

8. Mo, L. Y., and Cobbold, R. S. C., Speckle in C. W. Doppler ultrasound spectra: a simulation study, *IEEE Trans. Ultrason. Ferroelec. Freq. Control UFFC-33*:747-753, 1986.
9. Bascom, P. A. J., Cobbold, R. S. C., and Roelofs, B. H. M., Influence of spectral broadening on C. W. Doppler ultrasound spectra: a geometric approach, *Ultrasound Med. Biol.* **12**:387-395, 1986.
10. Reddy, S. M., and Kirlin, R. L., Spectral analysis of auditory evoked potentials with pseudorandom noise excitation, *IEEE Trans. Biomed. Eng.* **BME-26**(8):479-487, 1979.
11. Hampton, R. L., Experiments using pseudorandom noise, *Simulation* **4**:246-295, 1965.
12. Goldstein, R. F., "Electroencephalic Audiometry in Modern Developments in Audiology" (J. Berger, Ed.), Academic Press, New York, 1973.
13. Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Barger, A. C., and Cohen, R. J., Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat to beat cardiovascular control, *Science* **213**:220-222, 1981.
14. DeBoer, R. W., Karemaker, J. M., and Strackee, J., Comparing spectra of a series of point events particularly for heart rate variability spectra, *IEEE Trans. Biomed. Eng.* **BME-31**:384-387, 1984.
15. Eckberg, D. L., Human sinus arrhythmia as an index of vagal cardiac outflow, *J. Appl. Physiol.* **54**(4):961-966, 1983.
16. Craelius, W., Akay, M., and Tangella, M., Heart rate variability as an index of autonomic imbalance in patients with recent myocardial infarction, *Med. Biol. Comp.* **30**:385-388, 1992.
17. Myers G. A., Martin, G. J., Magid, N. M., Barnett, P. S., Schaad, J. W., Weiss, J. S., Lesch, M., and Singer, D. H., Power spectral analysis of heart rate variability in sudden cardiac deaths, *IEEE Trans. Biomed. Eng.* **BME-33**:1149-1158, 1986.

Biomedical Signal Processing  
M. Akay

## CHAPTER 6

# Cepstrum Analysis

In this chapter, a tutorial review of the cepstrum method is provided. While details of the power and complex cepstra are discussed, extensive derivations of the formulae are given elsewhere [1-7]. In addition to cepstrum signal processing, topics for review include biomedical applications in the areas of electrocardiogram (ECG) and heart sound signal analyses [1-7] and speech signal processing [2].

Previous studies suggest that cepstrum analysis is well suited to data which consist of wavelets [1]. This is true even if the shapes of the wavelets are not known prior to analysis. For instance, the power cepstrum was successfully applied in radar analysis, where the arrival time of the main wavelet was determined by reducing interference [4], and in marine exploration, where source depth was determined and the ocean bottom was mapped [4]. Considerable emphasis is given in this chapter to cepstrum applications in medicine, including diastolic heart sound analysis for the detection of coronary artery disease, ECG pattern classification [10,13], and speech signal decomposition for theoretical as well as bandwidth compression application purposes [2].

The cepstrum method serves as an alternative approach to linear prediction in that it does not make any assumption regarding the characteristics of the data sequence. Bogert *et al.* [1] developed the cepstrum approach to find echo arrival times in a composite signal by decomposing the nonadditive constituents. The term cepstrum represents the power spectrum; it is defined as a function of pseudo-time,  $t$ , the spectral ripple frequency or *quefrency*. The cepstrum terms defined by Bogert *et al.* are

summarized below [1,4]:

frequency	quefrency
spectrum	cepstrum
phase	saphe
amplitude	gamnitude
filtering	liftering
harmonic	rahmonic
period	repiod

Throughout the chapter, both terms are used so as not to confuse readers.

## 6.1 The Cepstra

Cepstrum analysis is concerned with the deconvolution of two signal types: the fundamental (basic) wavelet and a train of impulses (excitation function) [1,4]. The composite signal can be represented in terms of power, complex, or phase cepstra. In this chapter, emphasis is placed on the power and complex cepstra. Readers interested in phase cepstra are referred to [1,3,4].

## 6.2 The Power Cepstrum

The power cepstrum was first described and used by Bogert *et al.* [1] in 1963. The purpose of the study was to determine echo arrival times in a composite signal since the delayed echoes appear as ripples in the logarithmic spectrum of the input data sequence  $x(n)$ . In practice, the power cepstrum is an effective tool provided that the frequencies of the basic wavelet and excitation function do not overlap.

The power cepstrum of the signal is defined as the square of inverse  $z$ -transform of the logarithm of the magnitude squared of the  $z$ -transform of the data sequence, which can be written as

$$x_{pc}(nT) = (z^{-1}\{\log|X(z)|^2\})^2 \quad (6.1)$$

$$x_{pc}(nT) = \left( \frac{1}{2\pi j} \oint_c \log|X(z)|^2 z^{n-1} dz \right)^2, \quad (6.2)$$

where  $X(z)$  represents the  $z$ -transform of the data sequence  $x(nT)$ . Let us assume that the data sequence consists of two convoluted sequences  $y(nT)$  and  $v(nT)$ , which represent the basic wavelet and excitation function, respectively. The data sequence  $x(nT)$  can be written as

$$x(nT) = y(nT) * v(nT). \quad (6.3)$$

This equation can then be written as the multiplication of the Fourier transform of the two sequences,

$$|X(z)|^2 = |Y(z)|^2 \cdot |V(z)|^2. \quad (6.4)$$

Upon taking the logarithm of both sides of the equation, we obtain

$$\log|X(z)|^2 = \log|Y(z)|^2 + \log|V(z)|^2. \quad (6.5)$$

To further elaborate on the power spectrum analysis, let us assume that the excitation function (signal) is given as

$$v(nT) = \delta(nT) + c\delta(nT - n_0T), \quad (6.6)$$

where  $\delta(n)$  denotes the unit impulse function in a sampled data sequence. On the basis of this equation, Eq. (6.4) can be further written as

$$|X(z)|^2 = |Y(z)|^2[1 + cz^{-n_0}]^2. \quad (6.7)$$

By taking the logarithm of both sides of this equation and substituting  $z = e^{j\omega}$ , we expand Eq. (6.5) as

$$\log|X(e^{j\omega})|^2 = \log|Y(e^{j\omega})|^2 + \log(1 + c^2 + 2c \cos(\omega n_0 T)) \quad (6.8)$$

$$= \log|Y(e^{j\omega})|^2 + \log(1 + c^2) + \log\left(1 + \frac{2c}{1 + c^2} \cos(\omega n_0 T)\right). \quad (6.9)$$

The details of these derivations are described elsewhere [3]. It is obvious from Eq. (6.9) that the logarithm of the magnitude squared of the  $z$ -transform of the data sequence  $x(n)$  will have sinusoidal components (ripples). The amplitudes and frequencies of these ripples correspond to the amplitude  $c$  of the excitation function and the time delay,  $n_0 T$ .

By taking the inverse  $z$ -transform of Eq. (6.9), the data sequence  $x(nT)$  can now be expressed in terms of its components. It is assumed that the power cepstra of these components are additive, each corresponding to different frequency bands,

$$x_{pc} = y_{pc} + v_{pc}, \quad (6.10)$$

where  $y_{pc}$  is the power cepstrum of the basic wavelet, and  $v_{pc}$  is the power cepstrum of the excitation signal.

Note that in the above equation the cross-product term was neglected. If the data sequences  $y_{pc}$  and  $v_{pc}$  have different frequency ranges, they can be easily obtained by filtering in the pseudo-frequency domain.

In summary, after taking the inverse  $z$ -transform and obtaining the power cepstrum, the peaks produced by the excitation function can be identified at the quefrequencies (delays) of  $nT$ . Assuming that  $v_{pc}(nT)$  is an

impulse function, the peaks of the power cepstrum can be detected if the  $\log|Y(z)|^2$  is quefrency limited to less than  $n_0T$  and the ripples of the  $\log|Y(z)|^2$  have a period (repoid) less than  $(n_0T)^{-1}$ .

While power cepstrum methods have been successfully applied to biomedical signals including the ECG and diastolic heart sounds, the methods are limited by their failure to maintain the phase information required for precise recovery of analyzed signals.

### 6.3 The Complex Cepstrum

The complex cepstrum is an outgrowth of homomorphic system theory developed by Oppenheim [2]. Although the power cepstrum can be used for detecting echoes, it cannot be used for wavelet recovery since the phase information is lost [2–5]. The complex cepstrum of a data sequence can be defined as the inverse z-transform of the complex logarithm of the z-transform of the data sequence as follows,

$$\hat{x}(nT) = \frac{1}{2\pi j} \int_c \log(X(z)) z^{n-1} dz, \quad (6.11)$$

where  $\hat{x}(nT)$  represents the complex cepstrum and  $X(z)$  represents the z-transform of the data sequence  $x(nT)$ .

Let us assume that the input sequence is the convolution of two sequences as follows,

$$x(nT) = y(nT) * v(nT), \quad (6.12)$$

where  $y(nT)$  represents the basic wavelet and  $v(nT)$  represents the excitation function. This can be written in the z-domain as

$$X(z) = Y(z)V(z). \quad (6.13)$$

The logarithm of Eq. (6.13) is written as

$$\hat{X}(z) = \log X(z) = \log Y(z) + \log V(z). \quad (6.14)$$

The complex cepstrum can then be estimated by the inverse z-transform of this equation,

$$\hat{x}(nT) = \hat{f}(nT) + \hat{g}(nT), \quad (6.15)$$

where  $\hat{x}(nT)$  represents the complex cepstrum of composite signal  $x(n)$ ,  $\hat{f}(nT)$  represents the complex cepstrum of the wavelet component, and  $\hat{g}(nT)$  represents the complex cepstrum of the excitation component.

In an effort to account for the presence of the excitation function in the complex cepstra, we assume that the excitation function  $v(nT)$  is of the

### 6.3 The Complex Cepstrum

form

$$v(nT) = \delta nT + c\delta(nT - n_0T). \quad (6.16)$$

By taking the z-transform and substituting  $z = e^{j\omega}$ , we have

$$V(z) = V(e^{j\omega T}) = 1 + ce^{-j\omega n_0 T} \quad (6.17)$$

and

$$X(e^{j\omega T}) = Y(e^{j\omega T})(1 + ce^{-j\omega n_0 T}). \quad (6.18)$$

Taking the logarithm of both sides of Eq. (6.18),

$$\log X(e^{j\omega T}) = \log Y(e^{j\omega T}) + \log(1 + ce^{-j\omega n_0 T}). \quad (6.19)$$

Where  $c < 1$ , the wavelet component dominates and the data sequence exhibits minimum phase characteristics. This is most evident when Eq. (6.19) is expanded to the form

$$\log X(e^{j\omega T}) = \log Y(e^{j\omega T}) + ce^{-j\omega n_0 T} - \frac{c^2}{2} e^{-2j\omega n_0 T} \dots \quad (6.20)$$

Finally, the complex cepstrum of the data sequence  $x(n)$  is obtained by taking the inverse z-transform of Eq. (6.20),

$$\hat{x}(nT) = \hat{y}(nT) + c\delta(nT - n_0T) - \frac{c^2}{2} \delta(nT - 2n_0T) \dots \quad (6.21)$$

It is evident in Eq. (6.21) that the complex cepstrum includes the complex cepstrum of the wavelet as well as ripples of the excitation function at the positive frequencies ( $n_0T$ ). The amplitudes and frequencies of the ripples correspond to the amplitudes and delays of the excitation function  $v(nT)$ .

For the case  $c > 1$ , where the data sequence exhibits maximum phase characteristics, Eq. (6.19) can be further written as

$$\log(X(e^{j\omega T})) = \log(Y(e^{j\omega T})) - j\omega n_0 T + \log c + \frac{1}{c} e^{j\omega n_0 T} + \dots \quad (6.22)$$

Prior to filtering, the term  $-j\omega n_0 T$  in the above equation should be extracted in order for the excitation function to show maximum phase characteristics, which requires negative ripple frequencies. Note that the amplitudes of the ripples have been attenuated in Eq. (6.22) by the term  $1/c$ . The complex cepstrum can then be written as

$$\hat{x}(nT) = \hat{y}(nT) + (\log c)\delta(nT) + \frac{1}{c} \delta(nT + 2n_0T) + \dots \quad (6.23)$$

After filtering, the linear phase term  $-j\omega n_0 T$  should be included again in order to recover the excitation function [3].

The basic wavelet  $y(nT)$  can be recovered by lowpass filtering the complex cepstrum and taking the inverse  $z$ -transform of the resultant sequence. Note that for effective wavelet recovery it is essential that the frequencies of the wavelet and the excitation function do not overlap. If necessary, the excitation function can also be recovered by first highpass filtering the complex spectra and then taking the inverse  $z$ -transform of the resultant signal.

The recovery process requires that the filtered complex cepstrum be  $z$ -transformed, exponentiated, and inverse  $z$ -transformed. The  $y(nT)$  and  $v(nT)$  sequences can be restored since the necessary phase information has been retained. Figures 6.1a and 6.1b show the overall cepstrum (homomorphic deconvolution) wavelet recovery system. Figure 6.1c shows the typical filters utilized: shortpass (lowpass), longpass (highpass), and notch. These filters are defined in the pseudo-frequency domain and are analogous to lowpass, highpass, and notch filters in the frequency domain.

In summary, the complex cepstrum contains the phase information and therefore allows reconstruction of the composite signal. The power cepstrum can be calculated from the complex cepstrum as follows,

$$x_{pc}(nT) = (\hat{x}(nT) + \hat{x}(-nT))^2, \quad (6.24)$$

where  $x_{pc}(nT)$  represents the power cepstrum and  $\hat{x}(nT)$  represents the complex cepstrum.

### 6.3.1 Phase Unwrapping

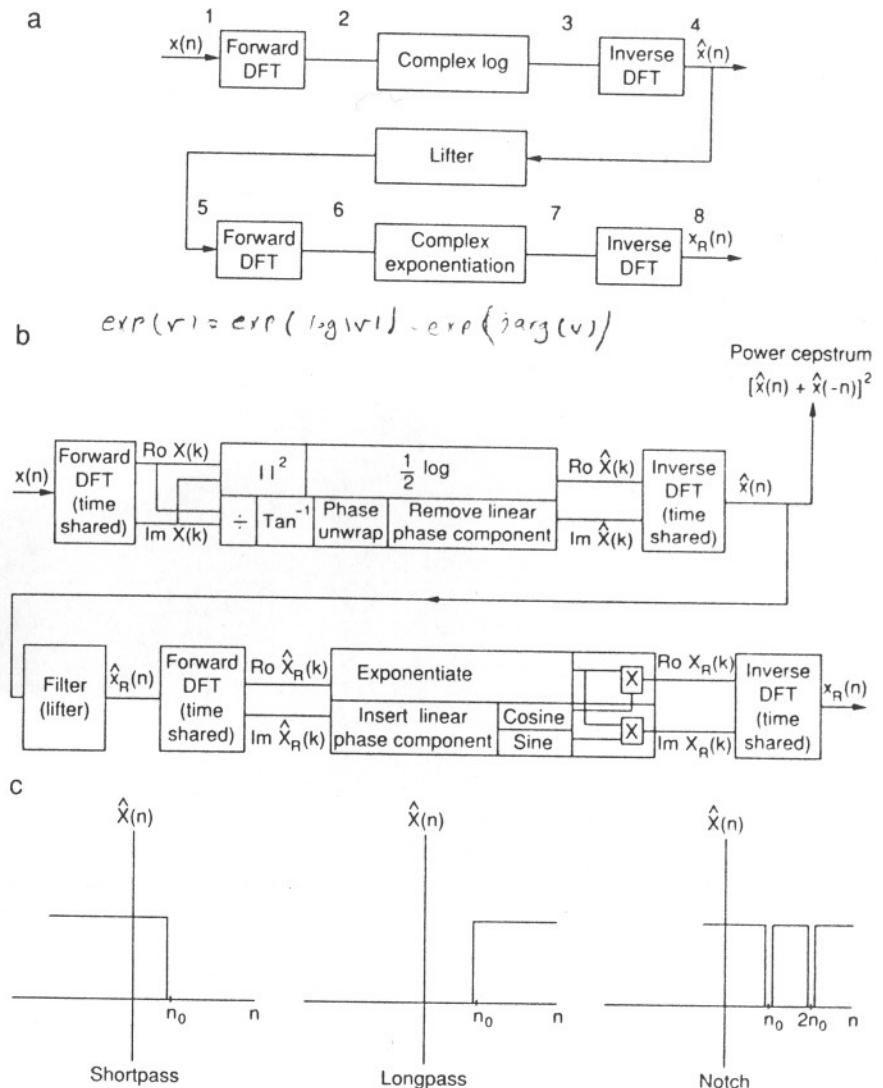
Calculation of the complex cepstrum is however complicated by the fact that it is multivalued. Where computers and commercial software are employed to compute the imaginary part of the complex cepstrum, the principal value is given as [4,6]

$$-\pi \leq \arg|X(z)| \leq \pi. \quad (6.25)$$

The term  $\arg|X(z)|$ , representing the imaginary part of the complex cepstrum, has discontinuities at multiples of  $2\pi$  radians. Since the function is discontinuous, calculation of the  $\log|X(z)|$  is inappropriate. Consequently, the imaginary part of the  $\log(X(z))$  must be continuous, periodic, and thus analytical in some annular region of the  $z$ -plane in order to perform the  $z$ -transformation of the  $\log(X(z))$ . Another important requirement is that the imaginary part of the  $\log(X(z))$  must be an odd function of  $\omega$  since the

$$\log(\sqrt{1+z^2}) \approx \log(z) \text{ for } z \gg 1$$

### 6.3 The Complex Cepstrum



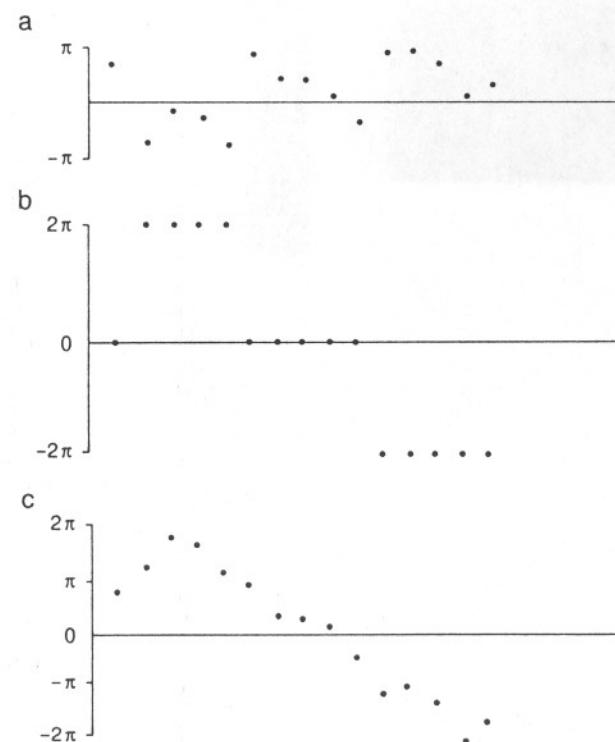
**Fig. 6.1.** Overall wavelet recovery system, also known as homomorphic deconvolution (filtering) or cepstrum system. The DFT is performed by an FFT algorithm.  $x_R(n)$  denotes the recovered wavelet. The input sequence is windowed and then appended with zeros. (a) Simplified block diagram. (b) More detailed block diagram which can be used to process data in real time. (c) Typical lifters for the single-echo, minimum phase ( $c < 1$ ) case where peaks occur at  $n_0$  and multiples thereof. (The notch lifter is sometimes called a comb lifter.) [From Childers et al. [4]].

complex cepstrum of a real function should be real. Therefore, unwrapped phase is required for calculation of the complex cepstrum.

Several approaches for computation of unwrapped phase values are available [4]. However, only two of them are discussed here. The first is based on the fact that the phase is sampled at a very high frequency. This sampling rate is required so that the phase never changes more than  $\pi$  between samples. As shown in Fig. 6.2, the correction term,  $c(k)$ , can be added if the phase difference between samples of the module  $2\pi$  phase sequence  $P(k)$  exceeds  $\pi$ ,

$$c(k) = \begin{cases} c(k-1) - 2\pi & \text{if } P(k) - P(k-1) > \pi \\ c(k-1) + 2\pi & \text{if } P(k) - P(k-1) > \pi \\ c(k-1) & \text{otherwise,} \end{cases}$$

where  $c(0) = 0$ .



**Fig. 6.2.** Phase unwrapping. (a) Phase modulo  $2\pi$ . (b)  $C(k)$ , the correction sequence. (c) Unwrapped phase. [From Childers *et al.* [4]].

Alternatively, the phase can be unwrapped by an adaptive numerical integration procedure proposed by Tribolet [6]. This approach combines information regarding the phase derivative and the principal value of the phase. For each frequency, a set of permissible phase values is defined by adding integer multiples of  $2\pi$  to the principal value of the phase. One of these values may be selected as the unwrapped phase with the help of a phase estimate. This phase estimate is formed by adaptive numerical integration of the phase derivative within a given step interval. The step interval is updated until the phase estimate approaches the permissible phase values [6].

A new approach based on finite length cepstrum modeling was proposed by Nadeu, details of which appear in Ref. [7].

### 6.3.2 Phase Unwrapping Using Adaptive Numerical Integration

The Fourier transform of the data sequence  $x(n)$  is given as

$$X(z) = X_R(z) + jX_I = |X(z)|e^{j\arg[X(z)]}, \quad (6.26)$$

where  $X_R(z)$  represents the real part of  $X(z)$ ;  $X_I(z)$  represents the imaginary part of  $X(z)$ ,  $|X(z)|$  represents the magnitude of the  $X(z)$ , and  $\arg[X(z)]$  represents the phase of the  $X(z)$ . The logarithm of the Fourier transform of the data sequence  $x(n)$  is written as

$$\hat{X}(z) = \log X(z) = \log|X(z)| + j \arg[X(z)]. \quad (6.27)$$

The derivative of  $\hat{X}(z)$  may be found by assuming that Eq. (6.27) has a valid Fourier transform,

$$\frac{\delta \hat{X}(z)}{\delta w} = \frac{\delta \log X(z)}{\delta w} = \frac{\Delta X(z)/\delta w}{X(z)}. \quad (6.28)$$

The derivative of  $\arg[X(z)]$  is obtained from

$$\frac{\delta \arg[X(z)]}{\delta w} = \frac{X_R(z)X'_I(z) - X_I(z)X'_R(z)}{|X(z)|^2}, \quad (6.29)$$

where the first derivative notation represents  $\delta/\delta w$ . Finally, the derivative of  $X(z)$  is written as

$$X'(z) = X'_R(z) + X'_I(z) = -jFT\{nx(n)\}. \quad (6.30)$$

It is evident then that the phase  $\arg[X(z)]$  can be defined as the integration of the derivative  $\arg'[X(z)]$  as follows [6],

$$\arg[X(z)] = \int_0^w \arg'[X(e^{j\eta})] d\eta, \quad (6.31)$$

based on the initial condition  $\arg[X(e^{j0})] = 0$ . The phase function exhibiting these properties is called the *unwrapped phase* function. On the basis of these properties, it is apparent that the unwrapped phase function is a continuous function and can also be defined as an odd function when the phase derivative of the mean is equal to zero. Otherwise, the linear phase component caused by the derivative of the  $\arg[X(z)]$  should be omitted before phase unwrapping.

To calculate the unwrapped phase function, the principal value of the phase is calculated at each frequency  $w_k$ . The limited phase value can then be found from the summation of the principal value and the correction factor  $2\pi l(w_k)$ ,

$$\arg[X(z)] + 2\pi l(w_k), \quad (6.32)$$

where  $\arg[X(z)]$  represents the principal value at the given frequency  $w_k$  and  $l$  is an integer value.

$$\hat{\arg}[X(e^{jw_k})] = \arg[X(e^{jw_k})] + 2\pi l(w_k). \quad (6.33)$$

The correction factor  $l(\omega_k)$  at a given frequency  $\omega_k$  can be estimated where  $\hat{\arg}[X(e^{jw_k})]$  represents the unwrapped phase function at frequency  $w_k$  by applying trapezoidal numerical integration to the phase derivative. The current value of the phase estimate can be calculated by utilizing the previous estimate of the phase value.

*Step 1.*

$$\begin{aligned} \hat{\arg}[X(e^{j\omega_k})] &= \hat{\arg}[X(e^{j\omega_{k-1}})] + \frac{\omega_k - \omega_{k-1}}{2} \\ &\quad \times [\arg'[X(e^{j\omega_{k-1}})] + \arg'[X(e^{j\omega_k})]]. \end{aligned} \quad (6.34)$$

Equation (6.34) improves as the step increment  $\Delta\omega = \omega_k - \omega_{k-1}$  becomes small.

*Step 2.* The correction factor  $l(\omega_k)$  at  $\omega_k$  can be assumed to be consistent if it falls into a predetermined range of one of the acceptable correction factors  $2\pi l(\omega_k)$  at  $\omega_k$ ,

$$|\hat{\arg}[X(e^{j\omega_k})] - \arg[X(e^{j\omega_k})] + 2\pi l(\omega_k)| < \text{THR} < \pi. \quad (6.35)$$

The resultant  $l(\omega_k)$  can be used in Eq. (6.33) to form the unwrapped phase at  $w_k$ . The unwrapped phase at  $w_{k+1}$  will be estimated using the recently unwrapped phase at  $w_k$ .

This process continues until all the unwrapped phase values have been determined.

## 6.4 Biomedical Applications of Cepstrum Analysis

### 6.4.1 Analysis of the ECG Signal Using the Cepstrum Technique

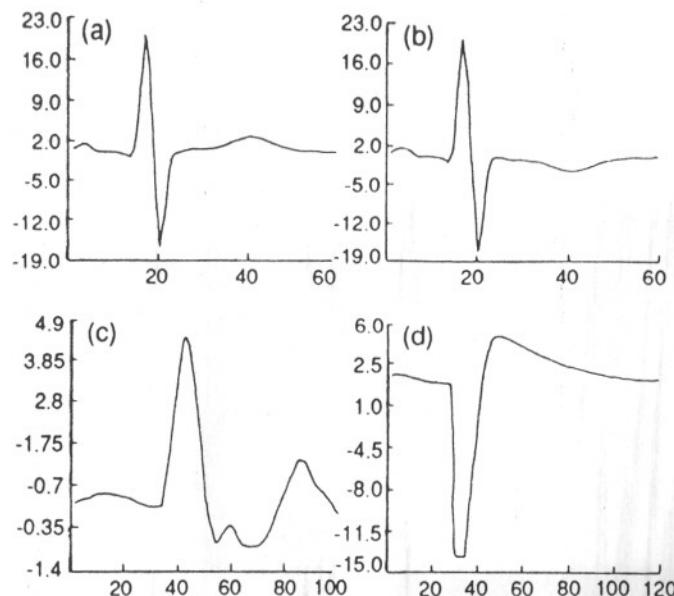
The analysis of ECG signals in the time and frequency domains has been utilized in clinical diagnosis and is well documented in the literature [8,9]. In the time domain analysis, the amplitudes and duration of the P, QRS, and T complexes have been used to construct the feature vector. Pattern recognition techniques were then applied to the feature vector for classification purposes [8,9]. However, extraction of the feature vector is often complicated by the presence of background noise and other artifacts interfering with the pattern.

In the frequency domain, analysis of the ECG signal has often resulted in unsuccessful discrimination among pathological states. In an effort to produce more conclusive spectral results, Amazeen *et al.* [9] utilized the phase information for distinguishing normal from abnormal ECGs.

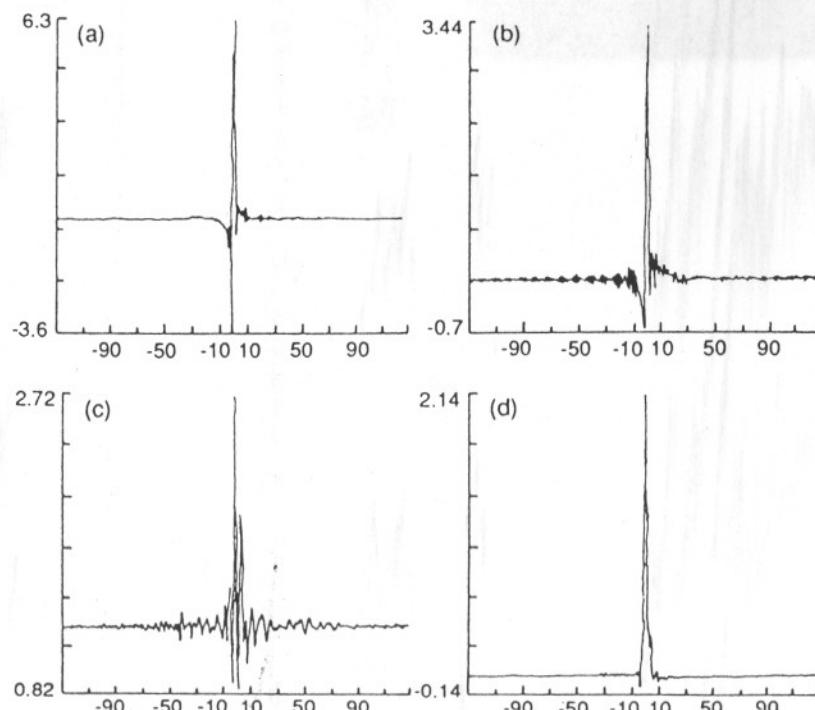
Murthy *et al.* [10] proposed homomorphic filtering and cepstrum analysis. Here we summarize their results and refer the reader to the Ref. [10] for details. The ECG signal was taken as the output of a system driven by excitation functions. Note that the ECG signal represented the convolution of the basic wavelet and excitatory functions. The complex cepstrum of the ECG signal was filtered to separate the system (basic wavelet) and the excitation functions. In this approach, the basic wavelet represented the action potential generated by the heart, while the excitation function represented the excitation pattern of the heart muscle during the cardiac cycle. In their study, Murthy *et al.* [10] analyzed ECG recordings that exhibited normal characteristics and recordings showing inverted T-waves and two types of ventricular premature beats as shown in Fig. 6.3.

Figure 6.4 shows the complex cepstra estimated by Eq. (6.11). As noted by Murthy *et al.*, the normal and inverted T-wave ECG patterns in Fig. 6.3 closely resemble one another. However, the complex cepstra of these two signals are clearly distinguishable. The complex cepstra of the two PVC patterns were also easily distinguished from the normal cepstrum. The minimum and maximum phase components of the complex cepstra presented in Fig. 6.4 were calculated as

$$\hat{x}(n) = \hat{x}_{\min}(n) + \hat{x}_{\max}(n), \quad (6.36)$$



**Fig. 6.3.** Typical ECG signals. (a) Normal. (b) Inverted T. (c) PVC1. (d) PVC2. [From Murthy *et al.* [3]].



**Fig. 6.4.** Complex cepstra for signals in Fig. 6.3. (a) Normal. (b) Inverted T. (c) PVC1. (d) PVC2. [From Murthy *et al.* [3]].

where  $\hat{x}(n)$  represents the complex cepstrum of the ECG signal and  $\hat{x}_{\min}(n)$  represents the minimum phase component of the signal  $x(n)$ :

$$\hat{x}_{\min}(n) = \begin{cases} 0, & n < 0 \\ \frac{1}{2}\hat{x}(0), & n = 0 \\ \hat{x}(n), & n > 0. \end{cases}$$

The maximum phase component that  $\hat{x}_{\max}(n)$  represents is defined as

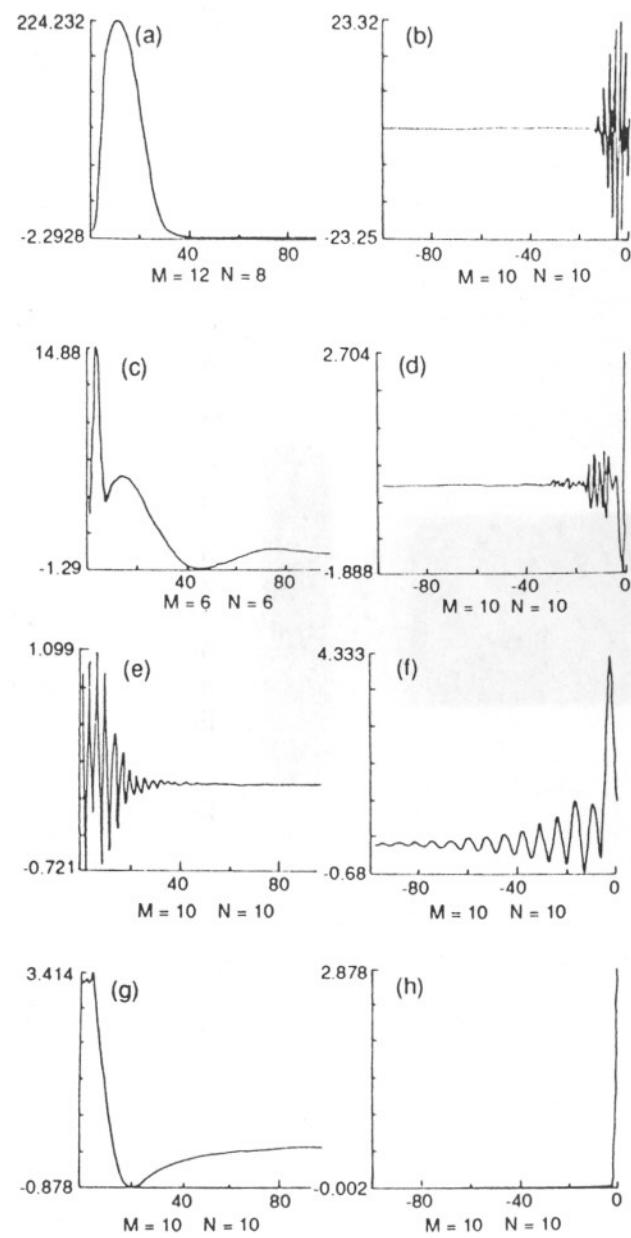
$$\hat{x}_{\max}(n) = \begin{cases} 0, & n > 0 \\ \frac{1}{2}\hat{x}(0), & n = 0 \\ \hat{x}(n), & n < 0. \end{cases}$$

By taking the inverse complex cepstrum of  $\hat{x}_{\min}(n)$  and  $\hat{x}_{\max}(n)$ , the minimum and maximum phase components of the signal  $x(n)$  can be constructed. Figure 6.5 shows the minimum and maximum phase components for the four ECG signals. It is obvious from this figure that the minimum, as opposed to the maximum, phase component acts as an effective decision criterion.

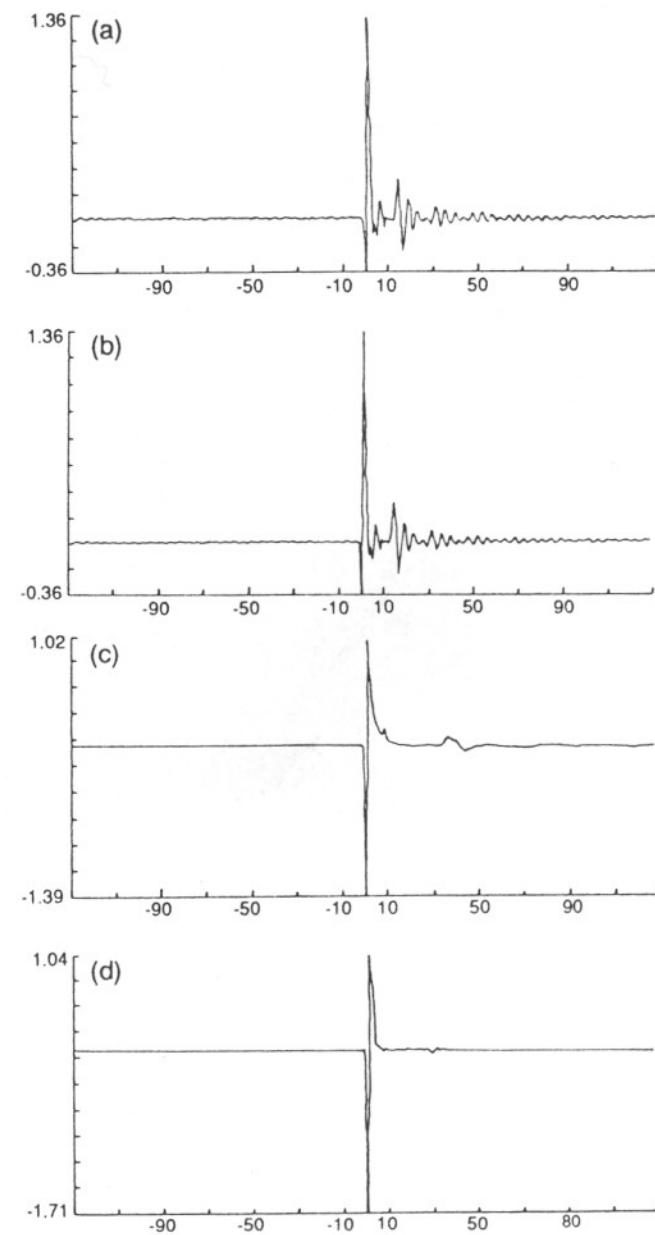
Finally, the basic wavelet of the ECG signal was recovered upon linear filtering of the complex cepstra. Prior to filtering, the complex cepstra were recalculated for an exponentially weighted input data sequence  $x_w(n)$ . The original input sequence  $x(n)$  was multiplied by an appropriate weighting factor  $\alpha^n$ , where  $\alpha$  represents some constant. This exponential weighting was provided in order for the sequence  $x(n)$  to exhibit minimal phase characteristics and ultimately ameliorate the basic wavelet recovered from the reconstructed cepstra [10].

Figure 6.6 shows the complex cepstra of the signals in Fig. 6.3, which have been weighted using the  $\alpha$  value appropriate for each signal. The complex cepstra recalculated on the basis of these weighting factors were lowpass filtered using a window 20 samples in width and centered at the origin. The basic wavelets presented in Fig. 6.7 were reconstructed in the time domain after filtering. The shapes of the wavelets in Figs. 6.7a and 6.7b, corresponding to the normal and inverted T-wave ECG patterns, respectively, similarly resemble that of an action potential generated in the heart muscle. Figure 6.8 shows the excitatory function reconstructed after highpass filtering of the complex cepstra shown in Fig. 6.6.

The results presented in Figs. 6.7 and 6.8 substantiate the existence of two excitation impulses: the first, a ventricular contraction manifested in the QRS complex, and the second, ventricular repolarization evident at the onset of the T-wave. In the case of PVC, the T-wave is absent since repolarization is inhibited by a more prominent premature ventricular

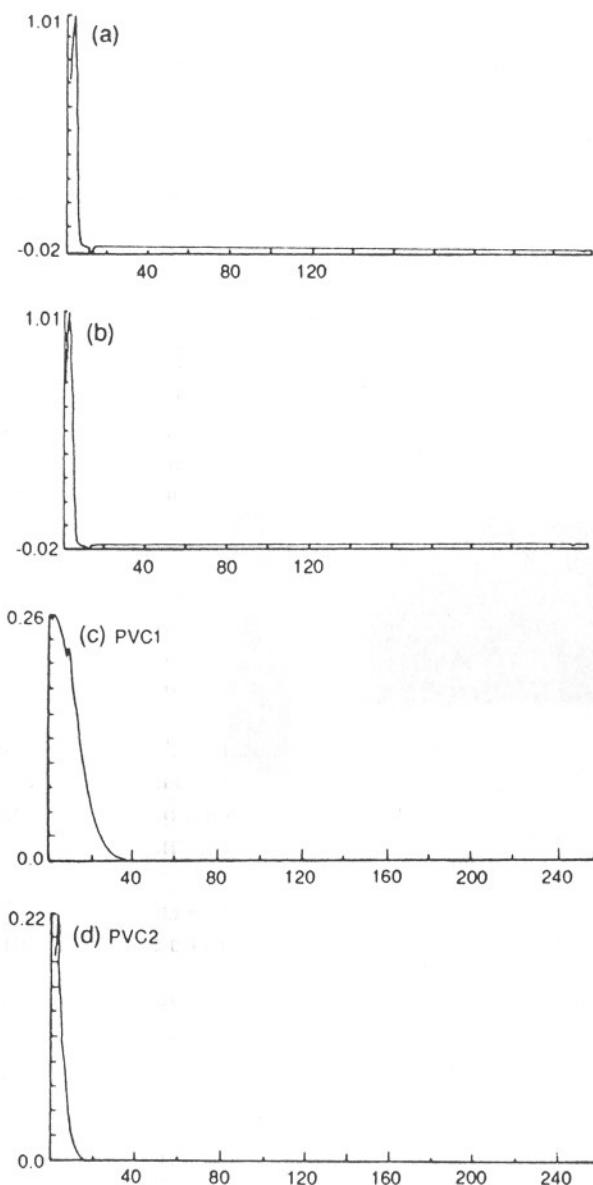


**Fig. 6.5.** (a–h) component signals. (a, c, e, and g) are minimum components of signals in Fig. 6.3a–6.3d. (b, d, f, and h) are maximum components of signals in Fig. 6.3a–d, 3d. [From Murthy *et al.* [10]].

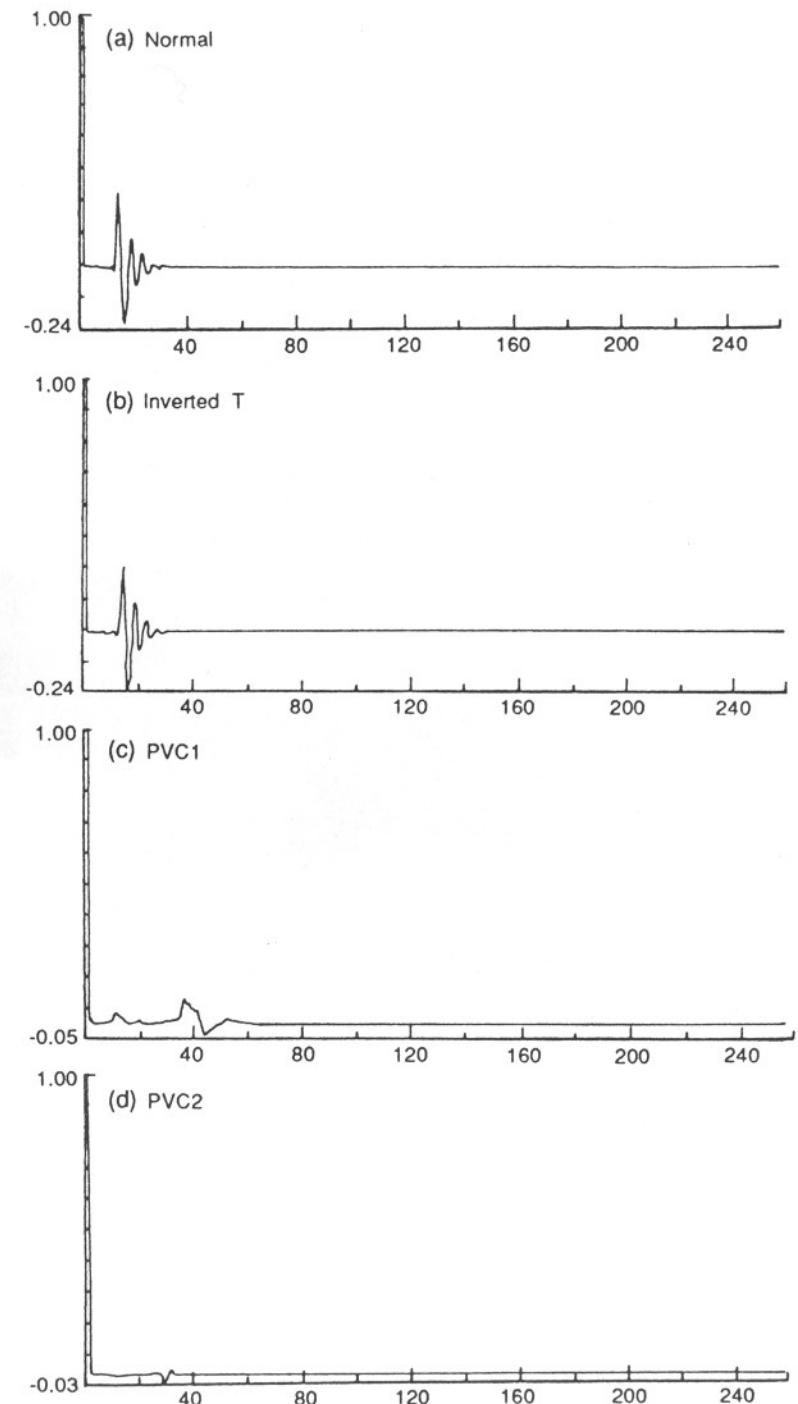


**Fig. 6.6.** Complex cepstra of the signals in Fig. 6.3 exponentially weighted with  $\alpha$ . (a) Normal  $\alpha = 0.76$ . (b) Inverted T,  $\alpha = 0.76$ . (c) PVC1,  $\alpha = 0.91$ . (d) PVC2,  $\alpha = 0.98$ . [From Murthy *et al.* [10]].

## 6. Cepstrum Analysis



**Fig. 6.7.** (a-d) Basic wavelets recovered by lowpass filtering the complex cepstra of Fig. 6.5a–6.5d, with a window  $w = 20$  samples centered at the origin. [From Murthy *et al.* [10]].



**Fig. 6.8.** Excitation function obtained by highpass filtering of the complex cepstra of Fig. 6.5a–6.5d. [From Murthy *et al.* [10]].

contraction. Therefore, the results demonstrate the fact that two excitatory impulses are present in the normal and inverted T-wave ECG patterns, while only one impulse can be identified in the presence of a PVC [10].

In summary, Murthy *et al.* [10] utilized homomorphic filtering and the complex cepstrum method to decompose the input data sequence  $x(n)$  into its wavelet and excitatory function components. The results of their analysis reveal that the basic wavelet components closely resemble action potentials in cardiac muscle fibers while the excitatory functions follow the excitation pattern evident in the heart muscle during the cardiac cycle.

Murthy *et al.* also applied the ARMA method (covered in Chapter 11) to analyze the ECG signal and its minimum and maximum phase components. The details of this study appear elsewhere [10].

#### 6.4.2 Analysis of Diastolic Heart Sound Using the Cepstrum Technique

Previous studies have shown that coronary stenoses produce sounds due to turbulent blood flow in partially occluded arteries [11,12]. During diastole, coronary blood flow is maximum and the sounds associated with turbulent blood flow through partially occluded coronary arteries are the loudest.

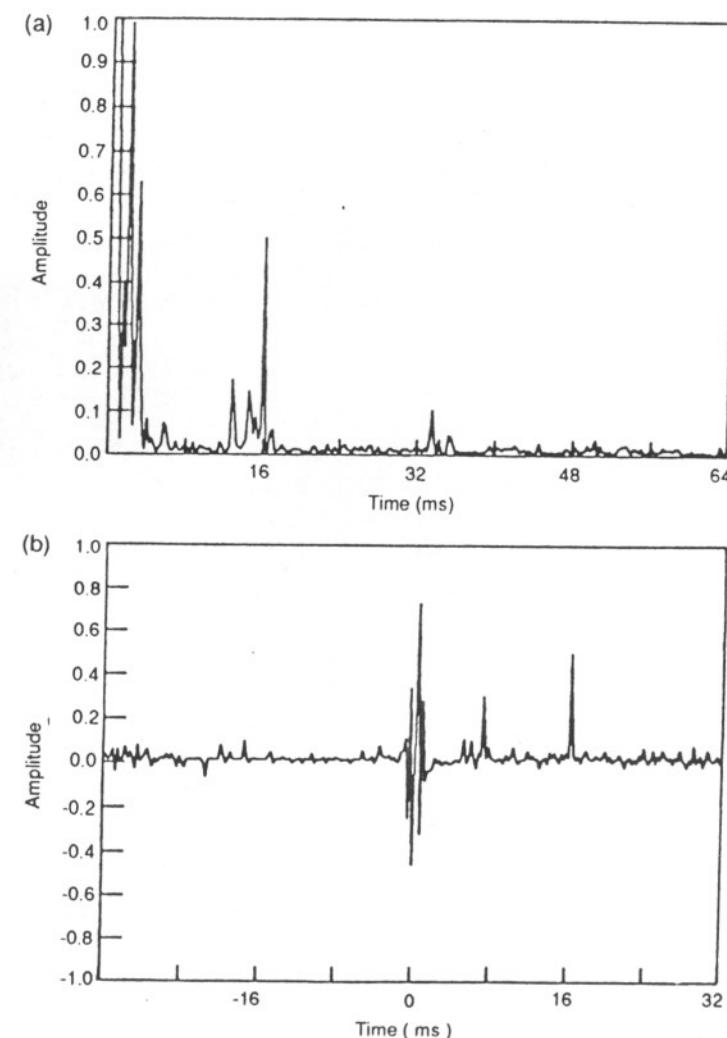
The cepstrum method was proposed by Shen [13] to analyze isolated diastolic heart sounds produced by partially occluded coronary arteries. The recorded diastolic heart sounds were assumed to be a convolution of the turbulent sound caused by coronary occlusions (excitation function) and the sound transmission mechanism (basic wavelet).

Isolated diastolic heart sound records, each 1024 samples in length (sampled at 4 kHz), were initially windowed using a Hanning window. The cepstrum method was applied to the heart sound recordings from five patients (one normal, one angioplasty, one young subject, and two coronary artery disease patients). The recordings were performed in a sound-proof room. Figures 6.9a and 6.10a show the power cepstra of diastolic heart sounds obtained from one diseased and one normal patient; Figs. 6.9b and 6.10b reveal the complex cepstra for the diseased and the normal patient respectively.

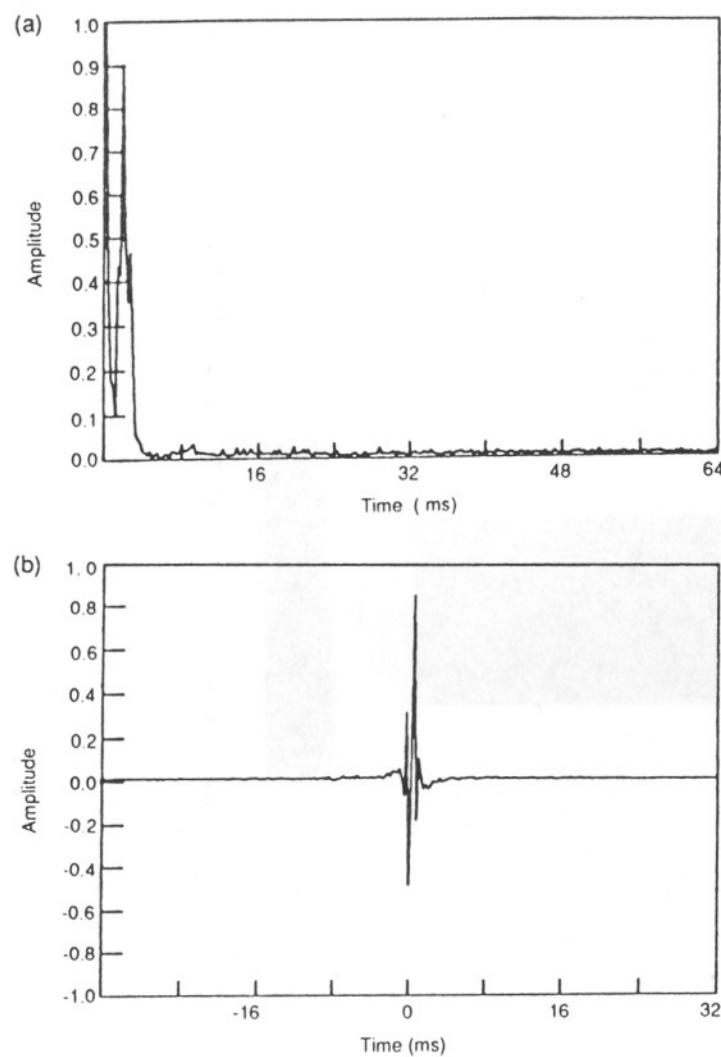
It is obvious from Fig. 6.9a that high time peaks can be found in the power spectrum of the abnormal patient. However, as shown in Fig. 6.10a, the power cepstrum of the normal patient did not include any high time peaks. The presence of the low time components (basic wavelets) in both normal and abnormal patients was attributed to the impulse response of the acoustic transmission system. The presence of high time compo-

nents in the abnormal power cepstra was associated with coronary artery disease.

It is evident upon comparison of Figs. 6.9b and 6.10b that the complex cepstra of the abnormal and normal patients exhibited marked differences at the high time region. The low time components (around the origin)



**Fig. 6.9.** (a) Power cepstrum from one CAD patient (137). (b) Complex cepstrum. [From Shen [13]].



**Fig. 6.10.** (a) Power cepstrum from one pseudonormal subject (125). (b) Complex cepstrum. [From Shen [13]].

appeared to be the same for both cepstra, while the high time components of the abnormal complex spectrum (as was also the case for the power spectrum) were an indication of the presence of coronary artery disease. Further details regarding the application of the cepstrum approach for the detection of coronary artery disease are described elsewhere [13].

#### 6.4.3 Analysis of Speech Signals Using Complex Cepstrum and Linear Filtering

Prolific research in the area of speech signal processing has greatly advanced our understanding of the speech signal waveform [2]. The speech signal may be regarded as the convolution of vocal cord timing, the glottal pulse, and the vocal tract impulse response. A new procedure for decomposition of the speech signal into its constituents has been proposed by Oppenheim and Schafer [2]. Their approach was based on the calculation of the complex cepstrum of the speech waveform, followed by linear filtering of the extracted components and recovery of the components in the time domain. This procedure is illustrated in Fig. 6.11, where  $D(\cdot)$  represents the complex cepstrum, and in Fig. 6.12 where  $D^{-1}$  is the inverse complex cepstrum [2].

The authors define the speech signal  $s(n)$  as the convolution of the pitch  $p(n)$ , the glottal pulse  $g(n)$ , and the vocal tract sequence  $v(n)$ ,

$$s(n) = [p(n) * g(n) * v(n)]w(n), \quad (6.37)$$

where  $w(n)$  represents the window.

The pitch data were taken as a sequence sampled at a rate  $\tau$ . The complex cepstrum of the train of pitch samples was weighted with the window  $w(n)$  such that

$$w_\tau(n) = p(n)w(n) \quad \text{for } n = 0, \pm\tau, \pm 2\tau \quad (6.38)$$

and the complex cepstrum of the pitch  $p(n)$  is

$$\hat{p}(n) = \hat{w}_\tau(n) \quad \text{for } n = 0, \pm\tau, \pm 2\tau. \quad (6.39)$$

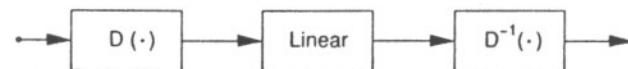
The minimum phase window allows the pitch data to exhibit minimum phase. Therefore, the complex cepstrum of  $p(n)$  is zero for  $n < 0$ .

The complex cepstrum  $\hat{v}(n)$  of the vocal tract is estimated on the basis of a cascade of damped resonators [2],

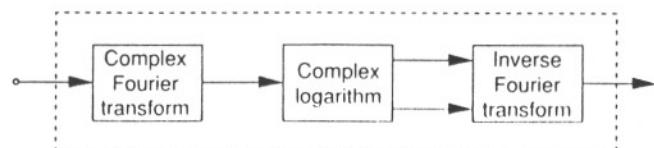
$$\hat{v}(n) = \sum_{i=1}^M \frac{|a_i|^n}{n} \cos \omega_i n \quad \text{for } n > 0, \quad (6.40)$$

$$\hat{v}(n) = 0 \quad \text{for } n < 0 \quad (6.41)$$

where  $a_i$  represents the poles of the  $z$ -transform of  $v(n)$ .



**Fig. 6.11.** Canonic form for homomorphic deconvolution. [From Oppenheim and Schafer [2]].



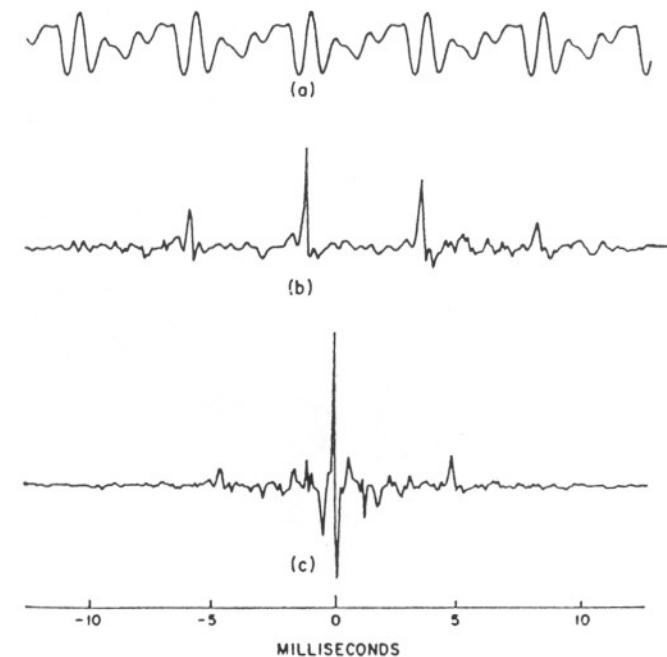
**Fig. 6.12.** Realization of the transformation  $D(\cdot)$  of Fig. 6.11. [From Oppenheim and Schafer [2]].

Finally, the glottal pulse  $g(n)$  was defined as the convolution of the minimum phase sequence  $g_1(n)$  and the maximum phase sequences  $g_2(n)$  ( $g(n) = g_1(n) * g_2(n)$ ). Both the maximum and minimum phase components are included in the convolution since details of the signal characteristics have not been well documented.

As far as speech signal decomposition is concerned, the complex cepstrum of the glottal pulse can be distinguished from the pitch pulse, given its much shorter duration. It has also been observed that the complex cepstrum of the vocal tract signal decays rapidly compared to the complex cepstrum of the pitch pulse [2]. Consequently, Oppenheim and Schafer divided the complex cepstrum into three regions. Any constituent pulse with a pitch period of  $\tau$  or  $n > \tau$ , was assumed to be the pitch pulse. For  $0 \leq n < \tau$ , the contribution to the composite was attributed to the minimal phase component of the glottal pulse and the vocal tract. Finally, any contribution for  $n < 0$  was said to be given by the maximum phase component of the glottal pulse [2].

To recover the three constituent pulses, Oppenheim and Schafer filtered the complex cepstrum as follows: for the pitch pulse period  $\tau$ , only those components of the complex cepstrum for  $n \geq \tau$  were maintained; otherwise, the complex cepstrum was multiplied by zero for  $|n| < \tau$  and by one for  $|n| > \tau$ . The convolution of the minimum phase component  $g_1(n)$  of the glottal pulse and the vocal tract signal was extracted by multiplying the complex cepstrum by zero for  $n < 0$  and  $n \geq \tau$ . The maximum phase component  $g_2(n)$  of the glottal pulse was obtained after multiplying the complex cepstrum by zero for  $n \geq 0$ . The filtering process was succeeded by the inverse  $z$ -transformation of the complex cepstrum in order to recover the pitch pulse, the convolution of the minimal phase portion of the glottal pulse and the vocal tract, and the maximum phase component of the glottal pulse.

Figure 6.13 shows the recovery of pitch for the vowel sound "ah" as in "father." The speech was sampled at 10 kHz and weighted with a Hanning window 2.56 msec in duration.



**Fig. 6.13.** (a) Sample of the vowel "ah." (b) Resulting output due to pitch. (c) Complex cepstrum of (a). [From Oppenheim and Schafer [2]].

Finally, the vocal tract sequences were extracted by matching the logarithmic spectrum of the vocal and glottal pulses with the logarithmic spectrum of a set of ideal cascade resonators [2]. For unabated signal recovery, the authors suggested that the same windowing process be used for calculating the complex cepstrum of the original speech waveform as was used for windowing the vocal tract signal. Details of the analysis using complex cepstrum and linear filtering are described in Ref. [2].

## 6.5 Computer Experiments

1. Create a data sequence consisting of three sinusoids with frequencies of 2, 3, and 5 kHz and a 20-kHz sampling frequency,

$$x(n) = \cos\left(2\pi \frac{f_1}{f_{sr}} n\right) + \cos\left(2\pi \frac{f_2}{f_{sr}} n\right) + \cos\left(2\pi \frac{f_3}{f_{sr}} n\right).$$