CHAPTER 10

Analysis and Processing of Random Signals

In this chapter we introduce methods for analyzing and processing random signals. We cover the following topics:

- Section 10.1 introduces the notion of power spectral density, which allows us to view random processes in the frequency domain.
- Section 10.2 discusses the response of linear systems to random process inputs and introduce methods for filtering random processes.
- Section 10.3 considers two important applications of signal processing: sampling and modulation.
- Sections 10.4 and 10.5 discuss the design of optimum linear systems and introduce the Wiener and Kalman filters.
- Section 10.6 addresses the problem of estimating the power spectral density of a random process.
- Finally, Section 10.7 introduces methods for implementing and simulating the processing of random signals.

10.1 POWER SPECTRAL DENSITY

The Fourier series and the Fourier transform allow us to view deterministic time functions as the weighted sum or integral of sinusoidal functions. A time function that varies slowly has the weighting concentrated at the low-frequency sinusoidal components. A time function that varies rapidly has the weighting concentrated at higher-frequency components. Thus the rate at which a deterministic time function varies is related to the weighting function of the Fourier series or transform. This weighting function is called the "spectrum" of the time function.

The notion of a time function as being composed of sinusoidal components is also very useful for random processes. However, since a sample function of a random process can be viewed as being selected from an ensemble of allowable time functions, the weighting function or "spectrum" for a random process must refer in some way to the average rate of change of the ensemble of allowable time functions. Equation (9.66) shows that, for wide-sense stationary processes, the autocorrelation function

$R_X(\tau)$ is an appropriate measure for the average rate of change of a random process. Indeed if a random process changes slowly with time, then it remains correlated with itself for a long period of time, and $R_X(\tau)$ decreases slowly as a function of τ. On the other hand, a rapidly varying random process quickly becomes uncorrelated with itself, and $R_X(\tau)$ decreases rapidly with τ.

We now present the **Einstein-Wiener-Khinchin theorem**, which states that the power spectral density of a wide-sense stationary random process is given by the Fourier transform of the autocorrelation function.[1]

10.1.1 Continuous-Time Random Processes

Let $X(t)$ be a continuous-time WSS random process with mean m_X and autocorrelation function $R_X(\tau)$. Suppose we take the Fourier transform of a sample of $X(t)$ in the interval $0 < t < T$ as follows

$$\widetilde{x}(f) = \int_0^T X(t')e^{-j2\pi ft'}\,dt'. \tag{10.1}$$

We then approximate the power density as a function of frequency by the function:

$$\widetilde{p}_T(f) = \frac{1}{T}|\widetilde{x}(f)|^2 = \frac{1}{T}\widetilde{x}(f)\widetilde{x}^*(f) = \frac{1}{T}\left\{\int_0^T X(t')e^{-j2\pi ft'}\,dt'\right\}\left\{\int_0^T X(t')e^{j2\pi ft'}\,dt'\right\}, \tag{10.2}$$

where * denotes the complex conjugate. X(t) is a random process, so $\widetilde{p}_T(f)$ is also a random process but over a different index set. $\widetilde{p}_T(f)$ is called the **periodogram estimate** and we are interested in the **power spectral density** of $X(t)$ which is defined by:

$$S_X(f) = \lim_{T\to\infty} E[\widetilde{p}_T(f)] = \lim_{T\to\infty}\frac{1}{T}E[|\widetilde{x}(f)|^2]. \tag{10.3}$$

We show at the end of this section that *the power spectral density of X(t) is given by the Fourier transform of the autocorrelation function:*

$$S_X(f) = \mathscr{F}\{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau)e^{-j2\pi f\tau}\,d\tau. \tag{10.4}$$

A table of Fourier transforms and its properties is given in Appendix B.

For real-valued random processes, the autocorrelation function is an even function of τ:

$$R_X(\tau) = R_X(-\tau). \tag{10.5}$$

[1]This result is usually called the Wiener-Khinchin theorem, after Norbert Wiener and A. Ya. Khinchin, who proved the result in the early 1930s. Later it was discovered that this result was stated by Albert Einstein in a 1914 paper (see Einstein).

AYER

WTF

Substitution into Eq. (10.4) implies that

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau)\{\cos 2\pi f\tau - j\sin 2\pi f\tau\}\, d\tau$$
$$= \int_{-\infty}^{\infty} R_X(\tau)\cos 2\pi f\tau\, d\tau, \tag{10.6}$$

since the integral of the product of an even function $(R_X(\tau))$ and an odd function $(\sin 2\pi f\tau)$ is zero. Equation (10.6) implies that $S_X(f)$ *is real-valued and an even function of f*. From Eq. (10.2) we have that $S_X(f)$ *is nonnegative*:

$$S_X(f) \geq 0 \qquad \text{for all } f. \tag{10.7}$$

The autocorrelation function can be recovered from the power spectral density by applying the inverse Fourier transform formula to Eq. (10.4):

$$R_X(\tau) = \mathscr{F}^{-1}\{S_X(f)\}$$
$$= \int_{-\infty}^{\infty} S_X(f)e^{j2\pi f\tau}\, df. \tag{10.8}$$

Equation (10.8) is identical to Eq. (4.80), which relates the pdf to its corresponding characteristic function. The last section in this chapter discusses how the FFT can be used to perform numerical calculations for $S_X(f)$ and $R_X(\tau)$.

In electrical engineering it is customary to refer to the second moment of $X(t)$ as the **average power of $X(t)$**.[2] Equation (10.8) together with Eq. (9.64) gives

$$E[X^2(t)] = R_X(0) = \int_{-\infty}^{\infty} S_X(f)\, df. \tag{10.9}$$

Equation (10.9) states that the average power of $X(t)$ is obtained by integrating $S_X(f)$ over all frequencies. This is consistent with the fact that $S_X(f)$ is the "density of power" of $X(t)$ at the frequency f.

Since the autocorrelation and autocovariance functions are related by $R_X(\tau) = C_X(\tau) + m_X^2$, the power spectral density is also given by

$$S_X(f) = \mathscr{F}\{C_X(\tau) + m_X^2\}$$
$$= \mathscr{F}\{C_X(\tau)\} + m_X^2\,\delta(f), \tag{10.10}$$

where we have used the fact that the Fourier transform of a constant is a delta function. We say the m_X is the "dc" component of $X(t)$.

The notion of power spectral density can be generalized to two jointly wide-sense stationary processes. The **cross-power spectral density $S_{X,Y}(f)$** is defined by

$$S_{X,Y}(f) = \mathscr{F}\{R_{X,Y}(\tau)\}, \tag{10.11}$$

[2]If $X(t)$ is a voltage or current developed across a 1-ohm resistor, then $X^2(t)$ is the instantaneous power absorbed by the resistor.

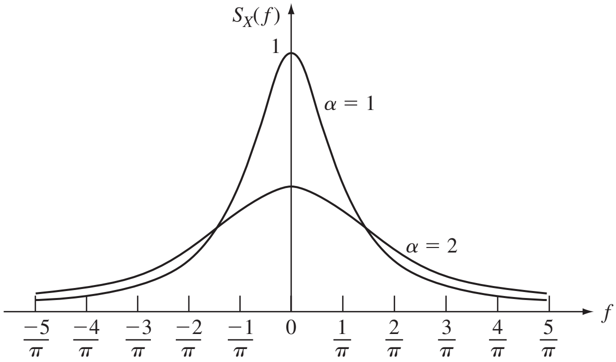

FIGURE 10.1
Power spectral density of a random telegraph signal with $\alpha = 1$ and $\alpha = 2$ transitions per second.

where $R_{X,Y}(\tau)$ is the cross-correlation between $X(t)$ and $Y(t)$:

$$R_{X,Y}(\tau) = E[X(t + \tau)Y(t)]. \tag{10.12}$$

In general, $S_{X,Y}(f)$ is a complex function of f even if $X(t)$ and $Y(t)$ are both real-valued.

Example 10.1 Random Telegraph Signal

Find the power spectral density of the random telegraph signal.

In Example 9.24, the autocorrelation function of the random telegraph process was found to be

$$R_X(\tau) = e^{-2\alpha|\tau|},$$

where α is the average transition rate of the signal. Therefore, the power spectral density of the process is

$$\begin{aligned} S_X(f) &= \int_{-\infty}^{0} e^{2\alpha\tau}e^{-j2\pi f\tau}\,d\tau + \int_{0}^{\infty} e^{-2\alpha\tau}e^{-j2\pi f\tau}\,d\tau \\ &= \frac{1}{2\alpha - j2\pi f} + \frac{1}{2\alpha + j2\pi f} \\ &= \frac{4\alpha}{4\alpha^2 + 4\pi^2 f^2}. \end{aligned} \tag{10.13}$$

Figure 10.1 shows the power spectral density for $\alpha = 1$ and $\alpha = 2$ transitions per second. The process changes two times more quickly when $\alpha = 2$; it can be seen from the figure that the power spectral density for $\alpha = 2$ has greater high-frequency content.

Example 10.2 Sinusoid with Random Phase

Let $X(t) = a\cos(2\pi f_0 t + \Theta)$, where Θ is uniformly distributed in the interval $(0, 2\pi)$. Find $S_X(f)$.

From Example 9.10, the autocorrelation for $X(t)$ is

$$R_X(\tau) = \frac{a^2}{2}\cos 2\pi f_0 \tau.$$

Thus, the power spectral density is

$$\begin{aligned} S_X(f) &= \frac{a^2}{2}\mathscr{F}\{\cos 2\pi f_0 \tau\} \\ &= \frac{a^2}{4}\delta(f - f_0) + \frac{a^2}{4}\delta(f + f_0), \end{aligned} \tag{10.14}$$

where we have used the table of Fourier transforms in Appendix B. The signal has average power $R_X(0) = a^2/2$. All of this power is concentrated at the frequencies $\pm f_0$, so the power density at these frequencies is infinite.

Example 10.3 White Noise

The power spectral density of a WSS white noise process whose frequency components are limited to the range $-W \le f \le W$ is shown in Fig. 10.2(a). The process is said to be "white" in analogy to white light, which contains all frequencies in equal amounts. The average power in this

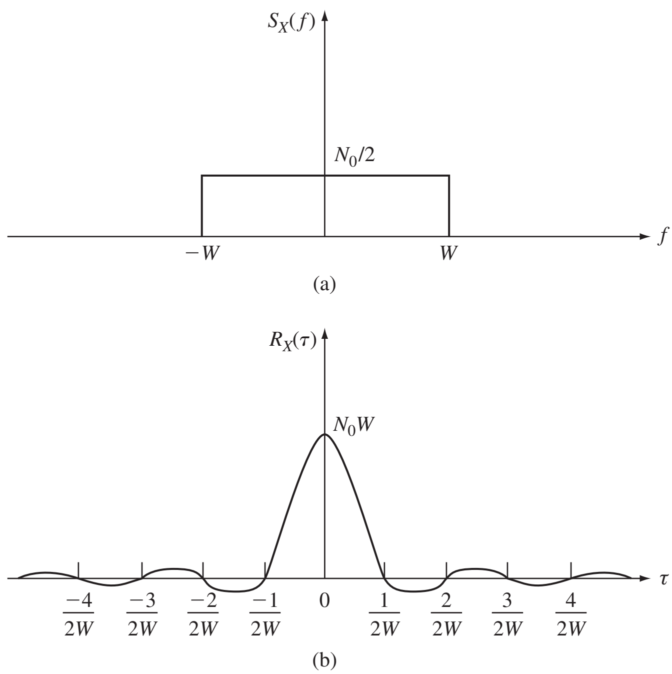

FIGURE 10.2
Bandlimited white noise: (a) power spectral density, (b) autocorrelation function.

process is obtained from Eq. (10.9):

$$E[X^2(t)] = \int_{-W}^{W} \frac{N_0}{2} df = N_0 W. \tag{10.15}$$

The autocorrelation for this process is obtained from Eq. (10.8):

$$\begin{aligned} R_X(\tau) &= \frac{1}{2} N_0 \int_{-W}^{W} e^{j2\pi f \tau} \, df \\ &= \frac{1}{2} N_0 \frac{e^{-j2\pi W\tau} - e^{j2\pi W\tau}}{-j2\pi\tau} \\ &= \frac{N_0 \sin(2\pi W_\tau)}{2\pi\tau}. \end{aligned} \tag{10.16}$$

$R_X(\tau)$ is shown in Fig. 10.2(b). Note that $X(t)$ and $X(t + \tau)$ are uncorrelated at $\tau = \pm k/2W$, $k = 1, 2, \ldots$.

The term white noise usually refers to a random process $W(t)$ whose power spectral density is $N_0/2$ for *all* frequencies:

$$S_W(f) = \frac{N_0}{2} \qquad \text{for all } f. \tag{10.17}$$

Equation (10.15) with $W = \infty$ shows that such a process must have infinite average power. By taking the limit $W \to \infty$ in Eq. (10.16), we find that the autocorrelation of such a process approaches

$$R_W(\tau) = \frac{N_0}{2} \delta(\tau). \tag{10.18}$$

If $W(t)$ is a Gaussian random process, we then see that $W(t)$ is the white Gaussian noise process introduced in Example 9.43 with $\alpha = N_0/2$.

Example 10.4 Sum of Two Processes

Find the power spectral density of $Z(t) = X(t) + Y(t)$, where $X(t)$ and $Y(t)$ are jointly WSS processes.

The autocorrelation of $Z(t)$ is

$$\begin{aligned} R_Z(\tau) &= E[Z(t + \tau)Z(t)] = E[(X(t + \tau) + Y(t + \tau))(X(t) + Y(t))] \\ &= R_X(\tau) + R_{YX}(\tau) + R_{XY}(\tau) + R_Y(\tau). \end{aligned}$$

The power spectral density is then

$$\begin{aligned} S_Z(f) &= \mathcal{F}\{R_X(\tau) + R_{YX}(\tau) + R_{XY}(\tau) + R_Y(\tau)\} \\ &= S_X(f) + S_{YX}(f) + S_{XY}(f) + S_Y(f). \end{aligned} \tag{10.19}$$

Example 10.5

Let $Y(t) = X(t - d)$, where d is a constant delay and where $X(t)$ is WSS. Find $R_{YX}(\tau)$, $S_{YX}(f)$, $R_Y(\tau)$, and $S_Y(f)$.

The definitions of $R_{YX}(\tau)$, $S_{YX}(f)$, and $R_Y(\tau)$ give

$$R_{YX}(\tau) = E[Y(t+\tau)X(t)] = E[X(t+\tau-d)X(t)] = R_X(\tau-d). \tag{10.20}$$

The time-shifting property of the Fourier transform gives

$$\begin{aligned} S_{YX}(f) &= \mathcal{F}\{R_X(\tau-d)\} = S_X(f)e^{-j2\pi fd} \\ &= S_X(f)\cos(2\pi fd) - jS_X(f)\sin(2\pi fd). \end{aligned} \tag{10.21}$$

Finally,

$$R_Y(\tau) = E[Y(t+\tau)Y(t)] = E[X(t+\tau-d)X(t-d)] = R_X(\tau). \tag{10.22}$$

Equation (10.22) implies that

$$S_Y(f) = \mathcal{F}\{R_Y(\tau)\} = \mathcal{F}\{R_X(\tau)\} = S_X(f). \tag{10.23}$$

Note from Eq. (10.21) that the cross-power spectral density is complex. Note from Eq. (10.23) that $S_X(f) = S_Y(f)$ despite the fact that $X(t) \neq Y(t)$. Thus, $S_X(f) = S_Y(f)$ *does not imply that* $X(t) = Y(t)$.

10.1.2 Discrete-Time Random Processes

Let X_n be a discrete-time WSS random process with mean m_X and autocorrelation function $R_X(k)$. The **power spectral density of X_n** is defined as the Fourier transform of the autocorrelation sequence

$$\begin{aligned} S_X(f) &= \mathcal{F}\{R_X(k)\} \\ &= \sum_{k=-\infty}^{\infty} R_X(k)e^{-j2\pi fk}. \end{aligned} \tag{10.24}$$

$-\pi < \omega < \pi$

Note that we need only consider frequencies in the range $-1/2 < f \leq 1/2$, since $S_X(f)$ is periodic in f with period 1. As in the case of continuous random processes, $S_X(f)$ can be shown to be a real-valued, nonnegative, even function of f.

The inverse Fourier transform formula applied to Eq. (10.23) implies that[3]

$$R_X(k) = \int_{-1/2}^{1/2} S_X(f)e^{j2\pi fk}\,df. \tag{10.25}$$

Equations (10.24) and (10.25) are similar to the discrete Fourier transform. In the last section we show how to use the FFT to calculate $S_X(f)$ and $R_X(k)$.

The **cross-power spectral density $S_{X,Y}(f)$** of two jointly WSS discrete-time processes X_n and Y_n is defined by

$$S_{X,Y}(f) = \mathcal{F}\{R_{X,Y}(k)\}, \tag{10.26}$$

where $R_{X,Y}(k)$ is the cross-correlation between X_n and Y_n:

$$R_{X,Y}(k) = E[X_{n+k}Y_n]. \tag{10.27}$$

[3] You can view $R_X(k)$ as the coefficients of the Fourier series of the periodic function $S_X(f)$.

Uy amigo q' turbio esto

Example 10.6 White Noise

Let the process X_n be a sequence of uncorrelated random variables with zero mean and variance σ_X^2. Find $S_X(f)$.

The autocorrelation of this process is

$$R_X(k) = \begin{cases} \sigma_X^2 & k = 0 \\ 0 & k \neq 0. \end{cases}$$

The power spectral density of the process is found by substituting $R_X(k)$ into Eq. (10.24):

$$S_X(f) = \sigma_X^2 \qquad -\frac{1}{2} < f < \frac{1}{2}. \tag{10.28}$$

Thus the process X_n contains all possible frequencies in equal measure.

Example 10.7 Moving Average Process

Let the process Y_n be defined by

$$Y_n = X_n + \alpha X_{n-1}, \tag{10.29}$$

where X_n is the white noise process of Example 10.6. Find $S_Y(f)$.

It is easily shown that the mean and autocorrelation of Y_n are given by

$$E[Y_n] = 0,$$

and

$$E[Y_n Y_{n+k}] = \begin{cases} (1 + \alpha^2)\sigma_X^2 & k = 0 \\ \alpha\sigma_X^2 & k = \pm 1 \\ 0 & \text{otherwise.} \end{cases} \tag{10.30}$$

The power spectral density is then

$$\begin{aligned} S_Y(f) &= (1 + \alpha^2)\sigma_X^2 + \alpha\sigma_X^2\{e^{j2\pi f} + e^{-j2\pi f}\} \\ &= \sigma_X^2\{(1 + \alpha^2) + 2\alpha \cos 2\pi f\}. \end{aligned} \tag{10.31}$$

$S_Y(f)$ is shown in Fig. 10.3 for $\alpha = 1$.

Example 10.8 Signal Plus Noise

Let the observation Z_n be given by

$$Z_n = X_n + Y_n,$$

where X_n is the signal we wish to observe, Y_n is a white noise process with power σ_Y^2, and X_n and Y_n are independent random processes. Suppose further that $X_n = A$ for all n, where A is a random variable with zero mean and variance σ_A^2. Thus Z_n represents a sequence of noisy measurements of the random variable A. Find the power spectral density of Z_n.

The mean and autocorrelation of Z_n are

$$E[Z_n] = E[A] + E[Y_n] = 0$$

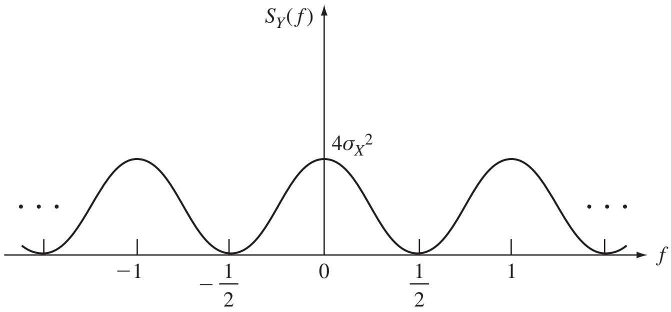

FIGURE 10.3
Power spectral density of moving average process discussed in Example 10.7.

and

$$\begin{aligned} E[Z_n Z_{n+k}] &= E[(X_n + Y_n)(X_{n+k} + Y_{n+k})] \\ &= E[X_n X_{n+k}] + E[X_n]E[Y_{n+k}] \\ &\quad + E[X_{n+k}]E[Y_n] + E[Y_n Y_{n+k}] \\ &= E[A^2] + R_Y(k). \end{aligned}$$

Thus Z_n is also a WSS process.

The power spectral density of Z_n is then

$$S_Z(f) = E[A^2]\delta(f) + S_Y(f),$$

where we have used the fact that the Fourier transform of a constant is a delta function.

10.1.3 Power Spectral Density as a Time Average

In the above discussion, we simply stated that the power spectral density is given as the Fourier transform of the autocorrelation without supplying a proof. We now show how the power spectral density arises naturally when we take Fourier transforms of realizations of random processes.

Let $X_0, \ldots, X_{k-1}$ be k observations from the discrete-time, WSS process X_n. Let $\tilde{x}_k(f)$ denote the discrete Fourier transform of this sequence:

Entiendo que esto es una señal, es decir, una realización de mi PA

$$\tilde{x}_k(f) = \sum_{m=0}^{k-1} X_m e^{-j2\pi fm}. \tag{10.32}$$

Note that $\tilde{x}_k(f)$ is a complex-valued random variable. The magnitude squared of $\tilde{x}_k(f)$ is a measure of the "energy" at the frequency f. If we divide this energy by the total "time" k, we obtain an estimate for the "power" at the frequency f:

$$\tilde{p}_k(f) = \frac{1}{k}|\tilde{x}_k(f)|^2. \tag{10.33}$$

$\tilde{p}_k(f)$ is called the **periodogram estimate** for the power spectral density.

Mi problema es este. ¿Cómo vas a tomar la transformada DE UN PROCESO????? ¿Cómo puede ser que sumes sobre las REALIZACIONES de un proceso y no sumes sobre el tiempo?????????? CONUSLTAR URGENTE

Es que sí está sumando sobre el tiempo. Es uso de la palabra observaciones es poco feliz

Consider the expected value of the periodogram estimate:

$$
\begin{aligned}
E[\widetilde{p}_k(f)] &= \frac{1}{k}E[\widetilde{x}_k(f)\widetilde{x}_k^*(f)] \\
&= \frac{1}{k}E\left[\sum_{m=0}^{k-1} X_m e^{-j2\pi fm} \sum_{i=0}^{k-1} X_i e^{j2\pi fi}\right] \\
&= \frac{1}{k}\sum_{m=0}^{k-1}\sum_{i=0}^{k-1} E[X_m X_i]e^{-j2\pi f(m-i)} \\
&= \frac{1}{k}\sum_{m=0}^{k-1}\sum_{i=0}^{k-1} R_X(m-i)e^{-j2\pi f(m-i)}. \qquad (10.34)
\end{aligned}
$$

Figure 10.4 shows the range of the double summation in Eq. (10.34). Note that all the terms along the diagonal $m' = m - i$ are equal, that m' ranges from $-(k-1)$ to $k-1$, and that .here are $k - |m'|$ terms along the diagonal $m' = m - i$. Thus Eq. (10.34) becomes

$$
\begin{aligned}
E[\widetilde{p}_k(f)] &= \frac{1}{k}\sum_{m'=-(k-1)}^{k-1} \{k - |m'|\}R_X(m')e^{-j2\pi fm'} \\
&= \sum_{m'=-(k-1)}^{k-1} \left\{1 - \frac{|m'|}{k}\right\}R_X(m')e^{-j2\pi fm'}. \qquad (10.35)
\end{aligned}
$$

Comparison of Eq. (10.35) with Eq. (10.24) shows that the mean of the periodogram estimate is not equal to $S_X(f)$ for two reasons. First, Eq. (10.34) does not have the term in brackets in Eq. (10.25). Second, the limits of the summation in Eq. (10.35) are not $\pm\infty$. We say that $\widetilde{p}_k(f)$ is a "biased" estimator for $S_X(f)$. However, as $k \to \infty$, we see

→ La putamadre

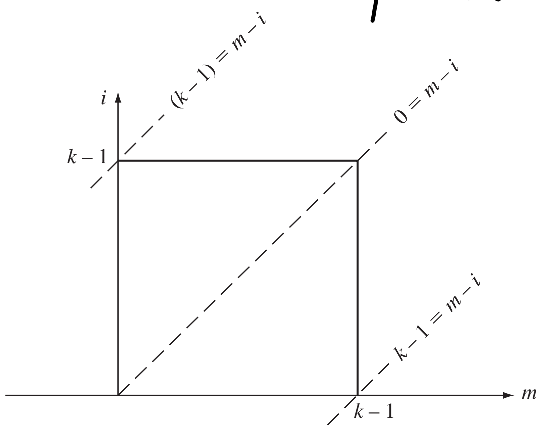

FIGURE 10.4
Range of summation in Eq. (10.34).

that the term in brackets approaches one, and that the limits of the summation approach $\pm\infty$. Thus

$$E[\widetilde{p}_k(f)] \to S_X(f) \qquad \text{as } k \to \infty, \tag{10.36}$$

No está completa la demostración

that is, the mean of the periodogram estimate does indeed approach $S_X(f)$. Note that Eq. (10.36) shows that $S_X(f)$ is nonnegative for all f, since $\widetilde{p}_k(f)$ is nonnegative for all f.

In order to be useful, the variance of the periodogram estimate should also approach zero. The answer to this question involves looking more closely at the problem of power spectral density estimation. We defer this topic to Section 10.6.

All of the above results hold for a continuous-time WSS random process $X(t)$ after appropriate changes are made from summations to integrals. The periodogram estimate for $S_X(f)$, for an observation in the interval $0 < t < T$, was defined in Eq. 10.2. The same derivation that led to Eq. (10.35) can be used to show that the mean of the periodogram estimate is given by

$$E[\widetilde{p}_T(f)] = \int_{-T}^{T} \left\{1 - \frac{|\tau|}{T}\right\} R_X(\tau) e^{-j2\pi f\tau}\, d\tau. \tag{10.37a}$$

It then follows that

$$E[\widetilde{p}_T(f)] \to S_X(f) \qquad \text{as } T \to \infty. \tag{10.37b}$$

Durísimo hermano.

10.2 RESPONSE OF LINEAR SYSTEMS TO RANDOM SIGNALS → Hasta acá llego

Many applications involve the processing of random signals (i.e., random processes) in order to achieve certain ends. For example, in prediction, we are interested in predicting future values of a signal in terms of past values. In filtering and smoothing, we are interested in recovering signals that have been corrupted by noise. In modulation, we are interested in converting low-frequency information signals into high-frequency transmission signals that propagate more readily through various transmission media.

Signal processing involves converting a signal from one form into another. Thus a signal processing method is simply a transformation or mapping from one time function into another function. If the input to the transformation is a random process, then the output will also be a random process. In the next two sections, we are interested in determining the statistical properties of the output process when the input is a wide-sense stationary random process.

10.2.1 Continuous-Time Systems

Consider a **system** in which an input signal $x(t)$ is mapped into the output signal $y(t)$ by the transformation

$$y(t) = T[x(t)].$$

The system is **linear** if superposition holds, that is,

$$T[\alpha x_1(t) + \beta x_2(t)] = \alpha T[x_1(t)] + \beta T[x_2(t)],$$

where $x_1(t)$ and $x_2(t)$ are arbitrary input signals, and α and β are arbitrary constants.[4] Let $y(t)$ be the response to input $x(t)$, then the system is said to be **time-invariant** if the response to $x(t-\tau)$ is $y(t-\tau)$. The **impulse response** $h(t)$ of a linear, time-invariant system is defined by

$$h(t) = T[\delta(t)]$$

where $\delta(t)$ is a unit delta function input applied at $t = 0$. The response of the system to an arbitrary input $x(t)$ is then

$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} h(s)x(t-s)\,ds = \int_{-\infty}^{\infty} h(t-s)x(s)\,ds. \tag{10.38}$$

Therefore a linear, time-invariant system is completely specified by its impulse response. The impulse response $h(t)$ can also be specified by giving its Fourier transform, the **transfer function** of the system:

$$H(f) = \mathscr{F}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\,dt. \tag{10.39}$$

A system is said to be **causal** if the response at time t depends only on past values of the input, that is, if $h(t) = 0$ for $t < 0$.

If the input to a linear, time-invariant system is a random process $X(t)$ as shown in Fig. 10.5, then the output of the system is the random process given by

$$Y(t) = \int_{-\infty}^{\infty} h(s)X(t-s)\,ds = \int_{-\infty}^{\infty} h(t-s)X(s)\,ds. \tag{10.40}$$

We assume that the integrals exist in the mean square sense as discussed in Section 9.7. We now show that if $X(t)$ is a wide-sense stationary process, then $Y(t)$ is also wide-sense stationary.[5]

The mean of $Y(t)$ is given by

$$E[Y(t)] = E\left[\int_{-\infty}^{\infty} h(s)X(t-s)\,ds\right] = \int_{-\infty}^{\infty} h(s)E[X(t-s)]\,ds.$$

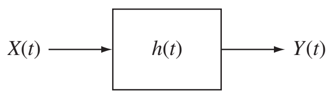

FIGURE 10.5
A linear system with a random input signal.

[4]For examples of nonlinear systems see Problems 9.11 and 9.56.

[5]Equation (10.40) supposes that the input was applied at an infinite time in the past. If the input is applied at $t = 0$, then $Y(t)$ is not wide-sense stationary. However, it becomes wide-sense stationary as the response reaches "steady state" (see Example 9.46 and Problem 10.29).

Now $m_X = E[X(t - \tau)]$ since $X(t)$ is wide-sense stationary, so

$$E[Y(t)] = m_X \int_{-\infty}^{\infty} h(\tau)\, d\tau = m_X H(0), \tag{10.41}$$

where $H(f)$ is the transfer function of the system. Thus the mean of the output $Y(t)$ is the constant $m_Y = H(0)m_X$.

The autocorrelation of $Y(t)$ is given by

$$\begin{aligned} E[Y(t)Y(t + \tau)] &= E\left[\int_{-\infty}^{\infty} h(s)X(t - s)\, ds \int_{-\infty}^{\infty} h(r)X(t + \tau - r)\, dr\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(s)h(r)E[X(t - s)X(t + \tau - r)]\, ds\, dr \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(s)h(r)R_X(\tau + s - r)\, ds\, dr, \end{aligned} \tag{10.42}$$

where we have used the fact that $X(t)$ is wide-sense stationary. The expression on the right-hand side of Eq. (10.42) depends only on τ. Thus the autocorrelation of $Y(t)$ depends only on τ, and since the $E[Y(t)]$ is a constant, we conclude that $Y(t)$ is a wide-sense stationary process.

We are now ready to compute the power spectral density of the output of a linear, time-invariant system. Taking the transform of $R_Y(\tau)$ as given in Eq. (10.42), we obtain

$$\begin{aligned} S_Y(f) &= \int_{-\infty}^{\infty} R_Y(\tau)e^{-j2\pi f\tau}\, d\tau \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(s)h(r)R_X(\tau + s - r)e^{-j2\pi f\tau}\, ds\, dr\, d\tau. \end{aligned}$$

Change variables, letting $u = \tau + s - r$:

$$\begin{aligned} S_Y(f) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(s)h(r)R_X(u)e^{-j2\pi f(u-s+r)}\, ds\, dr\, du \\ &= \int_{-\infty}^{\infty} h(s)e^{j2\pi fs}\, ds \int_{-\infty}^{\infty} h(r)e^{-j2\pi fr}\, dr \int_{-\infty}^{\infty} R_X(u)e^{-j2\pi fu}\, du \\ &= H^*(f)H(f)S_X(f) \\ &= |H(f)|^2\, S_X(f), \end{aligned} \tag{10.43}$$

where we have used the definition of the transfer function. *Equation (10.43) relates the input and output power spectral densities to the system transfer function.* Note that $R_Y(\tau)$ can also be found by computing Eq. (10.43) and then taking the inverse Fourier transform.

Equations (10.41) through (10.43) only enable us to determine the mean and autocorrelation function of the output process $Y(t)$. In general this is not enough to determine probabilities of events involving $Y(t)$. However, if the input process is a

Gaussian WSS random process, then as discussed in Section 9.7 the output process will also be a Gaussian WSS random process. Thus the mean and autocorrelation function provided by Eqs. (10.41) through (10.43) are enough to determine all joint pdf's involving the Gaussian random process $Y(t)$.

The cross-correlation between the input and output processes is also of interest:

$$\begin{aligned} R_{Y,X}(\tau) &= E[Y(t+\tau)X(t)] \\ &= E\left[X(t)\int_{-\infty}^{\infty} X(t+\tau-r)h(r)\,dr\right] \\ &= \int_{-\infty}^{\infty} E[X(t)X(t+\tau-r)]h(r)\,dr \\ &= \int_{-\infty}^{\infty} R_X(\tau-r)h(r)\,dr \\ &= R_X(\tau) * h(\tau). \end{aligned} \tag{10.44}$$

By taking the Fourier transform, we obtain the cross-power spectral density:

$$S_{Y,X}(f) = H(f)S_X(f). \tag{10.45a}$$

Since $R_{X,Y}(\tau) = R_{Y,X}(-\tau)$, we have that

$$S_{X,Y}(f) = S^*_{Y,X}(f) = H^*(f)S_X(f). \tag{10.45b}$$

Example 10.9 Filtered White Noise

Find the power spectral density of the output of a linear, time-invariant system whose input is a white noise process.

Let $X(t)$ be the input process with power spectral density

$$S_X(f) = \frac{N_0}{2} \qquad \text{for all } f.$$

The power spectral density of the output $Y(t)$ is then

$$S_Y(f) = |H(f)|^2\frac{N_0}{2}. \tag{10.46}$$

Thus the transfer function completely determines the shape of the power spectral density of the output process.

Example 10.9 provides us with a method for generating WSS processes with arbitrary power spectral density $S_Y(f)$. We simply need to filter white noise through a filter with transfer function $H(f) = \sqrt{S_Y(f)}$. In general this filter will be noncausal. We can usually, but not always, obtain a *causal* filter with transfer function $H(f)$ such that $S_Y(f) = H(f)H^*(f)$. For example, if $S_Y(f)$ is a rational function, that is, if it consists of the ratio of two polynomials, then it is easy to factor $S_X(f)$ into the above form, as shown in the next example. Furthermore any power spectral density can be approximated by a rational function. Thus filtered white noise can be used to synthesize WSS random processes with arbitrary power spectral densities, and hence arbitrary autocorrelation functions.

Example 10.10 Ornstein-Uhlenbeck Process

Find the impulse response of a causal filter that can be used to generate a Gaussian random process with output power spectral density and autocorrelation function

$$S_Y(f) = \frac{\sigma^2}{\alpha^2 + 4\pi^2 f^2} \qquad \text{and} \qquad R_Y(\tau) = \frac{\sigma^2}{2\alpha} e^{-\alpha|\tau|}$$

This power spectral density factors as follows:

$$S_Y(f) = \frac{1}{(\alpha - j2\pi f)} \frac{1}{(\alpha + j2\pi f)} \sigma^2.$$

If we let the filter transfer function be $H(f) = 1/(\alpha + j2\pi f)$, then the impulse response is

$$h(t) = e^{-\alpha t} \qquad \text{for } t \geq 0,$$

which is the response of a causal system. Thus if we filter white Gaussian noise with power spectral density σ^2 using the above filter, we obtain a process with the desired power spectral density.

In Example 9.46, we found the autocorrelation function of the transient response of this filter for a white Gaussian noise input (see Eq. (9.97a)). As was already indicated, when dealing with power spectral densities we assume that the processes are in steady state. Thus as $t \to \infty$ Eq. (9.97a) approaches Eq. (9.97b).

Example 10.11 Ideal Filters

Let $Z(t) = X(t) + Y(t)$, where $X(t)$ and $Y(t)$ are independent random processes with power spectral densities shown in Fig. 10.6(a). Find the output if $Z(t)$ is input into an ideal lowpass filter with transfer function shown in Fig. 10.6(b). Find the output if $Z(t)$ is input into an ideal bandpass filter with transfer function shown in Fig. 10.6(c).

The power spectral density of the output $W(t)$ of the lowpass filter is

$$S_W(f) = |H_{LP}(f)|^2 S_X(f) + |H_{LP}(f)|^2 S_Y(f) = S_X(f),$$

since $H_{LP}(f) = 1$ for the frequencies where $S_X(f)$ is nonzero, and $H_{LP}(f) = 0$ where $S_Y(f)$ is nonzero. Thus $W(t)$ has the same power spectral density as $X(t)$. As indicated in Example 10.5, this does not imply that $W(t) = X(t)$.

To show that $W(t) = X(t)$, in the mean square sense, consider $D(t) = W(t) - X(t)$. It is easily shown that

$$R_D(\tau) = R_W(\tau) - R_{WX}(\tau) - R_{XW}(\tau) + R_X(\tau).$$

The corresponding power spectral density is

$$\begin{aligned} S_D(f) &= S_W(f) - S_{WX}(f) - S_{XW}(f) + S_X(f) \\ &= |H_{LP}(f)|^2 S_X(f) - H_{LP}(f) S_X(f) - H^*_{LP}(f) S_X(f) + S_X(f) \\ &= 0. \end{aligned}$$

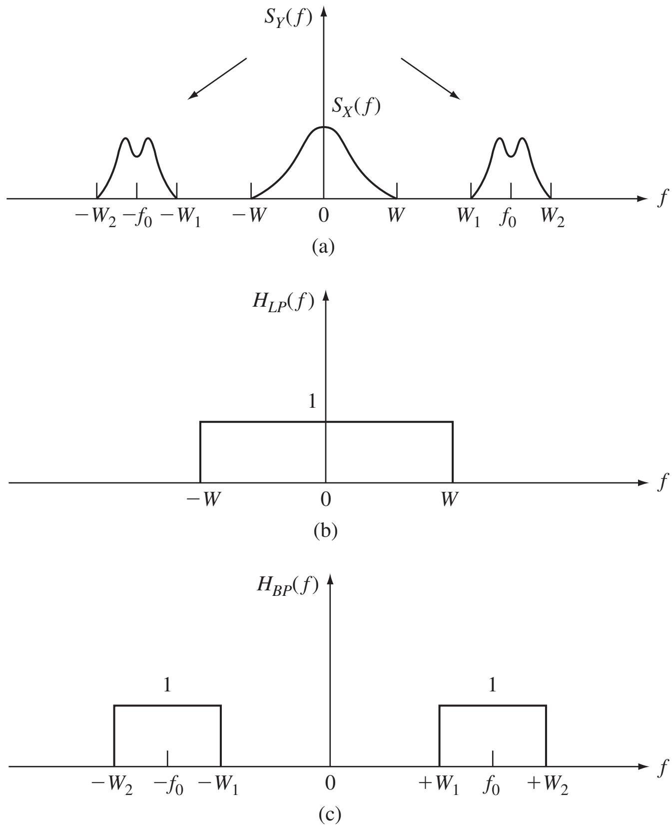

FIGURE 10.6
(a) Input signal to filters is $X(t) + Y(t)$, (b) lowpass filter, (c) bandpass filter.

Therefore $R_D(\tau) = 0$ for all τ, and $W(t) = X(t)$ in the mean square sense since

$$E[(W(t) - X(t))^2] = E[D^2(t)] = R_D(0) = 0.$$

Thus we have shown that the lowpass filter removes $Y(t)$ and passes $X(t)$. Similarly, the bandpass filter removes $X(t)$ and passes $Y(t)$.

Example 10.12

A random telegraph signal is passed through an RC lowpass filter which has transfer function

$$H(f) = \frac{\beta}{\beta + j2\pi f},$$

where $\beta = 1/\text{RC}$ is the time constant of the filter. Find the power spectral density and autocorrelation of the output.

In Example 10.1, the power spectral density of the random telegraph signal with transition rate α was found to be

$$S_X(f) = \frac{4\alpha}{4\alpha^2 + 4\pi^2 f^2}.$$

From Eq. (10.43) we have

$$\begin{aligned} S_Y(f) &= \left(\frac{\beta^2}{\beta^2 + 4\pi^2 f^2}\right)\left(\frac{4\alpha}{4\alpha^2 + 4\pi^2 f^2}\right) \\ &= \frac{4\alpha\beta^2}{\beta^2 - 4\alpha^2}\left\{\frac{1}{4\alpha^2 + 4\pi^2 f^2} - \frac{1}{\beta^2 + 4\pi^2 f^2}\right\}. \end{aligned}$$

$R_Y(\tau)$ is found by inverting the above expression:

$$R_Y(\tau) = \frac{1}{\beta^2 - 4\alpha^2}\{\beta^2 e^{-2\alpha|\tau|} - 2\alpha\beta e^{-\beta|\tau|}\}.$$

10.2.2 Discrete-Time Systems

The results obtained above for continuous-time signals also hold for discrete-time signals after appropriate changes are made from integrals to summations.

Let the **unit-sample response** h_n be the response of a discrete-time, linear, time-invariant system to a unit-sample input δ_n:

$$\delta_n = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases} \tag{10.47}$$

The response of the system to an arbitrary input random process X_n is then given by

$$Y_n = h_n * X_n = \sum_{j=-\infty}^{\infty} h_j X_{n-j} = \sum_{j=-\infty}^{\infty} h_{n-j} X_j. \tag{10.48}$$

Thus discrete-time, linear, time-invariant systems are determined by the unit-sample response h_n. The **transfer function** of such a system is defined by

$$H(f) = \sum_{i=-\infty}^{\infty} h_i e^{-j2\pi f i}. \tag{10.49}$$

The derivation from the previous section can be used to show that if X_n is a wide-sense stationary process, then Y_n is also wide-sense stationary. The mean of Y_n is given by

$$m_Y = m_X \sum_{j=-\infty}^{\infty} h_j = m_X H(0). \tag{10.50}$$

The autocorrelation of Y_n is given by

$$R_Y(k) = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} h_j h_i R_X(k + j - i). \tag{10.51}$$

By taking the Fourier transform of $R_Y(k)$ it is readily shown that the power spectral density of Y_n is

$$S_Y(f) = |H(f)|^2 S_X(f). \tag{10.52}$$

This is the same equation that was found for continuous-time systems.

Finally, we note that if the input process X_n is a Gaussian WSS random process, then the output process Y_n is also a Gaussian WSS random whose statistics are completely determined by the mean and autocorrelation function provided by Eqs. (10.50) through (10.52).

Example 10.13 Filtered White Noise

Let X_n be a white noise sequence with zero mean and average power σ_X^2. If X_n is the input to a linear, time-invariant system with transfer function $H(f)$, then the output process Y_n has power spectral density:

$$S_Y(f) = |H(f)|^2 \sigma_X^2. \tag{10.53}$$

Equation (10.53) provides us with a method for generating discrete-time random processes with arbitrary power spectral densities or autocorrelation functions. If the power spectral density can be written as a rational function of $z = e^{j2\pi f}$ in Eq. (10.24), then a causal filter can be found to generate a process with the power spectral density. Note that this is a generalization of the methods presented in Section 6.6 for generating vector random variables with arbitrary covariance matrix.

Example 10.14 First-Order Autoregressive Process

A first-order autoregressive (AR) process Y_n with zero mean is defined by

$$Y_n = \alpha Y_{n-1} + X_n, \tag{10.54}$$

where X_n is a zero-mean white noise input random process with average power σ_X^2. Note that Y_n can be viewed as the output of the system in Fig. 10.7(a) for an iid input X_n. Find the power spectral density and autocorrelation of Y_n.

The unit-sample response can be determined from Eq. (10.54):

$$h_n = \begin{cases} 0 & n < 0 \\ 1 & n = 0 \\ \alpha^n & n > 0. \end{cases}$$

Note that we require $|\alpha| < 1$ for the system to be stable.[6] Therefore the transfer function is

$$H(f) = \sum_{n=0}^{\infty} \alpha^n e^{-j2\pi f n} = \frac{1}{1 - \alpha e^{-j2\pi f}}.$$

[6] A system is said to be **stable** if $\sum_n |h_n| < \infty$. The response of a stable system to any bounded input is also bounded.

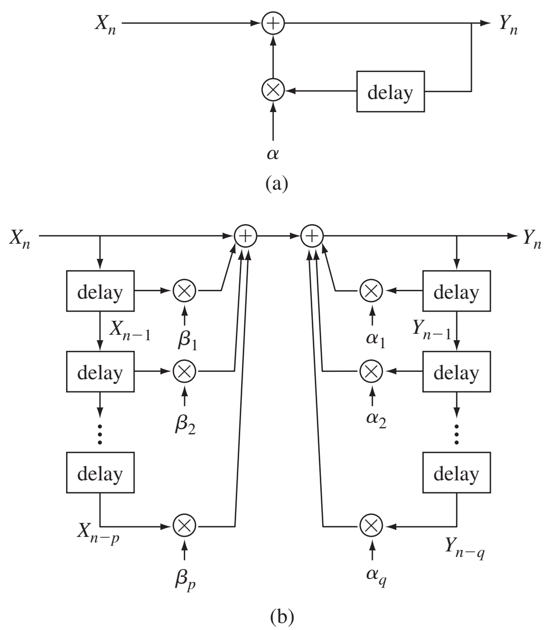

FIGURE 10.7
(a) Generation of AR process; (b) Generation of ARMA process.

Equation (10.52) then gives

$$\begin{aligned} S_Y(f) &= \frac{\sigma_X^2}{(1 - \alpha e^{-j2\pi f})(1 - \alpha e^{j2\pi f})} \\ &= \frac{\sigma_X^2}{1 + \alpha^2 - (\alpha e^{-j2\pi f} + \alpha e^{j2\pi f})} \\ &= \frac{\sigma_X^2}{1 + \alpha^2 - 2\alpha \cos 2\pi f}. \end{aligned}$$

Equation (10.51) gives

$$R_Y(k) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} h_j h_i \sigma_X^2 \delta_{k+j-i} = \sigma_X^2 \sum_{j=0}^{\infty} \alpha^j \alpha^{j+k} = \frac{\sigma_X^2 \alpha^k}{1 - \alpha^2}.$$

Example 10.15 ARMA Random Process

An **autoregressive moving average (ARMA) process** is defined by

$$Y_n = -\sum_{i=1}^{q} \alpha_i Y_{n-i} + \sum_{i'=0}^{p} \beta_{i'} W_{n-i'}, \tag{10.55}$$

where W_n is a WSS, white noise input process. Y_n can be viewed as the output of the recursive system in Fig. 10.7(b) to the input X_n. It can be shown that the transfer function of the linear system defined by the above equation is

$$H(f) = \frac{\sum_{i'=0}^{p} \beta_{i'} e^{-j2\pi f i'}}{1 + \sum_{i=1}^{q} \alpha_i e^{-j2\pi f i}}.$$

The power spectral density of the ARMA process is

$$S_Y(f) = |H(f)|^2 \sigma_W^2.$$

ARMA models are used extensively in random time series analysis and in signal processing. The general **autoregressive process** is the special case of the ARMA process with $\beta_1 = \beta_2 = \cdots = \beta_p = 0$. The general **moving average process** is the special case of the ARMA process with $\alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$. Octave has a function `filter(b, a, x)` which takes a set of coefficients $b = (\beta_1, \beta_2, \ldots, \beta_{p+1})$ and $a = (\alpha_1, \alpha_2, \ldots, \alpha_q)$ as coefficient for a filter as in Eq. (10.55) and produces the output corresponding to the input sequence x. The choice of a and b can lead to a broad range of discrete-time filters.

For example, if we let $a = (1/N, 1/N, \ldots, 1/N)$ we obtain a moving average filter:

$$Y_n = (W_n + W_{n-1} + \cdots + W_{n-N+1})/N.$$

Figure 10.8 shows a zero-mean, unit-variance Gaussian iid sequence W_n and the outputs from an $N = 3$ and an $N = 10$ moving average filter. It can be seen that the $N = 3$ filter moderates the extreme variations but generally tracks the fluctuations in X_n. The $N = 10$ filter on the other hand severely limits the variations and only tracks slower longer-lasting trends.

Figures 10.9(a) and (b) show the result of passing an iid Gaussian sequence X_n through first-order autoregressive filters as in Eq. (10.54). The AR sequence with $\alpha = 0.1$ has low correlation between adjacent samples and so the sequence remains similar to the underlying iid random process. The AR sequence with $\alpha = 0.75$ has higher correlation between adjacent samples which tends to cause longer lasting trends as evident in Fig. 10.9(b).

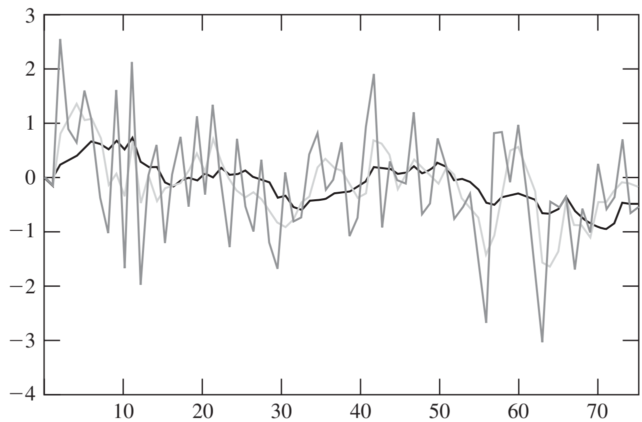

FIGURE 10.8
Moving average process showing iid Gaussian sequence and corresponding $N = 3$, $N = 10$ moving average processes.

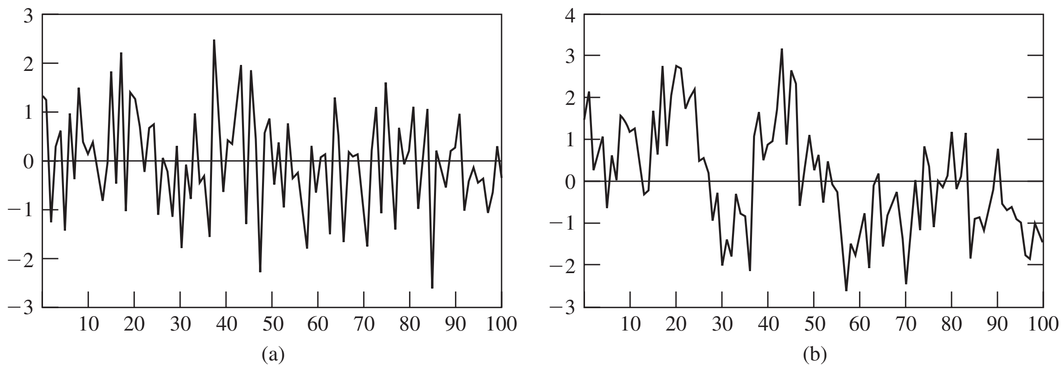

FIGURE 10.9
(a) First-order autoregressive process with $\alpha = 0.1$; (b) with $\alpha = 0.75$.

10.3 BANDLIMITED RANDOM PROCESSES

In this section we consider two important applications that involve random processes with power spectral densities that are nonzero over a finite range of frequencies. The first application involves the sampling theorem, which states that bandlimited random processes can be represented in terms of a sequence of their time samples. This theorem forms the basis for modern digital signal processing systems. The second application involves the modulation of sinusoidal signals by random information signals. Modulation is a key element of all modern communication systems.

10.3.1 Sampling of Bandlimited Random Processes

One of the major technology advances in the twentieth century was the development of digital signal processing technology. All modern multimedia systems depend in some way on the processing of digital signals. Many information signals, e.g., voice, music, imagery, occur naturally as analog signals that are continuous-valued and that vary continuously in time or space or both. The two key steps in making these signals amenable to digital signal processing are: (1). Convert the continuous-time signals into discrete-time signals by sampling the amplitudes; (2) Representing the samples using a fixed number of bits. In this section we introduce the sampling theorem for wide-sense stationary bandlimited random processes, which addresses the conversion of signals into discrete-time sequences.

Let $x(t)$ be a deterministic, finite-energy time signal that has Fourier transform $\widetilde{X}(f) = \mathscr{F}\{x(t)\}$ that is nonzero only in the frequency range $|f| \leq W$. Suppose we sample $x(t)$ every T seconds to obtain the sequence of sample values: $\{\ldots, x(-2T), x(-T), x(0), x(T), \ldots\}$. *The sampling theorem for deterministic signals states that $x(t)$ can be recovered exactly from the sequence of samples if $T \leq 1/2W$ or equivalently* $1/T \geq 2W$, that is, the sampling rate is at least twice the bandwidth of the signal. The minimum sampling rate $1/2W$ is called the **Nyquist sampling rate**. The sampling

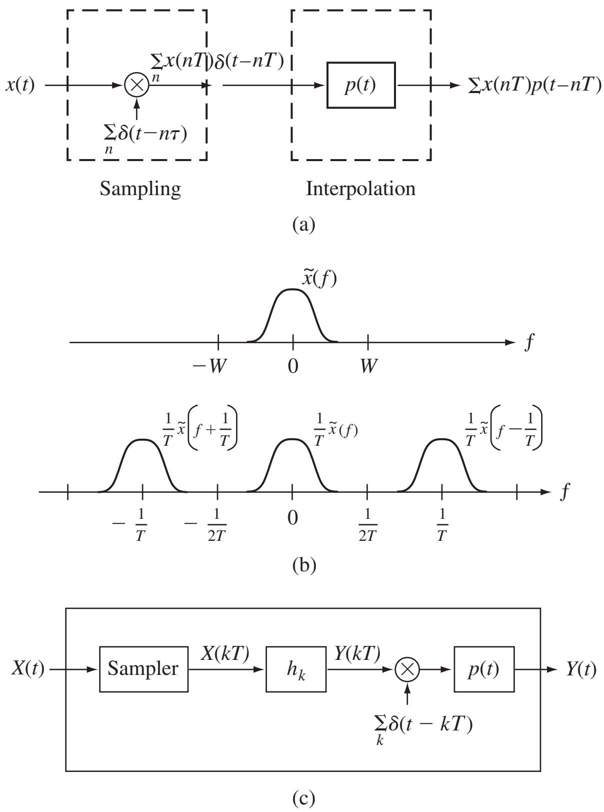

FIGURE 10.10
(a) Sampling and interpolation; (b) Fourier transform of sampled deterministic signal; (c) Sampling, digital filtering, and interpolation.

theorem provides the following interpolation formula for recovering $x(t)$ from the samples:

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT)p(t - nT) \quad \text{where} \quad p(t) = \frac{\sin(\pi t/T)}{\pi t/T}. \tag{10.56}$$

Eq. (10.56) provides us with the interesting interpretation depicted in Fig. 10.10(a). The process of sampling $x(t)$ can be viewed as the multiplication of $x(t)$ by a train of delta functions spaced T seconds apart. The sampled function is then represented by:

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT). \tag{10.57}$$

Eq. (10.56) can be viewed as the response of a linear system with impulse response $p(t)$ to the signal $x_s(t)$. It is easy to show that the $p(t)$ in Eq. (10.56) corresponds to the ideal lowpass filter in Fig. 10.6:

$$P(f) = \mathscr{F}\{p(t)\} = \begin{cases} 1 & -W \le f \le W \\ 0 & |f| > W. \end{cases}$$

The proof of the sampling theorem involves the following steps. We show that

$$\mathscr{F}\left\{\sum_{n=-\infty}^{\infty} x(nT)p(t - nT)\right\} = \frac{1}{T}P(f)\sum_{k=-\infty}^{\infty} \widetilde{X}(f - \frac{k}{T}), \tag{10.58}$$

which consists of the sum of translated versions of $\widetilde{X}(f) = \mathscr{F}\{x(t)\}$, as shown in Fig. 10.10(b). We then observe that as long as $1/T \geq 2W$, then $P(f)$ in the above expressions selects the $k = 0$ term in the summation, which corresponds to $X(f)$. See Problem 10.45 for details.

Example 10.16 Sampling a WSS Random Process

Let $X(t)$ be a WSS process with autocorrelation function $R_X(\tau)$. Find the mean and covariance functions of the discrete-time sampled process $X_n = X(nT)$ for $n = 0, \pm 1, \pm 2, \ldots$.

Since $X(t)$ is WSS, the mean and covariance functions are:

$$m_X(n) = E[X(nT)] = m$$

$$E[X_{n_1}X_{n_2}] = E[X(n_1T)X(n_2T)] = R_X(n_1T - n_2T) = R_X((n_1 - n_2)T).$$

This shows X_n is a WSS discrete-time process.

Let $X(t)$ be a WSS process with autocorrelation function $R_X(\tau)$ and power spectral density $S_X(f)$. Suppose that $S_X(f)$ is bandlimited, that is,

$$S_X(f) = 0 \quad |f| > W.$$

We now show that the sampling theorem can be extended to $X(t)$. Let

$$\hat{X}(t) = \sum_{n=-\infty}^{\infty} X(nT)p(t - nT) \quad \text{where} \quad p(t) = \frac{\sin(\pi t/T)}{\pi t/T}, \tag{10.59}$$

then $\hat{X}(t) = X(t)$ *in the mean square sense*. Recall that equality in the mean square sense does not imply equality for all sample functions, so this version of the sampling theorem is weaker than the version in Eq. (10.56) for finite energy signals.

To show Eq. (10.59) we first note that since $S_X(f) = \mathscr{F}\{R_X(\tau)\}$, we can apply the sampling theorem for deterministic signals to $R_X(\tau)$:

$$R_X(\tau) = \sum_{n=-\infty}^{\infty} R_X(nT)p(\tau - nT). \tag{10.60}$$

Next we consider the mean square error associated with Eq. (10.59):

$$\begin{aligned} E[\{X(t) - \hat{X}(t)\}^2] &= E[\{X(t) - \hat{X}(t)\}X(t)] - E[\{X(t) - \hat{X}(t)\}\hat{X}(t)] \\ &= \left\{E[X(t)X(t)] - E[\hat{X}(t)X(t)]\right\} - \\ &\quad \left\{E[X(t)\hat{X}(t)] - E[\hat{X}(t)\hat{X}(t)]\right\}. \end{aligned}$$

It is easy to show that Eq. (10.60) implies that each of the terms in braces is equal to zero. (See Problem 10.48.) We then conclude that $\hat{X}(t) = X(t)$ in the mean square sense.

Example 10.17 Digital Filtering of a Sampled WSS Random Process

Let $X(t)$ be a WSS process with power spectral density $S_X(f)$ that is nonzero only for $|f| \leq W$. Consider the sequence of operations shown in Fig. 10.10(c): (1) $X(t)$ is sampled at the Nyquist rate; (2) the samples $X(nT)$ are input into a digital filter in Fig. 10.7(b) with $\alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$; and (3) the resulting output sequence Y_n is fed into the interpolation filter. Find the power spectral density of the output $Y(t)$.

The output of the digital filter is given by:

$$Y(kT) = \sum_{n=0}^{p} \beta_n X((k - n)T)$$

and the corresponding autocorrelation from Eq. (10.51) is:

$$R_Y(kT) = \sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n \beta_i R_X((k + n - i)T).$$

The autocorrelation of $Y(t)$ is found from the interpolation formula (Eq. 10.60):

$$\begin{aligned}
R_Y(\tau) &= \sum_{k=-\infty}^{\infty} R_Y(kT)p(\tau - kT) = \sum_{k=-\infty}^{\infty}\sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n\beta_i R_X((k + n - i)T)p(\tau - kT) \\
&= \sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n\beta_i \left\{ \sum_{k=-\infty}^{\infty} R_X((k + n - i)T)p(\tau - kT) \right\} \\
&= \sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n\beta_i R_X(\tau + (n - i)T).
\end{aligned}$$

The output power spectral density is then:

$$\begin{aligned}
S_Y(f) &= \mathscr{F}\{R_Y(\tau)\} = \sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n\beta_i \mathscr{F}\{R_X(\tau + (n - i)T)\} \\
&= \sum_{n=0}^{p}\sum_{i=0}^{p} \beta_n\beta_i S_X(f)e^{-j2\pi f(n-i)T} \\
&= \left\{ \sum_{n=0}^{p} \beta_n e^{-j2\pi f nT} \right\}\left\{ \sum_{i=0}^{p} \beta_i e^{j2\pi f iT} \right\} S_X(f) \\
&= |H(fT)|^2\, S_X(f) \qquad (10.61)
\end{aligned}$$

where $H(f)$ is the transfer function of the digital filter as per Eq. (10.49). The key finding here is the appearance of $H(f)$ *evaluated at* fT. We have obtained a very nice result that characterizes the overall system response in Fig. 10.8 to the continuous-time input $X(t)$. This result is true for more general digital filters, see [Oppenheim and Schafer].

The sampling theorem provides an important bridge between continuous-time and discrete-time signal processing. It gives us a means for implementing the real as well as the simulated processing of random signals. First, we must sample the random process above its Nyquist sampling rate. We can then perform whatever digital processing is necessary. We can finally recover the continuous-time signal by interpolation. The only difference between real signal processing and simulated signal processing is that the former usually has real-time requirements, whereas the latter allows us to perform our processing at whatever rate is possible using the available computing power.

10.3.2 Amplitude Modulation by Random Signals

Many of the transmission media used in communication systems can be modeled as linear systems and their behavior can be specified by a transfer function $H(f)$, which passes certain frequencies and rejects others. Quite often the information signal $A(t)$ (i.e., a speech or music signal) is not at the frequencies that propagate well. The purpose of a **modulator** is to map the information signal $A(t)$ into a transmission signal $X(t)$ that is in a frequency range that propagates well over the desired medium. At the receiver, we need to perform an inverse mapping to recover $A(t)$ from $X(t)$. In this section, we discuss two of the amplitude modulation methods.

Let $A(t)$ be a WSS random process that represents an information signal. In general $A(t)$ will be "lowpass" in character, that is, its power spectral density will be concentrated at low frequencies, as shown in Fig. 10.11(a). An **amplitude modulation** (AM) system produces a transmission signal by multiplying $A(t)$ by a "carrier" signal $\cos(2\pi f_c t + \Theta)$:

$$X(t) = A(t)\cos(2\pi f_c t + \Theta), \tag{10.62}$$

where we assume Θ is a random variable that is uniformly distributed in the interval $(0, 2\pi)$, and Θ and $A(t)$ are independent.

The autocorrelation of $X(t)$ is

$$\begin{aligned} E[X(t+\tau)X(t)] & \\ &= E[A(t+\tau)\cos(2\pi f_c(t+\tau)+\Theta)A(t)\cos(2\pi f_c t+\Theta)] \\ &= E[A(t+\tau)A(t)]E[\cos(2\pi f_c(t+\tau)+\Theta)\cos(2\pi f_c t+\Theta)] \end{aligned}$$

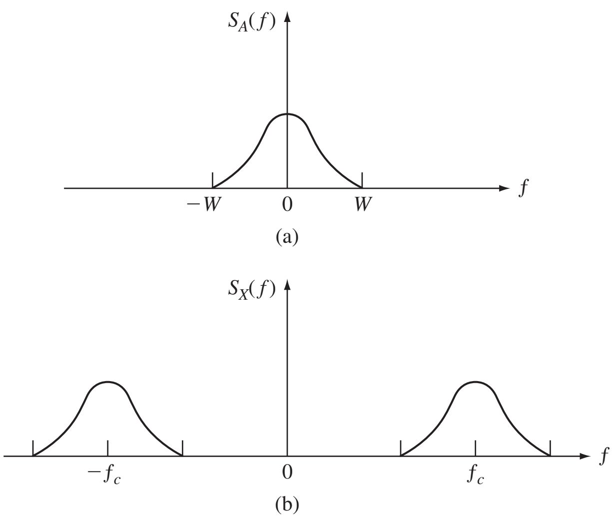

FIGURE 10.11
(a) A lowpass information signal; (b) an amplitude-modulated signal.

$$= R_A(\tau)E\left[\frac{1}{2}\cos(2\pi f_c\tau) + \frac{1}{2}\cos(2\pi f_c(2t + \tau) + 2\Theta)\right]$$
$$= \frac{1}{2}R_A(\tau)\cos(2\pi f_c\tau), \tag{10.63}$$

where we used the fact that $E[\cos(2\pi f_c(2t + \tau) + 2\Theta)] = 0$ (see Example 9.10). Thus $X(t)$ is also a wide-sense stationary random process.

The power spectral density of $X(t)$ is

$$S_X(f) = \mathcal{F}\left\{\frac{1}{2}R_A(\tau)\cos(2\pi f_c\tau)\right\}$$
$$= \frac{1}{4}S_A(f + f_c) + \frac{1}{4}S_A(f - f_c), \tag{10.64}$$

where we used the table of Fourier transforms in Appendix B. Figure 10.11(b) shows $S_X(f)$. It can be seen that the power spectral density of the information signal has been shifted to the regions around $\pm f_c$. $X(t)$ is an example of a **bandpass signal**. Bandpass signals are characterized as having their power spectral density concentrated about some frequency much greater than zero.

The transmission signal is demodulated by multiplying it by the carrier signal and lowpass filtering, as shown in Fig. 10.12. Let

$$Y(t) = X(t)2\cos(2\pi f_c t + \Theta). \tag{10.65}$$

Proceeding as above, we find that

$$S_Y(f) = \frac{1}{2}S_X(f + f_c) + \frac{1}{2}S_X(f - f_c)$$
$$= \frac{1}{2}\{S_A(f + 2f_c) + S_A(f)\} + \frac{1}{2}\{S_A(f) + S_A(f - 2f_c)\}.$$

The ideal lowpass filter passes $S_A(f)$ and blocks $S_A(f \pm 2f_c)$, which is centered about $\pm f$, so the output of the lowpass filter has power spectral density

$$S_Y(f) = S_A(f).$$

In fact, from Example 10.11 we know the output is the original information signal, $A(t)$.

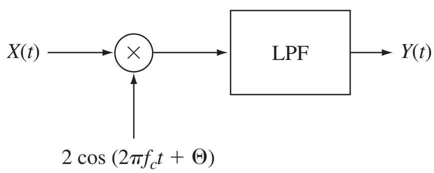

FIGURE 10.12
AM demodulator.

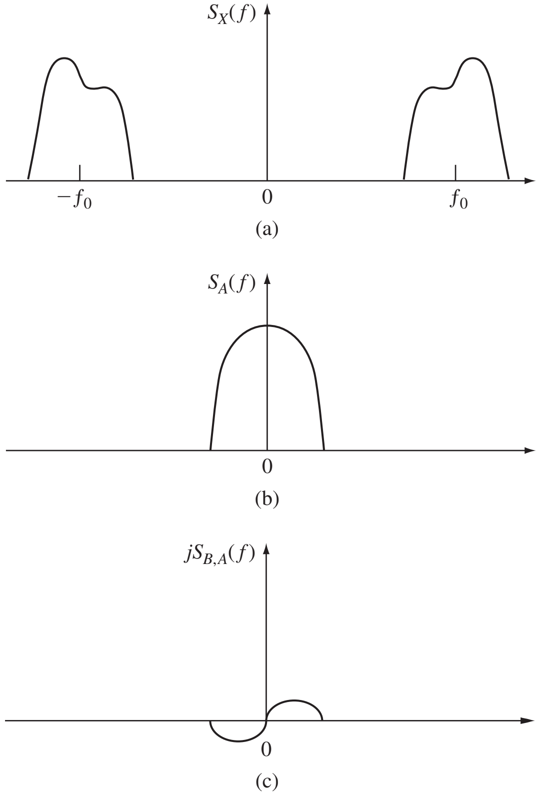

FIGURE 10.13
(a) A general bandpass signal. (b) a real-valued even function of *f*. (c) an imaginary odd function of *f*.

The modulation method in Eq. (10.56) can only produce bandpass signals for which $S_X(f)$ is locally symmetric about f_c, $S_X(f_c + \delta f) = S_X(f_c - \delta f)$ for $|\delta f| < W$, as in Fig. 10.11(b). The method cannot yield real-valued transmission signals whose power spectral density lack this symmetry, such as shown in Fig. 10.13(a). The following **quadrature amplitude modulation** (QAM) method can be used to produce such signals:

$$X(t) = A(t)\cos(2\pi f_c t + \Theta) + B(t)\sin(2\pi f_c t + \Theta), \tag{10.66}$$

where $A(t)$ and $B(t)$ are real-valued, jointly wide-sense stationary random processes, and we require that

$$R_A(\tau) = R_B(\tau) \tag{10.67a}$$

$$R_{B,A}(\tau) = -R_{A,B}(\tau). \tag{10.67b}$$

Note that Eq. (10.67a) implies that $S_A(f) = S_B(f)$, a real-valued, even function of *f*, as shown in Fig. 10.13(b). Note also that Eq. (10.67b) implies that $S_{B,A}(f)$ is a purely imaginary, odd function of *f*, as also shown in Fig. 10.13(c) (see Problem 10.57).

Proceeding as before, we can show that $X(t)$ is a wide-sense stationary random process with autocorrelation function

$$R_X(\tau) = R_A(\tau)\cos(2\pi f_c\tau) + R_{B,A}(\tau)\sin(2\pi f_c\tau) \tag{10.68}$$

and power spectral density

$$S_X(f) = \frac{1}{2}\{S_A(f - f_c) + S_A(f + f_c)\} + \frac{1}{2j}\{S_{BA}(f - f_c) - S_{BA}(f + f_c)\}. \tag{10.69}$$

The resulting power spectral density is as shown in Fig. 10.13(a). Thus QAM can be used to generate real-valued bandpass signals with arbitrary power spectral density.

Bandpass random signals, such as those in Fig. 10.13(a), arise in communication systems when wide-sense stationary white noise is filtered by bandpass filters. Let $N(t)$ be such a process with power spectral density $S_N(f)$. It can be shown that $N(t)$ can be represented by

$$N(t) = N_c(t)\cos(2\pi f_c t + \Theta) - N_s(t)\sin(2\pi f_c t + \Theta), \tag{10.70}$$

where $N_c(t)$ and $N_s(t)$ are jointly wide-sense stationary processes with

$$S_{N_c}(f) = S_{N_s}(f) = \{S_N(f - f_c) + S_N(f + f_c)\}_L \tag{10.71}$$

and

$$S_{N_c,N_s}(f) = j\{S_N(f - f_c) - S_N(f + f_c)\}_L, \tag{10.72}$$

where the subscript L denotes the lowpass portion of the expression in brackets. In words, every real-valued bandpass process can be treated as if it had been generated by a QAM modulator.

Example 10.18 Demodulation of Noisy Signal

The received signal in an AM system is

$$Y(t) = A(t)\cos(2\pi f_c t + \Theta) + N(t),$$

where $N(t)$ is a bandlimited white noise process with spectral density

$$S_N(f) = \begin{cases} \dfrac{N_0}{2} & |f \pm f_c| < W \\ 0 & \text{elsewhere.} \end{cases}$$

Find the signal-to-noise ratio of the recovered signal.

Equation (10.70) allows us to represent the received signal by

$$Y(t) = \{A(t) + N_c(t)\}\cos(2\pi f_c t + \Theta) - N_s(t)\sin(2\pi f_c t + \Theta).$$

The demodulator in Fig. 10.12 is used to recover $A(t)$. After multiplication by $2\cos(2\pi f_c t + \Theta)$, we have

$$\begin{aligned} 2Y(t)\cos(2\pi f_c t + \Theta) &= \{A(t) + N_c(t)\}2\cos^2(2\pi f_c t + \Theta) \\ &\quad - N_s(t)2\cos(2\pi f_c t + \Theta)\sin(2\pi f_c t + \Theta) \\ &= \{A(t) + N_c(t)\}(1 + \cos(4\pi f_c t + 2\Theta)) \\ &\quad - N_s(t)\sin(4\pi f_c t + 2\Theta). \end{aligned}$$

After lowpass filtering, the recovered signal is

$$A(t) + N_c(t).$$

The power in the signal and noise components, respectively, are

$$\sigma_A^2 = \int_{-W}^{W} S_A(f)\, df$$

$$\sigma_{N_c}^2 = \int_{-W}^{W} S_{N_c}(f)\, df = \int_{-W}^{W} \left(\frac{N_0}{2} + \frac{N_0}{2}\right) df = 2WN_0.$$

The output signal-to-noise ratio is then

$$\text{SNR} = \frac{\sigma_A^2}{2WN_0}.$$

10.4 OPTIMUM LINEAR SYSTEMS

Many problems can be posed in the following way. We observe a discrete-time, zero-mean process X_α over a certain time interval $I = \{t - a, \dots, t + b\}$, and we are required to use the $a + b + 1$ resulting observations $\{X_{t-a}, \dots, X_t, \dots, X_{t+b}\}$ to obtain an estimate Y_t for some other (presumably related) zero-mean process Z_t. The estimate Y_t is required to be linear, as shown in Fig. 10.14:

$$Y_t = \sum_{\beta=t-a}^{t+b} h_{t-\beta} X_\beta = \sum_{\beta=-b}^{a} h_\beta X_{t-\beta}. \tag{10.73}$$

The figure of merit for the estimator is the mean square error

$$E[e_t^2] = E[(Z_t - Y_t)^2], \tag{10.74}$$

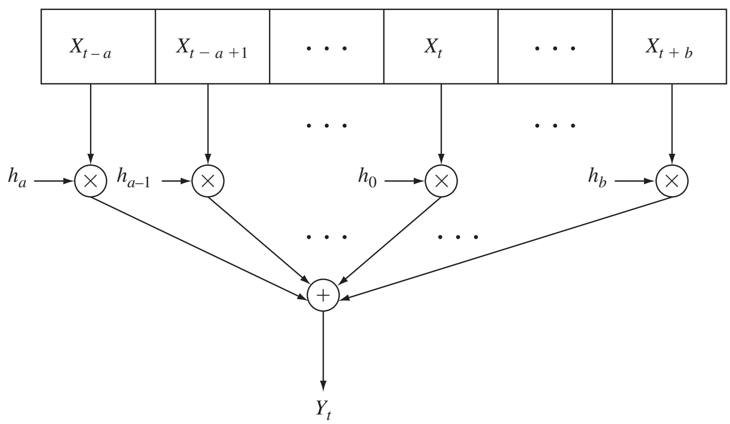

FIGURE 10.14
A linear system for producing an estimate Y_t.

and we seek to find the **optimum filter**, which is characterized by the impulse response h_β that minimizes the mean square error.

Examples 10.19 and 10.20 show that different choices of Z_t and X_α and of observation interval correspond to different estimation problems.

Example 10.19 Filtering and Smoothing Problems

Let the observations be the sum of a "desired signal" Z_α plus unwanted "noise" N_α:

$$X_\alpha = Z_\alpha + N_\alpha \qquad \alpha \in I.$$

We are interested in estimating the desired signal at time t. The relation between t and the observation interval I gives rise to a variety of estimation problems.

If $I = (-\infty, t)$, that is, $a = \infty$ and $b = 0$, then we have a **filtering** problem where we estimate Z_t in terms of noisy observations of the past and present. If $I = (t - a, t)$, then we have a filtering problem in which we estimate Z_t in terms of the $a + 1$ most recent noisy observations.

If $I = (-\infty, \infty)$, that is, $a = b = \infty$, then we have a **smoothing** problem where we are attempting to recover the signal from its entire noisy version. There are applications where this makes sense, for example, if the entire realization X_α has been recorded and the estimate Z_t is obtained by "playing back" X_α.

Example 10.20 Prediction

Suppose we want to predict Z_t in terms of its recent past: $\{Z_{t-a}, \ldots, Z_{t-1}\}$. The general estimation problem becomes this **prediction** problem if we let the observation X_α be the past a values of the signal Z_α, that is,

$$X_\alpha = Z_\alpha \qquad t - a \le \alpha \le t - 1.$$

The estimate Y_t is then a linear prediction of Z_t in terms of its most recent values.

10.4.1 The Orthogonality Condition

It is easy to show that the optimum filter must satisfy the **orthogonality condition** (see Eq. 6.56), which states that the error e_t must be orthogonal to all the observations X_α, that is,

$$\begin{aligned} 0 &= E[e_t X_\alpha] \qquad \text{for all } \alpha \in I \\ &= E[(Z_t - Y_t)X_\alpha] = 0, \end{aligned} \tag{10.75}$$

or equivalently,

$$E[Z_t X_\alpha] = E[Y_t X_\alpha] \qquad \text{for all } \alpha \in I. \tag{10.76}$$

If we substitute Eq. (10.73) into Eq. (10.76) we find

$$\begin{aligned} E[Z_t X_\alpha] &= E\left[\sum_{\beta=-b}^{a} h_\beta X_{t-\beta} X_\alpha\right] \qquad \text{for all } \alpha \in I \\ &= \sum_{\beta=-b}^{a} h_\beta E[X_{t-\beta} X_\alpha] \\ &= \sum_{\beta=-b}^{a} h_\beta R_X(t - \alpha - \beta) \qquad \text{for all } \alpha \in I. \end{aligned} \tag{10.77}$$

Equation (10.77) shows that $E[Z_t X_\alpha]$ depends only on $t - \alpha$, and thus X_α and Z_t are jointly wide-sense stationary processes. Therefore, we can rewrite Eq. (10.77) as follows:

$$R_{Z,X}(t - \alpha) = \sum_{\beta=-b}^{a} h_\beta R_X(t - \beta - \alpha) \qquad t - a \le \alpha \le t + b.$$

Finally, letting $m = t - \alpha$, we obtain the following key equation:

$$R_{Z,X}(m) = \sum_{\beta=-b}^{a} h_\beta R_X(m - \beta) \qquad -b \le m \le a. \tag{10.78}$$

The optimum linear filter must satisfy the set of $a + b + 1$ *linear equations given by Eq. (10.78).* Note that Eq. (10.78) is identical to Eq. (6.60) for estimating a random variable by a linear combination of several random variables. The wide-sense stationarity of the processes reduces this estimation problem to the one considered in Section 6.5.

In the above derivation we deliberately used the notation Z_t instead of Z_n to suggest that the same development holds for *continuous-time estimation*. In particular, suppose we seek a linear estimate $Y(t)$ for the continuous-time random process $Z(t)$ in terms of observations of the continuous-time random process $X(\alpha)$ in the time interval $t - a \le \alpha \le t + b$:

$$Y(t) = \int_{t-a}^{t+b} h(t - \beta) X(\beta)\, d\beta = \int_{-b}^{a} h(\beta) X(t - \beta)\, d\beta.$$

It can then be shown that the filter $h(\beta)$ that minimizes the mean square error is specified by

$$R_{Z,X}(\tau) = \int_{-b}^{a} h(\beta) R_X(\tau - \beta)\, d\beta \qquad -b \le \tau \le a. \tag{10.79}$$

Thus in the time-continuous case we obtain an integral equation instead of a set of linear equations. The analytic solution of this integral equation can be quite difficult, but the equation can be solved numerically by approximating the integral by a summation.[7]

We now determine the mean square error of the optimum filter. First we note that for the optimum filter, the error e_t and the estimate Y_t are orthogonal since

$$E[e_t Y_t] = E\left[e_t \sum h_{t-\beta} X_\beta\right] = \sum h_{t-\beta} E[e_t X_\beta] = 0,$$

where the terms inside the last summation are 0 because of Eq. (10.75). Since $e_t = Z_t - Y_t$, the mean square error is then

$$\begin{aligned} E[e_t^2] &= E[e_t(Z_t - Y_t)] \\ &= E[e_t Z_t], \end{aligned}$$

[7]Equation (10.79) can also be solved by using the Karhunen-Loeve expansion.

since e_t and Y_t are orthogonal. Substituting for e_t yields

$$\begin{aligned} E[e_t^2] &= E[(Z_t - Y_t)Z_t] = E[Z_tZ_t] - E[Y_tZ_t] \\ &= R_Z(0) - E[Z_tY_t] \\ &= R_Z(0) - E\left[Z_t \sum_{\beta=-b}^{a} h_\beta X_{t-\beta} \right] \\ &= R_Z(0) - \sum_{\beta=-b}^{a} h_\beta R_{Z,X}(\beta). \end{aligned} \tag{10.80}$$

Similarly, it can be shown that the mean square error of the optimum filter in the continuous-time case is

$$E[e^2(t)] = R_Z(0) = \int_{-b}^{a} h(\beta) R_{Z,X}(\beta)\, d\beta. \tag{10.81}$$

The following theorems summarize the above results.

Theorem

Let X_t and Z_t be discrete-time, zero-mean, jointly wide-sense stationary processes, and let Y_t be an estimate for Z_t of the form

$$Y_t = \sum_{\beta=t-a}^{t+b} h_{t-\beta} X_\beta = \sum_{\beta=-b}^{a} h_\beta X_{t-\beta}.$$

The filter that minimizes $E[(Z_t - Y_t)^2]$ satisfies the equation

$$R_{Z,X}(m) = \sum_{\beta=-b}^{a} h_\beta R_X(m - \beta) \qquad -b \le m \le a$$

and has mean square error given by

$$E[(Z_t - Y_t)^2] = R_Z(0) - \sum_{\beta=-b}^{a} h_\beta R_{Z,X}(\beta).$$

Theorem

Let $X(t)$ and $Z(t)$ be continuous-time, zero-mean, jointly wide-sense stationary processes, and let $Y(t)$ be an estimate for $Z(t)$ of the form

$$Y(t) = \int_{t-a}^{t+b} h(t - \beta) X(\beta)\, d\beta = \int_{-b}^{a} h(\beta) X(t - \beta)\, d\beta.$$

The filter $h(\beta)$ that minimizes $E[(Z(t) - Y(t))^2]$ satisfies the equation

$$R_{Z,X}(\tau) = \int_{-b}^{a} h(\beta) R_X(\tau - \beta)\, d\beta \qquad -b \le \tau \le a$$

and has mean square error given by

$$E[(Z(t) - Y(t))^2] = R_Z(0) - \int_{-b}^{a} h(\beta) R_{Z,X}(\beta)\, d\beta.$$

Example 10.21 Filtering of Signal Plus Noise

Suppose we are interested in estimating the signal Z_n from the $p + 1$ most recent noisy observations:

$$X_\alpha = Z_\alpha + N_\alpha \qquad \alpha \in I = \{n - p, \dots, n - 1, n\}.$$

Find the set of linear equations for the optimum filter if Z_α and N_α are independent random processes.

For this choice of observation interval, Eq. (10.78) becomes

$$R_{Z,X}(m) = \sum_{\beta=0}^{p} h_\beta R_X(m - \beta) \qquad m \in \{0, 1, \dots, p\}. \tag{10.82}$$

The cross-correlation terms in Eq. (10.82) are given by

$$R_{Z,X}(m) = E[Z_n X_{n-m}] = E[Z_n(Z_{n-m} + N_{n-m})] = R_Z(m).$$

The autocorrelation terms are given by

$$\begin{aligned} R_X(m - \beta) &= E[X_{n-\beta} X_{n-m}] = E[(Z_{n-\beta} + N_{n-\beta})(Z_{n-m} + N_{n-m})] \\ &= R_Z(m - \beta) + R_{Z,N}(m - \beta) \\ &\quad + R_{N,Z}(m - \beta) + R_N(m - \beta) \\ &= R_Z(m - \beta) + R_N(m - \beta), \end{aligned}$$

since Z_α and N_α are independent random processes. Thus Eq. (10.82) for the optimum filter becomes

$$R_Z(m) = \sum_{\beta=0}^{p} h_\beta \{R_Z(m - \beta) + R_N(m - \beta)\} \quad m \in \{0, 1, \dots, p\}. \tag{10.83}$$

This set of $p + 1$ linear equations in $p + 1$ unknowns h_β is solved by matrix inversion.

Example 10.22 Filtering of AR Signal Plus Noise

Find the set of equations for the optimum filter in Example 10.21 if Z_α is a first-order autoregressive process with average power σ_Z^2 and parameter r, $|r| < 1$, and N_α is a white noise process with average power σ_N^2.

The autocorrelation for a first-order autoregressive process is given by

$$R_Z(m) = \sigma_Z^2 r^{|m|} \qquad m = 0, \pm 1, \pm 2, \dots.$$

(See Problem 10.42.) The autocorrelation for the white noise process is

$$R_N(m) = \sigma_N^2\, \delta(m).$$

Substituting $R_Z(m)$ and $R_N(m)$ into Eq. (10.83) yields the following set of linear equations:

$$\sigma_Z^2 r^{|m|} = \sum_{\beta=0}^{p} h_\beta (\sigma_Z^2 r^{|m-\beta|} + \sigma_N^2 \delta(m - \beta)) \qquad m \in \{0, \dots, p\}. \tag{10.84}$$

If we divide both sides of Eq. (10.84) by σ_Z^2 and let $\Gamma = \sigma_N^2/\sigma_Z^2$, we obtain the following matrix equation:

$$\begin{bmatrix} 1+\Gamma & r & r^2 & \cdots & r^p \\ r & 1+\Gamma & r & \cdots & r^{p-1} \\ r^2 & r & 1+\Gamma & \cdots & r^{p-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r^p & r^{p-1} & r^{p-2} & \cdots & 1+\Gamma \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \cdot \\ \cdot \\ h_p \end{bmatrix} = \begin{bmatrix} 1 \\ r \\ \cdot \\ \cdot \\ r^p \end{bmatrix}. \tag{10.85}$$

Note that when the noise power is zero, i.e., $\Gamma = 0$, then the solution is $h_0 = 1, h_j = 0$, $j = 1, \ldots, p$, that is, no filtering is required to obtain Z_n.

Equation (10.85) can be readily solved using Octave. The following function will compute the optimum linear coefficients and the mean square error of the optimum predictor:

```
function [mse]= Lin_Est_AR (order,rho,varsig,varnoise)
n=[0:1:order-1]
r=varsig*rho.^n;
R=varnoise*eye(order)+toeplitz(r);
H=inv(R)*transpose(r)
mse=varsig-transpose(H)*transpose(r);
endfunction
```

Table 10.1 gives the values of the optimal predictor coefficients and the mean square error as the order of the estimator is increased for the first-order autoregressive process with $\sigma_Z^2 = 4$, $r = 0.9$, and noise variance $\sigma_N^2 = 4$. It can be seen that the predictor places heavier weight on more recent samples, which is consistent with the higher correlation of such samples with the current sample. For smaller values of r, the correlation for distant samples drops off more quickly and the coefficients place even lower weighting on them. The mean square error can also be seen to decrease with increasing order $p + 1$ of the estimator. Increasing the first few orders provides significant improvements, but a point of diminishing returns is reached around $p + 1 = 3$.

10.4.2 Prediction

The linear prediction problem arises in many signal processing applications. In Example 6.31 in Chapter 6, we already discussed the linear prediction of speech signals. In general, we wish to predict Z_n in terms of $Z_{n-1}, Z_{n-2}, \ldots, Z_{n-p}$:

$$Y_n = \sum_{\beta=1}^{p} h_\beta Z_{n-\beta}.$$

TABLE 10.1 Effect of predictor order on MSE performance.

$p + 1$	MSE	Coefficients				
1	2.0000	0.5				
2	1.4922	0.37304	0.28213			
3	1.3193	0.32983	0.22500	0.17017		
4	1.2549	0.31374	0.20372	0.13897	0.10510	
5	1.2302	0.30754	0.19552	0.12696	0.08661	0.065501

For this problem, $X_\alpha = Z_\alpha$, so Eq. (10.79) becomes

$$R_Z(m) = \sum_{\beta=1}^{p} h_\beta R_Z(m - \beta) \qquad m \in \{1, \ldots, p\}. \tag{10.86a}$$

In matrix form this equation becomes

$$\begin{bmatrix} R_Z(1) \\ R_Z(2) \\ . \\ . \\ R_Z(p) \end{bmatrix} = \begin{bmatrix} R_Z(0) & R_Z(1) & R_Z(2) & \cdots & R_Z(p-1) \\ R_Z(1) & R_Z(0) & R_Z(1) & \cdots & R_Z(p-2) \\ . & . & . & . & . \\ . & . & . & . & R_Z(1) \\ R_Z(p-1) & . & . & R_Z(1) & R_Z(0) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ . \\ . \\ h_p \end{bmatrix}$$

$$= \mathbf{R_Z h}. \tag{10.86b}$$

Equations (10.86a) and (10.86b) are called the **Yule-Walker equations**.

Equation (10.80) for the mean square error becomes

$$E[e_n^2] = R_Z(0) - \sum_{\beta=1}^{p} h_\beta R_Z(\beta). \tag{10.87}$$

By inverting the $p \times p$ matrix $\mathbf{R_Z}$, we can solve for the vector of filter coefficients h.

Example 10.23 Prediction for Long-Range and Short-Range Dependent Processes

Let $X_1(t)$ be a discrete-time first-order autoregressive process with $\sigma_X^2 = 1$ and $r = 0.7411$, and let $X_2(t)$ be a discrete-time long-range dependent process with autocovariance given by Eq. (9.109), $\sigma_X^2 = 1$, and $H = 0.9$. Both processes have $C_X(1) = 0.7411$, but the autocovariance of $X_1(t)$ decreases exponentially while that of $X_2(t)$ has long-range dependence. Compare the performance of the optimal linear predictor for these processes for short-term as well as long-term predictions.

The optimum linear coefficients and the associated mean square error for the long-range dependent process can be calculated using the following code. The function can be modified for the autoregressive case.

```
function mse= Lin_Pred_LR(order,Hurst,varsig)
n=[0:1:order-1]
H2=2*Hurst
r=varsig*((1+n).^H2−2*(n.^H2)+abs(n−1).^H2)/2
rz=varsig*((2+n).^H2−2*((n+1).^H2)+(n).^H2)/2
R=toeplitz(r);
H=transpose(inv(R)*transpose(rz))
mse=varsig-H*transpose(rz)
endfunction
```

Table 10.2 below compares the mean square errors and the coefficients of the two processes in the case of short-term prediction. The predictor for $X_1(t)$ attains all of the benefit of prediction with a $p = 1$ system. The optimum predictors for higher-order systems set the other coefficients to zero, and the mean square error remains at 0.4577. The predictor for $X_2(t)$

TABLE 10.2(a) Short-term prediction: autoregressive, $r = 0.7411, \sigma_X^2 = 1, C_X(1) = 0.7411$.

p	MSE	Coefficients	
1	0.45077	0.74110	
2	0.45077	0.74110	0

TABLE 10.2(b) Short-term prediction: long-range dependent process, Hurst $= 0.9, \sigma_X^2 = 1, C_X(1) = 0.7411$.

p	MSE	Coefficients				
1	0.45077	0.74110				
2	0.43625	0.60809	0.17948			
3	0.42712	0.582127	0.091520	0.144649		
4	0.42253	0.567138	0.082037	0.084329	0.103620	
5	0.41964	0.558567	0.075061	0.077543	0.056707	0.082719

achieves most of the possible performance with a $p = 1$ system, but small reductions in mean square error do accrue by adding more coefficients. This is due to the persistent correlation among the values in $X_2(t)$.

Table 10.3 shows the dramatic impact of long-range dependence on prediction performance. We modified Eq. (10.86) to provide the optimum linear predictor for X_t based on two observations X_{t-10} and X_{t-20} that are in the relatively remote past. $X_1(t)$ and its previous values are almost uncorrelated, so the best predictor has a mean square error of almost 1, which is the variance of $X_1(t)$. On the other hand, $X_2(t)$ retains significant correlation with its previous values and so the mean square error provides a significant reduction from the unit variance. Note that the second-order predictor places significant weight on the observation 20 samples in the past.

TABLE 10.3(a) Long-term prediction: autoregressive, $r = 0.7411, \sigma_X^2 = 1, C_X(1) = 0.7411$.

p	MSE	Coefficients	
1	0.99750	0.04977	
2	0.99750	0.04977	0

TABLE 10.3(b) Long-term prediction: long-range dependent process, Hurst $= 0.9, \sigma_X^2 = 1, C_X(1) = 0.7411$.

p	MSE	Coefficients	
10	0.79354	0.45438	
10;20	0.74850	0.34614	0.23822

10.4.3 Estimation Using the Entire Realization of the Observed Process

Suppose that Z_t is to be estimated by a linear function Y_t of the entire realization of X_t, that is, $a = b = \infty$ and Eq. (10.73) becomes

$$Y_t = \sum_{\beta=-\infty}^{\infty} h_\beta X_{t-\beta}.$$

In the case of continuous-time random processes, we have

$$Y(t) = \int_{-\infty}^{\infty} h(\beta) X(t - \beta)\, d\beta.$$

The optimum filters must satisfy Eqs. (10.78) and (10.79), which in this case become

$$R_{Z,X}(m) = \sum_{\beta=-\infty}^{\infty} h_\beta R_X(m - \beta) \qquad \text{for all } m \tag{10.88a}$$

$$R_{Z,X}(\tau) = \int_{-\infty}^{\infty} h(\beta) R_X(\tau - \beta)\, d\beta \qquad \text{for all } \tau. \tag{10.88b}$$

The Fourier transform of the first equation and the Fourier transform of the second equation both yield the same expression:

$$S_{Z,X}(f) = H(f) S_X(f),$$

which is readily solved for the transfer function of the optimum filter:

$$H(f) = \frac{S_{Z,X}(f)}{S_X(f)}. \tag{10.89}$$

The impulse response of the optimum filter is then obtained by taking the appropriate inverse transform. In general the filter obtained from Eq. (10.89) will be noncausal, that is, its impulse response is nonzero for $t < 0$. We already indicated that there are applications where this makes sense, namely, in situations where the entire realization X_α is recorded and the estimate Z_t is obtained in "nonreal time" by "playing back" X_α.

Example 10.24 Infinite Smoothing

Find the transfer function for the optimum filter for estimating $Z(t)$ from $X(\alpha) = Z(\alpha) + N(\alpha)$, $\alpha \in (-\infty, \infty)$, where $Z(\alpha)$ and $N(\alpha)$ are independent, zero-mean random processes.

The cross-correlation between the observation and the desired signal is

$$\begin{aligned} R_{Z,X}(\tau) &= E[Z(t + \tau)X(t)] = E[Z(t + \tau)(Z(t) + N(t))] \\ &= E[Z(t + \tau)Z(t)] + E[Z(t + \tau)N(t)] \\ &= R_Z(\tau), \end{aligned}$$

since $Z(t)$ and $N(t)$ are zero-mean, independent random processes. The cross-power spectral density is then

$$S_{Z,X}(t) = S_Z(f). \tag{10.90}$$

The autocorrelation of the observation process is

$$R_X(\tau) = E[(Z(t+\tau) + N(t+\tau))(Z(t) + N(t))]$$
$$= R_Z(\tau) + R_N(\tau).$$

The corresponding power spectral density is

$$S_X(f) = S_Z(f) + S_N(f). \tag{10.91}$$

Substituting Eqs. (10.90) and (10.91) into Eq. (10.89) gives

$$H(f) = \frac{S_Z(f)}{S_Z(f) + S_N(f)}. \tag{10.92}$$

Note that the optimum filter $H(f)$ is nonzero only at the frequencies where $S_Z(f)$ is nonzero, that is, where the signal has power content. By dividing the numerator and denominator of Eq. (10.92) by $S_Z(f)$, we see that $H(f)$ emphasizes the frequencies where the ratio of signal to noise power density is large.

*10.4.4 Estimation Using Causal Filters

Now, suppose that Z_t is to be estimated using only the past and present of X_α, that is, $I = (-\infty, t)$. Equations (10.78) and (10.79) become

$$R_{Z,X}(m) = \sum_{\beta=0}^{\infty} h_\beta R_X(m-\beta) \qquad \text{for all } m \tag{10.93a}$$

$$R_{Z,X}(\tau) = \int_0^{\infty} h(\beta) R_X(\tau-\beta)\, d\beta \qquad \text{for all } \tau. \tag{10.93b}$$

Equations (10.93a) and (10.93b) are called the **Wiener-Hopf equations** and, though similar in appearance to Eqs. (10.88a) and (10.88b), are considerably more difficult to solve.

First, let us consider the special case where the observation process is white, that is, for the discrete-time case $R_X(m) = \delta_m$. Equation (10.93a) is then

$$R_{Z,X}(m) = \sum_{\beta=0}^{\infty} h_\beta\, \delta_{m-\beta} = h_m \qquad m \geq 0. \tag{10.94}$$

Thus in this special case, the optimum causal filter has coefficients given by

$$h_m = \begin{cases} 0 & m < 0 \\ R_{Z,X}(m) & m \geq 0. \end{cases}$$

The corresponding transfer function is

$$H(f) = \sum_{m=0}^{\infty} R_{Z,X}(m) e^{-j2\pi f m}. \tag{10.95}$$

Note Eq. (10.95) is *not* $S_{Z,X}(f)$, since the limits of the Fourier transform in Eq. (10.95) do not extend from $-\infty$ to $+\infty$. However, $H(f)$ can be obtained from $S_{Z,X}(f)$ by finding $h_m = \mathscr{F}^{-1}[S_{Z,X}(f)]$, keeping the causal part (i.e., h_m for $m \geq 0$) and setting the noncausal part to 0.

We now show how the solution of the above special case can be used to solve the general case. It can be shown that under very general conditions, the power spectral density of a random process can be factored into the form

$$S_X(f) = |G(f)|^2 = G(f)G^*(f), \tag{10.96}$$

where $G(f)$ and $1/G(f)$ are *causal* filters.[8] This suggests that we can find the optimum filter in two steps, as shown in Fig. 10.15. First, we pass the observation process through a "whitening" filter with transfer function $W(f) = 1/G(f)$ to produce a white noise process X'_n, since

$$S_{X'}(f) = |W(f)|^2 S_X(f) = \frac{|G(f)|^2}{|G(f)|^2} = 1 \qquad \text{for all } f.$$

Second, we find the best estimator for Z_n using the whitened observation process X'_n as given by Eq. (10.95). The filter that results from the tandem combination of the whitening filter and the estimation filter is the solution to the Wiener-Hopf equations.

The transfer function of the second filter in Fig. 10.15 is

$$H_2(f) = \sum_{m=0}^{\infty} R_{Z,X'}(m) e^{-j2\pi f m} \tag{10.97}$$

by Eq. (10.95). To evaluate Eq. (10.97) we need to find

$$\begin{aligned} R_{Z,X'}(k) &= E[Z_{n+k} X'_n] \\ &= \sum_{i=0}^{\infty} w_i E[Z_{n+k} X_{n-i}] \\ &= \sum_{i=0}^{\infty} w_i R_{Z,X}(k+i), \end{aligned} \tag{10.98}$$

where w_i is the impulse response of the whitening filter. The Fourier transform of Eq. (10.98) gives an expression that is easier to work with:

$$S_{Z,X'}(f) = W^*(f) S_{Z,X}(f) = \frac{S_{Z,X}(f)}{G^*(f)}. \tag{10.99}$$

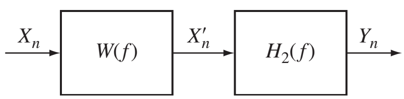

FIGURE 10.15
Whitening filter approach for solving Wiener-Hopf equations.

[8]The method for factoring $S_X(f)$ as specified by Eq. (10.96) is called **spectral factorization**. See Example 10.10 and the references at the end of the chapter.

The inverse Fourier transform of Eq. (10.99) yields the desired $R_{Z,X'}(k)$, which can then be substituted into Eq. (10.97) to obtain $H_2(f)$.

In summary, the optimum filter is found using the following procedure:

1. Factor $S_X(f)$ as in Eq. (10.96) and obtain a causal whitening filter $W(f) = 1/G(f)$.
2. Find $R_{Z,X'}(k)$ from Eq. (10.98) or from Eq. (10.99).
3. $H_2(f)$ is then given by Eq. (10.97).
4. The optimum filter is then

$$H(f) = W(f)H_2(f). \tag{10.100}$$

This procedure is valid for the continuous-time version of the optimum causal filter problem, after appropriate changes are made from summations to integrals. The following example considers a continuous-time problem.

Example 10.25 Wiener Filter

Find the optimum causal filter for estimating a signal $Z(t)$ from the observation $X(t) = Z(t) + N(t)$, where $Z(t)$ and $N(t)$ are independent random processes, $N(t)$ is zero-mean white noise density 1, and $Z(t)$ has power spectral density

$$S_Z(f) = \frac{2}{1 + 4\pi^2 f^2}.$$

The optimum filter in this problem is called the **Wiener filter**.

The cross-power spectral density between $Z(t)$ and $X(t)$ is

$$S_{Z,X}(f) = S_Z(f),$$

since the signal and noise are independent random processes. The power spectral density for the observation process is

$$\begin{aligned} S_X(f) &= S_Z(f) + S_N(f) \\ &= \frac{3 + 4\pi^2 f^2}{1 + 4\pi^2 f^2} \\ &= \left(\frac{j2\pi f + \sqrt{3}}{j2\pi f + 1}\right)\left(\frac{-j2\pi f + \sqrt{3}}{-j2\pi f + 1}\right). \end{aligned}$$

If we let

$$G(f) = \frac{j2\pi f + \sqrt{3}}{j2\pi f + 1},$$

then it is easy to verify that $W(f) = 1/G(f)$ is the whitening causal filter.

Next we evaluate Eq. (10.99):

$$\begin{aligned} S_{Z,X'}(f) &= \frac{S_{Z,X}(f)}{G^*(f)} = \frac{2}{1 + 4\pi^2 f^2}\frac{1 - j2\pi f}{\sqrt{3} - j2\pi f} \\ &= \frac{2}{(1 + j2\pi f)(\sqrt{3} - j2\pi f)} \\ &= \frac{c}{1 + j2\pi f} + \frac{c}{\sqrt{3} - j2\pi f}, \end{aligned} \tag{10.101}$$

where $c = 2/(1 + \sqrt{3})$. If we take the inverse Fourier transform of $S_{Z,X'}(f)$, we obtain

$$R_{Z,X'}(\tau) = \begin{cases} ce^{-\tau} & \tau > 0 \\ ce^{\sqrt{3}\tau} & \tau < 0. \end{cases}$$

Equation (10.97) states that $H_2(f)$ is given by the Fourier transform of the $\tau > 0$ portion of $R_{Z,X'}(\tau)$:

$$H_2(f) = \mathcal{F}\{ce^{-\tau}u(\tau)\} = \frac{c}{1 + j2\pi f}.$$

Note that we could have gotten this result directly from Eq. (10.101) by noting that only the first term gives rise to the positive-time (i.e., causal) component.

The optimum filter is then

$$H(f) = \frac{1}{G(f)}H_2(f) = \frac{c}{\sqrt{3} + j2\pi f}.$$

The impulse response of this filter is

$$h(t) = ce_t^{-\sqrt{3}} \qquad t > 0.$$

10.5 THE KALMAN FILTER

The optimum linear systems considered in the previous section have two limitations: (1) They assume wide-sense stationary signals; and (2) The number of equations grows with the size of the observation set. In this section, we consider an estimation approach that assumes signals have a certain structure. This assumption keeps the dimensionality of the problem fixed even as the observation set grows. It also allows us to consider certain nonstationary signals.

We will consider the class of signals that can be represented as shown in Fig. 10.16(a):

$$Z_n = a_{n-1}Z_{n-1} + W_{n-1} \qquad n = 1, 2, \ldots, \tag{10.102}$$

where Z_0 is the random variable at time 0, a_n is a known sequence of constants, and W_n is a sequence of zero-mean uncorrelated random variables with possibly time-varying variances $\{E[W_n^2]\}$. The resulting process Z_n is nonstationary in general. We assume that the process Z_n is not available to us, and that instead, as shown in Fig. 10.16(a), we observe

$$X_n = Z_n + N_n \qquad n = 0, 1, 2, \ldots, \tag{10.103}$$

where the observation noise N_n is a zero-mean, uncorrelated sequence of random variables with possibly time-varying variances $\{E[N_n^2]\}$. We assume that W_n and N_n are uncorrelated at all times n_1 and n_2. In the special case where W_n and N_n are Gaussian random processes, then Z_n and X_n will also be Gaussian random processes. We will develop the Kalman filter, which has the structure in Fig. 10.16(b).

Our objective is to find *for each time n* the minimum mean square estimate (actually prediction) of Z_n based on the observations $X_0, X_1, \ldots, X_{n-1}$ using a linear estimator that possibly varies with time:

$$Y_n = \sum_{j=i}^{n} h_j^{(n-1)} X_{n-j}. \tag{10.104}$$

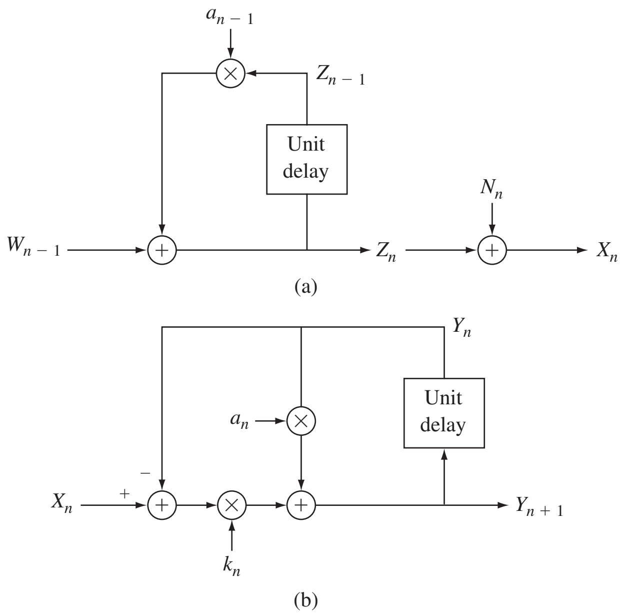

FIGURE 10.16
(a) Signal structure. (b) Kalman filter.

The orthogonality principle implies that the optimum filter $\{h_j^{(n-1)}\}$ satisfies

$$E\left[\left(Z_n - \sum_{j=1}^{n} h_j^{(n-1)} X_{n-j}\right) X_l\right] = 0 \qquad \text{for } l = 0, 1, \ldots, n-1,$$

which leads to a set of n equations in n unknowns:

$$R_{Z,X}(n, l) = \sum_{j=1}^{n} h_j^{(n-1)} R_X(n - j, l) \qquad \text{for } l = 0, 1, \ldots, n-1. \tag{10.105}$$

At the next time instant, we need to find

$$Y_{n+1} = \sum_{j=1}^{n+1} h_j^{(n)} X_{n+1-j} \tag{10.106}$$

by solving a system of $(n + 1) \times (n + 1)$ equations:

$$R_{Z,X}(n + 1, l) = \sum_{j=1}^{n+1} h_j^{(n)} R_X(n + 1 - j, l) \qquad \text{for } l = 0, 1, \ldots, n. \tag{10.107}$$

Up to this point we have followed the procedure of the previous section and we find that the dimensionality of the problem grows with the number of observations. We now use the signal structure to develop a recursive method for solving Eq. (10.106).

We first need the following two results: For $l < n$, we have

$$R_{Z,X}(n+1,l) = E[Z_{n+1}X_l] = E[(a_nZ_n + W_n)X_l]$$
$$= a_nR_{Z,X}(n,l) + E[W_nX_l] = a_nR_{Z,X}(n,l), \qquad (10.108)$$

since $E[W_nX_l] = E[W_n]E[X_l] = 0$, that is, W_n is uncorrelated with the past of the process and the observations prior to time n, as can be seen from Fig. 10.16(a). Also for $l < n$, we have

$$R_{Z,X}(n,l) = E[Z_nX_l] = E[(X_n - N_n)X_l]$$
$$= R_X(n,l) - E[N_nX_l] = R_X(n,l), \qquad (10.109)$$

since $E[N_nX_l] = E[N_n]E[X_l] = 0$, that is, the observation noise at time n is uncorrelated with prior observations.

We now show that the set of equations in Eq. (10.107) can be related to the set in Eq. (10.105). For $l < n$, we can equate the right-hand sides of Eqs. (10.108) and (10.107):

$$a_nR_{Z,X}(n,l) = \sum_{j=1}^{n+1} h_j^{(n)} R_X(n+1-j,l)$$
$$= h_1^{(n)}R_X(n,l) + \sum_{j=2}^{n+1} h_j^{(n)} R_X(n+1-j,l)$$
$$\text{for } l = 0, 1, \ldots, n-1. \qquad (10.110)$$

From Eq. (10.109) we have $R_X(n,l) = R_{Z,X}(n,l)$, so we can replace the first term on the right-hand of Eq. (10.110) and then move the resulting term to the left-hand side:

$$(a_n - h_1^{(n)})R_{Z,X}(n,l) = \sum_{j=2}^{n+1} h_j^{(n)} R_X(n+1-j,l)$$
$$= \sum_{j'=1}^{n} h_{j'+1}^{(n)} R_X(n-j',l). \qquad (10.111)$$

By dividing both sides by $a_n - h_1^{(n)}$ we finally obtain

$$R_{Z,X}(n,l) = \sum_{j'=1}^{n} \frac{h_{j'+1}^{(n)}}{a_n - h_1^{(n)}} R_X(n-j',l)$$
$$\text{for } l = 0, 1, \ldots, n-1. \qquad (10.112)$$

This set of equations is identical to Eq. (10.105) if we set

$$h_j^{(n-1)} = \frac{h_{j+1}^{(n)}}{a_n - h_1^{(n)}} \qquad \text{for } j = 1, \ldots, n. \qquad (10.113a)$$

Therefore, if at step n we have found $h_1^{(n-1)}, \ldots, h_n^{(n-1)}$, and if somehow we have found $h_1^{(n)}$, then we can find the remaining coefficients from

$$h_{j+1}^{(n)} = (a_n - h_1^{(n)})h_j^{(n-1)} \qquad j = 1, \ldots, n. \qquad (10.113b)$$

Thus the key question is how to find $h_1^{(n)}$.

Suppose we substitute the coefficients in Eq. (10.113b) into Eq. (10.106):

$$\begin{aligned}
Y_{n+1} &= h_1^{(n)} X_n + \sum_{j'=1}^{n} (a_n - h_1^{(n)}) h_{j'}^{(n-1)} X_{n-j'} \\
&= h_1^{(n)} X_n + (a_n - h_1^{(n)}) Y_n \\
&= a_n Y_n + h_1^{(n)} (X_n - Y_n),
\end{aligned} \tag{10.114}$$

where the second equality follows from Eq. (10.104). The above equation has a very pleasing interpretation, as shown in Fig. 10.16(b). Since Y_n is the prediction for time n, a_nY_n is the prediction for the next time instant, $n + 1$, based on the "old" information (see Eq. (10.102)). The term $(X_n - Y_n)$ is called the "innovations," and it gives the discrepancy between the old prediction and the observation. Finally, the term $h_1^{(n)}$ is called the *gain*, henceforth denoted by k_n, and it indicates the extent to which the innovations should be used to correct a_nY_n to obtain the "new" prediction Y_{n+1}. If we denote the **innovations** by

$$I_n = X_n - Y_n \tag{10.115}$$

then Eq. (10.114) becomes

$$Y_{n+1} = a_n Y_n + k_n I_n. \tag{10.116}$$

We still need to determine a means for computing the gain k_n.

From Eq. (10.115), we have that the innovations satisfy

$$I_n = X_n - Y_n = Z_n + N_n - Y_n = Z_n - Y_n + N_n = \varepsilon_n + N_n,$$

where $\varepsilon_n = Z_n - Y_n$ is the prediction error. A recursive equation can be obtained for the prediction error:

$$\begin{aligned}
\varepsilon_{n+1} &= Z_{n+1} - Y_{n+1} = a_n Z_n + W_n - a_n Y_n - k_n I_n \\
&= a_n (Z_n - Y_n) + W_n - k_n(\varepsilon_n + N_n) \\
&= (a_n - k_n)\varepsilon_n + W_n - k_n N_n,
\end{aligned} \tag{10.117}$$

with initial condition $\varepsilon_0 = Z_0$. Since X_0, W_n, and N_n are zero-mean, it then follows that $E[\varepsilon_n] = 0$ for all n. A recursive equation for the mean square prediction error is obtained from Eq. (10.117):

$$E[\varepsilon_{n+1}^2] = (a_n - k_n)^2 E[\varepsilon_n^2] + E[W_n^2] + k_n^2 E[N_n^2], \tag{10.118}$$

with initial condition $E[\varepsilon_0^2] = E[Z_0^2]$. We are finally ready to obtain an expression for the gain k_n.

The gain k_n must minimize the mean square error $E[\varepsilon_{n+1}^2]$. Therefore we can differentiate Eq. (10.118) with respect to k_n and set it equal to zero:

$$0 = -2(a_n - k_n)E[\varepsilon_n^2] + 2k_n E[N_n^2].$$

Then we can solve for k_n:

$$k_n = \frac{a_n E[\varepsilon_n^2]}{E[\varepsilon_n^2] + E[N_n^2]}. \tag{10.119}$$

The expression for the mean square prediction error in Eq. (10.118) can be simplified by using Eq. (10.119) (see Problem 10.72):

$$E[\varepsilon_{n+1}^2] = a_n(a_n - k_n)E[\varepsilon_n^2] + E[W_n^2]. \tag{10.120}$$

Equations (10.119), (10.116), and (10.120) when combined yield the recursive procedure that constitutes the Kalman filtering algorithm:

Kalman filter algorithm:[9]

Initialization: $Y_0 = 0 \qquad E[\varepsilon_0^2] = E[Z_0^2]$
For $n = 0, 1, 2, \ldots$

$$k_n = \frac{a_n E[\varepsilon_n^2]}{E[\varepsilon_n^2] + E[N_n^2]}$$

$$Y_{n+1} = a_n Y_n + k_n(X_n - Y_n)$$

$$E[\varepsilon_{n+1}^2] = a_n(a_n - k_n)E[\varepsilon_n^2] + E[W_n^2].$$

Note that the algorithm requires knowledge of the signal structure, i.e., the a_n, and the variances $E[N_n^2]$ and $E[W_n^2]$. The algorithm can be implemented easily and has consequently found application in a broad range of detection, estimation, and signal processing problems. The algorithm can be extended in matrix form to accommodate a broader range of processes.

Example 10.26 First-Order Autoregressive Process

Consider a signal defined by

$$Z_n = aZ_{n-1} + W_n \qquad n = 1, 2, \ldots \qquad Z_0 = 0,$$

where $E[W_n^2] = \sigma_W^2 = 0.36$, and $a = 0.8$, and suppose the observations are made in additive white noise

$$X_n = Z_n + N_n \qquad n = 0, 1, 2, \ldots,$$

where $E[N_n^2] = 1$. Find the form of the predictor and its mean square error as $n \to \infty$.

The gain at step n is given by

$$k_n = \frac{aE[\varepsilon_n^2]}{E[\varepsilon_n^2] + 1}.$$

The mean square error sequence is therefore given by

$$E[\varepsilon_0^2] = E[Z_0^2] = 0$$

[9] We caution the student that there are two common ways of defining the gain. The statement of the Kalman filter algorithm will differ accordingly in various textbooks.

$$E[\varepsilon_{n+1}^2] = a(a - k_n)E[\varepsilon_n^2] + \sigma_W^2$$
$$= a\left(\frac{a}{1 + E[\varepsilon_n^2]}\right)E[\varepsilon_n^2] + \sigma_W^2 \qquad \text{for } n = 1, 2, \ldots.$$

The steady state mean square error e_∞ must satisfy

$$e_\infty = \frac{a^2}{1 + e_\infty}e_\infty + \sigma_W^2.$$

For $a = 0.8$ and $\sigma_W^2 = 0.36$, the resulting quadratic equation yields $k_\infty = 0.3$ and $e_\infty = 0.6$. Thus at steady state the predictor is

$$Y_{n+1} = 0.8Y_n + 0.3(X_n - Y_n).$$

*10.6 ESTIMATING THE POWER SPECTRAL DENSITY

Let $X_0, \ldots, X_{k-1}$ be k observations of the discrete-time, zero-mean, wide-sense stationary process X_n. The periodogram estimate for $S_X(f)$ is defined as

$$\widetilde{p}_k(f) = \frac{1}{k}|\widetilde{x}_k(f)|^2, \tag{10.121}$$

where $\widetilde{x}_k(f)$ is obtained as a Fourier transform of the observation sequence:

$$\widetilde{x}_k(f) = \sum_{m=0}^{k-1} X_m e^{-j2\pi fm}. \tag{10.122}$$

In Section 10.1 we showed that the expected value of the periodogram estimate is

$$E[\widetilde{p}_k(f)] = \sum_{m'=-(k-1)}^{k-1}\left\{1 - \frac{|m'|}{k}\right\}R_X(m')e^{-j2\pi fm'}, \tag{10.123}$$

so $\widetilde{p}_k(f)$ is a biased estimator for $S_X(f)$. However, as $k \to \infty$,

$$E[\widetilde{p}_k(f)] \to S_X(f), \tag{10.124}$$

so the mean of the periodogram estimate approaches $S_X(f)$.

Before proceeding to find the variance of the periodogram estimate, we note that the periodogram estimate is equivalent to taking the Fourier transform of an estimate for the autocorrelation sequence; that is,

$$\widetilde{p}_k(f) = \sum_{m=-(k-1)}^{k-1} \hat{r}_k(m)e^{-j2\pi fm}, \tag{10.125}$$

where the estimate for the autocorrelation is

$$\hat{r}_k(m) = \frac{1}{k}\sum_{n=0}^{k-|m|-1} X_n X_{n+m}. \tag{10.126}$$

(See Problem 10.77.)

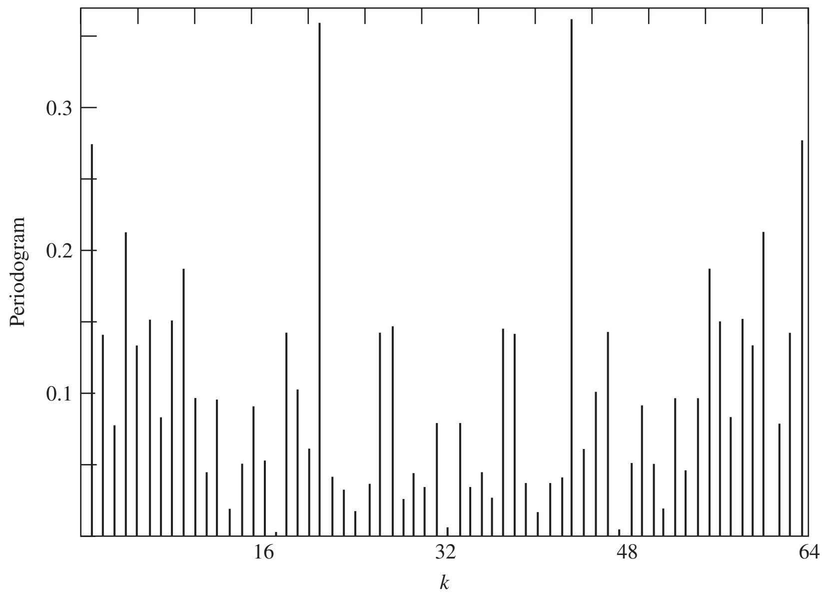

FIGURE 10.17
Periodogram for 64 samples of white noise sequence X_n iid uniform in (0, 1), $S_X(f) = \sigma_X^2 = 1/12 = 0.083$.

We might expect that as we increase the number of samples k, the periodogram estimate converges to $S_X(f)$. This does not happen. Instead we find that $\widetilde{p}_k(f)$ fluctuates wildly about the true spectral density, and that this random variation does not decrease with increased k (see Fig. 10.17). To see why this happens, in the next section we compute the statistics of the periodogram estimate for a white noise Gaussian random process. We find that the estimates given by the periodogram have a variance that does *not* approach zero as the number of samples is increased. This explains the lack of improvement in the estimate as k is increased. Furthermore, we show that the periodogram estimates are uncorrelated at uniformly spaced frequencies in the interval $-1/2 \le f < 1/2$. This explains the erratic appearance of the periodogram estimate as a function of f. In the final section, we obtain another estimate for $S_X(f)$ whose variance does approach zero as k increases.

10.6.1 Variance of Periodogram Estimate

Following the approach of [Jenkins and Watts, pp. 230–233], we consider the periodogram of samples of a white noise process with $S_X(f) = \sigma_X^2$ at the frequencies $f = n/k, -k/2 \le n < k/2$, which will cover the frequency range $-1/2 \le f < 1/2$. (In practice these are the frequencies we would evaluate if we were using the FFT algorithm to compute $\widetilde{x}_k(f)$.) First we rewrite Eq. (10.122) at $f = n/k$ as follows:

$$\begin{aligned}\widetilde{x}_k\left(\frac{n}{k}\right) &= \sum_{m=0}^{k-1} X_m\left(\cos\left(\frac{2\pi mn}{k}\right) - j\sin\left(\frac{2\pi mn}{k}\right)\right)\\ &= A_k(n) - jB_k(n) \qquad -k/2 \le n < k/2,\end{aligned} \tag{10.127}$$

where

$$A_k(n) = \sum_{m=0}^{k-1} X_m \cos\left(\frac{2\pi mn}{k}\right) \tag{10.128}$$

and

$$B_k(n) = \sum_{m=0}^{k-1} X_m \sin\left(\frac{2\pi mn}{k}\right). \tag{10.129}$$

Then it follows that the periodogram estimate is

$$\widetilde{p}_k\left(\frac{n}{k}\right) = \frac{1}{k}\left|\hat{x}_k\left(\frac{n}{k}\right)\right|^2 = \frac{1}{k}\{A_k^2(n) + B_k^2(n)\}. \tag{10.130}$$

We find the variance of $\widetilde{p}_k(n/k)$ from the statistics of $A_k(n)$ and $B_k(n)$.

The random variables $A_k(n)$ and $B_k(n)$ are defined as linear functions of the jointly Gaussian random variables $X_0, \dots, X_{k-1}$. Therefore $A_k(n)$ and $B_k(n)$ are also jointly Gaussian random variables. If we take the expected value of Eqs. (10.128) and (10.129) we find

$$E[A_k(n)] = 0 = E[B_k(n)] \qquad \text{for all } n. \tag{10.131}$$

Note also that the $n = -k/2$ and $n = 0$ terms are different in that

$$B_k(-k/2) = 0 = B_k(0) \tag{10.132a}$$

$$A_k(-k/2) = \sum_{i=0}^{k-1}(-1)^i X_i \qquad A_k(0) = \sum_{i=0}^{k-1} X_i. \tag{10.132b}$$

The correlation between $A_k(n)$ and $A_k(m)$ (for n, m not equal to $-k/2$ or 0) is

$$\begin{aligned}
E[A_k(n)A_k(m)] &= \sum_{i=0}^{k-1}\sum_{l=0}^{k-1} E[X_i X_l]\cos\left(\frac{2\pi ni}{k}\right)\cos\left(\frac{2\pi ml}{k}\right) \\
&= \sigma_X^2 \sum_{i=0}^{k-1} \cos\left(\frac{2\pi ni}{k}\right)\cos\left(\frac{2\pi mi}{k}\right) \\
&= \sigma_X^2 \sum_{i=0}^{k-1} \frac{1}{2}\cos\left(\frac{2\pi(n-m)i}{k}\right) + \sigma_X^2 \sum_{i=0}^{k-1} \frac{1}{2}\cos\left(\frac{2\pi(n+m)i}{k}\right),
\end{aligned}$$

where we used the fact that $E[X_i X_1] = \sigma_X^2 \delta_{il}$ since the noise is white. The second summation is equal to zero, and the first summation is zero except when $n = m$. Thus

$$E[A_k(n)A_k(m)] = \frac{1}{2}k\sigma_X^2\,\delta_{nm} \qquad \text{for all } n, m \neq -k/2, 0. \tag{10.133a}$$

It can similarly be shown that

$$E[B_k(n)B_k(m)] = \frac{1}{2}k\sigma_X^2\,\delta_{nm} \qquad n, m \neq 0 - k/2, 0 \tag{10.133b}$$

$$E[A_k(n)B_k(m)] = 0 \qquad \text{for all } n, m. \tag{10.133c}$$

When $n = -k/2$ or 0, we have

$$E[A_k(n)A_k(m)] = k\sigma_X^2\,\delta_{nm} \qquad \text{for all } m. \tag{10.133d}$$

Equations (10.133a) through (10.133d) imply that $A_k(n)$ and $B_k(m)$ are uncorrelated random variables. Since $A_k(n)$ and $B_k(n)$ are jointly Gaussian random variables, this implies that they are zero-mean, *independent* Gaussian random variables.

We are now ready to find the statistics of the periodogram estimates at the frequencies $f = n/k$. Equation (10.130) gives

$$\begin{aligned}\widetilde{p}_k\left(\frac{n}{k}\right) &= \frac{1}{k}\{A_k^2(n) + B_k^2(n)\} \qquad n \neq -k/2, 0 \\ &= \frac{1}{2}\sigma_X^2\left\{\frac{A_k^2(n)}{(1/2)k\sigma_X^2} + \frac{B_k^2(n)}{(1/2)k\sigma_X^2}\right\}.\end{aligned} \tag{10.134}$$

The quantity in brackets is the sum of the squares of two zero-mean, unit-variance, independent Gaussian random variables. This is a chi-square random variable with two degrees of freedom (see Problem 7.6). From Table 4.1, we see that a chi-square random variable with v degrees of freedom has variance $2v$. Thus the expression in the brackets has variance 4, and the periodogram estimate $\hat{p}_k(n/k)$ has variance

$$\text{VAR}\left[\widetilde{p}_k\left(\frac{n}{k}\right)\right] = \left(\frac{1}{2}\sigma_X^2\right)^2 4 = \sigma_X^4 = S_X(f)^2. \tag{10.135a}$$

For $n = -k/2$ and $n = 0$,

$$\widetilde{p}_k\left(\frac{n}{k}\right) = \sigma_X^2\left\{\frac{A_k^2(n)}{k\sigma_X^2}\right\}.$$

The quantity in brackets is a chi-square random variable with one degree of freedom and variance 2, so the variance of the periodogram estimate is

$$\text{VAR}\left[\widetilde{p}_k\left(\frac{n}{k}\right)\right] = 2\sigma_X^4 \qquad n = -k/2, 0. \tag{10.135b}$$

Thus we conclude from Eqs. (10.135a) and (10.135b) that *the variance of the periodogram estimate is proportional to the square of the power spectral density and does not approach zero as k increases*. In addition, Eqs. (10.133a) through (10.133d) imply that *the periodogram estimates at the frequencies $f = -n/k$ are uncorrelated random variables*. A more detailed analysis [Jenkins and Watts, p. 238] shows that for arbitrary f,

$$\text{VAR}[\widetilde{p}_k(f)] = S_X(f)^2\left\{1 + \left(\frac{\sin(2\pi fk)}{k\sin(2\pi f)}\right)^2\right\}. \tag{10.136}$$

Thus variance of the periodogram estimate does not approach zero as the number of samples is increased.

The above discussion has only considered the spectrum estimation for a white noise, Gaussian random process, but the general conclusions are also valid for non-white, non-Gaussian processes. If the X_i are not Gaussian, we note from Eqs. (10.128) and (10.129) that A_k and B_k are approximately Gaussian by the central limit theorem if k is large. Thus the periodogram estimate is then approximately a chi-square random variable.

If the process X_i is not white, then it can be viewed as filtered white noise:

$$X_n = h_n * W_n,$$

where $S_W(f) = \sigma_W^2$ and $|H(f)|^2 S_W(f) = S_X(f)$. The periodograms of X_n and W_n are related by

$$\frac{1}{k}\left|\widetilde{x}_k\left(\frac{n}{k}\right)\right|^2 = \frac{1}{k}\left|H\left(\frac{n}{k}\right)\right|^2\left|\widetilde{w}_k\left(\frac{n}{k}\right)\right|^2. \tag{10.137}$$

Thus

$$\left|\widetilde{w}_k\left(\frac{n}{k}\right)\right|^2 = \frac{|\widetilde{x}_k(n/k)|^2}{|H(n/k)|^2}. \tag{10.138}$$

From our previous results, we know that $|\widetilde{w}_k(n/k)|^2/k$ is a chi-square random variable with variance σ_W^4. This implies that

$$\text{VAR}\left[\frac{|\widetilde{x}_k(n/k)|^2}{k}\right] = \left|H\left(\frac{n}{k}\right)\right|^4 \sigma_W^4 = S_X(f)^2. \tag{10.139}$$

Thus we conclude that the variance of the periodogram estimate for nonwhite noise is also proportional to $S_X(f)^2$.

10.6.2 Smoothing of Periodogram Estimate

A fundamental result in probability theory is that the sample mean of a sequence of *independent* realizations of a random variable approaches the true mean with probability one. We obtain an estimate for $S_X(f)$ that goes to zero with the number of observations k by taking the average of N *independent* periodograms on samples of size k:

$$\langle \widetilde{p}_k(f)\rangle_N = \frac{1}{N}\sum_{i=1}^{N}\widetilde{p}_{k,i}(f), \tag{10.140}$$

where $\{\widetilde{p}_{k,i}(f)\}$ are N independent periodograms computed using separate sets of k samples each. Figures 10.18 and 10.19 show the $N = 10$ and $N = 50$ smoothed periodograms corresponding to the unsmoothed periodogram of Fig. 10.17. It is evident that the variance of the power spectrum estimates is decreasing with N.

The mean of the smoothed estimator is

$$\begin{aligned} E\langle \widetilde{p}_k(f)\rangle_N &= \frac{1}{N}\sum_{i=1}^{N}E[\widetilde{p}_{k,i}(f)] = E[\widetilde{p}_k(f)] \\ &= \sum_{m'=-(k-1)}^{k-1}\left\{1 - \frac{|m'|}{k}\right\}R_X(m')e^{-j2\pi f m'}, \end{aligned} \tag{10.141}$$

where we have used Eq. (10.35). Thus the smoothed estimator has the same mean as the periodogram estimate on a sample of size k.

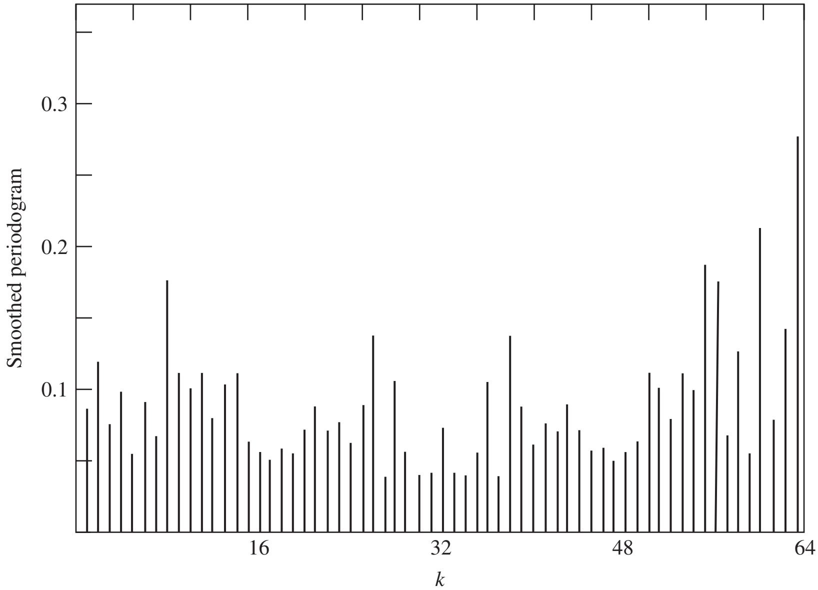

FIGURE 10.18
Sixty-four-point smoothed periodogram with $N = 10$, X_n iid uniform in (0, 1), $S_X(f) = 1/12 = 0.083$.

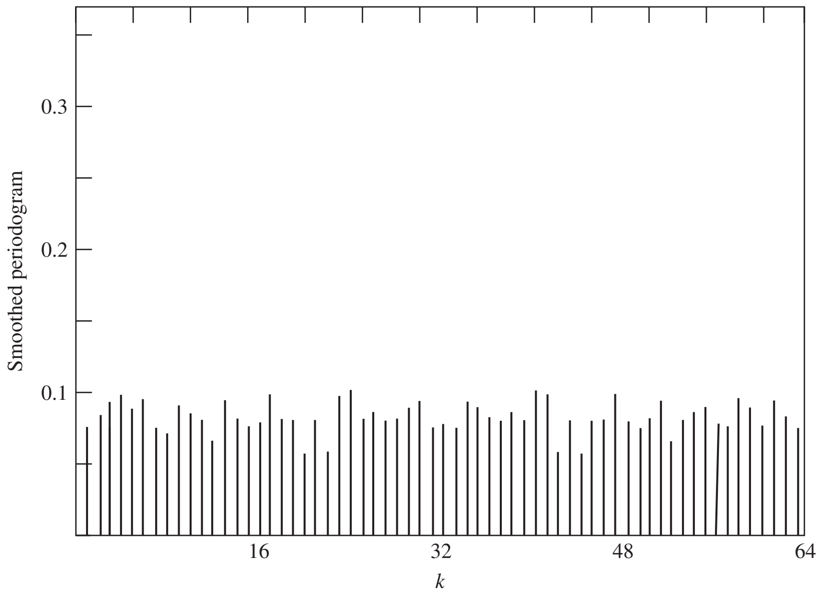

FIGURE 10.19
Sixty-four-point smoothed periodogram with $N = 50$, X_n iid uniform in (0, 1), $S_X(f) = 1/12 = 0.083$.

The variance of the smoothed estimator is

$$\begin{aligned}
\text{VAR}[\langle \widetilde{p}_k(f)\rangle_N] &= \frac{1}{N^2}\sum_{i=1}^{N}\text{VAR}[\widetilde{p}_{k,i}(f)] \\
&= \frac{1}{N}\text{VAR}[\widetilde{p}_k(f)] \\
&\simeq \frac{1}{N}S_X(f)^2.
\end{aligned}$$

Thus the variance of the smoothed estimator can be reduced by increasing N, the number of periodograms used in Eq. (10.140).

In practice, a sample set of size Nk, $X_0, \ldots, X_{Nk-1}$ is divided into N blocks and a separate periodogram is computed for each block. The smoothed estimate is then the average over the N periodograms. This method is called **Bartlett's smoothing procedure**. Note that, in general, the resulting periodograms are not independent because the underlying blocks are not independent. Thus this smoothing procedure must be viewed as an approximation to the computation and averaging of independent periodograms.

The choice of k and N is determined by the desired frequency resolution and variance of the estimate. The blocksize k determines the number of frequencies for which the spectral density is computed (i.e., the frequency resolution). The variance of the estimate is controlled by the number of periodograms N. The actual choice of k and N depends on the nature of the signal being investigated.

10.7 NUMERICAL TECHNIQUES FOR PROCESSING RANDOM SIGNALS

In this chapter our discussion has combined notions from random processes with basic concepts from signal processing. The processing of signals is a very important area in modern technology and a rich set of techniques and methodologies have been developed to address the needs of specific application areas such as communication systems, speech compression, speech recognition, video compression, face recognition, network and service traffic engineering, etc. In this section we briefly present a number of general tools available for the processing of random signals. We focus on the tools provided in Octave since these are quite useful as well as readily available.

10.7.1 FFT Techniques

The Fourier transform relationship between $R_X(\tau)$ and $S_X(f)$ is fundamental in the study of wide-sense stationary processes and plays a key role in random signal analysis. The fast fourier transform (FFT) methods we developed in Section 7.6 can be applied to the numerical transformation from autocorrelation functions to power spectral densities and back.

Consider the computation of $R_X(\tau)$ and $S_X(f)$ for continuous-time processes:

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f)e^{-j2\pi f\tau}\,df \approx \int_{-W}^{W} S_X(f)e^{-j2\pi f\tau}\,df.$$

First we limit the integral to the region where $S_X(f)$ has significant power. Next we restrict our attention to a discrete set of $N = 2M$ frequency values at kf_0 so that $-W = -Mf_0 < (-M + 1)f_0 < \cdots < (M - 1)f_0 < W$, and then approximate the integral by a sum:

$$R_X(\tau) \approx \sum_{m=-M}^{M-1} S_X(mf_0)e^{-j2\pi mf_0\tau}f_0.$$

Finally, we also focus on a set of discrete lag values: kt_0 so that $-T = -Mt_0 < (-M + 1\}$ $t_0 < \cdots < (M - 1)t_0 < T$. We obtain the DFT as follows:

$$R_X(kt_0) \approx f_0 \sum_{m=-M}^{M-1} S_X(mf_0)e^{-j2\pi mkt_0f_0} = f_0 \sum_{m=-M}^{M-1} S_X(mf_0)e^{-j2\pi mk/N}. \quad (10.142)$$

In order to have a discrete Fourier transform, *we must have* $t_0f_0 = 1/N$, which is equivalent to: $t_0 = 1/Nf_0$ and $T = Mt_0 = 1/2f_0$ and $W = Mf_0 = 1/2t_0$. We can use the FFT function introduced in Section 7.6 to perform the transformation in Eq. (10.142) to obtain the set of values $\{R_X(kt_0), k \in [-M, M - 1]\}$ from $\{S_X(mt_0), k \in [-M,M - 1]\}$. The transformation in the reverse direction is done in the same way. Since $R_X(\tau)$ and $S_X(f)$ are even functions various simplifications are possible. We discuss some of these in the problems.

Consider the computation of $S_X(f)$ and $R_X(k)$ for discrete-time processes. $S_X(f)$ spans the range of frequencies $|f| < 1/2$, so we restrict attention to N points $1/N$ apart:

$$S_X\left(\frac{m}{N}\right) = \sum_{k=-\infty}^{\infty} R_X(k)e^{-j2\pi kf}\bigg|_{f=m/N} \approx \sum_{k=-M}^{M-1} R_X(k)e^{-j2\pi km/N}. \quad (10.143)$$

The approximation here involves neglecting autocorrelation terms outside $[-M, M - 1\}$. Since $df \approx 1/N$, the transformation in the reverse direction is scaled differently:

$$R_X(k) = \int_{-1/2}^{1/2} S_X(f)e^{-j2\pi kf}\,df \approx \frac{1}{N}\sum_{k=-M}^{M-1} S_X\left(\frac{m}{N}\right)e^{-j2\pi km/N}. \quad (10.144)$$

We assume that the student has already tried the FFT exercises in Section 7.6, so we leave examples in the use of the FFT to the Problems.

The various frequency domain results for linear systems that relate input, output, and cross-spectral densities can be evaluated numerically using the FFT.

Example 10.27 Output Autocorrelation and Cross-Correlation

Consider Example 10.12, where a random telegraph signal $X(t)$ with $\alpha = 1$ is passed through a lowpass filter with $\beta = 1$ and $\beta = 10$. Find $R_Y(\tau)$.

The random telegraph has $S_X(f) = \alpha/(\alpha^2 + \pi^2 f^2)$ and the filter has transfer function $H(f) = \beta/(\beta + j2\pi f)$, so $R_Y(\tau)$ is given by:

$$R_Y(\tau) = \mathscr{F}^{-1}\left\{|H(f)|^2\, S_X(f)\right\} = \int_{-\infty}^{\infty} \frac{\beta^2}{\beta^2 + 4\pi^2 f^2}\frac{\alpha^2}{\alpha^2 + 4\pi^2 f^2}\,df.$$

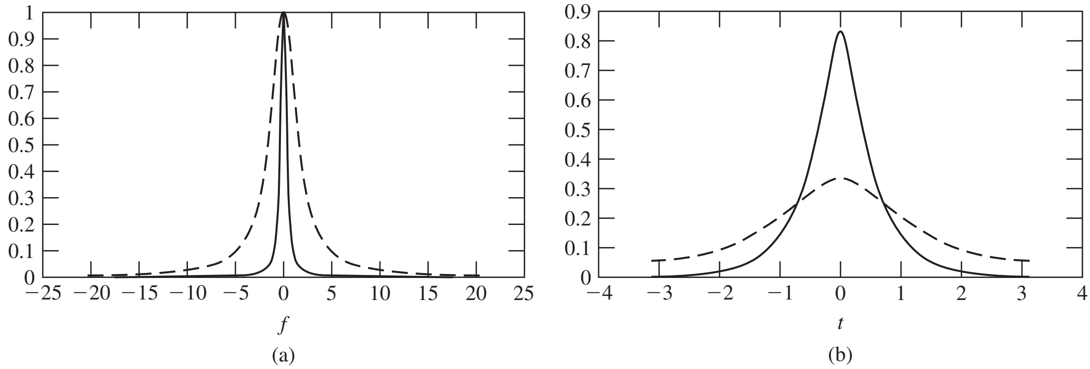

FIGURE 10.20
(a) Transfer function and input power spectral density; (b) Autocorrelation of filtered random telegraph with filter $\beta = 10$.

We used an $N = 256$ FFT to evaluate autocorrelation functions numerically for $\alpha = 1$ and $\beta = 1$ and $\beta = 10$. Figure 10.20(a) shows $|H(f)|^2$ and $S_X(f)$ for $\beta = 10$. It can be seen that the transfer function (the dashed line) is close to 1 in the region of f where $S_X(f)$ has most of its power. Consequently we expect the output for $\beta = 10$ to have an autocorrelation similar to that of the input. For $\beta = 1$, on the other hand, the filter will attenuate more of the significant frequencies of $X(t)$ and we expect more change in the output autocorrelation. Figure 10.20(b) shows the output autocorrelation and we see that indeed for $\beta = 10$ (the solid line), $R_Y(\tau)$ is close to the double-sided exponential of $R_X(\tau)$. For $\beta = 1$ the output autocorrelation differs significantly from $R_X(\tau)$.

10.7.2 Filtering Techniques

The autocorrelation and power spectral density functions provide us with information about the average behavior of the processes. We are also interested in obtaining sample functions of the inputs and outputs of systems. For linear systems the principal tools for signal processing are the convolution and Fourier transform.

Convolution in discrete-time (Eq. (10.48)) is quite simple and so convolution is the workhorse in linear signal processing. Octave provides several functions for performing convolutions with discrete-time signals. In Example 10.15 we encountered the function `filter(b,a,x)` which implements filtering of the sequence x with an ARMA filter with coefficients specified by vectors b and a in the following equation.

$$Y_n = -\sum_{i=1}^{q} \alpha_i Y_{n-i} + \sum_{j=0}^{p} \beta_j X_{n-j}.$$

Other functions use `filter(b,a,x)` to provide special cases of filtering. For example, `conv(a,b)` convolves the elements in the vectors a and b. We can obtain the output of a linear system by letting a be the impulse response and b the input random sequence. The moving average example in Fig. 10.7(b) is easily obtained using this `conv`. Octave provides other functions implementing specific digital filters.

We can also obtain the output of a linear system in the frequency domain. We take the FFT of the input sequence X_n and we then multiply it by the FFT of the transfer function. The inverse FFT will then provide Y_n of the linear system. The Octave function `fftconv(a,b,n)` implements this approach. The size of the FFT must be equal to the total number of samples in the input sequence, so this approach is not advisable for long input sequences.

10.7.3 Generation of Random Processes

Finally, we are interested in obtaining discrete-time and continuous-time sample functions of the inputs and outputs of systems. Previous chapters provide us with several tools for the generation of random signals that can act as inputs to the systems of interest.

Section 5.10 provides the method for generating independent pairs of Gaussian random variables. This method forms the basis for the generation of iid Gaussian sequences and is implemented in `normal_rnd=(M,V,Sz)`. The generation of sequences of WSS but correlated sequences of Gaussian random variables requires more work. One approach is to use the matrix approaches developed in Section 6.6 to generate individual vectors with a specified covariance matrix. To generate a vector $\mathbf{Y}$ of n outcomes with covariance $\mathbf{K_Y}$, we perform the following factorization:

$$\mathbf{K_Y} = \mathbf{A}^{\mathrm{T}}\,\mathbf{A}\;\mathbf{P}\,\boldsymbol{\lambda}\,\mathbf{P}\,\mathbf{T},$$

and we generate the vector

$$\mathbf{Y} = \mathbf{A}^{\mathrm{T}}\,\mathbf{X}$$

where $\mathbf{X}$ is vector of iid zero-mean, unit-variance Gaussian random variables. The Octave function `svd(B)` performs a singular value decomposition of the matrix B, see [Long]. When $B = \mathbf{K_Y}$ is a covariance matrix, `svd` returns the diagonal matrix $\mathbf{D}$ of eigenvalues of $\mathbf{K_Y}$ as well as the matrices $\mathbf{U} = \mathbf{P}$ and $\mathbf{V} = \mathbf{P}^{\mathrm{T}}$.

Example 10.28 Generation of Correlated Gaussian Random Variables

Generate 256 samples of the autoregressive process in Example 10.14 with $\alpha = -0.5, \sigma_X = 1$.

The autocorrelation of the process is given by $R_X(k) = (-1/2)^{|k|}$. We generate a vector r of the first 256 lags of $R_X(k)$ and use the function `toeplitz(r)` to generate the covariance matrix. We then call the `svd` to obtain A. Finally we produce the output vector $\mathbf{Y} = \mathbf{A}^{\mathrm{T}}\,\mathbf{X}$.

```
> n=[0:255]
> r=(−0.5).^n;
> K=toeplitz(r);
> [U,D,V]=svd(K);
> X=normal_rnd(0,1,1,256);
> y=V*(D^0.5)*transpose(X);
> plot(y)
```

Figure 10.21(a) shows a plot of $\mathbf{Y}$. To check that the sequence has the desired autocovariance we use the function `autocov(X,H)` which estimates the autocovariance function of the sequence X for the first H lag values. Figure 10.21(b) shows that the sample correlation coefficient that is obtained by dividing the autocovariance by the sample variance. The plot shows the alternating covariance values and the expected peak values of -0.5 and 0.25 to the first two lags.

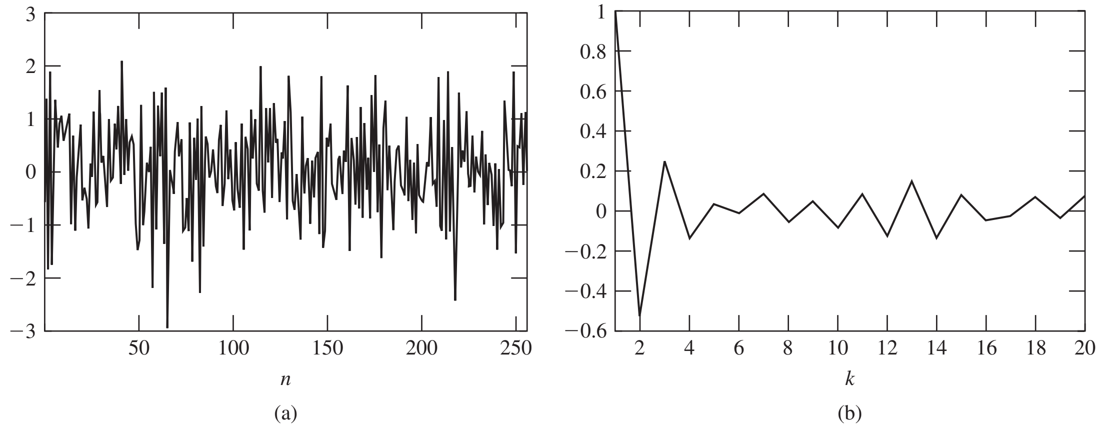

FIGURE 10.21
(a) Correlated Gaussian noise (b) Sample autocovariance.

An alternative approach to generating a correlated sequence of random variables with a specified covariance function is to input an uncorrelated sequence into a linear filter with a specific $H(f)$. Equation (10.46) allows us to determine the power spectral density of the output sequence. This approach can be implemented using convolution and is applicable to extremely long signal sequences. A large choice of possible filter functions is available for both continuous-time and discrete-time systems. For example, the ARMA model in Example 10.15 is capable of implementing a broad range of transfer functions. Indeed the entire discussion in Section 10.4 was focused on obtaining the transfer function of optimal linear systems in various scenarios.

Example 10.29 Generation of White Gaussian Noise

Find a method for generating white Gaussian noise for a simulation of a continuous-time communications system.

The generation of discrete-time white Gaussian noise is trivial and involves the generation of a sequence of iid Gaussian random variables. The generation of continuous-time white Gaussian noise is not so simple. Recall from Example 10.3 that true white noise has infinite bandwidth and hence infinite power and so is impossible to realize. Real systems however are bandlimited, and hence we always end up dealing with bandlimited white noise. If the system of interest is bandlimited to W Hertz, then we need to model white noise limited to W Hz. In Example 10.3 we found this type of noise has autocorrelation:

$$R_X(\tau) = \frac{N_0 \sin(2\pi W\tau)}{2\pi\tau}.$$

The sampling theorem discussed in Section 10.3 allows us to represent bandlimited white Gaussian noise as follows:

$$\hat{X}(t) = \sum_{n=-\infty}^{\infty} X(nT)p(t - nT) \quad \text{where} \quad p(t) = \frac{\sin(\pi t/T)}{\pi t/T},$$

where $1/T = 2W$. The coefficients $X(nT)$ have autocorrelation $R_X(nT)$ which is given by:

$$R_X(nT) = \frac{N_0 \sin(2\pi W nT)}{2\pi nT} = \frac{N_0 \sin(2\pi W n/2W)}{2\pi n/2W}$$

$$= \frac{N_0 W \sin(\pi n)}{\pi n} = \begin{cases} N_0 W & \text{for} \quad n = 0 \\ 0 & \text{for} \quad n \neq 0. \end{cases}$$

We thus conclude that $X(nT)$ is an iid sequence of Gaussian random variables with variance $N_0 W$. Therefore we can simulate sampled bandlimited white Gaussian noise by generating a sequence $X(nT)$. We can perform any processing required in the discrete-time domain, and we can then apply the result to an interpolator to recover the continuous-time output.

SUMMARY

- The power spectral density of a WSS process is the Fourier transform of its autocorrelation function. The power spectral density of a real-valued random process is a real-valued, nonnegative, even function of frequency.
- The output of a linear, time-invariant system is a WSS random process if its input is a WSS random process that is applied an infinite time in the past.
- The output of a linear, time-invariant system is a Gaussian WSS random process if its input is a Gaussian WSS random process.
- Wide-sense stationary random processes with arbitrary rational power spectral density can be generated by filtering white noise.
- The sampling theorem allows the representation of bandlimited continuous-time processes by the sequence of periodic samples of the process.
- The orthogonality condition can be used to obtain equations for linear systems that minimize mean square error. These systems arise in filtering, smoothing, and prediction problems. Matrix numerical methods are used to find the optimum linear systems.
- The Kalman filter can be used to estimate signals with a structure that keeps the dimensionality of the algorithm fixed even as the size of the observation set increases.
- The variance of the periodogram estimate for the power spectral density does not approach zero as the number of samples is increased. An average of several independent periodograms is required to obtain an estimate whose variance does approach zero as the number of samples is increased.
- The FFT, convolution, and matrix techniques are basic tools for analyzing, simulating, and implementing processing of random signals.

CHECKLIST OF IMPORTANT TERMS

Amplitude modulation
ARMA process
Autoregressive process
Bandpass signal
Causal system
Cross-power spectral density
Einstein-Wiener-Khinchin theorem
Filtering
Impulse response
Innovations

Kalman filter
Linear system
Long-range dependence
Moving average process
Nyquist sampling rate
Optimum filter
Orthogonality condition
Periodogram
Power spectral density
Prediction
Quadrature amplitude modulation
Sampling theorem
Smoothed periodogram
Smoothing
System
Time-invariant system
Transfer function
Unit-sample response
White noise
Wiener filter
Wiener-Hopf equations
Yule-Walker equations

ANNOTATED REFERENCES

References [1] through [6] contain good discussions of the notion of power spectral density and of the response of linear systems to random inputs. References [6] and [7] give accessible introductions to the spectral factorization problem. References [7] through [9] discuss linear filtering and power spectrum estimation in the context of digital signal processing. Reference [10] discusses the basic theory underlying power spectrum estimation.

1. A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 2002.
2. H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 3d ed., Prentice Hall, Upper Saddle River, N.J., 2002.
3. R. M. Gray and L. D. Davisson, *Random Processes: A Mathematical Approach for Engineers*, Prentice Hall, Englewood Cliffs, N.J., 1986.
4. R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes*, Wiley, New York, 2005.
5. J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineering*, Cambridge University Press, Cambridge, 2006.
6. G. R. Cooper and C. D. MacGillem, *Probabilistic Methods of Signal and System Analysis*, Holt, Rinehart & Winston, New York, 1986.
7. J. A. Cadzow, *Foundations of Digital Signal Processing and Data Analysis*, Macmillan, New York, 1987.
8. A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs, N.J., 1989.
9. M. Kunt, *Digital Signal Processing*, Artech House, Dedham, Mass., 1986.
10. G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*, Holden Day, San Francisco, 1968.
11. A. Einstein, "Method for the Determination of the Statistical Values of Observations Concerning Quantities Subject to Irregular Observations," reprinted in *IEEE ASSP Magazine,* October 1987, p. 6.
12. P. J. G. Long, "Introduction to Octave," University of Cambridge, September, 2005, available online.

PROBLEMS

Section 10.1: Power Spectral Density

10.1. Let $g(x)$ denote the triangular function shown in Fig. P10.1.

(a) Find the power spectral density corresponding to $R_X(\tau) = g(\tau/T)$.

(b) Find the autocorrelation corresponding to the power spectral density $S_X(f) = g(f/W)$.

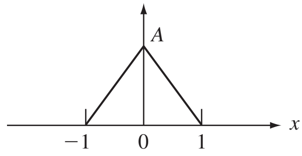

FIGURE P10.1

10.2. Let $p(x)$ be the rectangular function shown in Fig. P10.2. Is $R_X(\tau) = p(\tau/T)$ a valid autocorrelation function?

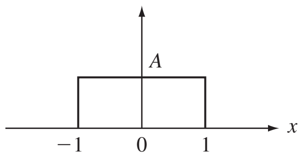

FIGURE P10.2

10.3. **(a)** Find the power spectral density $S_Y(f)$ of a random process with autocorrelation function $R_X(\tau)\cos(2\pi f_0 \tau)$, where $R_X(\tau)$ is itself an autocorrelation function.

(b) Plot $S_Y(f)$ if $R_X(\tau)$ is as in Problem 10.1a.

10.4. **(a)** Find the autocorrelation function corresponding to the power spectral density shown in Fig. P10.3.

(b) Find the total average power.

(c) Plot the power in the range $|f| > f_0$ as a function of $f_0 > 0$.

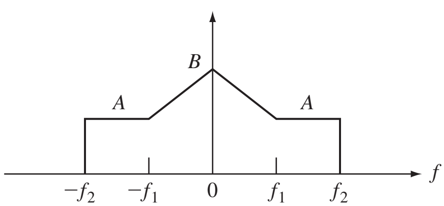

FIGURE P10.3

10.5. A random process $X(t)$ has autocorrelation given by $R_X(\tau) = \sigma_X^2 e^{-\tau^2/2\alpha^2}$, $\alpha > 0$.

(a) Find the corresponding power spectral density.

(b) Find the amount of power contained in the frequencies $|f| > k/2\pi\alpha$, where $k = 1, 2, 3$.

10.6. Let $Z(t) = X(t) + Y(t)$. Under what conditions does $S_Z(f) = S_X(f) + S_Y(f)$?

10.7. Show that

(a) $R_{X,Y}(\tau) = R_{Y,X}(-\tau)$.

(b) $S_{X,Y}(f) = S^*_{Y,X}(f)$.

10.8. Let $Y(t) = X(t) - X(t - d)$.

(a) Find $R_{X,Y}(\tau)$ and $S_{X,Y}(f)$.

(b) Find $R_Y(\tau)$ and $S_Y(f)$.

10.9. Do Problem 10.8 if $X(t)$ has the triangular autocorrelation function $g(\tau/T)$ in Problem 10.1 and Fig. P 10.1.

10.10. Let $X(t)$ and $Y(t)$ be independent wide-sense stationary random processes, and define $Z(t) = X(t)Y(t)$.

(a) Show that $Z(t)$ is wide-sense stationary.

(b) Find $R_Z(\tau)$ and $S_Z(f)$.

10.11. In Problem 10.10, let $X(t) = a\cos(2\pi f_0 t + \Theta)$ where Θ is a uniform random variable in $(0, 2\pi)$. Find $R_Z(\tau)$ and $S_Z(f)$.

10.12. Let $R_X(k) = 4\alpha^{|k|}$, $|\alpha| < 1$.

(a) Find $S_X(f)$.

(b) Plot $S_X(f)$ for $\alpha = 0.25$ and $\alpha = 0.75$, and comment on the effect of the value of α.

10.13. Let $R_X(k) = 4(\alpha)^{|k|} + 16(\beta)^{|k|}$, $\alpha < 1$, $\beta < 1$.

(a) Find $S_X(f)$.

(b) Plot $S_X(f)$ for $\alpha = \beta = 0.5$ and $\alpha = 0.75 = 3\beta$ and comment on the effect of value of α/β.

10.14. Let $R_X(k) = 9(1 - |k|/N)$, for $|k| < N$ and 0 elsewhere. Find and plot $S_X(f)$.

10.15. Let $X_n = \cos(2\pi f_0 n + \Theta)$, where Θ is a uniformly distributed random variable in the interval $(0, 2\pi)$. Find and plot $S_X(f)$ for $f_0 = 0.5, 1, 1.75, \pi$.

10.16. Let $D_n = X_n - X_{n-d}$, where d is an integer constant and X_n is a zero-mean, WSS random process.

(a) Find $R_D(k)$ and $S_D(f)$ in terms of $R_X(k)$ and $S_X(f)$. What is the impact of d?

(b) Find $E[D_n^2]$.

10.17. Find $R_D(k)$ and $S_D(f)$ in Problem 10.16 if X_n is the moving average process of Example 10.7 with $\alpha = 1$.

10.18. Let X_n be a zero-mean, bandlimited white noise random process with $S_X(f) = 1$ for $|f| < f_c$ and 0 elsewhere, where $f_c < 1/2$.

(a) Show that $R_X(k) = \sin(2\pi f_c k)/(\pi k)$.

(b) Find $R_X(k)$ when $f_c = 1/4$.

10.19. Let W_n be a zero-mean white noise sequence, and let X_n be independent of W_n.

(a) Show that $Y_n = W_n X_n$ is a white sequence, and find σ_Y^2.

(b) Suppose X_n is a Gaussian random process with autocorrelation $R_X(k) = (1/2)^{|k|}$. Specify the joint pmf's for Y_n.

10.20. Evaluate the periodogram estimate for the random process $X(t) = a\cos(2\pi f_0 t + \Theta)$, where Θ is a uniformly distributed random variable in the interval $(0, 2\pi)$. What happens as $T \to \infty$?

10.21. **(a)** Show how to use the FFT to calculate the periodogram estimate in Eq. (10.32).

(b) Generate four realizations of an iid zero-mean unit-variance Gaussian sequence of length 128. Calculate the periodogram.

(c) Calculate 50 periodograms as in part b and show the average of the periodograms after every 10 additional realizations.

Section 10.2: Response of Linear Systems to Random Signals

10.22. Let $X(t)$ be a differentiable WSS random process, and define

$$Y(t) = \frac{d}{dt}X(t).$$

Find an expression for $S_Y(f)$ and $R_Y(\tau)$. *Hint:* For this system, $H(f) = j2\pi f$.

10.23. Let $Y(t)$ be the derivative of $X(t)$, a bandlimited white noise process as in Example 10.3.

(a) Find $S_Y(f)$ and $R_Y(\tau)$.

(b) What is the average power of the output?

10.24. Repeat Problem 10.23 if $X(t)$ has $S_X(f) = \beta^2 e^{-\pi f^2}$.

10.25. Let $Y(t)$ be a short-term integration of $X(t)$:

$$Y(t) = \frac{1}{T}\int_{t-T}^{t} X(t')\,dt'.$$

(a) Find the impulse response $h(t)$ and the transfer function $H(f)$.

(b) Find $S_Y(f)$ in terms of $S_X(f)$.

10.26. In Problem 10.25, let $R_X(\tau) = (1 - |\tau|/T)$ for $|\tau| < T$ and zero elsewhere.

(a) Find $S_Y(f)$.

(b) Find $R_Y(\tau)$.

(c) Find $E[Y^2(t)]$.

10.27. The input into a filter is zero-mean white noise with noise power density $N_0/2$. The filter has transfer function

$$H(f) = \frac{1}{1 + j2\pi f}.$$

(a) Find $S_{Y,X}(f)$ and $R_{Y,X}(\tau)$.

(b) Find $S_Y(f)$ and $R_Y(\tau)$.

(c) What is the average power of the output?

10.28. A bandlimited white noise process $X(t)$ is input into a filter with transfer function $H(f) = 1 + j2\pi f$.

(a) Find $S_{Y,X}(f)$ and $R_{Y,X}(\tau)$ in terms of $R_X(\tau)$ and $S_X(f)$.

(b) Find $S_Y(f)$ and $R_Y(\tau)$ in terms of $R_X(\tau)$ and $S_X(f)$.

(c) What is the average power of the output?

10.29. **(a)** A WSS process $X(t)$ is applied to a linear system at $t = 0$. Find the mean and autocorrelation function of the output process. Show that the output process becomes WSS as $t \to \infty$.

10.30. Let $Y(t)$ be the output of a linear system with impulse response $h(t)$ and input $X(t)$. Find $R_{Y,X}(\tau)$ when the input is white noise. Explain how this result can be used to estimate the impulse response of a linear system.

10.31. **(a)** A WSS Gaussian random process $X(t)$ is applied to two linear systems as shown in Fig. P10.4. Find an expression for the joint pdf of $Y(t_1)$ and $W(t_2)$.

(b) Evaluate part a if $X(t)$ is white Gaussian noise.

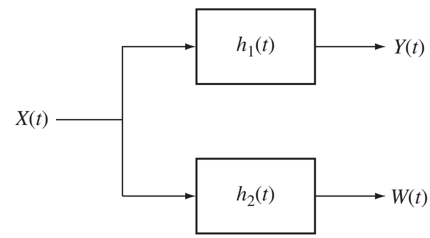

FIGURE P10.4

10.32. Repeat Problem 10.31b if $h_1(t)$ and $h_2(t)$ are ideal bandpass filters as in Example 10.11. Show that $Y(t)$ and $W(t)$ are independent random processes if the filters have nonoverlapping bands.

10.33. Let $Y(t) = h(t) * X(t)$ and $Z(t) = X(t) - Y(t)$ as shown in Fig. P10.5.

(a) Find $S_Z(f)$ in terms of $S_X(f)$.

(b) Find $E[Z^2(t)]$.

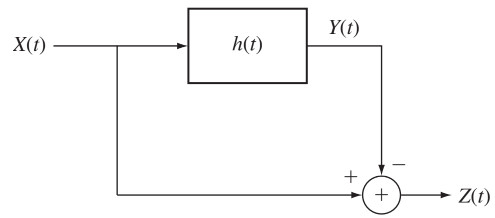

FIGURE P10.5

10.34. Let $Y(t)$ be the output of a linear system with impulse response $h(t)$ and input $X(t) + N(t)$. Let $Z(t) = X(t) - Y(t)$.

(a) Find $R_{X,Y}(\tau)$ and $R_Z(\tau)$.

(b) Find $S_Z(f)$.

(c) Find $S_Z(f)$ if $X(t)$ and $N(t)$ are independent random processes.

10.35. A random telegraph signal is passed through an ideal lowpass filter with cutoff frequency W. Find the power spectral density of the difference between the input and output of the filter. Find the average power of the difference signal.

10.36. Let $Y(t) = a\cos(2\pi f_c t + \Theta) + N(t)$ be applied to an ideal bandpass filter that passes the frequencies $|f - f_c| < W/2$. Assume that Θ is uniformly distributed in $(0, 2\pi)$. Find the ratio of signal power to noise power at the output of the filter.

10.37. Let $Y_n = (X_{n+1} + X_n + X_{n-1})/3$ be a "smoothed" version of X_n. Find $R_Y(k)$, $S_Y(f)$, and $E[Y^2{}_n]$.

10.38. Suppose X_n is a white Gaussian noise process in Problem 10.37. Find the joint pmf for (Y_n, Y_{n+1}, Y_{n+2}).

10.39. Let $Y_n = X_n + \beta X_{n-1}$, where X_n is a zero-mean, first-order autoregressive process with autocorrelation $R_X(k) = \sigma^2\alpha^k$, $|\alpha| < 1$.

(a) Find $R_{Y,X}(k)$ and $S_{Y,X}(f)$.

(b) Find $S_Y(f)$, $R_Y(k)$, and $E[Y_n^2]$.

(c) For what value of β is Y_n a white noise process?

10.40. A zero-mean white noise sequence is input into a cascade of two systems (see Fig. P10.6). System 1 has impulse response $h_n = (1/2)^n u(n)$ and system 2 has impulse response $g_n = (1/4)^n u(n)$ where $u(n) = 1$ for $n \geq 0$ and 0 elsewhere.

(a) Find $S_Y(f)$ and $S_Z(f)$.

(b) Find $R_{W,Y}(k)$ and $R_{W,Z}(k)$; find $S_{W,Y}(f)$ and $S_{W,Z}(f)$. *Hint:* Use a partial fraction expansion of $S_{W,Z}(f)$ prior to finding $R_{W,Z}(k)$.

(c) Find $E[Z_n^2]$.

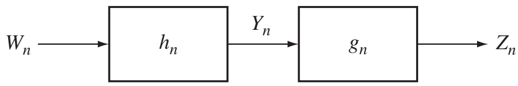

FIGURE P10.6

10.41. A moving average process X_n is produced as follows:

$$X_n = W_n + \alpha_1 W_{n-1} + \cdots + \alpha_p W_{n-p},$$

where W_n is a zero-mean white noise process.

(a) Show that $R_X(k) = 0$ for $|k| > p$.

(b) Find $R_X(k)$ by computing $E[X_{n+k}X_n]$, then find $S_X(f) = \mathcal{F}\{R_X(k)\}$.

(c) Find the impulse response h_n of the linear system that defines the moving average process. Find the corresponding transfer function $H(f)$, and then $S_X(f)$. Compare your answer to part b.

10.42. Consider the second-order autoregressive process defined by

$$Y_n = \frac{3}{4}Y_{n-1} - \frac{1}{8}Y_{n-2} + W_n,$$

where the input W_n is a zero-mean white noise process.

(a) Verify that the unit-sample response is $h_n = 2(1/2)^n - (1/4)^n$ for $n \geq 0$, and 0 otherwise.

(b) Find the transfer function.

(c) Find $S_Y(f)$ and $R_Y(k) = \mathcal{F}^{-1}\{S_Y(f)\}$.

10.43. Suppose the autoregressive process defined in Problem 10.42 is the input to the following moving average system:

$$Z_n = Y_n - 1/4Y_{n-1}.$$

(a) Find $S_Z(f)$ and $R_Z(k)$.

(b) Explain why Z_n is a first-order autoregressive process.

(c) Find a moving average system that will produce a white noise sequence when Z_n is the input.

10.44. An autoregressive process Y_n is produced as follows:

$$Y_n = \alpha_1 Y_{n-1} + \cdots + \alpha_q Y_{n-q} + W_n,$$

where W_n is a zero-mean white noise process.

(a) Show that the autocorrelation of Y_n satisfies the following set of equations:

$$R_Y(0) = \sum_{i=1}^{q} \alpha_i R_Y(i) + R_W(0)$$

$$R_Y(k) = \sum_{i=1}^{q} \alpha_i R_Y(k - i).$$

(b) Use these recursive equations to compute the autocorrelation of the process in Example 10.22.

Section 10.3: Bandlimited Random Processes

10.45. **(a)** Show that the signal $x(t)$ is recovered in Figure 10.10(b) as long as the sampling rate is above the Nyquist rate.

(b) Suppose that a deterministic signal is sampled at a rate below the Nyquist rate. Use Fig. 10.10(b) to show that the recovered signal contains additional signal components from the adjacent bands. The error introduced by these components is called aliasing.

(c) Find an expression for the power spectral density of the sampled bandlimited random process $X(t)$.

(d) Find an expression for the power in the aliasing error components.

(e) Evaluate the power in the error signal in part c if $S_X(f)$ is as in Problem 10.1b.

10.46. An ideal discrete-time lowpass filter has transfer function:

$$H(f) = \begin{cases} 1 & \text{for} \quad |f| < f_c < 1/2 \\ 0 & \text{for} \quad f_c < |f| < 1/2. \end{cases}$$

(a) Show that $H(f)$ has impulse response $h_n = \sin(2\pi f_c n)/\pi n$.

(b) Find the power spectral density of $Y(kT)$ that results when the signal in Problem 10.1b is sampled at the Nyquist rate and processed by the filter in part a.

(c) Let $Y(t)$ be the continuous-time signal that results when the output of the filter in part b is fed to an interpolator operating at the Nyquist rate. Find $S_Y(f)$.

10.47. In order to design a differentiator for bandlimited processes, the filter in Fig. 10.10(c) is designed to have transfer function:

$$H(f) = j2\pi f/T \text{ for } |f| < 1/2.$$

(a) Show that the corresponding impulse response is:

$$h_0 = 0, h_n = \frac{\pi n \cos \pi n - \sin \pi n}{\pi n^2 T} = \frac{(-1)^n}{nT} \quad n \neq 0$$

(b) Suppose that $X(t) = a\cos(2\pi f_0 t + \Theta)$ is sampled at a rate $1/T = 4f_0$ and then input into the above digital filter. Find the output $Y(t)$ of the interpolator.

10.48. Complete the proof of the sampling theorem by showing that the mean square error is zero. *Hint:* First show that $E[(X(t) - (\hat{X}((t)\, X(kT)] = 0$, all k.

10.49. Plot the power spectral density of the amplitude modulated signal $Y(t)$ in Example 10.18, assuming $f_c > W$; $f_c < W$. Assume that $A(t)$ is the signal in Problem 10.1b.

10.50. Suppose that a random telegraph signal with transition rate α is the input signal in an amplitude modulation system. Plot the power spectral density of the modulated signal assuming $f_c = \alpha/\pi$ and $f_c = 10\alpha/\pi$.

10.51. Let the input to an amplitude modulation system be $2\cos(2\pi f_1 + \Phi)$, where Φ is uniformly distributed in $(-\pi, \pi)$. Find the power spectral density of the modulated signal assuming $f_c > f_1$.

10.52. Find the signal-to-noise ratio in the recovered signal in Example 10.18 if $S_N(f) = \alpha f^2$ for $|f \pm f_c| < W$ and zero elsewhere.

10.53. The input signals to a QAM system are independent random processes with power spectral densities shown in Fig. P10.7. Sketch the power spectral density of the QAM signal.

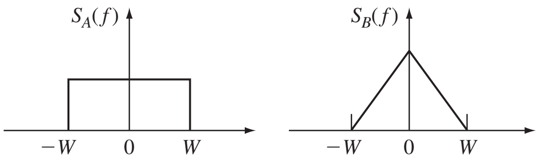

FIGURE P10.7

10.54. Under what conditions does the receiver shown in Fig. P10.8 recover the input signals to a QAM signal?

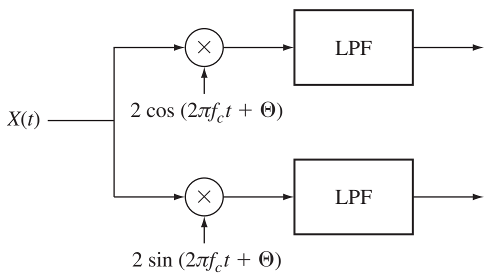

FIGURE P10.8

10.55. Show that Eq. (10.67b) implies that $S_{B,A}(f)$ is a purely imaginary, odd function of f.

Section 10.4: Optimum Linear Systems

10.56. Let $X_\alpha = Z_\alpha + N_\alpha$ as in Example 10.22, where Z_α is a first-order process with $R_Z(k) = 4(3/4)^{|k|}$ and N_α is white noise with $\sigma_N^2 = 1$.
- **(a)** Find the optimum $p = 1$ filter for estimating Z_α.
- **(b)** Find the mean square error of the resulting filter.

10.57. Let $X_\alpha = Z_\alpha + N_\alpha$ as in Example 10.21, where Z_α has $R_Z(k) = \sigma_Z^2(r_1)^{|k|}$ and N_α has $R_N(k) = \sigma_N^2 r_2^{|k|}$, where r_1 and r_2 are less than one in magnitude.
- **(a)** Find the equation for the optimum filter for estimating Z_α.
- **(b)** Write the matrix equation for the filter coefficients.
- **(c)** Solve the $p = 2$ case, if $\sigma_Z^2 = 9$, $r_1 = 2/3$, $\sigma_N^2 = 1$, and $r_2 = 1/3$.
- **(d)** Find the mean square error for the optimum filter in part c.
- **(e)** Use the matrix function of Octave to solve parts c and d for $p = 3, 4, 5$.

10.58. Let $X_\alpha = Z_\alpha + N_\alpha$ as in Example 10.21, where Z_α is the first-order moving average process of Example 10.7, and N_α is white noise.
- **(a)** Find the equation for the optimum filter for estimating Z_α.
- **(b)** For the $p = 1$ and $p = 2$ cases, write and solve the matrix equation for the filter coefficients.
- **(c)** Find the mean square error for the optimum filter in part b.

10.59. Let $X_\alpha = Z_\alpha + N_\alpha$ as in Example 10.19, and suppose that an estimator for Z_α uses observations from the following time instants: $I = \{n - p, \ldots, n, \ldots, n + p\}$.
- **(a)** Solve the $p = 1$ case if Z_α and N_α are as in Problem 10.56.
- **(b)** Find the mean square error in part a.
- **(c)** Find the equation for the optimum filter.
- **(d)** Write the matrix equation for the $2p + 1$ filter coefficients.
- **(e)** Use the matrix function of Octave to solve parts a and b for $p = 2, 3$.

10.60. Consider the predictor in Eq. (10.86b).
- **(a)** Find the optimum predictor coefficients in the $p = 2$ case when $R_Z(k) = 9(1/3)^{|k|}$.
- **(b)** Find the mean square error in part a.
- **(c)** Use the matrix function of Octave to solve parts a and b for $p = 3, 4, 5$.

10.61. Let $X(t)$ be a WSS, continuous-time process.
- **(a)** Use the orthogonality principle to find the best estimator for $X(t)$ of the form

$$\hat{X}(t) = aX(t_1) + bX(t_2),$$

  where t_1 and t_2 are given time instants.
- **(b)** Find the mean square error of the optimum estimator.
- **(c)** Check your work by evaluating the answer in part b for $t = t_1$ and $t = t_2$. Is the answer what you would expect?

10.62. Find the optimum filter and its mean square error in Problem 10.61 if $t_1 = t - d$ and $t_2 = t + d$.

10.63. Find the optimum filter and its mean square error in Problem 10.61 if $t_1 = t - d$ and $t_2 = t - 2d$, and $R_X(\tau) = e^{-\alpha|\tau|}$ Compare the performance of this filter to the performance of the optimum filter of the form $\hat{X}(t) = aX(t - d)$.

10.64. Modify the system in Problem 10.33 to obtain a model for the estimation error in the optimum infinite-smoothing filter in Example 10.24. Use the model to find an expression for the power spectral density of the error $e(t) = Z(t) - Y(t)$, and then show that the mean square error is given by:

$$E[e^2(t)] = \int_{-\infty}^{\infty} \frac{S_Z(f)S_N(f)}{S_Z(f) + S_N(f)} df.$$

Hint: $E[e^2(t)] = R_e(0)$.

10.65. Solve the infinite-smoothing problem in Example 10.24 if $Z(t)$ is the random telegraph signal with $\alpha = 1/2$ and $N(t)$ is white noise. What is the resulting mean square error?

10.66. Solve the infinite-smoothing problem in Example 10.24 if $Z(t)$ is bandlimited white noise of density $N_1/2$ and $N(t)$ is (infinite-bandwidth) white noise of noise density $N_0/2$. What is the resulting mean square error?

10.67. Solve the infinite-smoothing problem in Example 10.24 if $Z(t)$ and $N(t)$ are as given in Example 10.25. Find the resulting mean square error.

10.68. Let $X_n = Z_n + N_n$, where Z_n and N_n are independent, zero-mean random processes.

(a) Find the smoothing filter given by Eq. (10.89) when Z_n is a first-order autoregressive process with $\sigma_X^2 = 9$ and $\alpha = 1/2$ and N_n is white noise with $\sigma_N^2 = 4$.

(b) Use the approach in Problem 10.64 to find the power spectral density of the error $S_e(f)$.

(c) Find $R_e(k)$ as follows: Let $Z = e^{j2\pi f}$, factor the denominator $S_e(f)$, and take the inverse transform to show that:

$$R_e(k) = \frac{\sigma_X^2 z_1}{\alpha(1 - z_1^2)} z_1^{|k|} \quad \text{where} \quad 0 < z_1 < 1.$$

(d) Find an expression for the resulting mean square error.

10.69. Find the Wiener filter in Example 10.25 if $N(t)$ is white noise of noise density $N_0/2 = 1/3$ and $Z(t)$ has power spectral density

$$S_z(f) = \frac{4}{4 + 4\pi^2 f^2}.$$

10.70. Find the mean square error for the Wiener filter found in Example 10.25. Compare this with the mean square error of the infinite-smoothing filter found in Problem 10.67.

10.71. Suppose we wish to estimate (predict) $X(t + d)$ by

$$\hat{X}(t + d) = \int_0^{\infty} h(\tau)X(t - \tau)\, d\tau.$$

(a) Show that the optimum filter must satisfy

$$R_X(\tau + d) = \int_0^{\infty} h(x)R_X(\tau - x)\, dx \qquad \tau \geq 0.$$

(b) Use the Wiener-Hopf method to find the optimum filter when $R_X(\tau) = e^{-2|\tau|}$.

10.72. Let $X_n = Z_n + N_n$, where Z_n and N_n are independent random processes, N_n is a white noise process with $\sigma_N^2 = 1$, and Z_n is a first-order autoregressive process with $R_Z(k) = 4(1/2)^{|k|}$. We are interested in the optimum filter for estimating Z_n from $X_n, X_{n-1}, \ldots$.

(a) Find $S_X(f)$ and express it in the form:

$$S_X(f) = \frac{\frac{1}{2z_1}\left(1 - \frac{1}{z_1}e^{-j2\pi f}\right)\left(1 - z_1 e^{j2\pi f}\right)}{\left(1 - \frac{1}{2}e^{-j2\pi f}\right)\left(1 - \frac{1}{2}e^{j2\pi f}\right)}.$$

(b) Find the whitening causal filter.

(c) Find the optimal causal filter.

Section 10.5: The Kalman Filter

10.73. If W_n and N_n are Gaussian random processes in Eq. (10.102), are Z_n and X_n Markov processes?

10.74. Derive Eq. (10.120) for the mean square prediction error.

10.75. Repeat Example 10.26 with $a = 0.5$ and $a = 2$.

10.76. Find the Kalman algorithm for the case where the observations are given by

$$X_n = b_n Z_n + N_n$$

where b_n is a sequence of known constants.

*Section 10.6: Estimating the Power Spectral Density

10.77. Verify Eqs. (10.125) and (10.126) for the periodogram and the autocorrelation function estimate.

10.78. Generate a sequence X_n of iid random variables that are uniformly distributed in (0, 1).

(a) Compute several 128-point periodograms and verify the random behavior of the periodogram as a function of *f*. Does the periodogram vary about the true power spectral density?

(b) Compute the smoothed periodogram based on 10, 20, and 50 independent periodograms. Compare the smoothed periodograms to the true power spectral density.

10.79. Repeat Problem 10.78 with X_n a first-order autoregressive process with autocorrelation function: $R_X(k) = (.9)^{|k|}$; $R_X(k) = (1/2)^{|k|}$; $R_X(k) = (.1)^{|k|}$.

10.80. Consider the following estimator for the autocorrelation function

$$\hat{r}'_k(m) = \frac{1}{k - |m|} \sum_{n=0}^{k-|m|-1} X_n X_{n+m}.$$

Show that if we estimate the power spectrum of X_n by the Fourier transform of $\hat{r}'_k(m)$, the resulting estimator has mean

$$E[\widetilde{p}_k(f)] = \sum_{m'=-(k-1)}^{k-1} R_X(m') e^{-j2\pi f m'}.$$

Why is the estimator biased?

Section 10.7: Numerical Techniques for Processing Random Signals

10.81. Let $X(t)$ have power spectral density given by $S_X(f) = \beta^2 e^{-f^2/2W_0^2}/\sqrt{2\pi}$.

(a) Before performing an FFT of $S_X(f)$, you are asked to calculate the power in the aliasing error if the signal is treated as if it were bandlimited with bandwidth kW_0.

What value of W should be used for the FFT if the power in the aliasing error is to be less than 1% of the total power? Assume $W_0 = 1000$ and $\beta = 1$.

(b) Suppose you are to perform $N = 2M$ point FFT of $S_X(f)$. Explore how W, T, and t_0 vary as a function of f_0. Discuss what leeway is afforded by increasing N.

(c) For the value of W in part a, identify the values of the parameters f_0, T, and t_0 for $N = 128, 256, 512, 1024$.

(d) Find the autocorrelation $\{R_X(kt_0)\}$ by applying the FFT to $S_X(f)$. Try the options identified in part c and comment on the accuracy of the results by comparing them to the exact value of $R_X(\tau)$.

10.82. Use the FFT to calculate and plot $S_X(f)$ for the following discrete-time processes:

(a) $R_X(k) = 4\alpha^{|k|}$, for $\alpha = 0.25$ and $\alpha = 0.75$.

(b) $R_X(k) = 4(1/2)^{|k|} + 16(1/4)^{|k|}$.

(c) $X_n = \cos(2\pi f_0 n + \Theta)$, where Θ is a uniformly distributed in $(0, 2\pi]$ and $f_0 = 1000$.

10.83. Use the FFT to calculate and plot $R_X(k)$ for the following discrete-time processes:

(a) $S_X(f) = 1$ for $|f| < f_c$ and 0 elsewhere, where $f_c = 1/8, 1/4, 3/8$.

(b) $S_X(f) = 1/2 + 1/2 \cos 2\pi f$ for $|f| < 1/2$.

10.84. Use the FFT to find the output power spectral density in the following systems:

(a) Input X_n with $R_X(k) = 4\alpha^{|k|}$, for $\alpha = 0.25$, $H(f) = 1$ for $|f| < 1/4$.

(b) Input $X_n = \cos(2\pi f_0 n + \Theta)$, where Θ is a uniformly distributed random variable and $H(f) = j2\pi f$ for $|f| < 1/2$.

(c) Input X_n with $R_X(k)$ as in Problem 10.14 with $N = 3$ and $H(f) = 1$ for $|f| < 1/2$.

10.85. (a) Show that

$$R_X(\tau) = 2\text{Re}\left\{ \int_0^{\infty} S_X(f) e^{-j2\pi f \tau}\, df \right\}.$$

(b) Use approximations to express the above as a DFT relating N points in the time domain to N points in the frequency domain.

(c) Suppose we meet the $t_0 f_0 = 1/N$ requirement by letting $t_0 = f_0 = 1/\sqrt{N}$. Compare this to the approach leading to Eq. (10.142).

10.86. (a) Generate a sequence of 1024 zero-mean unit-variance Gaussian random variables and pass it through a system with impulse response $h_n = e^{-2n}$ for $n \geq 0$.

(b) Estimate the autocovariance of the output process of the digital filter and compare it to the theoretical autocovariance.

(c) What is the pdf of the continuous-time process that results if the output of the digital filter is fed into an interpolator?

10.87. (a) Use the covariance matrix factorization approach to generate a sequence of 1024 Gaussian samples with autocovariance $h(t) = e^{-2|\tau|}$.

(b) Estimate the autocovariance of the observed sequence and compare to the theoretical result.

Problems Requiring Cumulative Knowledge

10.88. Does the pulse amplitude modulation signal in Example 9.38 have a power spectral density? Explain why or why not. If the answer is yes, find the power spectral density.

10.89. Compare the operation and performance of the Wiener and Kalman filters for the signals discussed in Example 10.26.

10.90. **(a)** Find the power spectral density of the ARMA process in Example 10.15 by finding the transfer function of the associated linear system.

(b) For the ARMA process find the cross-power spectral density from $E[Y_n X_m]$, and then the power spectral density from $E[Y_n Y_m]$.

10.91. Let $X_1(t)$ and $X_2(t)$ be jointly WSS and jointly Gaussian random processes that are input into two linear time-invariant systems as shown below:

$$X_1(t) \rightarrow \boxed{h_1(t)} \rightarrow Y_1(t)$$

$$X_2(t) \rightarrow \boxed{h_2(t)} \rightarrow Y_2(t)$$

(a) Find the cross-correlation function of $Y_1(t)$ and $Y_2(t)$. Find the corresponding cross-power spectral density.

(b) Show that $Y_1(t)$ and $Y_2(t)$ are jointly WSS and jointly Gaussian random processes.

(c) Suppose that the transfer functions of the above systems are nonoverlapping, that is, $|H_1(f)||H_2(f)| = 0$. Show that $Y_1(t)$ and $Y_2(t)$ are independent random processes.

(d) Now suppose that $X_1(t)$ and $X_2(t)$ are nonstationary jointly Gaussian random processes. Which of the above results still hold?

10.92. Consider the communication system in Example 9.38 where the transmitted signal $X(t)$ consists of a sequence of pulses that convey binary information. Suppose that the pulses $p(t)$ are given by the impulse response of the ideal lowpass filter in Figure 10.6. The signal that arrives at the receiver is $Y(t) = X(t) + N(t)$ which is to be sampled and processed digitally.

(a) At what rate should $Y(t)$ be sampled?

(b) How should the bit carried by each pulse be recovered based on the samples $Y(nT)$?

(c) What is the probability of error in this system?