CHAPTER 6

FREQUENCY DOMAIN CHARACTERIZATION OF SIGNALS AND SYSTEMS

Many biomedical systems exhibit innate rhythms and periodicity that are more readily expressed and appreciated in terms of frequency than temporal measures. As a basic example, consider cardiac function: We express cardiac rhythm more conveniently in terms of *beats per minute* — a measure of the frequency of occurrence or the rate of repetition — than in terms of the duration of a beat or the interval between beats in seconds (the RR interval). A cardiac rhythm expressed as $72\ bpm$ is more easily understood than a statement of the corresponding RR interval as $0.833\ s$. By the same token, the notion of an EEG rhythm is more easily conveyed by description in *cycles per second* in lay terms, or in *Hertz (Hz)* in technical terms. Even engineers would find a frequency-domain expression easier to appreciate than a time-domain description, such as an alpha rhythm having a frequency of $11.5\ Hz$ versus the equivalent period of $0.087\ s$.

When the signal being studied is made up of discrete (that is, separate and distinct) events in time, such as the ECG or a train of SMUAPs, the basic rhythm or rate of activity present in the signal can be assessed directly in the time domain. On the other hand, signals such as the PCG display complex or complicated patterns in the time domain that do not facilitate ready appreciation of their frequency-domain char-

Biomedical Signal Analysis, Second Edition. By Rangaraj M. Rangayyan

acteristics; furthermore, the time-domain waveforms may differ from one occurrence of the signal (one heart beat) to another.

The PCG provides an interesting example of a signal with multiple frequency-domain features: In addition to the beat-to-beat periodicity or rhythm, the heart sounds within a cardiac cycle exhibit *resonance*. Due to the multicompartmental nature of the cardiac system, we should expect heart sounds to possess multiple resonance frequencies; this leads to the need to describe the PCG not only in terms of a rhythm (the heart rate) or a single resonance frequency, but also a composite *spectrum* of several dominant or resonance frequencies. Furthermore, constrained flow of blood through an orifice such as a septal defect or across a stenosed valve acting as a baffle could lead to turbulence, resulting in *wideband noise*. In the case of noise-like murmurs, we would be able to identify neither rhythms nor resonance frequencies: The need arises to consider the distribution of the signal's energy or power over a wide band of frequencies, leading to the notion of the power *spectral density* function.

We have seen in Chapter 3 that it is often more convenient and meaningful to describe filters in terms of their transfer function or frequency response — $H(s)$, $H(z)$, $H(\omega)$, or $H(f)$ — than in terms of their impulse response $h(t)$ or the time-domain input–output relationship (difference equation). Furthermore, we saw in Section 4.4 that it is easier to interpret the PSDs of EEG waves than it is to interpret their theoretically equivalent ACFs. The Fourier and other transforms provide an invertible or reversible transformation from the time domain to the frequency domain (and vice-versa). Therefore, it may be argued that no new information is created by taking a given signal from the time domain to the frequency domain. However, the distribution of the energy or power of the signal in the frequency domain that is provided by the Fourier transform — the spectrum or PSD of the signal — facilitates better analysis and description of the frequency-domain characteristics of the signal. The PSD of a signal is not only useful in analyzing the signal, but also in designing amplifiers, filters, data-acquisition and transmission systems, and signal processing systems to treat the signal appropriately. We have seen in Section 3.8 that we need not only the signal PSD but also the noise PSD in order to be able to implement the optimal Wiener filter.

The treatment of biomedical signals as stochastic processes provides flexibility and a sense of generality in analysis, but imposes conditions and requirements in the estimation of their statistics including the ACF and PSD. In the present chapter, we shall investigate methods to estimate the PSD and frequency-domain parameters of biomedical signals and systems. We shall also study methods to derive spectral parameters that can characterize the given signal as well as the system that generated the signal. The motivation for the study, as always, shall be to distinguish between normal and abnormal signals or systems, and the potential use of the methods in diagnosis.

6.1 Problem Statement

Investigate the potential use of the Fourier spectrum and parameters derived thereof in the analysis of biomedical signals. Identify physiological and pathological processes that could modify the frequency content of the corresponding signals. Outline the signal processing tasks needed to perform spectral analysis of biomedical signals and systems.

As in the preceding chapters, the problem statement given above is generic and represents the theme of the present chapter. The various signal analysis techniques described and the examples used for illustration in the following sections will address the points raised in the problem statement, with attention to specific problems and techniques.

6.2 Illustration of the Problem with Case Studies

6.2.1 The effect of myocardial elasticity on heart sound spectra

The first and second heart sounds — S1 and S2 — are typically composed of low-frequency components; this is to be expected due to the fluid-filled and elastic nature of the cardiohemic system. Sakai et al. [310] processed recorded heart sound signals by using tunable bandpass filters (with a bandwidth of 20 Hz, tuned over the range $20-40$ Hz to $400-420$ Hz), and estimated the frequency distributions of S1 and S2. They found the heart sound spectra to be maximum in the $20-40$ Hz band, that S1 had a tendency to demonstrate peaks at lower frequencies than those of S2, and that S2 exhibited a "gentle peaking" between 60 Hz and 220 Hz.

Gerbarg et al. [247, 248] developed a computer program to simulate a filter bank and obtained averaged power spectra of S1 and S2 of $1,000$ adult males, 32 high-school children, and 75 patients in a hospital. The averaged PSDs of S1 and S2 obtained by them indicated peak power in the range $60-70$ Hz, and relative power levels lower than -10 dB beyond 150 Hz. The PSD of S2 displayed slightly more high-frequency energy than that of S1.

Frome and Frederickson [311] applied the FFT to the analysis of first and second heart sounds. They described how segmented S1 and S2 data may be combined into a single complex signal and how a single FFT may be used to obtain the FFTs of the two signals. Computer data processing techniques were described to obtain smoothed, averaged periodograms (described in Section 6.3.1) of S1 and S2 separately.

Yoganathan et al. [312] applied the FFT for the analysis of S1 of 29 normal subjects. The FFT spectra of 250 ms windows containing S1 were averaged over 15 beats for each subject. It was found that the frequency spectrum of S1 contains peaks in a low-frequency range ($10-50$ Hz) and a medium-frequency range ($50-140$ Hz) [312]. In a related study [313], the spectrum of S2 was observed to contain peaks in low-frequency ($10-80$ Hz), medium-frequency ($80-220$ Hz), and high-frequency ranges ($220-400$ Hz). It has been suggested that the resonance peaks in the spectra

may be related to the elastic properties of the heart muscles and the dynamic events causing the various components of S1 and S2 (see Section 1.2.9).

Adolph et al. [314] used a dynamic spectrum analyzer to study the frequency content of S1 during the isovolumic contraction period. The center frequency of a filter with 20 Hz bandwidth was initially set to 30 Hz and then varied in 10 Hz increments up to 70 Hz. The outputs of the filters were averaged over the same (prerecorded) 10 consecutive beats. Finally, the ratios of the average peak voltage of the filtered outputs to that of the total S1 signal during the isovolumic contraction period were computed.

Adolph et al. hypothesized that the frequency content of S1 during the isovolumic contraction period should depend on the relative contributions of the mass and elasticity of the left ventricle. The mass of the left ventricle with its blood content remains constant during isovolumic contraction. Therefore, it was reasoned that the frequency content of S1 should decrease (that is, shift toward lower frequencies) in the case of diseases that reduce ventricular elasticity, such as myocardial infarction.

Figure 6.1 shows averaged S1 spectra for normal subjects and patients with acute or healed myocardial infarction; it is seen that the reduced elasticity due to myocardial infarction has reduced the relative content of power near 40 Hz. However, Adolph et al. also noted that an increase in ventricular mass as in the case of trained athletes, or a reduction in elasticity combined with an increase in the mass as in the case of myocardiopathy, could also cause a similar shift in the frequency content of S1. Regardless, they found that frequency analysis of S1 was of value in differentiating acute pulmonary embolism from acute myocardial infarction. Clarke et al. [315] also found reduction in the spectral energy of S1 to be a common accompaniment of myocardial ischemia.

6.2.2 Frequency analysis of murmurs to diagnose valvular defects

As we noted in Section 1.2.9, cardiovascular valvular defects and diseases cause high-frequency, noise-like sounds known as murmurs. Murmurs are often the only indicators of the early stages of certain cardiovascular diseases; prompt diagnosis could prevent further deterioration of the condition and possible complications.

We noted in Section 5.6.2 that zero-crossing analysis in the time domain was applied to assist in the detection of murmurs by Jacobs et al. [250] and Yokoi et al. [251]. Although ZCR increases with the presence of higher-frequency components, it does not yield a direct measure of the frequency content or the spectrum of the signal.

Application of electronic signal filtering techniques to analyze the frequency content of heart sounds and murmurs was initiated as early as the 1950s. Geckeler et al. [316] and McKusick et al. [317, 318] studied the applicability of the sound spectrograph for the analysis of heart sounds and murmurs. The sound spectrograph was developed in the late 1940s by Bell Telephone Laboratories as a tool to produce what was labeled as *visible speech*. The spectrograph used a bandpass filter (or a bank of bandpass filters) to determine the power of the given signal in each frequency band of interest. The signal was usually recorded and played back repeatedly

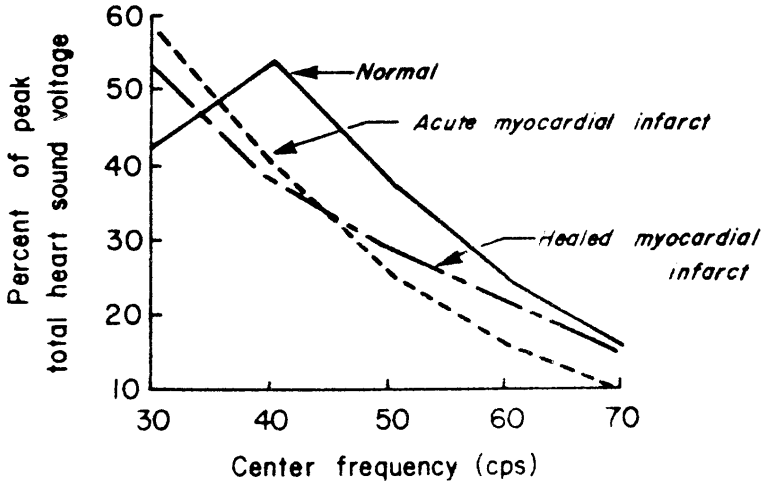

Figure 6.1 First heart sound spectra for normal, acute myocardial infarct, and healed myocardial infarct cases. The latter two cases exhibit an increased percentage of low-frequency components. Reproduced with permission from R.J. Adolph, J.F. Stephens, and K. Tanaka, The clinical value of frequency analysis of the first heart sound in myocardial infarction, *Circulation,* 41:1003–1014, 1970. ©American Heart Association.

as the center frequency of the bandpass filter was varied. The output was recorded on heat-sensitive or light-sensitive paper to produce a 2D distribution of frequency content of the signal at every instant of time as a gray-level image (essentially a TFD; see Section 8.4.1). Winer et al. [319] proposed isointensity contour plotting of the spectrogram instead of using variations in intensity (gray scale); they reported that, whereas normal heart sounds indicated the presence of regularity in the contours of equal intensity, abnormal sounds and murmurs produced irregular contour line structures with extensive "convolutions" and roughness. It was suggested that the cardiospectrograms (or spectral phonocardiography) could provide physiologic and pathologic information beyond that provided by auscultation, without suffering from the psychoacoustic impediments that affect human observers [33].

Yoshimura [320] used a tunable bandpass filter with low and high cutoff frequencies in the range $18 - 1,425\ Hz$ to process recorded PCG signals. It was determined that the diastolic *rumble* of mitral stenosis occupied the range $20 - 200\ Hz$, whereas the diastolic *blow* of aortic regurgitation spanned a much higher frequency range of $200 - 1,600\ Hz$ (although the characteristic range was $400 - 800\ Hz$).

Gerbarg et al. [247, 248] developed a computer program to simulate a filter bank and obtain power spectra of heart sounds and murmurs, with the aim of developing a system for screening to detect cardiovascular diseases. They argued that innocent (physiological) systolic murmur in children is limited to the first and middle thirds of the systolic interval between S1 and S2, whereas pathological systolic murmur due to mitral regurgitation is holosystolic. Therefore, they computed ratios of the

mean power of the *last third of systole* to the mean power of systole and also to a certain "standard" noise level. A ratio was also computed of the mean energy of systole to the mean energy of the PCG over the complete cardiac cycle. Gerbarg et al. obtained $78 - 91\%$ agreement of their computer classification based upon the three ratios defined above with clinical diagnosis of mitral regurgitation in different groups of subjects. Although they did not claim that a fully automated program for the diagnosis of mitral regurgitation had been developed, they indicated that the feasibility of computer-based diagnosis had been established, and that simulation of human auscultation had been partially achieved.

The specific problem of detection of the murmur due to aortic insufficiency in the presence of the murmur due to mitral stenosis was considered by van Vollenhoven et al. [321]. Aortic insufficiency causes an early diastolic murmur (with a blowing or hissing quality) that is best heard in the aortic area (second right-intercostal space, just right of the sternum), whereas the middiastolic rumbling murmur of mitral stenosis is best heard at the apex. A tunable bandpass filter with $50\ Hz$ bandwidth and center frequency tunable in steps of $50\ Hz$ was used by van Vollenhoven et al. to study the frequency content in a $100\ ms$ window during the diastolic phase of recorded PCG signals. They found that the murmur of mitral stenosis was limited in frequency content to less than $400\ Hz$, whereas the murmur in the case of aortic insufficiency combined with mitral stenosis had more high-frequency energy in the range $300 - 1,000\ Hz$.

Sarkady et al. [227] suggested synchronized averaging of the PSDs of PCG signals over several cardiac cycles computed using the FFT algorithm. Johnson et al. [322, 323] studied FFT-based PSDs of the systolic murmur due to aortic stenosis. They computed the PSDs of systolic windows of duration $86, 170$, and $341\ ms$, and averaged the results over 10 cardiac cycles. Johnson et al. hypothesized that higher murmur frequencies are generated as the severity of aortic stenosis increases. In their study of patients who underwent catheterization and cardiac fluoroscopy, the transvalvular systolic pressure gradient was measured during pull-back of the catheter from the left ventricle through the aortic valve and was found to be in the range $10 - 140\ mm\ of\ Hg$. Spectral power ratios (described in Section 6.4.2) were computed considering the band $25 - 75\ Hz$ as the constant area (CA) related to normal sounds and the band $75 - 150\ Hz$ as the predictive area (PA) related to murmurs.

Figure 6.2 illustrates the PSDs of four patients with aortic stenosis of different levels of severity. The PSDs in the figure are segmented into the CA and PA parts as described above; the transvalvular systolic pressure gradient (in $mm\ of\ Hg$) and the PA/CA spectral power ratio are also shown for each case. Johnson et al. found that the spectral power ratio increased linearly with the transvalvular systolic pressure gradient, and hence correlated well with the severity of aortic stenosis. The importance of recording the PCG in the aortic area, prefiltering the PCG to $25 - 1,500\ Hz$, and the selection of an appropriate systolic murmur window was discussed by Johnson et al. Although there were confounding factors, it was indicated that the noninvasive PCG-based technique could be useful in identifying the need for catheterization as well as follow-up of patients with aortic stenosis.

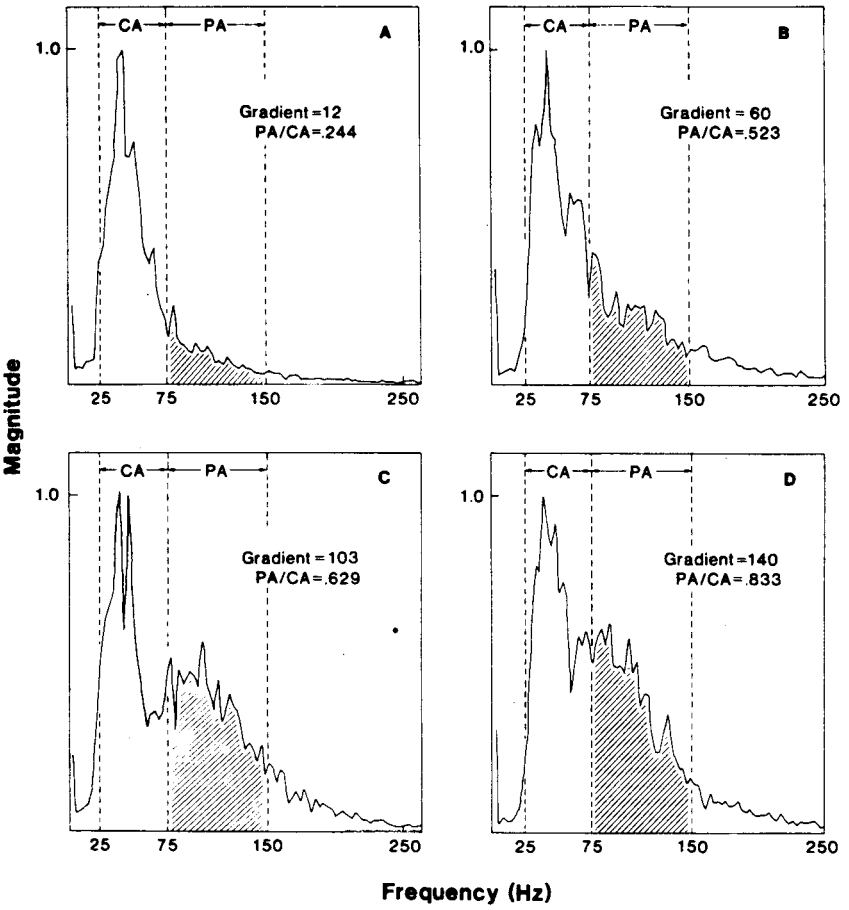

Figure 6.2 Averaged and normalized PSDs of four patients with aortic stenosis of different levels of severity. Each PSD is segmented into two parts: a constant area CA and a predictive area PA. The transvalvular systolic pressure gradient (measured via catheterization in $mm\ of\ Hg$) and the PA/CA spectral power ratio are shown for each case. Reproduced with permission from the American College of Cardiology, from G.R. Johnson, R.J. Adolph, and D.J. Campbell, Estimation of the severity of aortic valve stenosis by frequency analysis of the murmur, *Journal of the American College of Cardiology,* 1(5):1315–1323, 1983 ©Elsevier Science.

6.3 Estimation of the PSD

We encountered the ACF and CCF in Equations 3.17, 3.17, 3.20, and 4.27: The first equation cited here provides a general definition of the ACF as a statistical expectation or an integral over an indefinite duration; the last one treats the CCF as the projection of one signal onto another and neglects a scale factor that may be of no consequence in a given application. We shall now investigate more closely the procedures required to estimate the ACF, and hence the PSD, from finite-length signal records.

Let us consider a signal record of N samples: $x(n)$, $n = 0, 1, 2, \ldots, N-1$. In order to compute the time-averaged ACF $\phi_{xx}(m)$ for a delay of m samples, we need to form the product $x(n)\,x(n \pm m)$ and sum over the available range of data samples. The true ACF is given as $\phi_{xx}(m) = E[x(n)\,x(n+m)]$. Note that one of the copies of the signal entering the computation of the ACF should be conjugated if the signal is complex.

It is readily seen that we may sum from $n = 0$ to $n = N-1$ when computing $\phi_{xx}(0)$ with $x(n)\,x(n) = x^2(n)$. However, when computing $\phi_{xx}(1)$ with $x(n)\,x(n+1)$, we can only sum from $n = 0$ to $n = N-2$. As we apply a linear shift of m samples to one copy of the signal to compute $\phi_{xx}(\pm m)$, m samples of one of the copies of the signal drop out of the window of analysis indicated by the overlap between the two versions of the signal; see Figure 6.3. Therefore, only $N - |m|$ pairs of data samples are available to estimate the ACF for the delay of $\pm m$ samples. We then have a sample-mean estimate of the ACF given by

$$\phi_1(m) = \frac{1}{N-|m|} \sum_{n=0}^{N-|m|-1} x(n)\,x(n+m). \tag{6.1}$$

The subscript xx has been omitted in the equation given above; the subscript 1 indicates the use of one type of averaging scale factor in estimating the ACF. Oppenheim and Schafer [174] have shown that $\phi_1(m)$ is a consistent estimate of $\phi_{xx}(m)$: It has zero bias and has a variance that tends to zero as $N \to \infty$. However, the variance of the estimate becomes exceptionally large as m approaches N: Very few nonzero pairs of samples are then available to compute the ACF, and the estimate is useless.

An alternative definition of the ACF ignores the lack of $|m|$ nonzero pairs of samples, and applies the same scale factor for all delays, leading to

$$\phi_2(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x(n)\,x(n+m). \tag{6.2}$$

Note that the upper limit of summation in the above expression could be stated as $N-1$ with no effect on the result; the first or the last $|m|$ samples of $x(n)$ will not overlap with $x(n+m)$, and result in zero product terms. Oppenheim and Schafer [174] have shown that $\phi_2(m)$ has a bias equal to $\frac{m}{N}\phi_{xx}(m)$: The bias tends to the actual value being estimated as m approaches N, although the variance is almost independent of m and tends to zero as $N \to \infty$. Regardless, both the ACF

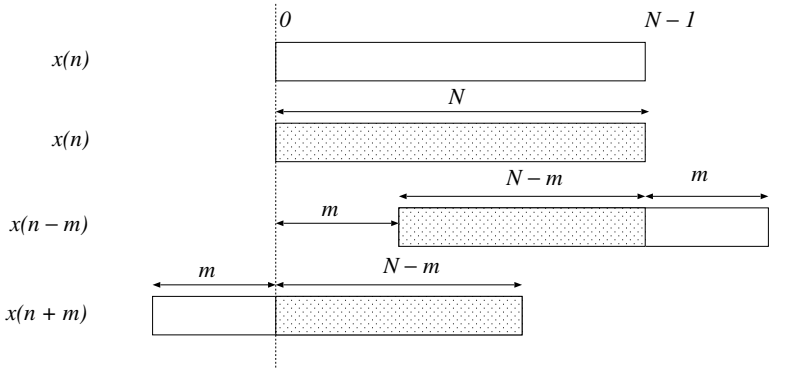

Figure 6.3 The effect of shifting in the computation of the ACF or CCF: as the shift increases, the number of available pairs of samples is reduced.

estimates are asymptotically unbiased [the bias of $\phi_2(m)$ tends to zero as $N \to \infty$], and they yield good estimates of the ACF as long as the number of samples N is large and $m << N$.

Note that the two ACF estimates $\phi_1(m)$ and $\phi_2(m)$ are interrelated as

$$\phi_2(m) = \frac{N - |m|}{N} \phi_1(m). \tag{6.3}$$

Thus, $\phi_2(m)$ is a scaled version of $\phi_1(m)$. However, since the scaling factor is a function of m, it is more commonly referred to as a *window*; more discussion on this interpretation is presented in Section 6.3.1. It should also be observed that the distinction between $\phi_1(m)$ and $\phi_2(m)$ is comparable to that between the unbiased and biased sample variance measures, where the division is by $N - 1$ or N, respectively, with N being the number of samples available.

See Bingham et al. [324], Welch [325], Kay and Marple [326], Thomson [327], Robinson [328], and Childers et al. [329] for discussions on several issues related to estimation of PSDs.

6.3.1 The periodogram

Since the PSD and the ACF are a Fourier transform pair, we may compute an estimate of the PSD as

$$S_2(\omega) = \sum_{m=-(N-1)}^{N-1} \phi_2(m) \exp(-j\omega m), \tag{6.4}$$

assuming that, indeed, the ACF is computed or available for $|m|$ up to $N - 1$. The Fourier transform of the signal $x(n)$, $n = 0, 1, 2, \ldots, N - 1$, is given as

$$X(\omega) = \sum_{n=0}^{N-1} x(n) \exp(-j\omega n). \tag{6.5}$$

It can be shown that

$$S_2(\omega) = \frac{1}{N}|X(\omega)|^2 . \tag{6.6}$$

The PSD estimate $S_2(\omega)$ is known as the *periodogram* of the signal $x(n)$ [174]. Oppenheim and Schafer [174] have shown that $S_2(\omega)$ is a biased estimate of the PSD, with

$$E[S_2(\omega)] = \sum_{m=-(N-1)}^{N-1} \frac{N-|m|}{N}\, \phi_{xx}(m) \exp(-j\omega m). \tag{6.7}$$

If we consider the Fourier transform of $\phi_1(m)$, we get a different estimate of the PSD as

$$S_1(\omega) = \sum_{m=-(N-1)}^{N-1} \phi_1(m) \exp(-j\omega m), \tag{6.8}$$

with the expected value [174]

$$E[S_1(\omega)] = \sum_{m=-(N-1)}^{N-1} \phi_{xx}(m) \exp(-j\omega m). \tag{6.9}$$

Because of the finite limits of the summation, $S_1(\omega)$ is a biased estimate of the PSD.

The two estimates $S_2(\omega)$ and $S_1(\omega)$ may be seen as the Fourier transforms of windowed ACFs, with the window functions being a triangular function, known as the Bartlett window, $w_B(m)$, in the first case and a rectangular window, $w_R(m)$, in the second case:

$$w_B(m) = \begin{cases} \frac{N-|m|}{N}, & |m| < N, \\ 0, & \text{otherwise}, \end{cases} \tag{6.10}$$

$$w_R(m) = \begin{cases} 1, & |m| < N, \\ 0, & \text{otherwise}. \end{cases} \tag{6.11}$$

Note that the windows defined above have a (nonzero) duration of $(2N-1)$ samples.

Since the ACF is multiplied with the window function, the PSD is convolved with the Fourier transform of the window function, leading to spectral leakage and loss of resolution (more details on windows follow in Section 6.3.3). The Fourier transforms of the Bartlett and rectangular windows are, respectively [174],

$$W_B(\omega) = \frac{1}{N}\left[\frac{\sin(\omega N/2)}{\sin(\omega/2)}\right]^2 \tag{6.12}$$

and

$$W_R(\omega) = \frac{\sin[\omega(2N-1)/2]}{\sin(\omega/2)}. \tag{6.13}$$

Oppenheim and Schafer [174] have shown that the periodogram has a variance that does not approach zero as $N \to \infty$; instead, the variance of the periodogram is of the order of σ_x^4 regardless of N. Thus, the periodogram is not a consistent estimate of the PSD.

6.3.2 The need for averaging

A common approach to reduce the variance of an estimate is to average over a number of independent estimates or observations. We have seen in Section 3.5.1 how the variance of the noise in noisy signals may be reduced by synchronized averaging over a number of observations of the corrupted signal. In a similar vein, a number of periodograms may be computed over multiple observations of a signal and averaged to obtain a better estimate of the PSD. It is necessary for the process to be stationary, at least during the period over which the periodograms are computed and averaged.

Problem: *How can we obtain an averaged periodogram when we are given only one signal record of finite duration?*

Solution: Oppenheim and Schafer [174] have described the following procedure, attributed to Bartlett, to average periodograms of segments of the given signal record:

1. Divide the given data sequence $x(n)$, $n = 0, 1, 2, \ldots, N-1$, into K segments of M samples each. We then have the segments given by

$$x_i(n) = x[n + (i-1)M], \;\; 0 \leq n \leq M-1, \;\; 1 \leq i \leq K. \tag{6.14}$$

2. Compute the periodogram of each segment as

$$S_i(\omega) = \frac{1}{M} \left| \sum_{n=0}^{M-1} x_i(n) \exp(-j\omega n) \right|^2, \;\; 1 \leq i \leq K. \tag{6.15}$$

   The Fourier transform in the equation given above is evaluated as a DFT (using the FFT algorithm) in practice.

3. If the ACF $\phi_{xx}(m)$ is negligible for $|m| > M$, the periodograms of the K segments of duration M samples each may be assumed to be mutually independent. Then, the Bartlett estimate $S_B(\omega)$ of the PSD is obtained as the sample mean of the K independent observations of the periodogram:

$$S_B(\omega) = \frac{1}{K} \sum_{i=1}^{K} S_i(\omega). \tag{6.16}$$

Oppenheim and Schafer [174] have shown that the expected value of the Bartlett estimate $S_B(\omega)$ is the convolution of the true PSD $S_{xx}(\omega)$ with the Fourier transform of the Bartlett window given in Equation 6.12 (with N replaced by M). The convolution relationship indicates the bias in the estimate, and has the effect of spectral smearing and leakage; the bias may, therefore, be interpreted as a loss in resolution. Although $S_B(\omega)$ is a biased estimate, its variance tends to zero as the number of segments K increases. Therefore, it is a consistent estimate.

When we have a (stationary) signal of fixed duration of N samples, we will face limitations on the number of segments K that we may obtain. While the variance of the estimate decreases as K is increased, it should be recognized that there is a

concomitant decrease in the number of samples M per segment. As M decreases, the main lobe of the Fourier transform of the Bartlett window (see Equation 6.12) widens; frequency resolution is lost because the estimate is the convolution of the true PSD with the window's frequency response. An illustration of application of the Bartlett procedure is provided at the end of Section 6.3.3.

Cyclostationary signals such as the PCG offer a unique and interesting approach to synchronized averaging of periodograms over a number of cycles, without the trade-off between the reduction of variance and the loss of resolution imposed by segmentation as described above. This is presented as an illustration of application in Section 6.3.5.

6.3.3 The use of windows: Spectral resolution and leakage

The Bartlett procedure may be viewed as an ensemble averaging approach to reduce the variance (which may be interpreted as noise) of the periodogram. Another approach to obtain a smooth spectrum is to convolve the periodogram $S(\omega)$ with a filter or smoothing function $W(\omega)$ in the frequency domain (similar to the use of an MA filter in the time domain). The smoothed estimate $S_s(\omega)$ is given by

$$S_s(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\nu)\, W(\omega - \nu)\, d\nu, \tag{6.17}$$

where ν is a temporary variable for integration.

As the PSD is a nonnegative function, the smoothing function $W(\omega)$ should satisfy $W(\omega) \geq 0$, $-\pi \leq \omega \leq \pi$. The Fourier transform of the Bartlett window $W_B(\omega)$ meets this requirement. Oppenheim and Schafer [174] have shown that the variance of the smoothed periodogram is reduced approximately by the factor

$$\frac{1}{N} \sum_{m=-(M-1)}^{M-1} w^2(m) = \frac{1}{2\pi N} \int_{-\pi}^{\pi} W^2(\omega)\, d\omega, \tag{6.18}$$

with reference to the variance of the original periodogram; here, N is the total number of samples in the signal and $(2M-1)$ is the number of samples in the smoothing window function. A rectangular window offers a variance-reduction factor of approximately $\frac{2M}{N}$, whereas the factor for the Bartlett window is $\frac{2M}{3N}$ [174]. It should be noted that smoothing of the spectrum (reduction of variance) is achieved at the price of loss of frequency resolution.

Since the periodogram is the Fourier transform of the ACF estimate $\phi(m)$, the convolution operation in the frequency domain in Equation 6.17 is equivalent to multiplying $\phi(m)$ with $w(m)$, the inverse Fourier transform of $W(\omega)$. This result suggests that the same PSD estimate as $S_s(\omega)$ may be obtained by applying a window to the ACF estimate and then taking the Fourier transform of the result. As the ACF is an even function, the window should also be even.

Based upon the arguments outlined above, Welch [325] (see also Oppenheim and Schafer [174]) proposed a method to average modified periodograms. In Welch's

procedure, the given signal is segmented as in the Bartlett procedure, but a window is applied directly to the original signal segments before Fourier transformation. The periodograms of the windowed segments are defined as

$$S_{Wi}(\omega) = \frac{1}{ME_w}\left|\sum_{n=0}^{M-1} x_i(n)w(n)\exp(-j\omega n)\right|^2, \quad i = 1, 2, \ldots, K, \tag{6.19}$$

where E_w is the average power of the window given by

$$E_w = \frac{1}{M}\sum_{n=0}^{M-1} w^2(n). \tag{6.20}$$

Note that the duration of the window defined above is M samples. The Welch PSD estimate $S_W(\omega)$ is obtained by averaging the modified periodograms as

$$S_W(\omega) = \frac{1}{K}\sum_{i=1}^{K} S_{Wi}(\omega). \tag{6.21}$$

Welch [325] showed that, if the segments are not overlapping, the variance of the averaged modified periodogram is inversely proportional to K, the number of segments used. Welch also suggested that the segments may be allowed to overlap, in which case the modified periodograms are not mutually independent. The spectral window that is effectively convolved with the PSD in the frequency domain is proportional to the squared magnitude of the Fourier transform of the time-domain data window applied. Therefore, no matter which type of a data window is used, the spectral smoothing function is nonnegative, thereby guaranteeing that the PSD estimate will be nonnegative as well.

Some of the commonly used data windows are defined below [174, 330]; the windows are of length N samples and causal, defined for $0 \le n \le N-1$.

Rectangular:

$$w(n) = 1. \tag{6.22}$$

Bartlett (triangular):

$$w(n) = \begin{cases} \frac{2n}{N-1}, & 0 \le n \le \frac{N-1}{2}, \\ 2 - \frac{2n}{N-1}, & \frac{N-1}{2} \le n \le N-1. \end{cases} \tag{6.23}$$

Hamming:

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right). \tag{6.24}$$

von Hann (also known as Hann or Hanning):

$$w(n) = \frac{1}{2}\left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right]. \tag{6.25}$$

Blackman:

$$w(n) = 0.42 - 0.5\cos\left(\frac{2\pi n}{N-1}\right) + 0.08\cos\left(\frac{4\pi n}{N-1}\right). \tag{6.26}$$

Figure 6.4 illustrates the rectangular, Bartlett, Hann, and Hamming windows with $N = 256$ samples. A Hann window with $N = 128$ samples is also illustrated (centered with reference to the longer-duration windows).

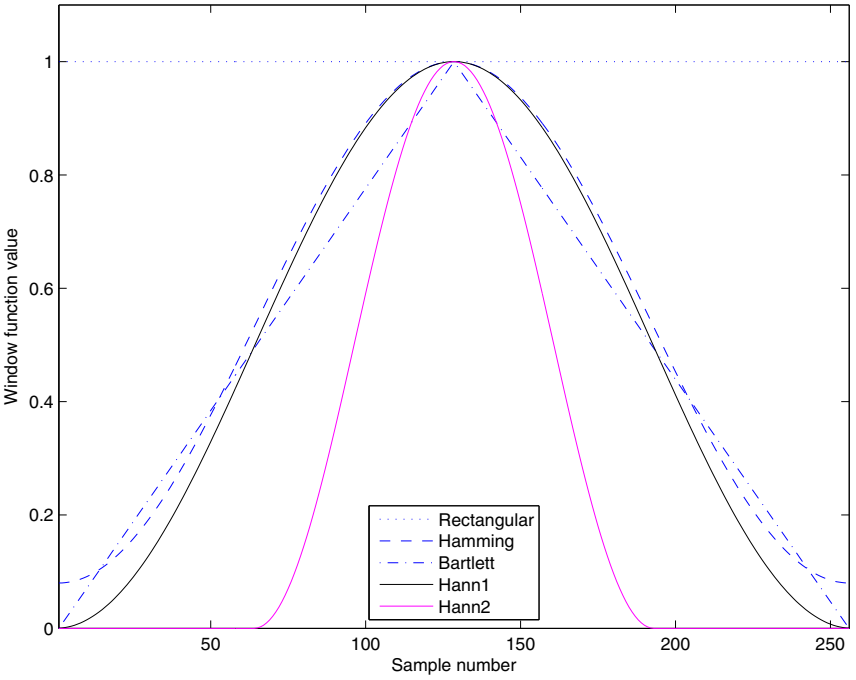

Figure 6.4 Commonly used window functions: rectangular, Bartlett, Hamming, and Hann windows with $N = 256$ (Hann1), and Hann window with $N = 128$ samples (Hann2). All windows are centered at the 128^{th} sample.

Use of the tapered windows (all of the above, except the rectangular window) provides the advantage that the ends of the given signal are reduced to zero (with the further exception of the Hamming window, for which the end-values are not zero but 0.08). This feature means that there are no discontinuities in the periodic version of the signal encountered in DFT-based procedures. All of the window functions listed above are symmetric (even) functions and, therefore, have a linear phase (or a real spectrum with zero phase if the window is centered at the origin).

Figures 6.5 to 6.9 illustrate the log-magnitude frequency responses of the window functions shown in Figure 6.4. The frequency responses were computed after padding the windows to a total duration of $L = 2,048$ samples for FFT computation.

The plots are on an expanded scale over the limited normalized-frequency range of $(0, 0.1)$ in order to illustrate clearly the characteristics of the main lobe and the side lobes. The plot for the frequency response of the rectangular window has been plotted to a different scale as compared to the rest. The discontinuities in the frequency responses of the rectangular and Bartlett windows in Figures 6.5 and 6.6 are due to the log of the zeros of the responses being $-\infty$.

The rectangular window has the narrowest main lobe of width $\frac{4\pi}{N}$; the main lobe is wider at $\frac{8\pi}{N}$ for the Bartlett, Hann, and Hamming windows; it is the widest at $\frac{12\pi}{N}$ for the Blackman window [174]. A reduction in window width will lead to an increase in the main-lobe width, as illustrated by the frequency responses of the two Hann windows in Figures 6.8 and 6.9. Note that the wider the main lobe, the greater the spectral smoothing, and hence the loss of spectral resolution is more severe.

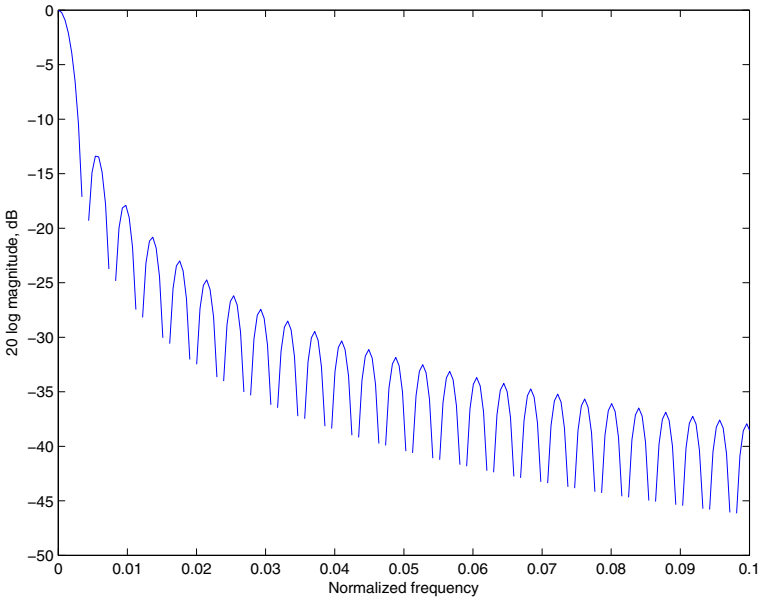

Figure 6.5 Log-magnitude frequency response of the rectangular window illustrated in Figure 6.4. The window width is $N = 256$ samples.

The rectangular window has the highest peak side-lobe levels of all of the windows listed at $-13\ dB$, with the Bartlett, Hann, Hamming, and Blackman windows having their peak side-lobe levels at $-25\ dB$, $-31\ dB$, $-41\ dB$, and $-57\ dB$, respectively [174]. Higher levels of the side lobes will cause increased spectral leakage (weighted summation of spectral components with significant weights over a wide range of frequencies due to convolution in the frequency domain), resulting in a more distorted spectrum. Note that reduction of leakage through the use of the tapered windows comes at the price of increased main-lobe width and, therefore, more severe loss of spectral resolution (more smoothing).

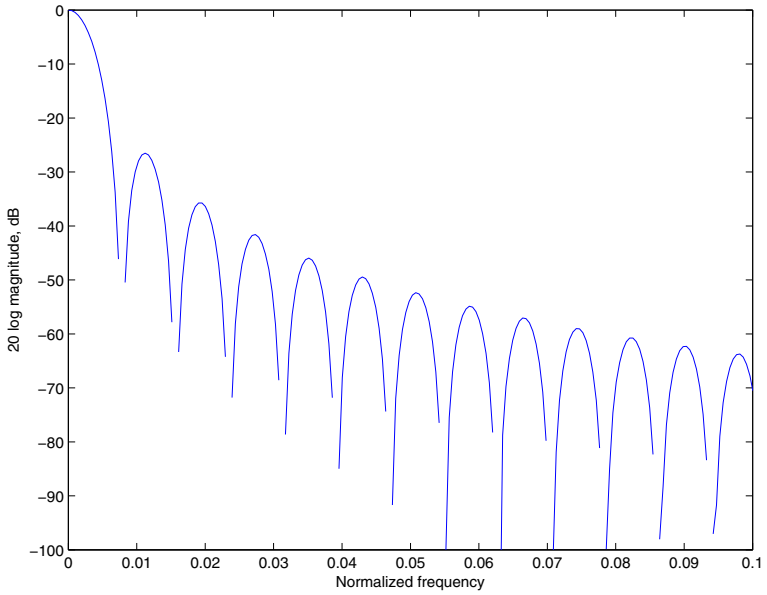

Figure 6.6 Log-magnitude frequency response of the Bartlett window illustrated in Figure 6.4. The window width is $N = 256$ samples.

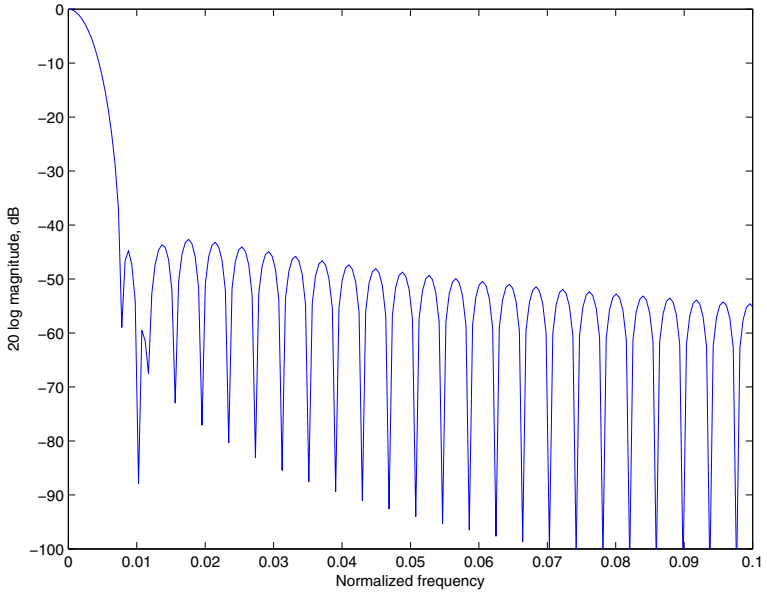

Figure 6.7 Log-magnitude frequency response of the Hamming window illustrated in Figure 6.4. The window width is $N = 256$ samples.

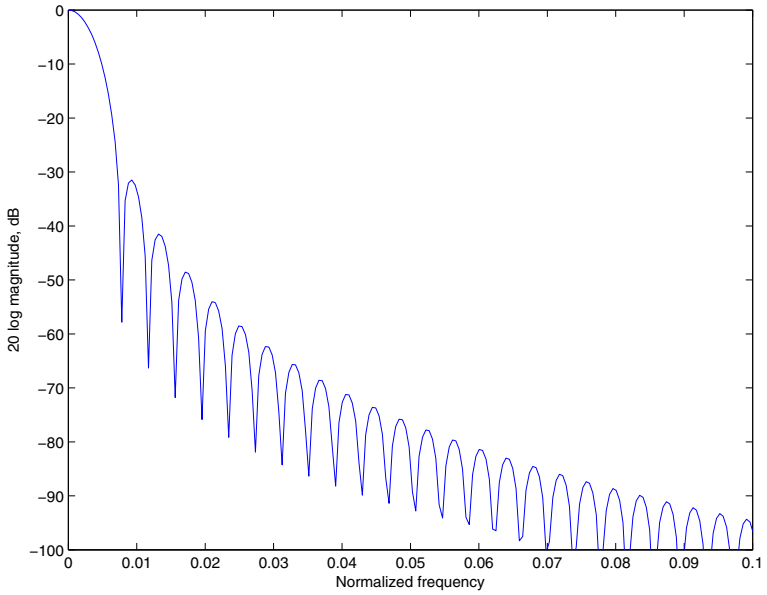

Figure 6.8 Log-magnitude frequency response of the Hann1 window illustrated in Figure 6.4. The window width is $N = 256$ samples.

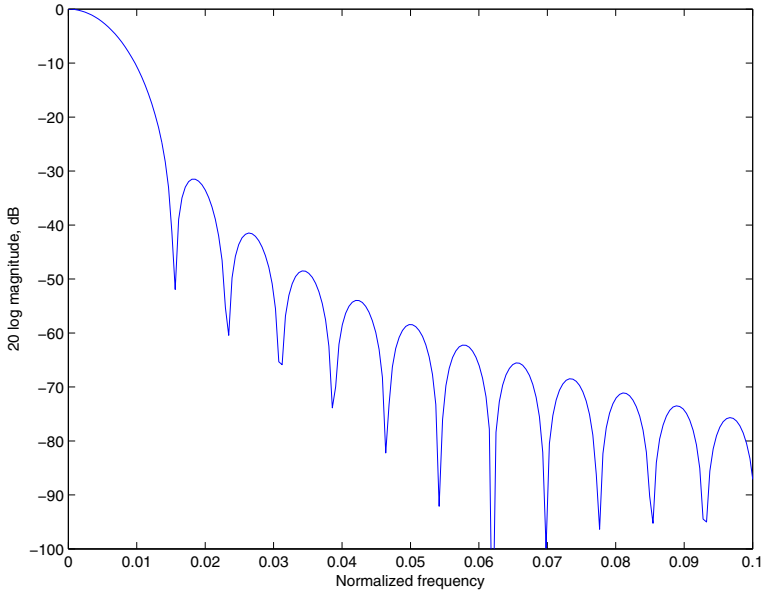

Figure 6.9 Log-magnitude frequency response of the Hann2 window illustrated in Figure 6.4. The window width is $N = 128$ samples.

Illustrations of application: The Welch method of windowing signal segments and averaging their PSDs was applied to the o2 channel of the EEG signal illustrated in Figure 1.39. The number of samples in the signal is $N = 750$, with the sampling frequency being $f_s = 100\ Hz$. Note that the specific EEG signal record may be assumed to be stationary over its relatively short duration of $7.5\ s$. The dominant activity in the signal is the alpha rhythm, which appears throughout the duration of the signal record.

The PSD of the entire signal was first computed using no window (that is, the rectangular window was applied implicitly); the FFT array was computed with $L = 1,024$ samples. The top trace in Figure 6.10 illustrates the PSD of the signal.

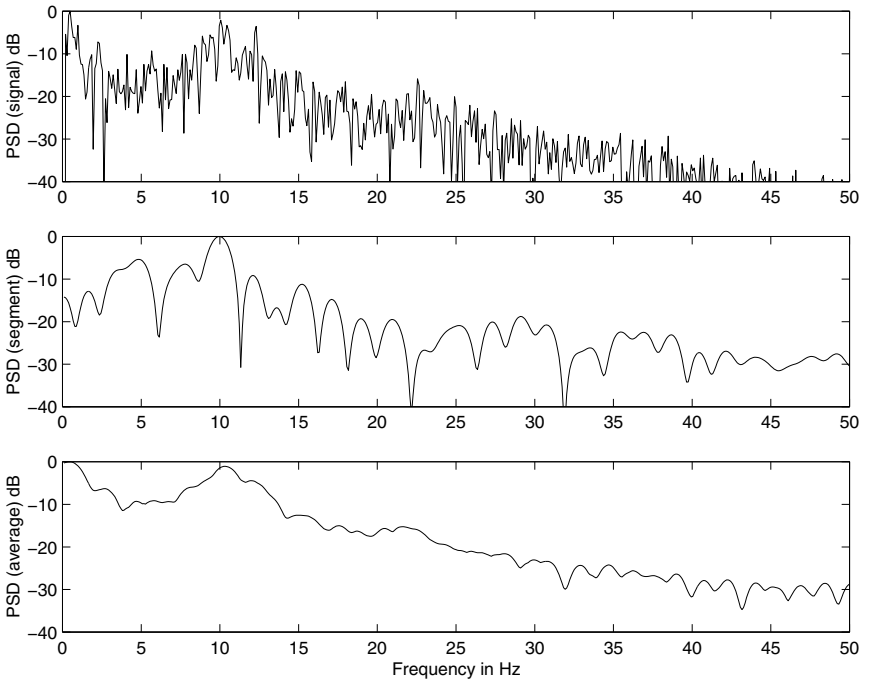

Figure 6.10 Bartlett PSD estimate of the o2 channel of the EEG signal in Figure 1.39. Top trace: PSD of the entire signal. Middle trace: PSD of the 11^{th} segment. Bottom trace: Averaged PSD using $K = 11$ segments of the signal. The rectangular window was (implicitly) used in all cases. Number of samples in the entire signal: $N = 750$. Number of samples in each segment: $M = 64$. All FFT arrays were computed with $L = 1,024$ samples. Sampling frequency $f_s = 100\ Hz$. See also Figure 6.11.

For the first averaged periodogram procedure, the EEG signal was segmented with $M = 64$ samples each, with implicit use of the rectangular window (equivalent to the Bartlett method). A total of $K = 11$ segments were obtained. Each segment was padded with zeros to a length of $L = 1,024$ for the sake of FFT computation. The PSDs of the segments were then averaged, followed by normalization and logarith-

mic transformation. The second and third plots in Figure 6.10 illustrate the PSD of a sample segment (the 11^{th} segment) and the averaged PSD (the Bartlett estimate), respectively. It is seen that the averaged PSD (third trace) provides a smooth spectral estimate with a clearly dominant peak at approximately $10\ Hz$, representing the alpha rhythm present in the signal. The PSD of the individual segment (middle trace) displays many peaks and valleys that are possibly spurious and not significant, and have been suppressed or smoothed by the averaging process, as evident in the third plot in the figure. The single PSD computed from the entire signal (top trace) exhibits numerous variations that may not be relevant and could confound visual or automated analysis. (*Note:* Direct comparison of the PSDs is possible since they have the same number of samples, that is, the same frequency sampling rate.)

Figure 6.11 illustrates a second set of PSDs similar to that in Figure 6.10, but with the use of the Hann window in the Welch procedure. The effect of the Hann window is not significant in the case of the PSD of the entire signal (top trace), as the window length is reasonably large ($N = 750$). However, the Hann window has clearly smoothed the multiple (possibly spurious) peaks and valleys in the PSD of the segment illustrated in the middle trace. The wider main lobe of the Hann window's frequency response has caused a more severe loss of frequency resolution (smoothing) than the rectangular window in the case of the corresponding PSD in Figure 6.10. Finally, the averaged PSD in the lowest trace of Figure 6.11 clearly illustrates the benefit of the Hann window in the significantly reduced power levels beyond $30\ Hz$. The lower side-lobe levels of the Hann window have resulted in less spectral leakage than that in the case of the rectangular window as illustrated by the corresponding PSD in Figure 6.10. The price paid, however, is evidenced by the wider peak in the averaged PSD with the Hann window, which spans the range $5 - 15\ Hz$ at the $-10\ dB$ level. The two distinct peaks at about $10\ Hz$ and $12\ Hz$ that are evident in the top traces of Figures 6.10 and 6.11 as well as in the smoothed PSD in the bottom trace of Figure 6.10 are no longer seen separately in the bottom trace of Figure 6.11. Regardless, the averaged PSD with the Hann window appears to be smoother and more amenable to analysis than the corresponding result with the rectangular window.

It should be noted that the segments of EEG signals used in the illustrations in the present section are of short duration. See Section 6.7 for an example of analysis of longer EEG records.

6.3.4 Estimation of the ACF

Good estimates of the ACF are required in applications such as the design of the optimal Wiener filter and estimation of the statistics of stochastic processes. Once a PSD estimate has been obtained by a method such as the Bartlett or Welch procedures, we may take the inverse Fourier transform of the result and use the result as an estimate of the ACF. We may also fit a smooth curve or a parametric model (such as an exponential or a Laplacian function) to the PSD or to the equivalent ACF model.

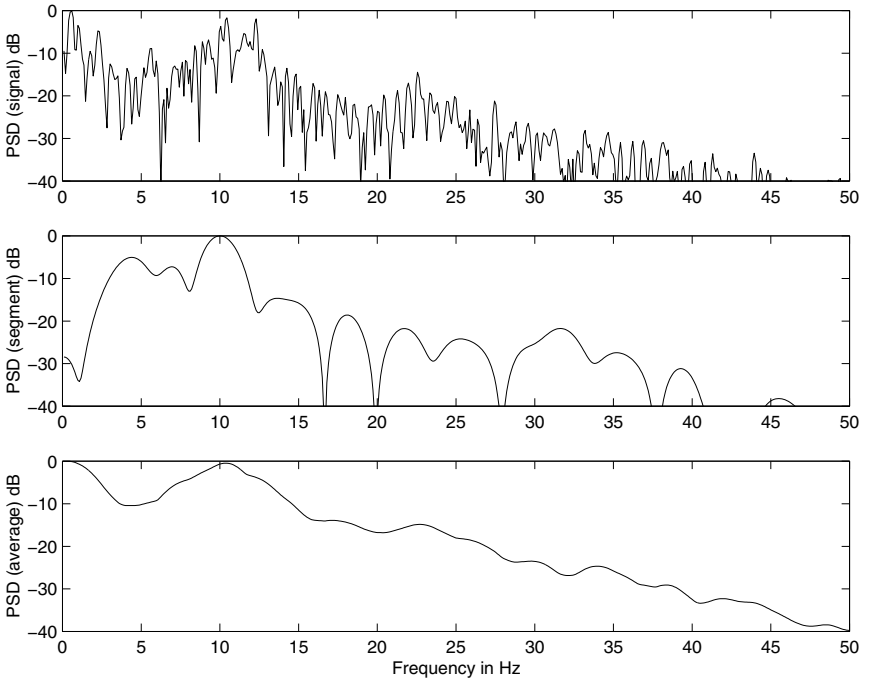

Figure 6.11 Welch PSD estimate of the o2 channel of the EEG signal in Figure 1.39. Top trace: PSD of the entire signal. Middle trace: PSD of the 11^{th} segment. Bottom trace: Averaged PSD using $K = 11$ segments of the signal. The Hann window was used in all cases. Number of samples in the entire signal and the size of the Hann window used in computing the PSD of the entire signal: $N = 750$. Number of samples in each segment and the size of the Hann window used in the averaged periodogram method: $M = 64$. All FFT arrays were computed with $L = 1,024$ samples. Sampling frequency $f_s = 100\ Hz$. See also Figure 6.10.

Let us consider again the expression

$$\phi_2(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x(n)x(n+m). \tag{6.27}$$

As the ACF is an even function, we need to compute it only for positive m. It is evident that the ACF estimate is simply the result of linear convolution of $x(n)$ with $x(-n)$ (with the scale factor $\frac{1}{N}$). If the DFT of $x(n)$ is $X(k)$, the DFT of $x(-n)$ is $X^*(k)$. Since convolution in the time domain is multiplication in the frequency domain, we could compute the DFT $X(k)$ of $x(n)$, obtain $X(k)X^*(k) = |X(k)|^2$, and take its inverse DFT. However, the DFT procedure provides circular convolution and not linear convolution. Therefore, we need to pad $x(n)$ with at least $M - 1$ zeros, where M is the largest lag for which the ACF is desired. The DFT must then be computed with at least $L = N + M - 1$ samples, where N is the number of

samples in the original signal. If this requirement is built into the periodogram or averaged periodogram procedure, the inverse DFT of the final PSD estimate may be used as an estimate of the ACF [with the scale factor $\frac{1}{N}$, or division by $\phi(0)$ to get the normalized ACF].

6.3.5 Synchronized averaging of PCG spectra

Every individual is familiar with the comforting *lub–dub* sounds of his or her heart beat; every prospective parent would have taken pleasure in listening to the throbbing heart of the yet-to-be-born baby. Use of the heart sounds is extremely common in clinical practice: the stethoscope is the most common sign and tool of a physician. Yet, behind this common signal lie many sophisticated and potentially complicating characteristics.

The PCG is a nonstationary signal due to the fact that the amount of blood in each cardiac chamber and the state of contraction of the muscles change continuously during each cardiac cycle. S2 usually has more high-frequency content than S1: The PSD of a normal PCG signal changes within about $300\ ms$. Valve opening or closing sounds, being of short duration of the order of $10\ ms$, are of a transient and high-frequency character. The presence of murmurs adds another dimension of nonstationarity, with frequency content well beyond that of the normal heart sounds: the PSD of an abnormal PCG could change every $100\ ms$ or less. Individual epochs of S1, S2, valve snaps, and murmurs are of limited durations of the order of $10 - 300\ ms$. These aspects of the PCG preclude segmented averaging as recommended by the Bartlett or Welch procedures.

Over and above all of the factors mentioned in the preceding paragraph, the transmission characteristics of the chest wall change during breathing. (Living systems are dynamic!) The PCG signals recorded at various locations on the chest are also subject to different transmission-path effects. While adult subjects may cooperate in PCG signal acquisition by holding their breath or performing other maneuvers, these possibilities cannot be considered in the case of infants and young children in poor states of health. The PCG signal presents more challenges in acquisition and analysis than most of the other biomedical signals we have encountered [71].

Problem: *Propose a method to obtain averaged PSD estimates of the systolic and diastolic heart sounds.*

Solution: The cyclostationarity of the PCG lends itself to a unique approach to averaging PCG segments corresponding to the same phase of the cardiac cycle extracted from multiple beats. If the subject were to hold his/her breath during the period of acquisition of the PCG record, the chest-wall transmission characteristics will be stationary over the multiple cardiac cycles in the record. Therefore, we may segment S1, S2, or any portion of the cardiac cycle of interest from as many beats as are available, and then average their PSD estimates in a procedure similar to the Bartlett or Welch procedures. (*Note:* Direct averaging of the PCG signals themselves or of their complex Fourier transforms could lead to undesired cancellation of noise-like murmurs or asynchronous frequency components and their disappearance from

the result! Refer to Sections 4.11 and 6.5 for discussions on intentional cancellation of asynchronous components in the PCG via synchronized averaging.)

We saw in Sections 5.5.2 and 5.5.3 how the envelope or the envelogram of the PCG may be averaged over several cardiac cycles. However, there was no need to segment parts of a cardiac cycle in envelope analysis: Nonstationarity of the signal within a cardiac cycle was not of concern. In the present application of PSD analysis, there is a need to segment the PCG further.

A procedure was described in Section 4.10 for segmentation of the systolic and diastolic parts of PCG signals based upon the detection of the QRS complex in the ECG and the detection of the dicrotic notch in the carotid pulse signal. Further segmentation of the systolic or diastolic parts into S1 and systolic murmur or S2 and diastolic murmur, respectively, would require more sophisticated methods, which are the topics of Chapter 8. For now, let us consider the task of obtaining averaged PSDs of the systolic and diastolic parts of a PCG signal.

Figure 6.12 shows the PCG signal over one cardiac cycle of a normal subject segmented using the procedure described in Section 4.10 and illustrated in Figure 4.30 (see also Figure 5.7). The periodograms of the systolic and diastolic parts of the PCG cycle illustrated are also shown in the figure. In order to obtain better PSD estimates, the periodogram of each systolic or diastolic segment was computed separately and averaged over 16 cardiac cycles. No data window was applied (the rectangular window was used, in effect); therefore, the procedure used is similar to the Bartlett procedure. Individual systolic or diastolic segments could be of different durations; for the present illustration, all periodograms were computed with the same number of samples, which was taken to be the maximum RR interval in the ECG record of the subject. The averaged systolic and diastolic PSD estimates are shown in Figure 6.12. (Averaging was performed using the nonnegative PSDs before conversion to dB.) The averaging procedure provides smoother estimates of the PSDs by removing beat-to-beat variations that are neither significant nor of interest. Spectral peaks may be clearly observed in the averaged periodograms and may be considered to be more reliable estimates of resonance than the peaks found in individual periodograms.

Figure 6.13 illustrates a PCG signal cycle as well as the individual and averaged systolic and diastolic PSD estimates for a patient with systolic murmur, split S2, and opening snap of the mitral valve (see also Figures 4.31 and 5.8). It is unlikely that the patient held her breath during data acquisition. The presence of increased high-frequency power in the range $120 - 250\ Hz$ due to the systolic murmur is evident in the averaged systolic PSD. The diastolic PSDs are comparable to the corresponding normal diastolic PSDs in Figure 6.12.

6.4 Measures Derived from Power Spectral Density Functions

The Fourier spectrum or PSD provides us with a density function of signal amplitude, power, or energy versus frequency. We would typically have a large number of samples of the PSD over a wide frequency range, which may not lend itself to easy

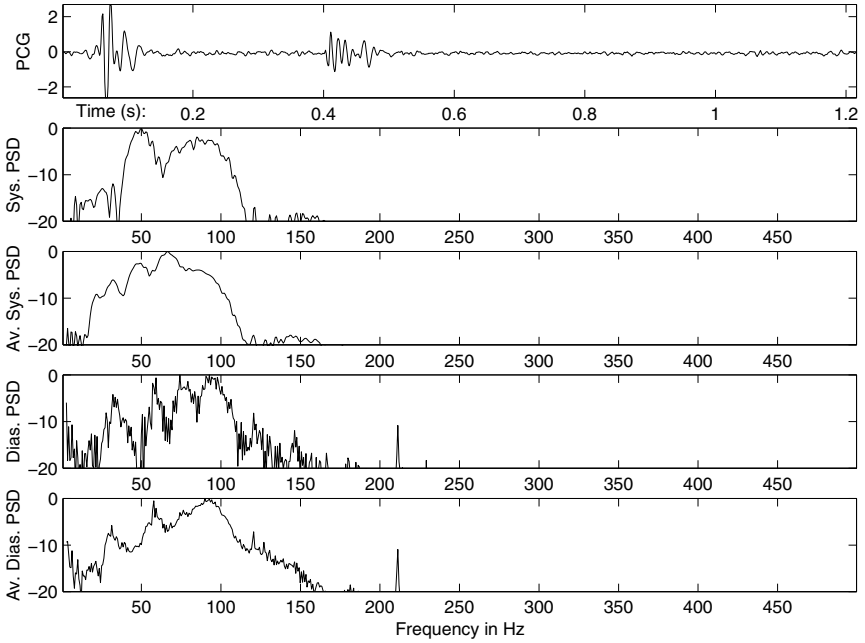

Figure 6.12 Top to bottom: A sample PCG signal over one cardiac cycle of a normal subject (male, 23 years; see also Figures 4.30 and 5.7); periodogram of the systolic portion of the signal (approximately $0 - 0.4$ s); averaged periodogram of the systolic parts of 16 cardiac cycles segmented as illustrated in Figure 4.30; periodogram of the diastolic portion of the signal shown in the first plot (approximately $0.4 - 1.2$ s); averaged periodogram of the diastolic parts of 16 cardiac cycles. The periodograms are on a log scale (dB). Av. = average. Sys. = systolic. Dias. = diastolic.

analysis. We may, of course, study the shape of the spectrum graphically and observe its general characteristics. Such an approach is often referred to as *nonparametric* spectral analysis. The spectral models described in Section 7.4 are characterized by a small number of parameters; hence, the procedure is known as *parametric* spectral analysis (or modeling).

Problem: *Derive parameters or measures from a Fourier spectrum or PSD that can help in the characterization of the spectral variations or features contained therein.*

Solution: Since the PSD is a nonnegative function as well as a density function, we may readily treat it as a PDF, and compute statistics using moments. We may also detect peaks corresponding to resonance, measure their bandwidth or quality factor, and derive measures of concentration of power in specific frequency bands of interest or concern. Although the PSD itself is nonparametric, we may derive several parameters that, while not completely representing the entire PSD, may facilitate the

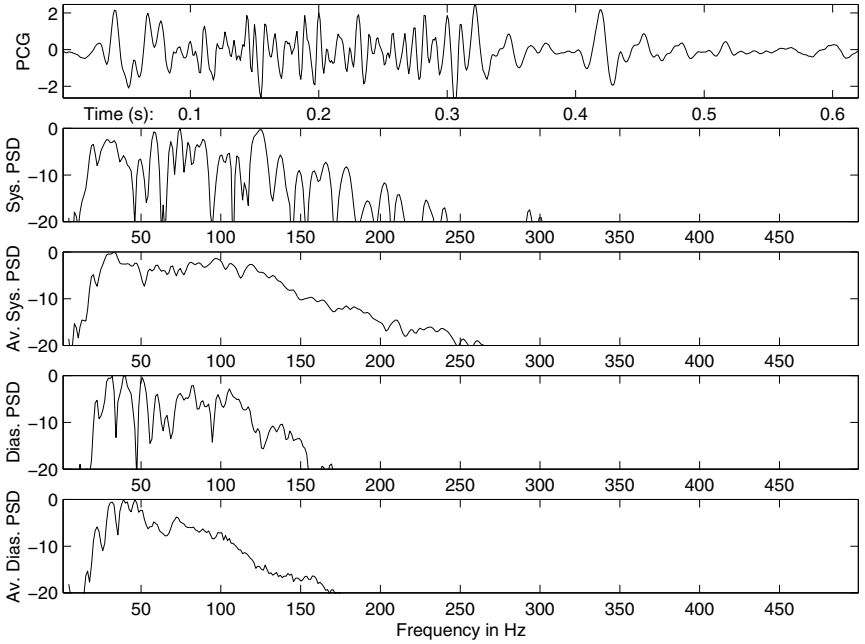

Figure 6.13 Top to bottom: A sample PCG signal over one cardiac cycle of a patient (female, 14 months; see also Figures 4.31 and 5.8) with systolic murmur, split S2, and opening snap of the mitral valve; periodogram of the systolic portion of the signal (approximately $0 - 0.28\ s$); averaged periodogram of the systolic parts of 26 cardiac cycles segmented as illustrated in Figure 4.31; periodogram of the diastolic portion of the signal shown in the first plot (approximately $0.28 - 0.62\ s$); averaged periodogram of the diastolic parts of 26 cardiac cycles. The periodograms are on a log scale (dB). Av. = average. Sys. = systolic. Dias. = diastolic.

identification of physiological and/or pathological phenomena. We shall investigate a few different approaches toward this end in the following sections.

6.4.1 Moments of PSD functions

As the area under the PSD curve represents the total signal power or energy which need not be unity, we have to normalize all moments by the total power or energy of the signal E_x, given by

$$\begin{aligned} E_x &= \sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 \\ &= \frac{1}{2\pi} \int_0^{2\pi} |X(\omega)|^2 \, d\omega = \int_{f_n=0}^{1} |X(f_n)|^2 \, df_n. \end{aligned} \tag{6.28}$$

The frequency variable f_n above is normalized as $f_n = f/f_s$, such that $0 \leq f_n \leq 1$. Assuming that the PSD has been obtained using one of the methods described in the preceding sections, we may replace $|X(.)|^2$ in the above expressions by $S_{xx}(.)$.

In all definitions of the moments of PSDs in the present section,

$$S_{xx}(k) = \frac{1}{N} |X(k)|^2,$$

where $X(k)$ is the DFT of $x(n)$. $S_{xx}(k)$ represents the PSD of $x(n)$. Note the presence of the factor $\frac{1}{N}$ in the definition of the PSD. An additional factor of $\frac{1}{N}$ is required in the definitions of the moments of PSDs in this section, in order to scale the frequency index k to the appropriate frequency in Hz in relation to the sampling frequency f_s. Two versions of the moments are provided: one using the normalized frequency variable $0 \leq f_n \leq 1$ and the other using the DFT index $0 \leq k \leq N-1$.

As a measure of the concentration or spread of the signal power over its frequency range, we may compute the mean frequency $\overline{f}$ as the first-order moment

$$\overline{f_n} = \frac{2}{E_x} \int_{f_n=0}^{0.5} f_n \, S_{xx}(f_n) \, df_n \tag{6.29}$$

or as

$$\overline{f_{Hz}} = \frac{f_s}{N} \frac{2}{E_x} \sum_{k=0}^{N/2} k \, S_{xx}(k), \tag{6.30}$$

where $\overline{f_n}$ and $\overline{f_{Hz}}$ indicate the mean frequency in its normalized form and in Hz, respectively, and N is the number of samples in the DFT-based representation of the PSD. The upper limit of integration of 0.5 represents integration from DC to the maximum frequency present in the signal, which is one-half of the sampling frequency, the frequency variable having been normalized to the range $0 \leq f_n \leq 1$. Note that the integration or summation is performed over one-half period of the periodic function $S_{xx}(f_n)$ or $S_{xx}(k)$, which possesses even symmetry about one-half of the sampling frequency for real signals.

The median frequency f_{med} is defined as that frequency which splits the PSD in half:

$$f_{\text{med}} = \frac{m}{N} f_s \text{ with the largest } m \text{ such that } \frac{2}{E_x} \sum_{k=0}^{m} S_{xx}(k) < \frac{1}{2}; \; 0 \leq m \leq \frac{N}{2}. \tag{6.31}$$

We may also compute higher-order statistics as follows:

- Variance f_{m2} as the second-order moment by using $(f_n - \overline{f_n})^2$ in place of f_n [the function of frequency that is multiplied with $S_{xx}(f_n)$] in Equation 6.29 or the equivalent expression in k in Equation 6.30; that is,

$$f_{m2} = \frac{2}{E_x} \int_{f_n=0}^{0.5} (f_n - \overline{f_n})^2 \, S_{xx}(f_n) \, df_n \tag{6.32}$$

or

$$f_{m2} = \left(\frac{f_s}{N}\right)^2 \frac{2}{E_x} \sum_{k=0}^{N/2} (k - \overline{k})^2 \, S_{xx}(k), \tag{6.33}$$

where $\overline{k} = N\,\overline{f_{Hz}}/f_s$ is the frequency sample index corresponding to $\overline{f_{Hz}}$.

- Skewness as

$$\text{skewness} = \frac{f_{m3}}{(f_{m2})^{3/2}}, \tag{6.34}$$

where the third-order moment f_{m3} is computed with $(f_n - \overline{f_n})^3$ in place of f_n in Equation 6.29, that is,

$$f_{m3} = \frac{2}{E_x} \int_{f_n=0}^{0.5} (f_n - \overline{f_n})^3 \, S_{xx}(f_n) \, df_n \tag{6.35}$$

or

$$f_{m3} = \left(\frac{f_s}{N}\right)^3 \frac{2}{E_x} \sum_{k=0}^{N/2} (k - \overline{k})^3 \, S_{xx}(k). \tag{6.36}$$

- Kurtosis as

$$\text{kurtosis} = \frac{f_{m4}}{(f_{m2})^2}, \tag{6.37}$$

where the fourth-order moment f_{m4} is computed with $(f_n - \overline{f_n})^4$ in place of f_n in Equation 6.29, that is,

$$f_{m4} = \frac{2}{E_x} \int_{f_n=0}^{0.5} (f_n - \overline{f_n})^4 \, S_{xx}(f_n) \, df_n \tag{6.38}$$

or

$$f_{m4} = \left(\frac{f_s}{N}\right)^4 \frac{2}{E_x} \sum_{k=0}^{N/2} (k - \overline{k})^4 \, S_{xx}(k). \tag{6.39}$$

In order to avoid notational clutter, the same symbols have been used for the two versions each of f_{m2}, f_{m3}, and f_{m4}; it is understood that similar versions of the moments are used to compute the ratios for skewness and kurtosis.

The mean frequency is a useful measure of the concentration of signal power, and could indicate the resonance frequency in the case of unimodal distributions. However, a nearly uniform PSD could lead to one-half of the maximum frequency as the mean frequency, which by itself may not be a useful representation of the PSD. The presence of multiple resonance frequencies could also lead to a mean frequency that may not be a useful measure. Multimodal PSDs may be characterized better by a series of peak frequencies, along with measures of their relative levels and bandwidths or quality factors (described in the next section).

The square root of f_{m2} provides a measure of spectral spread (SD about the mean) and an indication of the bandwidth (but not at -3 dB) about the mean frequency.

The skewness is zero if the density function is symmetric about the mean frequency; otherwise, it indicates the extent of asymmetry of the distribution. Kurtosis indicates if the PSD is a long-tailed function.

Moments of PSDs may be useful in characterizing the general trends in the distribution of the power of a signal over its bandwidth. The higher-order moments are sensitive to noise or spurious variations in the PSD estimate, and they may not yield reliable measures if the PSD pattern is not simple or if the PSD estimate is poor (has a high variance). The reliability of moments may be improved by smoothing the PSD estimate, or by fitting a smooth parametric curve (such as a Gaussian, a Laplacian, or a spline) as a model of the PSD estimate and computing the moments of the model. Saltzberg and Burch [249] discuss the relationship between moments of PSDs and ZCR, along with their application to EEG analysis.

6.4.2 Spectral power ratios

The moments described in the preceding section provide general statistical characterization of the PSD treated as a PDF. In the case of analysis of biomedical signals, it may be more advantageous to define specific measures based upon *a priori* information or empirical knowledge about the signals, the systems, and the physiological or pathological processes of concern. For example, in the case of PCG analysis for the detection of murmurs, we could specifically investigate the presence of signal power in the frequency range beyond that of S1 and/or S2. If a specific type of pathology of interest is known to cause a shift in the frequency content within a certain band of frequencies, we may measure spectral power ratios over partitions of the band of interest. We have seen in Sections 6.2.1 and 6.2.2 how such measures have been used for the analysis of ventricular elasticity, diagnosis of myocardial infarction, and detection of murmurs.

The fraction of signal power in a frequency band of interest $(f_1 : f_2)$ may be computed as

$$E_{(f_1:f_2)} = \frac{2}{E_x} \int_{f=f_1}^{f_2} |X(f)|^2 \, df = \frac{2}{NE_x} \sum_{k=k_1}^{k_2} |X(k)|^2, \qquad (6.40)$$

where k_1 and k_2 are the DFT indices corresponding to f_1 and f_2, respectively. Fractions of power as above may be computed for several bands of interest that may or may not span the entire signal bandwidth.

In a variation of the fractional-power measure defined above, Johnson et al. [322] compared the integral of the magnitude spectrum of the systolic murmurs due to aortic stenosis over the band $75 : 150 \ Hz$ to that over the band $25 : 75 \ Hz$. They considered the higher-frequency band to represent the *predictive area* (PA) of the spectrum related to aortic stenosis and considered the lower-frequency band to represent a *constant area* (CA) that would be common to all systolic PCG signal

segments. The ratio of PA to CA was defined as

$$\frac{PA}{CA} = \frac{\int_{f=f_2}^{f_3} |X(f)|\, df}{\int_{f=f_1}^{f_2} |X(f)|\, df}, \tag{6.41}$$

with $f_1 = 25\ Hz$, $f_2 = 75\ Hz$, and $f_3 = 150\ Hz$. The $\frac{PA}{CA}$ ratio is provided for the PSDs of systolic murmurs of four patients with aortic stenosis in Figure 6.2; Johnson et al. showed that the ratio correlates well with the severity of aortic stenosis.

Binnie et al. [258, 259] described the application of spectral analysis to EEG for the detection of epilepsy. Their method was based upon partitioning or banding of the EEG spectrum into not only the traditional δ, θ, α, and β bands, but also into seven other nonuniform bands specified as $1-2$, $2-4$, $4-6$, $6-8$, $8-11$, $11-14$, and $> 14\ Hz$. Additional features related to FF (see Section 5.6.4) were also used. In a study with 275 patients with suspected epilepsy, 90% of the signals of the patients with pathology were classified as abnormal by their methods; conversely, 86% of the patients whose EEGs were classified as abnormal had confirmed pathology.

When analyzing a spectral peak, we may also compute the $-3\ dB$ bandwidth of the peak and, furthermore, compute its quality factor as the ratio of the peak frequency to the bandwidth. Such measures may be computed for not only the dominant peak, but several peaks at progressively lower levels of signal power. Essentially, each potential resonance peak is treated and characterized as a bandpass filter. Durand et al. [331] used such measures to characterize the PSDs of sounds produced by prosthetic heart valves (discussed in Section 6.5).

6.5 Application: Evaluation of Prosthetic Heart Valves

Efficient opening and closing actions of cardiac valves are of paramount importance for proper pumping of blood by the heart. When native valves fail, they may be replaced by mechanical prosthetic valves or by bioprosthetic valves extracted from pigs. Mechanical prosthetic valves are prone to sudden failure due to fracture of their components. Bioprosthetic valves fail gradually due to tissue degeneration and calcification, and have been observed to last $7-12$ years [331]. Follow-up of the health of patients with prosthetic valves requires periodic, noninvasive assessment of the functional integrity of the valves.

Problem: *Deposition of calcium causes the normally pliant and elastic bioprosthetic valve leaflets to become stiff. Propose a method to assess the functional integrity of bioprosthetic valves.*

Solution: Based on the theory that valve opening and closure contribute directly to heart sounds, analysis of PCG components offers a noninvasive and passive approach to evaluation of prosthetic valves. The increased stiffness is expected to lead to higher-frequency components in the opening or closing sounds of the valve. Durand et al. [331] studied the spectra of the entire S1 signal segment to evaluate the sounds contributed by the closure of porcine (pig) bioprosthetic valves implanted in the mitral position in humans. They demonstrated that, whereas normal S1 spectra

were limited in bandwidth to about $100\ Hz$, degenerated bioprosthetic valves created significant spectral energy in the range $100 - 250\ Hz$. Figure 6.14 shows the power spectra of S1 in the case of a normal bioprosthetic valve and a degenerated bioprosthetic valve; increased high-frequency power is evident in the latter.

Durand et al. derived several parameters from S1 spectra and used them to discriminate normal from degenerated bioprosthetic valves. Some of the parameters used by them are the first and second dominant peak frequencies; the bandwidth and quality factor of the dominant peak; integrated mean area above $-20\ dB$; the highest frequency found at $-3\ dB$; total area and RMS value of the spectrum; area and RMS value in the $20 - 100\ Hz$, $100 - 200\ Hz$, and $200 - 300\ Hz$ bands; and the median frequency. Normal-versus-degenerated valve classification accuracies as high as 98% were achieved.

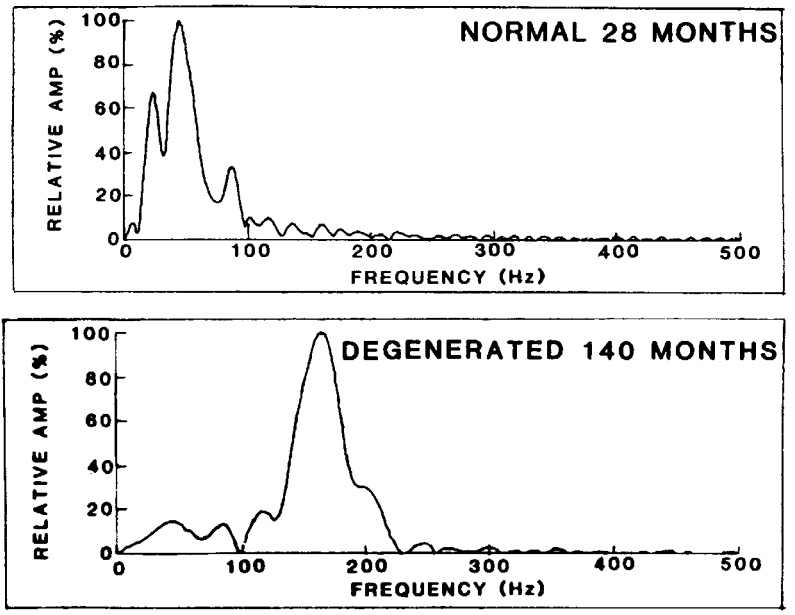

Figure 6.14 First heart sound spectra in the case of normal and degenerated porcine bioprosthetic valves implanted in the mitral position. Reproduced with permission from L.G. Durand, M. Blanchard, G. Cloutier, H.N. Sabbah, and P.D. Stein, Comparison of pattern recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position, *IEEE Transactions on Biomedical Engineering,* 37(12):1121–1129, 1990 ©IEEE.

Durand et al. [229] also studied the sounds of bioprosthetic valves in the aortic position. They argued that the aortic and pulmonary components (A2 and P2, respectively) of S2, each lasting about $50\ ms$, are not temporally correlated during normal breathing. The two components of S2 are separated by $30 - 60\ ms$ during inspiration, but get closer and could overlap during expiration. Furthermore, P2 is weaker

than A2 if the PCG is recorded in the aortic area. Thus, P2 may be suppressed and A2 strengthened by coherent detection and averaging of S2 over several cardiac and breath cycles; see Section 4.11. Durand et al. performed spectral analysis of A2 extracted as above for the purpose of evaluation of bioprosthetic valves in the aortic position. Among a selection of spectral analysis methods including the basic periodogram, Welch's averaged periodogram, all-pole modeling (see Section 7.5), and pole–zero modeling (see Section 7.6), they found the basic periodogram to provide the best compromise for estimating both the spectral distribution and the dominant frequency peaks of bioprosthetic valve sounds.

Cloutier et al. [332] studied the bias and variability of several diagnostic spectral parameters computed from simulated closing sounds of bioprosthetic valves in the mitral position. They found that the most-dominant spectral peak frequency and its quality factor were best estimated using an FFT-based PSD estimate with a rectangular window. However, the $-3\ dB$ bandwidth of the most-dominant spectral peak, the frequency of the second-dominant peak, and a few other parameters were best estimated by the Steiglitz–McBride method of pole–zero modeling (see Section 7.6.2). Some other parameters were best estimated by all-pole modeling using the covariance method (see Section 7.5). It was concluded that a single method would not provide the best estimates of all possible spectral parameters of interest.

6.6 Application: Fractal Analysis of VAG Signals

Problem: *Knee-joint VAG signals exhibit varying levels of waveform complexity depending upon the state of the cartilage covering the joint surfaces. Given that fractals demonstrate complex patterns depending upon certain parameters, can fractal analysis assist in parametric representation and classification of VAG signals?*

Solution: As explained in Section 1.2.14, normal knee joints with smooth cartilage surfaces produce little sound or vibration. However, when the cartilage surfaces are eroded or damaged due to pathological conditions, additional sounds could be generated. Rangayyan and Oloumi [333] and Rangayyan et al. [334] explored the use of FD as a parameter for screening of VAG signals. Given the nonstationary nature of VAG signals, multiple FD values were derived from the PSDs of segments of VAG signals. The related background and methods are described in the following paragraphs.

6.6.1 Fractals and the $1/f$ model

One of the several definitions of fractals is based on geometric structures that have self-similarity at different scales of length or size [275, 283, 335]; see Section 5.13. Fractals lack any single scale of length or size, and have a noninteger (fractional) dimension, referred to as FD. Fractal geometry has been used to represent and synthesize several natural forms such as leaves (especially ferns), trees, mountains, and clouds. Fractal analysis forms a part of advanced techniques for nonlinear analysis of biomedical signals [284–286].

Wiener's geometric model of physical Brownian motion has been used as the basis for algorithms to generate fractal signals or images. In this model, the unpredictable variation of a quantity, V, over time, t, is viewed as a noise process, $V(t)$. The PSD, $P_V(f)$, is used to estimate the power of fluctuations at a given frequency, f, and also of the variations over a time scale of the order of $1/f$.

A time-varying quantity, $V(t)$, with the best fitting line to its PSD, $P_V(f)$, varying as $1/f^{\beta}$ on a log–log scale, is referred to as $1/f$ noise; this is known as the inverse power law or the $1/f$ model. According to Voss [283], most natural phenomena under this model have β in the range of $[0.5, 1.5]$. The PSD of a noise process represented by Brownian motion or random walk varies as $1/f^2$; the power decreases quadratically with frequency. The trace of such a signal is a fractal curve. A direct relationship exists between the FD of the signal and the slope, β, of the best fitting line to its PSD on a log–log scale [283].

The fractional Brownian motion (fBm) model of Mandelbrot [275] has formed the basis of several mathematical models for computer generation of natural fractal scenery. Models based on fBm have been extended to 2D for the synthesis of Brownian surfaces and 3D to generate Brownian clouds [335].

Figure 6.15 shows examples of signals, $V_H(t)$, generated as functions of an arbitrary time variable, t, based on the fBm model. The scaling of the traces is characterized by the scaling parameter H, known as the Hurst coefficient, in the range $0 \leq H \leq 1$. A high value of H close to 1 results in a relatively smooth signal. A lower value of H produces a rougher trace. The variable H relates the changes in V, $\Delta V = V(t_2) - V(t_1)$, to differences in the time variable, $\Delta t = t_2 - t_1$, by the scaling law [283]

$$\Delta V \propto (\Delta t)^H. \tag{6.42}$$

Self-similar patterns repeat themselves at various levels of magnification. However, fBm traces repeat statistically only when t and V are magnified by different amounts. If t is magnified by a factor r as rt, the value of V is magnified by the factor r^H and becomes $r^H V$. Nonuniform scaling of this nature is known as self-affinity [275]. The zero-set of an fBm signal is the intersection of the signal with the horizontal axis. The zero-set is a set of disconnected points with a topological dimension of zero and fractal dimension, D_0, given by [283]

$$D_0 = 1 - H. \tag{6.43}$$

The zero-set of a self-affine signal is self-similar; different estimates of D_0 will yield the same result. The FD of the signal is related to D_0 as

$$FD = D_0 + 1 \tag{6.44}$$

and to the scaling parameter, H, as

$$FD = 2 - H. \tag{6.45}$$

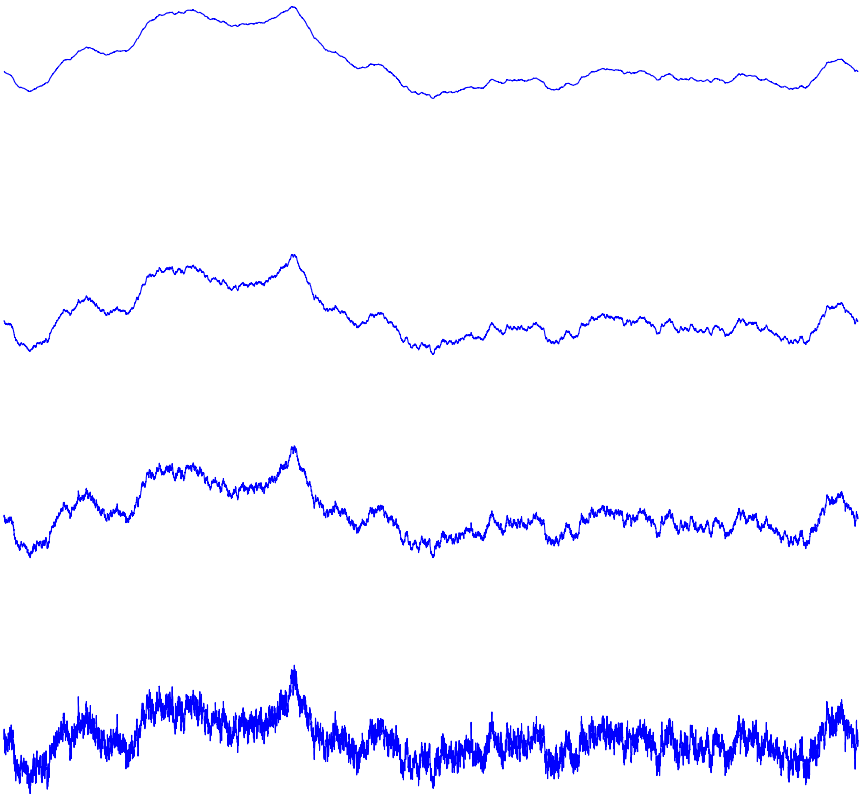

Figure 6.15 Examples of signals generated based on the fBm model for different values of H and FD. Top to bottom: $H = 0.9, 0.6, 0.4, 0.1$; model $FD = 1.1, 1.4, 1.6, 1.9$; estimated $FD = 1.103, 1.401, 1.600, 1.898$. Reproduced with permission from R.M. Rangayyan and F. Oloumi, Fractal analysis of knee-joint vibroarthrographic signals, Proceedings of the 10^{th} IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010), Corfu, Greece, November 2010, pp 1–4. ©IEEE.

6.6.2 FD via power spectral analysis

The best method available to estimate the FD of a self-affine signal is power spectral analysis (PSA). As explained in Section 6.6.1, an fBm signal has a PSD that follows the $1/f^{\beta}$ model. A high value of β indicates a rapid decrease in the high-frequency content of the signal. A self-affine fBm function in an E-dimensional Euclidean space has its PSD $P_V(f) \propto 1/f^{\beta}$, with [283]

$$FD = E + 1 - H, \tag{6.46}$$

where

$$H = \frac{\beta - 1}{2}. \tag{6.47}$$

The FD, in terms of the spectral component, β, for a 1D signal with $E = 1$, is

$$FD = \frac{5 - \beta}{2}. \tag{6.48}$$

In the studies of Rangayyan and Oloumi [333] and Rangayyan et al. [334], to estimate β, a Hann window was first applied to the given VAG signal; then, the DFT of the windowed signal was calculated, and the squared magnitude of the result was used to obtain the PSD of the signal. The slope of the best fitting line to the log–log plot of PSD was then determined. See Chang et al. [336] for related discussions.

When applying the PSA method, it is important to specify an appropriate frequency range to obtain the linear fit. Low-frequency details related to any slow drift present in the signal as well as high-frequency content dominated by noise or artifacts in the signal should be disregarded in order to obtain accurate estimates of β and FD. In the works of Rangayyan and Oloumi [333] and Rangayyan et al. [334], when applying the PSA method to the synthesized fBm signals, PSD components spanning 1% of the total frequency range were removed at both the low and high ends of the positive frequency axis. For VAG signals, the frequency range to obtain the linear fit was varied between $[10, 300]\ Hz$ and $[10, 380]\ Hz$ to study the effect of the frequency range used on classification accuracy. This was based upon experimentation and analysis of the band of frequencies over which appreciable differences were observed between normal and abnormal VAG signals. (The VAG signals had been prefiltered to the range $10\ Hz$ to $1\ kHz$ at the time of data acquisition; therefore, there is no useful information below $10\ Hz$.)

6.6.3 Examples of synthesized fractal signals

Two types of methods are available to generate fBm signals with known FD: approximation by spatial methods and approximation by spectral synthesis. The most reliable method is spectral synthesis by means of inverse Fourier filtering [337]. If a_k is the k^{th} complex coefficient of the DFT, to obtain $P(f) \propto 1/f^{\beta}$, the condition is

$$E[|a_k|^2] \propto \frac{1}{k^{\beta}}, \tag{6.49}$$

where k denotes the frequency index corresponding to f. By randomly selecting coefficients that satisfy the stated condition and then taking their inverse DFT, we can obtain the corresponding signal in the time domain. In the works of Rangayyan and Oloumi [333] and Rangayyan et al. [334], a direct implementation of the algorithm given by Saupe [337] was used to generate 110 synthetic signals, 10 for each value of H, with $0 \leq H \leq 1$ in intervals of 0.1. Figure 6.15 illustrates four examples of the synthesized signals. It is evident from the illustration that the variability and complexity of the signals increase as H decreases (or as FD increases).

The estimated FD values for the four synthesized signals shown in Figure 6.15 are given in the caption of the figure. The RMS error between the known and estimated values of FD for the 110 synthesized signals was computed to be 0.0198, which indicates accurate estimation of FD by the PSA method. The PSA method performed well with the synthesized signals for FD in the range of $[1.1, 1.8]$, but overestimated FD values in the range $[1.9, 2.0]$.

Estimates of FD were also obtained for the synthesized fBm signals using the well-known and popular box-counting and ruler methods [275, 276, 282] (see Section 5.13); the two methods led to slightly poorer results, with higher RMS errors of 0.1387 and 0.2243, respectively. The results indicate that the PSA method is the best suited method for the estimation of FD of fBm and self-affine signals, among the three methods mentioned.

6.6.4 Fractal analysis of segments of VAG signals

The database used in the studies of Rangayyan and Oloumi [333] and Rangayyan et al. [334] consists of 89 VAG signals; see Section 5.12.1 for details. Examples of a normal signal and an abnormal signal are shown in Figure 6.16 and 6.17. Each signal was normalized to the amplitude range $[0, 1]$ and resampled to the length of $8,192$ samples by linear interpolation.

It is known that the VAG signal is nonstationary and that it is appropriate to analyze segments of fixed or adaptive duration [96,264–266,269] or apply nonstationary signal processing techniques such as wavelets [270] and TFDs [273]. However, such methods increase the computational load. Furthermore, it is difficult, in practice, to associate parts of the knee joint affected by pathology to segments of VAG signals. The use of VAG signal segments of limited duration may lead to inaccuracies in the estimation of the PSD in low-frequency ranges. On the other hand, it has been observed that parts of VAG signals during certain portions of the swing cycle, especially extension, can provide features with higher discriminant capability than other parts [264–266, 269]. In the works of Rangayyan and Oloumi [333] and Rangayyan et al. [334], to facilitate separate analysis of parts of VAG signals during extension and flexion and to derive multiple features, the PSA method was applied not only to get a parameter for the full VAG signal in each case ($FD1$), but also to the first and second halves of the normalized duration ($FD1h$ and $FD2h$, corresponding to extension and flexion), as well as to four segments with each spanning one quarter of the total duration ($FD1q$, $FD2q$, $FD3q$, and $FD4q$).

Figures 6.18 and 6.19 show the PSDs and the linear fits derived to estimate the $FD1$ values of the normal and abnormal VAG signals shown in Figures 6.16 and 6.17, respectively. The estimated $FD1$ values for the two signals are given in the captions of Figures 6.18 and 6.19. The average and SD values of $FD1$ for the 51 normal signals were 1.8061 ± 0.2398; those for the 38 abnormal VAG signals were 1.6695 ± 0.2226 (using the frequency range $[10, 300]$ Hz). The discriminant capability of the FD values was assessed in terms of the area, A_z, under the receiver operating characteristic (ROC) curve by using ROCKIT [338]. The FD values obtained by the PSA method applied to the full VAG signals led to $A_z = 0.6872$.

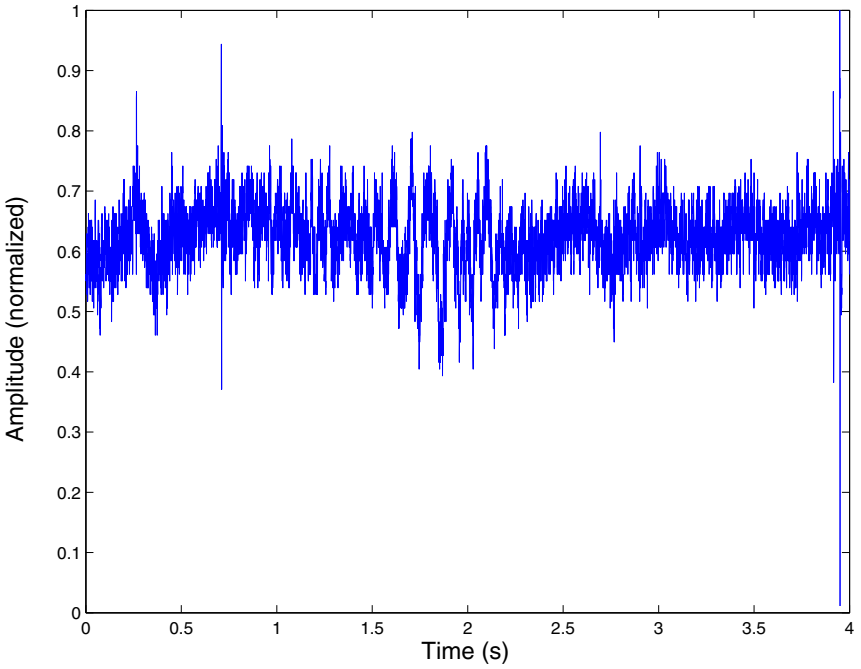

Figure 6.16 Example of a normal VAG signal. Reproduced with permission from R.M. Rangayyan and F. Oloumi, Fractal analysis of knee-joint vibroarthrographic signals, Proceedings of the 10^{th} IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010), Corfu, Greece, November 2010, pp 1–4. ©IEEE.

The values of FD estimated for the first half (extension) and second half (flexion) of each signal led to poorer A_z values of 0.6133 and 0.5916, respectively. With the FD values derived using quarter portions of the VAG signals, the four parts, in order, led to $A_z = 0.6546$, 0.7394, 0.5959, and 0.7023 (using the frequency range $[10, 380]$ Hz). Based on the p-values obtained via the t-test, it was noted that the differences between the values of $FD1$, $FD2q$, and $FD4q$ were statistically highly significant ($p < 0.01$), whereas the differences in the $FD1q$ values for the normal and abnormal categories were statistically significant ($0.01 < p < 0.05$). Thus, the usefulness of some of the FD-based parameters in discriminating between normal and abnormal VAG signals was demonstrated.

The FD values for abnormal VAG signals were observed to be, on the average, lower than those for normal signals, which was interpreted as follows. The presence of defects in the articular surfaces of the knee joint due to loss or shedding of cartilage leads to structured vibration components that disrupt or replace the random, fBm-like signals associated with the normal friction-free cartilage surfaces. Although the VAG signals of some normal knees could possibly contain nonrandom components, such as clicks and crepitus, they would be fewer and occur less often than the grinding

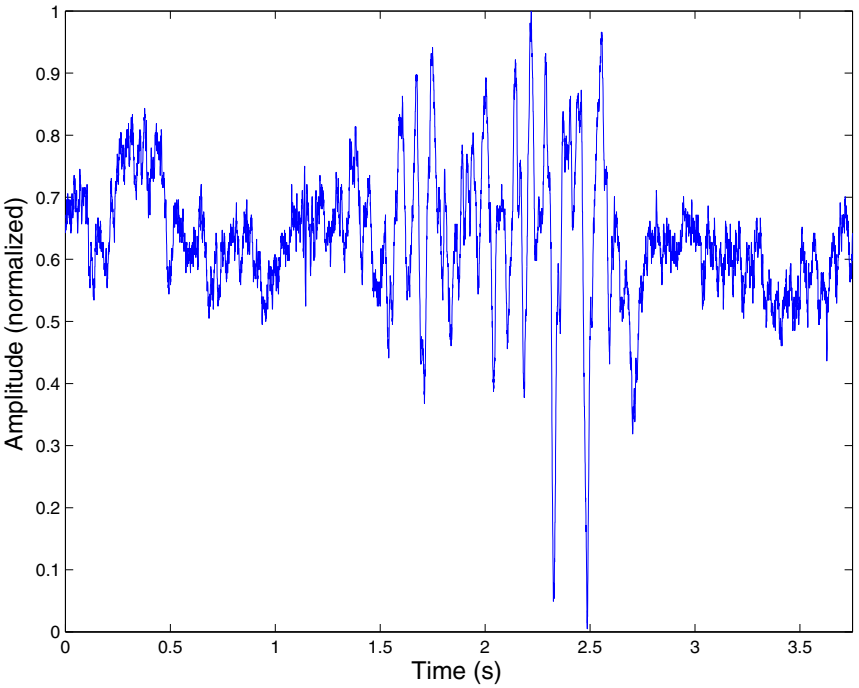

Figure 6.17 Example of an abnormal VAG signal. Reproduced with permission from R.M. Rangayyan and F. Oloumi, Fractal analysis of knee-joint vibroarthrographic signals, Proceedings of the 10^{th} IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010), Corfu, Greece, November 2010, pp 1–4. ©IEEE.

noise components found in abnormal knee joints. The disruption of the fBm-like characteristics of normal VAG signals by pathological conditions may be expected to result in lower FD values for abnormal VAG signals via the PSA approach and the $1/f$ model. The range of the FD values of the abnormal VAG signals obtained in the study of Rangayyan et al. [334] agrees with the ranges of FD of acceleration signals obtained from finger joints of patients with calcium pyrophosphate deposition disease, rheumatoid arthritis, or spondyloarthropathy in the study of Shah et al. [300] (see Section 5.13.3).

The classification performance of the features derived via fractal analysis, with the highest A_z value of about 0.74 in ROC analysis, was noted to be comparable to the performance provided by several features obtained by predictive modeling, cepstral analysis, measures of variability of power, turns count, probabilistic models, wavelets, and TFDs [96, 264–266, 270, 273]. Selected combinations of FD with some of the other features mentioned above led to better results with A_z values in the range $[0.92, 0.96]$. See Rangayyan et al. [334] for further details and additional results with various feature selection and pattern classification procedures.

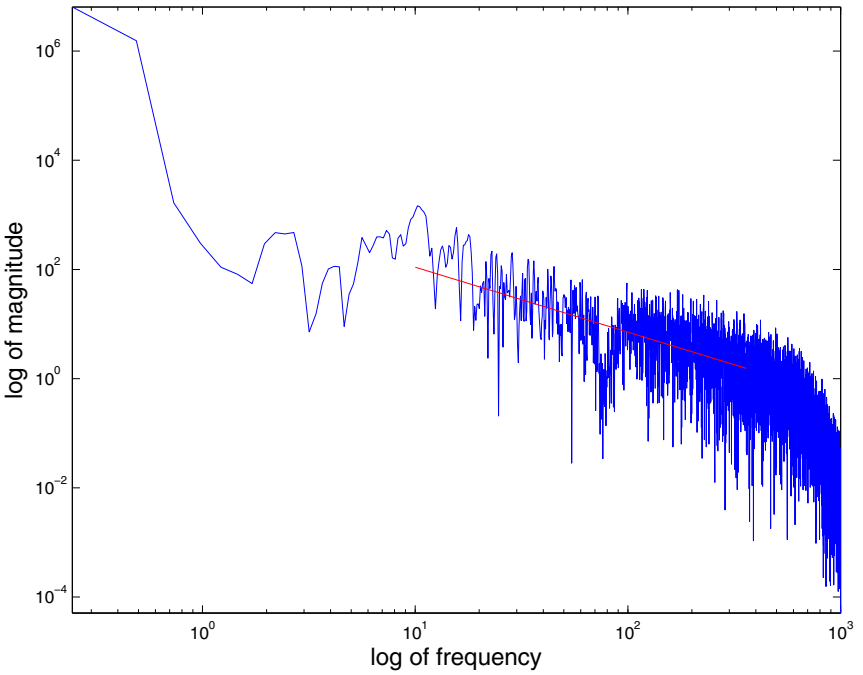

Figure 6.18 PSD of the normal VAG signal in Figure 6.16 with the straight-line fit to the range $[10, 360]$ Hz. $FD1 = 1.905$. Reproduced with permission from R.M. Rangayyan and F. Oloumi, Fractal analysis of knee-joint vibroarthrographic signals, Proceedings of the 10^{th} IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010), Corfu, Greece, November 2010, pp 1–4. ©IEEE.

6.7 Application: Spectral Analysis of EEG Signals

Problem: *It is known that the EEG signal exhibits shifts towards lower frequencies as the subject goes into deeper stages of sleep [57, 58]. Propose methods to obtain quantitative measures to characterize the spectral content of EEG signals and relate them to stages of sleep.*

Solution: As shown in Section 1.2.6, EEG signals exhibit rhythms in certain frequency bands. A subject with eyes closed and awake will typically demonstrate alpha waves in the EEG. As the subject goes to sleep, the EEG rhythms are expected to shift to lower bands of theta and delta rhythms (see Figure 1.38). Episodic EEG activities related to rapid eye movement (REM) and sleep spindles (see Figure 4.1) could contribute power in frequency bands that are higher than the theta and delta bands. Overnight PSG studies commonly include several channels of the EEG signal that are used to establish the stage of sleep by either manual (visual) analysis or computer-aided analysis [57]. The amount of low-frequency power in the delta and

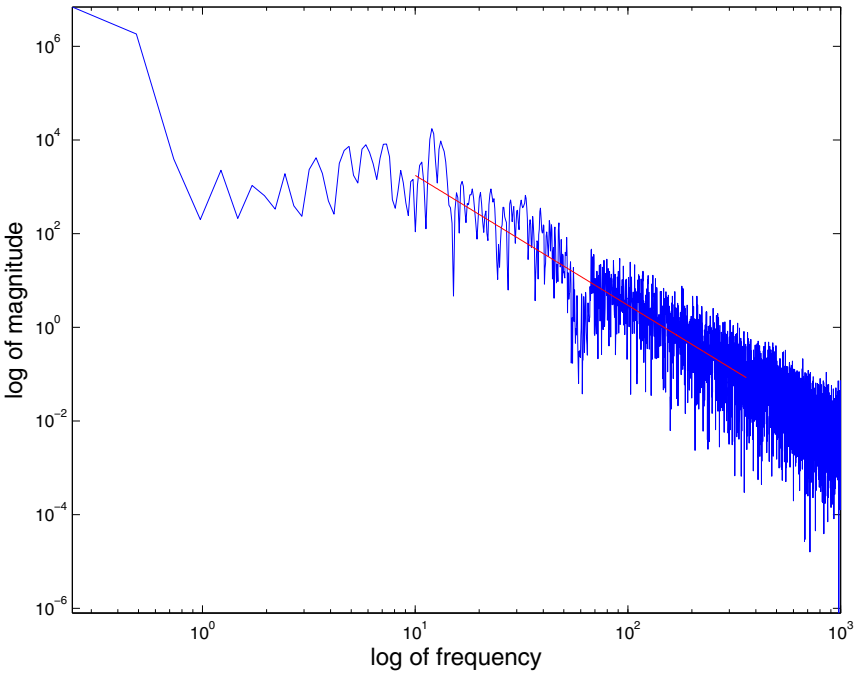

Figure 6.19 PSD of the abnormal VAG signal in Figure 6.17 with the straight-line fit to the range $[10, 360]$ Hz. $FD1 = 1.113$. Reproduced with permission from R.M. Rangayyan and F. Oloumi, Fractal analysis of knee-joint vibroarthrographic signals, Proceedings of the 10^{th} IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010), Corfu, Greece, November 2010, pp 1–4. ©IEEE.

theta bands in relation to the power in the alpha band is a useful feature in sleep staging [57].

Figure 6.20 illustrates three traces of the C3–A2 EEG channel of a subject during sleep in stages 0, 1, and 2. (Stage 0 indicates that the subject is awake.) The EEG signal was analyzed by an expert who assigned sleep stages to $5,988$-sample segments, corresponding to intervals of 29.94 s with $f_s = 200$ Hz [57]. Regardless of the noise present in the signals, it is evident that, in general, the signals shift to increasing power in lower frequencies as the sleep stage progresses from 0 to 1 to 2.

In order to obtain estimates of the PSDs of the EEG signals, the Welch method of windowing signal segments and averaging their PSDs was applied to $5,988$-sample segments of the EEG signal. Nonoverlapping segments of length 512 samples were obtained and the Hann window was applied before computing the DFT. The magnitude spectra of 10 segments were averaged. Figure 6.21 illustrates the average log-magnitude PSDs of three EEG segments corresponding to sleep stages 0, 1, and 2 (top to bottom). Sample traces of the same EEG segments with durations of 512 samples are shown in Figure 6.20. The PSDs clearly demonstrate a reduction in the

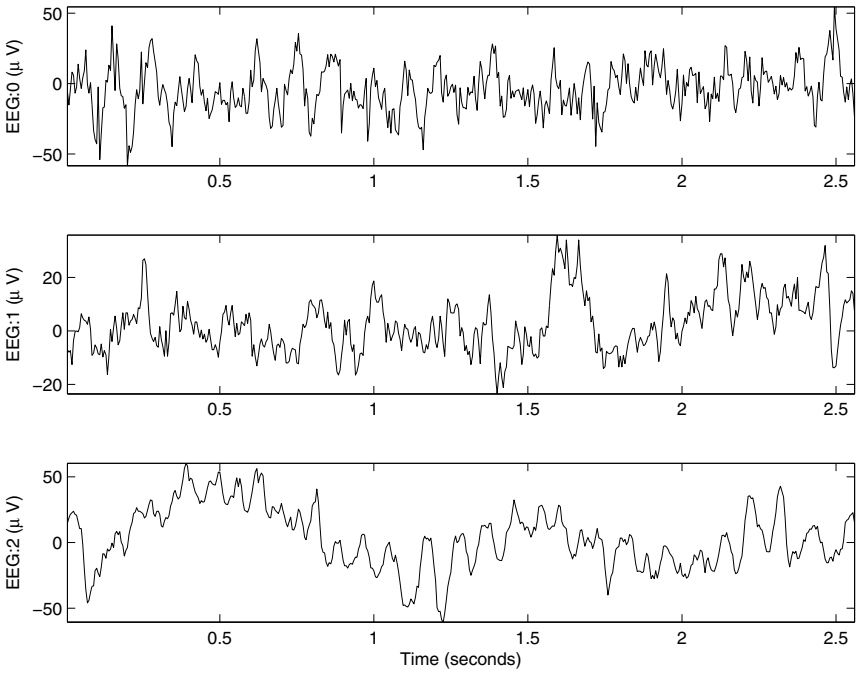

Figure 6.20 Top to bottom: Segments of the C3–A2 EEG channel of a subject during sleep in stages 0, 1, and 2. See also Figure 6.21. EEG data courtesy of R. Agarwal [57].

power of the EEG signals in the higher parts of the range of $0 - 25\ Hz$ plotted in the figure. The EEG signal for stage 0, which represents a wakeful state with the eyes closed (presumably), shows a clear peak at about $8\ Hz$ corresponding to the alpha rhythm. No peak in the alpha range is seen in the remaining two PSDs, which also show a reduction in high-frequency power. Peaks at lower frequencies (in the delta and theta bands) are evident in the PSDs of the EEG for sleep stages 1 and 2.

The mean frequencies of the PSDs were computed according to Equation 6.30 over the band of $[0, 100]\ Hz$. For the PSDs illustrated in Figure 6.21, the mean frequencies are $20.21, 10.09$, and $3.98\ Hz$, respectively, indicating a clear downward trend over the sleep stages of 0, 1, and 2. Lowpass filtering the signals to remove noise and reducing the frequency range used to compute the mean frequency could lead to lower values.

Instead of computing a single mean frequency value over an entire PSD, one could also compute fractions of the total power in several frequency bands, as illustrated in Figure 6.2 and in the discussion in Section 6.4.2. One of the features used by Agarwal and Gotman [57] is the the ratio of power in the alpha band (defined by them as 8.0 to $11\ Hz$) to the combined power in the delta (0.5 to $3.5\ Hz$) and theta (3.5 to $8.0\ Hz$) bands; this was referred to as the alpha-to-slow-wave index. The values of this index for the PSDs illustrated in Figure 6.21 are 1.68, 0.21, and 0.09,

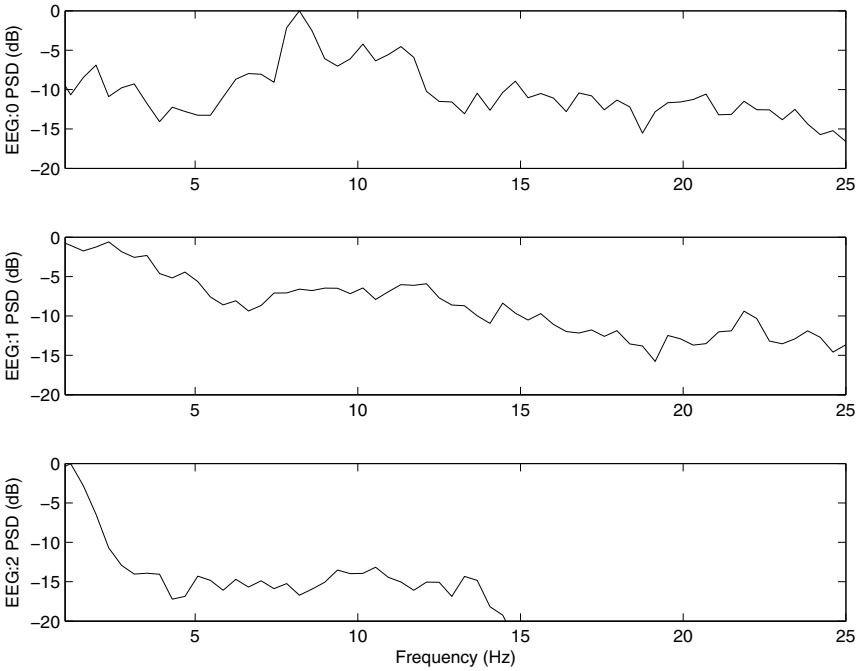

Figure 6.21 Top to bottom: Average PSDs of segments of the C3–A2 EEG channel of a subject during sleep in stages 0, 1, and 2. Figure 6.20 shows sample segments used in the derivation of the PSDs.

which indicate the expected reducing trend from stage 0 to 1 to 2, representative of the fact that alpha waves are replaced by slower theta and delta waves.

Figures 6.22 and 6.23 show cluster plots of the mean frequency and alpha-to-slow-wave index values for 702 segments of an overnight sleep EEG record with sleep stages of 0, 1, and 2. (Each segment has $5,988$ samples, corresponding to an interval of $29.94\ s$ with $f_s = 200\ Hz$. The total EEG record has a duration of about $6\ h$ and $50\ min$.) Each '$\times$' mark in the plots represents the value of the parameter and the corresponding sleep stage for one EEG segment. In general, the two parameters show trends of decreasing values as the sleep stage varies from 0 to 1 to 2. However, both parameters show substantial variance and overlap in their ranges for the three stages of sleep. It should be noted that EEG signals of a subject who is awake could contain substantial spectral variations. The presence of alpha waves in a stage-0 segment could cause the alpha-to-slow-wave index to be higher than that for a segment without alpha: This is indicated by the presence of multiple clusters of the index values for stage 0 in Figure 6.23.

Figure 6.24 illustrates a cluster plot of the mean frequency and alpha-to-slow-wave index values combined into a 2D vector for each EEG segment. The diamond marks for the feature vectors of stage 2 segments have formed a tight cluster with low

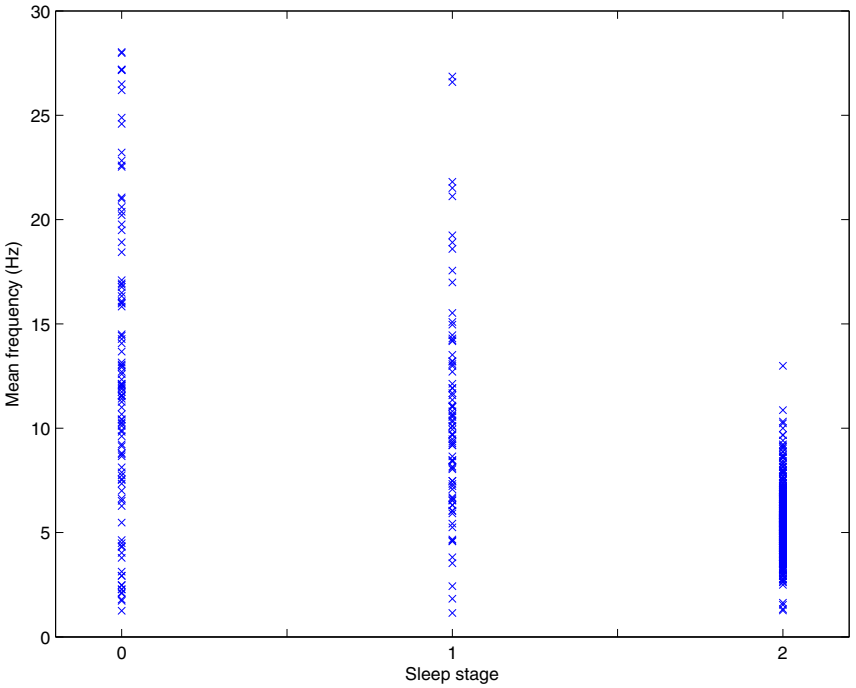

Figure 6.22 Cluster plot of mean frequency values for 702 EEG segments with sleep stages of 0, 1, and 2.

values for both features, as expected. However, the '○' and '+' marks representing the features for EEG segments corresponding to stages 0 and 1 overlap substantially. As mentioned in the preceding paragraph, stage-0 EEG segments could have a broad spread of the parameters depending upon the presence or absence of alpha; Figure 6.24 indicates the presence of two broad clusters for stage 0. Stage-1 EEG segments are also known to possess various types of waves and a broad spread of spectral features, as seen in Figure 6.24. These characteristics of EEG signals create difficulties in manual labeling of sleep stages and in automatic staging with a small number of parameters. The results shown in the present section indicate the need for additional features or measures to separate EEG segments of sleep stages 0 and 1.

EEG segments corresponding to sleep stages 3 and 4 were not present in the entire sleep record of the subject in the present case. Such segments could be expected to possess lower values for the two spectral parameters described in the present section.

Notwithstanding the results demonstrated in this section, it should be noted that staging of sleep requires several parameters derived from not only multiple channels of the EEG signal but also additional channels of the EOG and EMG signals; see Agarwal and Gotman [57] for descriptions of several measures for this application.

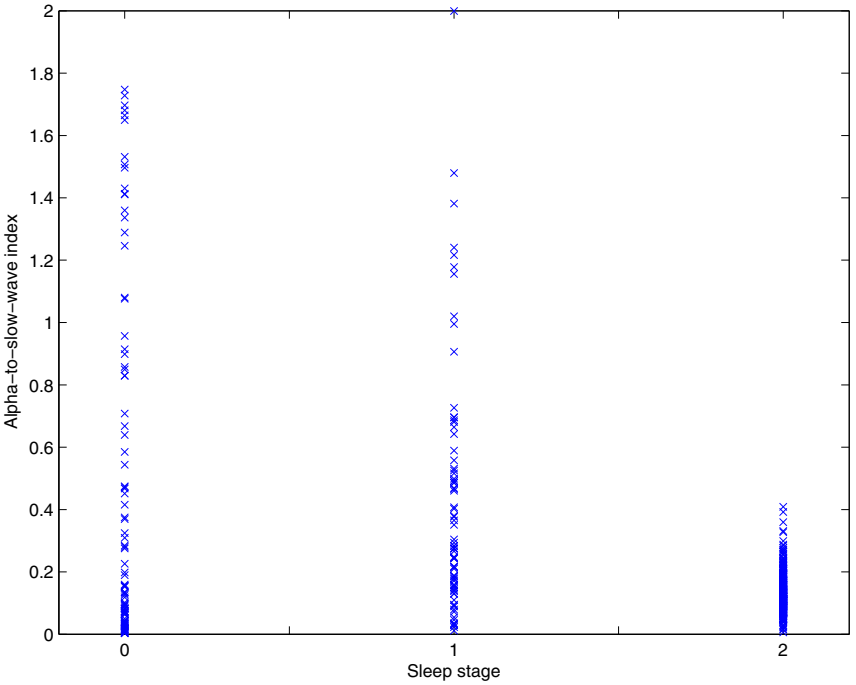

Figure 6.23 Cluster plot of the alpha-to-slow-wave index for 702 EEG segments with sleep stages of 0, 1, and 2.

6.8 Remarks

We have investigated the frequency-domain characteristics of a few biomedical signals and the corresponding physiological systems, with particular attention to the PCG and the cardiovascular system. Frequency-domain analysis via PSDs and parameters derived from PSDs can enable us to view the signal from a different perspective than the time domain. Certain signals such as the PCG and EEG may not lend themselves to easy interpretation in the time domain and, therefore, may benefit from a move to the frequency domain.

PSDs and their parameters facilitate investigation of the behavior of physiological systems in terms of rhythms, resonance, and parameters that could be related to the physical characteristics of anatomical entities (for example, the loss of elasticity of the myocardial muscles due to ischemia or infarction, the extent of aortic valvular stenosis, or the extent of calcification and stiffness of bioprosthetic valves). Pathological states may also be derived or simulated by modifying the spectral parameters or representations of the corresponding normal physiological states and signals.

It is worthwhile to pause at this stage of our study, and recognize the importance of the topics presented in the preceding chapters. A good understanding of the phy-

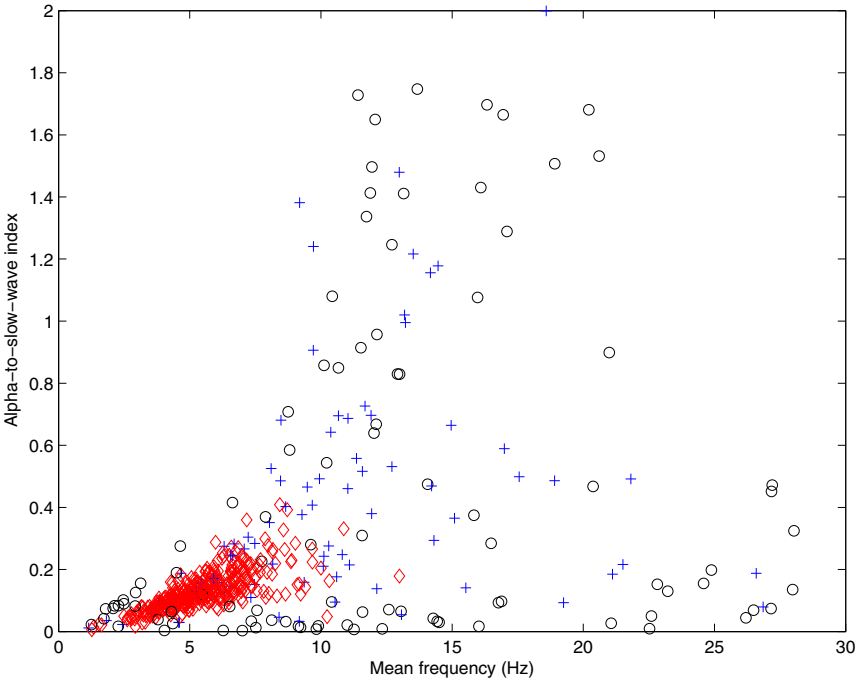

Figure 6.24 Cluster plot of the mean frequency and alpha-to-slow-wave index (ASI) for 702 EEG segments with sleep stages of 0 ('○' mark), 1 ('+' mark), and 2 (diamond mark).

siological systems that produce the biomedical signals we deal with, as well as of the pathological processes that alter their characteristics, is of paramount importance before we may process the signals. Preprocessing the signals to remove artifacts and detect events is essential before we may derive parameters to facilitate their analysis in the time and/or frequency domains. The design of biomedical signal analysis techniques requires a thorough understanding of the characteristics and properties of the biomedical systems behind the signals, in addition to detailed knowledge of mathematical principles, computer techniques, and digital signal processing algorithms.

6.9 Study Questions and Problems

1. The impulse response of a filter is specified by the series of sample values $\{3, 1, -1, 0\}$. (a) What will be the response of the filter to the input whose sample values are $\{4, 4, 2, 1\}$? (b) Is the filter response obtained by linear convolution or circular convolution of the input with the impulse response? (c) What will be the response with the type of convolution other than the one you indicated as the answer to the questions above? (d) How would you implement convolution of the two signals listed above using the FFT? Which type of convolution will this procedure provide? How would you get the other type of convolution for the signals in this problem via the FFT-based procedure?

2. A conjugate symmetric (even) signal $x_e(n)$ is defined as a signal with the property $x_e(n) = x_e^*(-n)$. A conjugate antisymmetric (odd) signal $x_o(n)$ is defined as a signal with the property $x_o(n) = -x_o^*(-n)$. An arbitrary signal $x(n)$ may be expressed as the sum of its conjugate symmetric and conjugate antisymmetric parts as $x(n) = x_e(n) + x_o(n)$, where $x_e(n) = \frac{1}{2}[x(n) + x^*(-n)]$ and $x_o(n) = \frac{1}{2}[x(n) - x^*(-n)]$. Prove that $FT[x_e(n)] = \text{real}[X(\omega)]$, and $FT[x_o(n)] = j\text{imag}[X(\omega)]$, where $FT[x(n)] = X(\omega)$, and FT stands for the Fourier transform [174].

3. A signal $x(t)$ is transmitted through a channel. The received signal $y(t)$ is a scaled, shifted, and noisy version of $x(t)$ given as $y(t) = \alpha x(t - t_0) + \eta(t)$ where α is a scale factor, t_0 is the time delay, and $\eta(t)$ is noise. Assume that the noise process has zero mean and is statistically independent of the signal process, and that all processes are stationary. Derive expressions for the PSD of $y(t)$ in terms of the PSDs of x and η [6, 212].

4. Consider a continuous-time sinusoidal signal of frequency $10\ Hz$. (a) Derive an analytical expression for the ACF of the signal. (b) Draw a schematic plot of the ACF, including detailed labeling of the time axis. (c) State the relationship of the PSD to the ACF. (d) Derive an analytical expression for the PSD of the given signal. (e) Draw a schematic plot of the PSD, including detailed labeling of the frequency axis.

5. Two real signals $x_1(n)$ and $x_2(n)$ are combined to form a complex signal defined as $y(n) = x_1(n) + jx_2(n)$. Derive a procedure to extract the DFTs $X_1(k)$ and $X_2(k)$ of $x_1(n)$ and $x_2(n)$, respectively, from the DFT $Y(k)$ of $y(n)$.

6. Distinguish between ensemble averages and temporal (time) averages. Identify applications of first-order and second-order averages of both types in PCG analysis.

7. Propose a procedure to process PCG signals to identify the possible presence of murmurs due to aortic stenosis.

8. Propose an algorithm to detect the presence of the alpha rhythm in an EEG signal. Propose an extension to the algorithm to detect the joint presence of the same rhythm in four simultaneously recorded EEG channels.

9. Give an equation to define the mean frequency (centroidal frequency) of a PSD function. Explain the role of each item of your equation. Explain how you would implement the procedure to obtain the mean frequency of a signal. Draw schematic sketches of two PSD functions and indicate their approximate mean frequencies. Explain the difference between the two examples.

10. Using continuous-time and continuous-frequency representation, the fraction of the energy of a signal $x(t)$ in the frequency band $[f_1 : f_2]\ Hz$ is given by

$$E_{f_1:f_2} = \frac{\int_{f_1}^{f_2} |X(f)|^2\, df}{\int_0^{\infty} |X(f)|^2\, df}\,.$$

Here, $X(f)$ is the Fourier transform of the signal $x(t)$ being processed. A researcher obtains a sampled version $x(n)$ of the signal $x(t)$ with the sampling frequency $f_s = 1,000\ Hz$. The number of samples in the signal is $N = 1,900$. The researcher then pads the signal with zeros to increase the length to $M = 2,048$ samples and applies the FFT routine to obtain the DFT $X(k)$ of the signal. The researcher wants to compute the fractions of the energy of the signal in the frequency bands $[0 : 50]\ Hz$, $[51 : 100]\ Hz$, and $[101 : 400]\ Hz$. Help the researcher by writing an algorithm to compute the three

fractions. Ensure that you give the index of the DFT array for each frequency listed above. Sketch a schematic representation of the spectrum $X(k)$ over the $M = 2,048$ samples in the DFT array. Indicate all of the frequency bands required to compute the three fractions of energy, in terms of the frequency in Hz and the index of the DFT array.

11. A researcher in biomedical signal processing wishes to compute averaged estimates of the PSD for the systolic parts and diastolic parts of PCG signals. The related ECG and carotid pulse (CP) signals are available. The researcher also wants to derive quantitative measures from the averaged PSD. Help the researcher with the following: (a) Give a step-by-step algorithm to identify the beginning and end points of the systolic part and the diastolic part for each cardiac cycle in a given PCG signal. Explain how the ECG and CP signals may be used for this purpose. No equation is required for this part. (b) Give a step-by-step algorithm to compute the PSD of each systolic or diastolic segment and to derive the averaged PSD for systole and diastole over an entire PCG signal including several cardiac cycles. (c) Explain the notions of cyclostationary signals and synchronized averaging. Relate these concepts to the PCG, ECG, and CP signals and the procedure to obtain averaged PSDs as above. (d) Given the averaged PSD, $S(k)$, $k = 1, 2, \ldots, N$, where N is the number of the samples used in the computation of the DFT, explain how the mean frequency of the PSD may be computed. Give an equation and explain each term or variable used.

12. A student new to the area of biomedical signal processing wishes to develop methods to detect the presence of murmurs in PCG signals. The student wishes to derive quantitative measures from segments of the PCG signal to facilitate characterization of murmurs. Help the student with the following: (a) Explain the notion of waveform complexity. Give equations to compute the mean, the variance, and the form factor (FF) of a signal $x(n)$ with N samples. Explain how FF may be expected to differ between normal heart sounds and murmurs. (b) Give a step-by-step algorithm to compute the PSD of a signal. Include mathematical formulas to compute the DFT of the signal and to obtain the PSD from the DFT. (c) Give typical bandwidths of normal heart sounds and murmurs. Draw schematic diagrams of the PSDs of normal heart sounds and murmurs, and explain the differences between them. (d) Give a step-by-step procedure to compute the ratio of the power of a given signal in a particular band of frequencies $[f_1, f_2]$ to the total power of the signal. (e) Assume that a signal segment has $N = 500$ samples, that the DFT is computed using the same number of samples, and that the sampling rate is $1,000\ Hz$. For the purpose of detecting murmurs in PCG signals, give a typical range, in Hz, for the band $[f_1, f_2]$ mentioned above. For the same range of frequencies, give the range of the DFT samples to be used to compute the ratio of power mentioned above. Explain how this measure could be used to detect murmurs and to distinguish murmurs from normal heart sounds.

6.10 Laboratory Exercises and Projects

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

`http://people.ucalgary.ca/~ranga/enel563`

1. Using MATLAB®, prepare a signal that contains the sum of two cosine waves of equal amplitude at $40\ Hz$ and $45\ Hz$. Let the sampling rate be $1\ kHz$. (a) Compute the power

spectrum of the signal with a rectangular window of duration 2 s. (b) Compute the power spectrum of the signal with a Hamming window of duration 2 s. (c) Compute the power spectrum of the signal with a rectangular window of duration 0.5 s. (d) Compute the power spectrum of the signal with a Hamming window of duration 0.5 s.

To obtain the power spectrum, you may take the FFT and square the absolute value of the result. Compare the spectra obtained in parts (a)–(d) and comment upon their similarities and/or differences. In order to visualize the differences clearly, use $2,048$-point FFTs and plot the logarithm of the magnitude-squared spectra with an expanded scale from 0 to 100 Hz only. Be sure to label the frequency axis in Hz!

What should the ideal spectrum look like?

2. Two VAG signals are given in the files vag1.dat and vag2.dat (see also the file vag.m). The sampling rate is 2 kHz. Obtain and plot their PSDs using MATLAB®. Label the frequency axis in Hz!

   Compute the mean frequency as the first moment of the PSD for each signal. Compute also the variance frequency (second central moment) of each PSD. What are the units of these parameters?

   Compare the spectra and the parameters derived and give your evaluation of the frequency content of the signals.

3. The file safety.wav contains the speech signal for the word "safety" uttered by a male speaker, sampled at 8 kHz (see also 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. Compute the PSD for each segment that you obtain and study its characteristics.

4. 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 pec33.dat has systolic murmur, and is of a patient suspected to have pulmonary stenosis, ventricular septal defect, and pulmonary hypertension.

   Apply the Pan–Tompkins method for QRS detection to the ECG channel and the Lehner and Rangayyan method to detect the dicrotic notch in the carotid pulse channel. Extrapolate the timing information from the ECG and carotid pulse channels to segment the PCG signal into two parts: the systolic part from the onset of an S1 and to the onset of the following S2, and the diastolic part from the onset of an S2 to the onset of the following S1. Compute the PSD of each segment.

   Extend the procedure to average the systolic and diastolic PSDs over several cardiac cycles. Compare the PSDs obtained for the three cases.

5. Compute the mean frequency and the ratio of the energy in the range $100 - 300$ Hz to the total energy for each PSD derived in the previous problem. What can you infer from these measures?

6. Compute the PSDs of a few channels of the EEG in the file eeg1-xx.dat using Welch's procedure (see also the file eeg1.m). Study the changes in the PSDs derived with variations in the window width, the number of segments averaged, and the type of the window used. Compare the results with the PSDs computed using the entire signal in each channel. Discuss the results in terms of the effects of the procedures and parameters on spectral resolution and leakage.

7. Compute the PSD of an ECG signal over a long duration including several cardiac cycles or beats. It would be good to select a segment with a steady heart rate (almost the same RR interval for each beat) and without any PVCs. Extract a segment over one cardiac cycle from the same signal and compute its PSD. Compare the two PSDs and explain the similarities and/or differences between them.