter

Suppose that $x[n]$ is a zero-mean white-noise sequence with $\phi_{xx}[m] = \sigma_x^2\delta[m]$ and power spectrum $\Phi_{xx}(e^{j\omega}) = \sigma_x^2$ for $|\omega| \leq \pi$, and furthermore, assume that $x[n]$ is the input to an ideal lowpass filter with cutoff frequency ω_c. Then from Eq. (A.51), it follows that the output $y[n]$ would be a bandlimited white noise process whose power spectrum would be

$$\Phi_{yy}(e^{j\omega}) = \begin{cases} \sigma_x^2, & |\omega| < \omega_c, \\ 0, & \omega_c < |\omega| \leq \pi. \end{cases} \tag{A.53}$$

Using the inverse Fourier transform, we obtain the autocorrelation sequence

$$\phi_{yy}[m] = \frac{\sin(\omega_c n)}{\pi n}. \tag{A.54}$$

Now, using Eq. (A.45a), we get for the average power of the output,

$$\mathcal{E}\{y^2[n]\} = \phi_{yy}[0] = \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c} \sigma_x^2 d\omega = \sigma_x^2 \frac{\omega_c}{\pi}. \tag{A.55}$$

A.5 USE OF THE z-TRANSFORM IN AVERAGE POWER COMPUTATIONS

To carry out average power calculations using Eq. (A.45a), we must evaluate an integral of the power spectrum as was done in Example A.1. While the integral in that Example was easy to evaluate, such integrals in general are difficult to evaluate as real integrals. However, a result based on the z-transform makes the calculation of average output power straightforward in the important case of systems that have rational system functions.

In general, the z-transform can be used to represent the covariance function but not a correlation function. This is because when a signal has nonzero average value, its correlation function will contain an additive constant component that does not have a z-transform representation. When the average value is zero, however, the covariance

and correlation functions are, of course, equal. If the z-transform of $\gamma_{xx}[m]$ exists, then since $\gamma_{xx}[-m] = \gamma_{xx}^*[m]$ it follows that in general

$$\Gamma_{xx}(z) = \Gamma_{xx}^*(1/z^*). \tag{A.56}$$

Furthermore, since $\gamma_{xx}[m]$ is two sided and conjugate-symmetric, it follows that the region of convergence of $\Gamma_{xx}(z)$ must be of the form

$$r_a < |z| < \frac{1}{r_a}$$

where necessarily $0 < r_a < 1$. In the important case when $\Gamma_{xx}(z)$ is a rational function of z, Eq. (A.56) implies that the poles and zeros of $\Gamma_{xx}(z)$ must occur in complex-conjugate reciprocal pairs.

The major advantage of the z-transform representation is that when $\Gamma_{xx}(z)$ is a rational function, the average power of the random signal can be computed easily using the relation

$$\mathcal{E}\{x^2[n]\} = \sigma_x^2 = \gamma_{xx}[0] = \left\{ \begin{array}{c} \text{Inverse } z\text{-transform} \\ \text{of } \Gamma_{xx}(z), \\ \text{evaluated for } m = 0 \end{array} \right\}. \tag{A.57}$$

It is straightforward to evaluate the right-hand side of this equation using a method based on the observation that when $\Gamma_{xx}(z)$ is a rational function of z, $\gamma_{xx}[m]$ can be computed for all m by employing a partial fraction expansion. Then to obtain the average power, we can simply evaluate $\gamma_{xx}[m]$ for $m = 0$.

The z-transform is also useful in determining the autocovariance and average power of the output of an LTI system when the input is a random signal. Generalizing Eq. (A.51) leads to

$$\Gamma_{yy}(z) = H(z)H^*(1/z^*)\Gamma_{xx}(z), \tag{A.58}$$

and from the properties of the z-transform and Eq. (A.58), it follows that the autocovariance of the output is the convolution

$$\gamma_{yy}[m] = h[m] * h^*[-m] * \gamma_{xx}[m]. \tag{A.59}$$

This result is particularly useful in quantization noise analysis where we need to compute the average output power when the input to a linear difference equation is a zero-mean white noise signal with average power σ_x^2. Since the autocovariance of such an input is $\gamma_{xx}[m] = \sigma_x^2\delta[m]$, it follows that the autocovariance of the output is $\gamma_{yy}[m] = \sigma_x^2(h[m] * h^*[-m])$, i.e., the covariance of the output is proportional to the deterministic autocorrelation of the impulse response of the LTI system. From this result it follows that

$$\mathcal{E}\{y^2[n]\} = \gamma_{yy}[0] = \sigma_x^2 \sum_{n=-\infty}^{\infty} |h[n]|^2. \tag{A.60}$$

As an alternative to computing the sum of squares of the impulse response sequence, which can be rather difficult for IIR systems, we can apply the method suggested in Eq. (A.57) to obtain $\mathcal{E}\{y^2[n]\}$ from a partial fraction expansion of $\Gamma_{yy}(z)$. Recall that for a white noise input with $\gamma_{xx}[m] = \sigma_x^2\delta[m]$, the z-transform is $\Gamma_{xx}(z) = \sigma_x^2$ so $\Gamma_{yy}(z) = \sigma_x^2 H(z)H^*(1/z^*)$. Therefore, Eq. (A.57) applied to the output of the system gives

$$\mathcal{E}\{y^2[n]\} = \gamma_{yy}[0] = \left\{ \begin{array}{c} \text{Inverse } z\text{-transform of} \\ \Gamma_{yy}(z) = H(z)H^*(1/z^*)\sigma_x^2, \\ \text{evaluated for } m = 0 \end{array} \right\}. \tag{A.61}$$

Now consider the special case of a stable and causal system having a rational system function of the form

$$H(z) = A\frac{\prod_{m=1}^{M}(1 - c_m z^{-1})}{\prod_{k=1}^{N}(1 - d_k z^{-1})} \qquad |z| > \max_k\{|d_k|\}, \tag{A.62}$$

where $\max_k\{|d_k|\} < 1$ and $M < N$. Such a system function might describe the relationship between an internal round-off noise source and the output of a system implemented with fixed-point arithmetic. Substituting Eq. (A.62) for $H(z)$ in Eq. (A.58) gives

$$\Gamma_{yy}(z) = \sigma_x^2 H(z)H^*(1/z^*) = \sigma_x^2|A|^2\frac{\prod_{m=1}^{M}(1 - c_m z^{-1})(1 - c_m^* z)}{\prod_{k=1}^{N}(1 - d_k z^{-1})(1 - d_k^* z)}. \tag{A.63}$$

Since we have assumed that $|d_k| < 1$ for all k, all of the original poles are inside the unit circle and therefore the other poles at $(d_k^*)^{-1}$ are at conjugate reciprocal locations outside the unit circle. The region of convergence for $\Gamma_{yy}(z)$ is therefore $\max_k |d_k| < |z| < \min_k |(d_k^*)^{-1}|$. For such rational functions, it can be shown that since $M < N$, the partial fraction expansion has the form

$$\Gamma_{yy}(z) = \sigma_x^2\left(\sum_{k=1}^{N}\left(\frac{A_k}{1 - d_k z^{-1}} - \frac{A_k^*}{1 - (d_k^*)^{-1}z^{-1}}\right)\right), \tag{A.64}$$

where the coefficients are found from

$$A_k = H(z)H^*(1/z^*)(1 - d_k z^{-1})\big|_{z=d_k}. \tag{A.65}$$

Since the poles at $z = d_k$ are inside the inner boundary of the region of convergence, each of them corresponds to a right-sided sequence, while the poles at $z = (d_k^*)^{-1}$ each correspond to a left-sided sequence. Thus, the autocovariance function corresponding to Eq. (A.64) is

$$\gamma_{yy}[n] = \sigma_x^2\sum_{k=1}^{N}(A_k(d_k)^n u[n] + A_k^*(d_k^*)^{-n}u[-n-1]),$$

from which it follows that we can obtain the average power from

$$\sigma_y^2 = \gamma_{yy}[0] = \sigma_x^2\left(\sum_{k=1}^{N} A_k\right), \tag{A.66}$$

where the quantities A_k are given by Eq: (A.65).

Thus, the computation of the total average power of the output of a system with rational system function and white noise input reduces to the straightforward problem of finding partial fraction expansion coefficients for the z-transform of the output autocorrelation function. The utility of this approach is illustrated by the following example.

Example A.2 Noise Power Output of a Second-Order IIR Filter

Consider a system with impulse response

$$h[n] = \frac{r^n \sin\theta(n+1)}{\sin\theta} u[n] \tag{A.67}$$

and system function

$$H(z) = \frac{1}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})}. \tag{A.68}$$

When the input is white noise with total average power σ_x^2, the z-transform of the autocovariance function of the output is

$$\Gamma_{yy}(z) = \sigma_x^2 \left(\frac{1}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})}\right)\left(\frac{1}{(1 - re^{-j\theta}z)(1 - re^{j\theta}z)}\right) \tag{A.69}$$

from which we obtain, using Eq. (A.65),

$$\begin{aligned}\mathcal{E}\{y^2[n]\} = \sigma_x^2 \Bigg[&\left(\frac{1}{(1 - re^{-j\theta}z^{-1})}\right)\left(\frac{1}{(1 - re^{-j\theta}z)(1 - re^{j\theta}z)}\right)\Bigg|_{z=re^{j\theta}} \\ &+ \left(\frac{1}{(1 - re^{j\theta}z^{-1})}\right)\left(\frac{1}{(1 - re^{-j\theta}z)(1 - re^{j\theta}z)}\right)\Bigg|_{z=re^{-j\theta}}\Bigg].\end{aligned} \tag{A.70}$$

Making the indicated substitutions, placing both terms over a common denominator, and doing some algebra leads to

$$\mathcal{E}\{y^2[n]\} = \sigma_x^2 \left(\frac{1 + r^2}{1 - r^2}\right)\left(\frac{1}{1 - 2r^2\cos(2\theta) + r^4}\right). \tag{A.71}$$

Thus, using the partial fraction expansion of $\Gamma_{yy}(z)$ we have effectively evaluated the expression

$$\mathcal{E}\{y^2[n]\} = \sigma_x^2 \sum_{n=-\infty}^{\infty} |h[n]|^2 = \sigma_x^2 \sum_{n=0}^{\infty} \left|\frac{r^n \sin\theta(n+1)}{\sin\theta}\right|^2,$$

which would be difficult to sum in closed form, and the expression

$$\mathcal{E}\{y^2[n]\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sigma_x^2 |H(e^{j\omega})|^2 d\omega = \frac{\sigma_x^2}{2\pi}\int_{-\pi}^{\pi} \frac{d\omega}{|(1 - re^{j\theta}e^{-j\omega})(1 - re^{-j\theta}e^{-j\omega})|^2},$$

which would be difficult to evaluate as a real integral.

The remarkably simple result of Example A.2 is an illustration of the power of the partial fraction method in evaluating average power formulas. In Chapter 6, we make use of this technique in the analysis of quantization effects in the implementation of digital filters.