CHAPTER 4

DETECTION OF EVENTS

Biomedical signals carry signatures of physiological events. The part of a signal related to a specific event of interest is often referred to as an *epoch*. Analysis of a signal for monitoring or diagnosis requires the identification of epochs and investigation of the corresponding events. Once an event has been identified, the corresponding waveform may be segmented and analyzed in terms of its amplitude, waveshape (morphology), time duration, intervals between events, energy distribution, frequency content, and other characteristics. Detection of events is, therefore, an important step in biomedical signal analysis.

4.1 Problem Statement

A generic problem statement applicable to the theme of this chapter may be formulated as follows:

> *Given a biomedical signal, identify discrete signal epochs and correlate them with events in the related physiological processes.*

In the sections to follow, we shall first study a few examples of epochs in different biomedical signals, with the aim of understanding the nature of the related physio-

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

logical events. Such an understanding will help in the subsequent development of signal processing techniques to emphasize, detect, and analyze epochs.

4.2 Illustration of the Problem with Case Studies

The following sections provide illustrations of several events in biomedical signals. The aim of the illustrations is to develop an appreciation of the nature of signal events. A good understanding of signal events will help in designing appropriate signal processing techniques for their detection.

4.2.1 The P, QRS, and T waves in the ECG

As shown in Section 1.2.5, a cardiac cycle is represented in a period of the repetitive ECG signal as the series of waves labeled P, QRS, and T. If we view the cardiac cycle as a series of events, we have the following epochs in an ECG waveform:

- **The P wave:** Contraction of the atria is triggered by the SA-node impulse. The atria do not possess any specialized conduction system as the ventricles do; as such, contraction of the atrial muscles takes place in a slow squeezing manner, with the excitation stimulus being propagated by the muscle cells themselves. For this reason, the P wave is a slow waveform, with a duration of about $80\ ms$. The P wave amplitude is much smaller (about $0.1 - 0.2\ mV$) than that of the QRS because the atria are smaller than the ventricles. The P wave is the epoch related to the event of atrial contraction. (Atrial relaxation does not produce any distinct waveform in the ECG as it is overshadowed by the following QRS wave.)

- **The PQ segment:** The AV node provides a delay to facilitate completion of atrial contraction and transfer of blood to the ventricles before ventricular contraction is initiated. The resulting PQ segment, of about $80\ ms$ duration, is thus a "nonevent"; however, it is important in recognizing the baseline as the interval is usually isoelectric.

- **The QRS wave:** The specialized system of Purkinje fibers stimulate contraction of ventricular muscles in a rapid sequence from the apex upwards. The almost-simultaneous contraction of the entire ventricular musculature results in a sharp and tall QRS complex of about $1\ mV$ amplitude and $80 - 100\ ms$ duration. The event of ventricular contraction is represented by the QRS epoch.

- **The ST segment:** The normally flat (isoelectric) ST segment is related to the plateau in the action potential of the left ventricular muscle cells (see Figures 1.7 and 1.24). The duration of the plateau in the action potential is about $200\ ms$; the ST segment duration is usually about $100 - 120\ ms$. As in the case of the PQ segment, the ST segment may also be called a nonevent. However, myocardial ischemia or infarction could change the action potentials of a portion of

the left ventricular musculature and cause the ST segment to be depressed (see Figure 1.46) or elevated. The PQ segment serves as a useful reference when the isoelectric nature of the ST segment needs to be verified.

- **The T wave:** The T wave appears in a normal ECG signal as a discrete wave separated from the QRS by an isoelectric ST segment. However, it relates to the last phase of the action potential of ventricular muscle cells, when the potential returns from the plateau of the depolarized state to the resting potential through the process of repolarization [33]. The T wave is commonly referred to as the wave corresponding to ventricular relaxation. While this is correct, it should be noted that relaxation through repolarization is simply the final phase of contraction: Contraction and relaxation are indicated by the upstroke and downstroke of the same action potential. For this reason, the T wave may be said to relate to a nonspecific event.

  The T wave is elusive, being low in amplitude ($0.1 - 0.3\ mV$) and being a slow wave extending over $120 - 160\ ms$. It is almost absent in many ECG recordings. Rather than attempt to detect the often obscure T wave, one may extract a segment of the ECG $80 - 360\ ms$ from the beginning of the QRS and use it to represent the ST segment and the T wave.

4.2.2 The first and second heart sounds

We observed in Section 1.2.9 that the normal cardiac cycle manifests as a series of the first and second heart sounds — S1 and S2. Murmurs and additional sounds may appear in the presence of cardiovascular diseases or defects. We shall concentrate on S1, S2, and murmurs only.

- **The first heart sound S1:** S1 reflects a sequence of events related to ventricular contraction — closure of the atrioventricular valves, isovolumic contraction, opening of the semilunar valves, and ejection of the blood from the ventricles [33]. The epoch of S1 is directly related to the event of ventricular contraction.

- **The second heart sound S2:** S2 is related to the end of ventricular contraction, signified by closure of the aortic and pulmonary valves. As we observed in the case of the T wave, the end of ventricular contraction cannot be referred to as a specific event per se. However, in the case of S2, we do have the specific events of closure of the aortic and pulmonary valves to relate to, as indicated by the corresponding A2 and P2 components of S2. Unfortunately, separate identification of A2 and P2 is confounded by the fact that they usually overlap in normal signals. If A2 and P2 are separated due to a cardiovascular disorder, simultaneous multisite PCG recordings will be required to identify each component definitively as they may be reversed in order (see Tavel [72] and Rushmer [33]).

- **Murmurs:** Murmurs, if present, could be viewed as specific events. For example, the systolic murmur of aortic stenosis relates to the event of turbulent

ejection of blood from the left ventricle through a restricted aortic valve opening. The diastolic murmur in the case of aortic insufficiency corresponds to the event of regurgitation of blood from the aorta back into the left ventricle through a leaky aortic valve.

4.2.3 The dicrotic notch in the carotid pulse

As we saw in Sections 1.2.10 and 1.2.11, closure of the aortic valve causes a sudden drop in aortic pressure that is already on a downward slope at the end of ventricular systole. The dicrotic notch inscribed in the carotid pulse is a delayed, upstream manifestation of the incisura in the aortic pressure wave. The dicrotic notch is a specific signature on the relatively nondescript carotid pulse signal, and may be taken as an epoch related to the event of aortic valve closure (albeit with a time delay); the same event also signifies the end of ventricular systole and ejection as well as the beginning of S2 and diastole.

4.2.4 EEG rhythms, waves, and transients

We have already studied a few basic characteristics of the EEG in Section 1.2.6, and noted the nature of the α, β, δ, and θ waves. We shall now consider a few events and transients that occur in EEG signals [54–56, 195–197]. Figure 4.1 shows typical manifestations of the activities described below [54].

- **K-complex:** This is a transient complex waveform with slow waves, sometimes associated with sharp components, and often followed by $14\ Hz$ waves. It occurs spontaneously or in response to a sudden stimulus during sleep, with an amplitude of about $200\ \mu V$.
- **Lambda waves:** These are monophasic, positive, sharp waves that occur in the occipital location with an amplitude of less than $50\ \mu V$. They are related to eye movement and are associated with visual exploration.
- **Mu rhythm:** This rhythm appears as a group of waves in the frequency range of $7-11\ Hz$ with an arcade or comb shape in the central location. The mu rhythm usually has an amplitude of less than $50\ \mu V$, and is blocked or attenuated by contralateral movement, thought of movement, readiness to move, or tactile stimulation.
- **Spike:** A spike is defined as a transient with a pointed peak, having a duration in the range of $20 - 30\ ms$.
- **Sharp wave:** A sharp wave is also a transient with a pointed peak, but with a longer duration than a spike, in the range of $70 - 200\ ms$.
- **Spike-and-wave rhythm:** A sequence of surface-negative slow waves in the frequency range of $2.5 - 3.5\ Hz$ and having a spike associated with each wave is referred to as a spike-and-wave rhythm. There could be several spikes of

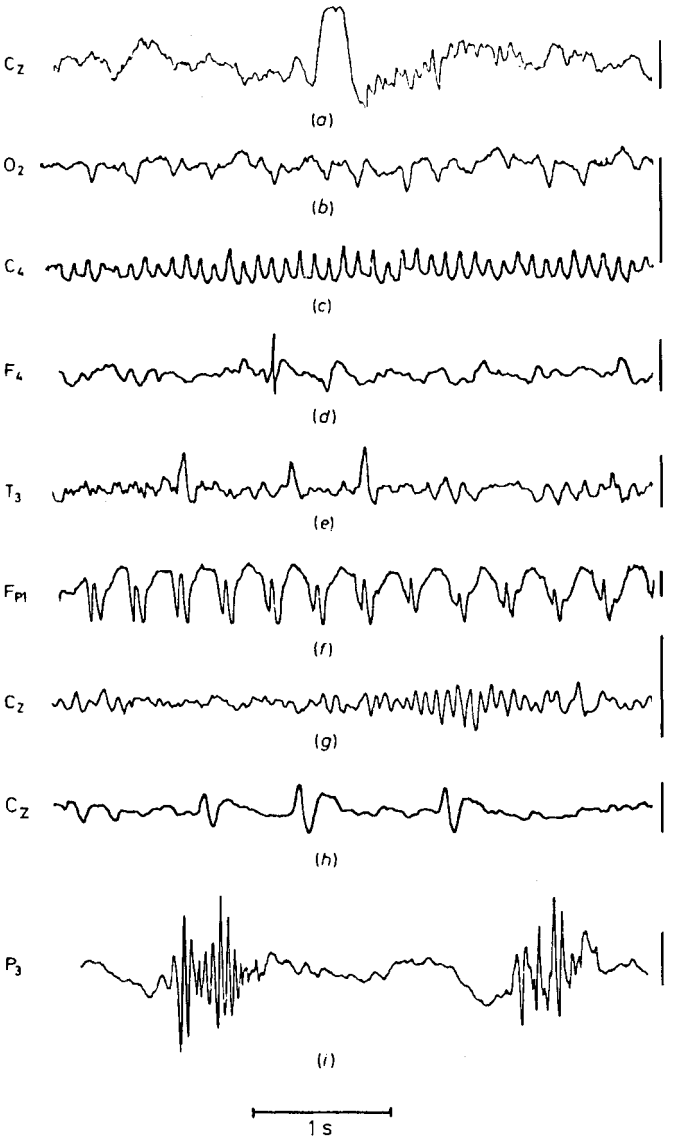

Figure 4.1 Top to bottom: (a) the K-complex; (b) the lambda wave; (c) the mu rhythm; (d) a spike; (e) sharp waves; (f) spike-and-wave complexes; (g) a sleep spindle; (h) vertex sharp waves; and (i) polyspike discharges. The horizontal bar at the bottom indicates a duration of 1 s; the vertical bars at the right indicate 100 μV. Reproduced with permission from R. Cooper, J.W. Osselton, and J.C. Shaw, *EEG Technology*, 3rd edition, 1980. ©Butterworth Heinemann Publishers, a division of Reed Educational & Professional Publishing Ltd., Oxford, UK.

amplitude up to $1,000\ \mu V$ in each complex, in which case the rhythm is called a polyspike-and-wave complex.

- **Sleep spindle:** This is an episodic rhythm at about $14\ Hz$ and $50\ \mu V$, occurring maximally over the frontocentral regions during certain stages of sleep. A spindle is defined, in general, as a short sequence of monomorphic waves having a fusiform appearance [55].

- **Vertex sharp transient** or **V-wave:** This wave represents a sharp potential that is maximal at the vertex at about $300\ \mu V$ and is negative in relation to the EEG in other areas. It occurs spontaneously during sleep or in response to a sensory stimulus during sleep or wakefulness.

In addition to the above, the term "burst" is used to indicate a phenomenon composed of two or more waves that are different from the principal (background) activity in terms of amplitude, frequency, or waveform. A burst is abrupt and has a relatively short duration [55].

An EEG record is described in terms of [54]

- the most persistent rhythm (for example, α);

- the presence of other rhythmic features, such as δ, θ, or β;

- discrete features of relatively long duration, such as an episode of spike-and-wave activity;

- discrete features of relatively short duration, such as isolated spikes or sharp waves;

- the activity remaining when all the previous features have been described, referred to as background activity; and

- artifacts, if any, giving rise to ambiguity in interpretation.

Each of the EEG waves or activities is described in chronological sequence in terms of amplitude; frequency, in the case of rhythmic features; waveform, in the case of both rhythmic and transient features; location or spatial distribution; incidence or temporal variability; the presence or absence of right–left symmetry in the location of activity; and responsiveness to stimuli, such as eye opening and closure. The EEG record at rest is first described as above; effects of evocative techniques are then specified in the same terms. Behavioral changes, such as the subject becoming drowsy or falling asleep, are also noted [54].

The EEG signals in Figure 1.39 demonstrate the presence of the α rhythm in all the channels. The EEG signals in Figure 1.40 depict spike-and-wave complexes in almost all the channels.

4.3 Detection of Events and Waves

We shall now see how the knowledge that we have gained so far of several biomedical signal events may be applied to develop signal processing techniques for their detection. Each of the following sections deals with the problem of detection of a specific type of event. The techniques described should find applications in the detection of other events of comparable characteristics.

4.3.1 Derivative based methods for QRS detection

Problem: *Develop signal processing techniques to facilitate detection of the QRS complex, given that it is the sharpest wave in an ECG cycle.*

Solution 1: We noted in Section 1.2.5 that the QRS complex has the largest slope (rate of change of voltage) in a cardiac cycle by virtue of the rapid conduction and depolarization characteristics of the ventricles. As the rate of change is given by the derivative operator, the $\frac{d}{dt}$ operation would be the most logical starting point in an attempt to develop an algorithm to detect the QRS complex.

We saw in Section 3.5.3 that the derivative operator enhances the QRS, although the resulting wave does not bear any resemblance to a typical QRS complex. Observe in Figures 3.54 and 3.55 that the slow P and T waves have been suppressed by the derivative operators, while the output is the highest at the QRS. However, given the noisy nature of the results of the derivative-based operators, it is also evident that significant smoothing will be required before further processing can take place.

Balda et al. [198] proposed a derivative-based algorithm for QRS detection, which was further studied and evaluated by Ahlstrom and Tompkins [199], Friesen et al. [200], and Tompkins [46]. The algorithm progresses as follows. In a manner similar to Equation 3.127, the smoothed three-point first derivative $y_0(n)$ of the given signal $x(n)$ is approximated as

$$y_0(n) = |x(n) - x(n-2)|. \tag{4.1}$$

The second derivative is approximated as

$$y_1(n) = |x(n) - 2x(n-2) + x(n-4)|. \tag{4.2}$$

The two results are weighted and combined to obtain

$$y_2(n) = 1.3y_0(n) + 1.1y_1(n). \tag{4.3}$$

The result $y_2(n)$ is scanned with a threshold of 1.0. Whenever the threshold is crossed, the subsequent eight samples are also tested against the same threshold. If at least six of the eight points pass the threshold test, the segment of eight samples is taken to be a part of a QRS complex. The procedure results in a pulse with its width proportional to that of the QRS complex; however, the method is sensitive to noise.

Illustration of application: Figure 4.2 illustrates, in the topmost trace, two cycles of a filtered version of the ECG signal shown in Figure 3.5. The signal was filtered

with an eighth-order Butterworth lowpass filter with $f_c = 90\ Hz$, down-sampled by a factor of five, and filtered with a notch filter with $f_o = 60\ Hz$. The effective sampling rate is $200\ Hz$. The signal was normalized by dividing by its maximum value.

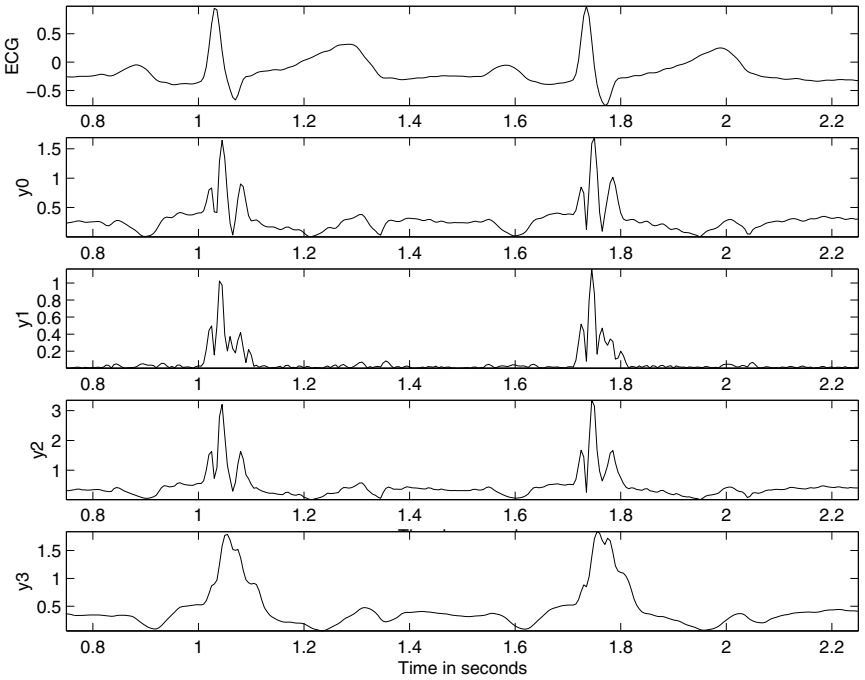

Figure 4.2 From top to bottom: two cycles of a filtered version of the ECG signal shown in Figure 3.5; output $y_0(n)$ of the first-derivative-based operator in Equation 4.1; output $y_1(n)$ of the second-derivative-based operator in Equation 4.2; the combined result $y_2(n)$ from Equation 4.3; and the result $y_3(n)$ of passing $y_2(n)$ through the 8-point MA filter in Equation 3.107.

The second and third plots in Figure 4.2 show the derivatives $y_0(n)$ and $y_1(n)$, respectively; the fourth plot illustrates the combined result $y_2(n)$. Observe the relatively high values in the derivative-based results at the QRS locations; the outputs are low or negligible at the P and T wave locations, in spite of the fact that the original signal possesses an unusually sharp and tall T wave. It is also seen that the results have multiple peaks over the duration of the QRS wave, due to the fact that the QRS complex includes three major swings: Q–R, R–S, and S–ST baseline in the present example (an additional PQ baseline–Q swing may also be present in other ECG signals).

The last plot in Figure 4.2 shows the smoothed result $y_3(n)$ obtained by passing $y_2(n)$ through the 8-point MA filter in Equation 3.107. We now have a single pulse with amplitude greater than 1.0 over the duration of the corresponding QRS complex.

A simple peak-searching algorithm may be used to detect each ECG beat. The net delay introduced by the filters should be subtracted from the detected peak location in order to obtain the corresponding QRS location.

Note that peak searching cannot be performed directly on an ECG signal: The QRS might not always be the highest wave in a cardiac cycle, and artifacts may easily upset the search procedure. Observe also that the ECG signal in the present illustration was filtered to a restricted bandwidth of 90 Hz before the derivatives were computed and that it is free of baseline drift.

Solution 2: Murthy and Rangaraj [201] proposed a QRS detection algorithm based upon a weighted and squared first-derivative operator and an MA filter. In this method, a filtered-derivative operator was defined as

$$g_1(n) = \sum_{i=1}^{N} |x(n-i+1) - x(n-i)|^2 (N-i+1), \tag{4.4}$$

where $x(n)$ is the ECG signal, and N is the width of a window within which first-order differences are computed, squared, and weighted by the factor $(N-i+1)$. The weighting factor decreases linearly from the current difference to the difference N samples earlier in time and provides a smoothing effect. Further smoothing of the result was performed by an MA filter over M points to obtain

$$g(n) = \frac{1}{M} \sum_{j=0}^{M-1} g_1(n-j). \tag{4.5}$$

With the sampling rate of 100 Hz, the filter window widths were set as $M = N = 8$ samples. The algorithm provides a single peak for each QRS complex and suppresses P and T waves.

Searching for the peak in a processed signal such as $g(n)$ may be accomplished by a simple peak-searching algorithm as follows:

1. Scan a portion of the signal $g(n)$ that may be expected to contain a peak and determine the maximum value $g_{\max}$, or let $g_{\max}$ be the maximum of $g(n)$ over its entire available duration.

2. Define a threshold as a fraction of the maximum, for example, $Th = 0.5\ g_{\max}$.

3. For all $g(n) > Th$, select those samples for which the corresponding $g(n)$ values are greater than a certain predefined number M of preceding and succeeding samples of $g(n)$, that is,

$$\begin{aligned} \{p\} = \{n \mid \quad & [\quad g(n) > Th\ ] \text{ AND} \\ & [\quad g(n) > g(n-i), i = 1, 2, \ldots, M\ ] \text{ AND} \\ & [\quad g(n) > g(n+i), i = 1, 2, \ldots, M\ ]\}. \end{aligned} \tag{4.6}$$

The set $\{p\}$ defined as above contains the indices of the peaks in $g(n)$.

Additional conditions may be imposed to reject peaks due to artifacts, such as a minimum interval between two adjacent peaks. A more elaborate peak-searching algorithm is described in Section 4.3.2.

Illustration of application: Figure 4.3 illustrates, in the topmost trace, two cycles of a filtered version of the ECG signal shown in Figure 3.5. The signal was filtered with an eighth-order Butterworth lowpass filter with $f_c = 40\ Hz$ and was down-sampled by a factor of ten. The effective sampling rate is $100\ Hz$ to match the parameters used by Murthy and Rangaraj [201]. The signal was normalized by dividing by its maximum value.

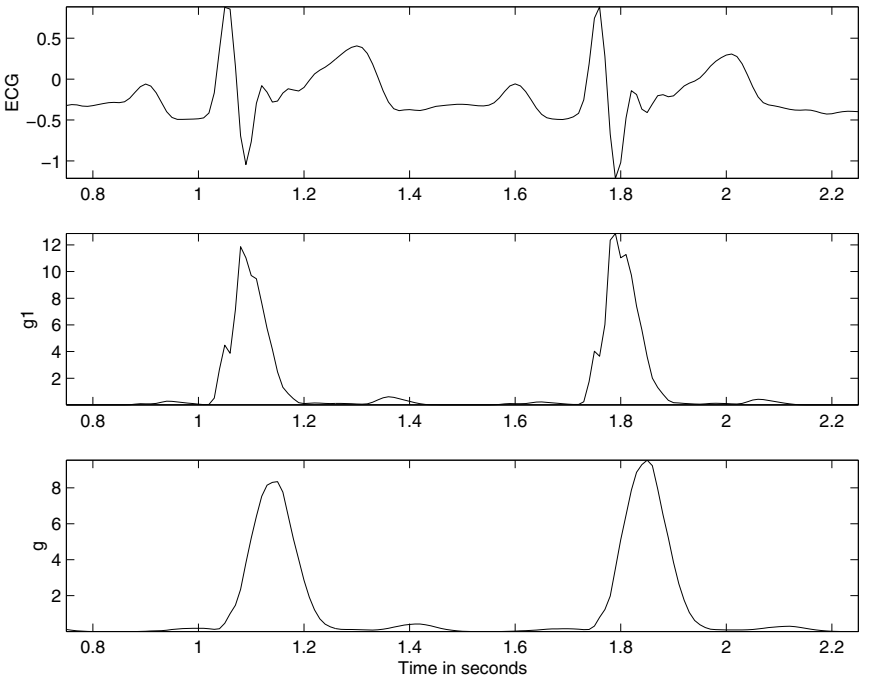

Figure 4.3 From top to bottom: two cycles of a filtered version of the ECG signal shown in Figure 3.5; output $g_1(n)$ of the weighted and squared first-derivative operator in Equation 4.4; output $g(n)$ of the smoothing filter in Equation 4.5.

The second and third plots in Figure 4.3 show the outputs of the derivative-based operator and the smoothing filter. Observe that the final output contains a single, smooth peak for each QRS, and that the P and T waves produce no significant output. A simple peak-searching algorithm may be used to detect and segment each beat [201]. The net delay introduced by the filter used should be subtracted from the detected peak location in order to obtain the corresponding QRS location.

4.3.2 The Pan–Tompkins algorithm for QRS detection

Problem: *Propose an algorithm to detect QRS complexes in an ongoing ECG signal.*

Solution: Pan and Tompkins [202] (see also Tompkins [46]) proposed a real-time QRS detection algorithm based on analysis of the slope, amplitude, and width of QRS complexes. The algorithm includes a series of filters and methods that perform lowpass, highpass, derivative, squaring, integration, adaptive thresholding, and search procedures. Figure 4.4 illustrates the steps of the algorithm in schematic form.

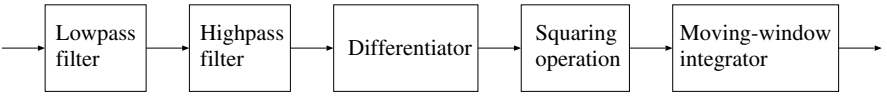

Figure 4.4 Block diagram of the Pan–Tompkins algorithm for QRS detection.

Lowpass filter: The recursive lowpass filter used in the Pan–Tompkins algorithm has integers as its coefficients to reduce computational complexity, with the transfer function defined as

$$H(z) = \frac{1}{32}\,\frac{(1-z^{-6})^2}{(1-z^{-1})^2}\,. \tag{4.7}$$

(See also Equations 3.119 and 3.120.) The output $y(n)$ is related to the input $x(n)$ as

$$y(n) = 2y(n-1) - y(n-2) + \frac{1}{32}\,[x(n) - 2x(n-6) + x(n-12)]. \tag{4.8}$$

With the sampling rate being $200\ Hz$, the filter has a low cutoff frequency of $f_c = 11\ Hz$ and introduces a delay of 5 samples or $25\ ms$. The filter provides an attenuation greater than $35\ dB$ at $60\ Hz$ and effectively suppresses power-line interference, if present.

Highpass filter: The highpass filter used in the algorithm is implemented as an allpass filter minus a lowpass filter. The lowpass component has the transfer function

$$H_{lp}(z) = \frac{(1-z^{-32})}{(1-z^{-1})}\,; \tag{4.9}$$

the input–output relationship is

$$y(n) = y(n-1) + x(n) - x(n-32). \tag{4.10}$$

The transfer function $H_{hp}(z)$ of the highpass filter is specified as

$$H_{hp}(z) = z^{-16} - \frac{1}{32}\,H_{lp}(z)\,. \tag{4.11}$$

Equivalently, the output $p(n)$ of the highpass filter is given by the difference equation

$$p(n) = x(n-16) - \frac{1}{32}\,[y(n-1) + x(n) - x(n-32)], \tag{4.12}$$

with $x(n)$ and $y(n)$ being related as in Equation 4.10. If all of the operations from Equation 4.9 to Equation 4.12 are combined, the input–output relationship of the overall highpass filter is

$$p(n) = p(n-1) - \frac{1}{32}x(n) + x(n-16) - x(n-17) + \frac{1}{32}x(n-32). \quad (4.13)$$

The highpass filter has a cutoff frequency of $5\ Hz$ and introduces a delay of $80\ ms$ for $f_s = 200\ Hz$.

Derivative operator: The derivative operation used by Pan and Tompkins is specified as

$$y(n) = \frac{1}{8}\left[2x(n) + x(n-1) - x(n-3) - 2x(n-4)\right] \quad (4.14)$$

and approximates the ideal $\frac{d}{dt}$ operator up to $30\ Hz$. The derivative procedure suppresses the low-frequency components of the P and T waves, and provides a large gain to the high-frequency components arising from the high slopes of the QRS complex. (See Section 3.5.3 for details on the properties of derivative-based filters.)

Figure 4.5 shows the magnitude of the frequency response for each of the filters in the Pan–Tompkins algorithm for QRS detection described to this point. Observe the response of the highpass filter that was derived from a lowpass filter. The combined procedure has a bandpass nature that is evident from the final response shown in the figure.

Squaring: The squaring operation makes the result positive and emphasizes large differences resulting from QRS complexes; the small differences arising from P and T waves are suppressed. The high-frequency components in the signal related to the QRS complex are further enhanced.

Integration: As observed in Section 4.3.1, the output of a derivative-based operation will exhibit multiple peaks within the duration of a single QRS complex. The Pan–Tompkins algorithm performs smoothing of the output of the preceding operations through a moving-window integration filter as

$$y(n) = \frac{1}{N}\left\{x[n-(N-1)] + x[n-(N-2)] + \cdots + x(n)\right\}. \quad (4.15)$$

The choice of the window width N is to be made with the following considerations: too large a value will result in the outputs due to the QRS and T waves being merged, whereas too small a value could yield multiple peaks for a single QRS. A window width of $N = 30$ samples was found to be suitable for $f_s = 200\ Hz$. Figure 4.6 illustrates the effect of the window width on the output of the integrator and its relationship to the QRS width. (See Section 3.5.2 for details on the properties of MA and integrating filters.)

Adaptive thresholding: The thresholding procedure in the Pan–Tompkins algorithm adapts to changes in the ECG signal by computing running estimates of signal and noise peaks. A peak is said to be detected whenever the final output $y(n)$ changes direction within a specified time interval. In the following discussion, $SPKI$ represents the peak level that the algorithm has learned to correspond to QRS peaks, and $NPKI$ represents the peak level related to non-QRS events (such as noise, EMG,

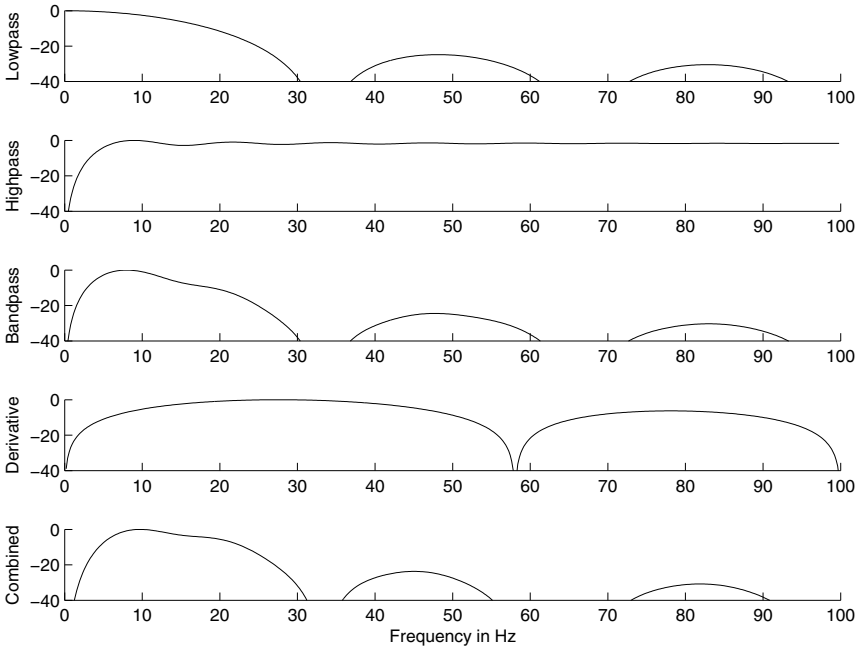

Figure 4.5 Frequency response (magnitude, in dB) of the filters used in the Pan–Tompkins algorithm for QRS detection. Top to bottom: the initial lowpass filter, the highpass filter, the bandpass filter resulting from the combination of the two filters, the derivative operator, and the combination of the bandpass filter and the derivative operator. The bandpass nature of the combined procedure is evident from the final response.

and various artifacts). $THRESHOLD\ I1$ and $THRESHOLD\ I2$ are two thresholds used to categorize peaks detected as signal (QRS) or noise.

Every new peak detected is categorized as a signal peak or a noise peak. If a peak exceeds $THRESHOLD\ I1$ during the first step of analysis, it is classified as a QRS (signal) peak. If the search-back technique (described in the next paragraph) is used, the peak should be above $THRESHOLD\ I2$ to be called a QRS. The peak levels and thresholds are updated after each peak is detected and classified as

$$\begin{aligned} SPKI = &\quad 0.125\,PEAKI + 0.875\,SPKI \quad \text{if } PEAKI \text{ is a signal peak;} \\ NPKI = &\quad 0.125\,PEAKI + 0.875\,NPKI \quad \text{if } PEAKI \text{ is a noise peak;} \end{aligned} \tag{4.16}$$

$$\begin{aligned} THRESHOLD\ I1 &= NPKI + 0.25\,(SPKI - NPKI); \\ THRESHOLD\ I2 &= 0.5\,THRESHOLD\ I1. \end{aligned} \tag{4.17}$$

The updating formula for $SPKI$ is changed to

$$SPKI = 0.25\ PEAKI + 0.75\ SPKI \tag{4.18}$$

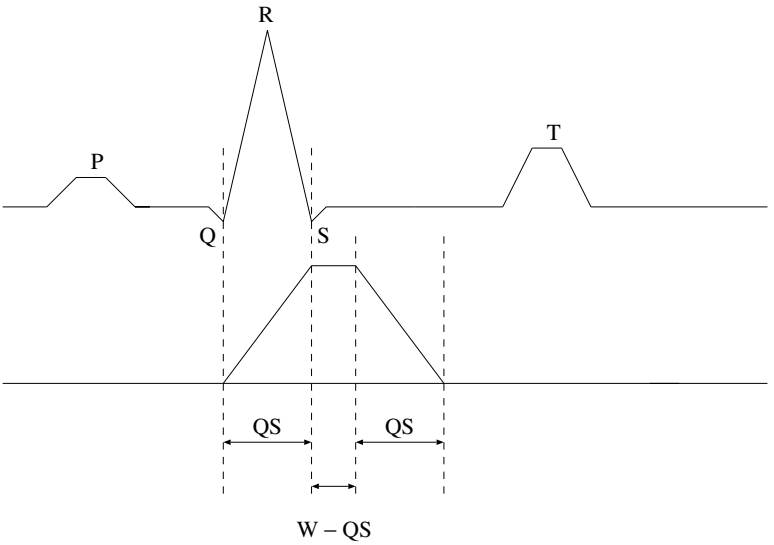

Figure 4.6 The relationship of a QRS complex to the moving-window integrator output. Upper plot: Schematic ECG signal. Lower plot: Schematic representation of the output of the moving-window integrator. QS: QRS complex width. W: width of the integrator window, given as N/f_s s. It is assumed that the QRS complex will provide a result in the form of a rectangular pulse of width QS before the integrator; then, the output of the integrator is a trapezoid as shown, with a total width of W+QS, assuming that W > QS. Adapted from Tompkins [46].

if a QRS is detected in the search-back procedure using $THRESHOLD\ I2$.

Search-back procedure: The Pan–Tompkins algorithm maintains two averages of RR intervals: $RR\ AVERAGE1$ is the average of the eight most-recent beats, and $RR\ AVERAGE2$ is the average of the eight most-recent beats having RR intervals within the range specified by $RR\ LOW\ LIMIT = 0.92 \times RR\ AVERAGE2$ and $RR\ HIGH\ LIMIT = 1.16 \times RR\ AVERAGE2$. Whenever a QRS is not detected for a certain interval specified as $RR\ MISSED\ LIMIT = 1.66 \times RR\ AVERAGE2$, the QRS is taken to be the peak between the established thresholds applied in the search-back procedure.

The Pan–Tompkins algorithm performed with a low error rate of 0.68%, or 33 beats per hour on a database of about $116,000$ beats obtained from 24-hour records of the ECGs of 48 patients (see Tompkins [46] for details).

If N_B QRS complexes or beats are detected in an ECG signal over a duration of T s, we have

$$HR = 60\,\frac{N_B}{T} \tag{4.19}$$

in bpm (on the average). If the average RR interval measured over several beats is RR_a s, we have

$$HR = \frac{60}{RR_a} \tag{4.20}$$

in bpm (on the average). The equation given above may also be used with the beat-to-beat RR interval instead of RR_a to obtain the instantaneous heart rate. In a practical ECG monitoring system, it would be desirable to derive the average heart rate and update its display a few times per minute, such as every 10 s.

Illustration of application: Figure 4.7 illustrates, in the topmost trace, the same ECG signal as in Figure 4.2. The Pan–Tompkins algorithm as above was implemented. The outputs of the various stages of the algorithm are illustrated in sequence in the same figure. The observations to be made are similar to those in the preceding section on the derivative-based method. The derivative operator suppresses the P and T waves and provides a large output at the QRS locations. The squaring operation preferentially enhances large values and boosts high-frequency components. The result still possesses multiple peaks for each QRS and hence needs to be smoothed. The final output of the integrator is a single smooth pulse for each QRS. Observe the shift between the actual QRS location and the pulse output due to the cumulative delays of the various filters. The thresholding and search procedures and their results are not illustrated. More examples of QRS detection are presented in Sections 4.9 and 4.10.

4.3.3 Detection of the dicrotic notch

Problem: *Propose a method to detect the dicrotic notch in the carotid pulse signal.*

Solution: Lehner and Rangayyan [131] proposed a method for detection of the dicrotic notch that used the least-squares estimate of the second derivative $p(n)$ of the carotid pulse signal $y(n)$, defined as

$$p(n) = 2y(n-2) - y(n-1) - 2y(n) - y(n+1) + 2y(n+2). \tag{4.21}$$

Observe that this expression is noncausal; it may be made causal by adding a delay of two samples.

The second derivative was used due to the fact that the dicrotic notch appears as a short wave riding on the downward slope of the carotid pulse signal (see also Starmer et al. [203]). A first-derivative operation would give an almost-constant output for the downward slope. The second-derivative operation removes the effect of the downward slope and enhances the notch itself. The result was squared and smoothed to obtain

$$s(n) = \sum_{k=1}^{M} p^2(n-k+1)w(k), \tag{4.22}$$

where $w(k) = (M-k+1)$, $k = 1, 2, \ldots, M$, is a linear weighting function, and $M = 16$ samples for $f_s = 256$ Hz.

The method yields two peaks for each period of the carotid pulse signal. The first peak in the result represents the onset of the carotid upstroke. The second peak that

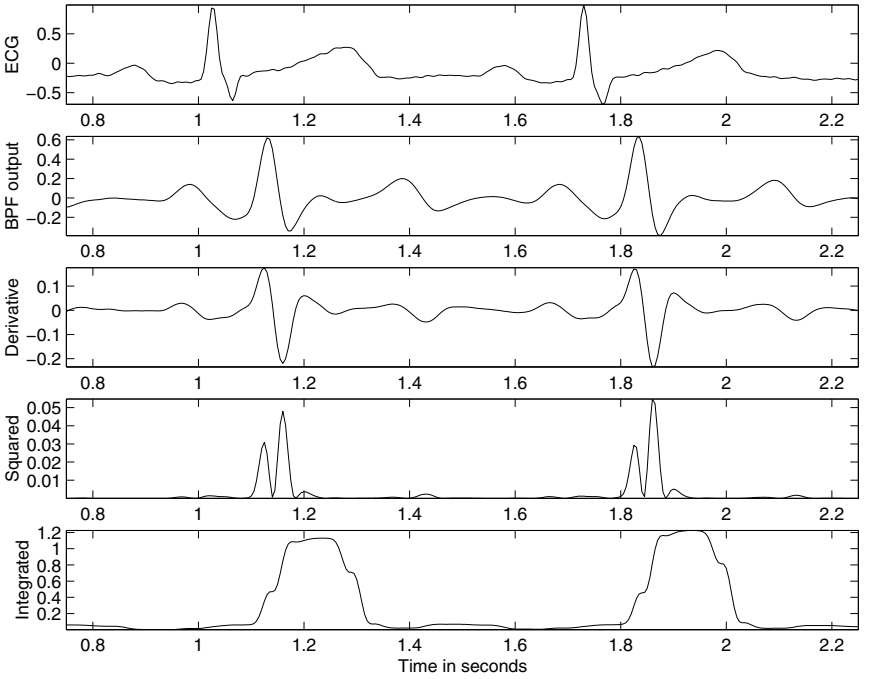

Figure 4.7 Results of the Pan–Tompkins algorithm. From top to bottom: two cycles of a filtered version of the ECG signal shown in Figure 3.5 (the same as that in Figure 4.2); output of the bandpass filter (BPF, a combination of lowpass and highpass filters); output of the derivative-based operator; the result of squaring; and $100\times$ the result of the final integrator.

appears in the result within a cardiac cycle is related to the dicrotic notch. To locate the dicrotic notch, the local minimum in the carotid pulse within a $\pm 20\ ms$ interval of the second peak was used.

Illustration of application: The upper plot in Figure 4.8 illustrates two cycles of a carotid pulse signal. The signal was lowpass filtered at $100\ Hz$ and sampled at $250\ Hz$. The result of application of the Lehner and Rangayyan method to the signal is shown in the lower plot. It is evident that the second derivative has successfully accentuated the dicrotic notch. A simple peak-searching algorithm may be used to detect the first and second peaks in the result. The dicrotic notch may then be located by searching for the minimum in the carotid pulse signal within a $\pm 20\ ms$ interval around the second peak location.

Observe that the result illustrated in Figure 4.8 may benefit from further smoothing by increasing the window width M in Equation 4.22. The window width needs to be chosen in accordance with the characteristics of the signal on hand as well as the lowpass filter and sampling rate used. Further illustration of the detection of the dicrotic notch is provided in Section 4.10.

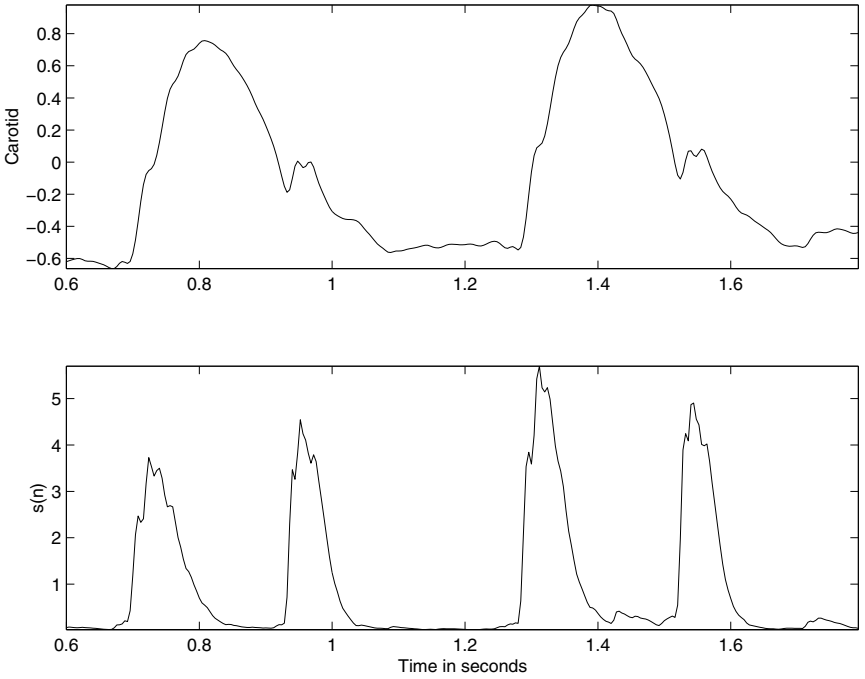

Figure 4.8 Two cycles of a carotid pulse signal and the result of the Lehner and Rangayyan method for detection of the dicrotic notch.

4.4 Correlation Analysis of EEG Rhythms

EEG signals are usually acquired simultaneously over multiple channels. Event detection in and epoch analysis of EEG signals becomes more complicated than the problems we have seen with the single-channel ECG and carotid pulse signals, due to the need to detect similar or related events across multiple channels. Autocorrelation and cross-correlation techniques in both the time and frequency domains serve such needs.

4.4.1 Detection of EEG rhythms

Problem: *Propose a method to detect the presence of the α rhythm in an EEG channel. How would you extend the method to detect the presence of the same rhythm simultaneously in two EEG channels?*

Solution: Two signals may be compared to detect the characteristics present in common between them via their dot product (also known as the inner or scalar prod-

uct), defined as

$$x \cdot y = \langle x, y \rangle = \sum_{n=0}^{N-1} x(n)\, y(n), \tag{4.23}$$

where the signals $x(n)$ and $y(n)$ have N samples each. The dot product represents the projection of one signal on to the other, with each signal being viewed as an N-dimensional vector. The dot product may be normalized by the geometric mean of the energies of the two signals to obtain a correlation coefficient as [132]

$$\gamma_{xy} = \frac{\sum_{n=0}^{N-1} x(n)\, y(n)}{\left[\sum_{n=0}^{N-1} x^2(n)\ \sum_{n=0}^{N-1} y^2(n)\right]^{1/2}}. \tag{4.24}$$

The means of the signals may be subtracted out, if desired, as in Equation 3.96.

In the case of two continuous-time signals $x(t)$ and $y(t)$, the projection of one signal on to the other is defined as

$$\theta_{xy} = \int_{-\infty}^{\infty} x(t)\, y(t)\, dt. \tag{4.25}$$

When a shift or time delay may be present in the occurrence of the epoch of interest in the two signals being compared, it becomes necessary to introduce a time-shift parameter to compute the projection for every possible position of overlap. The shift parameter facilitates searching one signal for the occurrence of an event matching that in the other signal at any time instant within the available duration of the signals. The CCF between two signals for a shift or delay of τ seconds or k samples may be obtained as

$$\theta_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)\, y(t+\tau)\, dt, \tag{4.26}$$

or

$$\theta_{xy}(k) = \sum_{n} x(n)\, y(n+k). \tag{4.27}$$

The range of summation in the latter case needs to be limited to the range of the available overlapped data. A scale factor, depending upon the number of data samples used, needs to be introduced to obtain the true CCF, but is neglected here (see Section 6.3). An extended version of the correlation coefficient γ_{xy} in Equation 4.24, to include time shift, is provided in Equation 3.96.

When the ACF or the CCF is computed for various shifts, a question arises about the data samples in one of the signal segments beyond the duration of the other. We may add zeros to one of the signals and increase its length by the maximum shift of interest, or we may use the true data samples from the original signal record, if available. The latter method was used wherever possible in the following illustrations.

In the case of random signals, we need to take the expectation or sample average of the outer product of the vectors formed by the available samples of the signals. Let $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T$ and $\mathbf{y}(n) = [y(n), y(n-1), \ldots, y(n-N+1)]^T$ represent the N-dimensional vectorial form of the two signals $x(n)$ and

$y(n)$ with the most-recent N samples being available in each signal at the time instant n. If $\mathbf{x}(n)$ and $\mathbf{y}(n)$ are sample observations of random processes, their CCF is defined as

$$\mathbf{\Theta}_{xy} = E[\mathbf{x}(n)\,\mathbf{y}^T(n)], \tag{4.28}$$

in a manner similar to Equations 3.162 and 3.163. The outer product, which is an $N \times N$ matrix, provides the cross-terms that include all possible delays (shifts) within the duration of the given signals.

All of the equations given above may be modified to obtain the ACF by replacing the second signal y with the first signal x. The signal x is then compared with itself.

The ACF displays peaks at intervals corresponding to the period (and integral multiples thereof) of any periodic or repetitive pattern present in the signal. This property facilitates the detection of rhythms in signals such as the EEG: The presence of the α rhythm would be indicated by a peak in the neighborhood of $0.1\ s$. The ACF of most signals decays and reaches negligible values after delays of a few milliseconds, except for periodic signals of infinite or indefinite duration for which the ACF will also exhibit periodic peaks. The ACF will also exhibit multiple peaks when the same event repeats itself at regular or irregular intervals. One may need to compute the ACF only up to certain delay limits depending upon the expected characteristics of the signal being analyzed.

The CCF displays peaks at the period of any periodic pattern present in *both* of the signals being analyzed. The CCF may, therefore, be used to detect rhythms present in common between between two signals, for example, between two channels of the EEG. When one of the functions being used to compute the CCF is a template representing an event, such as an ECG cycle as in the illustration in Section 3.5.1 or an EEG spike-and-wave complex as in Section 4.4.2, the procedure is known as *template matching.*

Illustration of application: Figure 4.9 shows, in the upper trace, the ACF of a segment of the p4 channel of the EEG in Figure 1.39 over the time interval $4.67 - 5.81\ s$. The ACF displays peaks at time delays of $0.11\ s$ and its integral multiples. The inverse of the delay of the first peak corresponds to $9\ Hz$, which is within the α rhythm range. (The PSD in the lower trace of Figure 4.9 and the others to follow are described in Section 4.5.) It is, therefore, obvious that the signal segment analyzed contains the α rhythm. A simple peak-searching algorithm may be applied to the ACF to detect the presence of peaks at specific delays of interest or over the entire range of the ACF.

To contrast with the preceding example, the upper trace of Figure 4.10 shows the ACF of the $4.2 - 4.96\ s$ segment of the f3 channel of the EEG in Figure 1.39. The ACF shows no peak in the $0.08 - 1.25\ s$ region, indicating absence of the α rhythm in the segment analyzed.

Figures 4.11, 4.12, and 4.13 illustrate the CCF results comparing the following portions of the EEG signal shown in Figure 1.39 in order: the p3 and p4 channels over the duration $4.72 - 5.71\ s$ when both channels exhibit the α rhythm; the o2 and c4 channels over the duration $5.71 - 6.78\ s$ when the former has the α rhythm but not the latter channel; and the f3 and f4 channels over the duration $4.13 - 4.96\ s$ when

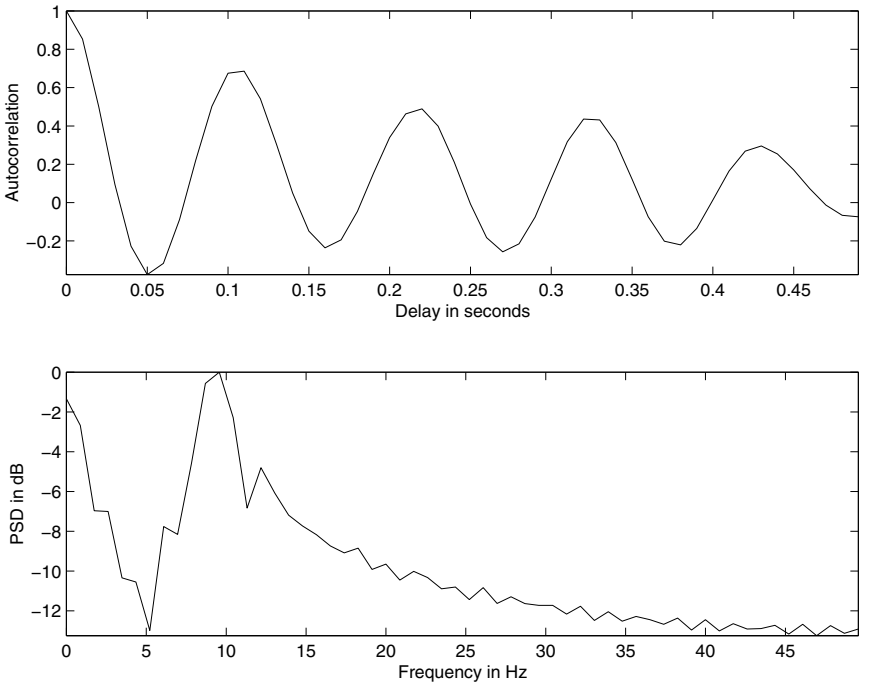

Figure 4.9 Upper trace: ACF of the $4.67 - 5.81$ s portion of the p4 channel of the EEG signal shown in Figure 1.39. Lower trace: The PSD of the signal segment in dB, given by the Fourier transform of the ACF.

neither channel has α activity. The relative strengths of the peaks in the α range, as described earlier, agree with the joint presence, singular presence, or absence of the α rhythm in the various segments (channels) analyzed.

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.

4.4.2 Template matching for EEG spike and wave detection

We have seen the use of template matching for the extraction of ECG cycles for use in synchronized averaging in Section 3.5.1. We shall now consider another application of template matching.

Problem: *Propose a method to detect spike-and-wave complexes in an EEG signal. You may assume that a sample segment of a spike-and-wave complex is available.*

Solution: A spike-and-wave complex is a well-defined event in an EEG signal. The complex is composed of a sharp spike followed by a wave with a frequency of about 3 Hz; the wave may contain a half period or a full period of an almost-

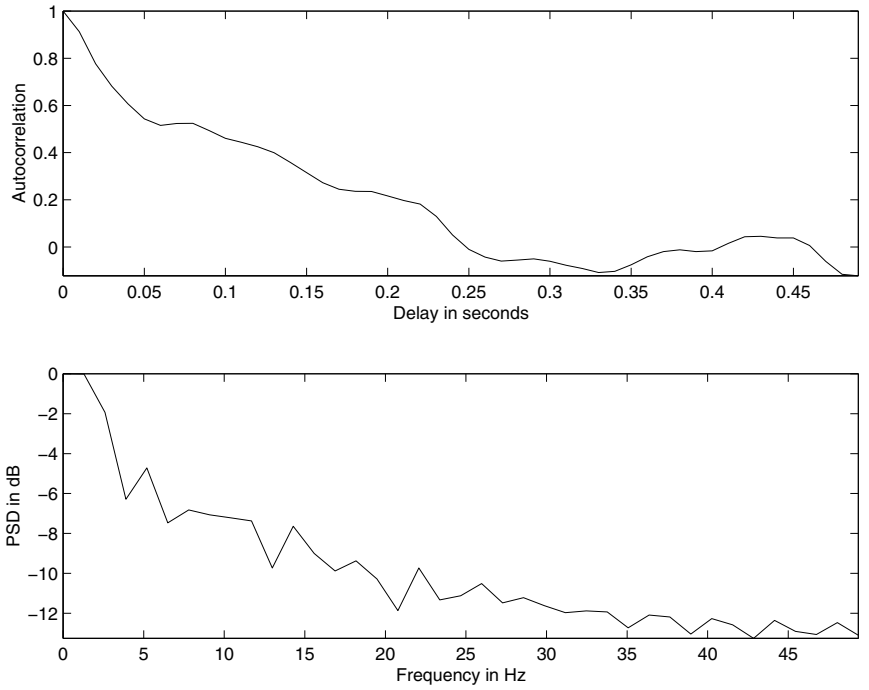

Figure 4.10 Upper trace: ACF of the $4.2 - 4.96$ s portion of the f3 channel of the EEG signal shown in Figure 1.39. Lower trace: The PSD of the signal segment in dB.

sinusoidal pattern. One may, therefore, extract an epoch of a spike-and-wave complex from an EEG channel and use it for template matching with the same formula as in Equation 3.96 (see also Barlow [196]). The template may be correlated with the same channel from which it was extracted to detect similar events that appear at a later time, or with another channel to search for similar events. A simple threshold on the result should yield the time instants where the events appear.

Illustration of application: The c3 channel of the EEG signal in Figure 1.40 is shown in the upper trace of Figure 4.14. The spike-and-wave complex between 0.60 s and 0.82 s in the signal was selected for use as the template, and template matching was performed with the same channel signal using the formula in Equation 3.96. The result in the lower trace of Figure 4.14 demonstrates strong and clear peaks at each occurrence of the spike-and-wave complex in the EEG signal. The peaks in the result occur at the same instants of time as the corresponding spike-and-wave complexes.

Figure 4.15 shows the f3 channel of the EEG signal in Figure 1.40, along with the result of template matching, using the same template that was used in the previous example from channel c3. The result shows that the f3 channel also has spike-and-wave complexes that match the template.

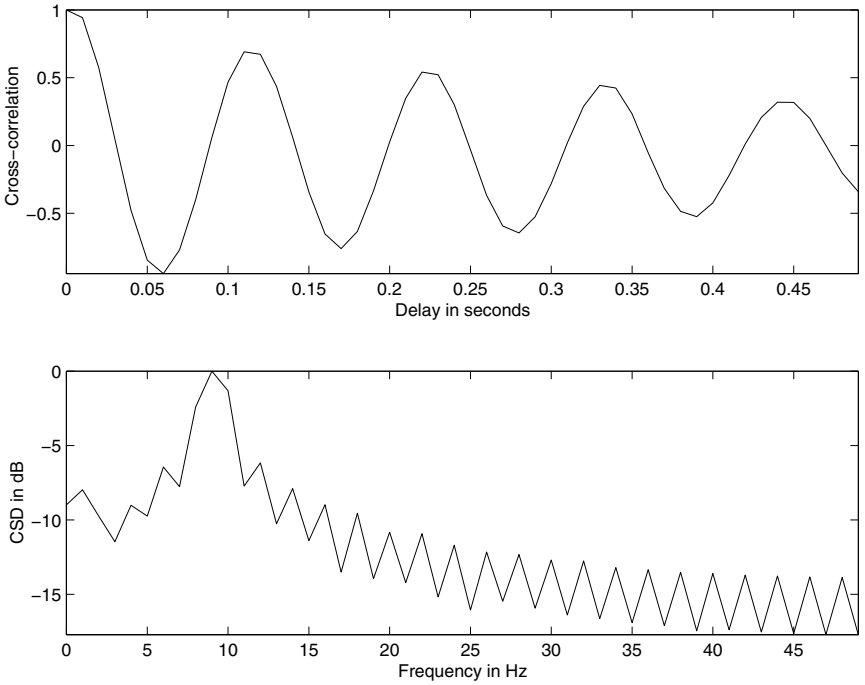

Figure 4.11 Upper trace: CCF between the $4.72 - 5.71$ s portions of the p3 and p4 channels of the EEG signal shown in Figure 1.39. Lower trace: The CSD of the signal segments in dB, computed as the Fourier transform of the CCF.

4.4.3 Detection of EEG rhythms related to seizure

Yadav et al. [59] noted that the recurring nature of seizures of a given patient could be used to design patient-specific templates for the recognition of seizures. They noted that, in most patients, a few types of seizures tend to occur repeatedly, with related similar, though not identical, patterns in the EEG. Figure 4.16 illustrates the temporal evolution of a seizure event on a single-channel EEG, including varying patterns of piecewise stationary rhythms. (Note that the seizure rhythms are substantially different from the α and other rhythms seen in the normal EEG.) Yadav et al. [59] suggested that it is possible to train a patient-specific seizure detector using previously identified templates, and they proposed a model-based patient-specific method using statistically optimal null filters for automatic detection of seizures in intracranial EEG. See Qu and Gotman [204] and Grewal and Gotman [60] for additional details on detection of seizures.

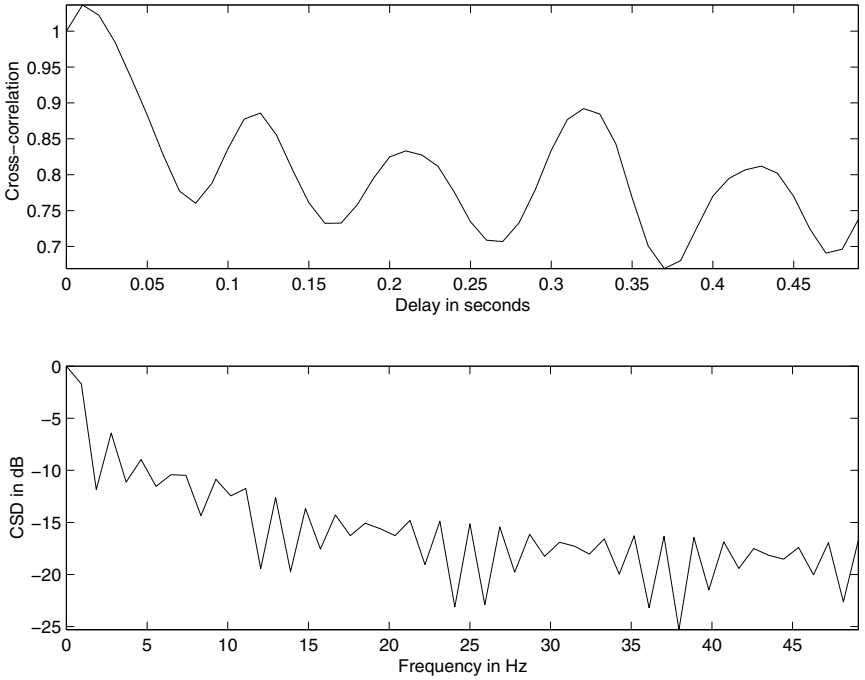

Figure 4.12 Upper trace: CCF between the $5.71 - 6.78$ s portions of the o2 and c4 channels of the EEG signal shown in Figure 1.39. Lower trace: The CSD of the signal segments in dB.

4.5 Cross spectral Techniques

The multiple peaks that arise in the ACF or CCF may cause confusion in the detection of rhythms; the analyst may be required to discount peaks that appear at integral multiples of the delay corresponding to a fundamental frequency and other delays. The Fourier-domain equivalents of the ACF or the CCF permit easier and more intuitive analysis in the frequency domain than in the time domain. The notion of rhythms would be easier to associate with frequencies in cps or Hz than with the corresponding inversely related periods (see also the introductory section of Chapter 6).

4.5.1 Coherence analysis of EEG channels

Problem: *Describe a frequency-domain approach to study the presence of rhythms in multiple channels of an EEG signal.*

Solution: The Fourier-domain equivalents of the ACF and CCF are the PSD (also known as the autospectrum) and the cross-spectrum (or cross-spectral density — CSD), respectively. The PSD $S_{xx}(f)$ of a signal is related to its ACF via the Fourier

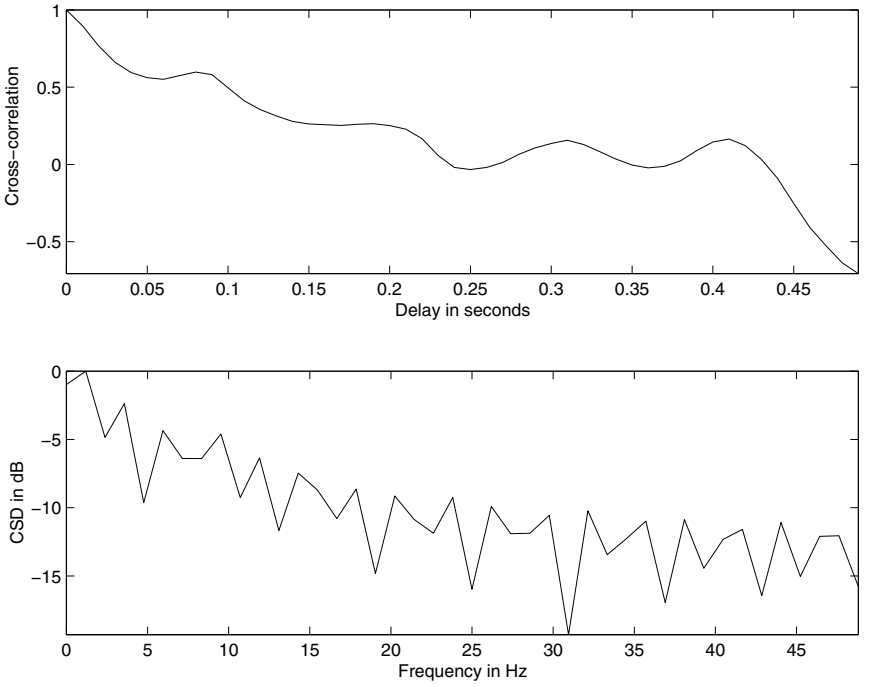

Figure 4.13 Upper trace: CCF between the $4.13 - 4.96$ s portions of the f3 and f4 channels of the EEG signal shown in Figure 1.39. Lower trace: The CSD of the signal segments in dB.

transform:

$$S_{xx}(f) = FT[\phi_{xx}(\tau)] = X(f)X^*(f) = |X(f)|^2, \tag{4.29}$$

where $FT[\ ]$ indicates the Fourier transform. The Fourier transform of the CCF between two signals gives the CSD:

$$S_{xy}(f) = FT[\theta_{xy}(\tau)] = X(f)Y^*(f). \tag{4.30}$$

(For the sake of simplicity, the double-symbol subscripts xx and yy may be replaced by their singular versions, or dropped entirely when not relevant in subsequent discussions.)

The PSD displays peaks at frequencies corresponding to periodic activities in the signal. This property facilitates the detection of rhythms in signals such as the EEG: The presence of the α rhythm would be indicated by a peak or multiple peaks in the neighborhood of $8 - 13$ Hz. The PSD may also be studied to locate the presence of activity spread over specific bands of frequencies, such as formants in the speech signal or murmurs in the PCG.

The CSD exhibits peaks at frequencies that are present in both of the signals being compared. The CSD may be used to detect rhythms present in common between two channels of the EEG.

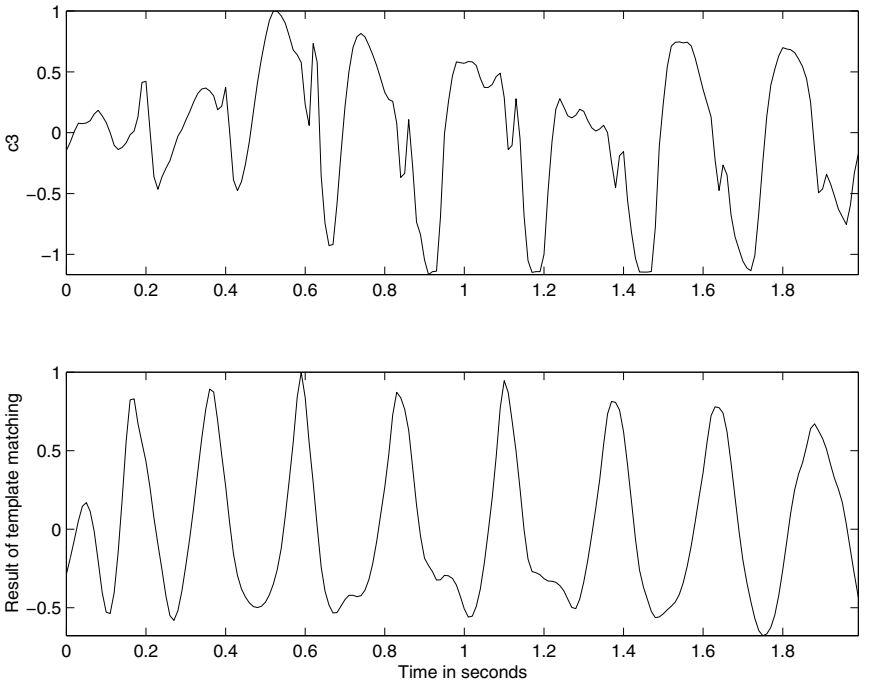

Figure 4.14 Upper trace: the c3 channel of the EEG signal shown in Figure 1.40. Lower trace: result of template matching. The spike-and-wave complex between $0.60\ s$ and $0.82\ s$ in the signal was used as the template.

The normalized magnitude of the *coherence spectrum* of two signals is given by [6, 54]

$$\Gamma_{xy}(f) = \left[\frac{|S_{xy}(f)|^2}{S_{xx}(f) S_{yy}(f)} \right]^{1/2}. \tag{4.31}$$

If this expression is computed for two individual signals directly, the magnitude of the result will be equal to unity for all f, which is incorrect. In order to evaluate the magnitude coherence spectrum as defined in Equation 4.31, each spectral density function involved — S_{xy}, S_{xx}, and S_{yy} — must be estimated using an averaging procedure applied to several observations of the processes generating the signals [6, 205, 206]. See Section 6.3 for procedures for the estimation of spectral density functions.

The phase of the coherence spectrum is given by $\psi_{xy}(f) = \angle S_{xy}(f)$, which represents the average phase difference (related to the time delay) between frequency components in the two signals.

Rosenberg et al. [207] applied coherence to identify functional coupling between neuronal spike trains. They observed that frequency-domain analysis using coher-

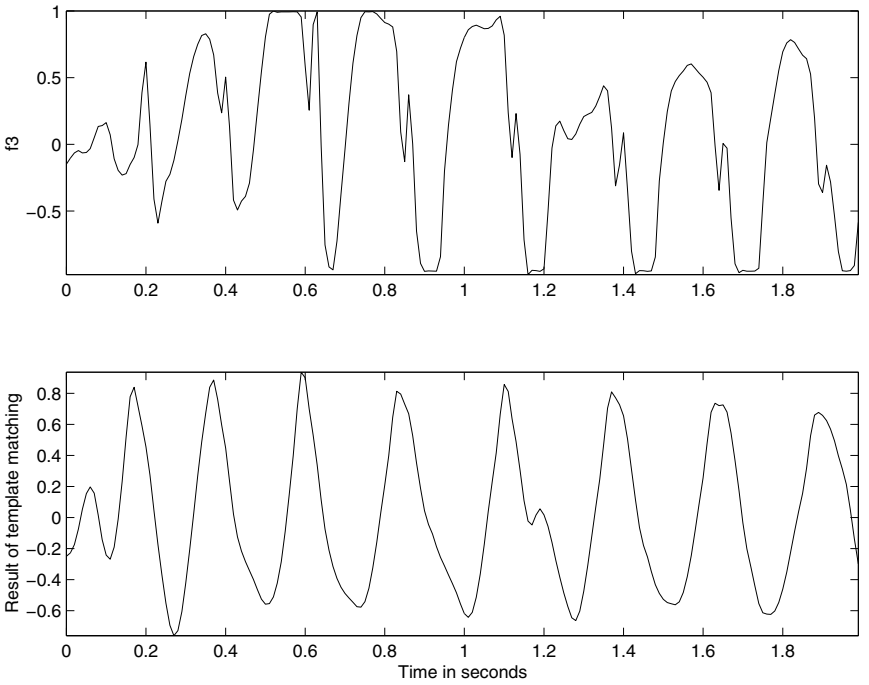

Figure 4.15 Upper trace: the f3 channel of the EEG signal shown in Figure 1.40. Lower trace: result of template matching. The spike-and-wave complex between $0.60\ s$ and $0.82\ s$ in the c3 channel (see Figure 4.14) was used as the template.

ence as a measure of association can lead to new insight and understanding of interactions between spike trains. Amjad et al. [208] described several methods to estimate the coherence function and its application to the analysis of SMUAP discharges and physiological tremor; see also Farmer et al. [209] and Farmer [210]. Johnston et al. [211] used coherence analysis to study the correlation between neural inputs to motor units innervating different muscles. See Section 6.3 for procedures for the estimation of spectral density functions.

Illustration of application: The coherence between EEG signals recorded from different positions on the scalp depends upon the structural connectivity or functional coupling between the corresponding parts of the brain. Investigations into the neurophysiology of seizure discharges and behavior attributable to disorganization of cerebral function may be facilitated by coherence analysis [54]. The symmetry, or lack thereof, between two EEG channels on the left and right sides of the same position (for example, c3 and c4) may be analyzed via the CSD or the coherence function.

The lower traces in Figures 4.9 and 4.10 illustrate the PSDs of EEG segments with and without the α rhythm, respectively. The former shows a strong and clear peak at about $9\ Hz$, indicating the presence of the α rhythm. Observe that the

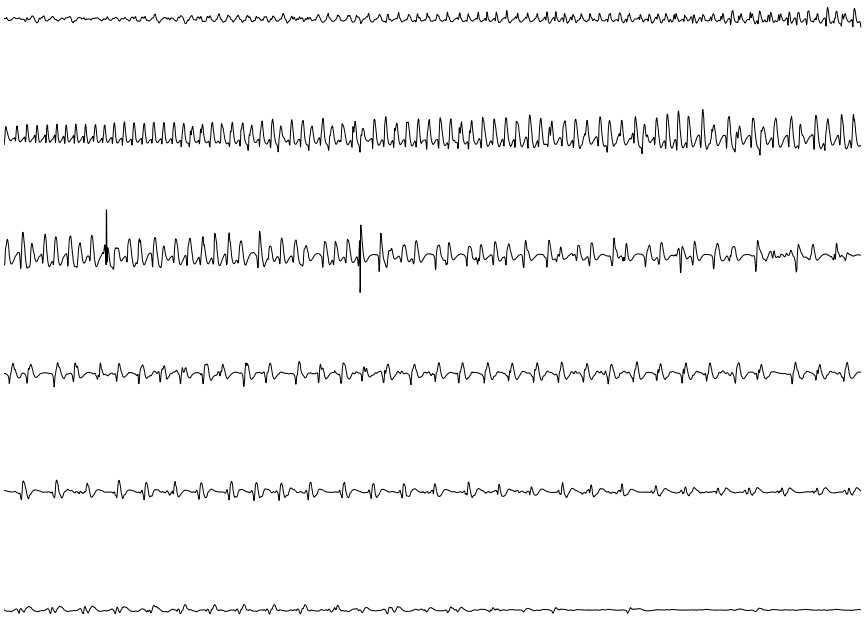

Figure 4.16 The evolution of various rhythms related to a seizure event as seen in an intracranial EEG signal. Each row represents a duration of $15\ s$ and is a continuation of the preceding row. The minimum to maximum range of the signal shown is about $120\ \mu V$. Data courtesy of R. Agarwal [59].

PSD displays a single peak although the corresponding ACF has multiple peaks at two, three, and four times the delay corresponding to the fundamental period of the α wave in the signal. The PSD in Figure 4.10 exhibits no peak in the α range, indicating the absence of the α rhythm in the signal.

The lower traces in Figures 4.11, 4.12, and 4.13 illustrate the CSDs corresponding to the CCFs in the respective upper traces. Once again, it is easier to deduce the common presence of strong α activity between channels p3 and p4 from the CSD rather than the CCF in Figure 4.11. The single peak at $9\ Hz$ in the CSD is more easily interpretable than the multiple peaks in the corresponding CCF. The CSD in Figure 4.12 lacks a clear peak in the α range, even though the corresponding CCF shows a peak at about $0.1\ s$, albeit less significant than that in Figure 4.11. The results agree with the fact that one channel has α activity while the other does not. Finally, the CSD in Figure 4.13 is clearly lacking a peak in the α range; the two signal segments have no α activity. Further methods for the analysis of α activity are presented in Sections 6.3.3, 6.7, and 7.5.2.

4.6 The Matched Filter

When a sample observation or template of a typical version of a signal event is available, it becomes possible to design a filter that is *matched* to the characteristics of the event and maximizes the SNR of the output. If a signal that contains repetitions of the event with almost the same characteristics is passed through the *matched filter*, the output should provide peaks at the time instants of occurrence of the event. Matched filters are commonly used for the detection of signals of known characteristics that are buried in noise [212, 213]. They are designed to perform a correlation operation between the input signal and the signal template, and hence are also known as *correlation filters*.

4.6.1 Derivation of the transfer function of the matched filter

In order to derive the transfer function, $H(\omega)$, of the matched filter [212], let the signal $x(t)$ be the input to the matched filter. The Fourier transform of $x(t)$ is

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\ \exp(-j\omega t)\ dt. \tag{4.32}$$

The output of the matched filter, $y(t)$, is given by the inverse Fourier transform of $Y(\omega) = X(\omega)H(\omega)$, as follows:

$$\begin{aligned} y(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\ H(\omega)\ \exp(+j\omega t)\ d\omega \\ &= \int_{-\infty}^{\infty} X(f)\ H(f)\ \exp(+j\,2\pi\, ft)\ df. \end{aligned} \tag{4.33}$$

In the second expression of the equation given above, the frequency variable has been changed from ω in radians per second to f in Hz.

Consider the presence of white noise at the input, with the PSD

$$S_{\eta i}(f) = \frac{P_{\eta i}}{2}, \tag{4.34}$$

where $P_{\eta i}$ is the average noise power at the input. Then, the noise PSD at the output is

$$S_{\eta o}(f) = \frac{P_{\eta i}}{2}\ |H(f)|^2. \tag{4.35}$$

The average output noise power is

$$P_{\eta o} = \frac{P_{\eta i}}{2}\ \int_{-\infty}^{\infty} |H(f)|^2\, df. \tag{4.36}$$

The RMS value of the noise in the absence of any signal is $\sqrt{P_{\eta o}}$.

Letting $t = t_0$ in Equation 4.33, the magnitude of the instantaneous output signal at $t = t_0$ is

$$M_y = |y(t_0)| = \left| \int_{-\infty}^{\infty} X(f)\ H(f)\ \exp(+j\,2\pi\,f t_0)\ df \right|. \tag{4.37}$$

Thus, the SNR at the output is $\frac{M_y}{\sqrt{P_{\eta o}}}$.

To derive the optimal transfer function of the matched filter, we could maximize the SNR, which is equivalent to maximizing the expression

$$\frac{M_y^2}{P_{\eta o}} = \frac{\text{instantaneous peak power of signal}}{\text{noise mean power}}, \tag{4.38}$$

which represents peak-power SNR [212].

For a given signal $x(t)$, the total energy is a constant, given by

$$E_x = \int_{-\infty}^{\infty} x^2(t)\ dt = \int_{-\infty}^{\infty} |X(f)|^2\ df. \tag{4.39}$$

Let us consider the following ratio:

$$\frac{M_y^2}{E_x\ P_{\eta o}} = \frac{\left| \int_{-\infty}^{\infty}\ H(f)\ X(f)\ \exp(+j\,2\pi\,f t_0)\ df \right|^2}{\frac{P_{\eta i}}{2}\ \int_{-\infty}^{\infty}\ |H(f)|^2\ df\ \int_{-\infty}^{\infty}\ |X(f)|^2\ df}. \tag{4.40}$$

The quantity E_x is a constant for a given input signal; hence, maximizing the expression in Equation 4.40 is equivalent to maximizing the expression in Equation 4.38.

In order to determine the condition for maximization of the expression in Equation 4.40, recall Schwarz's inequality for two arbitrary complex functions $A(f)$ and $B(f)$:

$$\left| \int_{-\infty}^{\infty} A(f)\ B(f)\ df \right|^2 \ \leq\ \left[\int_{-\infty}^{\infty} |A(f)|^2\ df \right]\ \left[\int_{-\infty}^{\infty} |B(f)|^2\ df \right]. \tag{4.41}$$

For any two real functions $a(t)$ and $b(t)$, the corresponding inequality is

$$\left[\int_{-\infty}^{\infty} a(t)\ b(t)\ dt \right]^2 \ \leq\ \left[\int_{-\infty}^{\infty} a^2(t)\ dt \right]\ \left[\int_{-\infty}^{\infty} b^2(t)\ dt \right]. \tag{4.42}$$

For any two vectors $\mathbf{a}$ and $\mathbf{b}$, Schwarz's inequality states that

$$|\mathbf{a}.\mathbf{b}| \leq |\mathbf{a}|\,|\mathbf{b}|\,, \tag{4.43}$$

and

$$|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|\,. \tag{4.44}$$

In the inequalities stated above, equality is achieved if $\mathbf{a} = K\,\mathbf{b}$, that is, $\mathbf{a}$ and $\mathbf{b}$ are collinear; if $a(t) = K\,b(t)$; or if $A(f) = K\,B^*(f)$, where K is a real constant.

The inequality in Equation 4.41 can be applied to Equation 4.40 by considering $A(f) = H(f)$ and $B(f) = X(f)\ \exp(+j\,2\pi\,ft_0)$. Then, we have

$$\frac{P_{\eta i}\ M_y^2}{2\ E_x\ P_{\eta o}} \leq 1, \tag{4.45}$$

because

$$\left| \int_{-\infty}^{\infty} X(f)\ H(f)\ \exp(+j\,2\pi\,ft_0)\ df \right|^2 \leq \int_{-\infty}^{\infty} |X(f)|^2\ df\ \int_{-\infty}^{\infty} |H(f)|^2\ df\,. \tag{4.46}$$

The LHS of Equation 4.46 represents the instantaneous output power, M_y^2, evaluated in the Fourier domain. Thus, the ratio in Equation 4.40 is maximized when equality is achieved, that is, $A(f) = K\,B^*(f)$ with the functions $A(f)$ and $B(f)$ as explained above, leading to the condition

$$\begin{aligned} H(f) &= K\ [X(f)\ \exp(+j\,2\pi\,ft_0)]^* \\ &= K\ X^*(f)\ \exp(-j\,2\pi\,ft_0), \end{aligned} \tag{4.47}$$

which leads to maximal peak output SNR.

Taking the inverse Fourier transform of the last expression given above, we have

$$h(t) = K\ x[-(t - t_0)]. \tag{4.48}$$

Therefore, the impulse response of the matched filter is a scaled, reversed, and shifted version of the signal of interest.

The output of the matched filter is then obtained by the following steps:

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} X(f)\ H(f)\ \exp(+j\,2\pi\,ft)\ df \\ &= \int_{-\infty}^{\infty} X(f)\ K\ X^*(f)\ \exp(-j\,2\pi\,ft_0)\ \exp(+j\,2\pi\,ft)\ df \\ &= K \int_{-\infty}^{\infty} |X(f)|^2\ \exp[+j\,2\pi\,f(t - t_0)]\ df \\ &= K\ \phi_x(t - t_0). \end{aligned} \tag{4.49}$$

The last relationship listed above is based on the property that the Fourier transform of the ACF is equal to the PSD of the signal. [*Note*: The Fourier transform of $x(-t)$ is $X^*(f)$, and that of $x(t - \tau)$ is $\exp(-j\,2\pi\,f\,\tau)\,X(f)$.]

It is seen from the derivations given above that, when the impulse response of the matched filter is related to the desired signal as in Equation 4.48, the output is maximized and is equal to a scaled and delayed version of the ACF of the signal.

Illustration of application: Figure 4.17 shows a signal $x(n)$ composed of three events. The signal may be defined in terms of impulses as

$$\begin{aligned} x(n) &= 3\,\delta(n-5) + 2\,\delta(n-6) + \delta(n-7) \\ &+ 1.5\,\delta(n-16) + \delta(n-17) + 0.5\,\delta(n-18) \\ &+ 0.75\,\delta(n-26) + 0.5\,\delta(n-27) + 0.25\,\delta(n-28). \end{aligned} \tag{4.50}$$

From the plot of the signal in Figure 4.17, it is seen that $x(n)$ is composed of three occurrences of a basic signal $g(n)$ that may be defined as

$$g(n) = 3\,\delta(n) + 2\,\delta(n-1) + \delta(n-2), \tag{4.51}$$

with the composite signal given by

$$x(n) = g(n-5) + 0.5\,g(n-16) + 0.25\,g(n-26). \tag{4.52}$$

If we consider $g(n)$ as the basic pattern to be detected by using a matched filter, we need the impulse response of the filter to be $h(n) = K\,g(-n+n_0)$, where K is a scale factor or gain and n_0 is a delay to obtain a causal $h(n)$. In the present case, we need $n_0 = 2$ samples. Let $K = 1$. Then, we have

$$h(n) = \delta(n) + 2\,\delta(n-1) + 3\,\delta(n-2), \tag{4.53}$$

which is plotted in the trace in the middle of Figure 4.17. The trace at the bottom in the same figure shows the output of the matched filter. It is evident that the output possesses peaks at the locations of occurrences of $g(n)$ in the input signal, with a delay of 2 samples introduced by the filtering process. The values of the peaks are in proportion to the amplitude scale factors of the repeated occurrences of the basic pattern. It can be verified that the output values for the samples $5-9$ represent the ACF of $g(n)$.

4.6.2 Detection of EEG spike and wave complexes

Problem: *Design a matched filter to detect spike-and-wave complexes in an EEG signal. A reference spike-and-wave complex is available.*

Solution: Let $x(t)$ be the given reference signal, representing an ideal observation of the event of interest. Let $X(f)$ be the Fourier transform of $x(t)$. Consider passing $x(t)$ through an LSI filter whose impulse response is $h(t)$; the transfer function of the filter is $H(f) = FT[h(t)]$. The output is given by $y(t) = x(t) * h(t)$ or $Y(f) = X(f)H(f)$.

Based on the derivation of the matched filter in Section 4.6.1, it is evident that the output energy is maximized when

$$H(f) = KX^*(f)\exp(-j2\pi f t_0), \tag{4.54}$$

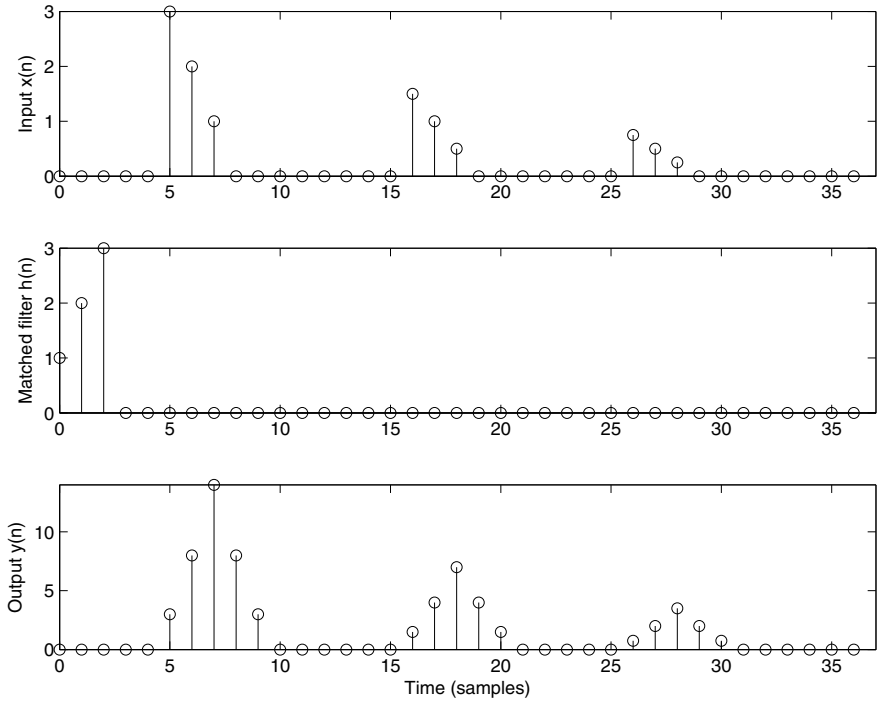

Figure 4.17 Top: A test signal with three events. Middle: Impulse response of the matched filter. Bottom: Output of the matched filter with peaks at the locations of occurrence of the basic signal pattern.

where K is a scale factor and t_0 is a time instant or delay [212]. This corresponds to the impulse response being

$$h(t) = Kx(t_0 - t). \tag{4.55}$$

Thus, the transfer function of the matched filter is proportional to the complex conjugate of the Fourier transform of the signal event to be detected. In the time domain, the impulse response is simply a *reversed* or *reflected* version of the reference signal that is scaled and delayed. A suitable delay will have to be added to make the filter causal, as determined by the duration of the reference signal.

If the derivation of the impulse response of the matched filter is based upon Equation 4.54 and implemented using the DFT, with the number of samples in the DFT equal to the number of samples N in the reference signal or template x, due to the periodicity of the DFT, the time delay or shift provided inherently will be $(N - 1)$. This is short by one sample; the desired shift is N samples to make the reversed signal causal.

Because the impulse response is a reversed version of $x(t)$, the convolution operation performed by the matched filter is equivalent to correlation: The output is then

equal to the cross-correlation between the input and the reference signal. When a portion of an input signal that is different from $x(t)$ matches the reference signal, the output approximates the ACF ϕ_x of the reference signal at the corresponding time delay. The corresponding frequency domain result is

$$Y(f) = X(f)H(f) = X(f)X^*(f) = S_x(f), \tag{4.56}$$

which is the PSD of the reference signal (ignoring the time delay and scale factors). The output is, therefore, maximum at the time instant of occurrence of an approximation to the reference signal. (See also Barlow [196] for related discussion.)

Illustration of application: To facilitate comparison with template matching, the spike-and-wave complex between 0.60 s and 0.82 s in the c3 channel of the EEG in Figure 1.40 was used as the reference signal to derive the matched filter. Figure 4.18 shows the extracted reference signal in the upper trace. The lower trace in the same figure shows the impulse response of the matched filter, which is simply a time-reversed version of the reference signal. The matched filter was implemented as an FIR filter using the MATLAB® *filter* command.

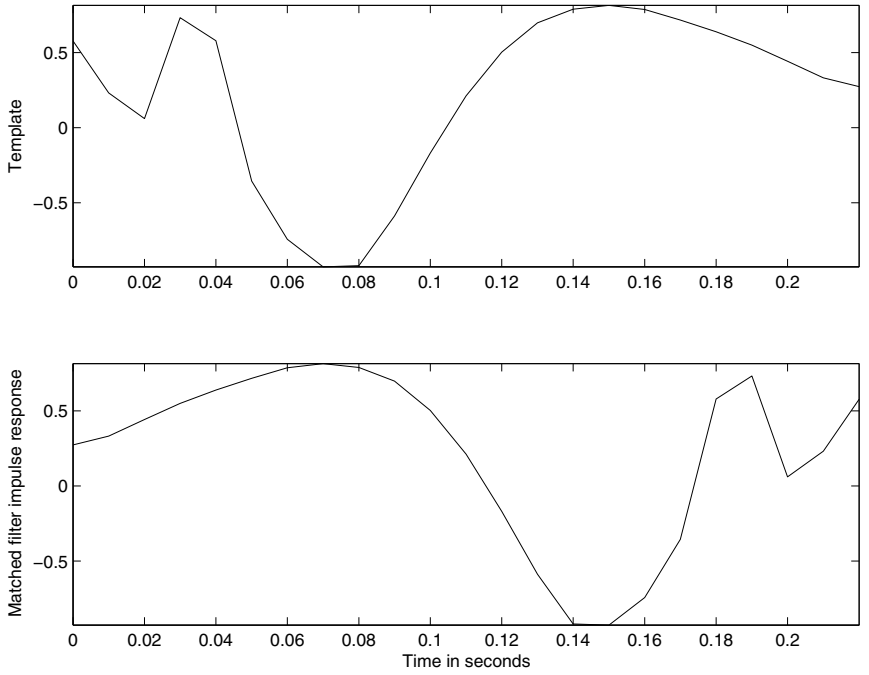

Figure 4.18 Upper trace: The spike-and-wave complex between 0.60 s and 0.82 s in the c3 channel of the EEG signal shown in Figure 1.40. Lower trace: Impulse response of the matched filter derived from the signal segment in the upper trace. Observe that the latter is a time-reversed version of the former.

Figures 4.19 and 4.20 show the outputs of the matched filter applied to the c3 and f3 channels of the EEG in Figure 1.40, respectively. The upper trace in each plot shows the signal, and the lower trace shows the output of the matched filter. It is evident that the matched filter provides a large output for each spike-and-wave complex. Comparing the outputs of the matched filter in Figures 4.19 and 4.20 with those of template matching in Figure 4.14 and 4.15, respectively, we observe that they are similar, with the exception that the results of the matched filter peak with a delay of $0.22\ s$ after the corresponding spike-and-wave complex. The delay corresponds to the duration of the impulse response of the filter. The values differ due to different normalization in Equation 3.96. (*Note:* MATLAB® provides the command *filtfilt* for zero-phase forward and reverse digital filtering; this method is not considered in the book.)

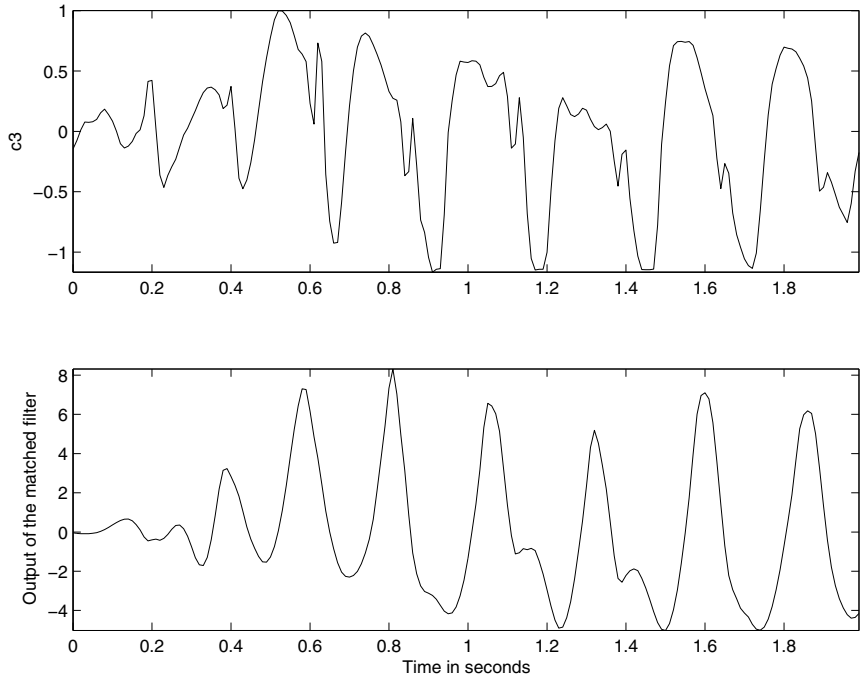

Figure 4.19 Upper trace: The c3 channel of the EEG signal shown in Figure 1.40, used as input to the matched filter in Figure 4.18. Lower trace: Output of the matched filter. See also Figure 4.14.

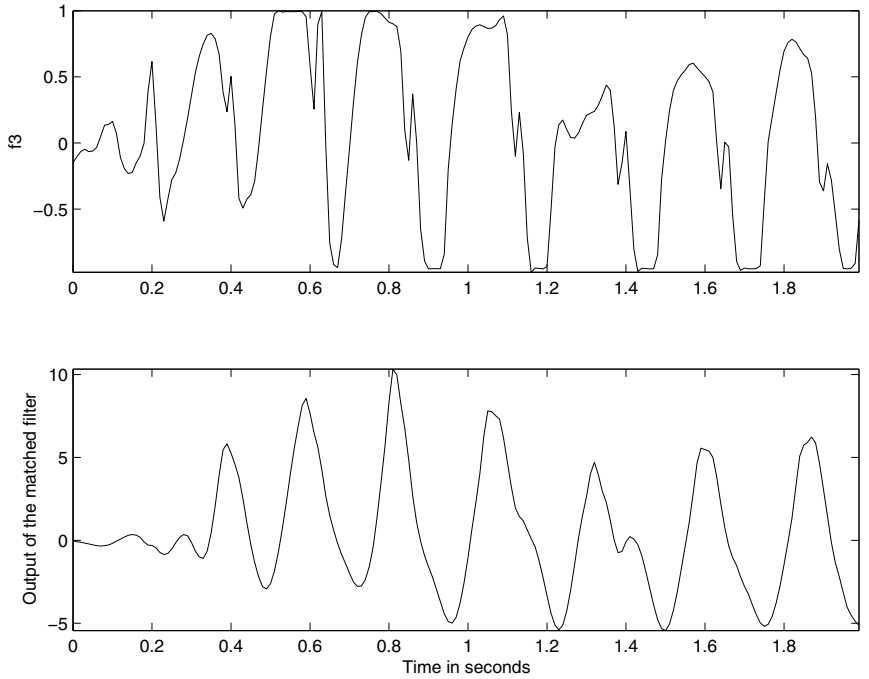

Figure 4.20 Upper trace: The f3 channel of the EEG signal shown in Figure 1.40, used as input to the matched filter in Figure 4.18. Lower trace: Output of the matched filter. See also Figure 4.15.

4.7 Detection of the P Wave in the ECG

Detection of the P wave in the ECG is difficult, because the P wave typically has a low amplitude, has an ill-defined and variable shape, and could be placed in a background of noise of varying size and origin.

Problem: *Propose an algorithm to detect the P wave in the ECG signal.*

Solution 1: In the method proposed by Hengeveld and van Bemmel [214], VCG signals are processed as follows:

1. The QRS is detected, deleted, and replaced with the baseline. The baseline is determined by analyzing a few samples preceding the QRS complex.
2. The resulting signal is bandpass filtered with $-3\ dB$ points at $3\ Hz$ and $11\ Hz$.
3. The end of the preceding T wave is estimated with reference to the current QRS by using $QT_{\max} = \frac{2}{9}RR + 250\ ms$, where RR is the interval between two successive QRS complexes.
4. The maximum and minimum values are found in all three VCG leads in the search interval from the preceding T wave to the current QRS.

5. The signal is rectified and thresholded at 50% and 75% of the maximum to obtain a ternary (three-level) signal.

6. The cross-correlation of the result is computed with a ternary template derived in a manner similar to the procedure in the previous step from a representative set of P waves.

7. The peak in the cross-correlation corresponds to the P location in the original ECG.

The algorithm overcomes the dominance of the QRS complex by first detecting the QRS and then deleting it. Observe that the cross-correlation is computed not with an original P wave, which we have noted could be obscure and variable, but with a ternary wave derived from the P wave. The ternary wave represents a simplified template of the P wave.

Figure 4.21 illustrates the results of the various stages of the P-finding algorithm of Hengeveld and van Bemmel [214]. Observe that the original ECG signal shown in part (a) of the figure has a P wave that is hardly discernible. The processed versions of the signal after deleting the QRS, filtering, and rectification are shown in parts (b), (c), and (d). The ternary version in part (e) shows that the P wave has been converted into two pulses corresponding to its upstroke and return parts. The result of cross-correlation with the template in part (f) is shown in part (g). A simple peak-picking algorithm with search limits may be used to detect the peak in the result, and hence determine the P wave position.

Note that the result in part (d) has other waves preceding those related to the P wave. An appropriate search interval should be used so as to disregard the unwanted components.

Solution 2: Gritzali et al. [215] proposed a common approach to detect the QRS, T, and P waves in multichannel ECG signals based upon a transformation they labeled as the "length" transformation. Given a collection of ECG signals from N simultaneous channels $x_1(t), x_2(t), \ldots, x_N(t)$, the length transformation was defined as

$$L(N, w, t) = \int_t^{t+w} \sqrt{\sum_{j=1}^{N} \left(\frac{dx_j}{dt}\right)^2}\, dt\,, \tag{4.57}$$

where w is the width of the time window over which the integration is performed. In essence, the procedure computes the total squared derivative of the signals across the various channels available, and integrates the summed quantity over a moving time window. The advantage of applying the derivative-based operator across multiple channels of an ECG signal is that P and T waves may be detected even when they well-defined in only one or a few of the channels used.

In the procedure for waveform detection proposed by Gritzali et al., the QRS is first detected by applying a threshold to $L(N, w, t)$, with w set equal to the average QRS width. The onset (start) and offset (end) points of the QRS are represented by a pulse waveform, as indicated in Figure 4.22. The QRS complexes in the signals are then replaced by the isoelectric baseline of the signals, the procedure is repeated

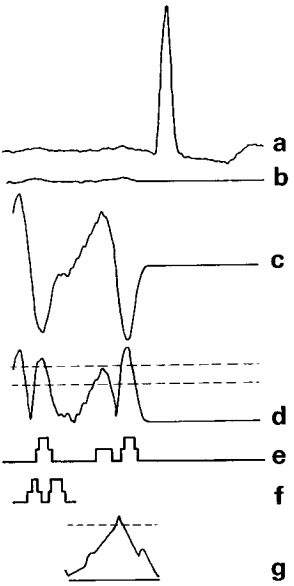

Figure 4.21 Illustration of the results at various stages of the Hengeveld and van Bemmel method for P wave detection. From top to bottom: (a) the original ECG signal; (b) after replacement of the QRS with the baseline; (c) after bandpass filtering; (d) after rectification, with the dashed lines indicating the thresholds; (e) the thresholded ternary signal; (f) the ternary P wave template; and (g) result of cross-correlation between the signals in (e) and (f). Reproduced with permission from S.J. Hengeveld and J.H. van Bemmel, Computer detection of P waves, *Computers and Biomedical Research*, 9:125–132, 1976. ©Academic Press.

with w set equal to the average T duration, and the T waves are detected. The same steps are repeated to detect the P waves. Figure 4.22 illustrates the detection of the QRS, T, and P waves in a three-channel ECG signal. Gritzali et al. also proposed a procedure based upon correlation analysis and least-squares modeling to determine the thresholds required.

4.8 Homomorphic Filtering and the Complex Cepstrum

In Chapter 3, we studied linear filters designed to separate signals that were added together. The question asked was the following: Given $y(t) = x(t) + \eta(t)$, how can one extract $x(t)$ only? Given that the Fourier transform is linear, we know that the Fourier transforms of the signals are also combined in an additive manner: $Y(\omega) = X(\omega) + \eta(\omega)$. Therefore, a linear filter will facilitate the separation of $X(\omega)$ and $\eta(\omega)$, with the assumption that they have substantial portions of their energy in different frequency bands.

Suppose now that we are presented with a signal that contains the product of two signals, for example, $y(t) = x(t)\,p(t)$. From the multiplication or convolution

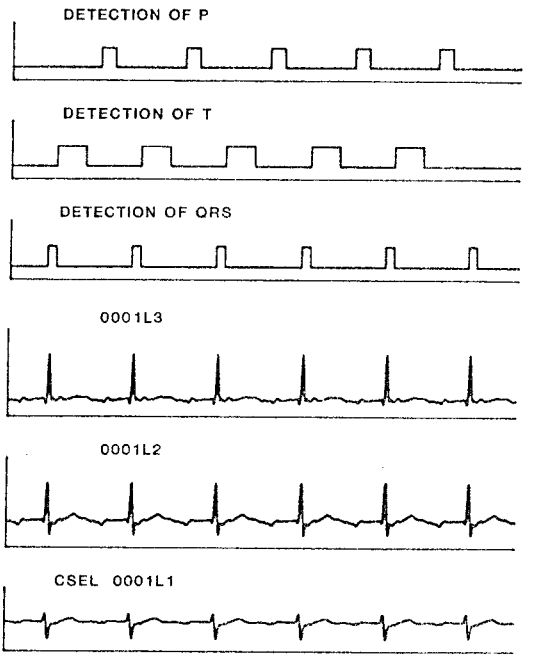

Figure 4.22 Detection of the P, QRS, and T waves in a three-channel ECG signal using the length transformation. The lower three traces show the three ECG channels. The upper three traces indicate the onset and end of the P, QRS, and T waves detected by the procedure in the form of pulse trains. The first P and the last T waves have not been processed. CSEL: CSE library. L1/L2/L3: three ECG leads. Reproduced with permission from F. Gritzali, G. Frangakis, and G. Papakonstantinou, Detection of the P and T waves in an ECG, *Computers and Biomedical Research*, 22:83–91, 1989. ©Academic Press. See Willems et al. [216, 217] for details on the ECG database used by Gritzali et al.

property of the Fourier transform we have $Y(\omega) = X(\omega) * P(\omega)$, where $*$ represents convolution in the frequency domain. How would we be able to separate $x(t)$ from $p(t)$?

Furthermore, suppose we have $y(t) = x(t) * h(t)$, where $*$ stands for convolution as in the case of the passage of the glottal pulse train or random excitation $x(t)$ through the vocal-tract system with the impulse response $h(t)$. The Fourier transforms of the signals are related as $Y(\omega) = X(\omega)\, H(\omega)$. How would we attempt to separate $x(t)$ and $h(t)$?

4.8.1 Generalized linear filtering

Given that linear filters are well established and understood, it is attractive to consider extending their application to signals that have been combined by operations other than addition, especially by multiplication and convolution as indicated in the

preceding paragraphs. An interesting possibility to achieve this is via conversion of the operation combining the signals into addition by one or more transformations. Under the assumption that the transformed signals occupy different portions of the transformed space, linear filters may be applied to separate them. The inverses of the transformations used initially would then take us back to the original space of the signals. This approach was proposed in a series of papers by Bogert et al. [218], Oppenheim et al. [219], and Oppenheim and Schafer [220, 221]. Because the procedure extends the application of linear filters to multiplied and convolved signals, it has been referred to as *generalized linear filtering*. Furthermore, as the operations can be represented by algebraically linear transformations between the input and output vector spaces, they have been called *homomorphic systems*.

As a simple illustration of a homomorphic system for multiplied signals, consider again the signal

$$y(t) = x(t)\, p(t). \tag{4.58}$$

Given the goal of converting the multiplication operation to addition, it is evident that a simple logarithmic transformation is appropriate:

$$\log[y(t)] = \log[x(t)\, p(t)] = \log[x(t)] + \log[p(t)]; \qquad x(t) \neq 0,\ p(t) \neq 0\ \forall t. \tag{4.59}$$

The logarithms of the two signals are now combined in an additive manner. Taking the Fourier transform, we get

$$Y_l(\omega) = X_l(\omega) + P_l(\omega), \tag{4.60}$$

where the subscript l indicates that the Fourier transform has been applied to a log-transformed version of the signal.

Assuming that the logarithmic transformation has not affected the separability of the Fourier components of the two signals $x(t)$ and $p(t)$, a linear filter (such as a lowpass or a highpass filter) may now be applied to $Y_l(\omega)$ to separate them. An inverse Fourier transform will yield the filtered signal in the time domain. An exponential operation will complete the reversal procedure (if required).

Figure 4.23 illustrates the operations involved in a multiplicative homomorphic system (or filter). The symbol at the input or output of each block indicates the operation that combines the signal components at the corresponding step. A system of this type is useful in image processing, where an image may be treated as the product of an illumination function and a transmittance or reflectance function. The homomorphic filter facilitates separation of the illumination function and correction for nonuniform lighting. The method has been used to achieve simultaneous dynamic range compression and contrast enhancement [18, 174, 219, 222].

4.8.2 Homomorphic deconvolution

Problem: *Propose a homomorphic filter to separate two signals that have been combined through the convolution operation.*

Solution: Consider the case expressed by the relation

$$y(t) = x(t) * h(t). \tag{4.61}$$

As in the case of the multiplicative homomorphic system, our goal is to convert the convolution operation to addition. From the convolution property of the Fourier transform, we know that

$$Y(\omega) = X(\omega)\, H(\omega). \tag{4.62}$$

Thus, application of the Fourier transform converts convolution to multiplication. Now, it is readily seen that the multiplicative homomorphic system may be applied to convert the multiplication to addition. Taking the complex logarithm of $Y(\omega)$, we have

$$\log[Y(\omega)] = \log[X(\omega)] + \log[H(\omega)]; \qquad X(\omega) \neq 0,\ H(\omega) \neq 0\ \forall\omega. \tag{4.63}$$

[*N*ote: $\log_e[X(\omega)] = \hat{X}(\omega) = \log_e[|X(\omega)|\angle X(\omega)] = \log_e[|X(\omega)|] + j\angle X(\omega)$, where $|X(\omega)|$ and $\angle X(\omega)$ are the magnitude and phase spectra of $x(t)$.]

A linear filter may now be used to separate the transformed components of x and h, with the assumption as before that they are separable in the transformed space. A series of the inverses of the transformations applied initially will take us back to the original domain.

While the discussion here has been in terms of application of the Fourier transform, the general formulation of the homomorphic filter by Oppenheim and Schafer [174] is in terms of the z-transform. However, the Fourier transform is equivalent to the z-transform evaluated on the unit circle in the z-plane, and the Fourier transform is more commonly used in signal processing than the z-transform.

Figure 4.24 gives a block diagram of the steps involved in a homomorphic filter for convolved signals. The path formed by the first three blocks (in the top row) transforms the convolution operation at the input to addition. The third block with the inverse Fourier transform is used to move back to a pseudo time domain. The last three blocks (in the bottom row) perform the reverse transformation, converting addition to convolution. The filter in between deals with (transformed) signals that are combined by simple addition. Further details of these operations and results are presented in the following section.

4.8.3 Extraction of the vocal tract response

Problem: *Design a homomorphic filter to extract the basic wavelet corresponding to the vocal-tract response from a voiced speech signal.*

Solution: We noted in Section 1.2.12 that voiced speech is generated by excitation of the vocal tract, while it is held in a particular form, with a glottal waveform that may be approximated as a series of pulses. The voiced speech signal may, therefore, be expressed in discrete-time terms as $y(n) = x(n) * h(n)$, where $y(n)$ is the speech signal, $x(n)$ is the glottal waveform (excitation sequence), and $h(n)$ is the impulse response of the vocal tract (basic wavelet). The $*$ symbol represents convolution, with the assumption that the vocal-tract filter may be approximated by an LSI filter. We may, therefore, use the homomorphic filter for convolved signals as introduced in the preceding section to separate $h(n)$ and $x(n)$.

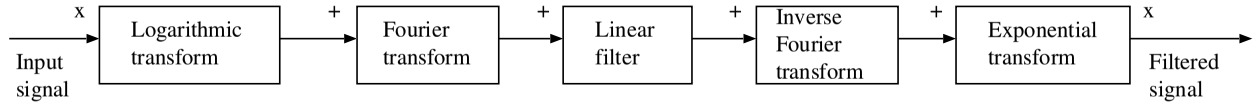

Figure 4.23 Operations involved in a multiplicative homomorphic system or filter. The symbol at the input or output of each block indicates the operation that combines the signal components at the corresponding step.

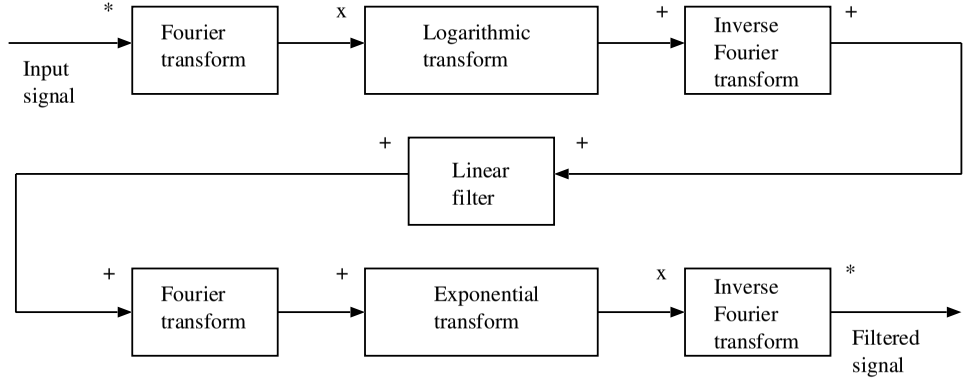

Figure 4.24 Operations involved in a homomorphic filter for convolved signals. The symbol at the input or output of each block indicates the operation that combines the signal components at the corresponding step.

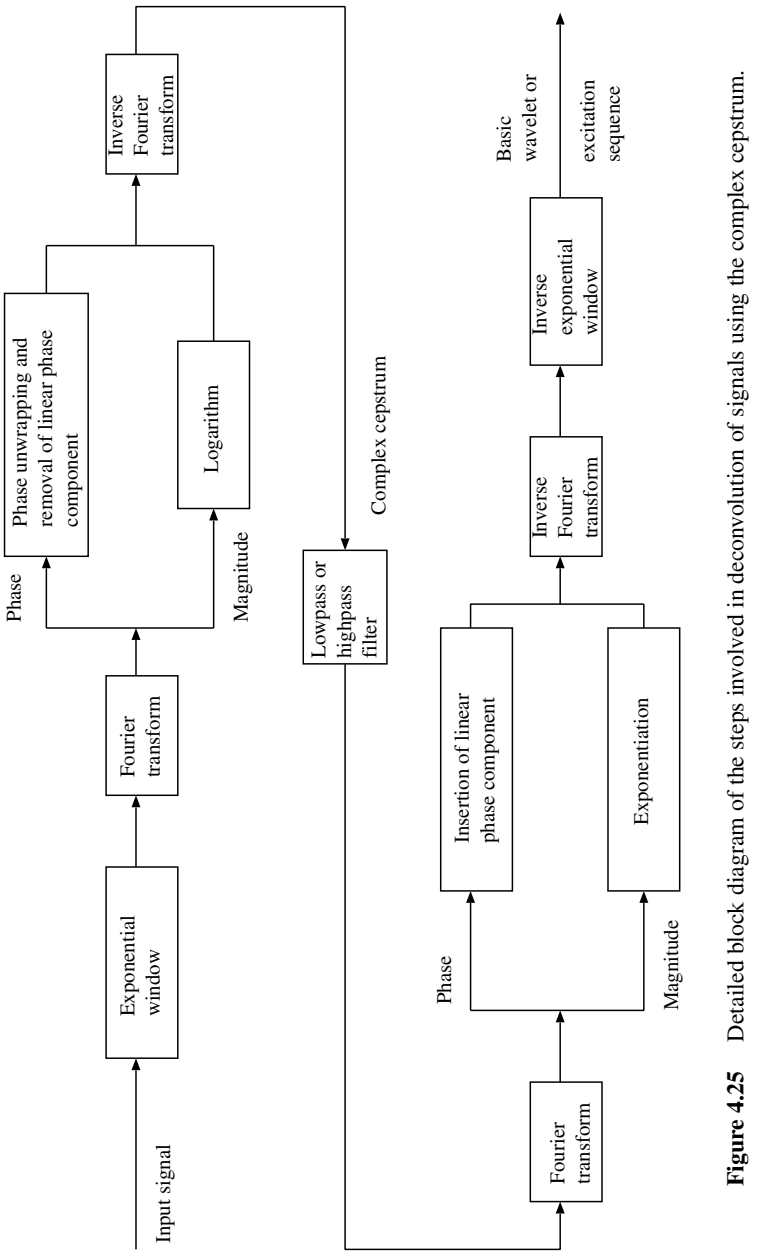

Figure 4.25 Detailed block diagram of the steps involved in deconvolution of signals using the complex cepstrum.

The glottal excitation sequence may be further expressed as $x(n) = p(n) * g(n)$, where $p(n)$ is a train of ideal impulses (Dirac delta functions) and $g(n)$ is a smoothing function, to indicate that the physical vocal-cord system cannot produce ideal impulses but rather pulses of finite duration and slope [174]. This aspect will be neglected in our discussions.

Practical application of the homomorphic filter is not simple. Figure 4.25 gives a detailed block diagram of the procedure [174, 223]. Some of the details and practical techniques are explained in the following paragraphs.

The complex cepstrum: The formal definition of the complex cepstrum states that it is the inverse z-transform of the complex logarithm of the z-transform of the input signal [174, 223]. (The name "cepstrum" was derived by transposing the syllables of the word "spectrum"; other transposed terms [174, 218, 223] are less commonly used.) If $y(n)$ is the input signal and $Y(z)$ is its z-transform, the complex cepstrum $\hat{y}(n)$ is defined as

$$\hat{y}(n) = \frac{1}{2\pi j} \oint \log[Y(z)]\, z^{n-1}\, dz. \tag{4.64}$$

The contour integral performs the inverse z-transform, and should be evaluated within an annular region in the complex z-plane where $\hat{Y}(z) = \log[Y(z)]$ is single-valued and analytic [3, 174]. The unit of n in $\hat{y}(n)$, that is, in the cepstral domain, is referred to as *quefrency*, a term obtained by switching the syllables in the term frequency; it is also referred to as pseudo time [218, 221].

Given $y(n) = x(n) * h(n)$, it follows that

$$\hat{Y}(z) = \hat{X}(z) + \hat{H}(z) \;\text{ or }\; \hat{Y}(\omega) = \hat{X}(\omega) + \hat{H}(\omega), \tag{4.65}$$

and further that the complex cepstra of the signals are related as

$$\hat{y}(n) = \hat{x}(n) + \hat{h}(n). \tag{4.66}$$

Here, the $\hat{}$ symbol over a function of z or ω indicates the complex logarithm of the corresponding function of z or ω, whereas the same symbol over a function of time (n) indicates the complex cepstrum of the corresponding signal. It should be noted that, if the original signal $y(n)$ is real, its complex cepstrum $\hat{y}(n)$ is *real*; the prefix *complex* is used to indicate the fact that the preceding z and logarithmic transformations are computed as complex functions. Furthermore, it should be noted that the complex cepstrum is a function of time.

An important consideration in the evaluation of the complex logarithm of $Y(z)$ or $Y(\omega)$ relates to the phase of the signal. The phase spectrum computed as its principal value in the range $[0, 2\pi]$, given by $\arctan\left[\frac{\text{imaginary}\{Y(\omega)\}}{\text{real}\{Y(\omega)\}}\right]$, will almost always have discontinuities that will conflict with the requirements of the inverse z-transformation or inverse Fourier transform to follow later. Thus, $Y(\omega)$ needs to be separated into its magnitude and phase components, the logarithmic operation applied to the magnitude, the phase corrected to be continuous by adding correction factors of $\pm 2\pi$ at discontinuities larger than π, and the two components combined

again before the subsequent inverse transformation. Correcting the phase spectrum as above is referred to as *phase unwrapping* [174, 223]. It has been shown that a linear phase term, if present in the spectrum of the input signal, may cause rapidly decaying oscillations in the complex cepstrum [223]. It is advisable to remove the linear phase term, if present, in the phase-unwrapping step. The linear phase term may be added to the filtered result (as a time shift) if necessary.

The complex cepstrum of exponential sequences: Exponential sequences are signals that have a rational z-transform, that is, their z-transforms may be expressed as ratios of polynomials in z [174]. Such signals are effectively represented as weighted sums of exponentials in the time domain. Consider a signal $x(n)$ whose z-transform is expressed as

$$X(z) = A\, z^r\, \frac{\prod_{k=1}^{M_I} (1 - a_k\, z^{-1})\ \prod_{k=1}^{M_O} (1 - b_k\, z)}{\prod_{k=1}^{N_I} (1 - c_k\, z^{-1})\ \prod_{k=1}^{N_O} (1 - d_k\, z)}, \tag{4.67}$$

with $|a_k|, |b_k|, |c_k|, |d_k| < 1$. The function $X(z)$ possesses M_I zeros that are located inside the unit circle, given by a_k, $k = 1, 2, \ldots, M_I$; M_O zeros outside the unit circle, given by $1/b_k$, $k = 1, 2, \ldots, M_O$; N_I poles inside the unit circle, given by c_k, $k = 1, 2, \ldots, N_I$; and N_O poles outside the unit circle, given by $1/d_k$, $k = 1, 2, \ldots, N_O$.

In the process of computing the complex cepstrum, we get [174]

$$\begin{aligned}
\hat{X}(z) &= \log[X(z)] \qquad (4.68)\\
&= \log[A] + \log[z^r]\\
&+ \sum_{k=1}^{M_I} \log(1 - a_k\, z^{-1}) + \sum_{k=1}^{M_O} \log(1 - b_k\, z)\\
&- \sum_{k=1}^{N_I} \log(1 - c_k\, z^{-1}) - \sum_{k=1}^{N_O} \log(1 - d_k\, z)\,.
\end{aligned}$$

For real sequences, A is real. If $A > 0$, $\hat{x}(0) = \log[A]$.

The factor z^r indicates a delay or advance by r samples, depending upon the sign of r.

Recall the power series

$$\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \quad \text{if } |x| < 1. \tag{4.69}$$

Thus, we have

$$\log(1 - \alpha\, z^{-1}) = -\sum_{n=1}^{\infty} \frac{\alpha^n}{n}\, z^{-n}, \quad \text{if } |z| > |\alpha|, \tag{4.70}$$

and

$$\log(1-\beta\, z) = -\sum_{n=1}^{\infty} \frac{\beta^n}{n}\, z^n\,, \quad \text{if} \;\; |z| < |\beta^{-1}|. \tag{4.71}$$

Expanding all of the log terms in Equation 4.68 into their equivalent power series and taking the inverse z-transform, we get [174]

$$\hat{x}(n) = \begin{cases} \log|A| & \text{for} \;\; n = 0; \\ -\sum_{k=1}^{M_I} \frac{a_k^n}{n} \;+\; \sum_{k=1}^{N_I} \frac{c_k^n}{n} & \text{for} \;\; n > 0; \\ \sum_{k=1}^{M_O} \frac{b_k^{-n}}{n} \;-\; \sum_{k=1}^{N_O} \frac{d_k^{-n}}{n} & \text{for} \;\; n < 0. \end{cases} \tag{4.72}$$

From the steps shown above, a few properties of the complex cepstrum that are evident and important are as follows [174]:

- The complex cepstrum, $\hat{x}(n)$, decays at least as fast as $1/n$. We have

$$|\hat{x}(n)| < K \left|\frac{\alpha^n}{n}\right|\,, \quad -\infty < n < \infty, \tag{4.73}$$

  where $\alpha = \max(|a_k|, |b_k|, |c_k|, |d_k|)$, and K is a constant.

- $\hat{x}(n)$ will be of infinite duration even if $x(n)$ is of finite duration, and exists for $-\infty < n < \infty$, in general.

- If $x(n)$ is a minimum-phase signal, all of its poles and zeros are inside the unit circle in the z-plane. Then, $\hat{x}(n) = 0$ for $n < 0$; that is, the complex cepstrum is causal.

- If $x(n)$ is a maximum-phase signal, it has no poles or zeros inside the unit circle in the z-plane. Then, $\hat{x}(n) = 0$ for $n > 0$; that is, the complex cepstrum is anticausal.

Limiting ourselves to causal signals of finite energy, we need not consider the presence of poles on or outside the unit circle in the z-plane. However, the z-transform of a finite-energy signal may have zeros outside the unit circle. Such a composite signal may be separated into its minimum-phase component and maximum-phase component by extracting the causal part ($n > 0$) and anticausal part ($n < 0$), respectively, of its complex cepstrum, followed by the inverse procedures. The composite signal is equal to the convolution of its minimum-phase component and maximum-phase component. (See also Section 5.4.2.)

Effect of echoes or repetitions of a wavelet: Let us consider a simplified signal $y(n) = x(n) * h(n)$, where

$$x(n) = \delta(n) + a\,\delta(n - n_0), \tag{4.74}$$

with a and n_0 being two constants. (The sampling interval T is ignored, or assumed to be normalized to unity in this example.) The signal may also be expressed as

$$y(n) = h(n) + a\,h(n - n_0). \tag{4.75}$$

The signal thus has two occurrences of the basic wavelet $h(n)$ at $n = 0$ and $n = n_0$. The coefficient a indicates the magnitude of the second appearance of the basic wavelet (called an *echo* in seismic applications), and n_0 indicates its delay (pitch in the case of a voiced speech signal). The topmost plot in Figure 4.26 shows a synthesized signal with a wavelet and an echo at one-half of the amplitude (that is, $a = 0.5$) of the first wavelet arriving at $n_0 = 0.01125\ s$.

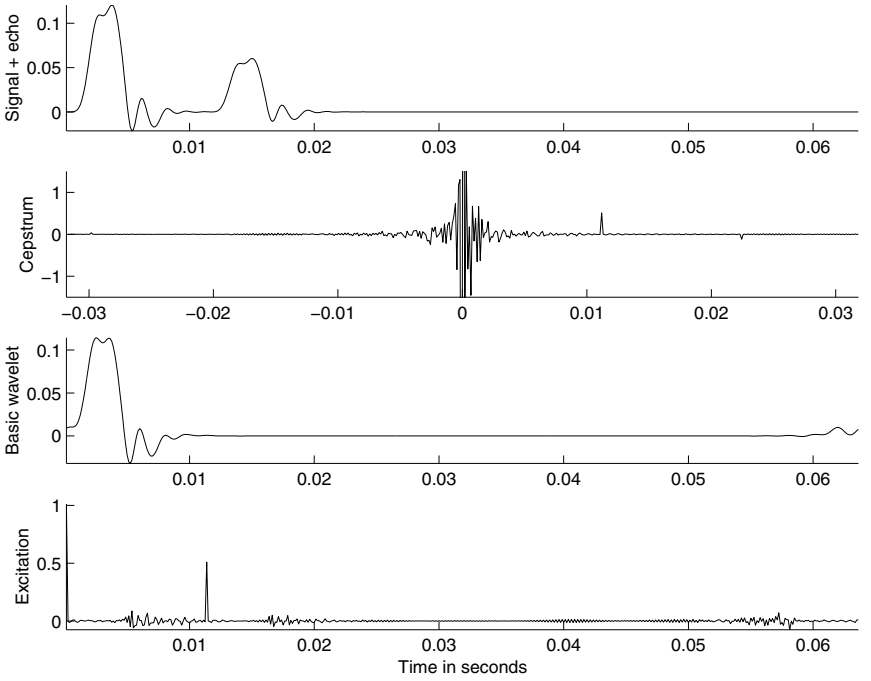

Figure 4.26 From top to bottom: a composite signal with a wavelet and an echo; the complex cepstrum of the signal (the amplitude axis has been stretched to make the peaks at the echo time and its multiples more readily visible; values outside the range ± 1.5 have been clipped); the basic wavelet extracted by shortpass filtering the cepstrum; and the excitation sequence extracted by longpass filtering the cepstrum.

Taking the z-transform of the signal in Equation 4.75, we have

$$Y(z) = (1 + az^{-n_0})H(z). \tag{4.76}$$

If the z-transform is evaluated on the unit circle, we get the Fourier-transform-based expression

$$Y(\omega) = [1 + a\,\exp(-j\omega n_0)]H(\omega). \tag{4.77}$$

Taking the logarithm, we have

$$\hat{Y}(\omega) = \hat{H}(\omega) + \log[1 + a\ \exp(-j\omega n_0)]. \tag{4.78}$$

If $a < 1$, the log term may be expanded in a power series, to get

$$\hat{Y}(\omega) = \hat{H}(\omega) + a\ \exp(-j\omega n_0) - \frac{a^2}{2}\ \exp(-2j\omega n_0) + \frac{a^3}{3} \exp(-3j\omega n_0) - \cdots. \tag{4.79}$$

Taking the inverse Fourier transform, we get

$$\hat{y}(n) = \hat{h}(n) + a\,\delta(n - n_0) - \frac{a^2}{2}\,\delta(n - 2n_0) + \frac{a^3}{3}\,\delta(n - 3n_0) - \cdots. \tag{4.80}$$

The derivation given above shows that the complex cepstrum of a signal with a basic wavelet and an echo is equal to the complex cepstrum of the basic wavelet plus a series of impulses at the echo delay and integral multiples thereof [174, 223]. The amplitudes of the impulses are proportional to the echo amplitude (the factor a) and decay for the higher-order repetitions (if $a < 1$). It may be readily seen that, if the signal has multiple echoes or repetitions of a basic wavelet, the cepstrum will possess multiple impulse trains, with an impulse at the arrival time of each wavelet and integral multiples thereof. In the case of a voiced speech signal, the location of the first peak will give the pitch. The second plot in Figure 4.26 shows the complex cepstrum of the signal in the first plot of the same figure. It is seen that the cepstrum has a peak at $0.01125\ s$, the echo arrival time; a smaller (negative) peak is also seen at twice the echo arrival time.

Under the assumption that the complex cepstrum of the basic wavelet decays to negligible values before the first impulse $a\,\delta(n-n_0)$ related to the echo, $\hat{h}(n)$ may be extracted from the complex cepstrum $\hat{y}(n)$ of the composite signal by a window that has unit value for $|n| < n_c$, n_c being the cutoff point. (This filter is referred to as a "shortpass" filter as the cepstrum is a function of pseudo time [218, 221]; it may also be called a lowpass filter.) The inverse procedures will yield $h(n)$. The remaining portion of the cepstrum (obtained by "longpass" or highpass filtering) will give $\hat{x}(n)$, which upon application of the inverse procedures will yield $x(n)$. The third and fourth plots in Figure 4.26 show the basic wavelet $h(n)$ and the excitation sequence $x(n)$ extracted by filtering the cepstrum with the cutoff point at $n_c = 0.005\ s$.

In the case where $a \geq 1$, it can be shown that the complex cepstrum will have a train of impulses on its negative time axis, that is, at $(n + kn_0),\ k = 1, 2, \ldots$ [174, 223]. An appropriate exponential weighting sequence may be used to achieve the condition $a < 1$, in which case the impulse train will appear on the positive axis of the cepstrum. If the weighted signal satisfies the minimum-phase condition, the cepstrum will be causal.

The power cepstrum: Several variants of the cepstrum have been proposed in the literature; Childers et al. [223] provided a review of the related techniques. One variant that is commonly used is the *real cepstrum* or the *power cepstrum*, which is defined as the square of the inverse z-transform of the logarithm of the squared magnitude of the z-transform of the given signal. In practice, the z-transform in the

definition stated is replaced by the DFT. The power cepstrum has the computational advantage of not requiring phase unwrapping, but does not facilitate separation of the components of the signal.

By evaluating the inverse z-transform on the unit circle in the z-plane, the power cepstrum $\hat{y}_p(n)$ of a signal $y(n)$ may be defined as

$$\hat{y}_p(n) = \left\{ \frac{1}{2\pi j} \oint \log |Y(z)|^2 z^{n-1} dz \right\}^2. \tag{4.81}$$

If, as before, we consider $y(n) = x(n) * h(n)$, we have $|Y(z)|^2 = |X(z)|^2 |H(z)|^2$, and it follows that $\log |Y(z)|^2 = \log |X(z)|^2 + \log |H(z)|^2$. Applying the inverse z-transform to this relationship, we get

$$\hat{y}_p(n) = \hat{x}_p(n) + \hat{h}_p(n), \tag{4.82}$$

where $\hat{h}_p(n)$ is the power cepstrum of the basic wavelet and $\hat{x}_p(n)$ is the power cepstrum of the excitation signal. Note that, in the equation given above, the cross-product term was neglected; the cross-term will be zero if the two component power cepstra occupy nonoverlapping quefrency ranges. The final squaring operation in Equation 4.81 is omitted in some definitions of the power cepstrum; in such a case, the cross-term does not arise, and Equation 4.82 is valid.

The power cepstrum does not retain the phase information of the original signal. However, it is useful in the identification of the presence of echoes in the signal, and in the estimation of their arrival times. The power cepstrum is related to the complex cepstrum as [223]

$$\hat{y}_p(n) = [\hat{y}(n) + \hat{y}(-n)]^2. \tag{4.83}$$

Let us again consider the situation of a signal with two occurrences of a basic wavelet $h(n)$ at $n = 0$ and $n = n_0$ as in Equations 4.74 and 4.75. Then [223],

$$|Y(z)|^2 = |H(z)|^2 |1 + az^{-n_0}|^2. \tag{4.84}$$

By taking the logarithm of both sides of the equation and substituting $z = \exp(j\omega)$, we get

$$\begin{aligned} \log |Y(\omega)|^2 &= \log |H(\omega)|^2 + \log[1 + a^2 + 2\,a\,\cos(\omega n_0)] \\ &= \log |H(\omega)|^2 + \log(1 + a^2) \\ &+ \log \left(1 + \frac{2a}{1 + a^2} \cos(\omega n_0) \right). \end{aligned} \tag{4.85}$$

It is now seen that the logarithm of the PSD of the signal has sinusoidal components (ripples) due to the presence of an echo. The amplitudes and frequencies of the ripples are related to the amplitude a of the echo and its time delay n_0.

Illustration of application: A voiced speech signal $y(n)$ is the result of convolution of a slowly varying vocal-tract response $h(n)$ with a relatively fast-varying glottal pulse train $x(n)$, expressed as $y(n) = x(n) * h(n)$. Under these conditions,

it may be demonstrated that the contributions of $h(n)$ to the complex cepstrum $\hat{y}(n)$ will be limited to low values of n within the range of the pitch period of the speech signal [174]. The complex cepstrum $\hat{y}(n)$ will possess impulses at the pitch period and integral multiples thereof. Therefore, a filter that selects a portion of the complex cepstrum for low values of n, followed by the inverse transformations, will yield an estimate of $h(n)$.

When the repetitions of the basic wavelet have magnitudes almost equal to (or even greater than) that of the first wavelet in the given signal record, the contributions of the pulse-train component to the complex cepstrum may not decay rapidly and may cause aliasing artifacts when the cepstrum is computed over a finite duration. A similar situation is caused when the delay between the occurrences of the multiple versions of the basic wavelet is a significant portion of the duration of the given signal. The problem may be ameliorated by applying an exponential weight α^n to the data sequence, with $\alpha < 1$. Childers et al. [223] recommended values of α in the range $[0.96, 0.99]$, depending upon the signal characteristics as listed above. Furthermore, they recommend appending or padding zeros to the input signal to facilitate computation of the cepstrum to a longer duration than the signal in order to avoid aliasing errors and ambiguities in time-delay estimates. (See Figure 4.25 for an illustration of the various steps in homomorphic filtering of convolved signals.)

Figure 4.27 illustrates a segment of a voiced speech signal (obtained by filtering the signal shown in Figure 1.51) and the basic wavelet extracted by shortpass filtering of its complex cepstrum with $n_c = 0.003125\ s$. The signal was padded with zeros to twice its duration; exponential weighting with $\alpha = 0.99$ was used. It is seen that the basic vocal-tract response wavelet has been successfully extracted. Extraction of the vocal-tract response facilitates spectral analysis without the effect of the quasiperiodic repetitions in the speech signal as indicated by Equation 4.85.

The fourth trace in Figure 4.27 shows the glottal (excitation) waveform extracted by longpass filtering of the cepstrum with the same parameters as in the preceding paragraph. The result shows impulses at the time of arrival of each wavelet in the composite speech signal. The peaks are decreasing in amplitude due to the use of exponential weighting (with $\alpha = 0.99$) prior to computation of the cepstrum. Inverse exponential weighting can restore the pulses to their original levels; however, the artifact at the end of the excitation signal gets amplified to much higher levels than the desired pulses due to progressively higher values of α^{-n} for large n. Hence, the inverse weighting operation was not applied in the present illustration. Regardless, the result indicates that pitch information may also be recovered by homomorphic filtering of voiced speech signals.

4.9 Application: ECG Rhythm Analysis

Problem: *Describe a method to measure the heart rate and average RR interval from an ECG signal.*

Solution: Algorithms for QRS detection such as the Pan–Tompkins method described in Section 4.3.2 are useful for ECG rhythm analysis or heart-rate monitoring.

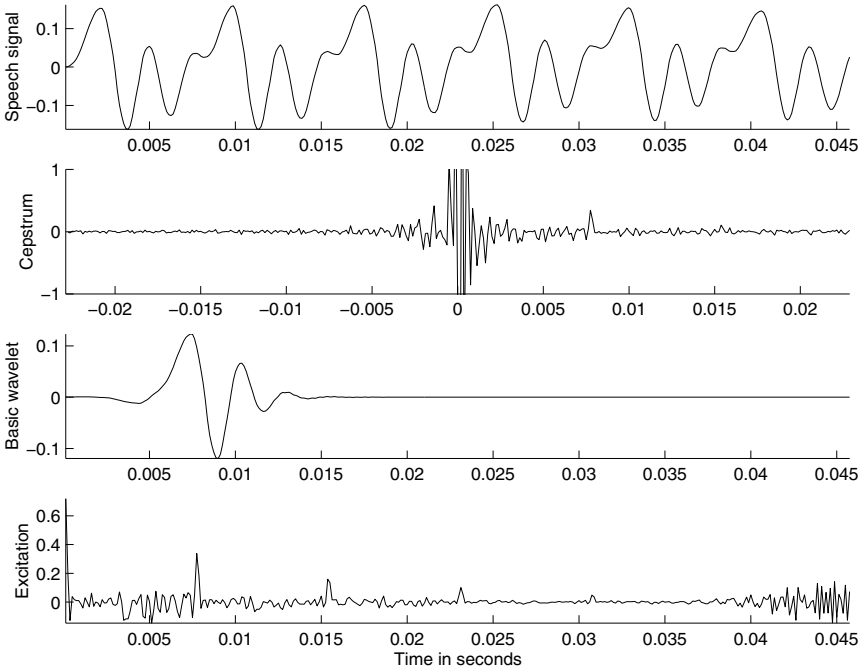

Figure 4.27 From top to bottom: a segment of a voiced speech signal over six pitch periods (extracted from the signal shown in Figure 1.51 and lowpass filtered); the complex cepstrum of the signal (the amplitude axis has been stretched to make the peaks at the echo time and its multiples more readily visible; values outside the range ± 1.0 have been clipped); the (shifted) basic wavelet extracted by shortpass filtering the cepstrum; and the excitation sequence extracted by longpass filtering the cepstrum.

The output of the final smoothing filter could be subjected to a peak-searching algorithm to obtain a time marker for each QRS or ECG beat. The search procedure proposed by Pan and Tompkins was explained in Section 4.3.2. The intervals between two such consecutive markers gives the RR interval, which could be averaged over a number of beats to obtain a good estimate of the interbeat interval. The heart rate may be computed in bpm as 60 divided by the average RR interval in seconds. The heart rate may also be obtained by counting the number of beats detected over a certain period, for example, $10\ s$, and multiplying the result with the required factor (6 in the present case) to get the number of beats over one minute.

The upper plot in Figure 4.28 shows a filtered version of the noisy ECG signal shown in Figure 3.5. The noisy signal was filtered with an eighth-order Butterworth lowpass filter with a cutoff frequency of $90\ Hz$, and the signal was down-sampled by a factor of five to an effective sampling rate of $200\ Hz$. The lower plot shows the output of the Pan–Tompkins method. The Pan–Tompkins result was normalized by dividing by its maximum over the data record available (as the present example was

computed off-line). A fixed threshold of 0.1 and a blanking interval of $250\ ms$ was used in a simple search procedure, which was successful in detecting all the beats in the signal. (The blanking interval indicates the period over which threshold checking is suspended once the threshold has been crossed.) The average RR interval was computed as $716\ ms$, leading to an effective heart rate of $84\ bpm$.

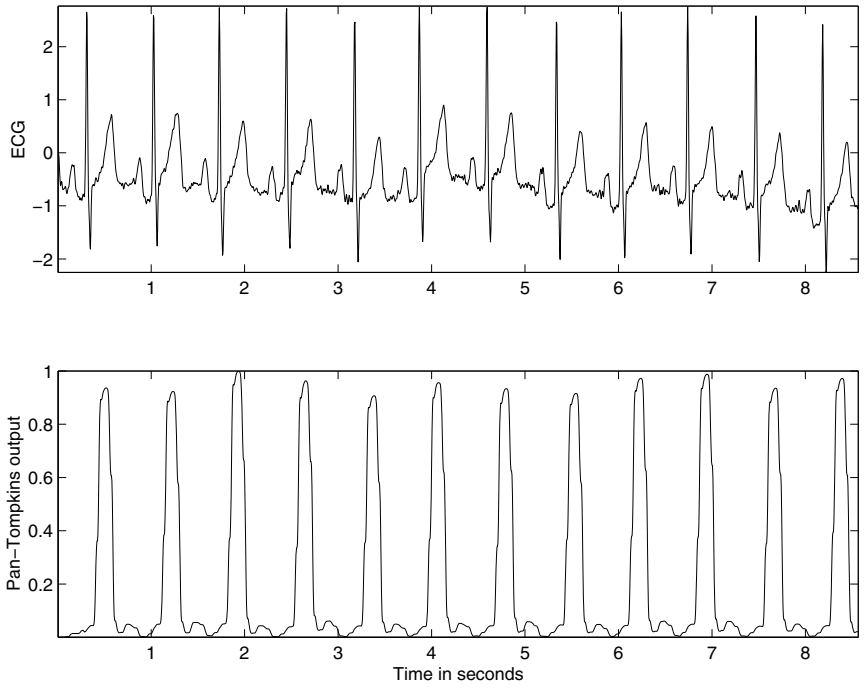

Figure 4.28 Results of the Pan–Tompkins algorithm. Top: lowpass-filtered version of the ECG signal shown in Figure 3.5. Bottom: normalized result of the final integrator.

Results at the various stages of the Pan–Tompkins algorithm for a noisy ECG signal sampled at $200\ Hz$ are shown in Figure 4.29. The bandpass filter has efficiently removed the low-frequency artifact in the signal. The final output has two peaks that are much larger than the others: one at the beginning of the signal due to filtering artifacts, and one at about $7.5\ s$ due to an artifact in the signal. Furthermore, the output has two peaks for the beat with an artifact at $7.5\ s$. The simple peak-searching procedure as explained in the previous paragraph was applied, which resulted in the detection of 46 beats: one more than the 45 present in the signal due to the artifact at about $7.5\ s$. The average RR interval was computed to be $446.6\ ms$, leading to an effective heart rate of $137\ bpm$.

The illustration demonstrates the need for prefiltering a given ECG signal to remove artifacts and the need to apply an adaptive threshold to the output of the Pan–Tompkins algorithm for QRS detection. It is readily seen that direct thresholding of

the original ECG signal will not be successful in detecting all of the QRS complexes in the signal.

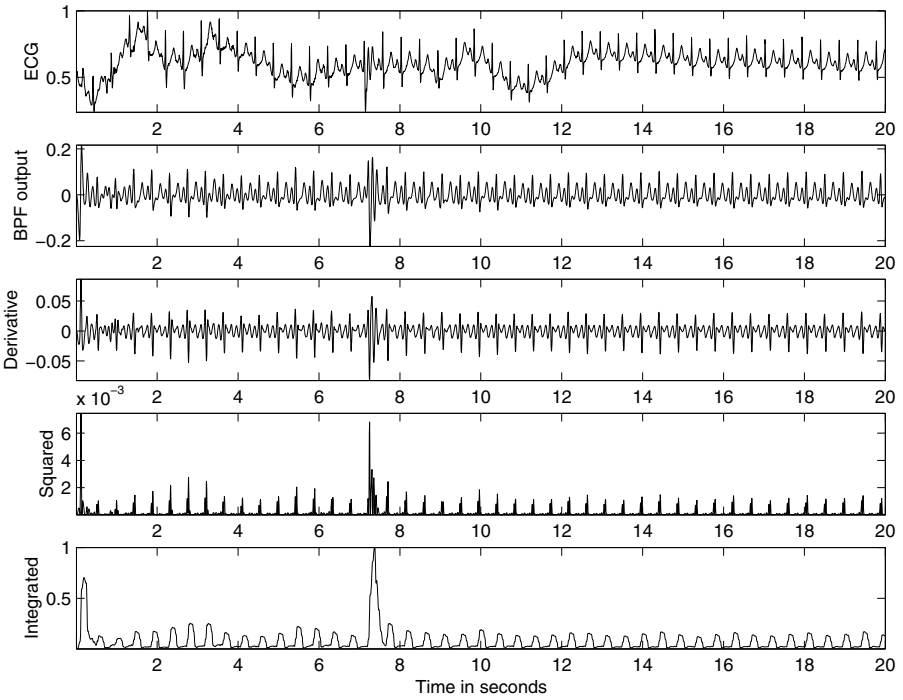

Figure 4.29 Results of the Pan–Tompkins algorithm with a noisy ECG signal. From top to bottom: ECG signal sampled at 200 Hz; output of the bandpass filter (BPF); output of the derivative-based operator; the result of squaring; and normalized result of the final integrator.

4.10 Application: Identification of Heart Sounds

Problem: *Outline a signal processing algorithm to identify S1 and S2 in a PCG signal, and further segment the PCG signal into its systolic and diastolic parts. The ECG and carotid pulse signals are available for reference.*

Solution: We saw in Section 2.3 how the ECG and carotid pulse signals could be used to demarcate the onset of S1 and S2 in the PCG; the procedure, however, was not based upon signal processing but upon visual analysis of the signals. We have, in the present chapter, studied signal processing techniques to detect the QRS complex in the ECG and the dicrotic notch in the carotid pulse signal. We may, therefore, use these methods to transfer the timing information from the ECG and carotid pulse signals to the PCG signal. In order to perform this task, we need to recognize a few timing relationships between the signals [72, 131].

The beginning of S1 may be taken to be the same instant as the beginning of the QRS. The QRS itself may be detected using one of the three methods described in the present chapter, such as the Pan–Tompkins method.

Detection of the beginning of S2 is more involved. Let the heart rate be $HR\ bpm$. The preejection period PEP is defined as the interval from the beginning of the QRS to the onset of the corresponding carotid upstroke. The rate-corrected PEP is defined as $PEPC = PEP + 0.4HR$, with the periods in ms. $PEPC$ is in the range of $131 \pm 13\ ms$ for normal adults [72].

The ejection time ET is the interval from the onset of the carotid upstroke to the dicrotic notch. The rate-corrected ejection time in ms is $ETC = ET + 1.6HR$, and is in the ranges of $395 \pm 13\ ms$ for normal adult males and $415 \pm 11\ ms$ for normal adult females.

Using $PEPC_{\max} = 144\ ms$ and $HR_{\min} = 60\ bpm$, we get $PEP_{\max} = 120\ ms$. With $HR_{\min} = 60\ bpm$ and $ETC_{\max} = 425\ ms$, we get $ET_{\max} = 329\ ms$. Based on the parameters defined in the preceding sentences, the maximum interval between the QRS and the dicrotic notch can be approximated to be $380\ ms$. The procedure proposed by Lehner and Rangayyan [131] for detection of the dicrotic notch recommends searching the output of the derivative-based method described in Section 4.3.3 in a $500\ ms$ interval after the QRS. After the dicrotic notch is detected, we need to subtract the time delay (S2–D) between the beginning of S2 and the dicrotic notch to get the time instant where S2 begins. Lehner and Rangayyan [131] measured the average S2–D delay over the PCG and carotid pulse signals of 60 pediatric patients to be $42.6\ ms$, with $SD = 5.0\ ms$.

The following procedure may be used to segment a PCG signal into its systolic and diastolic parts.

1. Use the Pan–Tompkins method described in Section 4.3.2 to locate the QRS complexes in the ECG.
2. Identify one period of the PCG as the interval between two successive QRS locations. Note that the delay introduced by the filters used in the Pan–Tompkins method needs to be subtracted from the detected peak locations to obtain the starting points of the QRS complexes.
3. Use the Lehner and Rangayyan method described in Section 4.3.3 to detect the dicrotic notch in the carotid pulse signal.
4. Let the standardized S2–D delay be its mean $+2 \times SD$ as reported by Lehner and Rangayyan [131], that is, $52.6\ ms$. Subtract the standardized S2–D delay from the detected dicrotic notch location to obtain the onset of S2.
5. Use the S1–S2 interval to obtain the systolic part of the PCG cycle.
6. Use the interval between the S2 point and the next detected S1 to obtain the diastolic part of the PCG cycle.

Figures 4.30 and 4.31 illustrate the results of application of the procedure described above to the PCG, ECG, and carotid pulse signals of a normal subject and

a patient with a split S2, systolic ejection murmur, and opening snap of the mitral valve. (Clinical diagnosis indicated the possibility of ventricular septal defect, pulmonary stenosis, or pulmonary hypertension for the 14-month-old female patient with murmur.) The peak positions detected in the output of the Pan–Tompkins method (the third trace in each figure) and the output of the Lehner and Rangayyan method (the fifth trace) have been marked with the $*$ symbol. A threshold of 0.75 times the maximum value was used to detect the peaks in the output of the Pan–Tompkins method, with a blanking interval of $250\ ms$.

The QRS and dicrotic notch positions have been marked on the ECG and carotid pulse traces with the triangle and diamond symbols, respectively. Finally, the S1 and S2 positions are marked on the PCG trace with triangles and diamonds, respectively. The filter delays and timing relationships between the three channels of signals described previously have been accounted for in the process of marking the events. Note how the results of event detection in the ECG and carotid pulse signals have been transferred to locate the corresponding events in the PCG. Lehner and Rangayyan [131] used a similar procedure to break PCG signals into systolic and diastolic segments; the segments were then analyzed separately in the time and frequency domains. (See also Sections 6.3.5 and 7.9.)

4.11 Application: Detection of the Aortic Component of the Second Heart Sound

Heart sounds are preferentially transmitted to different locations on the chest. The aortic and pulmonary components A2 and P2 of S2 are best heard at the aortic area (to the right of the sternum, in the second right intercostal space) and the pulmonary area (left parasternal line in the third left intercostal space), respectively (see Figure 1.31). A2 is caused by the closure of the aortic valve at the end of systole, and is usually louder than P2 at all locations on the chest. Earlier theories on the genesis of heart sounds attributed the sounds to the opening and closing actions of the valve leaflets *per se*. The more commonly accepted theory at the present time is that described by Rushmer [33]; see Section 1.2.9.

The relative timing of A2 and P2 depends upon the pressure differences across the corresponding valves in the left and right ventricular circulatory systems. In a normal individual, the timing of P2 with reference to A2 varies with respiration; the timing of A2 itself is independent of respiration. The pulmonary pressure (intrathoracic pressure) is decreased during inspiration, leading to a delayed closure of the pulmonary valve and hence an increased (audible and visible) gap between A2 and P2 [33, 72, 73]. The gap is closed and A2 and P2 overlap during expiration in normal individuals. A2 and P2 have individual durations of about $50\ ms$. The normal inspiratory gap between A2 and P2 is of the order of $30 - 40\ ms$, although splits as long as $100\ ms$ have been recorded [72].

A split in S2 longer than $40\ ms$ during sustained expiration is considered to be abnormal [72]. Complete right bundle-branch block could cause delayed activation of the right ventricle and, therefore, delayed pulmonary valve closure, a delayed P2,

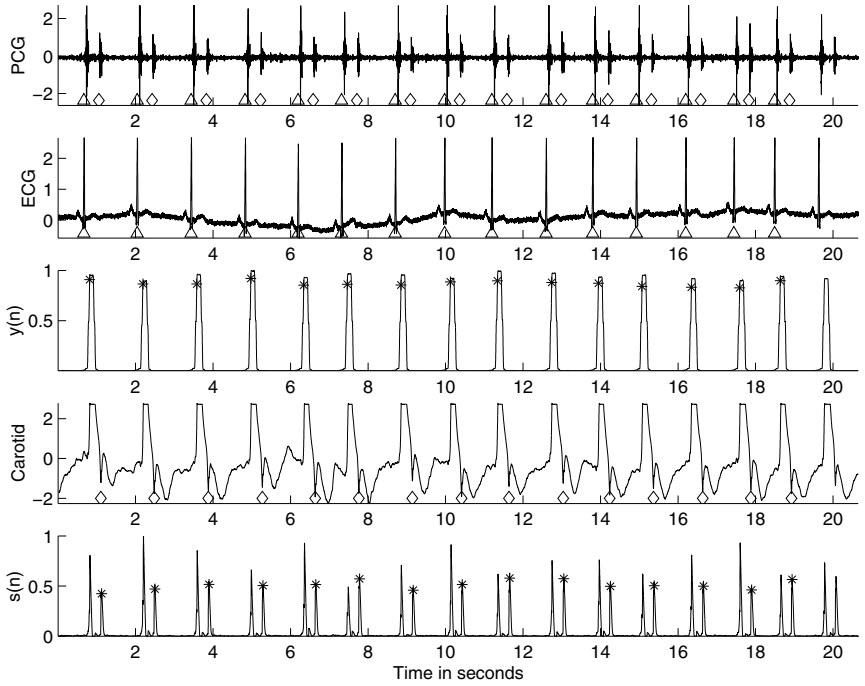

Figure 4.30 Results of segmentation of a PCG signal into systolic and diastolic parts using the ECG and carotid pulse signals for reference. From top to bottom: the PCG signal of a normal subject (male, 23 years); the ECG signal; $y(n)$, the output of the Pan–Tompkins method for detection of the QRS after normalization to the range $(0, 1)$; the carotid pulse signal; $s(n)$, the output of the Lehner and Rangayyan method for detection of the dicrotic notch, normalized to the range $(0, 1)$. The peaks detected in the outputs of the two methods have been identified with $*$ marks. The QRS and the dicrotic notch positions have been marked with the triangle and diamond symbols, respectively. The S1 and S2 positions are marked on the PCG trace with triangles and diamonds, respectively. The last cardiac cycle was not processed.

and hence a widely split S2. Some of the other conditions that cause a wide split in S2 are atrial septal defect, ventricular septal defect, and pulmonary stenosis. Left bundle-branch block leads to delayed left ventricular contraction and aortic valve closure (with reference to the right ventricle and the pulmonary valve), causing A2 to appear after P2, and *reversed splitting* of the two components. Some of the other conditions that could cause reversed splitting of S2 are aortic insufficiency and abnormally early pulmonary valve closure. It is thus seen that identification of A2 and P2 and their temporal relationships could assist in the diagnosis of several cardiovascular defects and diseases.

MacCanon et al. [224] conducted experiments on a dog for direct detection and timing of aortic valve closure. They developed a catheter with an electrical contacting device that could be placed at the aortic valve to detect the exact moment of

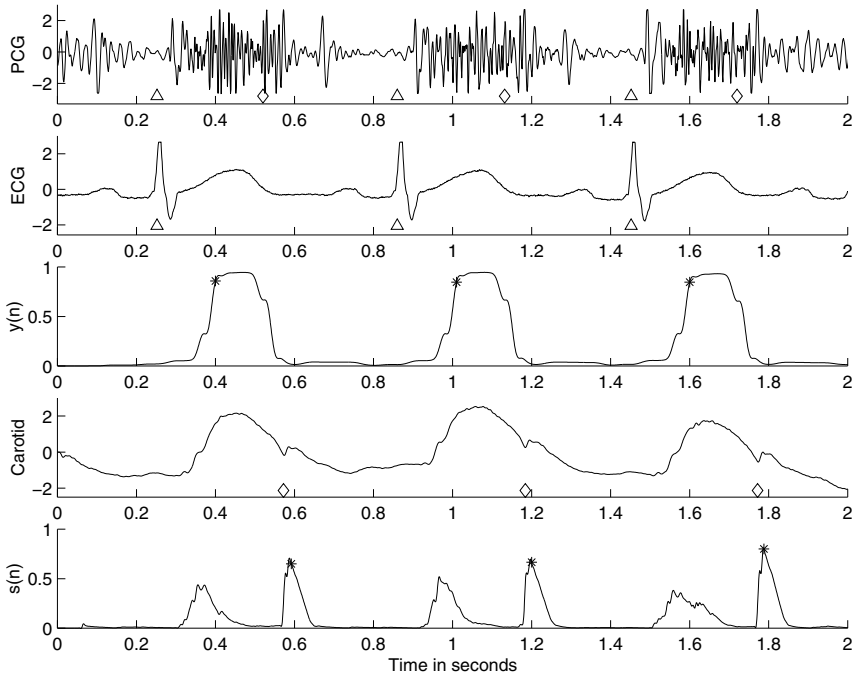

Figure 4.31 Results of segmentation of a PCG signal into systolic and diastolic parts using the ECG and carotid pulse signals for reference. From top to bottom: the PCG signal of a patient (female, 14 months)with a split S2, systolic ejection murmur, and opening snap of the mitral valve; the ECG signal; $y(n)$, the output of the Pan–Tompkins method for detection of the QRS; the carotid pulse signal; $s(n)$, the output of the Lehner and Rangayyan method for detection of the dicrotic notch. The peaks detected in the outputs of the two methods have been identified with $*$ marks. The QRS and the dicrotic notch positions have been marked with the triangle and diamond symbols, respectively. The S1 and S2 positions are marked on the PCG trace with triangles and diamonds, respectively.

closure of the aortic valve. They also measured the aortic pressure and the PCG at the third left intercostal space. It was demonstrated that the aortic valve closes at least $5 - 13$ ms before the incisura appears in the aortic pressure wave (see Figures 1.45 and 1.46 for illustrations of aortic pressure waves recorded from a dog). The conclusion reached was that S2 is caused not by the collision of the valve leaflets themselves, but due to the rebound of the aortic blood column and walls after valve closure. MacCanon et al. also hypothesized that the relative high-frequency characteristics of the incisura and S2 result from elastic recoil of the aortic wall and valve in reaction to the distention by the rebounding aortic blood column.

Stein et al. [225] and Stein and Sabbah [226] conducted experiments in which intracardiac and intraarterial sounds were recorded and analyzed. Their experiments indicated that S2 begins *after* the aortic valve closes. They argued that the intensity of

S2 depends upon, among other factors, the distensibility of the aortic and pulmonary valves; hemodynamic factors that cause the valves to distend and vibrate; viscosity of the blood and its ability to inhibit diastolic valve motion; and the configuration of the aorta, the pulmonary artery, and the ventricles. It was demonstrated that the pulmonary valve, due to its larger surface area than that of the aortic valve, is more distensible and hence produces a larger sound than the aortic valve even for the same pressure gradient across the valve. In the case of pulmonary hypertension, it was argued that the pulmonary valve would distend further at a higher speed: the rate of development of the diastolic pressure gradient across the closed pulmonary valve would be higher than in normal cases.

Problem: *Given that S2 is made up of an aortic component A2 and a pulmonary component P2 with variable temporal relationships, propose a method to detect only A2.*

Solution: We have seen in the preceding section how the dicrotic notch in the carotid pulse signal may be used to detect the beginning of S2. The technique is based upon the direct relationship between aortic valve closure and the aortic incisura, and consequently the dicrotic notch, as explained above. Now, if we were to detect and segment S2 over several cardiac cycles and several respiratory cycles, we could perform synchronized averaging of S2. A2 should appear at the same instant in every S2 segment and should be strengthened by the synchronized averaging process. P2, on the other hand, would appear at different times, and should be cancelled out (suppressed) by the averaging process.

Figure 4.32 shows segments of duration $300\ ms$ containing S2 segmented from nine successive cardiac cycles of the PCG of a patient with atrial septal defect. The PCG signal was segmented using the ECG and carotid pulse signals for reference in a method similar to that illustrated in Figures 4.30 and 4.31. The PCG signal was recorded at the second left intercostal space, which is closer to the pulmonary area than to the aortic area. The nine S2 segments clearly show the fixed timing of A2 and the variable timing of P2. The last plot is the average of S2 segments extracted from 21 successive cardiac cycles. The averaged signal displays A2 very well, while P2 has been suppressed.

The detection of A2 could have been better had the PCG been recorded at the aortic area, where A2 would be stronger than P2. Once A2 is detected, it could be subtracted from each S2 record to obtain individual estimates of P2. Sarkady et al. [227], Baranek et al. [228], and Durand et al. [229] proposed averaging techniques as above with or without envelope detection (but without the use of the carotid pulse); the methods were called aligned ensemble averaging to detect wavelets or coherent detection and averaging.

Nigam and Priemer [230] proposed a method to extract the A2 and P2 components from S2 via blind source separation (BSS) techniques. Xu et al. [231, 232] used time-frequency representations of transient nonlinear chirp signals to model A2 and P2 and derive the A2–P2 splitting interval; they showed that the splitting interval is strongly correlated with the pulmonary arterial pressure measured by catheterization.

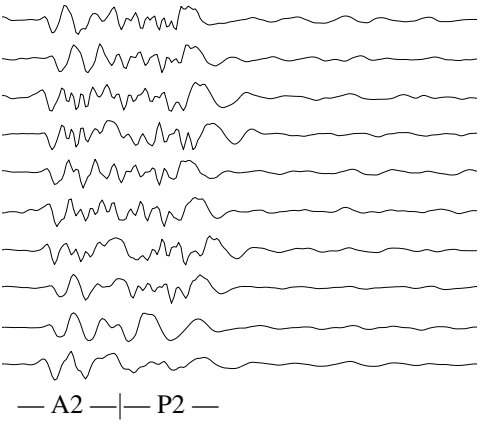

Figure 4.32 Synchronized averaging of S2 to detect A2 and suppress P2. The figure displays nine consecutive segments of S2 (duration = 300 ms) of a patient (female, 7 years) with atrial septal defect leading to a variable split in S2. The trace at the bottom is the average of S2 over 21 consecutive beats. The expected time intervals of A2 and P2 are indicated below the last trace.

4.12 Remarks

We have now established links between the characteristics of certain epochs in a number of biomedical signals and the corresponding physiological or pathological events in the biomedical systems of concern. We have seen how derivative-based operators may be applied to detect QRS complexes in the ECG signal as well as the dicrotic notch in the carotid pulse signal. The utility of correlation and spectral density functions in the detection of rhythms and events in EEG signals was also demonstrated. We have studied how signals with repetitions of a certain event or wavelet, such as a voiced speech signal, may be analyzed using the complex cepstrum and homomorphic filtering. Finally, we also saw how events detected in one signal may be used to locate the corresponding or related events in other signals: the task of detecting S1 and S2 in the PCG was made simpler by using the ECG and carotid pulse signals, where the QRS wave and the dicrotic notch can be detected more readily than the heart sounds themselves.

Event detection is an important step that is required before we may attempt to analyze the corresponding waves or wavelets in more detail. After a specific wave of interest has been detected, isolated, and extracted, methods targeted to the expected characteristics of the event may be applied for specifically focused analysis of the corresponding physiological or pathological event. Analysis of the event is then not hindered or obscured by other events or artifacts in the acquired signal.

4.13 Study Questions and Problems

1. Prove that the ACF $\phi_{xx}(\tau)$ of any stationary function $x(t)$ is maximum at $\tau = 0$. *Hint:* Start with $E[\{x(t+\tau) \pm x(t)\}^2] \geq 0$.
2. For a stationary process x, prove that the ACF is even symmetric, that is, $\phi_{xx}(\tau) = \phi_{xx}(-\tau)$. You may use the expectation or time-average definition of the ACF.
3. Starting with the continuous time-average definition of the ACF, prove that the Fourier transform of the ACF is the PSD of the signal.
4. What are the Fourier-domain equivalents of the ACF and the CCF? Describe their common features and differences. List their applications in biomedical signal analysis.
5. 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 mean and the ACF of $y(t)$ in terms of the statistics of x and η.
6. Derive an expression for the ACF of the signal $x(t) = \sin(\omega_0 t)$. Use the time-average definition of the ACF. From the ACF, derive an expression for the PSD of the signal. Show all steps.
7. A rhythmic episode of a theta wave in an EEG signal is approximated by a researcher to be a sine wave of frequency $5\ Hz$. The signal is sampled at $100\ Hz$. Draw a schematic representation of the ACF of the episode for delays up to $0.5\ s$. Label the time axis in samples and in seconds. Draw a schematic representation of the PSD of the episode. Label the frequency axis in Hz.
8. The values of a signal sampled at $100\ Hz$ are given by the series
   $\{0, 0, 0, 0, 10, 10, 10, 0, 0, 0, 0, 0, 5, 5, 5, 0, 0, 0, 0, 0, -3, -3, -3, 0, 0, 0\}$.
   An investigator performs template matching with the pattern $\{0, 5, 5, 5, 0\}$. The first sample in each array stands for zero time. Plot the output of the template matching operation and interpret the result. Label the time axis in seconds.
9. A biphasic signal $x(n)$ is represented by the series of samples
   $x(n) = \{0, 1, 2, 1, 0, -1, -2, -1, 0\}$ for $n = 0, 1, 2, \ldots, 8$. (a) Draw a plot of $x(n)$. (b) Compose a signal $y(n)$ defined as $y(n) = 3x(n) + 2x(n-12) - x(n-24)$. Draw a plot of $y(n)$. (c) Design a matched filter to detect the presence of $x(n)$ in $y(n)$. Explain how the impulse response $h(n)$ and the frequency response $H(\omega)$ of the filter are related to $x(n)$. Plot $h(n)$. (d) Compute the output of the filter and plot it. Interpret the output of the filter.
10. A researcher uses the derivative operator (filter) specified as $w(n) = x(n) - x(n-1)$, where $x(n)$ is the input and $w(n)$ is the output. The result is then passed through the MA filter $y(n) = \frac{1}{3}[w(n) + w(n-1) + w(n-2)]$, where $y(n)$ is the final output desired. (a) Derive the transfer functions (in the z-domain) of the two filters individually as well as that of the combination. (b) Does it matter which of the two filters is placed first? Why (not)? (c) Derive the impulse response of each filter and that of the combination. Plot the three signals. (d) The signal described by the samples
    $\{0, 0, \ldots, 0, 6, 6, 6, 6, 6, 6, 6, 0, 0, \ldots\}$ is applied to the system. Derive the values of the final output signal. Explain the effect of the operations on the signal.

11. You have been hired to develop two software packages for the analysis of 10-channel EEG signals for the following purposes: (a) Detection of the presence of the alpha rhythm: (i) in any one single channel and (ii) jointly in a pair of left-right channels. (b) Detection of spike-and-wave complexes of a prespecified shape in any channel.

    Design two signal processing packages to address the two problems mentioned above. For each package, provide the following details: (i) A schematic block diagram representing the various algorithms or signal processing steps that you recommend. Include at least three distinct and nontrivial procedures in your design. (ii) Explain each block or part of your package. Describe the reason or logic behind your recommendation. (iii) Write at least two nontrivial equations related to your procedures. (iv) Provide graphical illustrations representing a typical EEG signal as it is processed by each block of your package.

12. The template of a signal is specified by the series of samples $\{2, 3, -1\}$ for $n = 0, 1, 2$. (a) Design a matched filter to detect the signal. Give the impulse response and transfer function of the filter. (b) The signal specified by the series of samples $\{1, -1, -2, 4, 6, -2, 1, -1, 0\}$ is applied to the matched filter. Compute the output and explain its characteristics.

13. Describe the significance of the P wave in the analysis of ECG signals. Describe a method for the detection of P waves in an ECG signal. Explain the purpose and reasoning behind each step of your algorithm. Give at least one nontrivial equation representing a procedure in your algorithm. Draw schematic sketches representing a sample input signal and the corresponding output at each stage of your method.

14. Draw a block diagram representing the various steps in the Pan–Tompkins method to detect QRS complexes in ECG signals. Explain the purpose and nature of each step in the procedure, including detection of the peaks in the output corresponding to the QRS complexes. No equations are required in your answer to this question.

    Draw a schematic sketch of a noisy ECG signal including three cardiac cycles with high-frequency and low-frequency noise, and illustrate how it is modified by each step. Explain how the result could be used to measure the heart rate of a patient.

15. A researcher in signal processing seeks your help to design filters to obtain a smoothed estimate of the first derivative (difference) of a signal digitized at the sampling rate of 200 Hz. Help the researcher with the following: (a) Give the difference equation and impulse response of a filter to compute the first derivative (difference) of the signal. (b) Give the transfer function of the derivative filter. Derive the magnitude and phase parts of the frequency response of the filter. (c) Draw a sketch of the magnitude of the frequency response of the derivative filter. Explain the nature of the filter by deriving the gain at 0 Hz and 100 Hz. (d) Give the difference equation and impulse response of a filter to compute the average of the current input sample and the previous input sample (that is, an MA filter of length 2). (e) Give the transfer function of the MA filter. Derive the magnitude and phase parts of the frequency response of the filter. (f) Draw a sketch of the magnitude of the frequency response of the MA filter. Explain the nature of the filter by deriving the gain at 0 Hz and 100 Hz. (g) Derive the impulse response and the difference equation of a combined filter with the derivative and MA filters in series. (h) Derive the transfer function and plot the pole–zero diagram of the combined filter. (i) Derive the magnitude and phase parts of the frequency response of the combined filter. Draw a sketch of the magnitude of the frequency response of the filter. Explain the nature of the filter by deriving the gain at 0 Hz and 100 Hz. (j) Draw a signal-flow diagram of the combined filter.

16. Describe a procedure for the detection of QRS complexes in an ECG signal. Give a schematic block diagram of the important steps in the procedure. Include at least four steps. Give one suitable equation for each step of your algorithm either in the discrete-time domain as a difference equation or in the z-domain as a transfer function. Sketch an ECG signal with at least two beats and show the corresponding output at the various steps of your procedure. Explain the effect of each step of the procedure on the signal.

4.14 Laboratory Exercises and Projects

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

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

1. Implement the Pan–Tompkins method for QRS detection. You may employ a simple threshold-based method to detect QRS complexes as the procedure will be run off-line.

   Apply the procedure to the signals in the files ECG3.dat, ECG4.dat, ECG5.dat, and ECG6.dat, sampled at a rate of 200 Hz (see the file ECGS.m). Compute the averaged heart rate and QRS width for each record. Verify your results by measuring the parameters visually from plots of the signals.

2. The files eeg1-xx.dat (where xx indicates the channel name) give eight simultaneously recorded channels of EEG signals with the alpha rhythm. (You may read the signals using the MATLAB® program in the file eeg1.m.) The sampling rate is 100 Hz per channel. Cut out a portion of a signal with a clear presence of the alpha rhythm for use as a template or reference signal. Perform cross-correlation of the template with running (short-time) windows of various channels and study the use of the results for the detection of the presence of the alpha rhythm.

3. The files eeg2-xx.dat (where xx indicates the channel name) give ten simultaneously recorded channels of EEG signals with spike-and-wave complexes. (You may read the signals using the MATLAB® program in the file eeg2.m.) The sampling rate is 100 Hz per channel. Cut out one spike-and-wave complex from any EEG channel and use it as a template. Perform template matching by cross-correlation or by designing a matched filter. Apply the procedure to the same channel from which the template was selected as well as to other channels. Study the results and explain how they may be used to detect spike-and-wave complexes.

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 (adult male) and pec52.dat (male subject, 23 years) 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 (female, 14 months).

   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 detect the onset of S1 and S2 in the PCG channel. What are the corrections required to compensate the delays between the corresponding events in the three channels?