1

Basic Concepts

1.1 INTRODUCTION

The essence of the spectral estimation problem is captured by the following informal formulation.

From a finite record of a stationary data sequence, estimate how the total power is distributed over frequency.	(1.1.1)

Spectral analysis finds applications in many diverse fields. In *vibration monitoring*, the spectral content of measured signals gives information on the wear and other characteristics of mechanical parts under study. In *economics*, *meteorology*, *astronomy*, and several other fields, the spectral analysis may reveal "hidden periodicities" in the studied data, which are to be associated with cyclic behavior or recurring processes. In *speech analysis*, spectral models of voice signals are useful in better understanding the speech production process and—in addition—can be used for both speech synthesis (or compression) and speech recognition. In *radar and sonar systems*, the spectral contents of the received signals provide information on the location of the sources (or targets) situated in the field of view. In *medicine*, spectral analysis of various signals measured from a patient, such as electrocardiogram (ECG) or electroencephalogram (EEG) signals, can provide useful material for diagnosis. In *seismology*, the spectral analysis of the signals recorded prior to and during a seismic event (such as a volcano eruption or an earthquake) gives useful information on the ground movement associated with such events and could help in predicting them. Seismic spectral estimation is also used to predict subsurface geologic structure in gas and oil exploration. In *control systems*, there is a resurging interest in spectral analysis methods as a means of characterizing the dynamical behavior of a given system and ultimately synthesizing a controller for that system. The previous and other applications of spectral analysis are reviewed in [KAY 1988; MARPLE 1987; BLOOMFIELD 1976; BRACEWELL 1986; HAYKIN 1991; HAYKIN 1995; HAYES III 1996; KOOPMANS 1974; PRIESTLEY 1981; PERCIVAL AND WALDEN 1993; PORAT 1994; SCHARF 1991; THERRIEN 1992; PROAKIS, RADER, LING, AND NIKIAS 1992]. The textbook [MARPLE 1987] also contains a well-written historical perspective on spectral estimation, which is worth reading. Many of the classical articles on spectral analysis, both application-driven and theoretical, are reprinted in [CHILDERS 1978; KESLER 1986]; these excellent collections of reprints are well worth consulting.

There are *two broad approaches* to spectral analysis. One of these derives its basic idea directly from definition (1.1.1): The studied signal is applied to a bandpass filter with a narrow bandwidth, which is swept through the frequency band of interest, and the filter output power divided by the filter bandwidth is used as a measure of the spectral content of the input to the filter. This is essentially what the *classical* (or *nonparametric*) *methods* of spectral analysis do. These methods are described in Chapters 2 and 5 of this text. (The fact that the methods of Chapter 2 can be given the filter-bank interpretation is made clear in Chapter 5.) The second approach to spectral estimation, called the *parametric approach*, is to postulate a model for the data, which provides a means of parameterizing the spectrum, and to thereby reduce the spectral estimation problem to that of estimating the parameters in the assumed model. The parametric approach to spectral analysis is treated in Chapters 3, 4, and 6. Parametric methods offer more accurate spectral estimates than the nonparametric ones in the cases where the data indeed satisfy the model assumed by the former methods. However, in the more likely case that the data do not satisfy the assumed models, the nonparametric methods sometimes outperform the parametric ones, because of the sensitivity of the latter to model misspecifications. This observation has motivated renewed interest in the nonparametric approach to spectral estimation.

Many real-world signals can be characterized as being *random* (from the observer's viewpoint). Briefly speaking, this means that the variation of such a signal outside the observed interval cannot be determined exactly, but can only be specified in statistical terms of averages. In this text, we will be concerned with estimating the spectral characteristics of random signals. In spite of this fact, we find it useful to start the discussion by considering the spectral analysis of deterministic signals (as we do in Section 1.2). Throughout this work, we consider *discrete-index signals* (or *data sequences*). Such signals are most commonly obtained by the temporal or spatial sampling of a continuous (in time or space) signal. The main motivation for focusing on discrete signals lies in the fact that spectral analysis is most often performed by a digital computer or by digital circuitry. Chapters 2 to 5 of this text deal with *discrete-time signals*; Chapter 6 considers the case of *discrete-space data sequences*.

In the interest of notational simplicity, the discrete-time variable t, as used in this text, is assumed to be measured in units of sampling interval. A similar convention is adopted for spatial signals, whenever the sampling is uniform. Accordingly, the *units of frequency* are cycles per sampling interval.

The signals dealt with in the text are *complex valued*. Complex-valued data can appear in signal processing and spectral estimation applications—for instance, as a result of a "complex demodulation" process (explained in detail in Chapter 6). It should be noted that the treatment of complex-valued signals is not always more general or more difficult than the analysis of corresponding real-valued signals. A typical example that illustrates this claim is the case of sinusoidal signals considered in Chapter 4. A real-valued sinusoidal signal, $\alpha \cos(\omega t + \varphi)$, can be rewritten as a linear combination of two complex-valued sinusoidal signals, $\alpha_1 e^{i(\omega_1 t + \varphi_1)} + \alpha_2 e^{i(\omega_2 t + \varphi_2)}$, whose parameters are constrained as follows: $\alpha_1 = \alpha_2 = \alpha/2$, $\varphi_1 = -\varphi_2 = \varphi$, and $\omega_1 = -\omega_2 = \omega$. Here $i = \sqrt{-1}$. The fact that we need to consider *two constrained* complex sine waves to treat the case of *one unconstrained* real sine wave shows that the real-valued case of sinusoidal signals can actually be considered to be more complicated than the complex-valued case! Fortunately, it appears that the latter case is encountered more frequently in applications, where often both the *in-phase* and *quadrature* components of the studied signal are available. For more details and explanations on this aspect, see Section 6.2.

1.2 ENERGY SPECTRAL DENSITY OF DETERMINISTIC SIGNALS

Let $\{y(t); t = 0, \pm 1, \pm 2, \ldots\}$ denote a *deterministic* discrete-time data sequence. Most commonly, $\{y(t)\}$ is obtained by sampling a continuous-time signal. For notational convenience, the time index t is expressed in units of sampling interval—that is, $y(t) = y_c(t \cdot T_s)$, where $y_c(\cdot)$ is the continuous time signal and T_s is the sampling time interval.

Assume that $\{y(t)\}$ has *finite energy*, which means that

$$\sum_{t=-\infty}^{\infty} |y(t)|^2 < \infty \tag{1.2.1}$$

Then, under some additional regularity conditions, the sequence $\{y(t)\}$ possesses a *discrete-time Fourier transform* (DTFT) defined as

$$\boxed{Y(\omega) = \sum_{t=-\infty}^{\infty} y(t) e^{-i\omega t} \qquad \text{(DTFT)}} \tag{1.2.2}$$

In this text, we use the symbol $Y(\omega)$, in lieu of the more cumbersome $Y(e^{i\omega})$, to denote the DTFT. This notational convention is commented on a bit later, following equation (1.4.6). The corresponding inverse DTFT is then

$$\boxed{y(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(\omega) e^{i\omega t} d\omega \qquad \text{(Inverse DTFT)}} \tag{1.2.3}$$

which can be verified by substituting (1.2.3) into (1.2.2). The (angular) *frequency* ω is measured in radians per sampling interval. The conversion from ω to the *physical frequency variable* $\bar{\omega} = \omega / T_s$ [rad/sec] can be done in a straightforward manner, as described in Exercise 1.1.

Let

$$\boxed{S(\omega) = |Y(\omega)|^2 \qquad \text{(Energy Spectral Density)}} \tag{1.2.4}$$

A straightforward calculation gives

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} S(\omega)d\omega &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{t=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} y(t)y^*(s)e^{-i\omega(t-s)}d\omega \\
&= \sum_{t=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} y(t)y^*(s)\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\omega(t-s)}d\omega\right] \\
&= \sum_{t=-\infty}^{\infty} |y(t)|^2
\end{aligned} \tag{1.2.5}$$

To obtain the last equality in (1.2.5), we have used the fact that $\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\omega(t-s)}d\omega = \delta_{t,s}$ (the Kronecker delta). The symbol $(\cdot)^*$ will be used in this text to denote the complex conjugate of a scalar variable or the conjugate transpose of a vector or matrix. Equation (1.2.5) can be restated as

$$\boxed{\sum_{t=-\infty}^{\infty} |y(t)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} S(\omega)d\omega} \tag{1.2.6}$$

This equality is called *Parseval's theorem*. It shows that $S(\omega)$ represents the distribution of sequence energy as a function of frequency. For this reason, $S(\omega)$ is called the *energy spectral density*.

The previous interpretation of $S(\omega)$ also comes up in the following way: Equation (1.2.3) represents the sequence $\{y(t)\}$ as a weighted "sum" (actually, an integral) of orthonormal sequences $\{\frac{1}{\sqrt{2\pi}}e^{i\omega t}\}$ ($\omega \in [-\pi, \pi]$), with weighting $\frac{1}{\sqrt{2\pi}}Y(\omega)$. Hence, $\frac{1}{\sqrt{2\pi}}|Y(\omega)|$ "measures" the "length" of the projection of $\{y(t)\}$ on each of these basis sequences. In loose terms, therefore, $\frac{1}{\sqrt{2\pi}}|Y(\omega)|$ shows how much (or how little) of the sequence $\{y(t)\}$ can be "explained" by the orthonormal sequence $\{\frac{1}{\sqrt{2\pi}}e^{i\omega t}\}$ for some given value of ω.

Define

$$\rho(k) = \sum_{t=-\infty}^{\infty} y(t)y^*(t-k) \tag{1.2.7}$$

It is readily verified that

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} \rho(k)e^{-i\omega k} &= \sum_{k=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} y(t)y^*(t-k)e^{-i\omega t}e^{i\omega(t-k)} \\
&= \left[\sum_{t=-\infty}^{\infty} y(t)e^{-i\omega t}\right]\left[\sum_{s=-\infty}^{\infty} y(s)e^{-i\omega s}\right]^* \\
&= S(\omega)
\end{aligned} \tag{1.2.8}$$

which shows that $S(\omega)$ can be obtained as the DTFT of the "autocorrelation" (1.2.7) of the finite-energy sequence $\{y(t)\}$.

These definitions can be extended in a rather straightforward manner to the case of random signals treated throughout the remaining text. In fact, the only purpose for discussing the deterministic case in this section was to provide some motivation for the analogous definitions in the random case. As such, the discussion in this section has been kept brief. More insights into the meaning and properties of the previous definitions are provided by the detailed treatment of the random case in the next sections.

1.3 POWER SPECTRAL DENSITY OF RANDOM SIGNALS

Most of the signals encountered in applications are such that their future values cannot be determined exactly. We thus resort to probabilistic statements about future values. The mathematical device to describe such a signal is that of a *random sequence*, which consists of an ensemble of possible realizations, each of which has some associated probability of occurrence. Of course, from the whole ensemble of realizations, the experimenter can usually observe only one realization of the signal, and then it might be thought that the deterministic definitions of the previous section could be carried over unchanged to the present case. However, this is not possible, because the realizations of a random signal, viewed as discrete-time sequences, do not have finite energy and hence do not possess DTFTs. A random signal usually has finite *average* power and, therefore, can be characterized by an average power spectral density. For simplicity reasons, in what follows we will use the name *power spectral density* (PSD) for that quantity.

The discrete-time signal $\{y(t); t = 0, \pm 1, \pm 2, \ldots\}$ is assumed to be a sequence of random variables with *zero mean*:

$$E\{y(t)\} = 0 \qquad \text{for all} \quad t \tag{1.3.1}$$

Hereafter, $E\{\cdot\}$ denotes the expectation operator (which averages over the ensemble of realizations). The *autocovariance sequence* (ACS) or *covariance function* of $y(t)$ is defined as

$$\boxed{r(k) = E\left\{y(t)y^*(t-k)\right\}} \tag{1.3.2}$$

and it is assumed to depend only on the lag between the two samples averaged. The two assumptions (1.3.1) and (1.3.2) imply that $\{y(t)\}$ is a *second-order stationary sequence*. When it is required to distinguish between the autocovariance sequences of several signals, a lower index will be used to indicate the signal associated with a given covariance lag, such as $r_y(k)$.

The autocovariance sequence $r(k)$ enjoys some simple, but useful, properties:

$$\boxed{r(k) = r^*(-k)} \tag{1.3.3}$$

and

$$\boxed{r(0) \geq |r(k)| \qquad \text{for all } k} \tag{1.3.4}$$

The equality (1.3.3) directly follows from definition (1.3.2) and the stationarity assumption; (1.3.4) is a consequence of the fact that the *covariance matrix* of $\{y(t)\}$, defined as follows:

$$R_m = \begin{bmatrix} r(0) & r^*(1) & \dots & r^*(m-1) \\ r(1) & r(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & r^*(1) \\ r(m-1) & \dots & r(1) & r(0) \end{bmatrix}$$

$$= E\left\{ \begin{bmatrix} y^*(t-1) \\ \vdots \\ \vdots \\ y^*(t-m) \end{bmatrix} [y(t-1) \dots y(t-m)] \right\} \tag{1.3.5}$$

is positive semidefinite for all m. Recall that a Hermitian matrix M is positive semidefinite if $a^*Ma \geq 0$ for every vector a; see Section A.5 for details. Now,

$$a^*R_m a = a^*E\left\{ \begin{bmatrix} y^*(t-1) \\ \vdots \\ y^*(t-m) \end{bmatrix} [y(t-1) \dots y(t-m)] \right\} a$$

$$= E\left\{z^*(t)z(t)\right\} = E\left\{|z(t)|^2\right\} \geq 0 \tag{1.3.6}$$

where

$$z(t) = [y(t-1) \dots y(t-m)]a$$

so we see that R_m is indeed positive semidefinite for every m. Hence, (1.3.4) follows from the properties of positive semidefinite matrices. (See Definition D11 in Appendix A and Exercise 1.5.)

1.3.1 First Definition of Power Spectral Density

The PSD is defined as the DTFT of the covariance sequence:

$$\boxed{\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k} \qquad \text{(Power Spectral Density)}} \tag{1.3.7}$$

Note that the previous definition (1.3.7) of $\phi(\omega)$ is similar to the definition (1.2.8) in the deterministic case. The inverse transform, which recovers $\{r(k)\}$ from a given $\phi(\omega)$, is

$$\boxed{r(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\omega) e^{i\omega k} d\omega} \tag{1.3.8}$$

We readily verify that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\omega) e^{i\omega k} d\omega = \sum_{p=-\infty}^{\infty} r(p)\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega(k-p)} d\omega\right] = r(k)$$

which proves that (1.3.8) is the inverse transform for (1.3.7). Note that, to obtain the first equality described, the order of integration and summation has been inverted. This order inversion is possible under weak conditions, such as when $\phi(\omega)$ is square integrable—see Chapter 4 in [Priestley 1981] for a detailed discussion on this aspect.

From (1.3.8), we obtain

$$r(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\omega) d\omega \tag{1.3.9}$$

Since $r(0) = E\left\{|y(t)|^2\right\}$ measures the (average) power of $\{y(t)\}$, the equality (1.3.9) shows that $\phi(\omega)$ can indeed be named PSD, as it represents the distribution of the (average) signal power over frequencies. Put another way, it follows from (1.3.9) that $\phi(\omega)d\omega/2\pi$ is the infinitesimal power in the band $(\omega - d\omega/2,\ \omega + d\omega/2)$, and that the total power in the signal is obtained by integrating these infinitesimal contributions. Additional motivation for calling $\phi(\omega)$ a PSD is provided by the second definition of $\phi(\omega)$, given next, which resembles the usual definition (1.2.2), (1.2.4) in the deterministic case.

1.3.2 Second Definition of Power Spectral Density

The second definition of $\phi(\omega)$ is

$$\boxed{\phi(\omega) = \lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{t=1}^{N} y(t) e^{-i\omega t}\right|^2\right\}} \tag{1.3.10}$$

This definition is equivalent to (1.3.7) under the mild assumption that the covariance sequence $\{r(k)\}$ decays sufficiently rapidly that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=-N}^{N} |k||r(k)| = 0 \tag{1.3.11}$$

The equivalence of (1.3.7) and (1.3.10) can be verified as follows:

$$\begin{aligned}
\lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{-i\omega t}\right|^2\right\} &= \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\sum_{s=1}^{N} E\left\{y(t)y^*(s)\right\} e^{-i\omega(t-s)} \\
&= \lim_{N\to\infty}\frac{1}{N}\sum_{\tau=-(N-1)}^{N-1} (N-|\tau|)r(\tau)e^{-i\omega\tau} \\
&= \sum_{\tau=-\infty}^{\infty} r(\tau)e^{-i\omega\tau} - \lim_{N\to\infty}\frac{1}{N}\sum_{\tau=-(N-1)}^{N-1} |\tau| r(\tau)e^{-i\omega\tau} \\
&= \phi(\omega)
\end{aligned}$$

The second equality is proven in Exercise 1.6, and we used (1.3.11) in the last equality.

The second definition just mentioned of $\phi(\omega)$ resembles the definition (1.2.4) of energy spectral density in the deterministic case. The main difference between (1.2.4) and (1.3.10) consists of the appearance of the expectation operator in (1.3.10) and the normalization by $1/N$; the fact that the "discrete-time" variable in (1.3.10) runs over positive integers only is just for convenience and does not constitute an essential difference, compared with (1.2.2). In spite of these differences, the analogy between the deterministic formula (1.2.4) and (1.3.10) provides further motivation for calling $\phi(\omega)$ a PSD. The alternative definition (1.3.10) will also be quite useful when discussing the problem of estimating the PSD by nonparametric techniques in Chapters 2 and 5.

We can see, from either of these definitions, that $\phi(\omega)$ is a *periodic function*, with the period equal to 2π. Hence, $\phi(\omega)$ is completely described by its variation in the interval

$$\boxed{\omega \in [-\pi, \pi] \qquad \text{(radians per sampling interval)}} \tag{1.3.12}$$

Alternatively, the PSD can be viewed as a function of the frequency

$$\boxed{f = \frac{\omega}{2\pi} \qquad \text{(cycles per sampling interval)}} \tag{1.3.13}$$

which, according to (1.3.12), can be considered to take values in the interval

$$\boxed{f \in [-1/2, 1/2]} \tag{1.3.14}$$

We will generally write the PSD as a function of ω whenever possible, because doing so will simplify the notation.

As already mentioned, the discrete-time sequence $\{y(t)\}$ is most commonly derived by sampling a continuous-time signal. To avoid aliasing effects that might be incurred by the sampling process, the continuous-time signal should be (at least, approximately) bandlimited in the frequency domain. To ensure this, it may be necessary to lowpass filter the continuous-time signal before sampling. Let F_0 denote the largest ("significant") frequency component in the spectrum of the (possibly filtered) continuous signal, and let F_s be the *sampling frequency*. Then it follows from the Nyquist sampling theorem (sometimes called the Whittaker–Nyquist–Kotelnikov–Shannon sampling theorem) that the continuous-time signal can be exactly reconstructed from its samples $\{y(t)\}$, provided that

$$F_s > 2F_0 \tag{1.3.15}$$

In particular, no frequency aliasing will occur when (1.3.15) holds. (See, for example, [OPPENHEIM AND SCHAFER 1989].) The frequency variable, F, associated with the continuous-time signal is related to f by the equation

$$F = f \cdot F_s \tag{1.3.16}$$

so it follows that the interval of F corresponding to (1.3.14) is

$$\boxed{F \in \left[-\frac{F_s}{2}, \frac{F_s}{2}\right] \qquad \text{(cycles/sec)}} \tag{1.3.17}$$

which is quite natural in view of (1.3.15).

1.4 PROPERTIES OF POWER SPECTRAL DENSITIES

Since $\phi(\omega)$ is a power density, it should be real valued and nonnegative. That this is indeed the case is readily seen from definition (1.3.10) of $\phi(\omega)$. Hence,

$$\boxed{\phi(\omega) \geq 0 \qquad \text{for all} \ \ \omega} \tag{1.4.1}$$

From (1.3.3) and (1.3.7), we obtain

$$\phi(\omega) = r(0) + 2\sum_{k=1}^{\infty} \text{Re}\{r(k)e^{-i\omega k}\}$$

where Re$\{\cdot\}$ denotes the real part of the bracketed quantity. If $y(t)$, and hence $r(k)$, is real valued, then it follows that

$$\phi(\omega) = r(0) + 2\sum_{k=1}^{\infty} r(k)\cos(\omega k) \tag{1.4.2}$$

which shows that $\phi(\omega)$ is an even function in such a case. In the case of complex-valued signals, however, $\phi(\omega)$ is not necessarily symmetric about the $\omega = 0$ axis. Thus,

$$\boxed{\begin{array}{l} \text{For real-valued signals:} \\ \qquad\qquad \phi(\omega) = \phi(-\omega), \;\; \omega \in [-\pi, \pi] \\ \\ \text{For complex-valued signals:} \\ \qquad\qquad \text{in general } \phi(\omega) \neq \phi(-\omega), \;\; \omega \in [-\pi, \pi] \end{array}} \tag{1.4.3}$$

Remark: The reader might wonder why we did not define the ACS as

$$c(k) = E\left\{y(t)y^*(t+k)\right\}$$

Comparing with the ACS $\{r(k)\}$ used in this text, as defined in (1.3.2), we obtain $c(k) = r(-k)$. Consequently, the PSD associated with $\{c(k)\}$ is related to the PSD corresponding to $\{r(k)\}$ (see (1.3.7)) via

$$\psi(\omega) \triangleq \sum_{k=-\infty}^{\infty} c(k)e^{-i\omega k} = \sum_{k=-\infty}^{\infty} r(k)e^{i\omega k} = \phi(-\omega)$$

It may seem arbitrary as to which definition of the ACS (and corresponding definition of PSD) we choose. In fact, from a mathematical standpoint we can use either definition of the ACS, but the ACS definition $r(k)$ is preferred from a practical standpoint, as we now explain.

First, we should stress that the PSD describes the spectral content of the ACS, as seen from equation (1.3.7). The PSD $\phi(\omega)$ is sometimes perceived as showing the (infinitesimal) power at frequency ω in the signal itself, but that is not strictly true. If the PSD represented the power in the signal itself, then we should have had $\psi(\omega) = \phi(\omega)$, because the signal's spectral content should not depend on the ACS definition. However, as shown earlier, in the general complex case, $\psi(\omega) = \phi(-\omega) \neq \phi(\omega)$, which means that the signal power interpretation of the PSD is not (always) correct. Indeed, the PSD $\phi(\omega)$ "measures" the *power at frequency ω in the signal's ACS*.

On the other hand, our motivation for considering spectral analysis is to characterize the *average power* at frequency *ω in the signal*, as given by the second definition of the PSD in equation (1.3.10). If $c(k)$ is used as the ACS, its corresponding second definition of the PSD is

$$\psi(\omega) = \lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{t=1}^{N} y(t)e^{+i\omega t}\right|^2\right\}$$

which is the average power of $y(t)$ at frequency $-\omega$. Clearly, the second PSD definition corresponding to $r(k)$ aligns with this average-power motivation, whereas the one for $c(k)$ does not; it is for this reason that we use the definition $r(k)$ for the ACS. ■

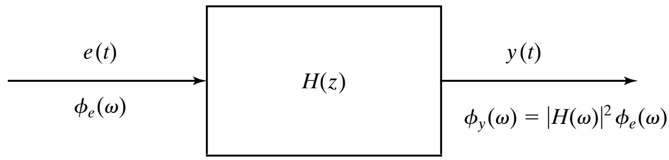

Figure 1.1 Relationship between the PSDs of the input and output of a linear system.

Next, we present a useful result that concerns the *transfer of PSD through an asymptotically stable linear system*. Let

$$H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k} \tag{1.4.4}$$

denote an asymptotically stable linear time-invariant system. The symbol z^{-1} denotes the unit delay operator defined by $z^{-1}y(t) = y(t-1)$. Also, let $e(t)$ be the stationary input to the system and $y(t)$ the corresponding output, as shown in Figure 1.1.Then $\{y(t)\}$ and $\{e(t)\}$ are related via the convolution sum

$$y(t) = H(z)e(t) = \sum_{k=-\infty}^{\infty} h_k e(t-k) \tag{1.4.5}$$

The transfer function of this filter is

$$H(\omega) = \sum_{k=-\infty}^{\infty} h_k e^{-i\omega k} \tag{1.4.6}$$

Throughout the text, we will follow the convention of writing $H(z)$ for the convolution operator of a linear system and its corresponding Z-transform and writing $H(\omega)$ for its transfer function. We obtain the transfer function $H(\omega)$ from $H(z)$ by the substitution $z = e^{i\omega}$:

$$\boxed{H(\omega) = H(z)\big|_{z=e^{i\omega}}}$$

We recognize the slight abuse of notation in writing $H(\omega)$ instead of $H(e^{i\omega})$ and in using z as both an operator and a complex variable, but we prefer the simplicity of notation it affords.

From (1.4.5) and (1.3.2), we obtain

$$\begin{aligned} r_y(k) &= \sum_{p=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} h_p h_m^* E\left\{e(t-p)e^*(t-m-k)\right\} \\ &= \sum_{p=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} h_p h_m^* r_e(m+k-p) \end{aligned} \tag{1.4.7}$$

Inserting (1.4.7) in (1.3.7) gives

$$\begin{aligned}\phi_y(\omega) &= \sum_{k=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_p h_m^* r_e(m+k-p) e^{-i\omega(k+m-p)} e^{i\omega m} e^{-i\omega p} \\ &= \left[\sum_{p=-\infty}^{\infty} h_p e^{-i\omega p}\right] \left[\sum_{m=-\infty}^{\infty} h_m^* e^{i\omega m}\right] \left[\sum_{\tau=-\infty}^{\infty} r_e(\tau) e^{-i\omega\tau}\right] \\ &= |H(\omega)|^2 \phi_e(\omega) \end{aligned} \tag{1.4.8}$$

From (1.4.8), we get the important formula

$$\boxed{\phi_y(\omega) = |H(\omega)|^2 \phi_e(\omega)} \tag{1.4.9}$$

which will be much used in the next chapters.

Finally, we derive a property that will be of use in Chapter 5. Let the signals $y(t)$ and $x(t)$ be related by

$$y(t) = e^{i\omega_0 t} x(t) \tag{1.4.10}$$

for some given ω_0. Then, it holds that

$$\boxed{\phi_y(\omega) = \phi_x(\omega - \omega_0)} \tag{1.4.11}$$

In other words, multiplication of a temporal sequence by $e^{i\omega_0 t}$ shifts its spectral density by the angular frequency ω_0. This interpretation motivates calling the process of constructing $y(t)$, as in (1.4.10), *complex (de)modulation*. The proof of (1.4.11) is immediate: Equations (1.4.10) and (1.3.2) imply that

$$r_y(k) = e^{i\omega_0 k} r_x(k) \tag{1.4.12}$$

so we obtain

$$\phi_y(\omega) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-i(\omega-\omega_0)k} = \phi_x(\omega - \omega_0) \tag{1.4.13}$$

which is the desired result.

1.5 THE SPECTRAL ESTIMATION PROBLEM

The spectral estimation problem can now be stated more formally as follows:

From a finite-length record $\{y(1), \ldots, y(N)\}$ of a second-order stationary random process, find an estimate $\hat{\phi}(\omega)$ of its power spectral density $\phi(\omega)$, for $\omega \in [-\pi, \pi]$.	(1.5.1)

It would, of course, be desirable that $\hat{\phi}(\omega)$ be as close to $\phi(\omega)$ as possible. As we shall see, the main limitation on the quality of most PSD estimates is due to the quite small number of data samples usually available for processing. Note that N will be used throughout this text to denote the number of points of the available data sequence. In some applications, N is small because the cost of obtaining large amounts of data is prohibitive. Most commonly, the value of N is limited by the fact that the signal under study can be considered second-order stationary only over short observation intervals.

As already mentioned in the introductory part of this chapter, there are two main approaches to the PSD estimation problem. The *nonparametric approach*, discussed in Chapters 2 and 5, proceeds to estimate the PSD by relying essentially only on the basic definitions (1.3.7) and (1.3.10) and on some properties that follow directly from these definitions. In particular, these methods do not impose any assumption on the functional form of $\phi(\omega)$. This is in contrast with the *parametric approach*, discussed in Chapters 3, 4, and 6. That approach makes assumptions on the signal under study, which leads to a parameterized functional form of the PSD, and then proceeds by estimating the parameters in the PSD model. The parametric approach can thus be used only when there is enough information about the studied signal to allow formulation of a model. Otherwise, the nonparametric approach should be used. Interestingly enough, the nonparametric methods are close competitors to the parametric ones, even when the model form assumed by the latter is a reasonable description of reality.

1.6 COMPLEMENTS

1.6.1 Coherence Spectrum

Let

$$\boxed{C_{yu}(\omega) = \frac{\phi_{yu}(\omega)}{[\phi_{yy}(\omega)\phi_{uu}(\omega)]^{1/2}}} \tag{1.6.1}$$

denote the so-called *complex coherence* of the stationary signals $y(t)$ and $u(t)$. In the previous definition, $\phi_{yu}(\omega)$ is the cross-spectrum of the two signals, defined as the DTFT of the cross-correlation sequence $r_{yu}(k) = E\{y(t)u^*(t-k)\}$, and $\phi_{yy}(\omega)$ and $\phi_{uu}(\omega)$ are their respective PSDs. (We implicitly assume in (1.6.1) that $\phi_{yy}(\omega)$ and $\phi_{uu}(\omega)$ are strictly positive for all ω.) Also, let

$$\epsilon(t) = y(t) - \sum_{k=-\infty}^{\infty} h_k u(t-k) \tag{1.6.2}$$

denote the residues of the least-squares problem in Exercise 1.11. Hence, $\{h_k\}$ in equation (1.6.2) satisfies

$$\sum_{k=-\infty}^{\infty} h_k e^{-i\omega k} \triangleq H(\omega) = \phi_{yu}(\omega)/\phi_{uu}(\omega).$$

In what follows, we will show that

$$E\left\{|\epsilon(t)|^2\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (1 - |C_{yu}(\omega)|^2)\phi_{yy}(\omega)\, d\omega \tag{1.6.3}$$

where $|C_{yu}(\omega)|$ is the so-called *coherence spectrum*. We will deduce from (1.6.3) that the coherence spectrum shows the extent to which $y(t)$ and $u(t)$ are linearly related to one another, hence providing a motivation for the name given to $|C_{yu}(\omega)|$. We will also show that $|C_{yu}(\omega)| \leq 1$, with equality, for all ω values, if and only if $y(t)$ and $u(t)$ are related as in equation (1.6.2) with $\epsilon(t) \equiv 0$ (in the mean-square sense). Finally, we will show that $|C_{yu}(\omega)|$ is invariant to linear filtering of $u(t)$ and $y(t)$ (possibly by different filters); that is, if $\tilde{u} = g * u$ and $\tilde{y} = f * y$, where f and g are linear filters and $*$ denotes convolution, then $|C_{\tilde{y}\tilde{u}}(\omega)| = |C_{yu}(\omega)|$.

Let $x(t) = \sum_{k=-\infty}^{\infty} h_k u(t-k)$. It can be shown that $u(t-k)$ and $\epsilon(t)$ are uncorrelated with one another for all k. (The reader is required to verify this claim; see also Exercise 1.11). Hence, $x(t)$ and $\epsilon(t)$ are also uncorrelated with each other. Now,

$$y(t) = \epsilon(t) + x(t), \tag{1.6.4}$$

so it follows that

$$\phi_{yy}(\omega) = \phi_{\epsilon\epsilon}(\omega) + \phi_{xx}(\omega) \tag{1.6.5}$$

By using the fact that $\phi_{xx}(\omega) = |H(\omega)|^2\phi_{uu}(\omega)$, we can write

$$\begin{aligned}
E\left\{|\epsilon(t)|^2\right\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_{\epsilon\epsilon}(\omega)\, d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[1 - |H(\omega)|^2 \frac{\phi_{uu}(\omega)}{\phi_{yy}(\omega)}\right] \phi_{yy}(\omega)\, d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[1 - \frac{|\phi_{yu}(\omega)|^2}{\phi_{uu}(\omega)\phi_{yy}(\omega)}\right] \phi_{yy}(\omega)\, d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[1 - |C_{yu}(\omega)|^2\right] \phi_{yy}(\omega)\, d\omega
\end{aligned}$$

which is (1.6.3).

Since the left-hand side in (1.6.3) is nonnegative and the PSD function $\phi_{yy}(\omega)$ is arbitrary, we must have $|C_{yu}(\omega)| \leq 1$ for all ω. It can also be seen from (1.6.3) that the closer $|C_{yu}(\omega)|$ is to 1, the smaller is the residual variance. In particular, if $|C_{yu}(\omega)| \equiv 1$, then $\epsilon(t) \equiv 0$ (in the mean-square sense) and hence $y(t)$ and $u(t)$ must be linearly related, as in (1.7.11). The previous interpretation leads to calling $C_{yu}(\omega)$ *the correlation coefficient in the frequency domain*.

Next, consider the filtered signals

$$\tilde{y}(t) = \sum_{k=-\infty}^{\infty} f_k y(t-k)$$

and

$$\tilde{u}(t) = \sum_{k=-\infty}^{\infty} g_k u(t-k)$$

where the filters $\{f_k\}$ and $\{g_k\}$ are assumed to be stable. As

$$\begin{aligned} r_{\tilde{y}\tilde{u}}(p) &\triangleq E\left\{\tilde{y}(t)\tilde{u}^*(t-p)\right\} \\ &= \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_k g_j^* E\left\{y(t-k)u^*(t-j-p)\right\} \\ &= \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_k g_j^* r_{yu}(j+p-k), \end{aligned}$$

it follows that

$$\begin{aligned} \phi_{\tilde{y}\tilde{u}}(\omega) &= \sum_{p=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_k e^{-i\omega k} \; g_j^* e^{i\omega j} \; r_{yu}(j+p-k) e^{-i\omega(j+p-k)} \\ &= \left(\sum_{k=-\infty}^{\infty} f_k e^{-i\omega k} \right) \left(\sum_{j=-\infty}^{\infty} g_j e^{-i\omega j} \right)^* \left(\sum_{s=-\infty}^{\infty} r_{yu}(s) e^{-i\omega s} \right) \\ &= F(\omega) G^*(\omega) \phi_{yu}(\omega) \end{aligned}$$

Hence,

$$|C_{\tilde{y}\tilde{u}}(\omega)| = \frac{|F(\omega)| \; |G(\omega)| \; |\phi_{yu}(\omega)|}{|F(\omega)|\phi_{yy}^{1/2}(\omega)|G(\omega)|\phi_{uu}^{1/2}(\omega)} = |C_{yu}(\omega)|$$

which is the desired result. Observe that the latter result is similar to the invariance of the modulus of the correlation coefficient in the time domain,

$$\frac{|r_{yu}(k)|}{[r_{yy}(0) r_{uu}(0)]^{1/2}}$$

to a scaling of the two signals: $\tilde{y}(t) = f \cdot y(t)$ and $\tilde{u}(t) = g \cdot u(t)$.

1.7 EXERCISES

Exercise 1.1: Scaling of the Frequency Axis
In this text, the time variable t has been expressed in units of the sampling interval T_s (say). Consequently, the frequency is measured in cycles per sampling interval. Assume we want the frequency units to be expressed in radians per second or in Hertz (Hz = cycles per second). Then we have to introduce the scaled frequency variables

$$\bar{\omega} = \omega/T_s \quad \bar{\omega} \in [-\pi/T_s,\ \pi/T_s] \text{ rad/sec} \tag{1.7.1}$$

and $\bar{f} = \bar{\omega}/2\pi$ (in Hz). It might be thought that the PSD in the new frequency variable is obtained by inserting $\omega = \bar{\omega}T_s$ into $\phi(\omega)$, but this is *wrong*. Show that the PSD, *as expressed in units of power per Hz*, is in fact given by

$$\boxed{\bar{\phi}(\bar{\omega}) = T_s\phi(\bar{\omega}T_s) \triangleq T_s \sum_{k=-\infty}^{\infty} r(k)e^{-i\bar{\omega}T_s k}, \qquad |\bar{\omega}| \le \pi/T_s} \tag{1.7.2}$$

(See [MARPLE 1987] for more details on this scaling aspect.)

Exercise 1.2: Time–Frequency Distributions
Let $y(t)$ denote a discrete-time signal, and let $Y(\omega)$ be its discrete-time Fourier transform. As explained in Section 1.2, $Y(\omega)$ shows how the energy in the *whole sequence* $\{y(t)\}_{t=-\infty}^{\infty}$ is distributed over frequency.

Assume that we want to characterize how the energy of the signal is distributed in *time and frequency*. If $D(t, \omega)$ is a function that characterizes the time–frequency distribution, then it should satisfy the so-called *marginal properties*:

$$\sum_{t=-\infty}^{\infty} D(t, \omega) = |Y(\omega)|^2 \tag{1.7.3}$$

and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D(t, \omega) d\omega = |y(t)|^2 \tag{1.7.4}$$

Use intuitive arguments to explain why the previous conditions are desirable properties of a time–frequency distribution. Next, show that the so-called Rihaczek distribution,

$$\boxed{D(t, \omega) = y(t)Y^*(\omega)e^{-i\omega t}} \tag{1.7.5}$$

satisfies conditions (1.7.3) and (1.7.4). (For treatments of the time–frequency distributions, the reader is referred to [THERRIEN 1992] and [COHEN 1995].)

Exercise 1.3: Two Useful Z-Transform Properties

(a) Let h_k be an absolutely summable sequence, and let $H(z) = \sum_{k=-\infty}^{\infty} h_k z^{-k}$ be its Z-transform. Find the Z-transforms of the following two sequences:
(i) h_{-k}
(ii) $g_k = \sum_{m=-\infty}^{\infty} h_m h_{m-k}^*$.

(b) Show that, if z_i is a zero of $A(z) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n}$, then $(1/z_i^*)$ is a zero of $A^*(1/z^*)$ (where $A^*(1/z^*) = [A(1/z^*)]^*$).

Exercise 1.4: A Simple ACS Example
Let $y(t)$ be the output of a linear system, as in Figure 1.1, with filter $H(z) = (1 + b_1 z^{-1})/(1 + a_1 z^{-1})$ whose input is zero-mean white noise with variance σ^2. (The ACS of such an input is $\sigma^2 \delta_{k,0}$.)

(a) Find $r(k)$ and $\phi(\omega)$ analytically in terms of a_1, b_1, and σ^2.

(b) Verify that $r(-k) = r^*(k)$ and that $|r(k)| \le r(0)$. Also verify that, when a_1 and b_1 are real, $r(k)$ can be written as a function of $|k|$.

Exercise 1.5: Alternative Proof that $|r(k)| \le r(0)$
We stated in the text that (1.3.4) follows from (1.3.6). Provide a proof of that statement. Also, find an alternative, simple proof of (1.3.4) by using (1.3.8).

Exercise 1.6: A Double Summation Formula
A result often used in the study of discrete-time random signals is the summation formula

$$\boxed{\sum_{t=1}^{N} \sum_{s=1}^{N} f(t-s) = \sum_{\tau=-N+1}^{N-1} (N - |\tau|) f(\tau)} \tag{1.7.6}$$

where $f(\cdot)$ is an arbitrary function. Provide a proof of this formula.

Exercise 1.7: Is a Truncated Autocovariance Sequence (ACS) a Valid ACS?
Suppose that $\{r(k)\}_{k=-\infty}^{\infty}$ is a valid ACS; thus, $\sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k} \ge 0$ for all ω. Is it possible that, for some integer p, the partial (or truncated) sum

$$\sum_{k=-p}^{p} r(k) e^{-i\omega k}$$

is negative for some ω? Justify your answer.

Exercise 1.8: When Is a Sequence an Autocovariance Sequence?

We showed in Section 1.3 that, if $\{r(k)\}_{k=-\infty}^{\infty}$ is an ACS, then $R_m \geq 0$ for $m = 0, 1, 2, \ldots$. We also implied that the first definition of the PSD in (1.3.7) satisfies $\phi(\omega) \geq 0$ for all ω; however, we did not prove this by using (1.3.7) solely. Show that

$$\phi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k} \geq 0 \text{ for all } \omega$$

if and only if

$$a^* R_m a \geq 0 \quad \text{for every } m \text{ and for every vector } a$$

Exercise 1.9: Spectral Density of the Sum of Two Correlated Signals

Let $y(t)$ be the output to the system shown in Figure 1.2. Assume $H_1(z)$ and $H_2(z)$ are linear, asymptotically stable systems. The inputs $e_1(t)$ and $e_2(t)$ are each zero-mean white noise, with

$$E\left\{\begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix}\begin{bmatrix} e_1^*(s) & e_2^*(s) \end{bmatrix}\right\} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\delta_{t,s}$$

(a) Find the PSD of $y(t)$.
(b) Show that, for $\rho = 0$, $\phi_y(\omega) = \phi_{x_1}(\omega) + \phi_{x_2}(\omega)$.
(c) Show that, for $\rho = \pm 1$ and $\sigma_1^2 = \sigma_2^2 = \sigma^2$, $\phi_y(\omega) = \sigma^2 |H_1(\omega) \pm H_2(\omega)|^2$.

Exercise 1.10: Least-Squares Spectral Approximation

Assume we are given an ACS $\{r(k)\}_{k=-\infty}^{\infty}$ or, equivalently, a PSD function $\phi(\omega)$, as in equation (1.3.7). We wish to find a finite-impulse response (FIR) filter, as in Figure 1.1, where $H(\omega) = h_0 + h_1 e^{-i\omega} + \ldots + h_m e^{-im\omega}$, whose input $e(t)$ is zero-mean unit-variance white noise and such

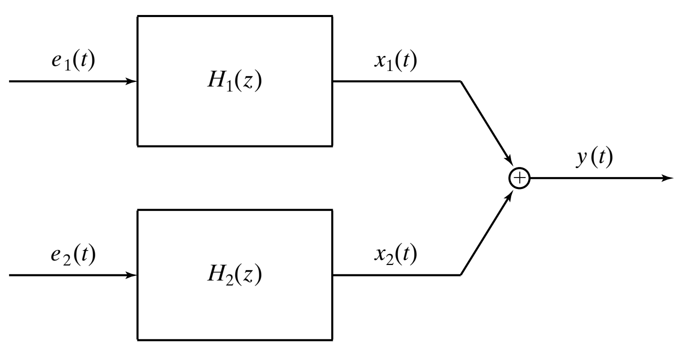

Figure 1.2 The system considered in Exercise 1.9.

that the output sequence $y(t)$ has PSD $\phi_y(\omega)$ "close to" $\phi(\omega)$. Specifically, we wish to find $h = [h_0 \dots h_m]^T$ so that the approximation error

$$\epsilon = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\phi(\omega) - \phi_y(\omega)]^2 \, d\omega \tag{1.7.7}$$

is minimum.

(a) Show that ϵ is a quartic (fourth-order) function of h and that thus no simple closed-form solution h exists to minimize (1.7.7).

(b) We attempt to reparameterize the minimization problem as follows: We note that $r_y(k) \equiv 0$ for $|k| > m$; thus,

$$\phi_y(\omega) = \sum_{k=-m}^{m} r_y(k) e^{-i\omega k} \tag{1.7.8}$$

Equation (1.7.8), and the fact that $r_y(-k) = r_y^*(k)$, mean that $\phi_y(\omega)$ is a function of $g = [r_y(0) \dots r_y(m)]^T$. Show that the minimization problem in (1.7.7) is quadratic in g; it thus admits a closed-form solution. Show that the vector g that minimizes ϵ in equation (1.7.7) gives

$$r_y(k) = \begin{cases} r(k), & |k| \le m \\ 0, & \text{otherwise} \end{cases} \tag{1.7.9}$$

(c) Can you identify any problems with the "solution" (1.7.9)?

Exercise 1.11: Linear Filtering and the Cross-Spectrum

For two stationary signals $y(t)$ and $u(t)$, with cross-covariance sequence $r_{yu}(k) = E\{y(t) u^*(t-k)\}$, the *cross-spectrum* is defined as

$$\boxed{\phi_{yu}(\omega) = \sum_{k=-\infty}^{\infty} r_{yu}(k) e^{-i\omega k}} \tag{1.7.10}$$

Let $y(t)$ be the output of a linear filter with input $u(t)$,

$$y(t) = \sum_{k=-\infty}^{\infty} h_k u(t-k) \tag{1.7.11}$$

Show that the input PSD, $\phi_{uu}(\omega)$, the filter transfer function

$$H(\omega) = \sum_{k=-\infty}^{\infty} h_k e^{-i\omega k}$$

and $\phi_{yu}(\omega)$ are related through the so-called *Wiener–Hopf equation*:

$$\boxed{\phi_{yu}(\omega) = H(\omega)\phi_{uu}(\omega)} \tag{1.7.12}$$

Next, consider the least-squares (LS) problem

$$\min_{\{h_k\}} E\left\{\left| y(t) - \sum_{k=-\infty}^{\infty} h_k u(t-k)\right|^2\right\} \tag{1.7.13}$$

where now $y(t)$ and $u(t)$ are no longer necessarily related through equation (1.7.11). Show that the filter minimizing the preceding LS criterion is still given by the Wiener–Hopf equation, by minimizing the expectation in (1.7.13) with respect to the real and imaginary parts of h_k. (Assume that $\phi_{uu}(\omega) > 0$ for all ω.)

COMPUTER EXERCISES

Exercise C1.12: Computer Generation of Autocovariance Sequences
Autocovariance sequences are two-sided sequences. In this exercise, we develop computer techniques for generating two-sided ACSs.

Let $y(t)$ be the output of the linear system in Figure 1.1, with filter $H(z) = (1 + b_1 z^{-1})/(1 + a_1 z^{-1})$, whose input is zero-mean white noise with variance σ^2.

(a) Find $r(k)$ analytically in terms of a_1, b_1, and σ^2. (See also Exercise 1.4.)
(b) Plot $r(k)$ for $-20 \le k \le 20$ and for various values of a_1 and b_1. Notice that the tails of $r(k)$ decay at a rate dictated by $|a_1|$.
(c) When $a_1 \simeq b_1$ and $\sigma^2 = 1$, then $r(k) \simeq \delta_{k,0}$. Verify this for $a_1 = -0.95$, $b_1 = -0.9$, and for $a_1 = -0.75$, $b_1 = -0.7$.
(d) A quick way to generate (approximately) $r(k)$ on the computer is to use the fact that $r(k) = \sigma^2 h(k) * h^*(-k)$, where $h(k)$ is the impulse response of the filter in Figure 1.1 (see equation (1.4.7)) and $*$ denotes convolution. Consider the case where

$$H(z) = \frac{1 + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}}$$

Write a MATLAB function `genacs.m` whose inputs are M, σ^2, a, and b, where a and b are the vectors of denominator and numerator coefficients, respectively, and whose output is a vector of ACS coefficients from 0 to M. Your function should make use of the MATLAB functions `filter` (to generate $\{h_k\}_{k=0}^{M}$) and `conv` (to compute $r(k) = \sigma^2 h(k) * h^*(-k)$ by using the truncated impulse response sequence).

(e) Test your function, using $\sigma^2 = 1$, $a_1 = -0.9$, and $b_1 = 0.8$. Try $M = 20$ and $M = 150$; why is the result more accurate for larger M? Suggest a "rule of thumb" about a good choice of M in relation to the poles of the filter.

This method is a "quick and simple" way to compute an approximation to the ACS, but it is sometimes not very accurate because the impulse response is truncated. Methods for computing the exact ACS from σ^2, a, and b are discussed in Exercise 3.2 and also in [Kinkel, Perl, Scharf, and Stubberud 1979; Demeure and Mullis 1989].

Exercise C1.13: DTFT Computations Using Two-Sided Sequences

In this exercise, we consider the DTFT of two-sided sequences (including autocovariance sequences); in doing so, we illustrate some basic properties of autocovariance sequences.

(a) We first consider how to use the DTFT to determine $\phi(\omega)$ from $r(k)$ on a computer. We are given an ACS:

$$r(k) = \begin{cases} \dfrac{M - |k|}{M}, & |k| \leq M \\ 0, & \text{otherwise} \end{cases} \tag{1.7.14}$$

Generate $r(k)$ for $M = 10$. Form, in MATLAB, a vector $\mathtt{x}$ of length $L = 256$ as

$$\mathtt{x} = [r(0), r(1), \ldots, r(M), 0 \ldots, 0, r(-M), \ldots, r(-1)]$$

Verify that `xf=fft(x)` gives $\phi(\omega_k)$ for $\omega_k = 2\pi k/L$. (Note that the elements of `xf` should be nonnegative and real.) Explain why this particular choice of `x` is needed, citing appropriate circular shift and zero-padding properties of the DTFT.

Note that `xf` often contains a very small imaginary part due to computer roundoff error; replacing `xf` by `real(xf)` truncates this imaginary component and leads to more expected results when plotting.

A word of caution—do not truncate the imaginary part unless you are sure it is negligible; the command `zf=real(fft(z))` when

$$\mathtt{z} = [r(-M), \ldots, r(-1), r(0), r(1), \ldots, r(M), 0 \ldots, 0]$$

gives erroneous "spectral" values; try it and explain why it does not work.

(b) Alternatively, since we can readily derive the analytical expression for $\phi(\omega)$, we can instead work backwards. Form a vector

$$\mathtt{yf} = [\phi(0), \phi(2\pi/L), \phi(4\pi/L), \ldots, \phi((L-1)2\pi/L)]$$

and find `y=ifft(yf)`. Verify that `y` closely approximates the ACS.

(c) Consider the ACS $r(k)$ in Exercise C1.12; let $a_1 = -0.9$ and $b_1 = 0$, and set $\sigma^2 = 1$. Form a vector `x` as before, with $M = 10$, and find `xf`. Why is `xf` not a good approximation of $\phi(\omega_k)$ in this case? Repeat the experiment for $M = 127$ and $L = 256$; is the approximation better for this case? Why?

We can again work backwards from the analytical expression for $\phi(\omega)$. Form a vector

$$\mathtt{yf} = [\phi(0), \phi(2\pi/L), \phi(4\pi/L), \ldots, \phi((L-1)2\pi/L)]$$

and find `y=ifft(yf)`. Verify that `y` closely approximates the ACS for large L (say, $L = 256$), but poorly approximates the ACS for small L (say, $L = 20$). By citing properties of inverse DTFTs of infinite, two-sided sequences, explain how the elements of `y` relate to the ACS $r(k)$ and why the approximation is poor for small L. Based on this explanation, give a "rule of thumb" for a choice of L that gives a good approximations of the ACS.

(d) We have seen that the `fft` command results in spectral estimates for $\omega \in [0, 2\pi)$ instead of the more commonly-used range $\omega \in [-\pi, \pi)$. The MATLAB command `fftshift` can be used to exchange the first and second halves of the `fft` output to make it correspond to the frequency range from $\omega \in [-\pi, \pi)$. Similarly, `fftshift` can be used on the output of the `ifft` operation to "center" the zero lag of an ACS. Experiment with `fftshift` to achieve both of these results. What frequency vector `w` is needed so that the command `plot(w, fftshift(fft(x)))` gives the spectral values at the proper frequencies? Similarly, what time vector `t` is needed to get a proper plot of the ACS with `stem(t,fftshift(ifft(xf)))`? Do the results depend on whether the vectors are even or odd in length?

Exercise C1.14: Relationship between the PSD and the Eigenvalues of the ACS Matrix

An interesting property of the ACS matrix R in equation (1.3.5) is that, for large dimensions m, its eigenvalues are close to the values of the PSD $\phi(\omega_k)$ for $\omega_k = 2\pi k/m$, $k = 0, 1, \ldots, m-1$. (See, for example, [Gray 1972].) We verify this property here:

Consider the ACS in Exercise C1.12, with the values $a_1 = -0.9$, $b_1 = 0.8$, and $\sigma^2 = 1$.

(a) Compute a vector `phi` that contains the values of $\phi(\omega_k)$ for $\omega_k = 2\pi k/m$, with $m = 256$ and $k = 0, 1, \ldots, m-1$. Plot a histogram of these values with `hist(phi)`. Also useful is the cumulative distribution of the values of `phi` (plotted on a logarithmic scale), which can be found with the command `semilogy( (1/m:1/m:1), sort(phi) )`.

(b) Compute the eigenvalues of R in equation (1.3.5) for various values of m. Plot the histogram of the eigenvalues and plot their cumulative distribution. Verify that, as m increases, the cumulative distribution of the eigenvalues approaches the cumulative distribution of the $\phi(\omega)$ values. Similarly, the histograms also approach the histogram for $\phi(\omega)$, but it is easier to see this result by using cumulative distributions than by using histograms.