CHAPTER 7

MODELING BIOMEDICAL SIGNAL GENERATING PROCESSES AND SYSTEMS

Up to this point in the book, we have concentrated on processing and analysis of biomedical signals. The signals were treated in their own right as conveyors of diagnostic information. While it was emphasized that the design and application of signal analysis procedures require an understanding of the physiological and pathological processes and systems that generate the signals, no specific mathematical model was used to represent the genesis of the signals in the methods we studied in the preceding chapters.

We shall now consider the modeling approach, where an explicit mathematical model is used to represent the process or the system that generates the signal of interest. The parameters of the model are then investigated for use in signal analysis, pattern recognition, and decision making. As we shall see, the model parameters may also be related to the physical or physiological aspects of the related systems. The *parametric modeling* approach often leads to succinct and efficient representation of signals and systems. Regardless of the emphasis on modeling, the final aim of the methods described in this chapter is analysis of the signal of interest.

7.1   Problem Statement

Propose mathematical models to represent the generation of biomedical signals. Identify the possible relationships between the mathematical models and the physiological and pathological processes and systems that generate the signals. Explore the potential use of the model parameters in signal analysis, pattern recognition, and classification.

Given the diversity of the biomedical signals that we have already encountered and the many others that exist, a generic model cannot be expected to represent a large number of signals. Indeed, a very specific model is often required for each signal. Bioelectric signals such as the ECG and EMG may be modeled using the basic action potential or SMUAP as the building block. Sound and vibration signals such as the PCG and speech may be modeled using fluid-filled resonating chambers, turbulent flow across a baffle or through a constriction, vibrating pipes, and acoustic or vibrational excitation of a tract of variable shape. We shall investigate a few representative signals and models in the following sections, and then study a few modeling techniques that facilitate signal analysis based upon the parameters extracted.

7.2   Illustration of the Problem with Case Studies

7.2.1   Motor unit firing patterns

We saw in Section 1.2.4 that the surface EMG of an active skeletal muscle is the spatiotemporal summation of the action potentials of a large number of motor units that have been recruited into action (see Figure 1.13). If we consider the EMG of a single motor unit, we have a train of SMUAPs; the same basic wave (spike, pulse, or wavelet) is repeated in a quasiperiodic sequence. For the sake of generality, we may represent the intervals between the SMUAPs by a random variable: Although an overall periodicity exists and is represented by the firing rate in pps, the intervals between the pulses, known as the *interpulse interval* or IPI, may not precisely be the same from one SMUAP to another.

Agarwal and Gottlieb [339] modeled the single-motor-unit EMG as the convolution of a series of unit impulses or Dirac delta functions — known as a *point process* [340–343] — with the basic SMUAP wave. The SMUAP train $y(t)$ may then be modeled as the output of a linear system whose impulse response $h(t)$ is the SMUAP, and the input is a point process $x(t)$:

$$y(t) = \int_0^t h(t - \tau)\, x(\tau)\, d\tau. \tag{7.1}$$

Physiological conditions dictate that successive action potentials of the same motor unit cannot overlap: The interval between any two pulses should be greater than the SMUAP duration. In normal muscle activation, SMUAP durations are of the order of $3 - 15\ ms$ and motor unit firing rates are in the range $6 - 30\ pps$; the IPI is, therefore, in the range of about $30 - 170\ ms$, which is significantly higher than

the SMUAP duration. An SMUAP train, therefore, consists of discrete (distinct and separated) events or waves.

The model as above permits independent analysis of SMUAP waveshape and firing pattern: The two are, indeed, physiologically separate entities. The SMUAP waveshape depends upon the spatial arrangement and characteristics of the muscle fibers that constitute the motor unit, whereas the firing pattern is determined by the motor neuron that stimulates the muscle fibers. Statistics of the point process representing the IPI may be used to study the muscle activation process independently of the SMUAP waveshape. Details on point processes and their application to EMG modeling are presented in Section 7.3.

7.2.2   Cardiac rhythm

The ECG is a quasiperiodic signal that is also cyclostationary in the normal case (see Section 1.2.5). Each beat is triggered by a pulse from the SA node. The P wave is the combined result of the action potentials of the atrial muscle cells, while the QRS and T waves are formed by the spatiotemporal summation of the action potentials of the ventricular muscle cells.

In rhythm analysis, one is more interested in the timing of the beats than in their individual waveshape (with the exception of PVCs). Diseases that affect the SA node could disturb the normal rhythm, and lead to abnormal variability in the RR intervals. Disregarding the details of atrial and ventricular ECG waves, an ECG rhythm may be modeled by a point process representing the firing pattern of the SA node. Sinus arrhythmia and HRV may then be investigated by studying the distribution and statistics of the RR interval.

Figure 7.1 illustrates the representation of ECG complexes in terms of the instantaneous heart rate values defined as the inverse of the RR interval of each beat, in terms of a series of RR interval values, and as a train of delta functions at the SA node firing instants [137]. A discrete-time signal may be derived by sampling the signal in Figure 7.1 (b) at equidistant points; the result, however, may not be continuous or differentiable [137]. The signal in Figure 7.1 (c), known as the interval series, has values $I_k = t_k - t_{k-1}$, where the instants t_k represent the time instants at which the QRS complexes occur in the ECG signal. The I_k series is defined as a function of interval number and not of time, and hence may pose difficulties regarding interpretation in the frequency domain. Finally, the signal in Figure 7.1 (d) is defined as a train of Dirac delta functions $s(t) = \sum \delta(t - t_k)$. The series of impulses represents a point process that may be analyzed and interpreted with relative ease, as shown in Section 7.3. The last two representations may be used to analyze cardiac rhythm and HRV, as demonstrated in Section 7.8 (see also Section 8.11).

7.2.3   Formants and pitch in speech

Speech signals are formed by exciting the vocal tract with either a pulse train or a random signal produced at the glottis, and possibly their combination as well (see Section 1.2.12). The shape of the vocal tract is varied according to the nature of the
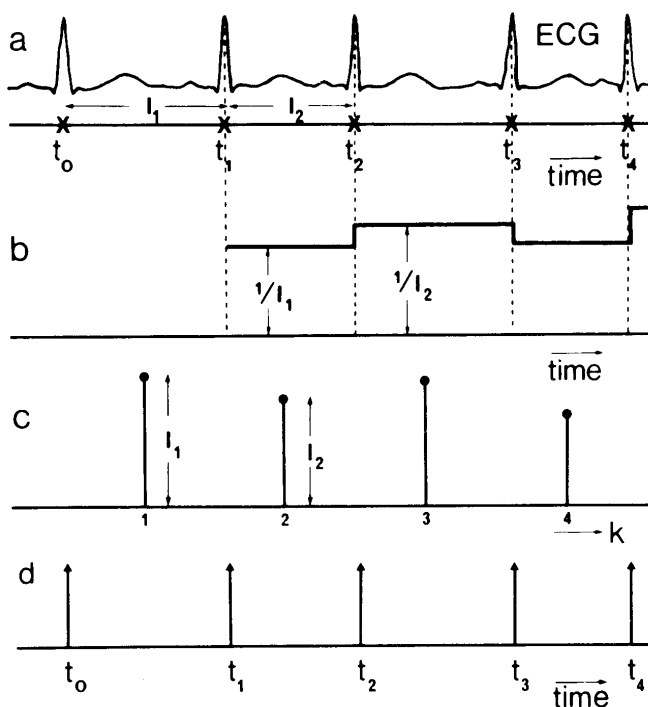
Figure 7.1   The train of ECG complexes in (a) is represented in terms of: (b) the instantaneous heart rate values defined as the inverse of the RR interval of each beat; (c) a series of RR interval values (known as the interval series); and (d) a train of delta functions at the SA node firing instants. Reproduced with permission from R.W. DeBoer, J.M. Karemaker, and J. Strackee, Comparing spectra of a series of point events particularly for heart rate variability studies, *IEEE Transactions on Biomedical Engineering,* 31(4):384–387, 1984. ©IEEE.

sound or phoneme to be produced; the system is, therefore, a time-variant system. We may model the output as the convolution of the (time-variant) impulse response of the vocal tract with the input glottal waveform. The input may be modeled by a random process for unvoiced speech and as a point process for voiced speech; see Figures 1.47, 1.48, 1.49, and 7.2. Clearly, the speech signal is a nonstationary signal; however, the signal may be considered to be quasistationary over short intervals of time during which the same phoneme is being produced.

Figure 7.2 illustrates the commonly used model for speech production [79]; see also Figures 1.47, 1.48, and 1.49. The speech signal may be modeled using the same convolutional relationship as in Equation 7.1, with the limitation that the expression is valid over durations of time when the vocal-tract shape is held fixed and the same glottal excitation is applied. Then, $h(t)$ represents the impulse response of the vocal-tract system (filter) for the time interval considered, and $x(t)$ represents the glottal waveform that is applied as input to the system. In the case of voiced speech, the IPI

statistics of the point-process input, in particular its mean, are related to the pitch. Furthermore, the frequency response of the filter $H(\omega)$ representing the vocal tract determines the spectral content of the speech signal: The dominant frequencies or peaks are known as formants in the case of voiced speech.
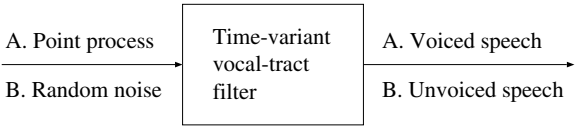

Figure 7.2   Model for production of speech, treating the vocal tract as a time-variant linear system. A point-process input generates quasiperiodic voiced speech, whereas a random-noise input generates unvoiced speech. See also Figures 1.47, 1.48, and 1.49.

Point processes are described in Section 7.3. Parametric spectral modeling and analysis techniques suitable for formant extraction are described in Sections 7.4, 7.5, and 7.6.

7.2.4   Patellofemoral crepitus

Among the various types of VAG signals produced by the knee joint (see Section 1.2.14), the most common is a signal known as physiological patellofemoral crepitus (PPC) [97, 344–347]. The PPC signal is a random sequence of vibrational pulses generated between the surfaces of the patella and the femur, typically observed during slow movement of the knee joint. The PPC signal may carry information on the state and lubrication of the knee joint. A mechanical model of the knee-joint surfaces that generate PPC, as proposed by Beverland et al. [346], is described in Section 7.7.3.

Zhang et al. [344] proposed a model for generation of the PPC signal based on point processes, similar to that for the SMUAP train described in Section 7.2.1. The effects of the repetition rate (or IPI) and the basic patellofemoral pulse (PFP) waveform on the spectrum of the PPC signal were analyzed separately. It was suggested that the model could represent the relationships between physiological parameters, such as the mean and SD of the IPI as well as the PFP waveshape, and parameters that could be measured from the PPC signal, such as its mean, RMS, and PSD-based features. Illustrations related to this application are provided at the end of Section 7.3.

7.3   Point Processes

Problem: *Formulate a mathematical model representing the generation of a train of SMUAPs and derive an expression for the PSD of the signal.*

Solution: In the model for EMG generation proposed by Agarwal and Gottlieb [339], a point process is used to represent the motor neuron firing sequence,

and the SMUAP train is modeled by the convolution integral as in Equation 7.1. The IPI is treated as a sequence of independent random variables with identical normal (Gaussian) PDFs.

Let the interval between the ith SMUAP and the preceding one be τ_i, and let the origin be set at the instant of appearance of the first SMUAP at $i = 0$ with $\tau_0 = 0$. The time of arrival of the ith SMUAP is then given by $t_i = \tau_1 + \tau_2 + \cdots + \tau_i$. The variable t_i is the sum of i independent random variables; note that $\tau_i > 0$. It is assumed that the mean μ and variance σ^2 of the random variable representing each IPI are the same. Then, the mean of t_i is $i\mu$, and its variance is $i\sigma^2$. Furthermore, t_i is also a random variable with the Gaussian PDF

$$p_{t_i}(t_i) = \frac{1}{\sqrt{2\pi i}\,\sigma}\,\exp\left[-\frac{(t_i - i\mu)^2}{2i\sigma^2}\right].\tag{7.2}$$

If the SMUAP train has $N + 1$ SMUAPs labeled as $i = 0, 1, 2, \ldots, N$, the motor neuron firing sequence is represented by the point process

$$x(t) = \sum_{i=0}^{N} \delta(t - t_i).\tag{7.3}$$

The Fourier transform of the point process is

$$\begin{aligned}X(\omega) &= \int_{-\infty}^{\infty} \sum_{i=0}^{N} \delta(t - t_i)\,\exp(-j\omega t)\,dt\tag{7.4}\\ &= \sum_{i=0}^{N} \exp(-j\omega t_i).\end{aligned}$$

$X(\omega)$ is a function of the random variable t_i, which is, in turn, a function of i random variables $\tau_1, \tau_2, \ldots, \tau_i$. Therefore, $X(\omega)$ is random. The ensemble average of $X(\omega)$ may be obtained by computing its expectation, taking into account the PDF of t_i, as follows [339]:

$$\overline{X}(\omega) = E[X(\omega)] = \sum_{i=0}^{N} E[\exp(-j\omega t_i)].\tag{7.5}$$

$$E[\exp(-j\omega t_i)] = \int_{-\infty}^{\infty} \exp(-j\omega t_i)\,p_{t_i}(t_i)\,dt_i.\tag{7.6}$$

Using the expression for $p_{t_i}(t_i)$ in Equation 7.2, we get

$$E[\exp(-j\omega t_i)] = \frac{1}{\sqrt{2\pi i}\sigma} \int_{-\infty}^{\infty} \exp(-j\omega t_i)\,\exp\left[-\frac{(t_i - i\mu)^2}{2i\sigma^2}\right]\,dt_i.\tag{7.7}$$

Substituting $t_i - i\mu = r$, where r is a temporary variable, we get

$$E[\exp(-j\omega t_i)] = \frac{1}{\sqrt{2\pi i}\sigma}\,\exp(-j\omega i\mu) \int_{-\infty}^{\infty} \exp\left[-\frac{r^2}{2i\sigma^2}\right]\exp(-j\omega r)\,dr.\tag{7.8}$$

Given that the Fourier transform of $\exp(-\frac{t^2}{2\sigma^2})$ is $\sigma\sqrt{2\pi}\exp(-\frac{\sigma^2\omega^2}{2})$ [1], it follows that

$$E[\exp(-j\omega t_i)] = \exp(-j\omega i\mu)\ \exp\left[-\frac{i\sigma^2\omega^2}{2}\right]. \tag{7.9}$$

Finally, we have

$$\overline{X}(\omega) = \sum_{i=0}^{N} \exp(-j\omega i\mu)\ \exp\left[-\frac{i\sigma^2\omega^2}{2}\right]. \tag{7.10}$$

(In the equations in the present context, i is an index and $j = \sqrt{-1}$.) The ensemble-averaged Fourier transform of the SMUAP train is given by

$$\overline{Y}(\omega) = \overline{X}(\omega)H(\omega), \tag{7.11}$$

where $H(\omega)$ is the Fourier transform of an individual SMUAP. The Fourier transform of an SMUAP train is, therefore, a multiplicative combination of the Fourier transform of the point process representing the motor neuron firing sequence and the Fourier transform of an individual SMUAP.

Illustration of application to EMG: Figure 7.3 illustrates EMG signals synthesized using the point-process model as above using $1, 20, 40$, and 60 motor units, all with the same biphasic SMUAP of $8\ ms$ duration and IPI statistics $\mu = 50\ ms$ and $\sigma = 6.27\ ms$ [339]. It is seen that the EMG signal complexity increases as more motor units are activated. The interference patterns resemble real EMG signals but obscure the shape of the SMUAP used to generate the signals.

Figure 7.4 shows the magnitude spectra of two synthesized EMG signals, one with one active motor unit and the other with 15 active motor units, with biphasic SMUAP duration of $8\ ms$, $\mu = 20\ ms$, and $\sigma = 4.36\ ms$ [339]. The smooth curve superimposed on the second spectrum in Figure 7.4 was derived from the mathematical model described in the preceding paragraphs. An important point to observe from the spectra is that the average magnitude spectrum of several identical motor units approaches the spectrum of a single MUAP. The spectral envelope of an SMUAP train or that of an interference pattern of several SMUAP trains with identical SMUAP waveshape is determined by the shape of an individual SMUAP.

Figure 7.5 shows the magnitude spectra of surface EMG signals recorded from the gastrocnemius-soleus muscle, averaged over $1, 5$, and 15 signal records [339]. The spectra in Figures 7.4 and 7.5 demonstrate comparable features. If all of the motor units active in a composite EMG record were to have similar or identical MUAPs, the spectral envelope of the signal could provide information on the MUAP waveshape (via an inverse Fourier transform). As we have noted earlier in Section 1.2.4, MUAP shape could be useful in the diagnosis of neuromuscular diseases. In reality, however, many motor units of different MUAP shapes could be contributing to the EMG signal even at low levels of effort, and analysis as above may have limited applicability. Regardless, the point-process model provides an interesting approach to model EMG signals. The same model is applicable to the generation of voiced-speech signals, as illustrated in Figure 7.2. For other models to represent the characteristics of
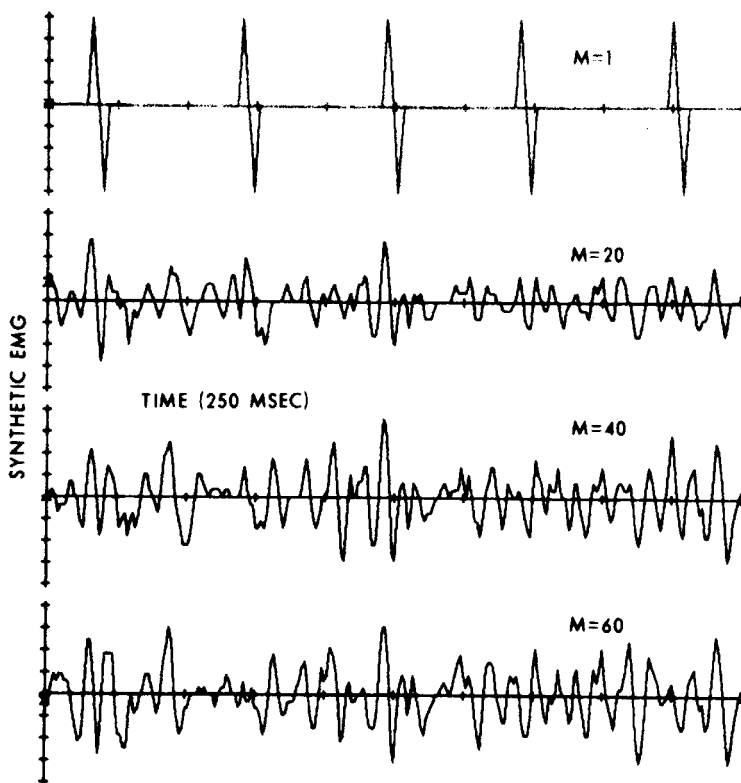
Figure 7.3   Synthesis of an SMUAP train and EMG interference pattern using the point-process model.   Top to bottom: SMUAP train of a single motor unit, and interference patterns of the activities of $20, 40$, and 60 motor units. SMUAP duration $= 8\ ms$. IPI statistics $\mu\ =\ 50\ ms$ and $\sigma\ =\ 6.27\ ms$.   The duration of each signal is $250\ ms$. Reproduced with permission from G.C. Agarwal and G.L. Gottlieb, An analysis of the electromyogram by Fourier, simulation and experimental techniques, *IEEE Transactions on Biomedical Engineering,* 22(3):225–229, 1975. ©IEEE.

the EMG signal, refer to the papers by Parker et al. [348], Lindström and Magnusson [349], Zhang et al. [350], Parker and Scott [351], Shwedyk et al. [352], Person and Libkind [353], Person and Kudina [354], de Luca [37, 355], Lawrence and de Luca [263], and de Luca and van Dyk [356].

Illustration of application to PPC: Zhang et al. [344] proposed a point-process model to represent knee-joint PPC signals, which they called PFP trains or signals (see Section 7.2.4). Figure 7.6 illustrates the PSDs of two point processes simulated with mean repetition rate $\mu_r = 21\ pps$ and coefficient of variation $CV_r = \sigma_r/\mu_r = 0.1$ and 0.05, where σ_r is the SD of the repetition rate. A Gaussian distribution was used to model the IPI statistics. The spectra clearly show the most-dominant peak at the mean repetition rate of the point process, followed by smaller peaks at its
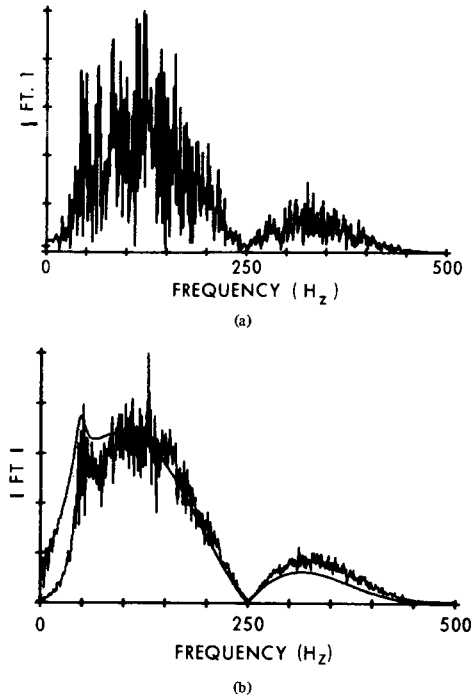
Figure 7.4 Magnitude spectra of synthesized EMG signals with (a) one motor unit, and (b) 15 motor units, with biphasic SMUAP duration of $8\ ms$, $\mu = 20\ ms$, and $\sigma = 4.36\ ms$. The smooth curve superimposed on the spectrum in (b) was derived from the point-process model with 10 SMUAPs. Reproduced with permission from G.C. Agarwal and G.L. Gottlieb, An analysis of the electromyogram by Fourier, simulation and experimental techniques, *IEEE Transactions on Biomedical Engineering,* 22(3):225–229, 1975. ©IEEE.

harmonics. The higher-order harmonics are better defined in the case with the lower CV_r; in the limit, the PSD will be a periodic impulse train with all impulses of equal strength when the point process is exactly periodic ($\sigma_r = 0, CV_r = 0$).

Zhang et al. [344] simulated PFP trains for different IPI statistics using a sample PFP waveform from a real VAG signal recorded at the patella of a normal subject using an accelerometer. The duration of the PFP waveform was $21\ ms$, and the IPI statistics μ_r and CV_r were limited such that the PFP trains synthesized would have nonoverlapping PFP waveforms and resemble real PFP signals. Figures 7.7 and 7.8 illustrate the PSDs of synthesized PFP signals for different μ_r but with the same CV_r, and for the same μ_r but with different CV_r, respectively. The PSDs clearly illustrate the influence of IPI statistics on the spectral features of signals generated by point processes. Some important observations to be made are:

- The PSD envelope of the PFP train remains the same, regardless of the IPI statistics.
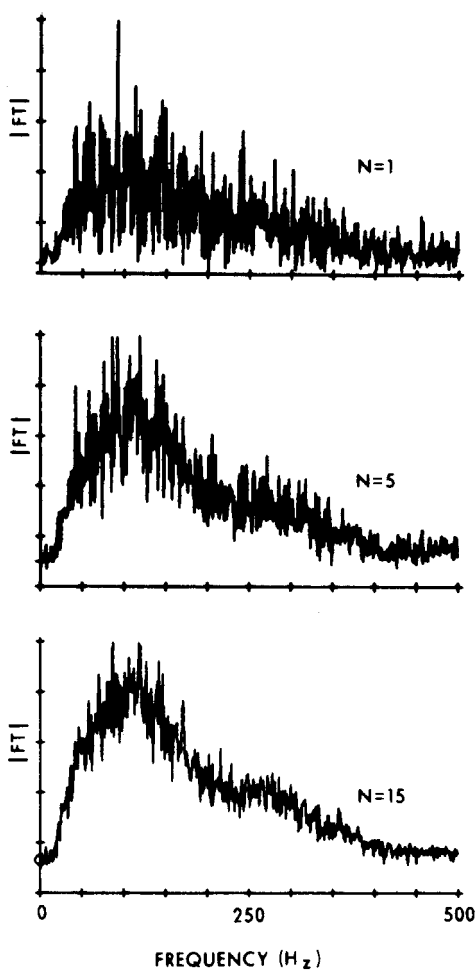
Figure 7.5   Magnitude spectra of surface EMG signals recorded from the gastrocnemius-soleus muscle, averaged over $1, 5$, and 15 signal records. Reproduced with permission from G.C. Agarwal and G.L. Gottlieb, An analysis of the electromyogram by Fourier, simulation and experimental techniques, *IEEE Transactions on Biomedical Engineering,* 22(3):225–229, 1975. ©IEEE.
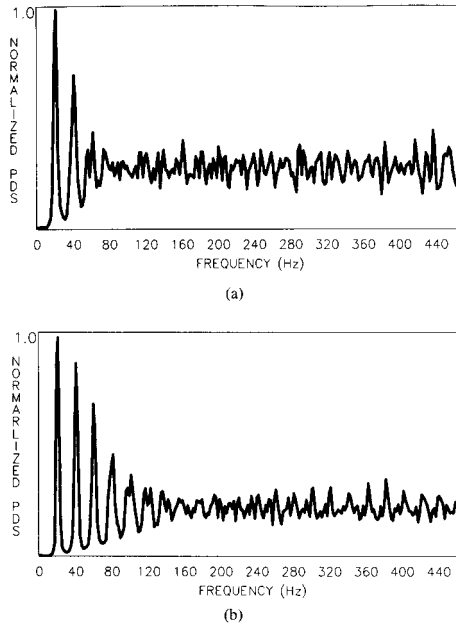
Figure 7.6 Normalized PSDs of synthesized point processes with (a) $\mu_r = 21$ *pps* and $CV_r = 0.1$, and (b) $\mu_r = 21$ *pps* and $CV_r = 0.05$. *Note:* PDS represents the power density spectrum, which is the same as the PSD. Reproduced with permission from Y.T. Zhang, C.B. Frank, R.M. Rangayyan, and G.D. Bell, Mathematical modelling and spectrum analysis of the physiological patello-femoral pulse train produced by slow knee movement, *IEEE Transactions on Biomedical Engineering,* 39(9):971–979, 1992. ©IEEE.

- The PSD envelope of the PFP train is determined by the PSD of an individual PFP waveform.

- The PSD envelope of the PFP train is modulated by a series of impulses with characteristics determined by the IPI statistics. The first impulse indicates the mean repetition rate.

- The point process has a highpass effect: low-frequency components of the PSD of the basic PFP are suppressed due to multiplication with the PSD of the point process.

- Physiological signals rarely exhibit precise periodicity. The CV_r value will be reasonably large, thereby limiting the effect of repetition to low frequencies in the PSD of the PFP train.

The observations made above are, in general, valid for all signals generated by point processes, including SMUAP trains and voiced-speech signals.

Zhang et al. [344] verified the point-process model for PFP signals by computing the IPI statistics and PSDs of real PFP signals recorded from normal subjects.

Figure 7.9 shows the IPI histograms computed from the PFP signals of two normal subjects. The IPI statistics computed for the two cases are $\mu_r = 25.2$ *pps* and $CV_r = 0.07$ for the first signal, and $\mu_r = 16.1$ *pps* and $CV_r = 0.25$ for the second signal. While the IPI histogram for the first signal appears to be close to a Gaussian distribution, the second is not. The PSDs of the two signals are shown in Figure 7.10. The PSDs of the real signals demonstrate features that are comparable to those observed from the PSDs of the synthesized signals and agree with the observations listed above. The envelopes of the two PSDs demonstrate minor variations: The basic PFP waveform in the two cases were not identical.

7.4   Parametric System Modeling

The importance of spectral analysis of biomedical signals is demonstrated in Chapter 6. The methods described are based on the computation and use of the Fourier spectrum; while this approach is, to begin with, nonparametric, a few parameters can be computed from Fourier spectra. The limitations of such an approach are also discussed in Chapter 6. We shall now study methods for parametric modeling and analysis that, although based on time-domain data and models at the outset, can facilitate parametric characterization of the spectral properties of signals and systems.

Problem: *Explore the possibility of parametric modeling of signal characteristics using the general linear system model.*

Solution: The difference equation that gives the output of a general discrete-time LSI system is

$$y(n) = -\sum_{k=1}^{P} a_k \, y(n-k) + G \sum_{l=0}^{Q} b_l \, x(n-l), \qquad (7.12)$$

with $b_0 = 1$. (*Note:* The advantage of the negative sign before the summation with a_k will become apparent later in this section; some model formulations use a positive sign, which does not make any significant difference in the rest of the derivation.) The input to the system is $x(n)$; the output is $y(n)$; the parameters $b_l, l = 0, 1, 2, \ldots, Q$, indicate how the present and Q past samples of the input are combined, in a linear manner, to generate the present output sample; the parameters $a_k, k = 1, 2, \ldots, P$, indicate how the past P samples of the output are linearly combined (in a feedback loop) to produce the current output; G is a gain factor; and P and Q determine the order of the system. The summation over x represents the *moving-average* or MA part of the system; the summation over y represents the *autoregressive* or AR part of the system; the entire system may be viewed as a combined *autoregressive, moving-average* or ARMA system. The feedback part typically makes the impulse response of the system infinitely long; the system may then be viewed as an IIR filter (see Figures 3.59 and 3.60).

Equation 7.12 indicates that the output of the system is simply a linear combination of the present input sample, a few past input samples, and a few past output samples. The use of the past input and output samples in computing the present output sample represents the memory of the system. The model also indicates that the
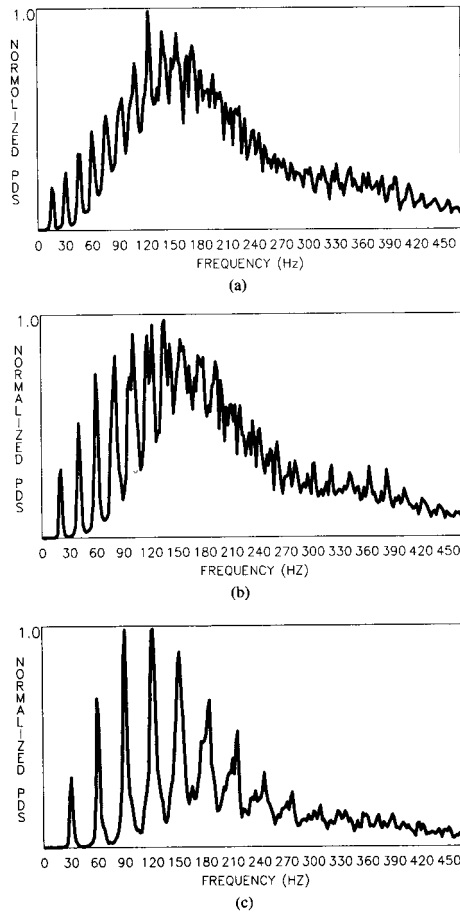
Figure 7.7 Normalized PSDs of synthesized PFP trains using a real PFP waveform with a duration of $21\ ms$, $CV_r = 0.05$, and (a) $\mu_r = 16\ pps$, (b) $\mu_r = 21\ pps$, and (c) $\mu_r = 31\ pps$. *Note:* PDS represents the power density spectrum, which is the same as the PSD. Reproduced with permission from Y.T. Zhang, C.B. Frank, R.M. Rangayyan, and G.D. Bell, Mathematical modelling and spectrum analysis of the physiological patello-femoral pulse train produced by slow knee movement, *IEEE Transactions on Biomedical Engineering,* 39(9):971–979, 1992. ©IEEE.
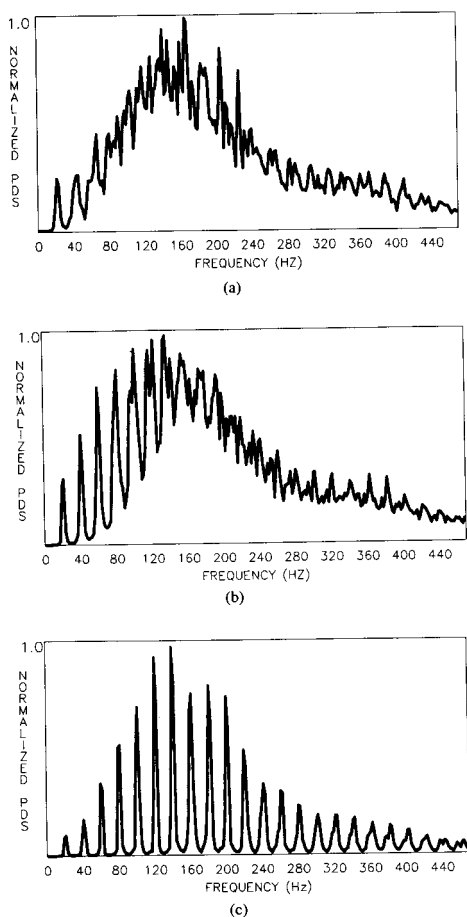
Figure 7.8   Normalized PSDs of synthesized PFP trains using a real PFP waveform with a duration of $21\ ms$, $\mu_r = 21\ pps$, and (a) $CV_r = 0.1$, (b) $CV_r = 0.05$, and (c) $CV_r = 0.01$. *Note:* PDS represents the power density spectrum, which is the same as the PSD. Reproduced with permission from Y.T. Zhang, C.B. Frank, R.M. Rangayyan, and G.D. Bell, Mathematical modelling and spectrum analysis of the physiological patello-femoral pulse train produced by slow knee movement, *IEEE Transactions on Biomedical Engineering,* 39(9):971–979, 1992. ©IEEE.
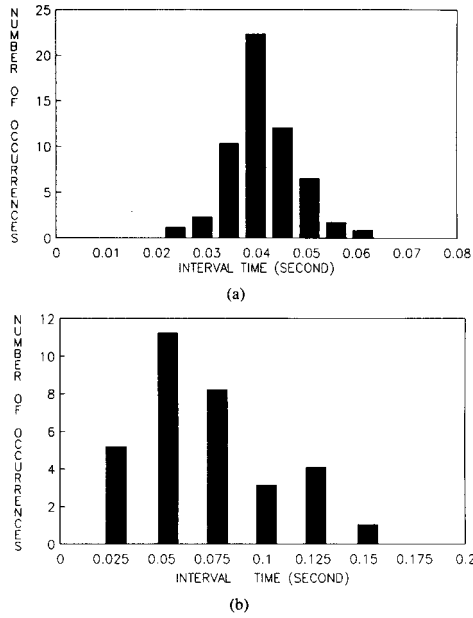
Figure 7.9   IPI histograms computed from real PFP trains recorded from two normal subjects. The statistics computed were (a) $\mu_r = 25.2$ *pps* and $CV_r = 0.07$, and (b) $\mu_r = 16.1$ *pps* and $CV_r = 0.25$. See also Figure 7.10. Reproduced with permission from Y.T. Zhang, C.B. Frank, R.M. Rangayyan, and G.D. Bell, Mathematical modelling and spectrum analysis of the physiological patello-femoral pulse train produced by slow knee movement, *IEEE Transactions on Biomedical Engineering,* 39(9):971–979, 1992. ©IEEE.

present output sample may be *predicted* as a *linear combination* of the present and a few past input samples, and a few past output samples. For this reason, the model is also known as the *linear prediction* or LP model [79, 148, 357, 358].

Applying the z-transform to Equation 7.12, we can obtain the transfer function of the system as

$$H(z) = \frac{Y(z)}{X(z)} = G \; \frac{1 + \sum_{l=1}^{Q} b_l \, z^{-l}}{1 + \sum_{k=1}^{P} a_k \, z^{-k}} \; . \tag{7.13}$$

(The advantage of the negative sign before the summation with a_k in Equation 7.12 is now apparent in the numerator – denominator symmetry of Equation 7.13.) The system is completely characterized by the parameters $a_k, k = 1, 2, \ldots, P$; $b_l, l = 1, 2, \ldots, Q$; and G. In most applications, the gain factor G is not important; the system is, therefore, completely characterized by the a and b parameters, with the exception of a gain factor. Furthermore, we may factorize the numerator and denominator polynomials in Equation 7.13 and express the transfer function as

$$H(z) = G \; \frac{\prod_{l=1}^{Q} (1 - z_l \, z^{-1})}{\prod_{k=1}^{P} (1 - p_k \, z^{-1})} \, , \tag{7.14}$$
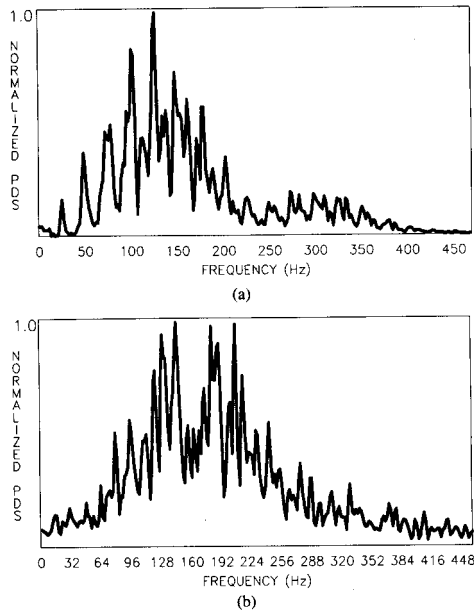
Figure 7.10   Normalized PSDs of the real PFP trains recorded from two normal subjects whose IPI histograms are shown in Figure 7.9. The IPI statistics of the two cases are (a) $\mu_r = 25.2$ *pps* and $CV_r = 0.07$, and (b) $\mu_r = 16.1$ *pps* and $CV_r = 0.25$. *Note:* PDS represents the power density spectrum, which is the same as the PSD. Reproduced with permission from Y.T. Zhang, C.B. Frank, R.M. Rangayyan, and G.D. Bell, Mathematical modelling and spectrum analysis of the physiological patello-femoral pulse train produced by slow knee movement, *IEEE Transactions on Biomedical Engineering,* 39(9):971–979, 1992. ©IEEE.

where $z_l, l = 1, 2, \ldots, Q$, are the zeros of the system and $p_k, k = 1, 2, \ldots, P$, are the poles of the system. The model may now be referred to as a *pole–zero model*. It is evident from Equation 7.14 that the system is completely characterized by its poles and zeros, but for a gain factor.

Equations 7.12, 7.13, and 7.14 demonstrate the applicability of the same conceptual model in the time and frequency domains. The a and b parameters are directly applicable in both the time and the frequency domains in expressing the input–output relationship or the system transfer function. The poles and zeros are more specific to the frequency domain, although the contribution of each pole or zero to the time-domain impulse response of the system may be derived directly from its coordinates in the z-plane [1].

Given a particular input signal $x(n)$ and the corresponding output of the system $y(n)$, we could derive their z-transforms $X(z)$ and $Y(z)$, and subsequently obtain the system transfer function $H(z)$ in some form. Difficulties arise at values of z for which $X(z) = 0$; as the system is linear and $Y(z) = H(z)X(z)$, we have $Y(z) = 0$ at such points as well. Then, $H(z)$ cannot be determined at the corresponding val-

ues of z. [The ideal test signal is the unit-impulse function $x(n) = \delta(n)$, for which $X(z) = 1$ for all z: The response of an LSI system to an impulse completely characterizes the system with the corresponding $y(n) = h(n)$ or its z-domain equivalent $H(z)$.] Methods to determine an AR or ARMA model for a given signal for which the corresponding input to the system is not known (or is assumed to be a point process or a random process) are described in the following sections.

7.5   Autoregressive or All pole Modeling

Problem: *How can we obtain an AR (or LP) model when the input to the system that caused the given signal as its output is unknown?*

   Solution: In the AR or all-pole model [79, 357], the output is modeled as a linear combination of P past values of the output and the present input sample as

$$y(n) = -\sum_{k=1}^{P} a_k \, y(n-k) + G \, x(n).$$

(7.15)

(The discussion on AR modeling here closely follows that in Makhoul [357], with permission.) Some model formulations use a positive sign in place of the negative sign before the summation in the equation given above. It should be noted that the model in Equation 7.15 does not account for the presence of noise.

   The all-pole transfer function corresponding to Equation 7.15 is

$$H(z) = \frac{G}{1 + \sum_{k=1}^{P} a_k \, z^{-k}}.$$

(7.16)

In the case of biomedical signals such as the EEG or the PCG, the input to the system is totally unknown. Then, we can only approximately predict the current sample of the output signal using its past values as

$$\tilde{y}(n) = -\sum_{k=1}^{P} a_k \, y(n-k),$$

(7.17)

where the $\tilde{\ }$ symbol indicates that the predicted value is only approximate. The error in the predicted value (also known as the residual) is

$$e(n) = y(n) - \tilde{y}(n) = y(n) + \sum_{k=1}^{P} a_k \, y(n-k).$$

(7.18)

The general signal-flow diagram of the AR model viewed as a prediction or error filter is illustrated in Figure 7.11.

   The least-squares method: In the least-squares method, the parameters a_k are obtained by minimizing the MSE with respect to all of the parameters. The procedure is similar to that used to derive the optimal Wiener filter (see Section 3.8 and Haykin [148]).
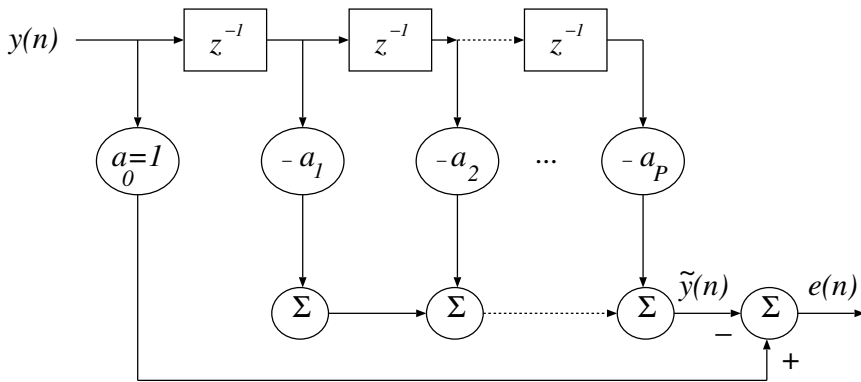
Figure 7.11   Signal-flow diagram of the AR model.

Given an observed signal $y(n)$, the following procedure is applicable for minimization of the MSE [357]. The total squared error (TSE) ε is given by

$$\varepsilon = \sum_n e^2(n) = \sum_n \left(y(n) + \sum_{k=1}^{P} a_k\, y(n-k) \right)^2. \qquad (7.19)$$

(*Note:* The TSE is the same as the MSE except for a scale factor.) Although the range of the summation in Equation 7.19 is important, we may minimize ε without specifying the range for the time being. Minimization of ε is performed by applying the conditions

$$\frac{\partial \varepsilon}{\partial a_k} = 0, \;\; 1 \leq k \leq P \qquad (7.20)$$

to Equation 7.19, which yields

$$\sum_{k=1}^{P} a_k \sum_n y(n-k)\, y(n-i) = -\sum_n y(n)\, y(n-i), \;\; 1 \leq i \leq P. \qquad (7.21)$$

For a given signal $y(n)$, Equation 7.21 provides a set of P equations in the P unknowns $a_k, k = 1, 2, \ldots, P$, known as the *normal equations*; the similarities between the normal equations here and those in the case of the Wiener filter (see Section 3.8) will become apparent later.

By expanding Equation 7.19 and using the relationship in Equation 7.21, the minimum TSE ε_P for the model of order P is obtained as

$$\varepsilon_P = \sum_n y^2(n) + \sum_{k=1}^{P} a_k \sum_n y(n)y(n-k). \qquad (7.22)$$

The expression for TSE will be simplified in the following sections.

The autocorrelation method: If the range of summation in Equations 7.19 and 7.21 is specified to be $-\infty < n < \infty$, the error is minimized over an infinite duration, and we have

$$\phi_y(i) = \sum_{n=-\infty}^{\infty} y(n)\, y(n-i), \tag{7.23}$$

where $\phi_y(i)$ is the ACF of $y(n)$. In practice, the signal $y(n)$ will be available only over a finite interval, such as $0 \le n \le N-1$; the given signal may then be assumed to be zero outside this range and treated as a windowed version of the true signal, as we have seen in Section 6.3. Then, the ACF may be expressed as

$$\phi_y(i) = \sum_{n=i}^{N-1-i} y(n)\, y(n-i), \quad i \ge 0, \tag{7.24}$$

where the scale factor $\frac{1}{N}$ is omitted. (It will become apparent later that the scale factor is immaterial in the derivation of the model coefficients.) The normal equations then become

$$\sum_{k=1}^{P} a_k\, \phi_y(i-k) = -\phi_y(i), \quad 1 \le i \le P. \tag{7.25}$$

We now see that an AR model may be derived for a signal with the knowledge of only its ACF; the signal samples themselves are not required. It is also obvious now that the scale factor $\frac{1}{N}$ that was omitted in defining the ACF is of no consequence in deriving the model coefficients. It may be advantageous to use the normalized ACF values given as $\overline{\phi}_y(i) = \phi_y(i)/\phi_y(0)$, which have the property that $|\overline{\phi}_y(i)| \le 1$.

The minimum TSE is given by

$$\varepsilon_P = \phi_y(0) + \sum_{k=1}^{P} a_k\, \phi_y(k). \tag{7.26}$$

Application to random signals: If the signal $y(n)$ is a sample of a random process, the error $e(n)$ is also a sample of a random process. We then have to use the expectation operation to obtain the MSE as follows:

$$\varepsilon = E[e^2(n)] = E\left[\left(y(n) + \sum_{k=1}^{P} a_k\, y(n-k)\right)^2\right]. \tag{7.27}$$

Applying the condition for minimum error as in Equation 7.20, we get the normal equations as

$$\sum_{k=1}^{P} a_k\, E[y(n-k)\, y(n-i)] = -E[y(n)\, y(n-i)], \quad 1 \le i \le P. \tag{7.28}$$

The minimum MSE is

$$\varepsilon_P = E[y^2(n)] + \sum_{k=1}^{P} a_k \, E[y(n) \, y(n-k)]. \tag{7.29}$$

If the signal is a sample of a stationary random process, we have $E[y(n-k) \, y(n-i)] = \phi_y(i-k)$. This leads to the same normal equations as in Equation 7.25. If the process is ergodic, the ACF may be computed as a time average as in Equation 7.24.

If the signal is a sample of a nonstationary random process, $E[y(n-k) \, y(n-i)] = \phi_y(n-k, n-i)$; the ACF is a function of not only the shift but also time. We would then have to compute the model parameters for every instant of time n; we then have a time-variant model. Methods for modeling and analysis of nonstationary signals are described in Chapter 8.

Computation of the gain factor G**:** Since we assumed earlier that the input to the system being modeled is unknown, the gain parameter G is not important. Regardless, the derivation of G demonstrates a few important points. Equation 7.18 may be rewritten as

$$y(n) = -\sum_{k=1}^{P} a_k \, y(n-k) + e(n). \tag{7.30}$$

Comparing this with Equation 7.15, we see that the only input signal $x(n)$ which can result in $y(n)$ at the output is given by the condition $Gx(n) = e(n)$. This condition indicates that the input signal is proportional to the error of prediction when the estimated model parameters are equal to the real system parameters. Regardless of the input, a condition that could be applied is that the energy of the output be equal to that of the signal $y(n)$ being modeled. Since the transfer function $H(z)$ is fixed, we then have the condition that the total energy of the input signal $Gx(n)$ be equal to the total energy of the error ε_P.

As illustrated in the model for speech generation in Figure 7.2, two types of input that are of interest are the impulse function and a random process that is stationary white noise. In the case when $x(n) = \delta(n)$, we have the impulse response $h(n)$ at the output, and

$$h(n) = -\sum_{k=1}^{P} a_k \, h(n-k) + G \, \delta(n). \tag{7.31}$$

Multiplying both sides of the expression above with $h(n-i)$ and summing over all n, we get expressions in terms of the ACF $\phi_h(i)$ of $h(n)$ as

$$\phi_h(i) = -\sum_{k=1}^{P} a_k \, \phi_h(i-k), \quad 1 \le |i| \le \infty \tag{7.32}$$

and

$$\phi_h(0) = -\sum_{k=1}^{P} a_k \, \phi_h(k) + G^2. \tag{7.33}$$

Due to the condition that the total energy of $h(n)$ be equal to that of $y(n)$, the condition $\phi_h(0) = \phi_y(0)$ must be satisfied. Comparing Equations 7.32 and 7.25, we then have

$$\phi_h(i) = \phi_y(i), \ \ 0 \leq i \leq P. \tag{7.34}$$

Therefore, for a model of order P, the first $(P+1)$ ACF terms of the impulse response $h(n)$ must be equal to the corresponding ACF terms of the signal $y(n)$ being modeled. It follows from Equations 7.33, 7.34, and 7.26 that

$$G^2 = \varepsilon_P = \phi_y(0) + \sum_{k=1}^{P} a_k \ \phi_y(k). \tag{7.35}$$

In the case when the input is a sequence of uncorrelated samples of a random process (white noise) with zero mean and unit variance, we could use the same procedure as for the impulse-input case, with the difference being that expectations are taken instead of summing over all n. {The conditions to be noted in this case are $E[x(n)] = 0$ and $E[x(n) \ x(n-i)] = \delta(i)$.} The same relations as above for the impulse-input case are obtained. The identical nature of the results for the two cases follows from the fact that the two types of input have identical ACFs and PSDs. These characteristics are relevant in the speech model shown in Figure 7.2.

Computation of the model parameters: For low orders of the model, Equation 7.25 may be solved directly. However, direct methods may not be feasible when P is large.

The normal equations in Equation 7.25 may be written in matrix form as

$$\begin{bmatrix} \phi_y(0) & \phi_y(1) & \cdots & \phi_y(P-1) \\ \phi_y(1) & \phi_y(0) & \cdots & \phi_y(P-2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_y(P-1) & \phi_y(P-2) & \cdots & \phi_y(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = - \begin{bmatrix} \phi_y(1) \\ \phi_y(2) \\ \vdots \\ \phi_y(P) \end{bmatrix}. \tag{7.36}$$

For real signals, the $P \times P$ ACF matrix is symmetric and the elements along any diagonal are identical, that is, it is a Toeplitz matrix.

It is worth noting the following similarities and differences between the normal equations in the case of the Wiener filter as given in Equation 3.169 and those given above in Equation 7.36:

- The matrix on the LHS is the ACF of the input to the filter in the case of the Wiener filter, whereas it is the ACF of the output of the prediction filter in the present case.

- The filter vector on the LHS contains the coefficients of the filter being designed in both cases — the optimal Wiener filter or the optimal prediction filter.

- The vector on the RHS is the CCF between the input and the desired response in the case of the Wiener filter, whereas it is the ACF of the output of the prediction filter in the present case.

Haykin [148] provides a more detailed correspondence between the AR model and the Wiener filter.

A procedure known as Durbin's method [359, 360] or the Levinson–Durbin algorithm (see Makhoul [357], Rabiner and Schafer [79], or Haykin [148]) provides a recursive method to solve the normal equations in Equation 7.36. The procedure starts with a model order of 1; computes the model parameters, the error, and a secondary set of parameters known as the reflection coefficients; updates the model order and the parameters; and repeats the procedure until the model of the desired order is obtained. The Levinson–Durbin algorithm is summarized below.

Initialize model order $i = 0$ and error $\varepsilon_0 = \phi_y(0)$. Perform the following steps recursively for $i = 1, 2, \ldots, P$.

1. Increment model order i and compute the ith reflection coefficient γ_i as

$$\gamma_i = -\frac{1}{\varepsilon_{i-1}} \left[\phi_y(i) + \sum_{j=1}^{i-1} a_{i-1,j} \, \phi_y(i-j) \right], \qquad (7.37)$$

   where $a_{i-1,j}$ denotes the jth model coefficient at iteration $(i-1)$; the iteration index is also the recursively updated model order.

2. Let $a_{i,i} = \gamma_i$.

3. Update the predictor coefficients as

$$a_{i,j} = a_{i-1,j} + \gamma_i \, a_{i-1,i-j}, \quad 1 \le j \le i - 1. \qquad (7.38)$$

4. Compute the error value as

$$\varepsilon_i = (1 - \gamma_i^2) \, \varepsilon_{i-1}. \qquad (7.39)$$

The final model parameters are given as $a_k = a_{P,k}, \ 1 \le k \le P$. The Levinson–Durbin algorithm computes the model parameters for all orders up to the desired order P. As the order of the model is increased, the TSE reduces, and hence we have $0 \le \varepsilon_i \le \varepsilon_{i-1}$. The reflection coefficients may also be used to test the stability of the model (filter) being designed: $|\gamma_i| < 1, i = 1, 2, \ldots, P$, is the required condition for stability of the model of order P.

The covariance method: In deriving the autocorrelation method, the range of summation of the prediction error in Equations 7.19 and 7.21 was specified to be $-\infty < n < \infty$. If, instead, we specify the range of summation to be a finite interval, such as $0 \le n \le N - 1$, we get

$$\sum_{k=1}^{P} a_k \, C(k, i) = -C(0, i), \quad 1 \le i \le P \qquad (7.40)$$

instead of Equation 7.25 based upon the ACF, and the minimum TSE is given by

$$\varepsilon_P = C(0, 0) + \sum_{k=1}^{P} a_k \, C(0, k) \qquad (7.41)$$

instead of Equation 7.26, where

$$C(i, k) = \sum_{n=0}^{N-1} y(n - i)y(n - k) \tag{7.42}$$

is the covariance of the signal $y(n)$ in the specified interval. The matrix formed by the covariance function is symmetric as $C(i, k) = C(k, i)$, similar to the ACF matrix in Equation 7.36; however, the elements along each diagonal will not be equal, as $C(i + 1, k + 1) = C(i, k) + y(-i - 1)y(-k - 1) - y(N - 1 - i)y(N - 1 - k)$. Computation of the covariance coefficients also requires $y(n)$ to be known for $-P \leq n \leq N - 1$. The distinctions disappear as the specified interval of summation (error minimization) tends to infinity.

7.5.1   Spectral matching and parameterization

The AR model was derived in the preceding section based upon time-domain formulations in the autocorrelation and covariance methods. We shall now see that equivalent formulations can be derived in the frequency domain, which can lead to a different interpretation of the model. Applying the z-transform to Equation 7.18, we get

$$E(z) = \left[1 + \sum_{k=1}^{P} a_k z^{-k}\right] Y(z) = A(z)Y(z) \tag{7.43}$$

and

$$H(z) = \frac{G}{A(z)} , \tag{7.44}$$

where

$$A(z) = 1 + \sum_{k=1}^{P} a_k z^{-k} , \tag{7.45}$$

and $E(z)$ is the z-transform of $e(n)$. We can now view the error $e(n)$ as the result of passing the signal being modeled $y(n)$ through the filter $A(z)$, which may be considered to be an *inverse filter*. In the case of $y(n)$ being a deterministic signal, applying Parseval's theorem, the TSE to be minimized may be written as

$$\varepsilon = \sum_{n=-\infty}^{\infty} e^2(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(\omega)|^2 \, d\omega, \tag{7.46}$$

where $E(\omega)$ is obtained by evaluating $E(z)$ on the unit circle in the z-plane. Using $S_y(\omega)$ to represent the PSD of $y(n)$, we have

$$\varepsilon = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_y(\omega)|A(\omega)|^2 \, d\omega, \tag{7.47}$$

where $A(\omega)$ is the frequency response of the inverse filter and is given by evaluating $A(z)$ on the unit circle in the z-plane.

From Equations 7.16 and 7.44, we get

$$\tilde{S}_y(\omega) = |H(\omega)|^2 = \frac{G^2}{|A(\omega)|^2} = \frac{G^2}{\left|1 + \sum_{k=1}^{P} a_k \, \exp(-jk\omega)\right|^2}. \qquad (7.48)$$

Here, $\tilde{S}_y(\omega)$ represents the PSD of the modeled signal $\tilde{y}(n)$ that is an approximation of $y(n)$ as in Equation 7.17. From Equation 7.43 we have

$$S_y(\omega) = \frac{|E(\omega)|^2}{|A(\omega)|^2}. \qquad (7.49)$$

Now, $\tilde{S}_y(\omega)$ is the model's approximation of $S_y(\omega)$. Comparing Equations 7.48 and 7.49, we see that the error PSD $|E(\omega)|^2$ is modeled by a uniform (or "flat" or "white") PSD equal to G^2. For this reason, the filter $A(z)$ is also known as a "whitening" filter.

From Equations 7.48 and 7.49, we get the TSE as

$$\varepsilon = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{S_y(\omega)}{\tilde{S}_y(\omega)} \, d\omega. \qquad (7.50)$$

As the model is derived by minimizing the TSE ε, we see now that the model is effectively minimizing the integrated ratio of the signal PSD $S_y(\omega)$ to its approximation $\tilde{S}_y(\omega)$. Makhoul [357] describes the equivalence of the model in the following terms:

- As the model order $P \to \infty$, the TSE is minimized, that is, $\varepsilon_P \to 0$.

- For a model of order P, the first $(P+1)$ ACF values of its impulse response are equal to those of the signal being modeled. Increasing P increases the range of the delay parameter (time) over which the model ACF is equal to the signal ACF.

- Given that the PSD and the ACF are Fourier-transform pairs, the preceding point leads to the frequency-domain statement that increasing P leads to a better fit of $\tilde{S}_y(\omega)$ to $S_y(\omega)$. As $P \to \infty$, the model ACF and PSD become identical to the signal ACF and PSD, respectively. Thus, *any spectrum may be approximated by an all-pole model of an appropriate order* (see Section 7.5.2 for a discussion on the optimal model order).

Noting from Equation 7.35 that $G^2 = \varepsilon_P$, Equation 7.50 yields another important property of the model as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{S_y(\omega)}{\tilde{S}_y(\omega)} \, d\omega = 1. \qquad (7.51)$$

Equations 7.50 and 7.51 lead to the following spectral-matching properties of the AR model [357]:

- Due to the fact that the TSE is determined by the ratio of the true PSD to the model PSD, the spectral-matching process performs uniformly over the entire

frequency range irrespective of the spectral shape. (Had the error measure been dependent on the difference between the true PSD and the model PSD, the spectral match would have been better at higher-energy frequency coordinates than at lower-energy frequency coordinates.)

- $S_y(\omega)$ will be greater than $\tilde{S}_y(\omega)$ at some frequencies and lower at others, while satisfying Equation 7.51 on the whole; the contribution to the TSE is more significant when $S_y(\omega) > \tilde{S}_y(\omega)$ than in the opposite case. Thus, when the error is minimized, the fitting of $\tilde{S}_y(\omega)$ to $S_y(\omega)$ is better at frequencies where $S_y(\omega) > \tilde{S}_y(\omega)$. Thus, the model PSD fits better at the *peaks* of the signal PSD.

- The preceding point leads to another interpretation: The AR-model spectrum $\tilde{S}_y(\omega)$ is a good estimate of the *spectral envelope* of the signal PSD. This is particularly useful when modeling quasiperiodic signals such as voiced speech, PCG, and other signals that have strong peaks in their spectra representing harmonics, formants, or resonance. By following the envelope, the effects of repetition, that is, the effects of the point-process excitation function (see Section 7.3), are removed.

Since the model PSD is entirely specified by the model parameters (as in Equation 7.48), we now have a *parametric representation* of the PSD of the given signal (subject to the error in the model). The TSE may be related to the signal PSD as follows [357].

$$\hat{\tilde{y}}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[\tilde{S}_y(\omega)] \, d\omega \qquad (7.52)$$

represents the zeroth coefficient of the (power or real) cepstrum (see Sections 4.8.3 and 5.4.2) of $\tilde{y}(n)$. Using the relationship in Equation 7.48, we get

$$\hat{\tilde{y}}(0) = \log G^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |A(\omega)|^2 \, d\omega. \qquad (7.53)$$

As all the roots (zeros) of $A(z)$ are inside the unit circle in the z-plane (for the AR model to be stable), the integral in the equation given above is zero [357]. We also have $G^2 = \varepsilon_P$. Therefore,

$$\varepsilon_P = \exp[\hat{\tilde{y}}(0)]. \qquad (7.54)$$

The minimum of ε_P is reached as $P \to \infty$, and is given by

$$\varepsilon_{\min} = \varepsilon_\infty = \exp[\hat{y}(0)]. \qquad (7.55)$$

This relationship means that the TSE ε_P is the geometric mean of the model PSD $\tilde{S}_y(\omega)$, which is always positive for a positive-definite PSD. The quantity ε_P represents that portion of the signal's information content that is not predictable by a model of order P.

7.5.2 Optimal model order

Given that the AR model performs better and better as the order P is increased, where do we stop? Makhoul [357] shows that if the given signal is the output of a P-pole system, then an AR model of order P would be the optimal model with the minimum error. But how would one find in practice if the given signal was indeed produced by a P-pole system?

One possibility to determine the optimal order for modeling a given signal is to follow the trend in the TSE as the model order P is increased. This is feasible in a recursive procedure such as the Levinson–Durbin algorithm, where models of all lower orders are computed in deriving a model of order P, and the error at each order is readily available. The procedure could be stopped when there is no significant reduction in the error as the model order is incremented.

Makhoul [357] describes the use of a normalized error measure $\overline{\varepsilon}_P$ defined as

$$\overline{\varepsilon}_P = \frac{\varepsilon_P}{\phi_y(0)} = \frac{\exp[\hat{\tilde{y}}(0)]}{\phi_y(0)}. \tag{7.56}$$

As the model order $P \to \infty$,

$$\overline{\varepsilon}_{\min} = \overline{\varepsilon}_\infty = \frac{\exp[\hat{y}(0)]}{\phi_y(0)}. \tag{7.57}$$

$\overline{\varepsilon}_{\min}$ is a monotonically decreasing function of P, with $\overline{\varepsilon}_0 = 1$ and $\overline{\varepsilon}_\infty = \overline{\varepsilon}_{\min}$; furthermore, it can be expressed as a function of the model PSD as

$$\overline{\varepsilon}_P = \frac{\exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \tilde{S}_y(\omega)\, d\omega\right]}{\frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{S}_y(\omega)\, d\omega}. \tag{7.58}$$

It is evident that $\overline{\varepsilon}_P$ depends only upon the shape of the model PSD and that $\overline{\varepsilon}_{\min}$ is determined solely by the signal PSD. The quantity $\overline{\varepsilon}_P$ may be viewed as the ratio of the geometric mean of the model PSD to its arithmetic mean, which is a measure of the spread of the PSD: The smaller the spread, the closer is the ratio to unity; the larger the spread, the closer is the ratio to zero. If the signal is the result of an all-pole system with P_0 poles, $\overline{\varepsilon}_P = \overline{\varepsilon}_{P_0}$ for $P \geq P_0$, that is, the curve remains flat. In practice, the incremental reduction in the normalized error may be checked with a condition such as

$$1 - \frac{\overline{\varepsilon}_{P+1}}{\overline{\varepsilon}_P} < \Delta, \tag{7.59}$$

where Δ is a small threshold. The optimal order may be considered to have been reached if the condition is satisfied for several consecutive model orders.

A measure based on an information-theoretic criterion proposed by Akaike [361] is expressed as [357]

$$I(P) = \log \overline{\varepsilon}_P + \frac{2P}{N_e}, \tag{7.60}$$

where N_e is the effective number of data points in the signal taking into account windowing (for example, $N_e = 0.4N$ for a Hamming window, where N is the

number of data samples). As P is increased, the first term in the equation given above decreases while the second term increases. Akaike's measure $I(P)$ may be computed up to the maximum order P of interest or the maximum that is feasible, and then the model of the order for which $I(P)$ is at its minimum could be taken as the optimal model.

Model parameters: The AR (all-pole) model $H(z)$ and its inverse $A(z)$ are uniquely characterized by any one set of the following sets of parameters [357]:

- The model parameters $a_k, k = 1, 2, \ldots, P$. The series of a_k parameters is also equal to the impulse response of the inverse filter.

- The impulse response $h(n)$ of the AR model.

- The poles of $H(z)$, which are also the roots (zeros) of $A(z)$.

- The reflection coefficients $\gamma_i, i = 1, 2, \ldots, P$.

- The ACF (or PSD) of the a_k coefficients.

- The ACF (or PSD) of $h(n)$.

- The cepstrum of a_k or $h(n)$.

With the inclusion of the gain factor G as required, all of the above sets have a total of $(P + 1)$ values and are equivalent in the sense that one set may be derived from another. Any particular set of parameters may be used, depending upon its relevance, interpretability, or relationship to the real-world system being modeled.

Illustration of application to EEG signals: Identification of the existence of rhythms of specific frequencies is an important aspect of EEG analysis. The direct relationship between the poles of an AR model and resonance frequencies makes this technique an attractive tool for the analysis of EEG signals.

Figure 7.12 shows the Fourier spectrum and AR-model spectra with $P = 6$ and $P = 10$ for the o1 channel of the EEG signal shown in Figure 1.39. The Fourier spectrum in the lowest trace of Figure 7.12 includes many spurious variations which make its interpretation difficult. On the other hand, the AR spectra indicate distinct peaks at about $10 \ Hz$ corresponding to an alpha rhythm; a peak at $10 \ Hz$ is clearly evident even with a low model order of $P = 6$ (the middle trace in Figure 7.12).

The poles of the AR model with order $P = 10$ are plotted in Figure 7.13. The dominant pole (closest to the unit circle in the z-plane) appears at $9.9 \ Hz$, corresponding to the peak observed in the spectrum in the topmost plot in Figure 7.12. The radius of the dominant pole is $|z| = 0.95$; the other complex-conjugate pole pairs have $|z| \leq 0.76$. The model with $P = 6$ resulted in two complex-conjugate pole pairs and two real poles, with the dominant pair at $10.5 \ Hz$ with $|z| = 0.91$; the magnitude of the other pole pair was 0.74. A simple search for the dominant (complex) pole can provide an indication of the prevalent EEG rhythm with fairly low AR model orders.

Illustration of application to PCG signals: Application of AR modeling is an attractive possibility for the analysis of PCG signals due to the need to identify significant frequencies of resonance in the presence of multiple components, artifacts,
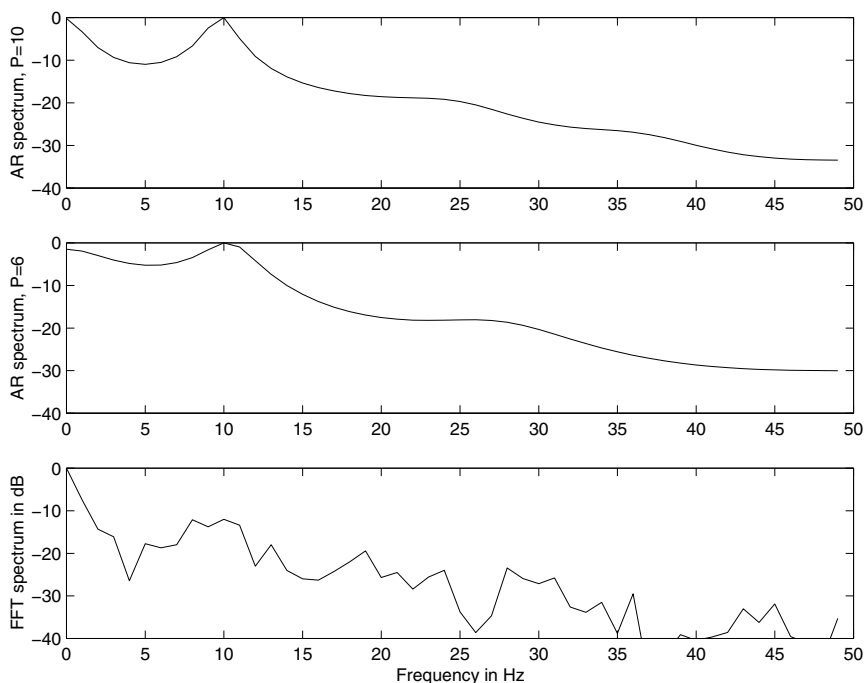
Figure 7.12   From bottom to top: Fourier and AR-model spectra with $P = 6$ and $P = 10$ for the o1 channel of the EEG signal shown in Figure 1.39.

and noise. Although the model coefficients themselves do not carry any physical correlates or significance, the poles may be related directly to the physical or physiological characteristics of hearts sounds and murmurs.

Figures 7.14 and 7.15 illustrate the Fourier spectrum of a segment containing one S1 and the subsequent systolic portion of the PCG signal of a normal subject, the AR-model spectra for order $P = 10$ and $P = 20$, and the poles of the model of order $P = 20$ (see also Figures 4.30, 5.7, and 6.12). Figures 7.16 and 7.17 illustrate the same items for a segment containing one S2 and the subsequent diastolic portion of the same subject. It is evident that the AR spectra follow the dominant peaks in the spectra of the original signals. The spectra for the models of order $P = 20$ provide closer fits than those for $P = 10$; peaks in the $P = 10$ spectra gloss over multiple peaks in the original spectra. Observe the presence of poles close to the unit circle in the z-plane at frequencies corresponding to the peaks in the spectra of the signals. The AR-model spectra are smoother and easier to interpret than the periodogram-based spectra illustrated in Figure 6.12 for the same subject. The spectra for the diastolic portion indicate more medium-frequency energy than those for the systolic portion, as expected. The model coefficients or poles provide a compact parametric representation of the signals and their spectra.
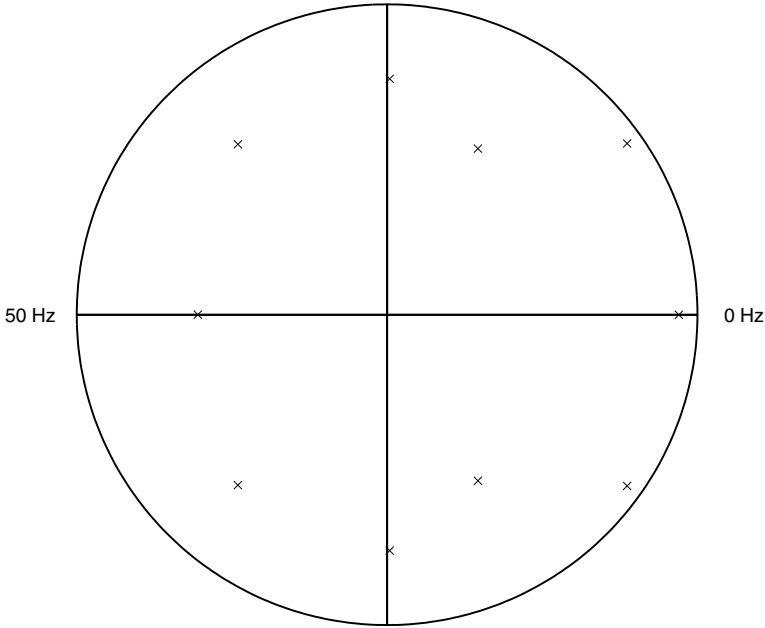
Figure 7.13   Poles of the AR model with $P = 10$ for the o1 channel of the EEG signal shown in Figure 1.39. See also Figure 7.12.

Figures 7.18 and 7.19 illustrate the Fourier spectrum of a segment containing one S1 and the subsequent systolic portion of the PCG signal of a subject with systolic murmur, split S2, and opening snap of the mitral valve (see also Figures 4.31, 5.8, and 6.13); the AR-model spectra for order $P = 10$ and $P = 20$; and the poles of the model of order $P = 20$. Figures 7.20 and 7.21 illustrate the same items for a segment containing one S2 and the subsequent diastolic portion of the same subject. The systolic murmur has given rise to more medium-frequency components than in the case of the normal subject in Figure 7.14. The AR-model spectra clearly indicate additional and stronger peaks at $150\ Hz$ and $250\ Hz$, which are confirmed by poles close to the unit circle at the corresponding frequencies in Figure 7.19.

7.5.3   AR and cepstral coefficients

If the poles of $H(z)$ are inside the unit circle in the complex z-plane, from the theory of complex variables, $\ln H(z)$ can be expanded into a Laurent series as

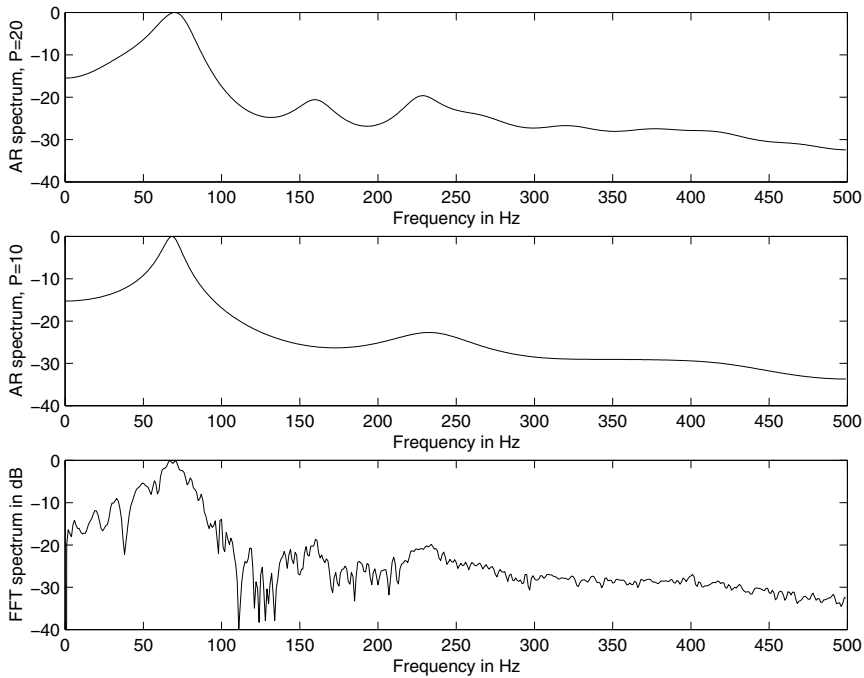$$\ln H(z) = \sum_{n=1}^{\infty} \hat{h}(n)\, z^{-n}. \tag{7.61}$$

Figure 7.14 Bottom to top: Fourier spectrum of the systolic portion of the PCG of a normal subject (male, 23 years); AR-model spectrum with order $P = 10$; AR-model spectrum with order $P = 20$. (See also Figures 4.30, 5.7, and 6.12.)

Considering the definition of the complex cepstrum as the inverse z-transform of the logarithm of the z-transform of the signal, and the fact that the LHS of the equation given above represents the z-transform of $\hat{h}(n)$, it is clear that the coefficients of the series $\hat{h}(n)$ are the cepstral coefficients of $h(n)$. If $H(z)$ has been approximated by an AR model with coefficients a_k, $1 \leq k \leq P$, we have

$$\ln\left(\frac{1}{1 + \sum_{k=1}^{P} a_k\, z^{-k}}\right) = \sum_{n=1}^{\infty} \hat{h}(n)\, z^{-n}. \tag{7.62}$$

Differentiating both sides of the equation given above with respect to z^{-1}, we get

$$\frac{-\left(\sum_{k=1}^{P} k\, a_k\, z^{-k+1}\right)}{1 + \sum_{k=1}^{P} a_k\, z^{-k}} = \sum_{n=1}^{\infty} n\, \hat{h}(n)\, z^{-n+1}, \tag{7.63}$$

or

$$-\sum_{k=1}^{P} k\, a_k\, z^{-k+1} = \left(1 + \sum_{k=1}^{P} a_k\, z^{-k}\right) \sum_{n=1}^{\infty} n\, \hat{h}(n)\, z^{-n+1}. \tag{7.64}$$
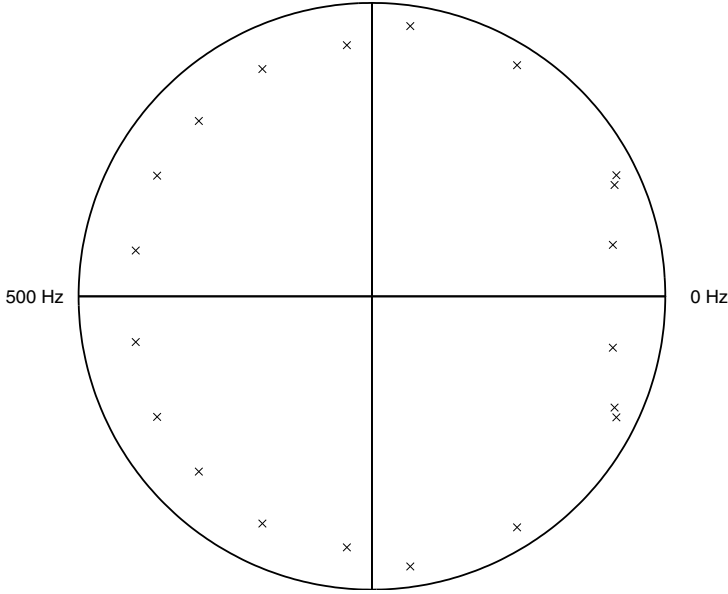
Figure 7.15   Poles of the AR model with order $P = 20$ of the systolic portion of the PCG of a normal subject. (See also Figure 7.14.)

By equating the constant term and the like powers of z^{-1} on both sides, the following relationship can be obtained [362]:

$$\hat{h}(1) = -a_1,$$

$$\hat{h}(n) = -a_n - \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k \, \hat{h}(n-k), \; 1 < n \le P. \qquad (7.65)$$

As noted in Section 4.8.3, phase unwrapping could cause difficulties in estimating the cepstral coefficients using the inverse Fourier transform of the logarithm of the Fourier transform of a given signal [223]. Estimation of the cepstral coefficients using the AR coefficients has the advantage that it does not require phase unwrapping. Although the cepstral coefficients are deduced from the AR coefficients, it is expected that the nonlinear characteristics of the transformation could lead to an improvement in signal classification using the former than the latter set of coefficients. Cepstral coefficients have provided better classification than AR coefficients in speech [362], EMG [363], and VAG [96] signal analysis.

See Section 7.9 for a discussion on the application of AR modeling for the analysis of PCG signals. See Chisci et al. [61] for details on the application of AR modeling for the extraction of features from intracranial EEG signals and prediction of epileptic seizures.
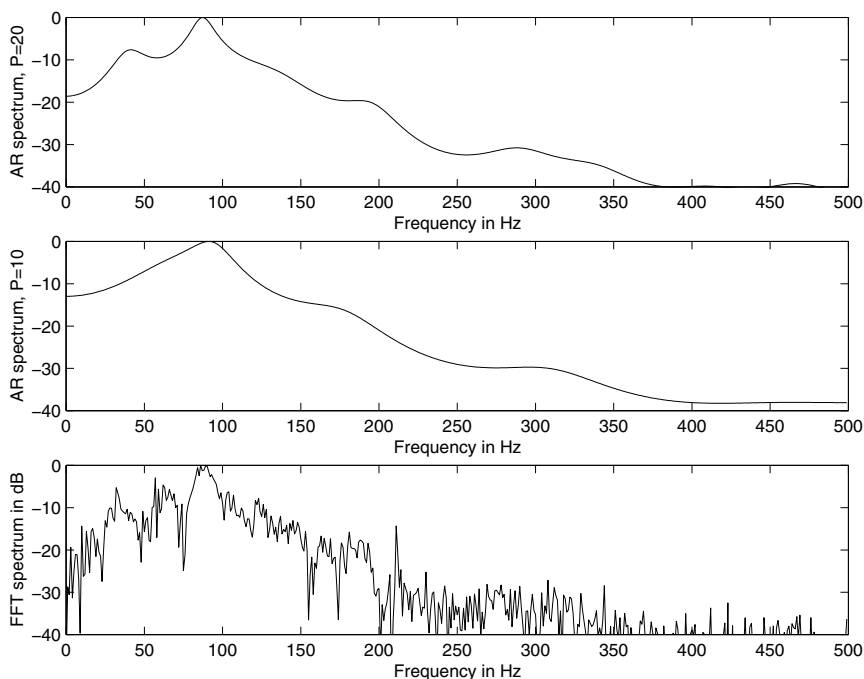
Figure 7.16   Bottom to top: Fourier spectrum of the diastolic portion of the PCG of a normal subject (male, 23 years); AR-model spectrum with order $P = 10$; AR-model spectrum with order $P = 20$. (See also Figures 4.30, 5.7, and 6.12.)

7.6 Pole zero Modeling

Although AR or all-pole modeling can provide good spectral models for any kind of spectra with appropriately high model orders, it has a few limitations. The AR model essentially follows the peaks in the PSD of the signal being modeled, and thus resonance characteristics are represented well. However, if the signal has spectral nulls or valleys (antiresonance), the AR-model spectrum will not provide a good fit in such spectral segments. Spectral zeros are important in modeling certain signals, such as nasal speech signals [364]. Furthermore, an all-pole model assumes the signal to be a minimum-phase signal or a maximum-phase signal, and does not allow mixed-phase signals [239].

The main conceptual difficulty posed by pole–zero modeling is that it is inherently nonunique, because a zero can be approximated by a large number of poles, and viceversa [357]. However, if the system being modeled has a number of influential zeros, the number of poles required for an all-pole model can become large. For these reasons, ARMA or pole–zero modeling [176, 239, 357, 364–366] is important in certain applications.
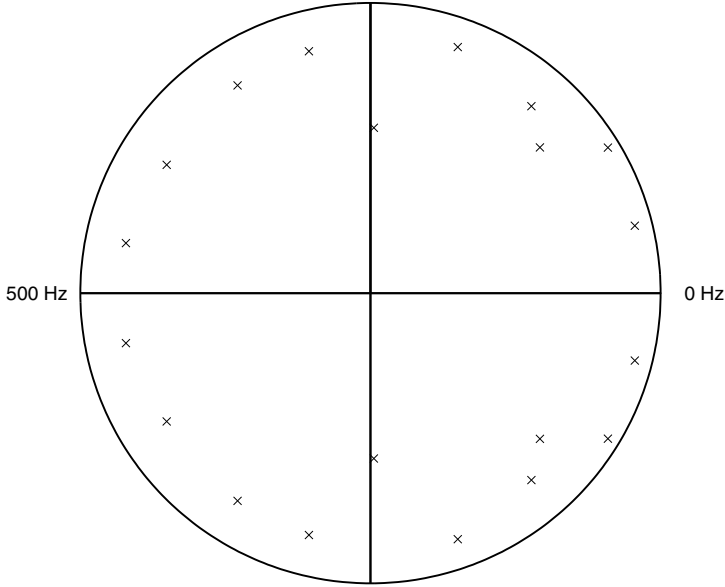
Figure 7.17   Poles of the AR model with order $P = 20$ of the diastolic portion of the PCG of a normal subject. (See also Figure 7.16.)

The ARMA normal equations: From the ARMA model represented by Equation 7.13, we can write the model PSD as [357]

$$\tilde{S}_y(\omega) = |H(\omega)|^2 = G^2 \frac{|B(\omega)|^2}{|A(\omega)|^2} = G^2 \frac{S_b(\omega)}{S_a(\omega)}, \tag{7.66}$$

where

$$S_a(\omega) = \left|1 + \sum_{k=1}^{P} a_k \, \exp(-jk\omega)\right|^2 \tag{7.67}$$

and

$$S_b(\omega) = \left|1 + \sum_{l=1}^{Q} b_l \, \exp(-jl\omega)\right|^2. \tag{7.68}$$

The total spectral-matching error is given by

$$\varepsilon = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_y(\omega) \frac{S_a(\omega)}{S_b(\omega)} \, d\omega, \tag{7.69}$$
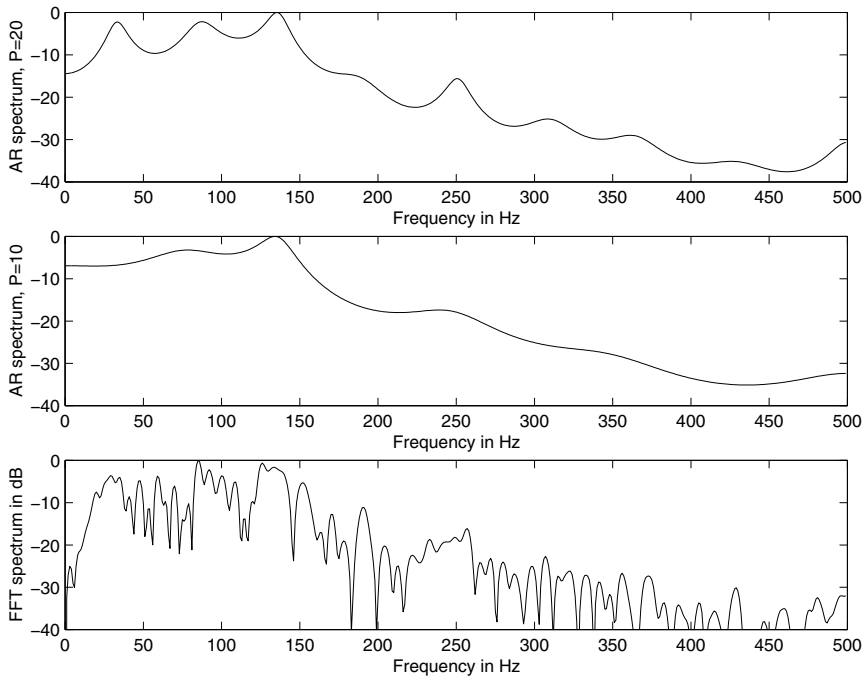
Figure 7.18   Bottom to top: Fourier spectrum of the systolic portion of the PCG of a subject (female, 14 months) with systolic murmur, split S2, and opening snap of the mitral valve; AR-model spectrum with order $P = 10$; AR-model spectrum with order $P = 20$. (See also Figures 4.31, 5.8, and 6.13.)

which may be viewed as the residual energy after passing the modeled signal through the inverse filter $\frac{A(z)}{B(z)}$. In order to obtain the optimal pole–zero model, we need to determine the coefficients a_k and b_l such that the error ε is minimized.

Before taking the derivatives of ε with respect to a_k and b_l, the following relationships are worth noting. Taking the partial derivative of $S_a(\omega)$ in Equation 7.67 with respect to a_i, we get

$$\frac{\partial S_a(\omega)}{\partial a_i} = 2 \sum_{k=0}^{P} a_k \, \cos[(i-k)\omega], \tag{7.70}$$

with $a_0 = 1$. Similarly, from Equation 7.68 we get

$$\frac{\partial S_b(\omega)}{\partial b_i} = 2 \sum_{l=0}^{Q} b_l \, \cos[(i-l)\omega]. \tag{7.71}$$

Let

$$\phi_{y\alpha\beta}(i) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_y(\omega) \frac{[S_a(\omega)]^\beta}{[S_b(\omega)]^\alpha} \, \cos(i\omega) \, d\omega. \tag{7.72}$$
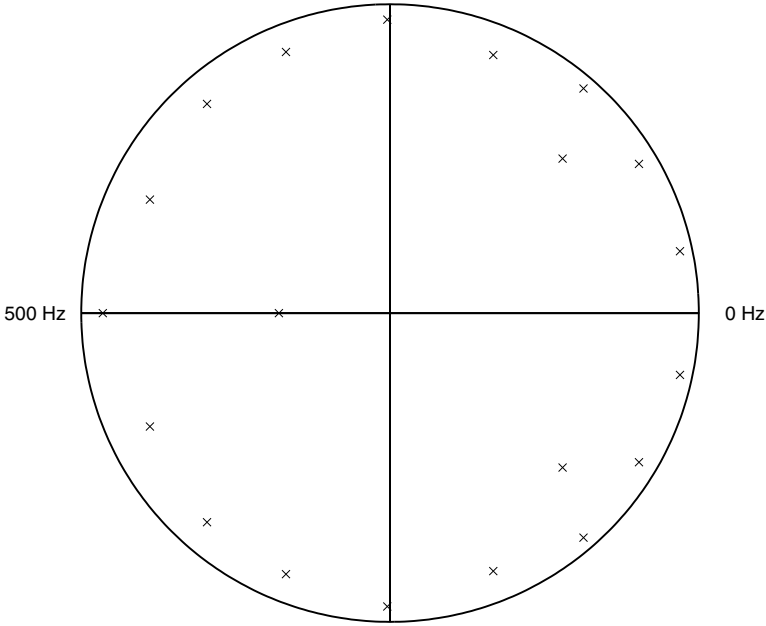
Figure 7.19 Poles of the AR model with order $P = 20$ of the systolic portion of the PCG of a subject with systolic murmur, split S2, and opening snap of the mitral valve. (See also Figure 7.18.)

$\phi_{y00}(i)$ is the inverse Fourier transform of $S_y(\omega)$ and hence simply $\phi_y(i)$.

Now, we can take the partial derivative of ε in Equation 7.69 with respect to a_i as

$$
\begin{aligned}
\frac{\partial \varepsilon}{\partial a_i} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{S_y(\omega)}{S_b(\omega)} \frac{\partial}{\partial a_i} S_a(\omega) \, d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{S_y(\omega)}{S_b(\omega)} \, 2 \sum_{k=0}^{P} a_k \, \cos[(i-k)\omega] \, d\omega \\
&= 2 \sum_{k=0}^{P} a_k \, \phi_{y10}(i-k), \quad 1 \le i \le P.
\end{aligned}
\tag{7.73}
$$

In the same manner, we can obtain

$$
\frac{\partial \varepsilon}{\partial b_i} = -2 \sum_{l=0}^{Q} b_l \, \phi_{y21}(i-l), \quad 1 \le i \le Q.
\tag{7.74}
$$

Because $\phi_{y10}(i-k)$ in Equation 7.73 is not a function of a_k, we obtain a set of linear equations by setting the final expression in Equation 7.73 to zero, which could
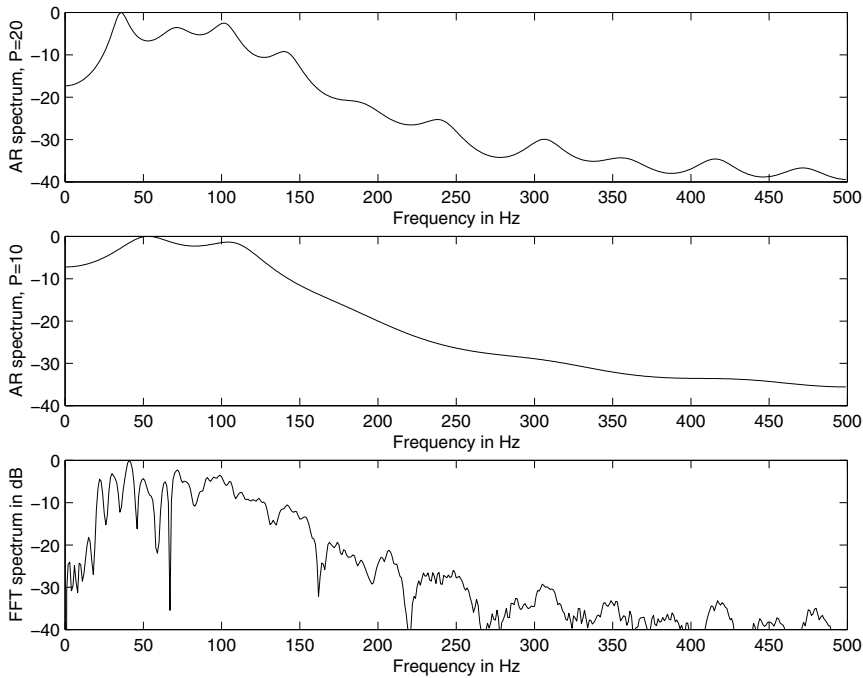
Figure 7.20   Bottom to top: Fourier spectrum of the diastolic portion of the PCG of a subject (female, 14 months) with systolic murmur, split S2, and opening snap of the mitral valve; AR-model spectrum with order $P = 10$; AR-model spectrum with order $P = 20$. (See also Figures 4.31, 5.8, and 6.13.)

be solved to obtain the a_k coefficients in a manner similar to the procedures used in AR modeling. However, $\phi_{y21}(i - l)$ in Equation 7.74 is a function of the b_l coefficients, which leads to a set of nonlinear equations that must be solved to obtain the b_l coefficients; the zeros of the model may then be derived from the b_l coefficients. Obtaining the ARMA model, therefore, requires solving P linear equations and Q nonlinear equations.

Iterative solution of the ARMA normal equations: Makhoul [357] describes the following iterative procedure to solve the $(P + Q)$ ARMA model normal equations based on the Newton–Raphson procedure:

Let $\mathbf{a} = [a_1, a_2, \ldots, a_P]^T$, $\mathbf{b} = [b_1, b_2, \ldots, b_Q]^T$, and $\mathbf{c} = [a_1, a_2, \ldots, a_P,$ $b_1, b_2, \ldots, b_Q]^T$ represent the model coefficients to be derived in vectorial form. The vector at iteration $(m + 1)$ is derived from that at iteration m as

$$\mathbf{c}_{m+1} = \mathbf{c}_m - \mathbf{J}^{-1} \left. \frac{\partial \varepsilon}{\partial \mathbf{c}} \right|_{\mathbf{c}=\mathbf{c}_m}, \tag{7.75}$$

where $\mathbf{J}$ is the $(P + Q) \times (P + Q)$ symmetric Hessian matrix defined as $\mathbf{J} = \frac{\partial^2 \varepsilon}{\partial \mathbf{c} \, \partial \mathbf{c}^T}$. The vector $\mathbf{c}$ may be partitioned as $\mathbf{c}^T = [\mathbf{a}^T, \mathbf{b}^T]$, and the iterative procedure may
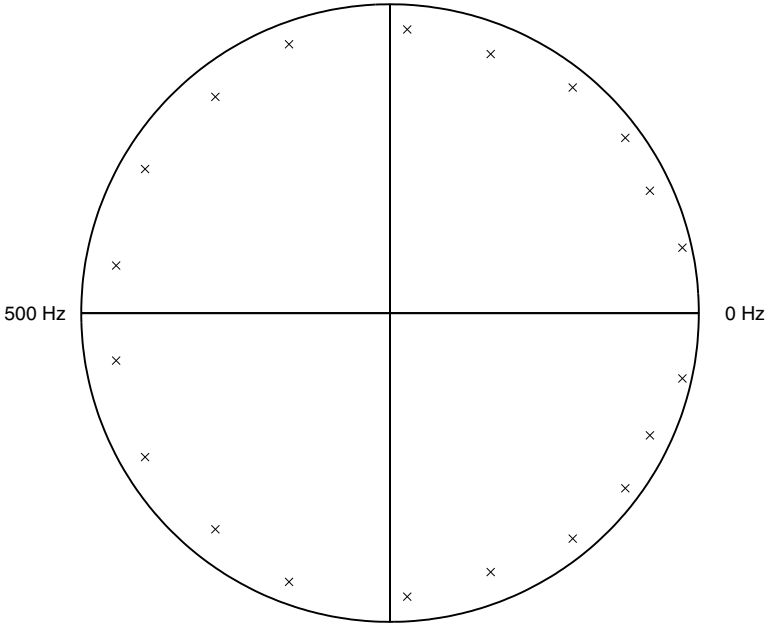
Figure 7.21   Poles of the AR model with order $P = 20$ of the diastolic portion of the PCG of a subject with systolic murmur, split S2, and opening snap of the mitral valve. (See also Figure 7.20.)

be expressed as

$$
\begin{bmatrix} \mathbf{a}_{m+1} \\ \mathbf{b}_{m+1} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_m \\ \mathbf{b}_m \end{bmatrix}
$$

$$
- \begin{bmatrix} \frac{\partial^2 \varepsilon}{\partial \mathbf{a}\, \partial \mathbf{a}^T} & \frac{\partial^2 \varepsilon}{\partial \mathbf{a}\, \partial \mathbf{b}^T} \\ \frac{\partial^2 \varepsilon}{\partial \mathbf{b}\, \partial \mathbf{a}^T} & \frac{\partial^2 \varepsilon}{\partial \mathbf{b}\, \partial \mathbf{b}^T} \end{bmatrix}^{-1}_{\substack{\mathbf{a}=\mathbf{a}_m \\ \mathbf{b}=\mathbf{b}_m}} \begin{bmatrix} \frac{\partial \varepsilon}{\partial \mathbf{a}} \\ \frac{\partial \varepsilon}{\partial \mathbf{b}} \end{bmatrix}_{\substack{\mathbf{a}=\mathbf{a}_m \\ \mathbf{b}=\mathbf{b}_m}} . \quad (7.76)
$$

Equations 7.73 and 7.74 give the first-order partial derivatives required above. The second-order partial derivatives are given as follows [357]:

$$
\frac{\partial^2 \varepsilon}{\partial a_i\, \partial a_j} = 2\phi_{y10}(i - j), \quad (7.77)
$$

$$
\frac{\partial^2 \varepsilon}{\partial a_i\, \partial b_j} = -2 \sum_{k=0}^{P} \sum_{l=0}^{Q} a_k\, b_l\, [\phi_{y20}(j + i - l - k) + \phi_{y20}(j - i - l + k)], \quad (7.78)
$$

and

$$\frac{\partial^2 \varepsilon}{\partial b_i \, \partial b_j} = -2\phi_{y21}(i-j) + 4\sum_{k=0}^{P}\sum_{l=0}^{Q} b_k \, b_l \, [\phi_{y31}(j+i-l-k) + \phi_{y31}(j-i-l+k)].$$

(7.79)

(In the present context, i and j are both indices with integral values.) The iterative procedure works well if the initial estimate is close to the optimal model; otherwise, one of the noniterative methods described in the following sections may be considered.

7.6.1   Sequential estimation of poles and zeros

Given the difficulties with the nonlinear nature of direct pole–zero modeling, a few methods have been proposed to split the problem into two parts: Identify the poles first by AR modeling, and then treat the residual error in some manner to estimate the zeros [176, 239, 357, 364–366]. (*Note:* Several notational differences exist between the various references cited here. The following derivations use notations consistent with those used so far in the present chapter.)

Shanks' method: Let us consider a slightly modified version of Equation 7.13 as

$$H(z) = \frac{Y(z)}{X(z)} = \frac{B(z)}{A(z)} = \frac{1 + \sum_{l=1}^{Q} b_l \, z^{-l}}{1 + \sum_{k=1}^{P} a_k \, z^{-k}},$$

(7.80)

where the gain factor G has been set to be unity: $G = 1$. The difference equation relating the output to the input is given by a small change to Equation 7.12 as

$$y(n) = -\sum_{k=1}^{P} a_k \, y(n-k) + \sum_{l=0}^{Q} b_l \, x(n-l).$$

(7.81)

The effect of the numerator and denominator polynomials in Equation 7.80 may be separated by considering $Y(z) = V(z)B(z)$, where $V(z) = \frac{X(z)}{A(z)}$. This leads to the all-zero or MA part of the system

$$y(n) = \sum_{l=0}^{Q} b_l \, v(n-l),$$

(7.82)

with $v(n)$ given by the all-pole or AR part of the model as

$$v(n) = -\sum_{k=1}^{P} a_k \, v(n-k) + x(n).$$

(7.83)

Let us consider the case of determining the a_k and b_l coefficients (equivalently, the poles and zeros) of the system $H(z)$ given its impulse response. Recollect that

$y(n) = h(n)$ when $x(n) = \delta(n)$; consequently, we have $X(z) = 1$, and $Y(z) = H(z)$. The impulse response of the system is given by

$$h(n) = -\sum_{k=1}^{P} a_k \, h(n-k) + \sum_{l=0}^{Q} b_l \, \delta(n-l), \qquad (7.84)$$

which simplifies to

$$h(n) = -\sum_{k=1}^{P} a_k \, h(n-k), \; n > Q. \qquad (7.85)$$

The effect of the impulse input does not last beyond the number of zeros in the system: the system output is then perfectly predictable from the preceding P samples, and hence an AR or all-pole model is adequate to model $h(n)$ for $n > Q$. As a consequence, Equation 7.32 is modified to

$$\phi_h(i) = -\sum_{k=1}^{P} a_k \, \phi_h(i-k), \;\; i > Q. \qquad (7.86)$$

This system of equations may be solved by considering P equations with $Q+1 \leq i \leq Q+P$. Thus, the a_k coefficients, and hence the poles of the system, may be computed independently of the b_l coefficients or the zeros by restricting the AR error analysis to $n > Q$.

In a practical application, the error of prediction needs to be considered, as the model order P will not be known or some noise will be present in the estimation. Kopec et al. [364] recommend that the covariance method described in Section 7.5 be used to derive the AR model by considering the error of prediction as

$$e(n) = h(n) + \sum_{k=1}^{P} a_k \, h(n-k) \qquad (7.87)$$

and minimizing the TSE defined as

$$\varepsilon = \sum_{n=Q+1}^{\infty} |e(n)|^2. \qquad (7.88)$$

The first Q points are left out as they are not predictable with an all-pole model.

Let us assume that the AR modeling part has been successfully performed by the procedure described above. Let

$$\tilde{A}(z) = 1 + \sum_{k=1}^{P} \tilde{a}_k \, z^{-k} \qquad (7.89)$$

represent the denominator polynomial of the system that has been estimated. The TSE in modeling $h(n)$ is given by

$$\varepsilon = \sum_{n=0}^{\infty} \left| h(n) - \sum_{l=0}^{Q} b_l \, \tilde{v}(n-l) \right|^2, \qquad (7.90)$$

where $\tilde{v}(n)$ is the impulse response of the AR model derived, with $\tilde{V}(z) = \frac{1}{\tilde{A}(z)}$. Minimization of ε as above leads to the set of linear equations

$$\sum_{l=0}^{Q} b_l \, \phi_{\tilde{v}\tilde{v}}(l, j) = \phi_{h\tilde{v}}(0, j), \quad 0 \leq j \leq Q, \tag{7.91}$$

where

$$\phi_{h\tilde{v}}(l, j) = \sum_{n=0}^{\infty} h(n - l) \, \tilde{v}(n - j) \tag{7.92}$$

is the CCF between $h(n)$ and $\tilde{v}(n)$, and $\phi_{\tilde{v}\tilde{v}}$ is the ACF of $\tilde{v}(n)$.

The frequency-domain equivalents of the steps above may be analyzed as follows. The TSE is

$$\begin{aligned}
\varepsilon &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| H(\omega)\tilde{A}(\omega) - \sum_{l=0}^{Q} b_l \, \exp(-jl\omega) \right|^2 |\tilde{V}(\omega)|^2 \, d\omega \tag{7.93} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} |E_h(\omega) - B(\omega)|^2 \, |\tilde{V}(\omega)|^2 \, d\omega,
\end{aligned}$$

where $E_h(z) = H(z)\tilde{A}(z)$ is the AR-model error in the z-domain. [Recall that the Fourier spectrum of a signal is obtained by evaluating the corresponding function of z with $z = \exp(j\omega)$.] The method described above, which is originally attributed to Shanks [176] and has been described as above by Kopec et al. [364], therefore estimates the numerator polynomial of the model by fitting a polynomial (spectral function) to the z-transform of the error of the AR or all-pole model.

Makhoul [357] and Kopec et al. [364] suggest another method labeled as *inverse LP* modeling, where the inverse of the AR-model error $e_h^{-1}(n)$ given as the inverse z-transform of $E_h^{-1}(z)$ is subjected to all-pole modeling. The poles so obtained are the zeros of the original system being modeled.

7.6.2 Iterative system identification

Problem: *Given a noisy observation of the output of a linear system in response to a certain input, develop a method to estimate the numerator and denominator polynomials of a rational z-domain model of the system.*

Solution: In consideration of the difficulty in solving the nonlinear problem inherent in ARMA modeling or pole–zero estimation, Steiglitz and McBride [365] proposed an iterative procedure based upon an initial estimate of the denominator (AR) polynomial. Since their approach to system identification is slightly different from the LP approach we have been using so far in this chapter, it is appropriate to restate the problem.

The Steiglitz–McBride method: Figure 7.22 shows a block diagram illustrating the problem of system identification. The system is represented by its transfer function $H(z)$, input $x(n)$, output $y(n) = h(n) * x(n)$, and the noisy observation

$w(n) = y(n) + \eta(n)$, where $\eta(n)$ is a noise process that is statistically independent of the signals being considered. $H(z)$ is represented as a rational function of z, as

$$H(z) = \frac{Y(z)}{X(z)} = \frac{B(z)}{A(z)} = \frac{\sum_{l=0}^{Q} b_l z^{-l}}{\sum_{k=0}^{P} a_k z^{-k}}. \tag{7.94}$$
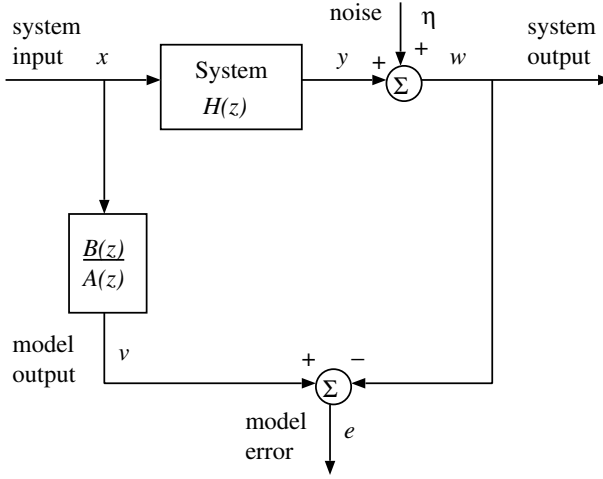

Figure 7.22   Schematic representation of system identification. Adapted from Steiglitz and McBride [365].

The error to be minimized may be written as [365]

$$\sum_{n=0}^{N-1} e^2(n) = \frac{1}{2\pi j} \oint \left| X(z)\frac{B(z)}{A(z)} - W(z) \right|^2 \frac{dz}{z}, \tag{7.95}$$

where the RHS represents the inverse z-transform of the function of z involved, $W(z)$ is the z-transform of $w(n)$, and N is the number of data samples available. The functions of z within the integral essentially compare the predicted model output with the observed output of the physical system.

As seen earlier, this approach leads to a nonlinear problem. The problem may be simplified (linearized) by taking the approach of separate identification of the numerator and denominator polynomials: the estimation problem illustrated in Figure 7.23 treats $A(z)$ and $B(z)$ as separate systems. The error to be minimized may then be written as [365]

$$\sum_{n=0}^{N-1} e^2(n) = \frac{1}{2\pi j} \oint |X(z)B(z) - W(z)A(z)|^2 \frac{dz}{z}. \tag{7.96}$$

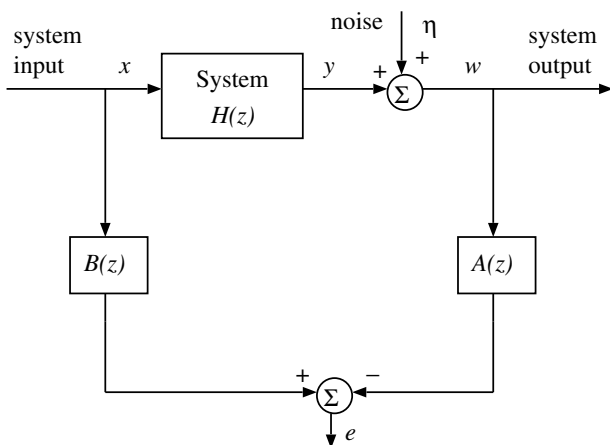(This approach was originally proposed by Kalman [367].)

Figure 7.23 Schematic representation of system identification with separate estimation of $A(z)$ and $B(z)$. Adapted from Steiglitz and McBride [365].

The sample-model error is given by

$$e(n) = \sum_{l=0}^{Q} b_l \, x(n-l) - \sum_{k=1}^{P} a_k \, w(n-k) - w(n). \qquad (7.97)$$

The model coefficients and the input–output data samples may be written in vectorial form as

$$\mathbf{c} = [b_0, b_1, \ldots, b_Q, -a_1, -a_2, \ldots, -a_P]^T \qquad (7.98)$$

and

$$\mathbf{d}(n) = [x(n), x(n-1), \ldots, x(n-Q), w(n-1), w(n-2), \ldots, w(n-P)]^T, \quad (7.99)$$

with the vectors being of size $(P + Q + 1)$. Then, the error is given by

$$e(n) = \mathbf{d}^T(n)\,\mathbf{c} - w(n). \qquad (7.100)$$

The condition for minimum TSE is given by

$$\frac{\partial}{\partial \mathbf{c}} \sum_{n=0}^{N-1} e^2(n) = 2 \sum_{n=0}^{N-1} \frac{\partial e(n)}{\partial \mathbf{c}} e(n) = 2 \sum_{n=0}^{N-1} \mathbf{d}(n)\,e(n) = 0. \qquad (7.101)$$

Substitution of the expression for the error in Equation 7.100 in the condition stated above gives

$$\left(\sum_{n=0}^{N-1} \mathbf{d}(n)\,\mathbf{d}^T(n) \right) \mathbf{c} = \sum_{n=0}^{N-1} w(n)\,\mathbf{d}(n). \qquad (7.102)$$

If we let

$$\mathbf{\Phi} = \sum_{n=0}^{N-1} \mathbf{d}(n)\,\mathbf{d}^T(n) \tag{7.103}$$

represent the $(P + Q + 1) \times (P + Q + 1)$ correlation matrix of the combined string of input–output data samples $\mathbf{d}(n)$ and let

$$\mathbf{\Theta} = \sum_{n=0}^{N-1} w(n)\,\mathbf{d}(n) \tag{7.104}$$

represent the correlation between the signal $w(n)$ and the data vector $\mathbf{d}(n)$ of size $(P + Q + 1)$, we get the solution to the estimation problem as

$$\mathbf{c} = \mathbf{\Phi}^{-1}\,\mathbf{\Theta}. \tag{7.105}$$

Although the vectors and matrices related to the filter coefficients and the signal correlation functions are defined in a different manner, the solution obtained above is comparable to that of the optimal Wiener filter (see Section 3.8 and Equation 3.168).

The limitation of the approach described above is that the error used has no physical significance. The separation of the numerator and denominator functions as in Figure 7.23, while simplifying the estimation problem, has led to a situation that is far from reality.

To improve upon the match between the real physical situation and the estimation problem, Steiglitz and McBride [365] proposed an iterative procedure which is schematically illustrated in Figure 7.24. The basic approach is to treat the system identified using the simplified procedure described above as an initial estimate, labeled as $A_1(z)$ and $B_1(z)$; filter the original signals $x(n)$ and $w(n)$ with the system $\frac{1}{A_1(z)}$; use the filtered signals to obtain new estimates $A_2(z)$ and $B_2(z)$; and iterate the procedure until convergence is achieved.

The error to be minimized may be written as [365]

$$\sum_{n=0}^{N-1} e^2(n) = \frac{1}{2\pi j} \oint \left| X(z)\frac{B_i(z)}{A_{i-1}(z)} - W(z)\frac{A_i(z)}{A_{i-1}(z)} \right|^2 \frac{dz}{z} \tag{7.106}$$

$$= \frac{1}{2\pi j} \oint \left| X(z)\frac{B_i(z)}{A_i(z)} - W(z) \right|^2 \left| \frac{A_i(z)}{A_{i-1}(z)} \right|^2 \frac{dz}{z},$$

with $A_0(z) = 1$. It is obvious that, upon convergence, when $A_i(z) = A_{i-1}(z)$, the minimization problem stated above reduces to the ideal (albeit nonlinear) situation expressed in Equation 7.95 and illustrated in Figure 7.22.

Steiglitz and McBride [365] suggest a modified iterative procedure to improve the estimate further, by imposing the condition that the partial derivatives of the true error criterion with respect to the model coefficients be equal to zero at convergence. The error of the true (ideal) model is given in the z-domain as (refer to Figure 7.22)

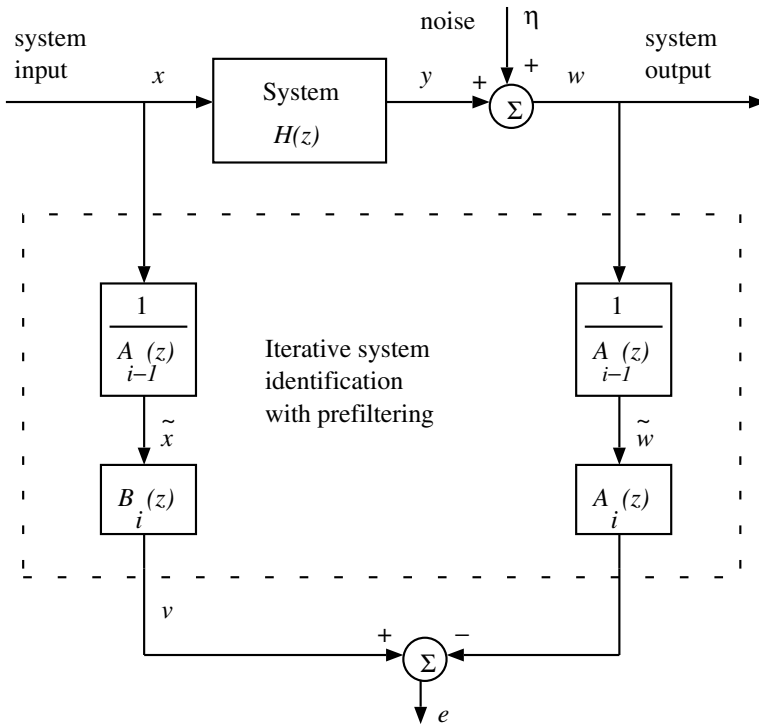$$E(z) = X(z)\frac{B(z)}{A(z)} - W(z) = V(z) - W(z). \tag{7.107}$$

Figure 7.24 Schematic representation of system identification via iterative prefiltering. Adapted from Steiglitz and McBride [365].

$V(z)$ is the output of the model for the input $x(n)$. The derivatives of $E(z)$ with respect to the model coefficients are given by

$$\frac{\partial \varepsilon}{\partial a_i} = -\frac{X(z)B(z)}{A^2(z)} z^{-i} = -\frac{V(z)}{A(z)} z^{-i} = -\tilde{V}(z)z^{-i} \qquad (7.108)$$

and

$$\frac{\partial \varepsilon}{\partial b_i} = \frac{X(z)}{A(z)} z^{-i} = \tilde{X}(z)z^{-i}, \qquad (7.109)$$

where the superscript $\tilde{\ }$ represents a filtered version of the corresponding signal, the filter transfer function being $\frac{1}{A(z)}$. A new data vector is defined as

$$\mathbf{d}_1(n) = [\tilde{x}(n), \tilde{x}(n-1), \ldots, \tilde{x}(n-Q), \tilde{v}(n-1), \tilde{v}(n-2), \ldots, \tilde{v}(n-P)]^T. \quad (7.110)$$

The error gradient in Equation 7.101 is modified to

$$\frac{\partial}{\partial \mathbf{c}} \sum_{n=0}^{N-1} e^2(n) = 2 \sum_{n=0}^{N-1} \frac{\partial e(n)}{\partial \mathbf{c}} e(n) = 2 \sum_{n=0}^{N-1} \mathbf{d}_1(n) e(n) \qquad (7.111)$$

$$= 2 \sum_{n=0}^{N-1} [\mathbf{d}_1(n) \, \mathbf{d}_1^T(n) \, \mathbf{c} - w(n) \, \mathbf{d}_1(n)],$$

where the last equality is true only at convergence. The rest of the procedure remains the same as before, but with the correlation functions defined as

$$\mathbf{\Phi_1} = \sum_{n=0}^{N-1} \mathbf{d}_1(n) \, \mathbf{d}_1^T(n) \tag{7.112}$$

and

$$\mathbf{\Theta_1} = \sum_{n=0}^{N-1} w(n) \, \mathbf{d}_1(n). \tag{7.113}$$

Once the a_k and b_l coefficients are obtained, the related polynomials may be solved to obtain the poles and zeros of the system being modeled. Furthermore, the polynomials may be used to derive spectral models of the system or the signal of interest. Note that the procedures given above are applicable to the special case of system identification when the impulse response $h(n)$ is given: We just need to change $x(n) = \delta(n)$ and $X(z) = 1$. Steiglitz and McBride [365] did not provide any proof of convergence of their methods; however, it was indicated that the method performed successfully in many practical applications.

The Steiglitz–McBride method was applied to the modeling and classification of PCG signals by Joo et al. [368]. The first and second peak frequencies were detected from the model spectra and used to analyze porcine prosthetic valve function. Murthy and Prasad [369] applied the Steiglitz–McBride method to ECG signals. Pole–zero models derived from ECG strips including a few cardiac cycles were used to reconstruct and identify the ECG waveform over a single cycle, and also to reconstruct separately (that is, to segment) the P, QRS, and T waves.

7.6.3 Homomorphic prediction and modeling

Problem: *Given the relative ease of all-pole modeling, is it possible to convert the zeros of a system to poles?*

Solution: As mentioned earlier, an all-pole model assumes the signal being modeled to be a minimum-phase signal or a maximum-phase signal, and does not allow mixed-phase signals [239]. We have seen in Sections 4.8.3 and 5.4.2 that homomorphic filtering can facilitate the separation of the minimum-phase and maximum-phase components of a mixed-phase signal, and further facilitate the derivation of a minimum-phase version or correspondent (MPC) of a mixed-phase signal. Makhoul [357], Oppenheim et al. [239], and Kopec et al. [364] suggested methods to combine homomorphic filtering and LP into a procedure that has been labeled *homomorphic prediction* or *cepstral prediction*.

An intriguing property that arises in homomorphic prediction is that if a signal $x(n)$ has a rational z-transform, then $n\hat{x}(n)$ [where $\hat{x}(n)$ is the complex cepstrum of $x(n)$] has a rational z-transform whose poles correspond to the poles and zeros

of $x(n)$. The basic property of the z-transform we need to recollect here is that if $X(z)$ is the z-transform of $x(n)$, then the z-transform of $nx(n)$ is $-z \frac{dX(z)}{dz}$. Now, the complex cepstrum $\hat{x}(n)$ of $x(n)$ is defined as the inverse z-transform of $\hat{X}(z) = \log X(z)$. Therefore, we have

$$ZT[n\hat{x}(n)] = -z \frac{d\hat{X}(z)}{dz} = -z \frac{1}{X(z)} \frac{dX(z)}{dz}, \qquad (7.114)$$

where $ZT[\,]$ represents the z-transform operator. If $X(z) = \frac{B(z)}{A(z)}$, we get

$$ZT[n\hat{x}(n)] = -z \frac{A(z)B'(z) - B(z)A'(z)}{A(z)B(z)}, \qquad (7.115)$$

where the prime $'$ denotes the derivative of the associated function with respect to z. A general representation of a rational function of z (which represents an exponential signal in the z-domain) in terms of its poles and zeros is given by [239]

$$X(z) = A \, z^r \, \frac{\prod_{l=1}^{Q_i} (1 - z_{il} \, z^{-1}) \prod_{n=1}^{Q_o} (1 - z_{on} \, z)}{\prod_{k=1}^{P_i} (1 - p_{ik} \, z^{-1}) \prod_{m=1}^{P_o} (1 - p_{om} \, z)}, \qquad (7.116)$$

with the magnitudes of all of the z_i, z_o, p_i, and p_o coefficients being less than unity (see also Sections 3.4.3, 4.8.3, and 5.4.2). The p_i and z_i values above give the P_i poles and Q_i zeros, respectively, of the system that are inside the unit circle in the z-plane; $\frac{1}{p_o}$ and $\frac{1}{z_o}$ give the P_o poles and Q_o zeros, respectively, that lie outside the unit circle. A causal and stable system will not have any poles outside the unit circle; regardless, the general representation given above permits the analysis and modeling of maximum-phase signals that are anticausal. Computation of the complex cepstrum requires the removal of any linear phase component that may be present, and hence we could impose the condition $r = 0$. We then have

$$
\begin{aligned}
\hat{X}(z) \;=\; & \log X(z) = \log A \\
& + \sum_{l=1}^{Q_i} \log(1 - z_{il} \, z^{-1}) + \sum_{n=1}^{Q_o} \log(1 - z_{on} \, z) \\
& - \sum_{k=1}^{P_i} \log(1 - p_{ik} \, z^{-1}) - \sum_{m=1}^{P_o} \log(1 - p_{om} \, z), \qquad (7.117)
\end{aligned}
$$

and, furthermore,

$$
\begin{aligned}
-z \frac{d\hat{X}(z)}{dz} \;=\; & - \sum_{l=1}^{Q_i} \frac{z_{il} \, z^{-1}}{(1 - z_{il} \, z^{-1})} + \sum_{n=1}^{Q_o} \frac{z_{on} \, z}{(1 - z_{on} \, z)} \\
& + \sum_{k=1}^{P_i} \frac{p_{ik} \, z^{-1}}{(1 - p_{ik} \, z^{-1})} - \sum_{m=1}^{P_o} \frac{p_{om} \, z}{(1 - p_{om} \, z)}. \qquad (7.118)
\end{aligned}
$$

From the expression given above, it is evident that $n\hat{x}(n)$ has simple (first-order) poles at every pole as well as every zero of $x(n)$. Therefore, we could apply an all-pole modeling procedure to $n\hat{x}(n)$, and then separate the poles so obtained into the desired poles and zeros of $x(n)$. An initial all-pole model of $x(n)$ can assist in the task of separating the poles from the zeros. Oppenheim et al. [239] show further that even if $X(z)$ is irrational, $n\hat{x}(n)$ has a rational z-transform with first-order poles corresponding to each irrational factor in $X(z)$.

Illustration of application to a synthetic speech signal: Figures 7.25 and 7.26 show examples of the application of several pole–zero and all-pole modeling techniques to a synthetic speech signal [239]. The impulse response of a synthetic system with two poles at $292\ Hz$ and $3,500\ Hz$ with bandwidth equal to $79\ Hz$ and $100\ Hz$, respectively, and one zero at $2,000\ Hz$ with bandwidth equal to $200\ Hz$ is shown in Figure 7.25 (a), with its log-magnitude spectrum in Figure 7.26 (a). The formant or resonance structure of the signal is evident in the spectral peaks. (The sampling rate is $12\ kHz$.) Excitation of the system with a pulse train with the repetition rate of $120\ Hz$ resulted in the signal in Figure 7.25 (b), whose spectrum is shown in Figure 7.26 (b); the spectrum clearly shows the effect of repetition of the basic wavelet in the series of waves that are superimposed on the basic spectrum of the wavelet. Application of homomorphic filtering to the signal in Figure 7.25 (b) provided an estimate of the basic wavelet as shown in Figure 7.25 (c), with the corresponding spectrum in Figure 7.26 (c).

The pole–zero modeling method of Shanks was applied to the result of homomorphic filtering in Figure 7.25 (c) with four poles and two zeros. The impulse response of the model and the corresponding spectrum are shown in Figure 7.25 (d) and Figure 7.26 (d), respectively. It is seen that the two peaks and the valley in the original spectrum are faithfully reproduced in the modeled spectrum. The frequencies of the poles (and their bandwidths) given by the model are $291\ Hz$ ($118\ Hz$) and $3,498\ Hz$ ($128\ Hz$), and those of the zero are $2,004\ Hz$ ($242\ Hz$), which compare well with those of the synthesized system listed in the preceding paragraph.

Application of the autocorrelation method of LP modeling with six poles to the original signal in Figure 7.25 (a) resulted in the model impulse response and spectrum illustrated in Figures 7.25 (e) and 7.26 (e). While the all-pole model spectrum has followed the spectral peaks well, it has failed to represent the valley or null related to the zero.

Illustration of application to a real speech signal: Figure 7.27 (a) shows the log-magnitude spectrum of a real speech signal (preemphasized) of the nasalized vowel /U/ in the word "moon" [364]. Part (b) of the same figure shows the spectrum after homomorphic filtering to remove the effects of repetition of the basic wavelet. Parts (c) and (d) show 10-pole, 6-zero model spectra obtained using Shanks' method and inverse LP modeling, respectively. The spectra of the models have successfully followed the peaks and valleys in the signal spectrum.

Shanks' method was applied to the minimum-phase and maximum-phase components of ECG signals obtained via homomorphic filtering by Murthy et al. [370]. Akay et al. [371] used ARMA techniques to model diastolic heart sounds for the
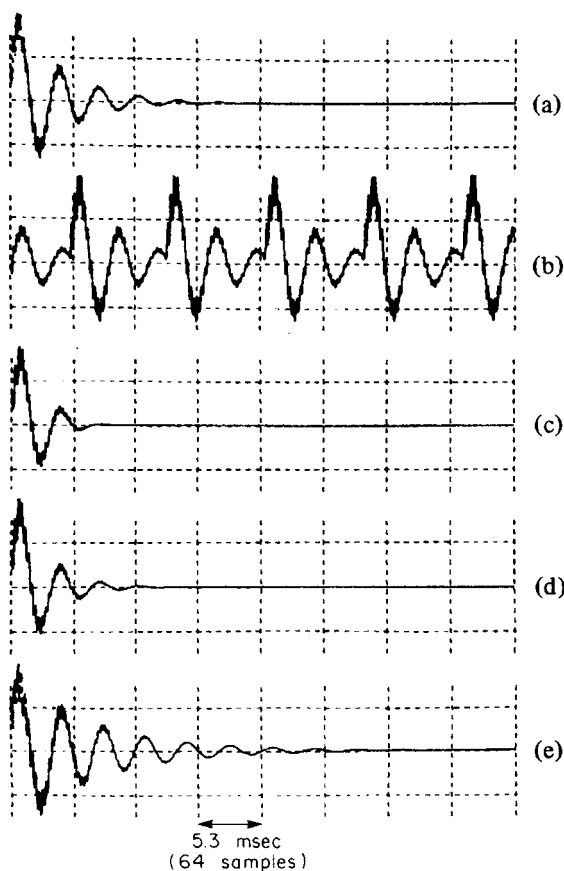
Figure 7.25   Time-domain signals: (a) impulse response of a 4-pole, 2-zero synthetic system; (b) synthesized voiced-speech signal obtained by triggering the system with an impulse train; (c) result of basic wavelet extraction via application of homomorphic filtering to the signal in (b); (d) impulse response of a 4-pole, 2-zero model of the signal in (c) obtained by Shanks' method; (e) impulse response of a 6-pole AR model. msec = ms. See also Figure 7.26. Reproduced with permission from A.V. Oppenheim, G.E. Kopec, and J.M. Tribolet, Signal analysis by homomorphic prediction, *IEEE Transactions on Acoustics, Speech, and Signal Processing,* 24(4):327–332, 1976. ©IEEE.
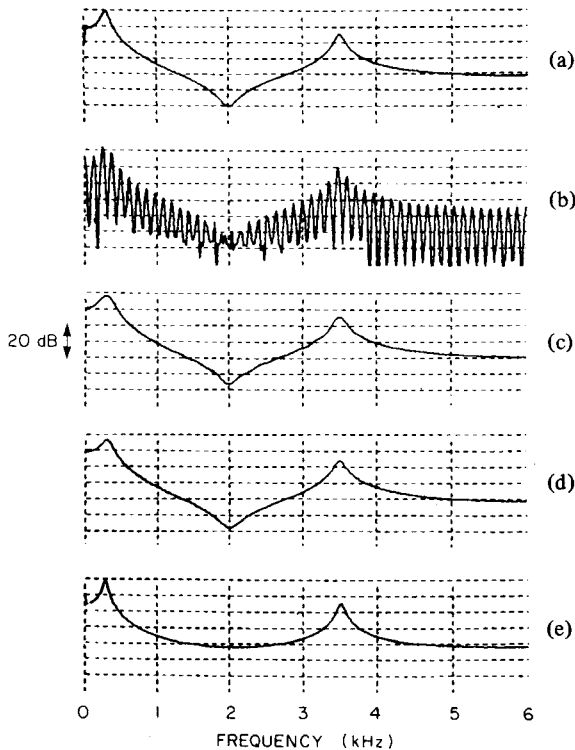
Figure 7.26   Log-magnitude spectra of the time-domain signals in Figure 7.25: (a) actual spectral response of the 4-pole, 2-zero synthetic system; (b) spectrum of the synthesized voiced-speech signal obtained by triggering the system with an impulse train; (c) spectrum of the basic wavelet extracted via application of homomorphic filtering to the signal corresponding to (b); (d) spectral response of a 4-pole, 2-zero model of the signal in (c) obtained by Shanks' method; (e) spectral response of a 6-pole AR model. Reproduced with permission from A.V. Oppenheim, G.E. Kopec, and J.M. Tribolet, Signal analysis by homomorphic prediction, *IEEE Transactions on Acoustics, Speech, and Signal Processing,* 24(4):327–332, 1976. ©IEEE.

detection of coronary heart disease; however, only the dominant poles of the model were used in pattern analysis (see Section 7.10 for details of this application).

7.7   Electromechanical Models of Signal Generation

While purely mathematical models of signal generation, such as point processes and linear system models, provide the advantage of theoretical elegance and convenience, they may not be able to represent directly certain physical and physiological aspects of the systems that generate the signals. For example, the models described in the preceding sections cannot directly accommodate the physical dimensions of blood
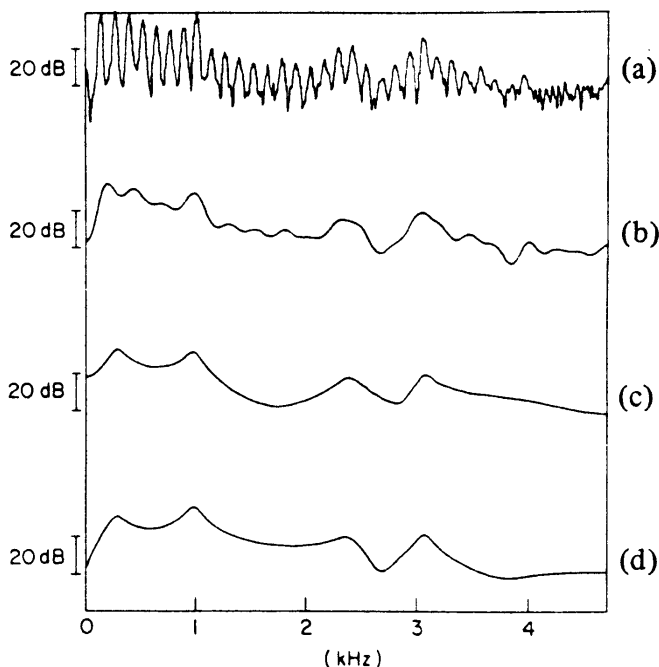
Figure 7.27   (a) Log-magnitude spectrum of the preemphasized, real speech signal of the nasalized vowel /U/ in the word "moon"; (b) spectrum after homomorphic filtering to remove the effects of repetition of the basic wavelet; (c) spectral response of a 10-pole, 6-zero model obtained by Shanks' method; (d) spectral response of a 10-pole, 6-zero model obtained by inverse LP modeling. Reproduced with permission from G.E. Kopec, A.V. Oppenheim, and J.M. Tribolet, Speech analysis by homomorphic prediction, *IEEE Transactions on Acoustics, Speech, and Signal Processing,* 25(1):40–49, 1977. ©IEEE.

vessels or valves, the loss in the compliance of a valve leaflet due to stenosis, or the lubrication (or the lack thereof) or friction between joint surfaces.

Sikarskie et al. [372] proposed a model to characterize aortic valve vibration for the analysis of its contribution to S2; in addition to mathematical relationships, they included physical factors, such as the valve forcing function, valve mass, and valve stiffness. It was shown that the amplitude and frequency of A2 depend strongly on the valve forcing function and valve stiffness. Valve mass was shown to have little effect on the amplitude and frequency of A2; blood density was shown to have no effect on the same parameters.

We shall now study three representative applications of electromechanical modeling, where mechanical models and their electrical counterparts are used to represent the generation and altered characteristics of sounds in the respiratory system, coronary arteries, and knee joints.

7.7.1 Modeling of respiratory sounds

Problem: *Given that parts of the respiratory system are composed of pipe-like segments, propose an electromechanical model to characterize the sounds produced due to respiration.*

Solution: Moussavi [108] proposed mechanical as well as electrical circuit models to characterize parts of the respiratory system. The models were used in the analysis of sounds related to breathing.

Regardless of the complex structure of the human vocal and respiratory tract, its main characteristics can be represented by mechanical models with tubes or pipes and their equivalent electrical circuit models [108, 373]; see also Section 1.2.12. A simple model of the vocal tract is given by a pipe closed at one end by the glottis and open at the other end at the lips. A pipe with length L has resonance frequencies given by $f_n = \frac{nv}{4L}$, $n = 1, 3, 5, \ldots$, where n is the harmonic number and v is the velocity of air through the pipe. In order to take into account the variations of the respiratory tract along its length, the model could be modified to include multiple short segments of pipes or tubes with different diameter and length. Due to the mass of the air in each segment of the pipe model, the related inertance that opposes acceleration needs to be taken into account. Furthermore, the compliance related to the compressibility of air and the elasticity of the pipes' walls need to be represented. Overall, these considerations lead to a lossy transmission line or circuit model [108, 373]. Figure 7.28 shows a mechanical model and an electrical circuit model for a segment of the respiratory system given by Moussavi [108]; see also Figures 1.47 and 1.48 as well as Flanagan [373].

If a tube is smooth and its walls are hard, loss of energy occurs via viscous friction at the walls of the tube and heat conduction. The viscous loss is proportional to the square of particle velocity and the loss due to heat conduction is proportional to the square of the pressure [373]. Some of the similarities that can be drawn between mechanical and electrical models for the system described above are as follows [108, 373]: The sound pressure, P, is analogous to voltage, V; the acoustic volume velocity, U, is comparable to current, I; viscous loss can be represented by $I^2 R$ loss, where R is resistance; heat conduction loss can be related to V^2/R or $V^2 G$ loss, where G is conductance; inertance of air mass is analogous to inductance; and compliance can be modeled by capacitance.

Moussavi [108] gives the following derivations for various parameters of the models; see also Flanagan [373]. Starting with Newton's second law, we have force $F = ma$, where m is the mass and a is the acceleration of an object. In a tube with cross-sectional area A and pressure P, we have $F = PA$; $m = \rho A l$, where ρ is the density of air and l is the length of the pipe; and $a = \frac{du}{dt}$, where u is the particle velocity and t is time. The volume velocity U is given by $U = Au$. Then, we have

$$PA = \rho A l \frac{du}{dt} = \rho l \frac{dU}{dt}, \qquad (7.119)$$

which leads to
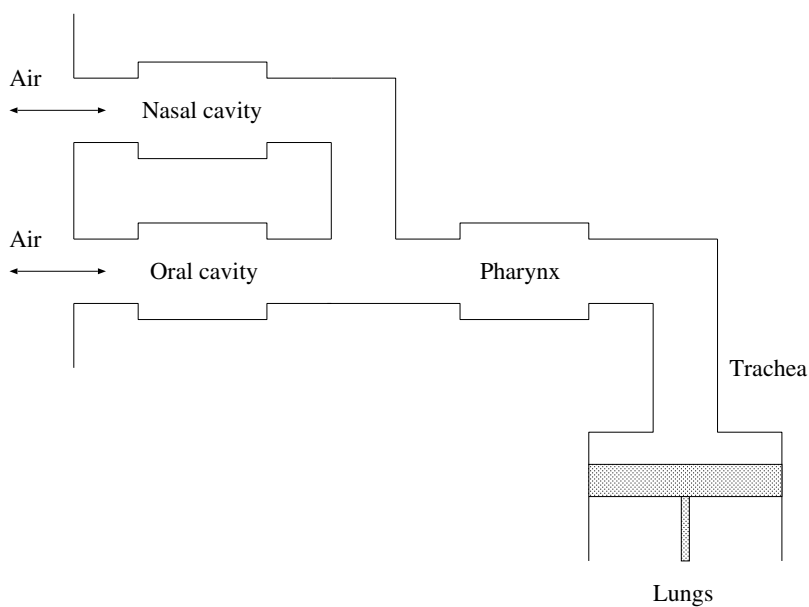
(a)
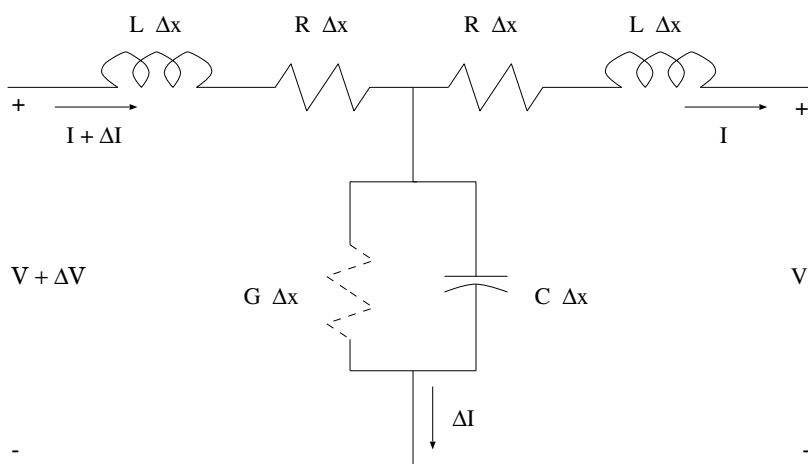
(b)

Figure 7.28   (a) Mechanical and (b) electrical circuit models of parts of the respiratory system. Adapted with permission from Z. Moussavi, *Fundamentals of Respiratory Sounds and Analysis*, Morgan & Claypool, San Francisco, CA, 2006, ©Morgan & Claypool.

$$P = \rho \frac{l}{A} \frac{dU}{dt}.$$ (7.120)

Comparing the equation given above with that for the voltage across an inductance L given by

$$V = L \frac{dI}{dt},$$ (7.121)

we have the acoustic equivalent of inductance as

$$L_a = \rho \frac{l}{A}.$$ (7.122)

Now, consider the adiabatic gas law $PV_a^\eta = $ a constant, where V_a is the volume of the gas under pressure P, and η is the adiabatic constant. Differentiating this equation with respect to time, we have

$$P \eta V_a^{\eta-1} \frac{dV_a}{dt} + V_a^\eta \frac{dP}{dt} = 0,$$ (7.123)

which leads to

$$\frac{1}{P} \frac{dP}{dt} = -\frac{\eta}{V_a} \frac{dV_a}{dt} = \frac{\eta}{V_a} U$$ (7.124)

and

$$U = \frac{V_a}{P \eta} \frac{dP}{dt}.$$ (7.125)

Comparing the last equation given above with the equation for the current through a capacitor C, expressed as

$$I = C \frac{dV}{dt},$$ (7.126)

we have the acoustic equivalent of capacitance or compliance as

$$C_a = \frac{V_a}{P \eta}.$$ (7.127)

Acoustic resistance is given by

$$R_a = \frac{l\,S}{A^2} \sqrt{\frac{\omega \rho \mu}{2}},$$ (7.128)

where ω is the frequency component of air flow, μ is the viscosity coefficient, and S is the circumference of the pipe, and acoustic conductance is given by

$$G_a = S\,l \frac{\eta - l}{\rho c^2} \sqrt{\frac{\lambda \omega}{2c_p \rho}},$$ (7.129)

where λ is the coefficient of heat conduction and c_p is the specific heat of air at constant pressure; see Flanagan [373] and Moussavi [108] for detailed derivations of these expressions.

Parameters derived as above may be used to model segments of tubes or pipes that can be combined to form more elaborate models of parts of the respiratory system. The effects of various segments could be analyzed as filtering of a certain input signal (such as an impulse, a pulse, or random noise) as it passes through the segments. Various respiratory diseases cause narrowing of parts of the respiratory tract as well changes in their smoothness, stiffness, and other mechanical properties. If such pathological changes can be included in the mechanical or electrical model, it becomes possible to analyze the related respiratory sounds and their characteristics. See Moussavi [108] for examples of sounds related to various diseases that affect the respiratory system and their characteristics.

7.7.2 Sound generation in coronary arteries

Problem: *Propose an electromechanical model to characterize the sounds produced due to blood flow in stenosed arteries.*

Solution: Blood vessels are normally flexible, elastic, and pliant, with smooth internal surfaces. When a segment of a blood vessel is hardened due to the deposition of calcium and other minerals, the segment becomes rigid. Furthermore, the development of plaque inside the vessel causes narrowing or constriction of the vessel, which impedes the flow of blood. The result is turbulent flow of blood, with accompanying high-frequency sounds.

Wang et al. [374, 375] proposed a sound-source model combining an incremental-network model of the left coronary artery tree with a transfer-function model describing the resonance characteristics of arterial chambers. The network model, illustrated in Figure 7.29, predicts flow in normal and stenosed arteries. It was noted that stenotic branches may require division into multiple segments in the model due to greater geometric variations. Furthermore, it was observed that a stenotic segment may exhibit poststenotic dilation as illustrated in Figure 7.30, due to increased pressure fluctuations caused by turbulence at the point of stenosis.

The resonance frequency of a segment of an artery depends upon the length and diameter of the segment, as well as upon the distal (away from the heart) hydraulic pressure loading the segment. The physical parameters required for the model were obtained from arteriograms of the patient being examined. The terminal resistances, labeled Z in Figure 7.29, represent loading of the resistive arteriolar beds, assumed to be directly related to the areas that the terminal branches serve.

Wang et al. related the network elements (resistance R, inertance or inductance L, and capacitance C) to the physical parameters of the artery segments as

$$R = 8\pi\nu \, \frac{l}{A^2}, \tag{7.130}$$
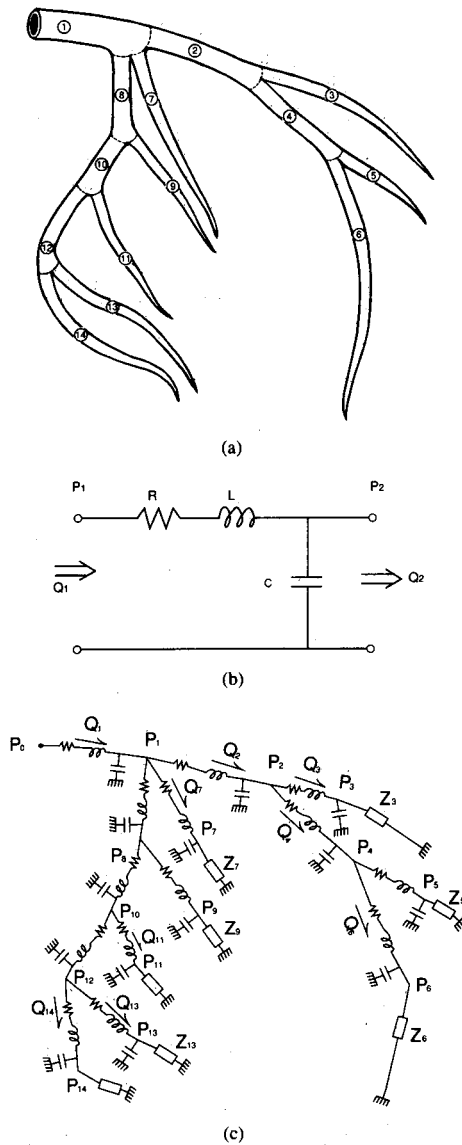
$$L = \rho \, \frac{l}{A}, \tag{7.131}$$

Figure 7.29   Electromechanical model of a coronary artery tree. (a) The left coronary artery tree is divided into 14 branches. (b) Circuit model of a segment. (c) Circuit model of the artery tree. Reproduced with permission from J.Z. Wang, B. Tie, W. Welkowitz, J.L. Semmlow, and J.B. Kostis, Modeling sound generation in stenosed coronary arteries, *IEEE Transactions on Biomedical Engineering,* 37(11):1087–1094, 1990, ©IEEE; and J.Z. Wang, B. Tie, W. Welkowitz, J. Kostis, and J. Semmlow, Incremental network analogue model of the coronary artery, *Medical & Biological Engineering & Computing,* 27:416–422, 1989. ©IFMBE.
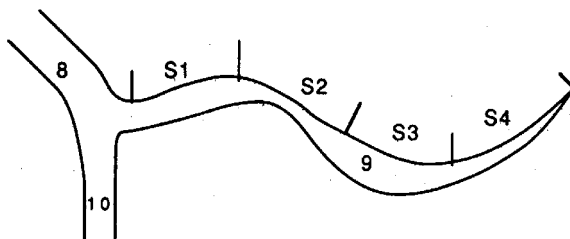
S1: **Proximal to Stenosis Segment**
S2: **Stenotic Segment**
S3: **Poststenotic Dilation Segment**
S4: **Distal to Dilation Segment**

Figure 7.30 Hypothetical example of stenosis in coronary artery branch number 9. Reproduced with permission from J.Z. Wang, B. Tie, W. Welkowitz, J.L. Semmlow, and J.B. Kostis, Modeling sound generation in stenosed coronary arteries, *IEEE Transactions on Biomedical Engineering,* 37(11):1087–1094, 1990. ©IEEE.

and

$$C = A\, l\, h\, \frac{D}{E}, \tag{7.132}$$

where $\nu = 0.04\ g\, cm^{-1}\, s$ is the viscosity of blood, $\rho = 1.0\ g\, cm^{-3}$ is the density of blood, $E = 2 \times 10^6\ g\, cm^{-1}\, s^2$ is the Young's modulus of the blood vessel, D is the diameter of the segment, $A = \pi \frac{D^2}{4}$ is the cross-sectional area of the segment, $h \approx 0.08D$ is the wall thickness of the segment, and l is the length of the segment (see also Section 7.7.1). Wang et al. [375] remarked that, while the network elements may be assumed to be approximately constant during diastole, the assumption would not be valid during systole due to variations in the parameters of the segments.

In analyzing the artery–network model, voltage is analogous to pressure (P) and current is analogous to blood flow (Q). State-variable differential equations were used by Wang et al. [375] to derive the flow through the artery tree model for various pressure waveforms. It was hypothesized that turbulence at the point of stenosis would provide the excitation power, and that the stenotic segment and the dilated segment distal to the point of stenosis (see Figure 7.30) would act as resonance chambers.

Wang et al. [374] used the following relationships to compute the RMS pressure fluctuation (see also Fredberg [376]):

$$\langle P^2 \rangle_{\max} = 10^{-4} \rho u^2 f(x), \tag{7.133}$$

$$f(x) = 25.1 - 37.1x + 15.5x^2 - 0.08x^3 - 0.89x^4 + 0.12x^5, \tag{7.134}$$

and

$$x = 10^{-3} \frac{ud}{\nu} \left(\frac{D}{d}\right)^{0.75},\qquad(7.135)$$

where u is the blood velocity in the stenotic segment, and d is the diameter of the stenotic segment. The incremental network model was used to estimate the blood velocity in each segment.

The wideband spectrum of the sound associated with turbulent flow was modeled as (see also Fredberg [376])

$$S(f) = \frac{0.7 \frac{d}{U} \langle P^2 \rangle_{\max}}{1 + 0.5 \left[f\frac{d}{U}\right]^{\frac{10}{3}}}\qquad(7.136)$$

where U is the velocity of blood in a normal segment and f is frequency in Hz. Wang et al. used the function $S(f)$ given above as the source of excitation power to derive the response of their network model. It was observed that the model spectra indicated two resonance frequencies, the magnitude and frequency of which depended upon the geometry and loading of the segments. Wang et al. cautioned that the results of the model are sensitive to errors in the estimation of the required parameters from arteriograms or other sources.

Figure 7.31 illustrates the model spectra for segment number 12 of the artery tree model in Figure 7.29 with no stenosis and with stenosis of two grades. Narrowing of the segment with increasing stenosis is seen to shift the second peak in the spectrum to higher frequencies, while the magnitude and frequency of the first peak are both reduced. The results were confirmed by comparing the model spectra with spectra of signals recorded from a few patients with stenosed coronary arteries. Examples of spectral analysis of signals recorded from patients before and after angioplasty to correct for stenosis are presented in Section 7.10.

7.7.3   Sound generation in knee joints

Problem: *Develop a mechanical analog of the knee joint to model the generation of the pulse train related to PPC.*

Solution: Beverland et al. [346] studied the PPC signals produced during very slow movement of the leg (at about $4°/s$). The signals were recorded by taping accelerometers to the skin above the upper pole and/or the lower pole of the patella. Reproducible series of bursts of vibration were recorded in their experiments. Figure 7.32 illustrates two channels of simultaneously recorded PPC signals from the upper and lower poles of the patella during extension and flexion of the leg. The signals display reversed similarity when extension versus flexion or upper-pole versus lower-pole recordings are compared.

Beverland et al. proposed a mechanical model to explain the generation of the PPC signals. The patella was considered to behave like a seesaw in the model, which was supported by the observation that a pivot point exists at the midpoint of the patella. The apparatus constructed, as illustrated in Figure 7.33, included a rubber
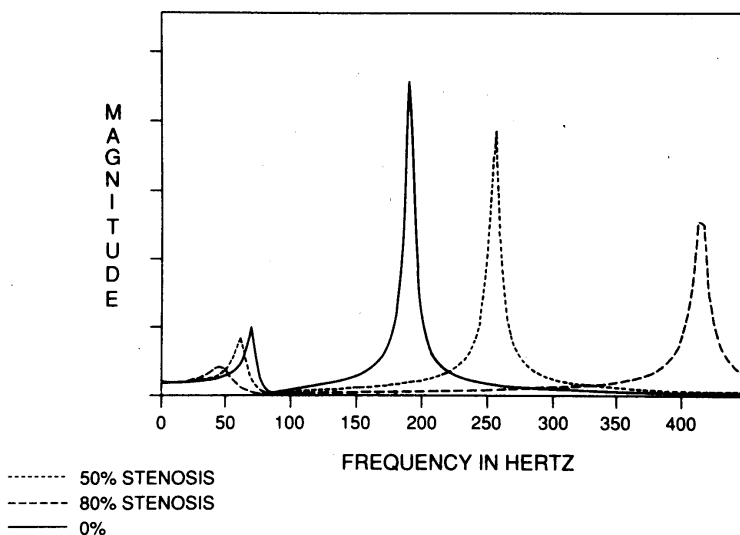
Figure 7.31   Shift in frequency components predicted by the transfer-function model for the case of stenosis in element number 12 in the model of the coronary artery in Figure 7.29. Reproduced with permission from J.Z. Wang, B. Tie, W. Welkowitz, J.L. Semmlow, and J.B. Kostis, Modeling sound generation in stenosed coronary arteries, *IEEE Transactions on Biomedical Engineering,* 37(11):1087–1094, 1990. ©IEEE.
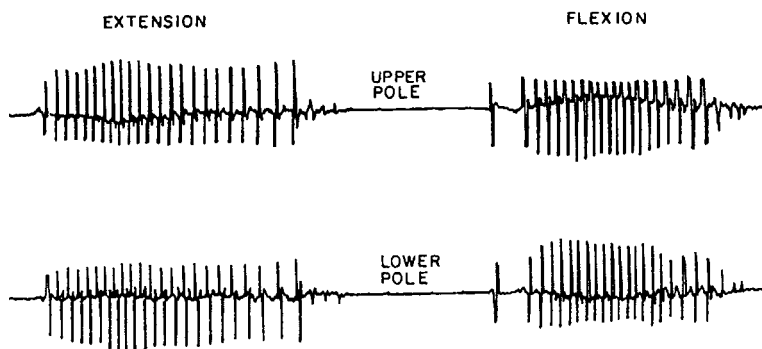

Figure 7.32   Simultaneously recorded PPC signals from the upper and lower poles of the patella during extension and flexion of the leg. The duration of the signal was not specified. Reproduced with permission from D.E. Beverland, W.G. Kernohan, G.F. McCoy, and R.A.B. Mollan, What is physiological patellofemoral crepitus?, *Proceedings of XIV International Conference on Medical and Biological Engineering and VII International Conference on Medical Physics,* pp 1249–1250, Espoo, Finland, 1985. ©IFMBE

wheel to represent the trochlear surface of the femur, on top of which was tensioned a rectangular piece of hardboard to represent the patella.

It was argued that, as the wheel in the model is slowly rotated clockwise (representing extension), it would initially stick to the overlying patella (hardboard) due to static friction. This would tend to impart an anticlockwise rotatory motion, as a rotating cogwheel would impart an opposite rotation to a cog in contact with it (as illustrated in the upper right-hand corner of Figure 7.33). The upper end of the patella would then move toward the wheel. A point would be reached where the static friction would be overcome, when the patella would slip and the rotation is suddenly reversed, with the upper pole jerking outward and the lower pole jerking inward. The actions would be the opposite to those described above in the case of flexion. The mechanical model was shown to generate signals similar to those recorded from subjects, thereby confirming the *stick–slip* frictional model for the generation of PPC signals.
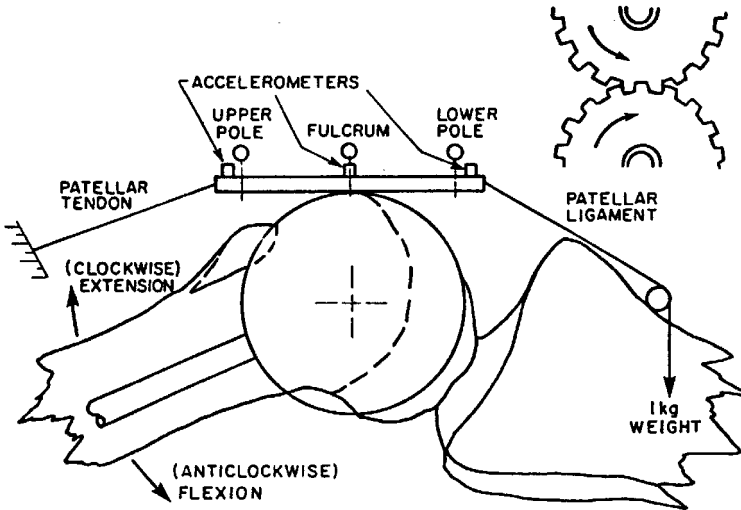

Figure 7.33   Apparatus to mimic the generation of PPC signals via a stick–slip frictional model. Reproduced with permission from D.E. Beverland, W.G. Kernohan, G.F. McCoy, and R.A.B. Mollan, What is physiological patellofemoral crepitus?, *Proceedings of XIV International Conference on Medical and Biological Engineering and VII International Conference on Medical Physics,* pp 1249–1250, Espoo, Finland, 1985. ©IFMBE

7.8   Application: Analysis of Heart rate Variability

Problem: *Explore the applicability of Fourier spectral analysis methods to study HRV.*

Solution: DeBoer et al. [137] applied Fourier analysis techniques to two types of measures derived from heart-rate data (see also Akselrod et al. [377]). They noted that the standard Fourier analysis methods cannot be applied directly to a series of point events. Therefore, they derived three types of signals from trains of ECG beats as illustrated in Figure 7.1.

The *interval spectrum* was derived by computing the Fourier spectrum of the interval series, normalized as $\tilde{I}_k = (I_k - \bar{I}) / \bar{I}$, where $\bar{I}$ is the mean interval length. The frequency axis was scaled by considering the time-domain data to be spaced at distances equal to the mean interval length $\bar{I}$, that is, the effective sampling frequency is $1 / \bar{I}$.

The *spectrum of counts* was derived by taking the Fourier transform of the impulse-train representation, derived from RR interval data as shown in Figure 7.1. The signal was normalized and scaled as $\tilde{s}(t) = \sum [\bar{I} \ \delta(t - t_k)] - N$, where N is the number of data samples, and the Fourier transform was computed. The spectra computed were smoothed with a 27-point rectangular window. DeBoer et al. demonstrated that the two spectra exhibit similar characteristics under certain conditions of slow or slight modulation of the data about the mean heart rate.

The RR interval data of a subject breathing freely and the two spectra derived from the data are shown in Figure 7.34. Three peaks are seen in both the spectra, which were explained as follows [137]:

- the effect of respiration at about $0.3 \ Hz$;

- the peak at $0.1 \ Hz$ related to $10 \ s$ waves seen in the BP; and

- a peak at a frequency lower than $0.1 \ Hz$ caused by the thermoregulatory system.

Figure 7.35 shows the RR interval data and spectra for a subject breathing at a fixed rate of $0.16 \ Hz$. The spectra display well-defined peaks at both the average heart rate ($1.06 \ Hz$) and at the breathing rate, as well as their harmonics. The spectra clearly illustrate the effect of respiration on the heart rate, and may be used to analyze the coupling between the cardiovascular and respiratory systems (see Section 2.2.4).

Note that direct Fourier analysis of a stream of ECG signals will not provide the same information as above. The reduced representation (model) of the RR interval data, as illustrated in Figure 7.1, has permitted Fourier analysis of the heart rate and its relationship with respiration. The methods have application in studies on HRV [134–136, 138, 139, 145, 146].

As described in Section 2.4.3, Mendez et al. [166] proposed a method for the detection of OSA based on the ECG recorded during sleep. Features that gave good results in the detection of OSA included coherence between the RR interval data and the QRS area as well as measures of HRV.
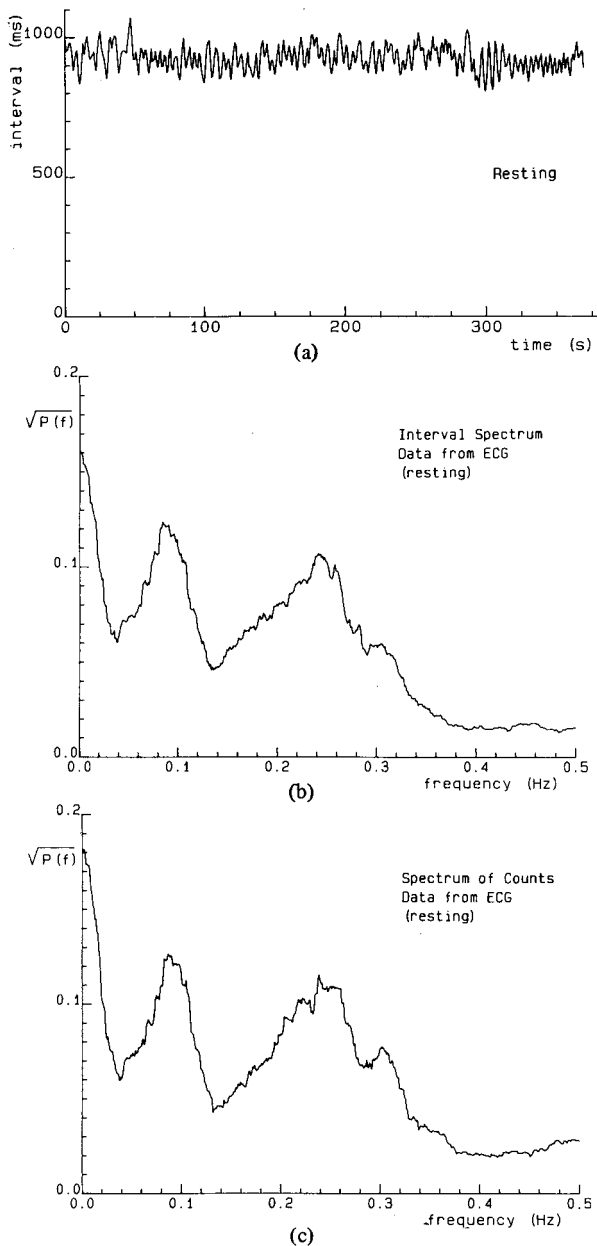
Figure 7.34   (a) $400 \ RR$ interval values from a healthy subject breathing freely. (b) Interval spectrum computed from a total of 940 intervals, including the 400 shown in (a) at the beginning.   (c) Spectrum of counts.   The spectra are shown for the range $0 - 0.5 \ Hz$ only.   Reproduced with permission from R.W. DeBoer, J.M. Karemaker, and J. Strackee, Comparing spectra of a series of point events particularly for heart rate variability studies, *IEEE Transactions on Biomedical Engineering,* 31(4):384–387, 1984. ©IEEE.
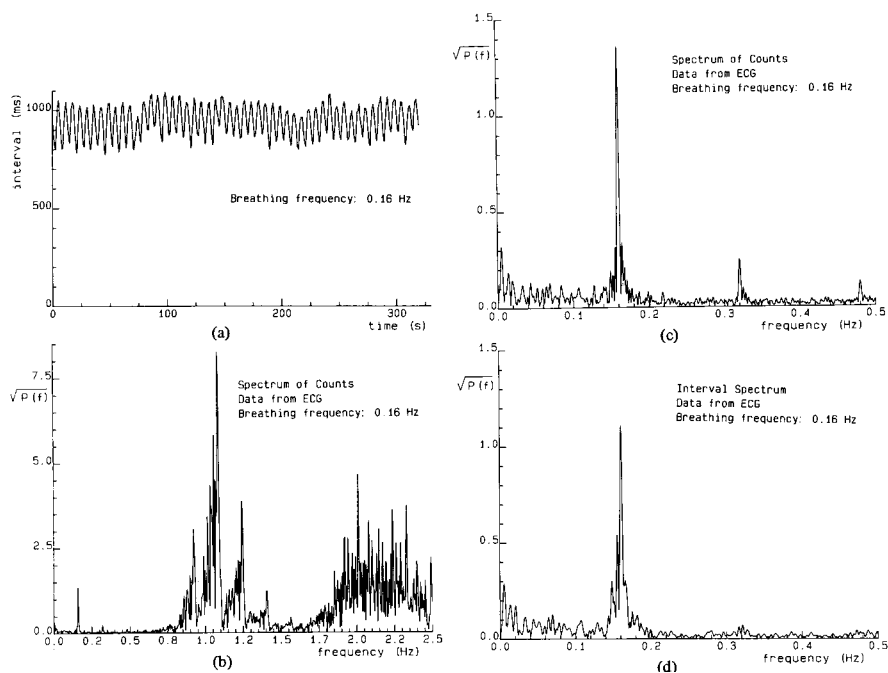
Figure 7.35   (a) 340 RR interval values from a healthy subject breathing at a fixed rate of 0.16 Hz. (b) Spectrum of counts for the range $0 - 2.5$ Hz. (c) Spectrum of counts for the range $0 - 0.5$ Hz. (d) Interval spectrum. Reproduced with permission from R.W. DeBoer, J.M. Karemaker, and J. Strackee, Comparing spectra of a series of point events particularly for heart rate variability studies, *IEEE Transactions on Biomedical Engineering,* 31(4):384–387, 1984. ©IEEE.

7.9 Application: Spectral Modeling and Analysis of PCG Signals

Iwata et al. [378, 379] applied AR modeling and parametric spectral analysis techniques to PCG signals for the detection of murmurs as well as the detection of the onset of S1 and S2. Their techniques included AR modeling, extraction of the dominant poles for pattern classification, and spectral tracking, which are explained in the following paragraphs.

Dominant poles: After the $a_k, k = 1, 2, \ldots, P$, coefficients of an all-pole or AR model of order P have been computed, the polynomial $A(z)$ may be factorized and solved to obtain the locations of the poles $p_k, k = 1, 2, \ldots, P$, of the system. The closer a pole is to the unit circle in the z-plane, the narrower its bandwidth and the stronger its contribution to the impulse response of the system. Poles that are close to the unit circle may be related to the resonance frequencies of the system, and used in system identification and pattern recognition.

In view of the nonstationary nature of the signal, Iwata et al. [378] computed a new model with order $P = 8$ for every window or frame of duration $25\ ms$, allowing an overlap of $12.5\ ms$ between adjacent frames (with the sampling rate $f_s = 2\ kHz$). The frequency of a pole p_k was calculated as

$$f_k = \frac{\angle p_k}{2\pi}\ f_s, \tag{7.137}$$

and its bandwidth was calculated as

$$bw_k = \frac{\log |p_k|}{\pi}\ f_s. \tag{7.138}$$

Conditions based upon the difference in the spectral power estimate of the model from one frame to the next, and the existence of poles with $f_k < 300\ Hz$ with the minimal bandwidth for the model considered, were used to segment each PCG signal into four phases: S1, a systolic phase spanning the S1 – S2 interval, S2, and a diastolic phase spanning the interval from one S2 to the following S1. (See also Section 4.10.)

Figures 7.36 and 7.37 show the PCG signals, spectral contours, the spectral power estimate, and the dominant poles for a normal subject and a patient with murmur due to patent ductus arteriosus. Most of the dominant poles of the model for the normal subject are below $300\ Hz$. The model for the patient with murmur indicates many dominant poles above $300\ Hz$.

The mean and SD of the poles with $bw_k < 80\ Hz$ of the model of each PCG phase were computed and used for pattern classification. The five coefficients of a fourth-order polynomial fitted to the series of spectral power estimates of the models for each phase were also used as features. Twenty-six out of 29 design samples and 14 out of 19 test samples were correctly classified. However, the number of cases was low compared to the number of features used in most of the six categories present in the samples (normal, aortic insufficiency, mitral insufficiency, mitral stenosis, aortic stenosis combined with insufficiency, and atrial septal defect).

Spectral tracking: In another application of AR modeling for the analysis of PCG signals, Iwata et al. [379] proposed a spectral-tracking procedure based upon AR modeling to detect S1 and S2. PCG signals were recorded at the apex with a high-pass filter that, at $100\ Hz$, had a gain $-40\ dB$ below the peak gain at $300\ Hz$ (labeled Ap-H). The signals were lowpass filtered with a gain of $-20\ dB$ at $1,000\ Hz$ and sampled at $2\ kHz$. The AR model was computed with order $P = 8$ for frames of length $25\ ms$; the frame-advance interval was only $5\ ms$. The model PSD was computed as

$$\tilde{S}(\omega) = \frac{\sigma_r^2}{2\pi} \frac{1}{\sum_{k=0}^{P} \phi_a(k)\ \cos(\omega T)}, \tag{7.139}$$

where

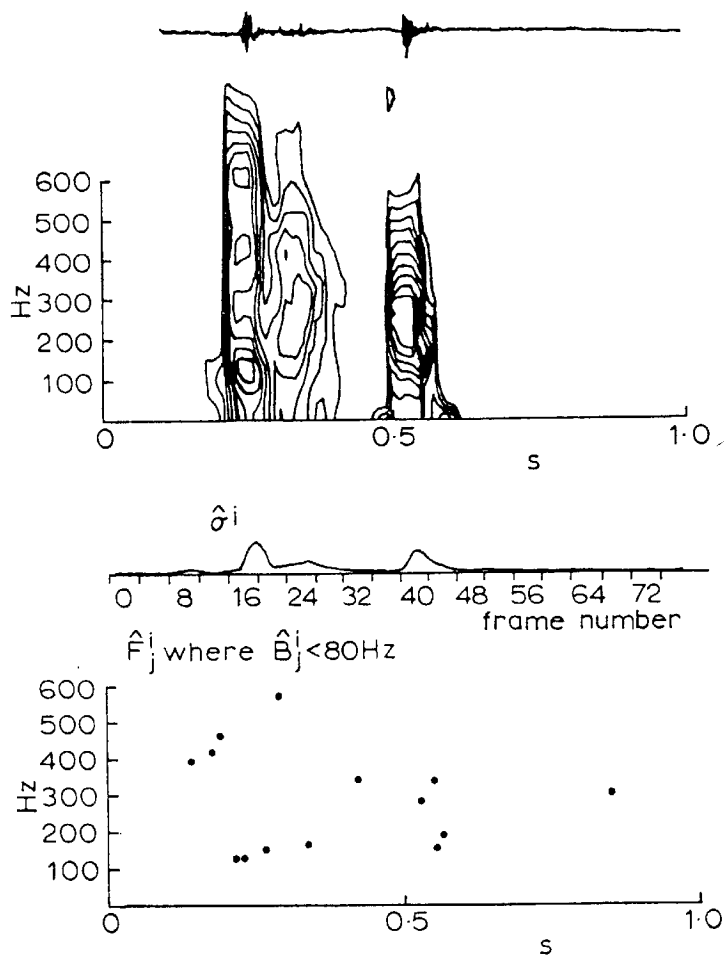$$\phi_a(k) = \sum_{j=0}^{P-k} a_j\ a_{j+k}, \tag{7.140}$$

Figure 7.36   Illustration of feature extraction based upon all-pole modeling of a normal PCG signal. From top to bottom: PCG signal; model spectrum in the form of isointensity contours; model spectral power estimate $\hat{\sigma}^i$, where i refers to the frame number; the frequencies $\hat{F}_j^i$ of the dominant poles with bandwidth $\hat{B}_j^i < 80 \; Hz$. Reproduced with permission from A. Iwata, N. Suzumara, and K. Ikegaya, Pattern classification of the phonocardiogram using linear prediction analysis, *Medical & Biological Engineering & Computing,* 15:407–412, 1977 ©IFMBE.
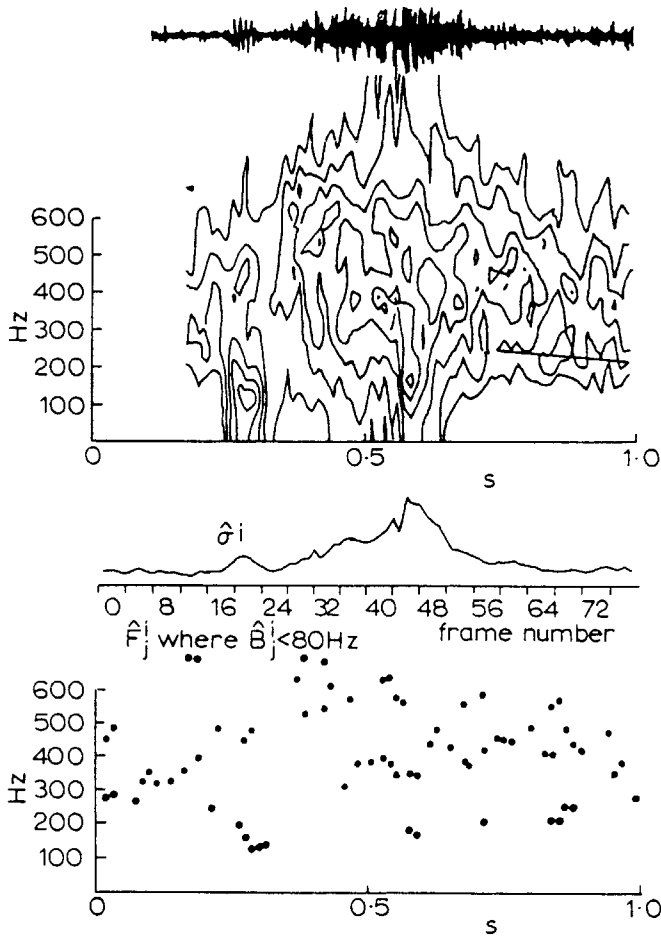
Figure 7.37   Illustration of feature extraction based upon all-pole modeling of the PCG signal of a patient with murmur due to patent ductus arteriosus. From top to bottom: PCG signal; model spectrum in the form of isointensity contours; model spectral power estimate $\hat{\sigma}^i$, where i refers to the frame number; the frequencies $\hat{F}^i_j$ of the dominant poles with bandwidth $\hat{B}^i_j < 80 \ Hz$. Reproduced with permission from A. Iwata, N. Suzumara, and K. Ikegaya, Pattern classification of the phonocardiogram using linear prediction analysis, *Medical & Biological Engineering & Computing,* 15:407–412, 1977 ⓒIFMBE.

with a_k being the AR-model coefficients, $P = 8$, $T = 0.5 \; ms$, σ_r^2 being the model residual energy (error), and $a_0 = 1$.

Based upon a study of the spectra of 69 normal and abnormal PCG signals, Iwata et al. [379] found the mean peak frequency of S1 to be $127 \; Hz$ and that of S2 to be $170 \; Hz$; it should be noted that the PCG signals were highpass filtered (as described in the preceding paragraph) at the time of data acquisition. The model spectral power at $100 \; Hz$ was used as the tracking function to detect S1: The peak in the tracking function after the location t_R of the R wave in the ECG was taken to be the position of S1. The tracking function to detect S2 was based upon the spectral power at $150 \; Hz$; the peak in the interval $t_R + 0.25RR \leq t \leq t_R + 0.6RR$, where RR is the interbeat interval, was treated as the position of S2. The use of a normalized spectral density function based upon the AR-model coefficients but without the σ_r^2 factor in Equation 7.139 was recommended, in order to overcome problems due to the occurrence of murmurs close to S2.

Figure 7.38 illustrates the performance of the tracking procedure with a normal PCG signal. The peaks in the $100 \; Hz$ and $150 \; Hz$ spectral-tracking functions (lowest traces) coincide well with the timing instants of S1 and S2, respectively. Figure 7.39 illustrates the application of the tracking procedure to the PCG signal of a patient with mitral insufficiency. The systolic murmur completely fills the interval between S1 and S2, and no separation is seen between the sounds and the murmur. Whereas the $150 \; Hz$ spectral-tracking function labeled (b) in the figure does not demonstrate a clear peak related to S2, the normalized spectral-tracking function labeled (c) shows a clear peak corresponding to S2. The two additional PCG traces shown at the bottom of the figure (labeled Ap-L for the apex channel including more low-frequency components with a gain of $-20 \; dB$ at $40 \; Hz$, and 3L-H for a channel recorded at the third left intercostal space with the same bandwidth as the Ap-H signal) illustrate S2 more distinctly than the Ap-H signal, confirming the peak location of the spectral-tracking function labeled (c) in the figure.

7.10   Application: Detection of Coronary Artery Disease

The diastolic segment of a normal PCG signal after S2 is typically silent; in particular, the central portion of the diastolic segment after the possible occurrence of AV valve-opening snaps is silent. Akay et al. [130] conjectured that blood flow in the coronary arteries is at its maximum during middiastole, and further that the effects of coronary artery disease, such as occlusion or stenosis, could present high-frequency sounds in this period due to turbulent blood flow (see Section 7.7.2).

Akay et al. [130] studied the spectra of middiastolic segments of the PCGs, averaged over $20 - 30$ beats, of normal subjects and patients with coronary artery disease confirmed by angiography. It was found that the PCG signals in the case of coronary artery disease exhibited greater portions of their energy above $300 \; Hz$ than the normal signals.

Figure 7.40 illustrates the AR-model spectra of two normal subjects and two patients with coronary artery disease. The signals related to coronary artery disease are
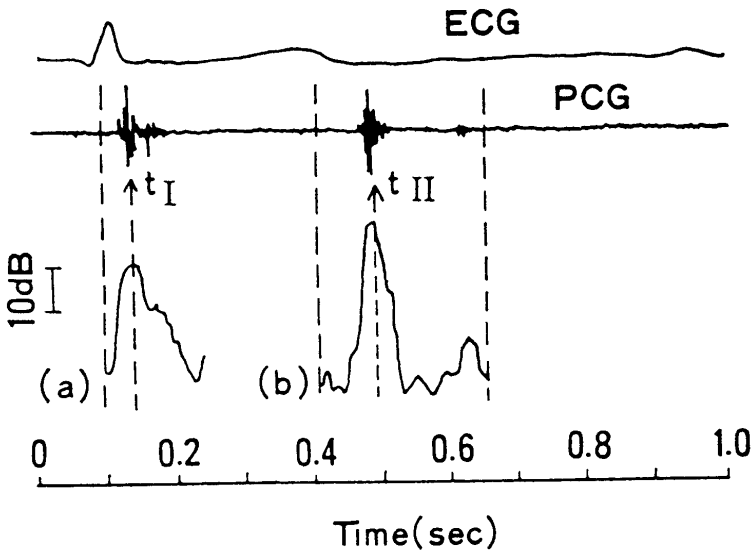
Figure 7.38 Illustration of the detection of S1 and S2 via spectral tracking based upon all-pole modeling of a normal PCG signal. From top to bottom: ECG signal; PCG signal; spectral-tracking functions at (a) 100 Hz for S1 and (b) 150 Hz for S2. The S1 and S2 locations detected are marked as t_I and t_{II}, respectively. Reproduced with permission from A. Iwata, N. Ishii, N. Suzumara, and K. Ikegaya, Algorithm for detecting the first and the second heart sounds by spectral tracking, *Medical & Biological Engineering & Computing,* 18:19–26, 1980 ©IFMBE.

seen to possess a high-frequency peak in the range $400 - 600$ Hz that is not evident in the normal cases.

Akay et al. [380] further found that the high relative-power levels of resonance frequencies in the range of $400 - 600$ Hz that were evident in patients with coronary artery disease were reduced after angioplasty. Figure 7.41 shows the spectra of the diastolic heart sounds of a patient before and after coronary artery occlusion was corrected by angioplasty. It may be readily observed that the high-frequency components that were present before surgery ("preang.") are not present after the treatment ("postang."). (The minimum-norm method of PSD estimation used by Akay et al. [380] — labeled as "MINORM" in the figure — is not discussed in this book.)

7.11 Remarks

In this chapter, we have studied how mathematical models may be derived to represent physiological processes that generate biomedical signals and, furthermore, how the models may be related to changes in signal characteristics due to functional and
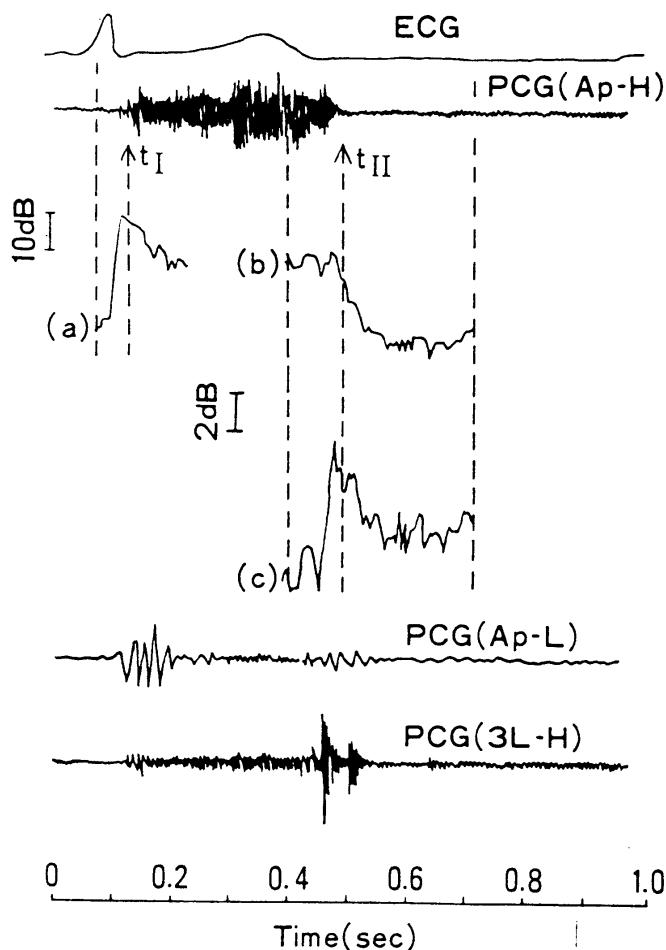
Figure 7.39   Illustration of the detection of S1 and S2 via spectral tracking based upon all-pole modeling of a PCG signal with systolic murmur due to mitral insufficiency. From top to bottom: ECG signal; PCG (Ap-H) signal; spectral-tracking functions at (a) $100~Hz$ for S1, (b) $150~Hz$ for S2, and (c) normalized spectral-tracking function at $150~Hz$ for S2; PCG (Ap-L) signal from the apex with more low-frequency components included; PCG (3L-H) signal from the third left intercostal space with the same filters as for Ap-H. The S1 and S2 locations detected are marked as t_I and t_{II}, respectively. Reproduced with permission from A. Iwata, N. Ishii, N. Suzumara, and K. Ikegaya, Algorithm for detecting the first and the second heart sounds by spectral tracking, *Medical & Biological Engineering & Computing,* 18:19–26, 1980 ©IFMBE.
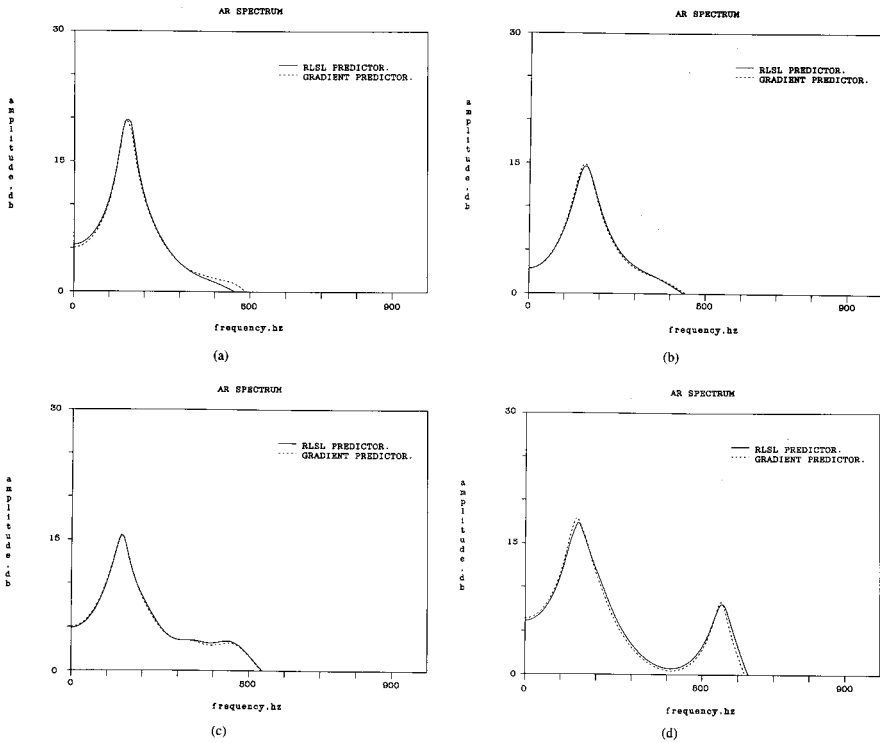
Figure 7.40 Diastolic heart sound spectra of (a, b) two normal subjects and (c, d) two patients with coronary artery disease. The method of estimating AR models identified in the figure as "RLSL" is described in Section 8.6.2; the gradient predictor method is not discussed in this book. Reproduced with permission from A.M. Akay, J.L. Semmlow, W. Welkowitz, M.D. Bauer, and J.B. Kostis, Detection of coronary occlusions using autoregressive modeling of diastolic heart sounds, *IEEE Transactions on Biomedical Engineering,* 37(4):366–373, 1990. ©IEEE.

pathological processes. The important point to note in the modeling approach is that the models provide a small number of *parameters* that characterize the signal and/or system of interest; the modeling approach is, therefore, useful in *parametric analysis* of signals and systems. As the number of parameters derived is usually much smaller than the number of signal samples, the modeling approach could also assist in data compression and compact representation of signals and related information.

Pole–zero models could be used to view physiological systems as control systems. Pathological states may be derived or simulated by modifying the parameters of the related models. Models of signals and systems are also useful in the design and control of prostheses.

A combination of mathematical modeling with electromechanical modeling can allow the inclusion of physical parameters, such as the diameter of a blood vessel,
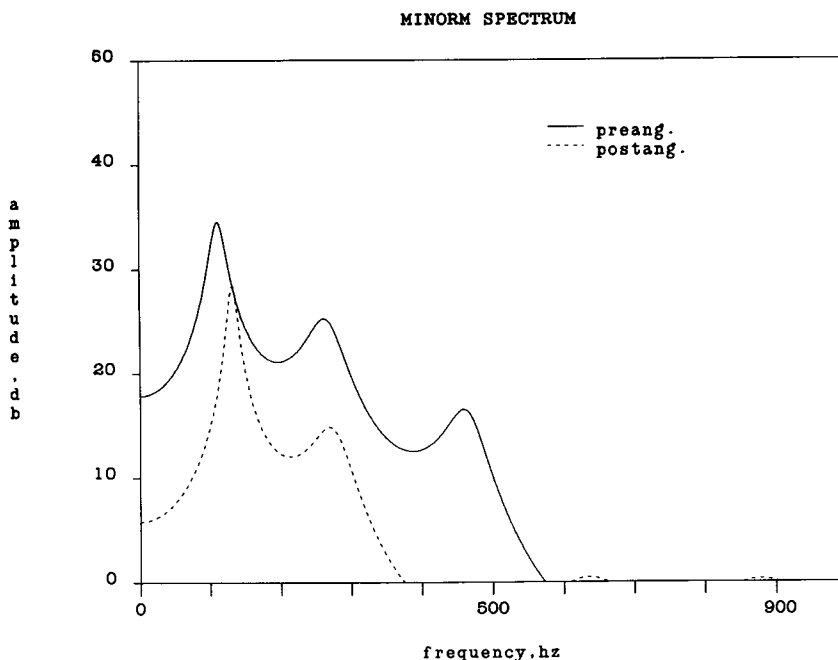
MINORM SPECTRUM

Figure 7.41   Diastolic heart sound spectra before (preang.) and after angioplasty (postang.) of a patient for whom coronary artery occlusion was corrected. (The minimum-norm method of PSD estimation used by Akay et al. [380] — labeled as "MINORM" in the figure — is not discussed in this book.) Reproduced with permission from A.M. Akay, J.L. Semmlow, W. Welkowitz, M.D. Bauer, and J.B. Kostis, Noninvasive detection of coronary stenoses before and after angioplasty using eigenvector methods, *IEEE Transactions on Biomedical Engineering,* 37(11):1095–1104, 1990. ©IEEE.

constriction due to plaque, stiffness due to stenosis, and friction coefficient. Although accurate estimation of such parameters for human subjects may not always be possible, the models could lead to better understanding of the related biomedical signals and systems.

7.12   Study Questions and Problems

1. Consider the LP model given by $\tilde{y}(n) = a\,y(n-1)$. Define the prediction error, and derive the optimal value for a by minimizing the TSE.

2. The AR-model coefficients of a signal are $a_0 = 1, a_1 = 1, a_2 = 0.5$. What is the transfer function of the model? Draw the pole–zero diagram of the model. What are the resonance frequencies of the system?

3. The AR-model coefficient vectors of a number of signals are made available to you. Propose two measures to compare the signals for (a) similarity, and (b) dissimilarity.

4. In AR modeling of signals, show why setting the derivative of the TSE with respect to any coefficient to zero will always lead to the minimum error (and not the maximum).

5. What type of a filter can convert the autocorrelation matrix of a signal to a diagonal matrix?

6. A biomedical signal is sampled at $500\ Hz$ and subjected to AR modeling. The poles of the model are determined to be at $0.4 \pm j0.5$ and $-0.7 \pm j0.6$. (a) Derive the transfer function of the model. (b) Derive the difference equation in the time domain. (c) What are the resonance frequencies of the system that is producing the signal?

7. A model is described by the relationship $y(n) = x(n) + 0.5\,x(n-1) + 0.25\,x(n-2)$, where $x(n)$ is the input and $y(n)$ is the output. What is the type of this system among AR, MA, and ARMA systems? What is the model order? What is its transfer function? Draw the pole–zero diagram of the system. Comment upon the stability of the system.

8. A model is described by the relationship $y(n) = -0.5\,y(n-1) - y(n-2) + x(n) + 0.5\,x(n-1) - x(n-2)$, where $x(n)$ is the input and $y(n)$ is the output. What is the type of this system among AR, MA, and ARMA systems? What is the model order? What is its transfer function? Draw the pole–zero diagram of the system. Comment upon the stability of the system.

7.13   Laboratory Exercises and Projects

Note: Data files related to the exercises are available at the site

$$\texttt{http://people.ucalgary.ca/\~{}ranga/enel563}$$

1. The file safety.wav contains the speech signal for the word "safety" uttered by a male speaker, sampled at $8\ kHz$ (see the file safety.m). The signal has a significant amount of background noise (as it was recorded in a computer laboratory). Develop procedures to segment the signal into voiced, unvoiced, and silence (background noise) portions using short-time RMS, turns count, or ZCR measures.

   Apply the AR modeling procedure to each segment; the command *lpc* in MATLAB$^{\circledR}$ may be used for this purpose. Compute the model PSD for each segment. Compare the model PSD with the Fourier PSD for each segment. What are the advantages and disadvantages of the model-based PSD in the case of voiced and unvoiced sounds?

2. Derive the poles of the models you obtained in the preceding problem. Express each pole in terms of not only its z-plane coordinates but also its frequency and bandwidth. Study the variations in the pole positions as the type of the sound varies from one segment to the next over the duration of the signal.

3. The files pec1.dat, pec33.dat, and pec52.dat give three-channel recordings of the PCG, ECG, and carotid pulse signals (sampled at $1,000\ Hz$; you may read the signals using the program in the file plotpec.m). The signals in pec1.dat and pec52.dat are normal; the PCG signal in pecg33.dat has systolic murmur, and is of a patient suspected to have pulmonary stenosis, ventricular septal defect, and pulmonary hypertension.

   Segment each signal into its systolic and diastolic parts. Apply the AR modeling procedure to each segment and derive the model PSD. Compare the result with the corresponding PSDs obtained using Welch's procedure.

4. Derive the poles of the models you obtained in the preceding problem. Express each pole in terms of not only its z-plane coordinates but also its frequency and bandwidth. Study the variations in the pole positions from the systolic part to the diastolic part of each signal. What are the major differences between the pole plots for the normal cases and the case with murmur?

5. The files ECG3, ECG4, ECG5, and ECG6 contain ECG signals sampled at the rate of $200\ Hz$ (see the file ECGS.m). Apply the Pan–Tompkins method for QRS detection to each signal. Create impulse sequences including a delta function at every QRS location for the four signals. Create also the interval series for each signal as illustrated in Figure 7.1. Compute the spectra corresponding to the two representations of cardiac rhythm and study their relationship to the heart rate and its variability in each case.