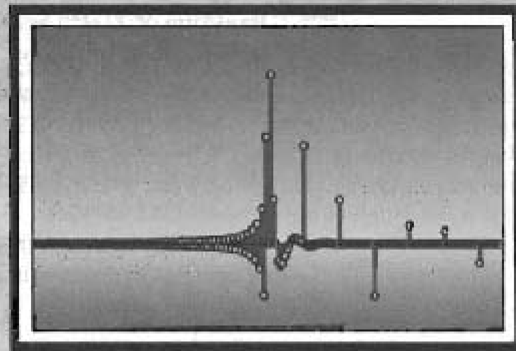
13

Cepstrum Analysis and Homomorphic Deconvolution



13.0 INTRODUCTION

Throughout this text, we have focused primarily on linear signal processing methods. In this chapter, we introduce a class of nonlinear techniques referred to as *cepstrum analysis* and *homomorphic deconvolution*. These methods have proven to be effective and useful in a variety of applications. In addition, they further illustrate the considerable flexibility and sophistication offered by discrete-time signal processing technologies.

In 1963, Bogert, Healy, and Tukey published a paper with the unusual title "The Quefrency Analysis of Time Series for Echoes: Cepstrum, Pseudoautocovariance, Cross-Cepstrum, and Saphe Cracking."(See Bogert, Healy and Tukey, 1963.) They observed that the logarithm of the power spectrum of a signal containing an echo has an additive periodic component due to the echo, and thus, the power spectrum of the logarithm of the power spectrum should exhibit a peak at the echo delay. They called this function the *cepstrum*, interchanging letters in the word *spectrum* because "in general, we find ourselves operating on the frequency side in ways customary on the time side and vice versa." Bogert et al. went on to define an extensive vocabulary to describe this new signal processing technique; however, only the terms cepstrum and quefrency have been widely used.

At about the same time, Oppenheim (1964, 1967, 1969a) proposed a new class of systems called *homomorphic systems*. Although nonlinear in the classic sense, these systems satisfy a generalization of the principle of superposition; i.e., input signals and their corresponding responses are superimposed (combined) by an operation having the same algebraic properties as addition. The concept of homomorphic systems is very general, but it has been studied most extensively for the combining operations of multiplication and convolution, because many signal models involve these operations. The transformation of a signal into its cepstrum is a homomorphic transformation that maps convolution into addition, and a refined version of the cepstrum is a fundamental part of the theory of homomorphic systems for processing signals that have been combined by convolution.

Since the introduction of the cepstrum, the concepts of the cepstrum and homomorphic systems have proved useful in signal analysis and have been applied with success in processing speech signals (Oppenheim, 1969b, Oppenheim and Schafer, 1968 and Schafer and Rabiner, 1970), seismic signals (Ulrych, 1971 and Tribolet, 1979), biomedical signals (Senmoto and Childers, 1972), old acoustic recordings (Stockham, Cannon and Ingebretsen, 1975), and sonar signals (Reut, Pace and Heator, 1985). The cepstrum has also been proposed as the basis for spectrum analysis (Stoica and Moses, 2005). This chapter provides a detailed treatment of the properties and computational issues associated with the cepstrum and with deconvolution based on homomorphic systems. A number of these concepts are illustrated in Section 13.10 in the context of speech processing.

13.1 DEFINITION OF THE CEPSTRUM

The original motivation for the cepstrum as defined by Bogert et al. is illustrated by the following simple example. Consider a sampled signal $x[n]$ that consists of the sum of a signal $v[n]$ and a shifted and scaled copy (echo) of that signal; i.e.,

$$x[n] = v[n] + \alpha v[n - n_0] = v[n] * (\delta[n] + \alpha\delta[n - n_0]). \quad (13.1)$$

Noting that $x[n]$ can be represented as a *convolution*, it follows that the discrete-time Fourier transform of such a signal has the form of a product

$$X(e^{j\omega}) = V(e^{j\omega})[1 + \alpha e^{-j\omega n_0}]. \quad (13.2)$$

The magnitude of $X(e^{j\omega})$ is

$$|X(e^{j\omega})| = |V(e^{j\omega})|(1 + \alpha^2 + 2\alpha\cos(\omega n_0))^{1/2}, \quad (13.3)$$

a real even function of ω. The basic observation motivating the cepstrum was that the logarithm of the product such as in Eq. (13.3) would be a sum of two corresponding terms, specifically

$$\log|X(e^{j\omega})| = \log|V(e^{j\omega})| + \tfrac{1}{2}\log(1 + \alpha^2 + 2\alpha\cos(\omega n_0)). \quad (13.4)$$

For convenience, we define $C_x(e^{j\omega}) = \log|X(e^{j\omega})|$. Also, in anticipation of a discussion in which we will want to stress the duality between the time- and frequency- domains, we substitute $\omega = 2\pi f$ to obtain

$$C_x(e^{j2\pi f}) = \log|X(e^{j2\pi f})| = \log|V(e^{j2\pi f})| + \tfrac{1}{2}\log(1 + \alpha^2 + 2\alpha\cos(2\pi f n_0)). \quad (13.5)$$

There are two components to this real function of normalized frequency f. The term $\log|V(e^{j2\pi f})|$ is due solely to the signal $v[n]$, and the second term, $\log(1 + \alpha^2 + 2\alpha\cos(2\pi f n_0))$ is due to the combination (echoing) of the signal with itself. We can think of $C_x(e^{j2\pi f})$ as a waveform with continuous independent variable f. The part due to the echo will be periodic in f with period $1/n_0$.[1] We are used to the notion

[1]Because $\log(1 + \alpha^2 + 2\alpha\cos(2\pi f n_0))$ is the log-magnitude of a DTFT, it is also periodic in f with period one (2π in ω), as well as $1/n_0$.

that a periodic time waveform has a line spectrum, i.e., its spectrum is concentrated at integer multiples of a common fundamental frequency, which is the reciprocal of the fundamental period. In this case, we have a "waveform" that is a real, even function of f (i.e., frequency). Fourier analysis appropriate for a continuous-variable periodic function such as $C_x(e^{j2\pi f})$ would naturally be the inverse DTFT; i.e.,

$$c_x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} C_x(e^{j\omega})e^{j\omega n}d\omega = \int_{-1/2}^{1/2} C_x(e^{j2\pi f})e^{j2\pi f n}df. \tag{13.6}$$

In the terminology of Bogert et al., $c_x[n]$ is referred to as the *cepstrum* of $C_x(e^{j2\pi f})$ (or equivalently, of $x[n]$ since $C_x(e^{j2\pi f})$ is derived directly from $x[n]$). Although the cepstrum defined as in Eq. (13.6) is clearly a function of a discrete-time index n, Bogert et al. introduced the term "quefrency" to draw a distinction between the cepstrum time domain and that of the original signal. Because the term $\log(1 + \alpha^2 + 2\alpha\cos(2\pi f n_0))$ in $C_x(e^{j2\pi f})$ is periodic in f with period $1/n_0$, the corresponding component in $c_x[n]$ will be nonzero only at integer multiples of n_0, the fundamental quefrency of the term $\log(1 + \alpha^2 + 2\alpha\cos(2\pi f n_0))$. Later in this chapter, we will show that for this example of a simple echo with $|\alpha| < 1$, the cepstrum has the form

$$c_x[n] = c_v[n] + \sum_{k=1}^{\infty}(-1)^{k+1}\frac{\alpha^k}{2k}(\delta[n + kn_0] + \delta[n - kn_0]), \tag{13.7}$$

where $c_v[n]$ is the inverse DTFT of $\log|V(e^{j\omega})|$, (i.e., the cepstrum of $v[n]$), and the discrete impulses involve only the echo parameters α and n_0. It was this result that led Bogert et al. to observe that the cepstrum of a signal with an echo had a "peak" at the echo delay time n_0 that stands out clearly from $c_v[n]$. Thus the cepstrum could be used as the basis for *detecting* echoes. As mentioned above, the strange-sounding terms "cepstrum" and "quefrency" and other terms were created to call attention to a new way of thinking about Fourier analysis of signals wherein the time and frequency domains were interchanged. In the remainder of this chapter, we will generalize the concept of cepstrum by using the *complex* logarithm, and we will show many interesting properties of the resulting mathematical definition. Furthermore, we will see that the complex cepstrum can also serve as the basis for *separating* signals that are combined by convolution.

13.2 DEFINITION OF THE COMPLEX CEPSTRUM

As the basis for generalizing the concept of the cepstrum, consider a stable sequence $x[n]$ whose z-transform expressed in polar form is

$$X(z) = |X(z)|e^{j\angle X(z)}, \tag{13.8}$$

where $|X(z)|$ and $\angle X(z)$ are the magnitude and angle, respectively, of the complex function $X(z)$. Since $x[n]$ is stable, the ROC for $X(z)$ includes the unit circle, and the DTFT of $x[n]$ exists and is equal to $X(e^{j\omega})$. The *complex cepstrum* associated with $x[n]$

is defined to be the stable sequence $\hat{x}[n]$,[2] whose z-transform is

$$\hat{X}(z) = \log[X(z)]. \tag{13.9}$$

Although any base can be used for the logarithm, the natural logarithm (i.e., base e) is typically used and will be assumed throughout the remainder of the discussion. The logarithm of a complex quantity $X(z)$ expressed as in Eq. (13.8) is defined as

$$\log[X(z)] = \log[|X(z)|e^{j\angle X(z)}] = \log|X(z)| + j\angle X(z). \tag{13.10}$$

Since in the polar representation of a complex number the angle is unique only to within integer multiples of 2π, the imaginary part of Eq. (13.10) is not well defined. We will address that issue shortly; for now we assume that an appropriate definition is possible and has been used.

The complex cepstrum exists if $\log[X(z)]$ has a convergent power series representation of the form

$$\hat{X}(z) = \log[X(z)] = \sum_{n=-\infty}^{\infty} \hat{x}[n]z^{-n}, \qquad |z| = 1, \tag{13.11}$$

i.e., $\hat{X}(z) = \log[X(z)]$ must have all the properties of the z-transform of a stable sequence. Specifically, the ROC for the power series representation of $\log[X(z)]$ must be of the form

$$r_R < |z| < r_L, \tag{13.12}$$

where $0 < r_R < 1 < r_L$. If this is the case, $\hat{x}[n]$, the sequence of coefficients of the power series, is what we call the *complex cepstrum* of $x[n]$.

Since we require $\hat{x}[n]$ to be stable, the ROC of $\hat{X}(z)$ includes the unit circle, and the complex cepstrum can be represented using the inverse DTFT as

$$\begin{aligned}\hat{x}[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \log[X(e^{j\omega})]e^{j\omega n}d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} [\log|X(e^{j\omega})| + j\angle X(e^{j\omega})]e^{j\omega n}d\omega.\end{aligned} \tag{13.13}$$

The term complex cepstrum distinguishes our more general definition from the original definition of the cepstrum by Bogert et al. (1963), which was originally stated in terms of the power spectrum of continuous-time signals. The use of the word *complex* in this context implies that the complex logarithm is used in the definition. It does not imply that the complex cepstrum is necessarily a complex-valued sequence. Indeed, as we will see shortly, the definition we choose for the complex logarithm ensures that the complex cepstrum of a real sequence will also be a real sequence.

The operation of mapping a sequence $x[n]$ into its complex cepstrum $\hat{x}[n]$ is denoted as a discrete-time system operator $D_*[\cdot]$; i.e., $\hat{x} = D_*[x]$. This operation is depicted as the block diagram on the left in Figure 13.1. Similarly, since Eq. (13.9) is invertible with the complex exponential function, we can also define the inverse system $D_*^{-1}[\cdot]$

[2] In a somewhat more general definition of the complex cepstrum, $x[n]$ and $\hat{x}[n]$ need not be restricted to be stable. However, with the restriction of stability the important concepts can be illustrated with simpler notation than in the general case.

Figure 13.1 System notation for the mapping and inverse mapping between a signal and its complex cepstrum.

which recovers $x[n]$ from $\hat{x}[n]$. The block diagram representation of $D_*^{-1}[\cdot]$ is shown on the right in Figure 13.1. Specifically, $D_*[\cdot]$ and $D_*^{-1}[\cdot]$ in Figure 13.1 are defined so that if $\hat{y}[n] = \hat{x}[n]$ in Figure 13.1, then $y[n] = x[n]$. In the context of homomorphic filtering of convolved signals to be discussed in Section 13.8, $D_*[\cdot]$ is called the *characteristic system* for convolution.

As introduced in Section 13.1, the cepstrum $c_x[n]$ of a signal[3] is defined as the inverse Fourier transform of the logarithm of the magnitude of the Fourier transform; i.e.,

$$c_x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(e^{j\omega})|e^{j\omega n}d\omega. \tag{13.14}$$

Since the Fourier transform magnitude is real and nonnegative, no special considerations are involved in defining the logarithm in Eq. (13.14). By comparing Eq. (13.14) and Eq. (13.13), we see that $c_x[n]$ is the inverse transform of the real part of $\hat{X}(e^{j\omega})$. Consequently $c_x[n]$ is equal to the conjugate-symmetric part of $\hat{x}[n]$; i.e.,

$$c_x[n] = \frac{\hat{x}[n] + \hat{x}^*[-n]}{2}. \tag{13.15}$$

The cepstrum is useful in many applications, and since it does not depend on the phase of $X(e^{j\omega})$, it is much easier to compute than the complex cepstrum. However, since it is based on only the Fourier transform magnitude, it is not invertible, i.e., $x[n]$ cannot in general be recovered from $c_x[n]$, except in special cases. The complex cepstrum is somewhat more difficult to compute, but it is invertible. Since the complex cepstrum is a more general concept than the cepstrum, and since the properties of the cepstrum can be derived from the properties of the complex cepstrum using Eq. (13.15), we will emphasize the complex cepstrum in this chapter.

The additional difficulties encountered in defining and computing the complex cepstrum are worthwhile for a variety of reasons. First, we see from Eq. (13.10) that the complex logarithm has the effect of creating a new Fourier transform whose real and imaginary parts are $\log|X(e^{j\omega})|$ and $\angle X(e^{j\omega})$, respectively. Thus, we can obtain Hilbert transform relations between these two quantities when the complex cepstrum is causal. We discuss this point further in Section 13.5.2 and see in particular how it relates to minimum-phase sequences. A second more general motivation, developed in Section 13.8, stems from the role that the complex cepstrum plays in defining a class of systems for separating and filtering signals that are combined by convolution.

13.3 PROPERTIES OF THE COMPLEX LOGARITHM

Since the complex logarithm plays a key role in the definition of the complex cepstrum, it is important to understand its definition and properties. Ambiguity in the definition

[3] $c_x[n]$ is also referred to as the *real cepstrum* to emphasize that it corresponds to only the real part of the complex logarithm.

of the complex logarithm causes serious computational issues. These will be discussed in detail in Section 13.6. A sequence has a complex cepstrum if the logarithm of its z-transform has a power series expansion, as in Eq. (13.11), where we have specified the ROC to include the unit circle. This means that the Fourier transform

$$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\angle X(e^{j\omega}) \tag{13.16}$$

must be a continuous, periodic function of ω, and consequently, both $\log|X(e^{j\omega})|$ and $\angle X(e^{j\omega})$ must be continuous functions of ω. Provided that $X(z)$ does not have zeros on the unit circle, the continuity of $\log|X(e^{j\omega})|$ is guaranteed, since $X(e^{j\omega})$ is assumed to be analytic on the unit circle. However, as previously discussed in Section 5.1.1, $\angle X(e^{j\omega})$ is in general ambiguous, since at each ω, any integer multiple of 2π can be added, and continuity of $\angle X(e^{j\omega})$ is dependent on how the ambiguity is resolved. Since $\text{ARG}[X(e^{j\omega})]$ can be discontinuous, it is generally necessary to specify $\angle X(e^{j\omega})$ explicitly in Eq. (13.16) as the unwrapped (i.e., continuous) phase curve $\arg[X(e^{j\omega})]$.

It is important to note that if $X(z) = X_1(z)X_2(z)$, then

$$\arg[X(e^{j\omega})] = \arg[X_1(e^{j\omega})] + \arg[X_2(e^{j\omega})]. \tag{13.17}$$

A similar additive property will not hold for $\text{ARG}[X(e^{j\omega})]$, i.e., in general,

$$\text{ARG}[X(e^{j\omega})] \neq \text{ARG}[X_1(e^{j\omega})] + \text{ARG}[X_2(e^{j\omega})]. \tag{13.18}$$

Therefore, in order that $\hat{X}(e^{j\omega})$ be analytic (continuous) and have the property that if $X(e^{j\omega}) = X_1(e^{j\omega})X_2(e^{j\omega})$, then

$$\hat{X}(e^{j\omega}) = \hat{X}(e^{j\omega}) + \hat{X}_2(e^{j\omega}), \tag{13.19}$$

we must define $\hat{X}(e^{j\omega})$ as

$$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\arg[X(e^{j\omega})]. \tag{13.20}$$

With $x[n]$ real, $\arg[X(e^{j\omega})]$ can always be specified so that it is an odd periodic function of ω. With $\arg[X(e^{j\omega})]$ an odd function of ω and $\log|X(e^{j\omega})|$ an even function of ω, the complex cepstrum $\hat{x}[n]$ is guaranteed to be real.[4]

13.4 ALTERNATIVE EXPRESSIONS FOR THE COMPLEX CEPSTRUM

So far we have defined the complex cepstrum as the sequence of coefficients in the power series representation of $\hat{X}(z) = \log[X(z)]$, and we have also given an integral formula in Eq. (13.13) for determining $\hat{x}[n]$ from $\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + \angle X(e^{j\omega})$, where $\angle X(e^{j\omega})$ is the unwrapped phase function $\arg[X(e^{j\omega})]$. The logarithmic derivative can be used to derive other relations for the complex cepstrum that do not explicitly involve the complex logarithm. Assuming that $\log[X(z)]$ is analytic, then

$$\hat{X}'(z) = \frac{X'(z)}{X(z)} \tag{13.21}$$

[4]The approach outlined above to the problems presented by the complex logarithm can be developed more formally through the concept of the Riemann surface (Brown and Churchill, 2008).

where $'$ denotes differentiation with respect to z. From property 4 in Table 3.2, $z\hat{X}'(z)$ is the z-transform of $-n\hat{x}[n]$, i.e.,

$$-n\hat{x}[n] \stackrel{\mathcal{Z}}{\longleftrightarrow} z\hat{X}'(z) \tag{13.22}$$

Consequently, from Eq. (13.21),

$$-n\hat{x}[n] \stackrel{\mathcal{Z}}{\longleftrightarrow} \frac{zX'(z)}{X(z)}. \tag{13.23}$$

Beginning with Eq. (13.21) we can also derive a difference equation that is satisfied by $x[n]$ and $\hat{x}[n]$. Rearranging Eq. (13.21) and multiplying by z, we obtain

$$zX'(z) = z\hat{X}'(z) \cdot X(z). \tag{13.24}$$

Using Eq. (13.22), the inverse z-transform of this equation is

$$-nx[n] = \sum_{k=-\infty}^{\infty} (-k\hat{x}[k])x[n-k]. \tag{13.25}$$

Dividing both sides by $-n$, we obtain

$$x[n] = \sum_{k=-\infty}^{\infty} \left(\frac{k}{n}\right) \hat{x}[k]x[n-k], \qquad n \neq 0. \tag{13.26}$$

The value of $\hat{x}[0]$ can be obtained by noting that

$$\hat{x}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{X}(e^{j\omega})d\omega. \tag{13.27}$$

Since the imaginary part of $\hat{X}(e^{j\omega})$ is an odd function of ω, Eq. (13.27) becomes

$$\hat{x}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(e^{j\omega})|d\omega. \tag{13.28}$$

In summary, a signal and its complex cepstrum satisfy a nonlinear difference equation (Eq. (13.26)). Under certain conditions, this implicit relation between $\hat{x}[n]$ and $x[n]$ can be rearranged into a recursion formula that can be used in computation. Formulas of this type are discussed in Section 13.6.4.

13.5 THE COMPLEX CEPSTRUM FOR EXPONENTIAL, MINIMUM-PHASE AND MAXIMUM-PHASE SEQUENCES

13.5.1 Exponential Sequences

If a sequence $x[n]$ consists of a sum of complex exponential sequences, its z-transform $X(z)$ is a rational function of z. Such sequences are both useful and amenable to analysis. In this section, we consider the complex cepstrum for stable sequences $x[n]$ whose z-transforms are of the form

$$X(z) = \frac{Az^r \prod_{k=1}^{M_i}(1-a_k z^{-1}) \prod_{k=1}^{M_o}(1-b_k z)}{\prod_{k=1}^{N_i}(1-c_k z^{-1}) \prod_{k=1}^{N_o}(1-d_k z)}, \tag{13.29}$$

where $|a_k|$, $|b_k|$, $|c_k|$, and $|d_k|$ are all less than unity, so that factors of the form $(1-a_kz^{-1})$ and $(1-c_kz^{-1})$ correspond to the M_i zeros and the N_i poles inside the unit circle, and the factors $(1-b_kz)$ and $(1-d_kz)$ correspond to the M_o zeros and the N_o poles outside the unit circle. Such z-transforms are characteristic of sequences composed of a sum of stable exponential sequences. In the special case where there are no poles (i.e., the denominator of Eq. (13.29) is unity), then the corresponding sequence $x[n]$ is a sequence of finite length $(M+1=M_o+M_i+1)$.

Through the properties of the complex logarithm, the product of terms in Eq. (13.29) is transformed to the sum of logarithmic terms:

$$\begin{aligned}\hat{X}(z) = \log(A) + \log(z^r) &+ \sum_{k=1}^{M_i}\log(1-a_kz^{-1}) + \sum_{k=1}^{M_o}\log(1-b_kz)\\ &- \sum_{k=1}^{N_i}\log(1-c_kz^{-1}) - \sum_{k=1}^{N_o}\log(1-d_kz).\end{aligned} \tag{13.30}$$

The properties of $\hat{x}[n]$ depend on the composite properties of the inverse transforms of each term.

For real sequences, A is real, and if A is positive, the first term $\log(A)$ contributes only to $\hat{x}[0]$. Specifically, (see Problem 13.15),

$$\hat{x}[0] = \log|A|. \tag{13.31}$$

If A is negative, it is less straightforward to determine the contribution to the complex cepstrum due to the term $\log(A)$. The term z^r corresponds only to a delay or advance of the sequence $x[n]$. If $r=0$, this term vanishes from Eq. (13.30). However, if $r\neq 0$, then the unwrapped phase function $\arg[X(e^{j\omega})]$ will include a linear term with slope r. Consequently, with $\arg[X(e^{j\omega})]$ defined to be odd and periodic in ω and continuous for $|\omega|<\pi$, this linear-phase term will force a discontinuity in $\arg[X(e^{j\omega})]$ at $\omega=\pm\pi$, and $\hat{X}(z)$ will no longer be analytic on the unit circle. Although the cases of A negative and/or $r\neq 0$ can be formally accommodated, doing so seems to offer no real advantage, because if two transforms of the form of Eq. (13.29) are multiplied together, we would not expect to be able to determine how much of either A or r was contributed by each component. This is analogous to the situation in ordinary linear filtering where two signals, each with dc levels, have been added. Therefore, this question can be avoided in practice by first determining the algebraic sign of A and the value of r and then altering the input, so that its z-transform is of the form

$$X(z) = \frac{|A|\prod_{k=1}^{M_i}(1-a_kz^{-1})\prod_{k=1}^{M_o}(1-b_kz)}{\prod_{k=1}^{N_i}(1-c_kz^{-1})\prod_{k=1}^{N_o}(1-d_kz)}. \tag{13.32}$$

Correspondingly, Eq. (13.30) becomes

$$\begin{aligned}\hat{X}(z) = \log|A| + \sum_{k=1}^{M_i}\log(1-a_k z^{-1}) + \sum_{k=1}^{M_o}\log(1-b_k z)\\ - \sum_{k=1}^{N_i}\log(1-c_k z^{-1}) - \sum_{k=1}^{N_o}\log(1-d_k z).\end{aligned} \tag{13.33}$$

With the exception of the term $\log|A|$, which we have already considered, all the terms in Eq. (13.33) are of the form $\log(1-\alpha z^{-1})$ and $\log(1-\beta z)$. Bearing in mind that these factors represent z-transforms with regions of convergence that include the unit circle, we can make the power series expansions

$$\log(1-\alpha z^{-1}) = -\sum_{n=1}^{\infty}\frac{\alpha^n}{n}z^{-n}, \qquad |z| > |\alpha|, \tag{13.34}$$

$$\log(1-\beta z) = -\sum_{n=1}^{\infty}\frac{\beta^n}{n}z^{n}, \qquad |z| < |\beta^{-1}|. \tag{13.35}$$

Using these expressions, we see that for signals with rational z-transforms as in Eq. (13.32), $\hat{x}[n]$ has the general form

$$\hat{x}[n] = \begin{cases} \log|A|, & n = 0, \quad (13.36a)\\ -\sum_{k=1}^{M_i}\frac{a_k^n}{n} + \sum_{k=1}^{N_i}\frac{c_k^n}{n}, & n > 0, \quad (13.36b)\\ \sum_{k=1}^{M_o}\frac{b_k^{-n}}{n} - \sum_{k=1}^{N_o}\frac{d_k^{-n}}{n}, & n < 0. \quad (13.36c)\end{cases}$$

Note that for the special case of a finite-length sequence, the second term would be missing in each of Eqs. (13.36b) and (13.36c). Equations (13.36a) to (13.36c) suggest the following general properties of the complex cepstrum:

Property 1: The complex cepstrum decays at least as fast as $1/|n|$: Specifically,

$$|\hat{x}[n]| < C\frac{\alpha^{|n|}}{|n|}, \qquad -\infty < n < \infty,$$

where C is a constant and α equals the maximum of $|a_k|$, $|b_k|$, $|c_k|$, and $|d_k|$.[5]

Property 2: $\hat{x}[n]$ will have infinite duration, even if $x[n]$ has finite duration.

Property 3: If $x[n]$ is real, $\hat{x}[n]$ is also real.

[5]In practice, we generally deal with finite-length signals, which are represented by polynomials in z^{-1}; i.e., the numerator in Eq. (13.32). In many cases, the sequence may be hundreds or thousands of samples long. For such sequences, as the sequence length increases, it is increasingly likely that almost all of the zeros of the polynomial will cluster around the unit circle (Hughes and Nikeghbali, 2005). This implies that for long finite-length sequences, the decay of the complex cepstrum is due primarily to the factor $1/n$.

Properties 1 and 2 follow directly from Eqs. (13.36a) to (13.36c). We have suggested property 3 earlier on the basis that for $x[n]$ real, $\log|X(e^{j\omega})|$ is even and $\arg[X(e^{j\omega})]$ is odd, so that the inverse transform of

$$\hat{X}(e^{j\omega}) = \log|X(e^{j\omega})| + j\arg[X(e^{j\omega})]$$

is real. To see property 3 in the context of this section, we note that if $x[n]$ is real, then the poles and zeros of $X(z)$ are in complex conjugate pairs. Therefore, for every complex term of the form α^n/n in Eqs. (13.36a) to (13.36c) there will be a complex conjugate term $(\alpha^*)^n/n$, so that their sum will be real.

13.5.2 Minimum-Phase and Maximum-Phase Sequences

As discussed in Chapters 5 and 12, a minimum-phase sequence is a real, causal, and stable sequence with all the poles and zeros of the z-transform inside the unit circle. Note that $\log[X(z)]$ has singularities at both the poles and the zeros of $X(z)$. Since we require that the ROC of $\log[X(z)]$ include the unit circle so that $\hat{x}[n]$ is stable, and since causal sequences have an ROC of the form $r_R < |z|$, it follows that there can be no singularities of $\log[X(z)]$ on or outside the unit circle if $\hat{x}[n] = 0$ for $n < 0$. Conversely, if all the singularities of $\hat{X}(z) = \log[X(z)]$ are inside the unit circle, then it follows that $\hat{x}[n] = 0$ for $n < 0$. Since the singularities of $\hat{X}(z)$ are the poles and the zeros of $X(z)$, the complex cepstrum of $x[n]$ will be causal ($\hat{x}[n] = 0$ for $n < 0$) if and only if the poles and zeros of $X(z)$ are inside the unit circle. In other words, $x[n]$ is a minimum-phase sequence if and only if its complex cepstrum is causal.

This is easily seen for the case of exponential or finite-length sequences by considering Eqs. (13.36a)–(13.36c). Clearly, all terms in Eq. (13.36c) will be zero if all the coefficients b_k and d_k are zero, i.e., if there are no poles or zeros outside or on the unit circle. Thus, another property of the complex cepstrum is

Property 4: The complex cepstrum $\hat{x}[n] = 0$ for $n < 0$ if and only if $x[n]$ is minimum phase, i.e., $X(z)$ has all its poles and zeros inside the unit circle.

Therefore, causality of the complex cepstrum is equivalent to the minimum phase lag, minimum group delay, and minimum energy delay properties that also characterize minimum-phase sequences.

Example 13.1 Complex Cepstrum of a Minimum-Phase Echo System

The concept of the cepstrum arose initially from a consideration of echoes. As we showed in Section 13.1, a signal with an echo is represented by a convolution $x[n] = v[n] * p[n]$, where

$$p[n] = \delta[n] + \alpha\delta[n - n_0] \stackrel{\mathcal{Z}}{\longleftrightarrow} P(z) = 1 + \alpha z^{-n_0}. \tag{13.37}$$

The zeros of $P(z)$ are at locations $z_k = \alpha^{1/n_0} e^{j2\pi(k+1/2)/n_0}$, and if $|\alpha| < 1$, all the zeros will lie inside the unit circle, in which case $p[n]$ is a minimum-phase system. To find the complex cepstrum $\hat{p}[n]$, we can use the power series expansion of $\log[P(z)]$ as in Section 13.5.1 to obtain

$$\hat{P}(z) = \log[1 + \alpha z^{-n_0}] = -\sum_{n=1}^{\infty} \frac{(-\alpha)^n}{n} z^{-nn_0}, \tag{13.38}$$

from which it follows that

$$\hat{p}[n] = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\alpha^m}{m} \delta[n - mn_0]. \tag{13.39}$$

From Eq. (13.39), we see that $\hat{p}[n] = 0$ for $n < 0$ for $|\alpha| < 1$ as it should be for a minimum-phase system. Furthermore, we see that the nonzero values of the complex cepstrum for the minimum-phase echo system occur at positive integer multiples of n_0.

Maximum-phase sequences are stable sequences whose poles and zeros are all *outside* the unit circle. Thus, maximum-phase sequences are left-sided, and, by analogous arguments, it follows that the complex cepstrum of a maximum-phase sequence is also left-sided. Thus, another property of the complex cepstrum is:

Property 5: The complex cepstrum $\hat{x}[n] = 0$ for $n > 0$ if and only if $x[n]$ is maximum phase; i.e., $X(z)$ has all its poles and zeros outside the unit circle.

This property of the complex cepstrum is easily verified for exponential or finite-length sequences by noting that if all the c_ks and a_ks are zero (i.e., no poles or zeros *inside* the unit circle), then Eq. (13.36b) shows that $\hat{x}[n] = 0$ for $n > 0$.

In Example 13.1, we determined the complex cepstrum of the impulse response of the echo system when $|\alpha| < 1$; i.e., when the echo is smaller than the direct signal. If $|\alpha| > 1$, the echo is larger than the direct signal, and the zeros of the system function $P(z) = 1 + \alpha z^{-n_0}$ lie outside the unit circle. In this case, the echo system is a maximum-phase system.[6] The corresponding complex cepstrum is

$$\hat{p}[n] = \log|\alpha|\delta[n] + \sum_{m=1}^{\infty} (-1)^{m+1} \frac{\alpha^{-m}}{m} \delta[n + mn_0]. \tag{13.40}$$

From Eq. (13.40) we see that $\hat{p}[n] = 0$ for $n > 0$ for $|\alpha| > 1$ as it should be for a maximum-phase system. In this case, we see that the nonzero values of the complex cepstrum for the maximum-phase echo system occur at negative integer multiples of n_0.

13.5.3 Relationship Between the Real Cepstrum and the Complex Cepstrum

As discussed in Sections 13.1 and 13.2, the Fourier transform of the real cepstrum $c_x[n]$ is the real part of the Fourier transform of the complex cepstrum $\hat{x}[n]$, and equivalently, $c_x[n]$ corresponds to the even part of $\hat{x}[n]$, i.e.,

$$c_x[n] = \frac{\hat{x}[n] + \hat{x}[-n]}{2}. \tag{13.41}$$

[6] $P(z) = z^{-n_0}(\alpha + z^{n_0})$ has n_0 poles at $z = 0$, which are ignored in computing $\hat{p}[n]$

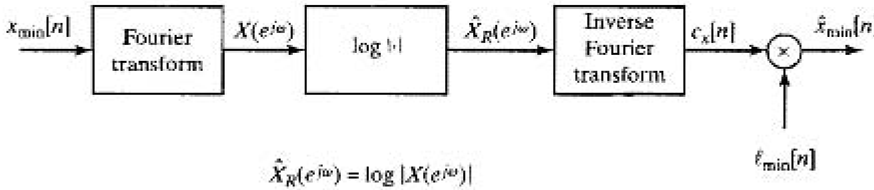


Figure 13.2 Determination of the complex cepstrum for minimum-phase signals.

If $\hat{x}[n]$ is causal, as it is if $x[n]$ is minimum phase, then Eq. (13.41) is reversible, i.e., $\hat{x}[n]$ can be recovered from $c_x[n]$ by applying an appropriate window to $c_x[n]$. Specifically,

$$\hat{x}[n] = c_x[n]\ell_{min}[n], \tag{13.42a}$$

where

$$\ell_{min}[n] = 2u[n] - \delta[n] = \begin{cases} 2 & n > 0 \\ 1 & n = 0 \\ 0 & n < 0 \end{cases}. \tag{13.42b}$$

Equations (13.42a) and (13.42b) indicate how the complex cepstrum can be obtained from the cepstrum and consequently also from the log magnitude alone if $x[n]$ is known to be minimum phase. This is also illustrated in block diagram form in Figure 13.2.

In the following example, we illustrate Eqs. (13.41) and (13.42a) for the minimum-phase echo system of Example 13.1.

Example 13.2 Real Cepstrum of a Minimum-Phase Echo System

Consider the complex cepstrum of the minimum-phase echo system as given in Eq. (13.39) in Example 13.1. From Eq. (13.41) it follows that the real cepstrum for the minimum-phase echo system is

$$c_p[n] = \frac{1}{2}\Big(\sum_{m=1}^{\infty}(-1)^{m+1}\frac{\alpha^m}{m}\delta[n - mn_0] + \sum_{m=1}^{\infty}(-1)^{m+1}\frac{\alpha^m}{m}\delta[-n - mn_0]. \tag{13.43}$$

Since $\delta[-n] = \delta[n]$, Eq. (13.43) can be written in the more compact form

$$c_p[n] = \sum_{m=1}^{\infty}(-1)^{m+1}\frac{\alpha^m}{2m}\left(\delta[n - mn_0] + \delta[n + mn_0]\right). \tag{13.44}$$

Also note that if $c_p[n]$ is given by Eq. (13.44) and $\ell_{min}[n]$ is given by Eq. (13.42b), then $\ell_{min}[n]c_p[n]$ is equal to $\hat{p}[n]$ in Eq. (13.39).

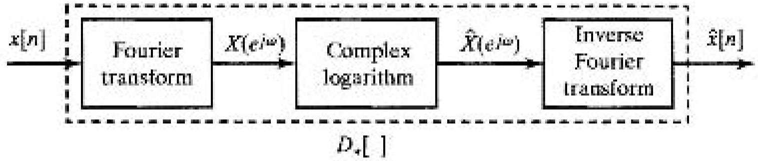


Figure 13.3 Cascade of three systems implementing the computation of the complex cepstrum operation $D_*[\ ]$.

13.6 COMPUTATION OF THE COMPLEX CEPSTRUM

The practical use of the complex cepstrum requires accurate and efficient computational methods to obtain it from a sampled signal. Implicit in all of the previous discussions has been the assumption of uniqueness and continuity of the complex logarithm of the Fourier transform of the input signal. If the mathematical representations obtained above are to serve as the basis for computation of the complex cepstrum, or equivalently, as the basis for realizations of the system $D_*[\cdot]$, then we must deal with the issues associated with computing the Fourier transform and the complex logarithm.

The system $D_*[\cdot]$ is represented in terms of the Fourier transform by the equations

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}, \tag{13.45a}$$

$$\hat{X}(e^{j\omega}) = \log[X(e^{j\omega})], \tag{13.45b}$$

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{X}(e^{j\omega})e^{j\omega n}d\omega. \tag{13.45c}$$

These equations correspond to the cascade of three systems as depicted in Figure 13.3.

In computing the complex cepstrum numerically, we are limited to finite-length input sequences, and we can compute the Fourier transform at only a finite number of frequencies. That is, instead of using the DTFT, we must use the DFT. Thus, instead of Eqs. (13.45a) to (13.45c), we have the computational realization

$$X[k] = X(e^{j\omega})\Big|_{\omega=(2\pi/N)k} = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)kn}, \tag{13.46a}$$

$$\hat{X}[k] = \log[X(e^{j\omega})]\Big|_{\omega=(2\pi/N)k}, \tag{13.46b}$$

$$\hat{x}_p[n] = \frac{1}{N}\sum_{k=0}^{N-1} \hat{X}[k]e^{j(2\pi/N)kn}. \tag{13.46c}$$

These operations are depicted in Figure 13.4(a), and the corresponding operations for realizing the inverse system are depicted in Figure 13.4(b).

Since in Eq. (13.46b) $\hat{X}[k]$ is a sampled version of $\hat{X}(e^{j\omega})$, it follows from the discussion in Section 8.4 that $\hat{x}_p[n]$ will be a time-aliased version of $\hat{x}[n]$, i.e., that $\hat{x}_p[n]$

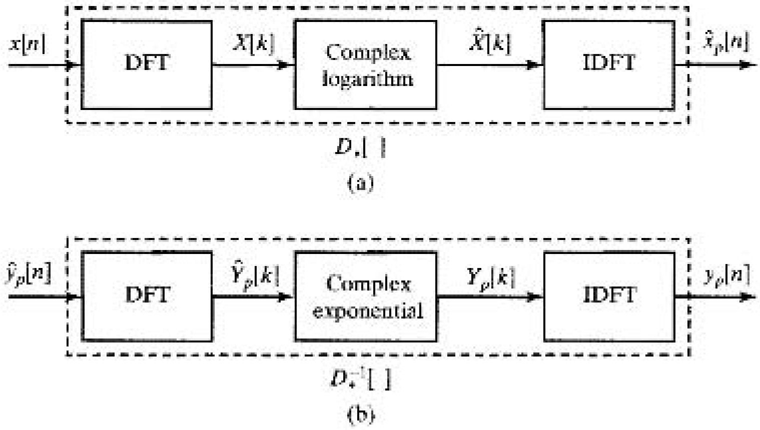


Figure 13.4 Approximate realization using the DFT of (a) $D_*[\cdot]$ and (b) $D_*^{-1}[\cdot]$.

is related to the desired $\hat{x}[n]$ by

$$\hat{x}_p[n] = \sum_{r=-\infty}^{\infty} \hat{x}[n + rN]. \tag{13.47}$$

However, we noted in Property 1 in Section 13.5 that $\hat{x}[n]$ decays faster than an exponential sequence, so it is to be expected that the approximation would become increasingly better as N increases. By appending zeros to an input sequence, it is generally possible to increase the sampling rate of the complex logarithm of the Fourier transform so that severe time aliasing does not occur in the computation of the complex cepstrum.

13.6.1 Phase Unwrapping

Samples of $\hat{X}(e^{j\omega})$ as given by Eq. (13.46b) require samples of $\log|X(e^{j\omega})|$ and $\arg[X(e^{j\omega})]$. Samples of $\log|X(e^{j\omega})|$ at a suitable sampling rate can be computed by computing the DFT of $x[n]$ with zero padding. Samples $\text{ARG}[X(e^{j\omega})]$, i.e., the phase modulo 2π are likewise straightforward to compute from samples of $X(e^{j\omega})$ by using standard inverse tangent routines available in most high-level computer languages. However, to obtain the complex cepstrum or its aliased version $\hat{x}_p[n]$, we require samples of the unwrapped phase $\arg[X(e^{j\omega})]$. Consequently, effective procedures for *unwrapping* the phase, that is, obtaining samples of the unwrapped phase from samples of the phase modulo 2π, become an important computational aspect of obtaining the complex cepstrum.

To illustrate the issues, consider a finite-length causal input sequence whose Fourier transform is of the form

$$\begin{aligned} X(e^{j\omega}) &= \sum_{n=0}^{M} x[n]e^{-j\omega n} \\ &= Ae^{-j\omega M_o} \prod_{k=1}^{M_i}(1 - a_k e^{-j\omega}) \prod_{k=1}^{M_o}(1 - b_k e^{j\omega}), \end{aligned} \tag{13.48}$$

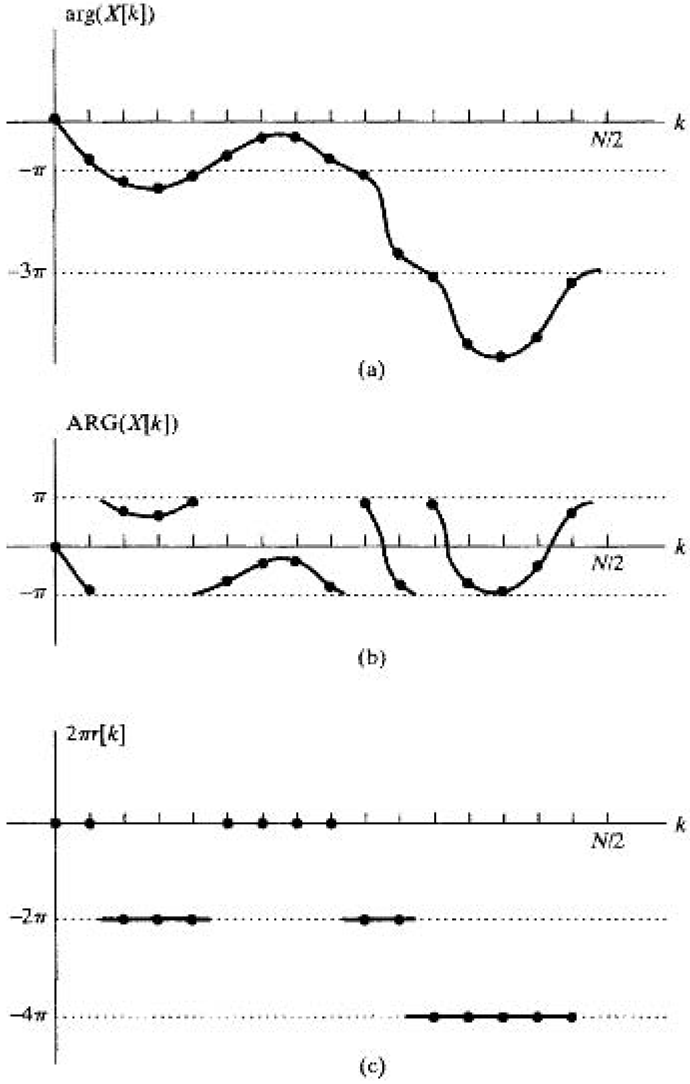


Figure 13.5 (a) Samples of $\arg[X(e^{j\omega})]$. (b) Principal value of part (a). (c) Correction sequence for obtaining arg from ARG.

where $|a_k|$ and $|b_k|$ are less than unity, $M = M_o + M_i$, and A is positive. A continuous-phase curve for a sequence of this form is shown in Figure 13.5(a). The dots indicate samples at frequencies $\omega_k = (2\pi/N)k$. Figure 13.5(b) shows the principal value and its samples as computed from the DFT of the input sequence. One approach to unwrapping the principal-value phase is based on the relation

$$\arg(X[k]) = \text{ARG}(X[k]) + 2\pi r[k], \tag{13.49}$$

where $r[k]$ denotes an integer that determines the appropriate multiple of 2π to add to the principal value at frequency $\omega_k = 2\pi k/N$. Figure 13.5(c) shows $2\pi r[k]$ required to obtain Figure 13.5(a) from 13.5(b). This example suggests the following algorithm for computing $r[k]$ from $\text{ARG}(X[k])$ starting with $r[0] = 0$:

1. If $\text{ARG}(X[k]) - \text{ARG}(X[k-1]) > 2\pi - \varepsilon_1$, then $r[k] = r[k-1] - 1$.
2. If $\text{ARG}(X[k]) - \text{ARG}(X[k-1]) < -(2\pi - \varepsilon_1)$, then $r[k] = r[k-1] + 1$.
3. Otherwise, $r[k] = r[k-1]$.
4. Repeat steps $1 - 3$ for $1 \leq k < N/2$.

After $r[k]$ is determined, Eq. (13.49) can be used to compute $\arg(X[k])$ for $0 \leq k < N/2$. At this stage, $\arg(X[k])$ will contain a large linear-phase component due to the factor $e^{-j\omega M_o}$ in Eq. (13.48). This can be removed by adding $2\pi k M_o/N$ to the unwrapped phase over the interval $0 \leq k < N/2$. The values of $\arg(X[k])$ for $N/2 < k \leq N-1$ can be obtained by using symmetry. Finally, $\arg(X[N/2]) = 0$.

The above algorithm works well if the samples of $\text{ARG}(X[k])$ are close enough together so that the discontinuities can be detected reliably. The parameter ε_1 is a tolerance recognizing that the magnitude of the difference between adjacent samples of the principal-value phase will always be less than 2π. If ε_1 is too large, a discontinuity will be indicated where there is none. If ε_1 is too small, the algorithm will miss a discontinuity falling between two adjacent samples of a rapidly varying unwrapped phase function $\arg[X(e^{j\omega})]$. Obviously, increasing the sampling rate of the DFT by increasing N improves the chances of correctly detecting discontinuities, and thus, correctly computing $\arg(X[k])$. If $\arg[X(e^{j\omega})]$ varies rapidly, then we expect $\hat{x}[n]$ to decay less rapidly than if $\arg[X(e^{j\omega})]$ varied more slowly. Therefore, aliasing of $\hat{x}[n]$ is more of a problem for rapidly varying phase. Increasing the value of N reduces the aliasing of the complex cepstrum and also improves the chances of being able to correctly unwrap the phase of $X[k]$ by the previously described algorithm.

In some cases, the simple algorithm we just developed may fail because it is impossible or impractical to use a large enough value for N. Often, the aliasing for a given N is acceptable, but principal-value discontinuities cannot be reliably detected. Tribolet (1977, 1979) proposed a modification of the algorithm that uses both the principal value of the phase and the phase derivative to compute the unwrapped phase. As above, Eq. (13.49) gives the set of permissible values at frequency $\omega_k = (2\pi/N)k$, and we seek to determine $r[k]$. It is assumed that we know the phase derivative,

$$\arg'(X[k]) = \frac{d}{d\omega}\arg[X(e^{j\omega})]\bigg|_{\omega=2\pi k/N}$$

at all values of k. (A procedure for computing these samples of the phase derivative will be developed in Section 13.6.2.) To compute $\arg(X[k])$ we further assume that $\arg(X[k-1])$ is known. Then, $\widetilde{\arg}(X[k])$, the estimate of $\arg(X[k])$, is defined as

$$\widetilde{\arg}(X[k]) = \arg(X[k-1]) + \frac{\Delta\omega}{2}\{\arg'(X[k]) + \arg'(X[k-1])\}. \tag{13.50}$$

Equation (13.50) is obtained by applying trapezoidal numerical integration to the samples of the phase derivative. This estimate is said to be *consistent* if for some ε_2 an integer $r[k]$ exists such that

$$|\widetilde{\arg}(X[k]) - \text{ARG}(X[k]) - 2\pi r[k]| < \varepsilon_2 < \pi. \tag{13.51}$$

Obviously, the estimate improves with decreasing numerical integration step size $\Delta\omega$. Initially, $\Delta\omega = 2\pi/N$ as provided by the DFT. If Eq. (13.51) cannot be satisfied by an integer $r[k]$, then $\Delta\omega$ is halved, and a new estimate of $\arg(X[k])$ is computed with the new step size. Then, Eq. (13.51) is evaluated with the new estimate. Increasingly accurate estimates of $\arg(X[k])$ are computed by numerical integration until Eq. (13.51) can be satisfied by an integer $r[k]$. That resulting $r[k]$ is used in Eq. (13.49) to finally compute $\arg(X[k])$. This unwrapped phase is then used to compute $\arg(X[k+1])$, and so on.

Another approach to phase unwrapping for a finite-length sequence is based on the fact that the z-transform of a finite-length sequence is a finite-order polynomial, and therefore can be viewed as consisting of a product of 1^{st}-order factors. For each such factor, $\text{ARG}[X(e^{j\omega})]$ and $\arg[X(e^{j\omega})]$ are equal, i.e., the phase for a single factor will never require unwrapping. Furthermore, the unwrapped phase for the product of the individual factors is the sum of the unwrapped phases of the individual factors. Consequently, by treating a finite-length sequence of length N as the coefficients in an N^{th}-order polynomial, and by first factoring that polynomial into its 1^{st}-order factors, the unwrapped phase can be easily computed. For small values of N, conventional polynomial-rooting algorithms can be applied. For large values, an effective algorithm has been developed by Sitton et al. (2003) and has been successfully demonstrated with polynomials of order in the millions. However, there are cases in which that algorithm also fails, particularly in identifying roots that are not close to the unit circle.

In the discussion above, we have briefly described several algorithms for obtaining the unwrapped phase. Karam and Oppenheim (2007) have also proposed combining these algorithms to exploit their various advantages.

Other issues in computing the complex cepstrum from a sampled input signal $x[n]$ relate to the linear-phase term in $\arg[X(e^{j\omega})]$ and the sign of the overall scale factor A. In our definition of the complex cepstrum, $\arg[X(e^{j\omega})]$ is required to be continuous, odd and periodic in ω. Therefore, the sign of A must be positive, since if negative, a phase discontinuity would occur at $\omega = 0$. Furthermore, $\arg[X(e^{j\omega})]$ cannot contain a linear term, since that would impose a discontinuity at $\omega = \pi$. Consider, for example, a finite-length causal sequence of length $M+1$. The corresponding z-transform will be of the form of Eq. (13.29) with N_o=N_i=0, and M=M_o+M_i. Also, since $x[n] = 0, n < 0$, it follows that $r = -M_o$. Consequently, the Fourier transform takes the form

$$\begin{aligned} X(e^{j\omega}) &= \sum_{n=0}^{M} x[n]e^{-j\omega n} \\ &= Ae^{-j\omega M_o} \prod_{k=1}^{M_i}(1 - a_k e^{-j\omega}) \prod_{k=1}^{M_o}(1 - b_k e^{j\omega}), \end{aligned} \tag{13.52}$$

with $|a_k|$ and $|b_k|$ less than unity. The sign of A is easily determined, since it will correspond to the sign of $X(e^{j\omega})$ at $\omega = 0$, which, in turn, is easily computed as the sum of all the terms in the input sequence.

13.6.2 Computation of the Complex Cepstrum Using the Logarithmic Derivative

As an alternative to the explicit computation of the complex logarithm, a mathematical representation based on the logarithmic derivative can be exploited. For real sequences, the derivative of $\hat{X}(e^{j\omega})$ can be represented in the equivalent forms

$$\hat{X}'(e^{j\omega}) = \frac{d\hat{X}(e^{j\omega})}{d\omega} = \frac{d}{d\omega}\log|X(e^{j\omega})| + j\frac{d}{d\omega}\arg[X(e^{j\omega})] \tag{13.53a}$$

and

$$\hat{X}'(e^{j\omega}) = \frac{X'(e^{j\omega})}{X(e^{j\omega})}, \tag{13.53b}$$

where $'$ represents differentiation with respect to ω. Since the DTFT of $x[n]$ is

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}, \tag{13.54}$$

its derivative with respect to ω is

$$X'(e^{j\omega}) = \sum_{n=-\infty}^{\infty} (-jnx[n])e^{-j\omega n}; \tag{13.55}$$

i.e., $X'(e^{j\omega})$ is the DTFT of $-jnx[n]$. Likewise, $\hat{X}'(e^{j\omega})$ is the Fourier transform of $-jn\hat{x}[n]$. Thus, $\hat{x}[n]$ can be determined for $n \neq 0$ from

$$\hat{x}[n] = \frac{-1}{2\pi nj}\int_{-\pi}^{\pi} \frac{X'(e^{j\omega})}{X(e^{j\omega})}e^{j\omega n}d\omega, \qquad n \neq 0. \tag{13.56}$$

The value of $\hat{x}[0]$ can be determined from the log magnitude as

$$\hat{x}[0] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|X(e^{j\omega})|d\omega. \tag{13.57}$$

Equations (13.54) to (13.57) represent the complex cepstrum in terms of the DTFTs of $x[n]$ and $nx[n]$ and thus do not explicitly involve the unwrapped phase. For finite-length sequences, samples of these transforms can be computed using the DFT, thereby leading to the corresponding equations

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)kn} = X(e^{j\omega})\Big|_{\omega=(2\pi/N)k}, \tag{13.58a}$$

$$X'[k] = -j\sum_{n=0}^{N-1} nx[n]e^{-j(2\pi/N)kn} = X'(e^{j\omega})\Big|_{\omega=(2\pi/N)k}, \tag{13.58b}$$

$$\hat{x}_{dp}[n] = -\frac{1}{jnN}\sum_{k=0}^{N-1}\frac{X'[k]}{X[k]}e^{j(2\pi/N)kn}, \qquad 1 \leq n \leq N-1, \tag{13.58c}$$

$$\hat{x}_{dp}[0] = \frac{1}{N}\sum_{k=0}^{N-1}\log|X[k]|, \tag{13.58d}$$

where the subscript d refers to the use of the logarithmic derivative and the subscript p is a reminder of the inherent periodicity of the DFT calculations. With the use of

Eqs. (13.58a) to (13.58d), we avoid the problems of computing the complex logarithm at the cost, however, of more severe aliasing, since now

$$\hat{x}_{dp}[n] = \frac{1}{n} \sum_{r=-\infty}^{\infty} (n + rN)\hat{x}[n + rN], \quad n \neq 0. \tag{13.59}$$

Thus, assuming that the sampled continuous phase curve is accurately computed, we would expect that for a given value of N, $\hat{x}_p[n]$ in Eq. (13.46c) would be a better approximation to $\hat{x}[n]$ than would $\hat{x}_{dp}[n]$ in Eq. (13.58c).

13.6.3 Minimum-Phase Realizations for Minimum-Phase Sequences

In the special case of minimum-phase sequences, the mathematical representation is simplified, as indicated in Figure 13.2. A computational realization based on using the DFT in place of the Fourier transform in Figure 13.2 is given by the equations

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)kn}, \tag{13.60a}$$

$$c_{xp}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \log|X[k]|e^{j(2\pi/N)kn}. \tag{13.60b}$$

In this case, it is the cepstrum that is aliased; i.e.,

$$c_{xp}[n] = \sum_{r=-\infty}^{\infty} c_x[n + rN]. \tag{13.61}$$

To compute the complex cepstrum from $c_{xp}[n]$ based on Figure 13.2, we write:

$$\hat{x}_{cp}[n] = \begin{cases} c_{xp}[n], & n = 0, \quad N/2, \\ 2c_{xp}[n], & 1 \le n < N/2, \\ 0, & N/2 < n \le N - 1. \end{cases} \tag{13.62}$$

Clearly, $\hat{x}_{cp}[n] \neq \hat{x}_p[n]$, since it is the even part of $\hat{x}[n]$ that is aliased, rather than $\hat{x}[n]$ itself. Nevertheless, for large N, $\hat{x}_{cp}[n]$ can be expected to be a reasonable approximation to $\hat{x}[n]$ over the finite interval $0 \le n < N/2$. Similarly, if $x[n]$ is maximum phase, an approximation to the complex cepstrum would be obtained from

$$\hat{x}_{cp}[n] = \begin{cases} c_{xp}[n], & n = 0, \quad N/2, \\ 0, & 1 \le n < N/2, \\ 2c_{xp}[n], & N/2 < n \le N - 1. \end{cases} \tag{13.63}$$

13.6.4 Recursive Computation of the Complex Cepstrum for Minimum- and Maximum-Phase Sequences

For minimum-phase sequences, the difference Eq. (13.26) can be rearranged to provide a recursion formula for $\hat{x}[n]$. Since for minimum-phase sequences both $\hat{x}[n] = 0$ and

$x[n] = 0$ for $n < 0$. Eq. (13.26) becomes

$$\begin{aligned} x[n] &= \sum_{k=0}^{n} \left(\frac{k}{n}\right) \hat{x}[k]x[n-k], \qquad n > 0, \\ &= \hat{x}[n]x[0] + \sum_{k=0}^{n-1} \left(\frac{k}{n}\right) \hat{x}[k]x[n-k], \end{aligned} \tag{13.64}$$

which is a recursion for $D_*[\ ]$ for minimum-phase signals. Solving for $\hat{x}[n]$ yields the recursion formula

$$\hat{x}[n] = \begin{cases} 0, & n < 0, \\ \dfrac{x[n]}{x[0]} - \displaystyle\sum_{k=0}^{n-1} \left(\frac{k}{n}\right) \hat{x}[k] \frac{x[n-k]}{x[0]}, & n > 0. \end{cases} \tag{13.65}$$

Assuming $x[0] > 0$, the value of $\hat{x}[0]$ can be shown to be (see Problem 13.15)

$$\hat{x}[0] = \log(|A|) = \log(|x[0]|). \tag{13.66}$$

Therefore, Eqs. (13.65) and (13.66) constitute a procedure for computing the complex cepstrum for minimum-phase signals. It also follows from Eq. (13.65) that this computation is causal for minimum-phase inputs; i.e., the output at time n_0 is dependent only on the input for $n \le n_0$, where n_0 is arbitrary (see Problem 13.20). Similarly, Eqs. (13.64) and (13.66) represent the computation of the minimum-phase sequence from its complex cepstrum.

For maximum-phase signals, $\hat{x}[n] = 0$, and $x[n] = 0$ for $n > 0$. Thus, in this case Eq. (13.26) becomes

$$\begin{aligned} x[n] &= \sum_{k=n}^{0} \left(\frac{k}{n}\right) \hat{x}[k]x[n-k], \qquad n < 0, \\ &= \hat{x}[n]x[0] + \sum_{k=n+1}^{0} \left(\frac{k}{n}\right) \hat{x}[k]x[n-k]. \end{aligned} \tag{13.67}$$

Solving for $\hat{x}[n]$, we have

$$\hat{x}[n] = \begin{cases} \dfrac{x[n]}{x[0]} - \displaystyle\sum_{k=n+1}^{0} \left(\frac{k}{n}\right) \hat{x}[k] \frac{x[n-k]}{x[0]}, & n < 0, \\ \log(x[0]), & n = 0, \\ 0, & n > 0. \end{cases} \tag{13.68}$$

Equation (13.68) serves as a procedure for computing the complex cepstrum for a maximum-phase sequence and Eq. (13.67) is a computational procedure for the inverse characteristic system for convolution.

Thus we see that in the case of minimum-phase or maximum-phase sequences, we also have the recursion formulas of Eqs. (13.64)–(13.68) as possible realizations of the characteristic system and its inverse. These equations can be quite useful when the input sequence is very short or when only a few samples of the complex cepstrum are desired. With these formulas, of course, there is no aliasing error.

13.6.5 The Use of Exponential Weighting

Exponential weighting of a sequence can be used to avoid or mitigate some of the problems encountered in computing the complex cepstrum. Exponential weighting of a sequence $x[n]$ is defined by

$$w[n] = \alpha^n x[n]. \tag{13.69}$$

The corresponding z-transform is

$$W(z) = X(\alpha^{-1}z). \tag{13.70}$$

If the ROC of $X(z)$ is $r_R < |z| < r_L$, then the ROC of $W(z)$ is $|\alpha|r_R < |z| < |\alpha|r_L$, and the poles and zeros of $X(z)$ are shifted radially by the factor $|\alpha|$; i.e., if z_0 is a pole or zero of $X(z)$, then $z_0\alpha$ is the corresponding pole or zero of $W(z)$.

A convenient property of exponential weighting is that it commutes with convolution. That is, if $x[n] = x_1[n] * x_2[n]$ and $w[n] = a^n x[n]$, then

$$W(z) = X(\alpha^{-1}z) = X_1(\alpha^{-1}z)X_2(\alpha^{-1}z), \tag{13.71}$$

so that

$$\begin{aligned} w[n] &= (a^n x_1[n]) * (a^n x_2[n]) \\ &= w_1[n] * w_2[n]. \end{aligned} \tag{13.72}$$

Thus, in computing the complex cepstrum, if $X(z) = X_1(z)X_2(z)$,

$$\begin{aligned} \hat{W}(z) &= \log[W(z)] \\ &= \log[W_1(z)] + \log[W_2(z)]. \end{aligned} \tag{13.73}$$

Exponential weighting can be exploited with cepstrum computation in a variety of ways. For example, poles or zeros of $X(z)$ on the unit circle require special care in computing the complex cepstrum. It can be shown (Carslaw, 1952) that a factor $\log(1 - e^{j\theta}e^{-j\omega})$ has a Fourier series

$$\log(1 - e^{j\theta}e^{-j\omega}) = -\sum_{n=1}^{\infty} \frac{e^{j\theta n}}{n} e^{-j\omega n} \tag{13.74}$$

and thus, the contribution of such a term to the complex cepstrum is $(e^{j\theta n}/n)u[n-1]$. However, the log magnitude is infinite, and the phase is discontinuous with a jump of π radians at $\omega = \theta$. This presents obvious computational difficulties that we would prefer to avoid. By exponential weighting with $0 < \alpha < 1$, all poles and zeros are moved radially inward. Therefore, a pole or zero on the unit circle will move inside the unit circle.

As another example, consider a causal, stable signal $x[n]$ that is nonminimum phase. The exponentially weighted signal, $w[n] = \alpha^n x[n]$, can be converted into a minimum-phase sequence if α is chosen, so that $|z_{max}\alpha| < 1$, where z_{max} is the location of the zero with the greatest magnitude.

13.7 COMPUTATION OF THE COMPLEX CEPSTRUM USING POLYNOMIAL ROOTS

In Section 13.6.1, we discussed the fact that for finite-length sequences, we could exploit the fact that the z-transform is a finite-order polynomial, and that the total unwrapped phase can be obtained by summing the unwrapped phases for each of the factors. If the polynomial is first factored into its 1^{st}-order terms using a polynomial rooting algorithm, then the unwrapped phase for each factor is easily specified analytically. In a similar manner the complex cepstrum for the finite-length sequence can be obtained by first factoring the polynomial, and then summing the complex cepstra for each of the factors.

The basic approach is suggested by Section 13.5.1. If the sequence $x[n]$ has finite length, as is essentially always the case with signals obtained by sampling, then its z-transform is a polynomial in z^{-1} of the form

$$X(z) = \sum_{n=0}^{M} x[n]z^{-n}. \tag{13.75}$$

Such an M^{th}-order polynomial in z^{-1} can be represented as

$$X(z) = x[0]\prod_{m=1}^{M_i}(1 - a_m z^{-1})\prod_{m=1}^{M_o}(1 - b_m^{-1} z^{-1}), \tag{13.76}$$

where the quantities a_m are the (complex) zeros that lie inside the unit circle, and the quantities b_m^{-1} are the zeros that are outside the unit circle; i.e., $|a_m| < 1$ and $|b_m| < 1$. We assume that no zeros lie precisely on the unit circle. If we factor a term $-b_m^{-1}z^{-1}$ out of each factor of the product on the right in Eq. (13.76), that equation can be expressed as

$$X(z) = Az^{-M_o}\prod_{m=1}^{M_i}(1 - a_m z^{-1})\prod_{m=1}^{M_o}(1 - b_m z), \tag{13.77a}$$

where

$$A = x[0](-1)^{M_o}\prod_{m=1}^{M_o} b_m^{-1}. \tag{13.77b}$$

This representation can be computed by using a polynomial rooting algorithm to find the zeros a_m and $1/b_m$ that lie inside and outside the unit circle, respectively, for the polynomial whose coefficients are the sequence $x[n]$.[7]

Given the numeric representation of the z-transform polynomial as in Eqs. (13.77a) and (13.77b), numeric values of the complex cepstrum sequence can be computed from

[7] Perhaps not surprisingly, it is rare that a computed root of a polynomial is precisely on the unit circle. In cases where this occurs, such roots can be moved by exponential weighting, as described in Section 13.6.5.

Eqs. (13.36a)–(13.36c) as

$$\hat{x}[n] = \begin{cases} \log|A|, & n = 0, \\ -\displaystyle\sum_{m=1}^{M_i} \frac{a_m^n}{n}, & n > 0, \\ \displaystyle\sum_{m=1}^{M_o} \frac{b_m^{-n}}{n}, & n < 0. \end{cases} \tag{13.78}$$

If $A < 0$, this fact can be recorded separately, along with the value of M_o, the number of roots that are outside the unit circle. With this information and $\hat{x}[n]$, we have all that is needed to reconstruct the original signal $x[n]$. Indeed, in Section 13.8.2, it will be shown that, in principle, $x[n]$ can be computed recursively from just $M + 1 = M_o + M_i + 1$ samples of $\hat{x}[n]$.

This method of computation is particularly useful when $M = M_o + M_i$ is small, but it is not limited to small M. Steiglitz and Dickinson (1982) first proposed this method and reported successful rooting of polynomials with degree as high as $M = 256$, which was a practical limit imposed by computational resources readily available at that time. With the polynomial rooting algorithm of Sitton et al. (2003), the complex cepstrum of extremely long finite-length sequences can be accurately computed. Among the advantages of this method are the fact that there is no aliasing and there are none of the uncertainties associated with phase unwrapping.

13.8 DECONVOLUTION USING THE COMPLEX CEPSTRUM

The complex cepstrum operator $D_*[\ ]$, plays a key role in the theory of homomorphic systems, which is based on a generalization of the principle of superposition (Oppenheim, 1964, 1967, 1969a, Schafer, 1969 and Oppenheim, Schafer and Stockham, 1968). In homomorphic filtering of convolved signals, the operator $D_*[\ ]$ is termed the *characteristic system for convolution* since it has the special property of transforming convolution into addition. To see this, suppose

$$x[n] = x_1[n] * x_2[n] \tag{13.79}$$

so that the corresponding z-transform is

$$X(z) = X_1(z) \cdot X_2(z). \tag{13.80}$$

If the complex logarithm is computed as we have prescribed in the definition of the complex cepstrum, then

$$\begin{aligned} \hat{X}(z) &= \log[X(z)] = \log[X_1(z)] + \log[X_2(z)] \\ &= \hat{X}_1(z) + \hat{X}_2(z), \end{aligned} \tag{13.81}$$

which implies that the complex cepstrum is

$$\hat{x}[n] = D_*[x_1[n] * x_2[n]] = \hat{x}_1[n] + \hat{x}_2[n]. \tag{13.82}$$

A similar analysis shows that if $\hat{y}[n] = y_1[n] + y_2[n]$, then it follows that $D_*^{-1}[\hat{y}_1[n] + \hat{y}_2[n]] = \hat{y}_1[n] * \hat{y}_2[n]$. If the cepstral components $\hat{x}_1[n]$ and $\hat{x}_2[n]$ occupy

Figure 13.6 Canonic form for homomorphic systems where inputs and corresponding outputs are combined by convolution.

different quefrency ranges, linear filtering can be applied to the complex cepstrum to remove either $x_1[n]$ or $x_2[n]$. If this is followed by transformation through the inverse system $D_*^{-1}[\ ]$, the corresponding component will be removed in the output. This procedure for separating convolved signals (deconvolution) is depicted in Figure 13.6, where the system $L[\ ]$ is a linear (although not necessarily time invariant) system. The symbols $*$ and $+$ at the inputs and outputs of the component systems in Figure 13.6 denote the operations of superposition that hold at each point in the diagram. Figure 13.6 is a general representation of a class of systems that obey a generalized principle of superposition with convolution as the operation for combining signals. All members of this class of systems differ only in the linear part $L[\ ]$.

In the remainder of this section, we illustrate how cepstral analysis can be used for the special deconvolution problems of decomposing a signal into either a convolution of a minimum-phase and allpass component or minimum-phase and maximum-phase component. In Section 13.9, we illustrate how cepstral analysis can be applied to deconvolution of a signal convolved with an impulse train, representing for example, an idealization of a multipath environment. In Section 13.10, we generalize this example to illustrate how cepstral analysis has been successfully applied to speech processing.

13.8.1 Minimum-Phase/Allpass Homomorphic Deconvolution

Any sequence $x[n]$ for which the complex cepstrum exists can always be expressed as the convolution of a minimum-phase and an allpass sequence as in

$$x[n] = x_{min}[n] * x_{ap}[n]. \tag{13.83}$$

In Eq. (13.83) $x_{min}[n]$ and $x_{ap}[n]$ denote minimum-phase and allpass components respectively.

If $x[n]$ is not minimum phase, then the system of Figure 13.2 with input $x[n]$ and $\ell_{min}[n]$ given by Eq. (13.42b) produces the complex cepstrum of the minimum-phase sequence that has the same Fourier transform magnitude as $x[n]$. If $\ell_{max}[n] = \ell_{min}[-n]$ is used, the output will be the complex cepstrum of the maximum-phase sequence having the same Fourier transform magnitude as $x[n]$.

We can obtain the complex cepstrum $\hat{x}_{min}[n]$ of the sequence $x_{min}[n]$ in Eq. (13.83) through the operations of Figure 13.2. The complex cepstrum $\hat{x}_{ap}[n]$ can be obtained from $\hat{x}[n]$ by subtracting $\hat{x}_{min}[n]$ from $\hat{x}[n]$, i.e.,

$$\hat{x}_{ap}[n] = \hat{x}[n] - \hat{x}_{min}[n].$$

To obtain $x_{min}[n]$ and $x_{ap}[n]$, we apply the transformation D_*^{-1} to $\hat{x}_{min}[n]$ and $\hat{x}_{ap}[n]$.

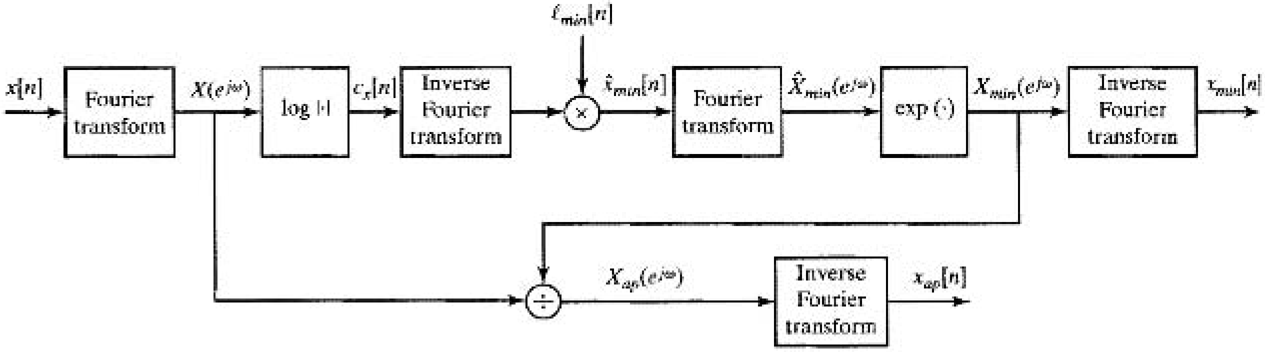


Figure 13.7 Deconvolution of a sequence into minimum-phase and allpass components using the cepstrum.

Although the approach outlined above to obtain $x_{min}[n]$ and $x_{ap}[n]$ is theoretically correct, explicit evaluation of the complex cepstrum $\hat{x}[n]$ is required in its implementation. If we are interested only in obtaining $x_{min}[n]$ and $x_{ap}[n]$, evaluation of the complex cepstrum and the associated need for phase unwrapping can be avoided. The basic strategy is incorporated in the block diagram of Figure 13.7. This system relies on the fact that

$$X_{ap}(e^{j\omega}) = \frac{X(e^{j\omega})}{X_{min}(e^{j\omega})}. \tag{13.84a}$$

The magnitude of $X_{ap}(e^{j\omega})$ is therefore

$$|X_{ap}(e^{j\omega})| = \frac{|X(e^{j\omega})|}{|X_{min}(e^{j\omega})|} = 1 \tag{13.84b}$$

and

$$\angle X_{ap}(e^{j\omega}) = \angle X(e^{j\omega}) - \angle X_{min}(e^{j\omega}). \tag{13.84c}$$

Since $x_{ap}[n]$ is obtained as the inverse Fourier transform of $e^{j\angle X_{ap}(e^{j\omega})}$, (that is, $|X_{ap}(e^{j\omega})| = 1$), each of the phase functions in Eq. (13.84c) need only be known or specified to within integer multiples of 2π. Therefore, even though as a natural consequence of the procedure outlined in Figure 13.7, $\angle X_{min}(e^{j\omega}) = \mathcal{I}m\{\hat{X}_{min}(e^{j\omega})\}$ will be an unwrapped phase function, $\angle X(e^{j\omega})$ in Eq. (13.84c) can be computed modulo 2π.

13.8.2 Minimum-Phase/Maximum-Phase Homomorphic Deconvolution

Another representation of a sequence is as the convolution of a minimum-phase sequence with a maximum-phase sequence as in

$$x[n] = x_{mn}[n] * x_{mx}[n], \tag{13.85}$$

where $x_{mn}[n]$ and $x_{mx}[n]$ denote minimum-phase and maximum-phase components, respectively.[8] In this case, the corresponding complex cepstrum is

$$\hat{x}[n] = \hat{x}_{mn}[n] + \hat{x}_{mx}[n]. \tag{13.86}$$

[8]In general the minimum-phase component $x_{mn}[n]$ in Eq. (13.85) will be different from $x_{min}[n]$ in Eq. (13.83).

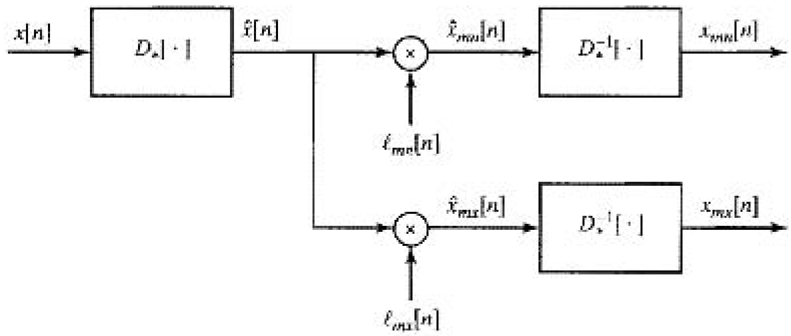


Figure 13.8 The use of homomorphic deconvolution to separate a sequence into minimum-phase and maximum-phase components.

To extract $x_{mn}[n]$ and $x_{mx}[n]$ from $x[n]$, we specify $\hat{x}_{mn}[n]$ as

$$\hat{x}_{mn}[n] = \ell_{mn}[n]\hat{x}[n], \tag{13.87a}$$

where

$$\ell_{mn}[n] = u[n]. \tag{13.87b}$$

Similarly, we specify $\hat{x}_{mx}[n]$ as

$$\hat{x}_{mx}[n] = \ell_{mx}[n]\hat{x}[n] \tag{13.88a}$$

where

$$\ell_{mx}[n] = u[-n-1]. \tag{13.88b}$$

$x_{mn}[n]$ and $x_{mx}[n]$ can be obtained from $\hat{x}_{mn}[n]$ and $\hat{x}_{mx}[n]$, respectively, as the output of the inverse characteristic system $D_*^{-1}[\cdot]$. The operations required for the decomposition of Eq. (13.85) are depicted in Figure 13.8. This method of factoring a sequence into its minimum- and maximum-phase parts has been used by Smith and Barnwell (1986) in the design of filter banks. Note that we have arbitrarily assigned all of $\hat{x}[0]$ to $\hat{x}_{mn}[0]$, and we have set $\hat{x}_{mx}[0] = 0$. Obviously, other combinations are possible, since all that is required is that $\hat{x}_{mn}[0] + \hat{x}_{mx}[0] = \hat{x}[0]$.

The recursion formulas of Section 13.6.4 can be combined with the representation of Eq. (13.85) to yield an interesting result for finite-length sequences. Specifically, in spite of the infinite extent of the complex cepstrum of a finite-length sequence, we can show that for an input sequence of length $M+1$, we need only $M+1$ samples of $\hat{x}[n]$ to determine $x[n]$. To see this, consider the z-transform of Eq. (13.85), i.e.,

$$X(z) = X_{mn}(z)X_{mx}(z), \tag{13.89a}$$

where

$$X_{mn}(z) = A\prod_{k=1}^{M_i}(1-a_k z^{-1}), \tag{13.89b}$$

$$X_{mx}(z) = \prod_{k=1}^{M_o}(1-b_k z), \tag{13.89c}$$

with $|a_k| < 1$ and $|b_k| < 1$. Note that we have neglected the delay of M_o samples that would be needed for a causal sequence, so that $x_{mn}[n] = 0$ outside the interval $0 \le n \le M_i$ and $x_{mx}[n] = 0$ outside the interval $-M_o \le n \le 0$. Since the sequence $x[n]$ is the convolution of $x_{mn}[n]$ and $x_{mx}[n]$, it is nonzero in the interval $-M_o \le n \le M_i$. Using the previous recursion formulas, we can write

$$x_{mn}[n] = \begin{cases} 0, & n < 0, \\ e^{\hat{x}[0]}, & n = 0, \\ \hat{x}[n]x_{mn}[0] + \displaystyle\sum_{k=0}^{n-1} \left(\frac{k}{n}\right) \hat{x}[k]x_{mn}[n-k], & n > 0, \end{cases} \tag{13.90}$$

and

$$x_{mx}[n] = \begin{cases} \hat{x}[n] + \displaystyle\sum_{k=n+1}^{0} \left(\frac{k}{n}\right) \hat{x}[k]x_{mx}[n-k], & n < 0, \\ 1, & n = 0, \\ 0, & n > 0. \end{cases} \tag{13.91}$$

Clearly, we require $M_i + 1$ values of $\hat{x}[n]$ to compute $x_{mn}[n]$ and M_o values of $\hat{x}[n]$ to compute $x_{mx}[n]$. Thus, only $M_i + M_o + 1$ values of the infinite sequence $\hat{x}[n]$ are required to completely recover the minimum-phase and maximum-phase components of the finite-length sequence $x[n]$.

As mentioned in Section 13.7, the result that we have just obtained could be used to implement the inverse characteristic system for convolution when the cepstrum has been computed by polynomial rooting. We simply need to compute $x_{mn}[n]$ and $x_{mx}[n]$ by the recursions of Eqs. (13.90) and (13.91) and then reconstruct the original signal by the convolution $x[n] = x_{mn}[n] * x_{mx}[n]$.

13.9 THE COMPLEX CEPSTRUM FOR A SIMPLE MULTIPATH MODEL

As discussed in Example 13.1, a highly simplified model of multipath or reverberation consists of representing the received signal as the convolution of the transmitted signal with an impulse train. Specifically, with $v[n]$ denoting a transmitted signal and $p[n]$ the impulse response of a multipath channel or other system generating multiple echoes,

$$x[n] = v[n] * p[n], \tag{13.92a}$$

or, in the z-transform domain,

$$X(z) = V(z)P(z). \tag{13.92b}$$

In our analysis in this section, we choose $p[n]$ to be of the form

$$p[n] = \delta[n] + \beta\delta[n - N_0] + \beta^2\delta[n - 2N_0], \tag{13.93a}$$

and its z-transform is then

$$P(z) = 1 + \beta z^{-N_0} + \beta^2 z^{-2N_0} = \frac{1 - \beta^3 z^{-3N_0}}{1 - \beta z^{-N_0}}. \tag{13.93b}$$

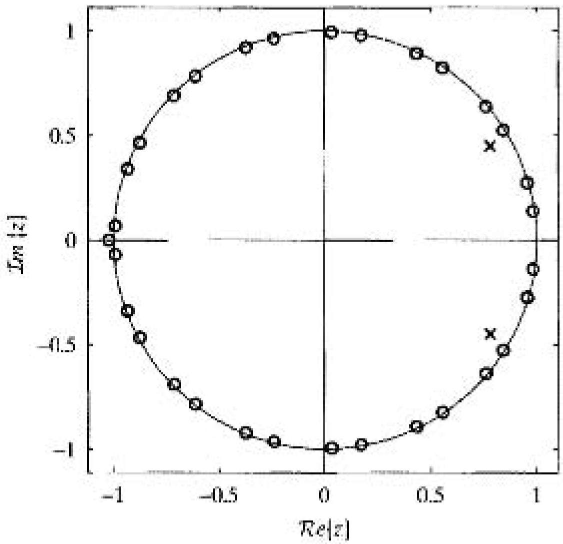


Figure 13.9 Pole-zero plot of the z-transform $X(z) = V(z)P(z)$ for the example signal of Figure 13.10.

For example, $p[n]$ might correspond to the impulse response of a multipath channel or other system that generates multiple echoes at a spacing of N_0 and $2N_0$. The component $v[n]$ will be taken to be the response of a 2nd-order system, such that

$$V(z) = \frac{b_0 + b_1 z^{-1}}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})}, \qquad |z| > |r|. \tag{13.94a}$$

In the time domain, $v[n]$ can be expressed as

$$v[n] = b_0 w[n] + b_1 w[n-1], \tag{13.94b}$$

where

$$w[n] = \frac{r^n}{4\sin^2\theta}[\cos(\theta n) - \cos[\theta(n+2)]]u[n], \qquad \theta \neq 0, \pi. \tag{13.94c}$$

Figure 13.9 shows the pole–zero plot of the z-transform $X(z) = V(z)P(z)$ for the specific set of parameters $b_0 = 0.98$, $b_1 = 1$, $\beta = r = 0.9$, $\theta = \pi/6$, and $N_0 = 15$. Figure 13.10 shows the signals $v[n]$, $p[n]$, and $x[n]$ for these parameters. As seen in Figure 13.10, the convolution of the pulse-like signal $v[n]$ with the impulse train $p[n]$ results in a series of superimposed delayed copies (echoes) of $v[n]$.

This signal model is a simplified version of models that are used in the analysis and processing of signals in a variety of contexts, including communications systems, speech processing, sonar, and seismic data analysis. In a communications context, $v[n]$ in Eqs. (13.92a) and (13.92b) might represent a signal transmitted over a multipath channel, $x[n]$ the received signal, and $p[n]$ the channel impulse response. In speech processing, $v[n]$ would represent the combined effects of the glottal pulse shape and the resonance effects of the human vocal tract, while $p[n]$ would represent the periodicity of the vocal excitation during voiced speech such as a vowel sound (Flanagan, 1972; Rabiner and Schafer, 1978; Quatieri, 2002). Equation (13.94a) incorporates only one resonance, while in the general speech model, the denominator would generally include at least ten complex poles. In seismic data analysis, $v[n]$ would represent the waveform

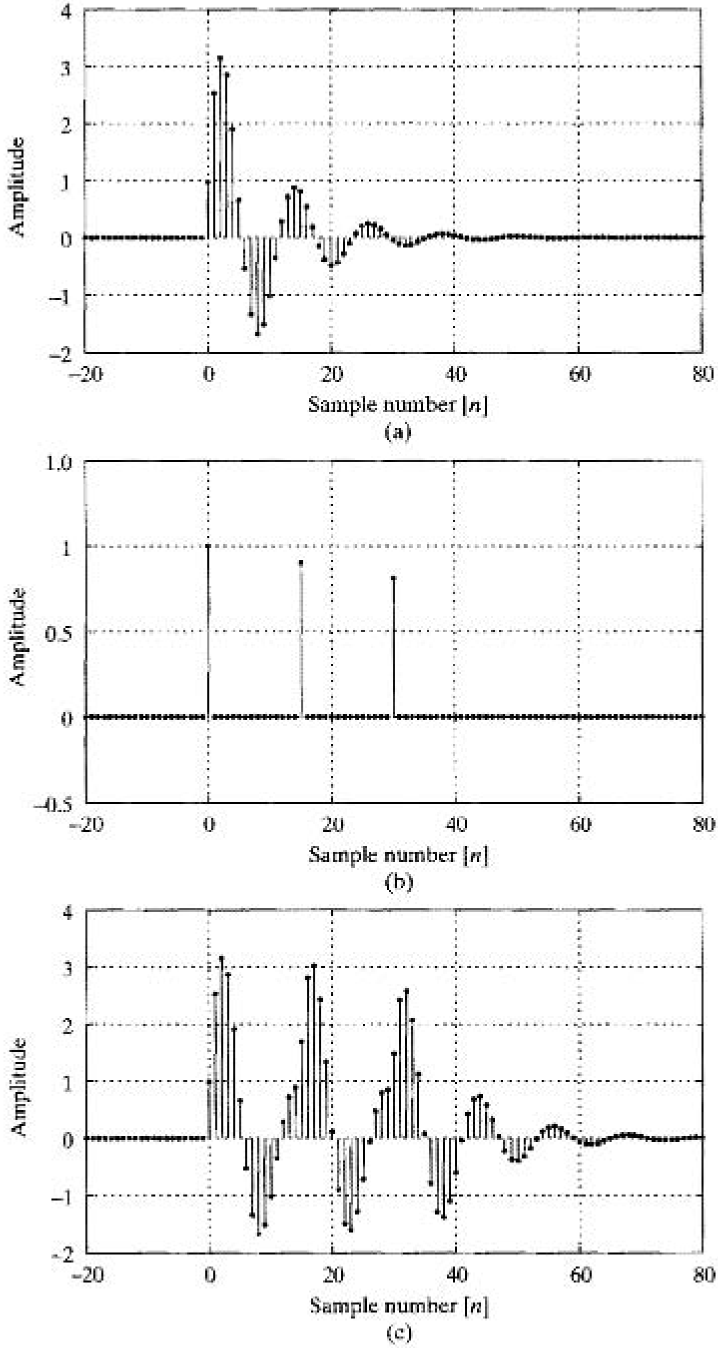


Figure 13.10 The sequences: (a) $v[n]$, (b) $p[n]$, and (c) $x[n]$ corresponding to the pole–zero plot of Figure 13.9.

of an acoustic pulse propagating in the earth due to a dynamite explosion or similar disturbance. The impulsive component $p[n]$ would represent reflections at boundaries between layers having different propagation characteristics. In the practical use of such a model, there would be more impulses in $p[n]$ than we assumed in Eq. (13.93a), and they would be unequally spaced. Also, the component $V(z)$ would generally involve many more zeros, and often no poles are included in the model (Ulrych, 1971; Tribolet, 1979; Robinson and Treitel, 1980).

Although the model discussed above is a highly simplified representation of that encountered in typical applications, it is analytically convenient and useful to obtain exact formulas to compare with computed results obtained for sampled signals. Furthermore, we will see that this simple model illustrates all the important properties of the cepstrum of a signal with a rational z-transform.

In Section 13.9.1, we evaluate analytically the complex cepstrum for the received signal $x[n]$. In Section 13.9.2, we illustrate the computation of the complex cepstrum using the DFT, and in Section 13.9.3 illustrate the technique of homomorphic deconvolution.

13.9.1 Computation of the Complex Cepstrum by z-Transform Analysis

To determine an equation for $\hat{x}[n]$, the complex cepstrum of $x[n]$ for the simple model of Eq. (13.92a), we use the relations

$$\hat{x}[n] = \hat{v}[n] + \hat{p}[n], \tag{13.95a}$$

$$\hat{X}(z) = \hat{V}(z) + \hat{P}(z), \tag{13.95b}$$

$$\hat{X}(z) = \log[X(z)], \tag{13.96a}$$

$$\hat{V}(z) = \log[V(z)], \tag{13.96b}$$

and

$$\hat{P}(z) = \log[P(z)]. \tag{13.96c}$$

To determine $\hat{v}[n]$, we can directly apply the results in Section 13.5. Specifically, to express $V(z)$ in the form of Eq. (13.29), we first note that for the specific signal $X(z)$ in Figure 13.9, the poles of $V(z)$ are inside the unit circle and the zero is outside ($r = 0.9$ and $b_0/b_1 = 0.98$), so that in accordance with Eq. (13.29), we rewrite $V(z)$ as

$$V(z) = \frac{b_1 z^{-1}(1 + (b_0/b_1)z)}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})}, \qquad |z| > |r|. \tag{13.97}$$

As discussed in Section 13.5, the factor z^{-1} contributes a linear component to the unwrapped phase that will force a discontinuity at $\omega = \pm\pi$ in the Fourier transform of $\hat{v}[n]$, so $\hat{V}(z)$ will not be analytic on the unit circle. To avoid this problem, we can alter $v[n]$ (and therefore also $x[n]$) with a one-sample time shift so that we evaluate instead the complex cepstrum of $v[n+1]$ and, consequently, also $x[n+1]$. If $x[n]$ or $v[n]$ is to be resynthesized after some processing of the complex cepstrum, we can remember this time shift and compensate for it at the final output.

With $v[n]$ replaced by $v[n+1]$, and correspondingly $V(z)$ replaced by $zV(z)$, we now consider $V(z)$ to have the form

$$V(z) = \frac{b_1(1 + (b_0/b_1)z)}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})}. \tag{13.98}$$

From Eqs. (13.36a) to (13.36c), we can write $\hat{v}[n]$ exactly as

$$\hat{v}[n] = \begin{cases} \log b_1, & n = 0, \quad (13.99a) \\ \frac{1}{n}[(re^{j\theta})^n + (re^{-j\theta})^n], & n > 0, \quad (13.99b) \\ \frac{1}{n}\left(\frac{-b_0}{b_1}\right)^{-n}, & n < 0. \quad (13.99c) \end{cases}$$

To determine $\hat{p}[n]$, we can evaluate the inverse z-transform of $\hat{P}(z)$, which, from Eq. (13.93b), is

$$\hat{P}(z) = \log(1 - \beta^3 z^{-3N_0}) - \log(1 - \beta z^{-N_0}), \tag{13.100}$$

where for our example $\beta = 0.9$, and consequently, $|\beta| < 1$. One approach to determining the inverse z-transform of Eq. (13.100) is to use the power series expansion of $\hat{P}(z)$. Specifically, since $|\beta| < 1$,

$$\hat{P}(z) = -\sum_{k=1}^{\infty} \frac{\beta^{3k}}{k} z^{-3N_0 k} + \sum_{k=1}^{\infty} \frac{\beta^k}{k} z^{-N_0 k}, \tag{13.101}$$

from which it follows that $\hat{p}[n]$ is

$$\hat{p}[n] = -\sum_{k=1}^{\infty} \frac{\beta^{3k}}{k} \delta[n - 3N_0 k] + \sum_{k=1}^{\infty} \frac{\beta^k}{k} \delta[n - N_0 k]. \tag{13.102}$$

An alternative approach to obtaining $\hat{p}[n]$ is to use the property developed in Problem 13.28.

From Eq. (13.95a), the complex cepstrum of $x[n]$ is

$$\hat{x}[n] = \hat{v}[n] + \hat{p}[n], \tag{13.103}$$

where $\hat{v}[n]$ and $\hat{p}[n]$ are given by Eqs. (13.99a) to (13.99c) and (13.102), respectively. The sequences $\hat{v}[n]$, $\hat{p}[n]$, and $\hat{x}[n]$ are shown in Figure 13.11.

The cepstrum of $x[n]$, $c_x[n]$, is the even part of $\hat{x}[n]$, i.e.,

$$c_x[n] = \tfrac{1}{2}(\hat{x}[n] + \hat{x}[-n]) \tag{13.104}$$

and furthermore

$$c_x[n] = c_v[n] + c_p[n]. \tag{13.105}$$

From Eqs. (13.99a) to (13.99c),

$$\begin{aligned} c_v[n] = \log(b_1)\delta[n] &+ \sum_{k=1}^{\infty} \frac{(-1)^k (b_0/b_1)^{-k}}{2k} (\delta[n-k] + \delta[n+k]) \\ &+ \sum_{k=1}^{\infty} \frac{r^k \cos(\theta k)}{k} (\delta[n-k] + \delta[n+k]), \end{aligned} \tag{13.106a}$$

and from Eq. (13.102),

$$\begin{aligned} c_p[n] = &-\frac{1}{2} \sum_{k=1}^{\infty} \frac{\beta^{3k}}{k} \{\delta[n - 3N_0 k] + \delta[n + 3N_0 k]\} \\ &+ \frac{1}{2} \sum_{k=1}^{\infty} \frac{\beta^k}{k} \{\delta[n - N_0 k] + \delta[n + N_0 k]\}. \end{aligned} \tag{13.106b}$$

The sequences $c_v[n]$, $c_p[n]$, and $c_x[n]$ for this example are shown in Figure 13.12.

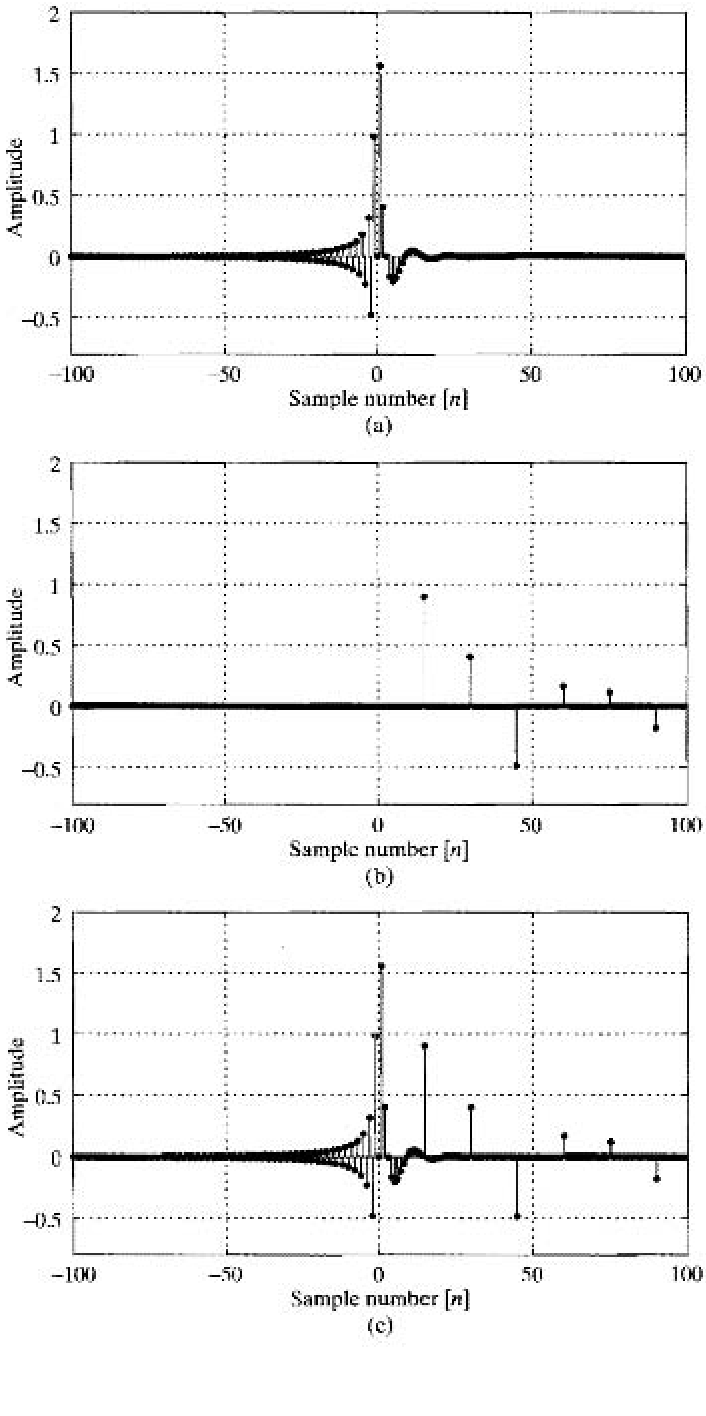


Figure 13.11 The sequences: (a) $\hat{v}[n]$, (b) $\hat{p}[n]$, and (c) $\hat{x}[n]$.

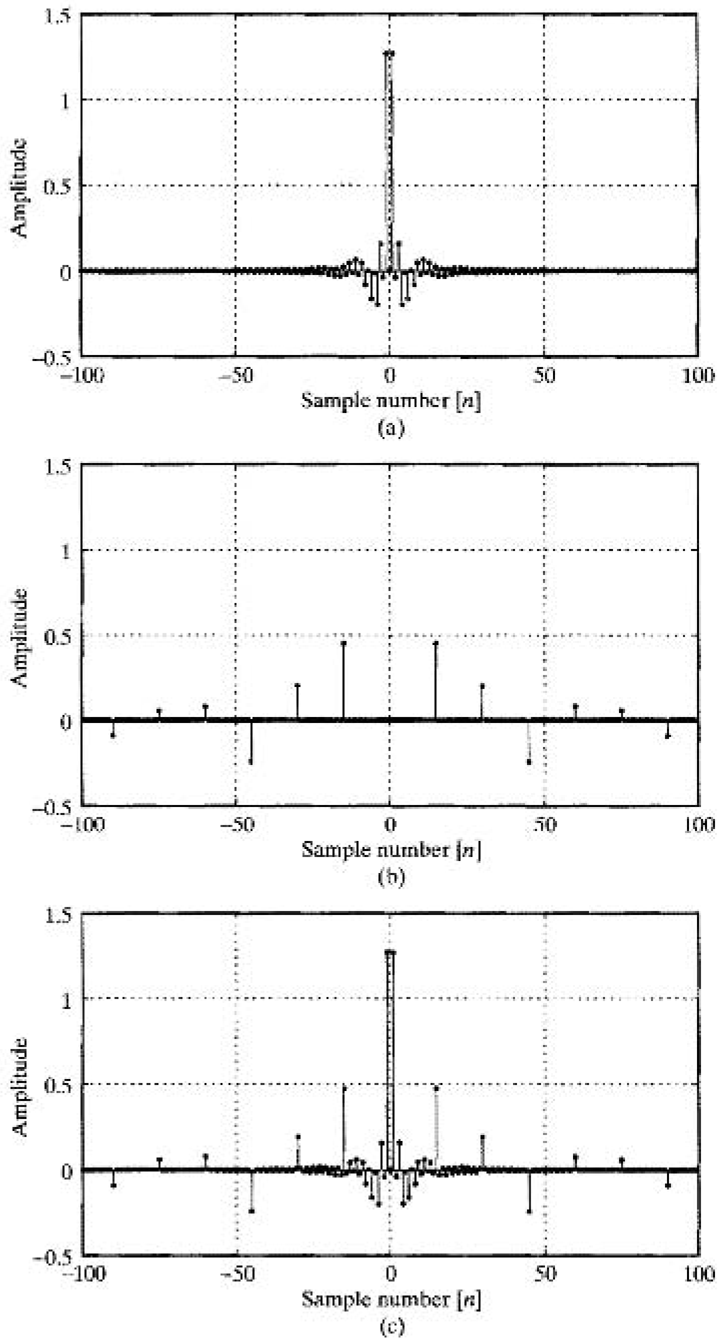


Figure 13.12 The sequences: (a) $c_V[n]$, (b) $c_p[n]$, and (c) $c_X[n]$.

13.9.2 Computation of the Cepstrum Using the DFT

In Figures 13.11 and 13.12, we showed the complex cepstra and the cepstra corresponding to evaluating the analytical expressions obtained in Section 13.9.1. In most applications, we do not have simple mathematical formulas for the signal values, and consequently, we cannot analytically determine $\hat{x}[n]$ or $c_x[n]$. However, for finite-length sequences, we can use either polynomial rooting or the DFT to compute the complex cepstrum. In this section, we illustrate the use of the DFT in the computation of the complex cepstrum and the cepstrum of $x[n]$ for the example of this section.

To compute the complex cepstrum or the cepstrum using the DFT as in Figure 13.4(a), it is necessary that the input be of finite extent. Thus, for the signal model discussed at the beginning of this section, $x[n]$ must be truncated. In the examples discussed in this section, the signal $x[n]$ in Figure 13.10(c) was truncated to $N = 1024$ samples and 1024-point DFTs were used in the system of Figure 13.4(a) to compute the complex cepstrum and the cepstrum of the signal. Figure 13.13 shows the Fourier transforms that are involved in the computation of the complex cepstrum. Figure 13.13(a) shows the logarithm of the magnitude of the DFT of 1024 samples of $x[n]$ in Figure 13.10, with the DFT samples connected in the plot to suggest the appearance of the DTFT of the finite-length input sequence. Figure 13.13(b) shows the principal value of the phase. Note the discontinuities as the phase exceeds $\pm\pi$ and wraps around modulo 2π. Figure 13.13(c) shows the continuous "unwrapped" phase curve obtained as discussed in Section 13.6.1. As discussed above, and as is evident by carefully comparing Figures 13.13(b) and 13.13(c), a linear-phase component corresponding to a delay of one sample has been removed so that the unwrapped phase curve is continuous at 0 and π. Thus, the unwrapped phase of Figure 13.13(c) corresponds to $x[n+1]$ rather than $x[n]$.

Figures 13.13(a) and 13.13(c) correspond to the computation of samples of the real and imaginary parts, respectively, of the DTFT of the complex cepstrum. Only the frequency range $0 \le \omega \le \pi$ is shown, since the function of Figure 13.13(a) is even and periodic with period 2π, and the function of Figure 13.13(c) is odd and periodic with period 2π. In examining the plots in Figures 13.13(a) and 13.13(c), we note that they have the general appearance of a rapidly varying, periodic (in frequency) component added to a more slowly varying component. The periodically varying component in fact corresponds to $\hat{P}(e^{j\omega})$ and the more slowly varying component to $\hat{V}(e^{j\omega})$.

In Figure 13.14(a), we show the inverse Fourier transform of the complex logarithm of the DFT, i.e., the time-aliased complex cepstrum $\hat{x}_p[n]$. Note the impulses at integer multiples of $N_0 = 15$. These are contributed by $\hat{p}[n]$ and correspond to the rapidly varying periodic component observed in the logarithm of the DFT. We see also that since the input signal is not minimum phase, the complex cepstrum is nonzero for $n < 0$.[9]

Since a large number of points were used in computing the DFTs, the time-aliased complex cepstrum differs very little from the exact values that would be obtained by

[9]In using the DFT to obtain the inverse Fourier transform of Figures 13.13(a) and 13.13(c), the values associated with $n < 0$ would normally appear in the interval $N/2 < n \le N-1$. Traditionally, time sequences are displayed with $n = 0$ in the center, so we have repositioned $\hat{x}_p[n]$ accordingly and have shown only a total of 201 points symmetrically about $n = 0$.

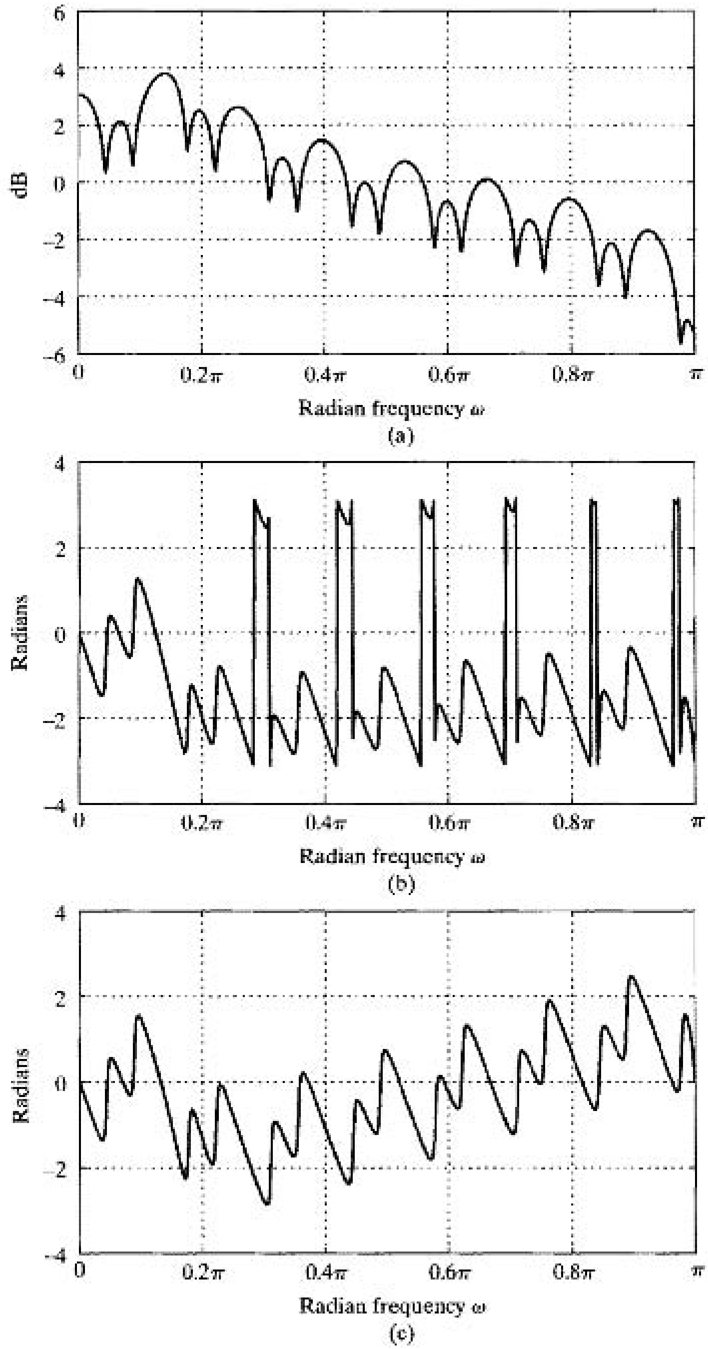


Figure 13.13 Fourier transforms of $x[n]$ in Figure 13.10. (a) Log magnitude. (b) Principal value of the phase. (c) Continuous "unwrapped" phase after removing a linear-phase component from part (b). The DFT samples are connected by straight lines.

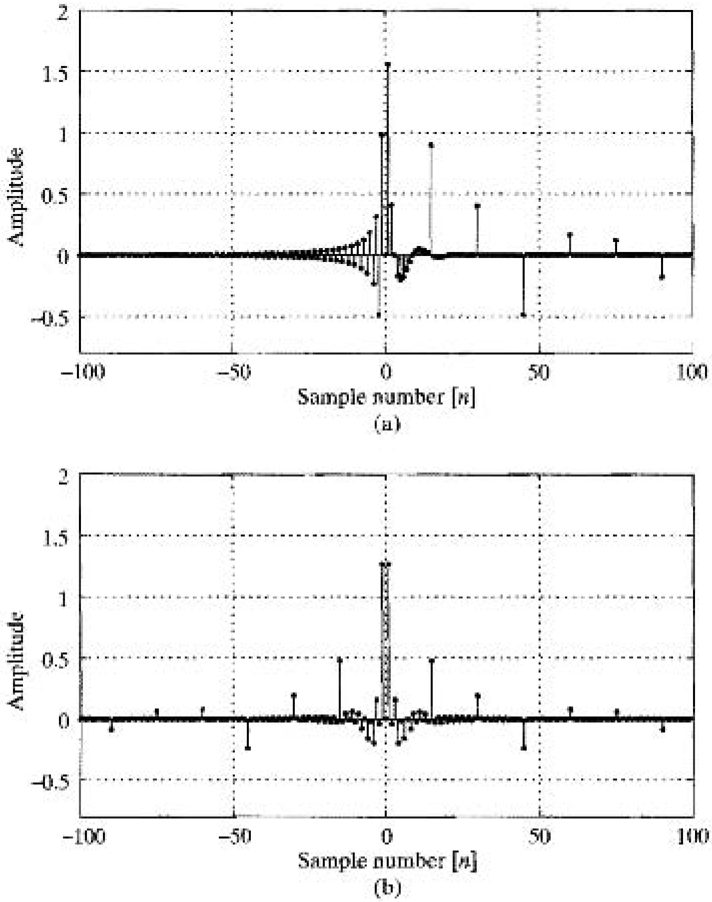


Figure 13.14 (a) Complex cepstrum $\hat{x}_p[n]$ of sequence in Figure 13.10(c). (b) Cepstrum $c_x[n]$ of sequence in Figure 13.10(c).

evaluating Eqs. (13.99a) to (13.99c), (13.102), and (13.103) for the specific values of the parameters used to generate the input signal of Figure 13.10.

The time-aliased cepstrum $c_{xp}[n]$ for this example is shown in Figure 13.14(b). As with the complex cepstrum, impulses at multiples of 15 are evident, corresponding to the periodic component of the logarithm of the magnitude of the Fourier transform.

As mentioned at the beginning of this section, convolution of a signal $v[n]$ with an impulse train such as $p[n]$ is a model for a signal containing multiple echoes. Since $x[n]$ is a convolution of $v[n]$ and $p[n]$, the echo times are often not easily detected by examining $x[n]$. In the cepstral domain, however, the effect of $p[n]$ is present as an additive impulse train, and consequently, the presence and location of the echoes are often more evident. As discussed in Section 13.1, it was this observation that motivated the proposal by Bogert, Healy and Tukey (1963) that the cepstrum be used as a means for detecting echoes. This same idea was later used by Noll (1967) as a basis for detecting vocal pitch in speech signals.

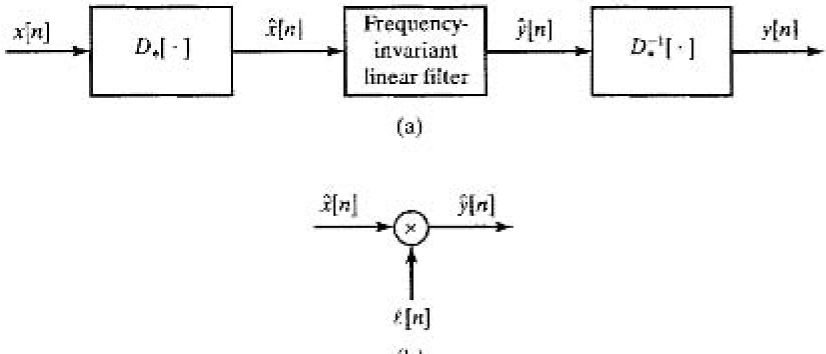


Figure 13.15 (a) System for homomorphic deconvolution. (b) Time-domain representation of frequency-invariant filtering.

13.9.3 Homomorphic Deconvolution for the Multipath Model

For the multipath model that is the basis for Section 13.9, the slowly varying component of the complex logarithm, and equivalently the "low-time" (low-quefrency) portion of the complex cepstrum, were mainly due to $v[n]$. Correspondingly, the more rapidly varying component of the complex logarithm and the "high-time" (high-quefrency) portion of the complex cepstrum were due primarily to $p[n]$. This suggests that the two convolved components of $x[n]$ can be separated by applying linear filtering to the logarithm of the Fourier transform (i.e., frequency invariant filtering), or, equivalently the complex cepstrum components can be separated by windowing or time gating the complex cepstrum.

Figure 13.15(a) depicts the operations involved in separation of the components of a convolution by filtering the complex logarithm of the Fourier transform of a signal. The frequency-invariant linear filter can be implemented by convolution in the frequency domain or, as indicated in Figure 13.15(b), by multiplication in the time domain. Figure 13.16(a) shows the time response of a lowpass frequency-invariant linear system as required for recovering an approximation to $v[n]$, and Figure 13.16(b) shows the time response of a highpass frequency-invariant linear system for recovering an approximation to $p[n]$.[10]

Figure 13.17 shows the result of lowpass frequency-invariant filtering. The more rapidly varying curves in Figures 13.17(a) and 13.17(b) are the complex logarithm of the Fourier transform of the input signal, i.e., the Fourier transform of the complex cepstrum. The slowly varying (dashed) curves in Figures 13.17(a) and 13.17(b) are the real and imaginary parts, respectively, of the Fourier transform of $\hat{y}[n]$, when the frequency-invariant linear system $\ell[n]$ is of the form of Figure 13.16(a) with $N_1 = 14$, $N_2 = 14$, and with the system of Figure 13.15 implemented using DFTs of length $N = 1024$. Figure 13.17(c) shows the corresponding output $y[n]$. This sequence is the approximation to

[10]Figure 13.16 assumes that the systems $D_*[\cdot]$ and $D_*^{-1}[\cdot]$ are implemented using the DFT as in Figure 13.4.

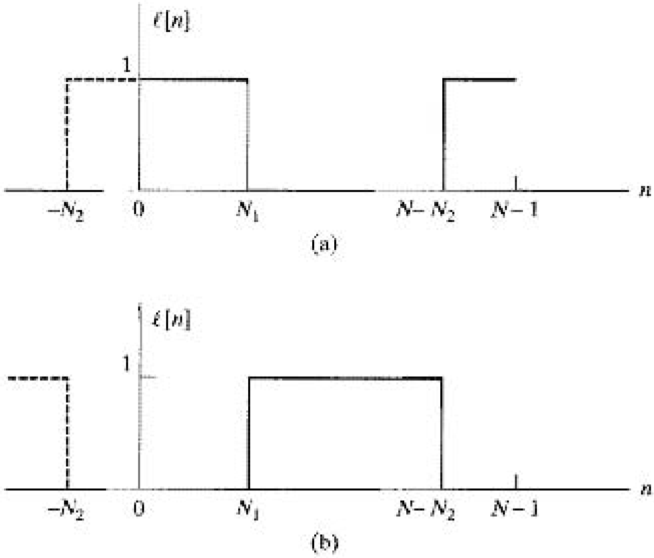


Figure 13.16 Time response of frequency-invariant linear systems for homomorphic deconvolution. (a) Lowpass system. (b) Highpass system. (Solid line indicates envelope of the sequence $\ell[n]$ as it would be applied in a DFT implementation. The dashed line indicates the periodic extension.)

$v[n]$ obtained by homomorphic deconvolution. To relate this output $y[n]$ to $v[n]$, recall that in computing the unwrapped phase, a linear-phase component was removed, corresponding to a one-sample time shift of $v[n]$. Consequently, $y[n]$ in Figure 13.17(c) corresponds to an approximation to $v[n+1]$ obtained by homomorphic deconvolution.

This type of filtering has been successfully used in speech processing to recover the vocal tract response information (Oppenheim, 1969b; Schafer and Rabiner, 1970) and in seismic signal analysis to recover seismic wavelets (Ulrych, 1971; Tribolet, 1979).

Figure 13.18 shows the result of highpass frequency-invariant filtering. The rapidly varying curves in Figures 13.18(a) and (b) are the real and imaginary parts, respectively, of the Fourier transform of $\hat{y}[n]$ when the frequency-invariant linear system $\ell[n]$ is of the form of Figure 13.16(b) with $N_1 = 14$ and $N_2 = 512$ (i.e., the negative-time parts are completely removed). Again, the system is implemented using a 1024-point DFT. Figure 13.18(c) shows the corresponding output $y[n]$. This sequence is the approximation to $p[n]$ obtained by homomorphic deconvolution. In contrast to the use of the cepstrum to *detect* echoes or periodicity, this approach seeks to obtain the impulse train that specifies the location and size of the repeated copies of $v[n]$.

13.9.4 Minimum-Phase Decomposition

In Section 13.8.1, we discussed ways that homomorphic deconvolution could be used to decompose a sequence into minimum-phase and allpass components or minimum-phase and maximum-phase components. We will apply these techniques to the signal model of Section 13.9. Specifically, for the parameters of the example, the z-transform of the input is

$$X(z) = V(z)P(z) = \frac{(0.98 + z^{-1})(1 + 0.9z^{-15} + 0.81z^{-30})}{(1 - 0.9e^{j\pi/6}z^{-1})(1 - 0.9e^{-j\pi/6}z^{-1})}. \tag{13.107}$$

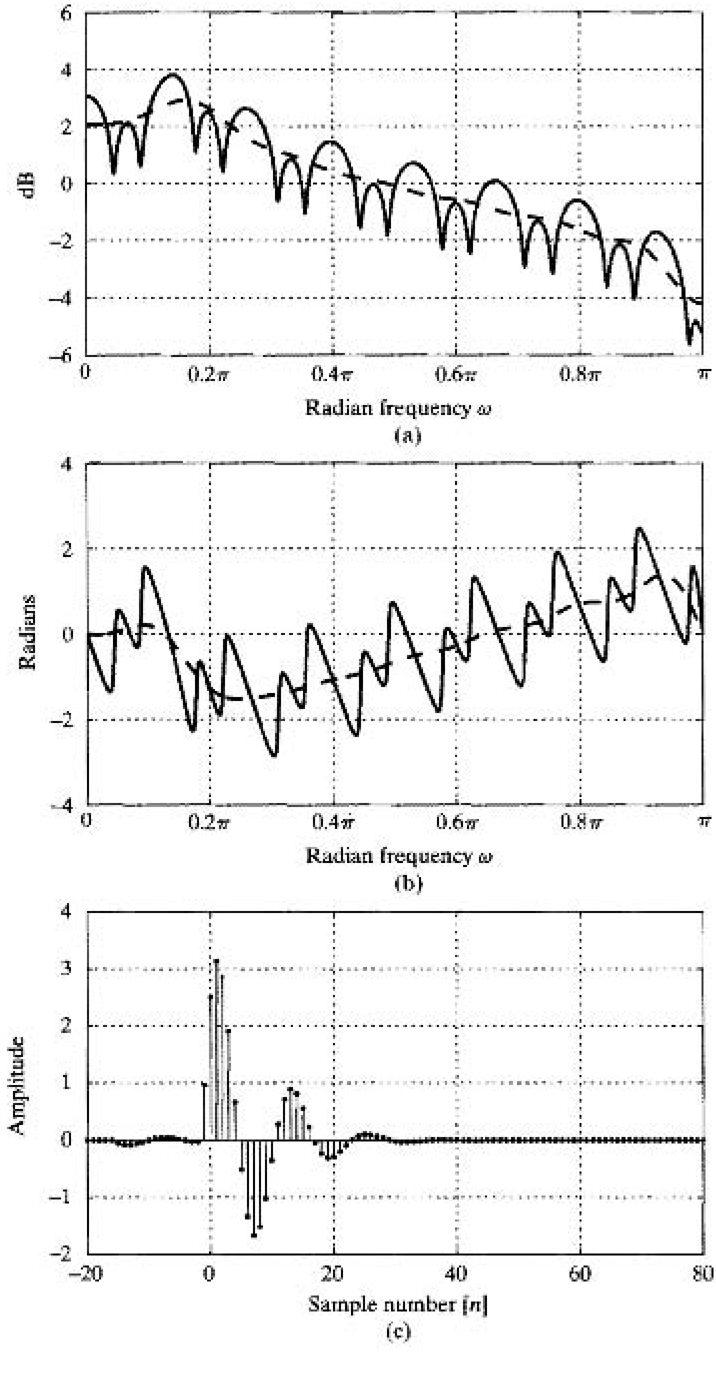


Figure 13.17 Lowpass frequency-invariant linear filtering in the system of Figure 13.15. (a) Real parts of the Fourier transforms of the input (solid line) and output (dashed line) of the lowpass system with $N_1 = 14$ and $N_2 = 14$ in Figure 13.16(a). (b) Imaginary parts of the input (solid line) and output (dashed line). (c) Output sequence $y[n]$ for the input of Figure 13.10(c).

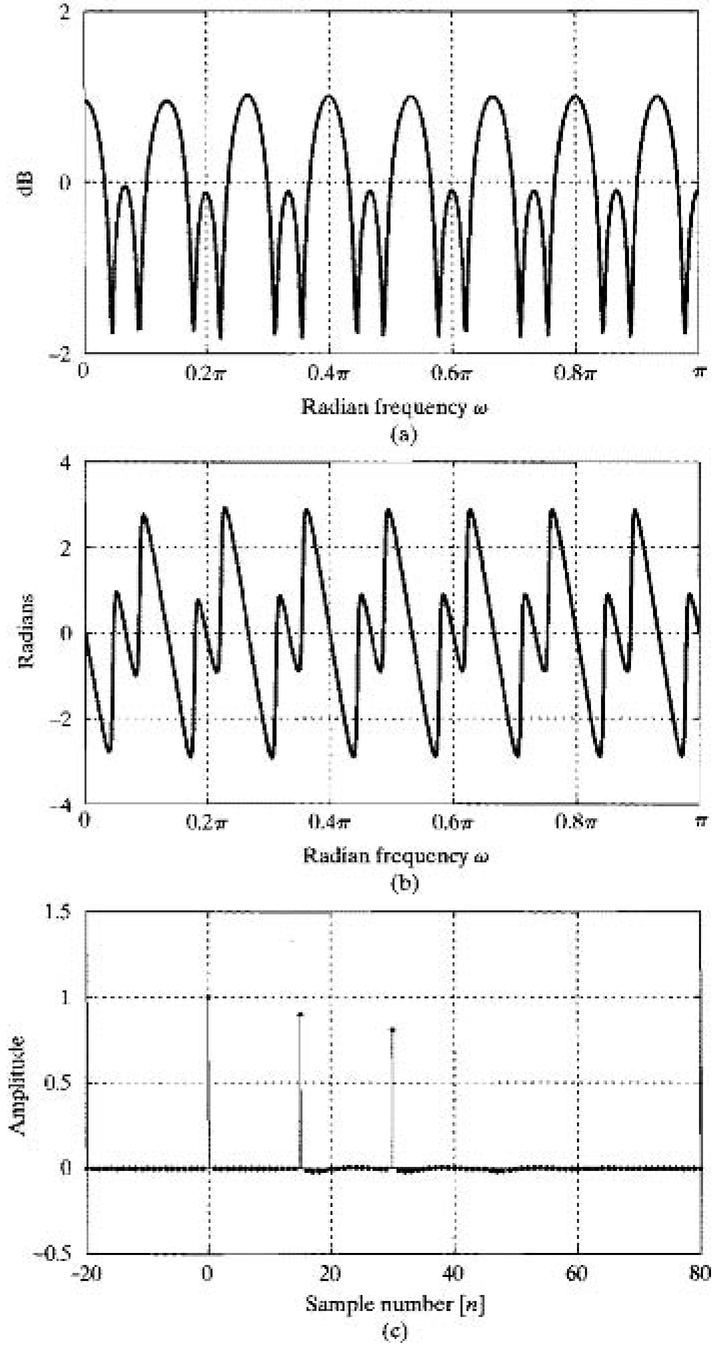


Figure 13.18 Illustration of highpass frequency-invariant linear filtering in the system of Figure 13.15. (a) Real part of the Fourier transform of the output of the highpass frequency-invariant system with $N_1 = 14$ and $N_2 = 512$ in Figure 13.16(b). (b) Imaginary part for conditions of part (a). (c) Output sequence $y[n]$ for the input of Figure 13.10.

First, we can write $X(z)$ as the product of a minimum-phase z-transform and an allpass z-transform; i.e.,

$$X(z) = X_{min}(z)X_{ap}(z), \tag{13.108}$$

where

$$X_{min}(z) = \frac{(1+0.98z^{-1})(1+0.9z^{-15}+0.81z^{-30})}{(1-0.9e^{j\pi/6}z^{-1})(1-0.9e^{-j\pi/6}z^{-1})} \tag{13.109}$$

and

$$X_{ap}(z) = \frac{0.98+z^{-1}}{1+0.98z^{-1}}. \tag{13.110}$$

The sequences $x_{min}[n]$ and $x_{ap}[n]$ can be found using the partial fraction expansion methods of Chapter 3, and the corresponding complex cepstra $\hat{x}_{min}[n]$ and $\hat{x}_{ap}[n]$ can be found using the power series technique of Section 13.5 (see Problem 13.25). Alternatively, $\hat{x}_{min}[n]$ and $\hat{x}_{ap}[n]$ can be obtained exactly from $\hat{x}[n]$ by the operations discussed in Section 13.8.1 and as depicted in Figure 13.7. If the characteristic systems in Figure 13.7 are implemented using the DFT, then the separation is only approximate since $x_{ap}[n]$ is infinitely long, but the approximation error can be small over the interval where $x_{ap}[n]$ is large if the DFT length is large enough. Figure 13.19(a) shows the complex cepstrum for $x[n]$ as computed using a 1024-point DFT, again with a one-sample time delay removed from $v[n]$ so that the phase is continuous at π. Figure 13.19(b) shows the complex cepstrum of the minimum-phase component $\hat{x}_{min}[n]$, and Figure 13.19(c) shows the complex cepstrum of the allpass component $\hat{x}_{ap}[n]$ as obtained by the operations of Figure 13.7 with $D_*[\cdot]$ implemented as in Figure 13.4(a).

Using the DFT as in Figure 13.4(b) to implement the system $D_*^{-1}[\cdot]$ gives the approximations to the minimum-phase and allpass components shown in Figures 13.20(a) and 13.20(b), respectively. Since all the zeros of $P(z)$ are inside the unit circle, all of $P(z)$ is included in the minimum-phase z-transform or, equivalently, $\hat{p}[n]$ is entirely included in $\hat{x}_{min}[n]$. Thus, the minimum-phase component consists of delayed and scaled replicas of the minimum-phase component of $v[n]$. Therefore, the minimum-phase component of Figure 13.20(a) appears very similar to the input shown in Figure 13.10(c). From Eq. (13.110), the allpass component can be shown to be

$$x_{ap}[n] = 0.98\delta[n] + 0.0396(-0.98)^{n-1}u[n-1]. \tag{13.111}$$

The result of Figure 13.20(b) is very close to this ideal result for small values of n where the sequence values are of significant amplitude. This example illustrates a technique of decomposition that has been applied by Bauman, Lipshitz and Vanderkooy (1985) in the analysis and characterization of the response of electroacoustic transducers. A similar decomposition technique can be used to factor magnitude-squared functions as required in digital filter design (see Problem 13.27).

As an alternative to the minimum-phase/allpass decomposition, we can express $X(z)$ as the product of a minimum-phase z-transform and a maximum-phase z-transform; i.e.,

$$X(z) = X_{mn}(z)X_{mx}(z), \tag{13.112}$$

where

$$X_{mn}(z) = \frac{z^{-1}(1+0.9z^{-15}+0.81z^{-30})}{(1-0.9e^{j\pi/6}z^{-1})(1-0.9e^{-j\pi/6}z^{-1})} \tag{13.113}$$

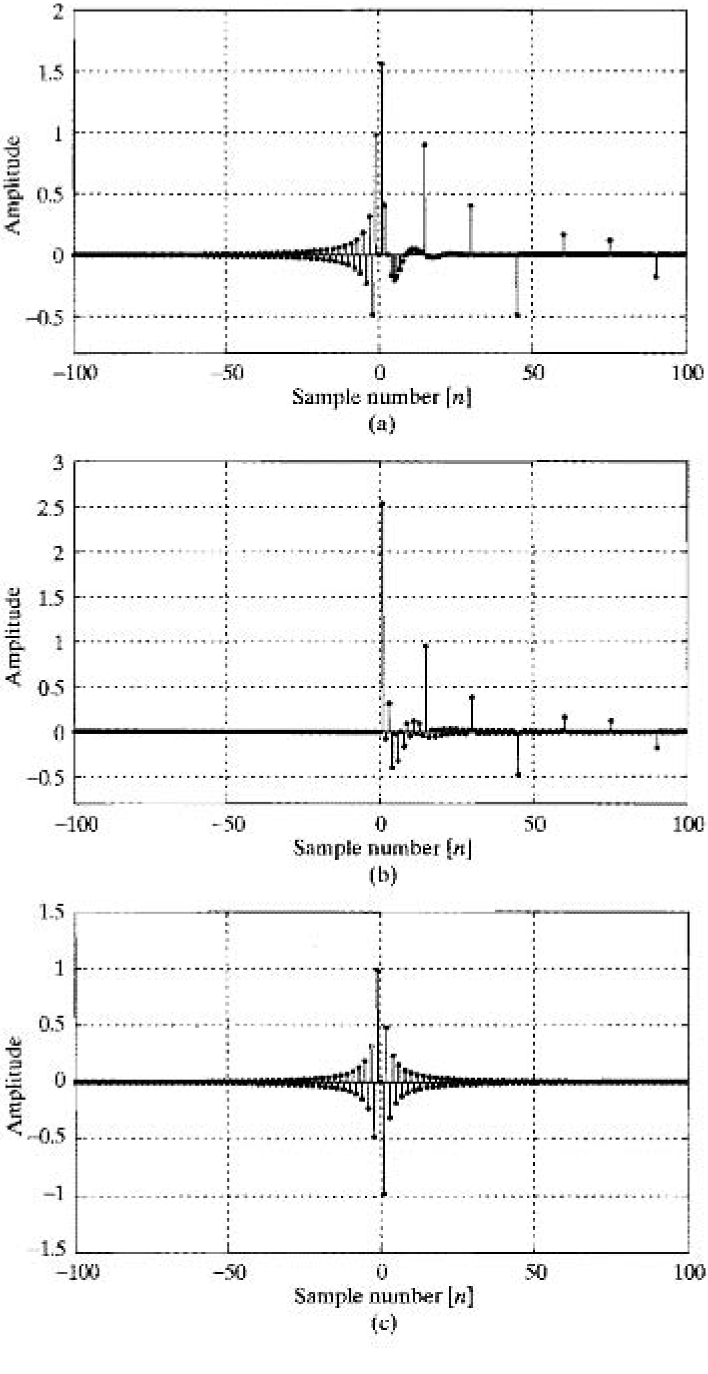


Figure 13.19 (a) Complex cepstrum of $x[n] = x_{min} n * x_{ap}[n]$. (b) Complex cepstrum of $x_{min}[n]$. (c) Complex cepstrum of $x_{ap}[n]$.

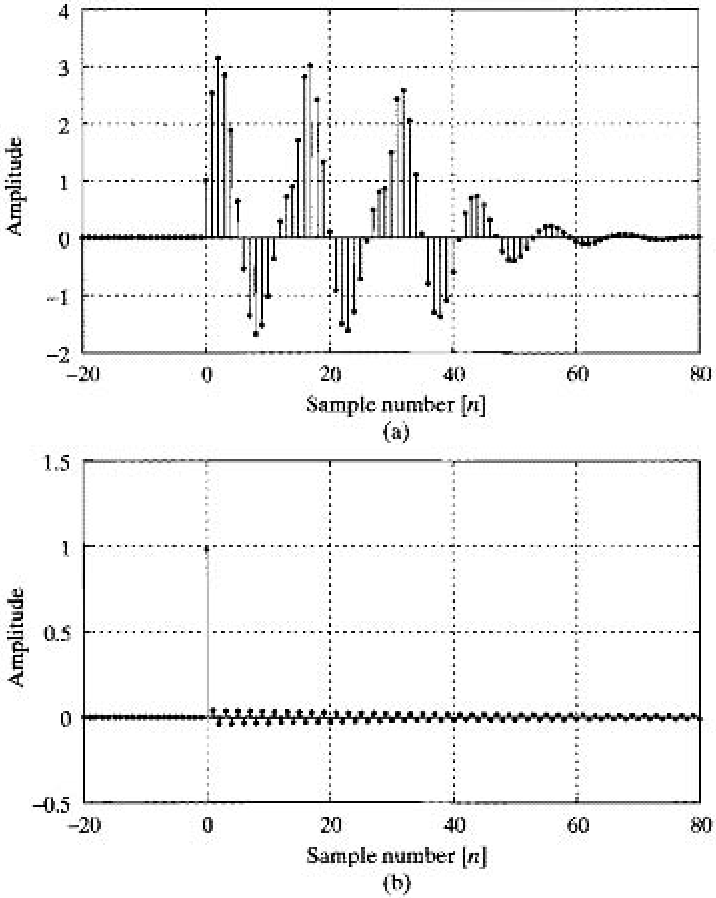


Figure 13.20 (a) Minimum-phase output. (b) Allpass output obtained as depicted in Figure 13.7.

and

$$X_{mx}(z) = 0.98z + 1. \tag{13.114}$$

The sequences $x_{mn}[n]$ and $x_{mx}[n]$ can be found using the partial fraction expansion methods of Chapter 3, and the corresponding complex cepstra $\hat{x}_{mn}[n]$ and $\hat{x}_{mx}[n]$ can be found using the power series technique of Section 13.5 (see Problem 13.25). Alternatively, $\hat{x}_{mn}[n]$ and $\hat{x}_{mx}[n]$ can be obtained exactly from $\hat{x}[n]$ by the operations discussed in Section 13.8.2 and as depicted in Figure 13.8, where

$$\ell_{mn}[n] = u[n] \tag{13.115}$$

and

$$\ell_{mx}[n] = u[-n-1]. \tag{13.116}$$

That is, the minimum-phase sequence is now defined by the positive time part of the complex cepstrum and the maximum-phase part is defined by the negative time part of the complex cepstrum. If the characteristic systems in Figure 13.8 are implemented using the DFT, the negative time part of the complex cepstrum is positioned in the last half of the DFT interval. In this case, the separation of the minimum-phase and maximum-phase components is only approximate because of time aliasing, but the time-aliasing error can be made small by choosing a sufficiently large DFT length. Figure 13.19(a) shows the complex cepstrum of $x[n]$ as computed using a 1024-point DFT. Figure 13.21 shows the two output sequences that are obtained from the complex cepstrum of Figure 13.19(a) using Eqs. (13.87) and (13.88) as in Fig 13.8 with the inverse characteristic system being implemented using the DFT as in Figure 13.4(b). As before, since $\hat{p}[n]$ is entirely included in $\hat{x}_{mn}[n]$, the corresponding output $x_{mn}[n]$ consists of delayed and scaled replicas of a minimum-phase sequence, thus, it also looks very much like the input sequence. However, a careful comparison of Figures 13.20(a) and 13.21(a) shows that $x_{min}[n] \neq x_{mn}[n]$. From Eq. (13.114), the maximum-phase sequence is

$$x_{mx}[n] = 0.98\delta[n+1] + \delta[n]. \tag{13.117}$$

Figure 13.21(b) is very close to this ideal result. (Note the shift due to the linear phase removed in the phase unwrapping.) This technique of minimum-phase/maximum-phase decomposition was used by Smith and Barnwell (1984) in the design and implementation of exact reconstruction filter banks for speech analysis and coding.

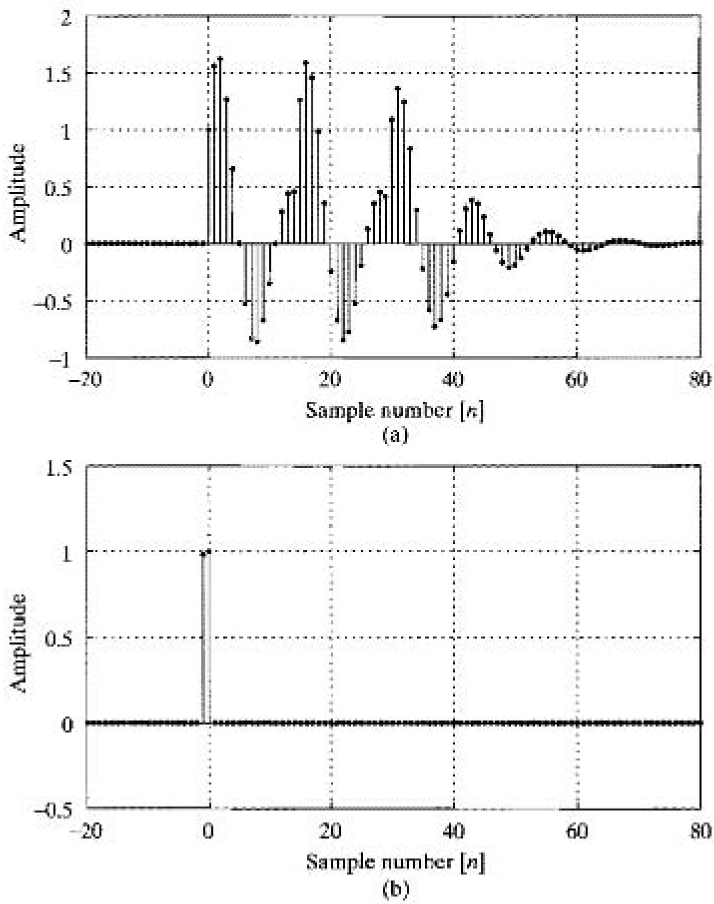


Figure 13.21 (a) Minimum-phase output. (b) Maximum-phase output obtained as depicted in Figure 13.8.

13.9.5 Generalizations

The example in Section 13.9 considered a simple exponential signal that was convolved with an impulse train to produce a series of delayed and scaled replicas of the exponential signal. This model illustrates many of the features of the complex cepstrum and of homomorphic filtering.

In particular, in more general models associated with speech, communication, and seismic applications an appropriate signal model consists of the convolution of two components. One component has the characteristics of $v[n]$, specifically a Fourier transform that is slowly varying in frequency. The second has the characteristics of $p[n]$, i.e., an echo pattern or impulse train for which the Fourier transform is more rapidly varying and quasiperiodic in frequency. Thus, the contributions of the two components would be separated in the complex cepstrum or the cepstrum, and, furthermore, the complex cepstrum or the cepstrum would contain impulses at multiples of the echo delays. Thus, homomorphic filtering can be used to separate the convolutional components of the signal, or the cepstrum can be used to detect echo delays. In the next section, we will illustrate the use of these general properties of the cepstrum in applications to speech analysis.

13.10 APPLICATIONS TO SPEECH PROCESSING

Cepstrum techniques have been applied successfully to speech analysis in a variety of ways. As discussed briefly in this section, the previous theoretical discussion and the extended example of Section 13.9 apply in a relatively straightforward way to speech analysis.

13.10.1 The Speech Model

As we briefly described in Section 10.4.1, there are three basic classes of speech sounds corresponding to different forms of excitation of the vocal tract. Specifically:

- *Voiced sounds* are produced by exciting the vocal tract with quasiperiodic pulses of airflow caused by the opening and closing of the glottis.
- *Fricative sounds* are produced by forming a constriction somewhere in the vocal tract and forcing air through the constriction so that turbulence is created, thereby producing a noise-like excitation.
- *Plosive sounds* are produced by completely closing off the vocal tract, building up pressure behind the closure, and then abruptly releasing the pressure.

In each case, the speech signal is produced by exciting the vocal tract system (an acoustic transmission system) with a wideband excitation. The vocal tract shape changes relatively slowly with time, thus, it can be modeled as a slowly time-varying filter that imposes its frequency-response properties on the spectrum of the excitation. The vocal tract is characterized by its natural frequencies (called *formants*), which correspond to resonances in its frequency response.

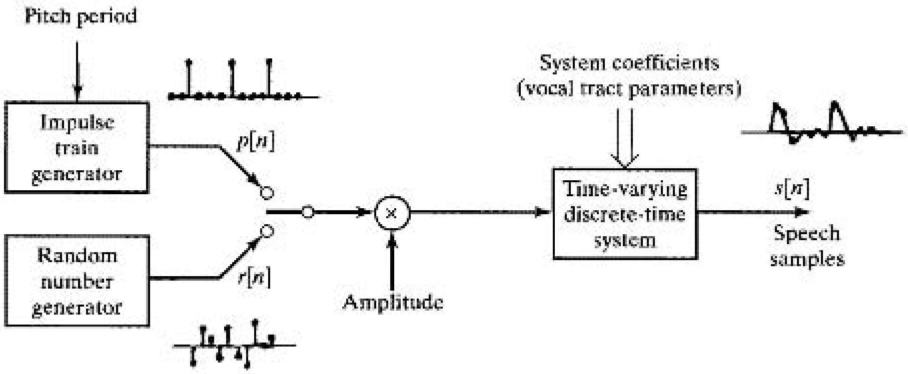


Figure 13.22 Discrete-time model of speech production.

If we assume that the excitation sources and the vocal tract shape are independent, we arrive at the discrete-time model of Figure 13.22 as a representation of the sampled speech waveform. In this model, samples of the speech signal are assumed to be the output of a time-varying discrete-time system that models the resonances of the vocal tract system. The mode of excitation of the system switches between periodic impulses and random noise, depending on the type of sound being produced.

Since the vocal tract shape changes rather slowly in continuous speech, it is reasonable to assume that the discrete-time system in the model has fixed properties over a time interval on the order of 10 ms. Thus, the discrete-time system may be characterized in each such time interval by an impulse response or a frequency response or a set of coefficients for an IIR system. Specifically, a model for the system function of the vocal tract takes the form

$$V(z) = \frac{\sum_{k=0}^{K} b_k z^{-k}}{\sum_{k=0}^{P} a_k z^{-k}} \tag{13.118}$$

or, equivalently,

$$V(z) = \frac{Az^{-K_o} \prod_{k=1}^{K_i}(1-\alpha_k z^{-1}) \prod_{k=1}^{K_o}(1-\beta_k z)}{\prod_{k=1}^{[P/2]}(1-r_k e^{j\theta_k} z^{-1})(1-r_k e^{-j\theta_k} z^{-1})}, \tag{13.119}$$

where the quantities $r_k e^{j\theta_k}$ (with $|r_k| < 1$ are the complex natural frequencies of the vocal tract, which, of course, are dependent on the vocal tract shape and consequently are time varying. The zeros of $V(z)$ account for the finite-duration glottal pulse waveform and for the zeros of transmission caused by the constrictions of the vocal tract in the creation of nasal voiced sounds and fricatives. Such zeros are often not included, because it is very difficult to estimate their locations from only the speech waveform. Also, it has been shown (Atal and Hanauer, 1971) that the spectral shape of the speech signal can be accurately modeled using no zeros, if we include extra poles beyond the number needed just to account for the vocal tract resonances. The zeros are included in our analysis, because they are necessary for an accurate representation of the complex cepstrum of speech. Note that we include the possibility of zeros outside the unit circle.

The vocal tract system is excited by an excitation sequence $p[n]$, which is a train of impulses when modeling voiced speech sounds and $r[n]$, which is a pseudorandom noise sequence when modeling unvoiced speech sounds, such as fricatives and plosives.

Many of the fundamental problems of speech processing reduce to the estimation of the parameters of the model of Figure 13.22. These parameters are as follows:

- The coefficients of $V(z)$ in Eq. (13.118) or the pole and zero locations in Eq. (13.119)
- The mode of excitation of the vocal tract system; i.e., a *periodic impulse train* or *random noise*
- The amplitude of the excitation signal
- The pitch period of the speech excitation for voiced speech.

Homomorphic deconvolution can be applied to the estimation of the parameters if it is assumed that the model is valid over a short time interval, so that a short segment of length L samples of the sampled speech signal can be thought of as the convolution

$$s[n] = v[n] * p[n] \qquad \text{for } 0 \leq n \leq L-1. \tag{13.120}$$

where $v[n]$ is the impulse response of the vocal tract and $p[n]$ is either periodic (for voiced speech) or random noise (for unvoiced speech). Obviously, Eq. (13.120) is not valid at the edges of the interval, because of pulses that occur before the beginning of the analysis interval and pulses that end after the end of the interval. To mitigate the effect of the "discontinuities" of the model at the ends of the interval, the speech signal $s[n]$ can be multiplied by a window $w[n]$ that tapers smoothly to zero at both ends. Thus, the input to the homomorphic deconvolution system is

$$x[n] = w[n]s[n]. \tag{13.121}$$

Let us first consider the case of voiced speech. If $w[n]$ varies slowly with respect to the variations of $v[n]$, then the analysis is greatly simplified if we assume that

$$x[n] = v[n] * p_w[n], \tag{13.122}$$

where

$$p_w[n] = w[n]p[n]. \tag{13.123}$$

(See Oppenheim and Schafer, 1968.) A more detailed analysis without this assumption leads to essentially the same conclusions as below (Verhelst and Steenhaut, 1986). For voiced speech, $p[n]$ is a train of impulses of the form

$$p[n] = \sum_{k=0}^{M-1} \delta[n - kN_0] \tag{13.124}$$

so that

$$p_w[n] = \sum_{k=0}^{M-1} w[kN_0]\delta[n - kN_0], \tag{13.125}$$

where we have assumed that the pitch period is N_0 and that M periods are spanned by the window.

The complex cepstra of $x[n]$, $v[n]$, and $p_w[n]$ are related by

$$\hat{x}[n] = \hat{v}[n] + \hat{p}_w[n]. \tag{13.126}$$

To obtain $\hat{p}_w[n]$, we define a sequence

$$w_{N_0}[k] = \begin{cases} w[kN_0], & k = 0, 1, \ldots, M-1, \\ 0, & \text{otherwise}, \end{cases} \tag{13.127}$$

whose Fourier transform is

$$P_w(e^{j\omega}) = \sum_{k=0}^{M-1} w[kN_0]e^{-\omega k N_0} = W_{N_0}(e^{j\omega N_0}). \tag{13.128}$$

Thus, $P_w(e^{j\omega})$ and $\hat{P}_w(e^{j\omega})$ are both periodic with period $2\pi/N_0$, and the complex cepstrum of $p_w[n]$ is

$$\hat{p}_w[n] = \begin{cases} \hat{w}_{N_0}[n/N_0], & n = 0, \pm N_0, \pm 2N_0, \ldots, \\ 0, & \text{otherwise}. \end{cases} \tag{13.129}$$

The periodicity of the complex logarithm resulting from the periodicity of the voiced speech signal is manifest in the complex cepstrum as impulses spaced at integer multiples of N_0 samples (the pitch period). If the sequence $w_{N_0}[n]$ is minimum phase, then $\hat{p}_w[n]$ will be zero for $n < 0$. Otherwise, $\hat{p}_w[n]$ will have impulses spaced at intervals of N_0 samples for both positive and negative values of n. In either case, the contribution of $\hat{p}_w[n]$ to $\hat{x}[n]$ will be found in the interval $|n| \geq N_0$.

From the power series expansion of the complex logarithm of $V(z)$, it can be shown that the contribution to the complex cepstrum due to $v[n]$ is

$$\hat{v}[n] = \begin{cases} \displaystyle\sum_{k=1}^{K_0} \frac{\beta_k^{-n}}{n}, & n < 0, \\ \log|A|, & n = 0, \\ \displaystyle -\sum_{k=1}^{K_i} \frac{\alpha_k^n}{n} + \sum_{k=1}^{\lfloor P/2 \rfloor} \frac{2r_k^n}{n} \cos(\theta_k n), & n > 0. \end{cases} \tag{13.130}$$

As with the simpler example in Section 13.9.1, the term z^{-K_0} in Eq. (13.119) represents a linear-phase factor that would be removed in obtaining the unwrapped phase and the complex cepstrum. Consequently, $\hat{v}[n]$ in Eq. (13.130) more accurately is the complex cepstrum of $v[n + K_0]$.

From Eq. (13.130), we see that the contributions of the vocal tract response to the complex cepstrum occupy the full range $-\infty < n < \infty$, but they are concentrated around $n = 0$. We note also that since the vocal tract resonances are represented by poles inside the unit circle, their contribution to the complex cepstrum is zero for $n < 0$.

13.10.2 Example of Homomorphic Deconvolution of Speech

For speech sampled at 10,000 samples/s, the pitch period N_0 will range from about 25 samples for a high-pitched voice up to about 150 samples for a very low-pitched voice. Since the vocal tract component of the complex cepstrum $\hat{v}[n]$ decays rapidly, the peaks of $\hat{p}_w[n]$ stand out from $\hat{v}[n]$. In other words, in the complex logarithm, the vocal tract components are slowly varying, and the excitation components are rapidly varying. This is illustrated by the following example. Figure 13.23(a) shows a segment of a speech wave

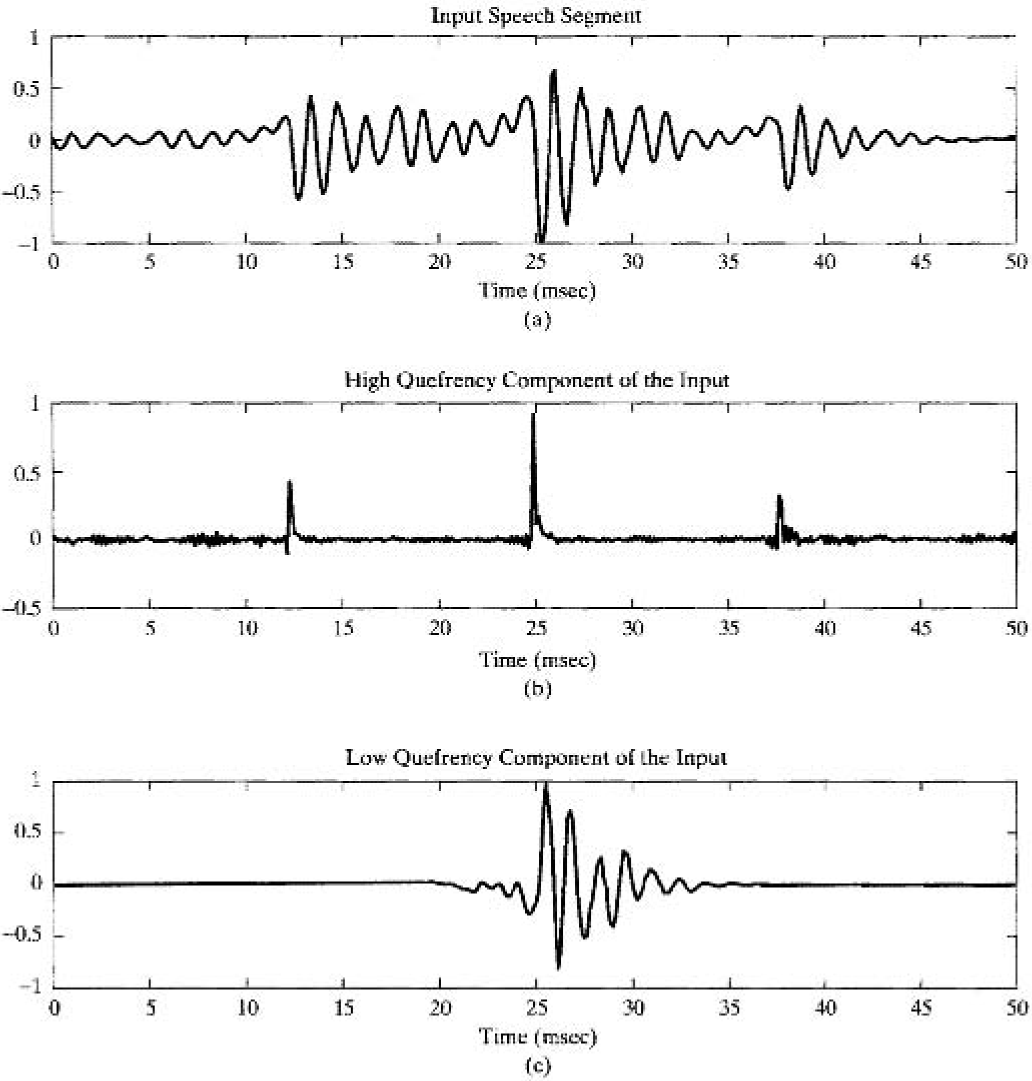


Figure 13.23 Homomorphic deconvolution of speech. (a) Segment of speech weighted by a Hamming window. (b) High quefrency component of the signal in (a). (c) Low quefrency component of the signal in (a).

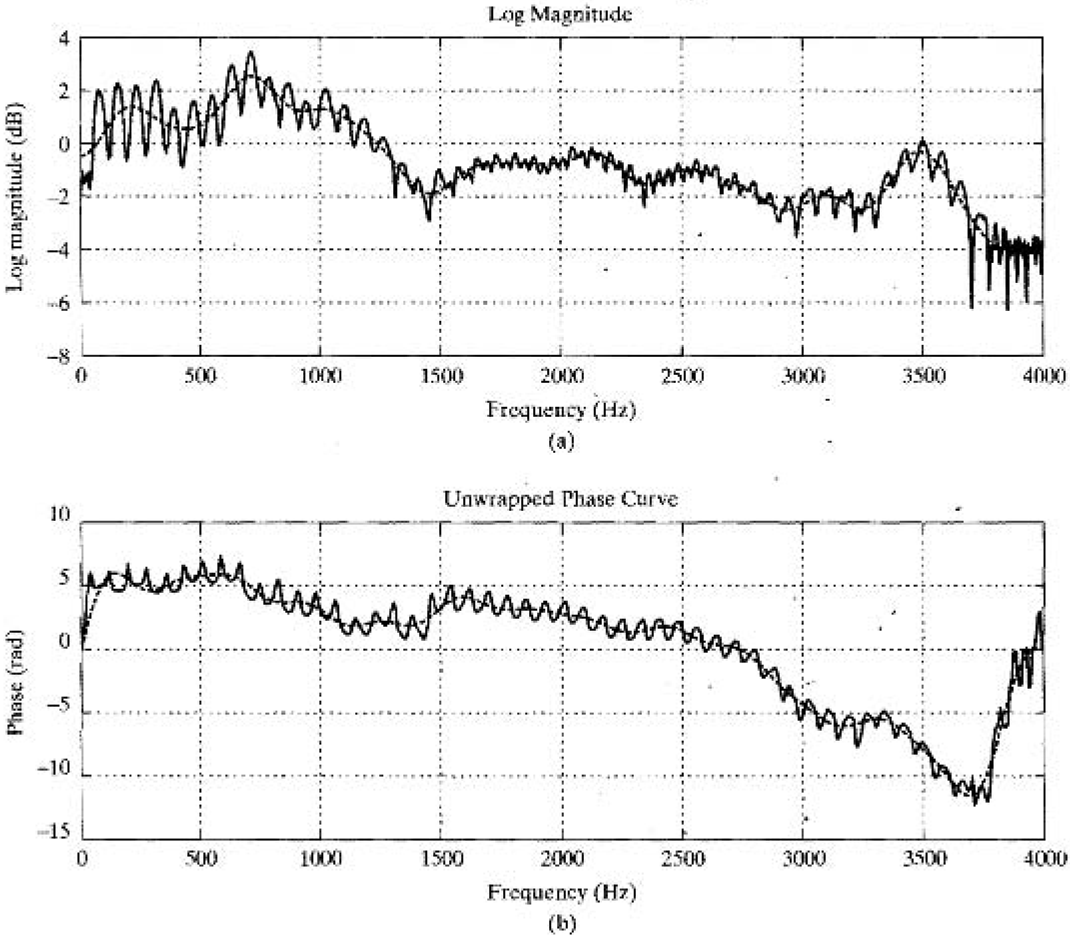


Figure 13.24 Complex logarithm of the signal of Figure 13.23(a): (a) Log magnitude. (b) Unwrapped phase.

multiplied by a Hamming window of length 401 samples (50 ms time duration at a sampling rate of 8000 samples/s). Figure 13.24 shows the complex logarithm (log magnitude and unwrapped phase) of the DFT of the signal in Figure 13.23(a).[11] Note the rapidly varying, almost periodic component due to $p_w[n]$ and the slowly varying component due to $v[n]$. These properties are manifest in the complex cepstrum of Figure 13.25 in the form of impulses at multiples of approximately 13 ms (the period of the input speech segment) due to $\hat{p}_w[n]$ and in the samples in the region $|nT| < 5$ ms, which we attribute to $\hat{v}[n]$. As in the previous section, frequency-invariant filtering can be used

[11] In all the figures of this section, the samples of all sequences were connected for ease in plotting.

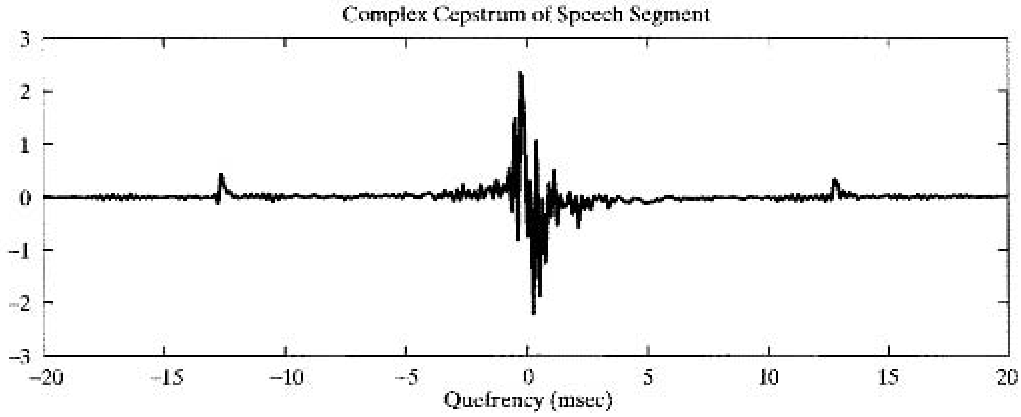


Figure 13.25 Complex cepstrum of the signal in Figure 13.23(a) (inverse DTFT of the complex logarithm in Figure 13.24).

to separate the components of the convolutional model of the speech signal. Lowpass filtering of the complex logarithm can be used to recover an approximation to $v[n]$, and highpass filtering can be used to obtain $p_w[n]$. Figure 13.23(c) shows an approximation to $v[n]$ obtained by using a lowpass frequency-invariant filter as in Figure 13.16(a) with $N_1 = 30$ and $N_2 = 30$. The slowly varying dotted curves in Figure 13.24 show the complex logarithm of the DTFT of the low quefrency component shown in Figure 13.23(c). On the other hand, Figure 13.23(b) is an approximation to $p_w[n]$ obtained by applying to the complex cepstrum a symmetrical highpass frequency-invariant filter as in Figure 13.16(b) with $N_1 = 95$ and $N_2 = 95$. In both cases, the inverse characteristic system was implemented by using 1024-point DFTs, as in Figure 13.4(b).

13.10.3 Estimating the Parameters of the Speech Model

Although homomorphic deconvolution can be successfully applied in *separating* the components of a speech waveform, in many speech processing applications we are interested only in *estimating* the parameters in a parametric representation of the speech signal. Since the properties of the speech signal change rather slowly with time, it is common to estimate the parameters of the model of Figure 13.22 at intervals of about 10 ms (100 times/s). In this case, the time-dependent Fourier transform discussed in Chapter 10 serves as the basis for time-dependent homomorphic analysis. For example, it may be sufficient to examine segments of speech selected about every 10 ms (100 samples at 10,000 Hz sampling rate) to determine the mode of excitation of the model (voiced or unvoiced) and, for voiced speech, the pitch period. Or we may wish to track the variation of the vocal tract resonances (formants). For such problems, the phase computation can be avoided by using the cepstrum, which requires only the logarithm of the magnitude of the Fourier transform. Since the cepstrum is the even part of the

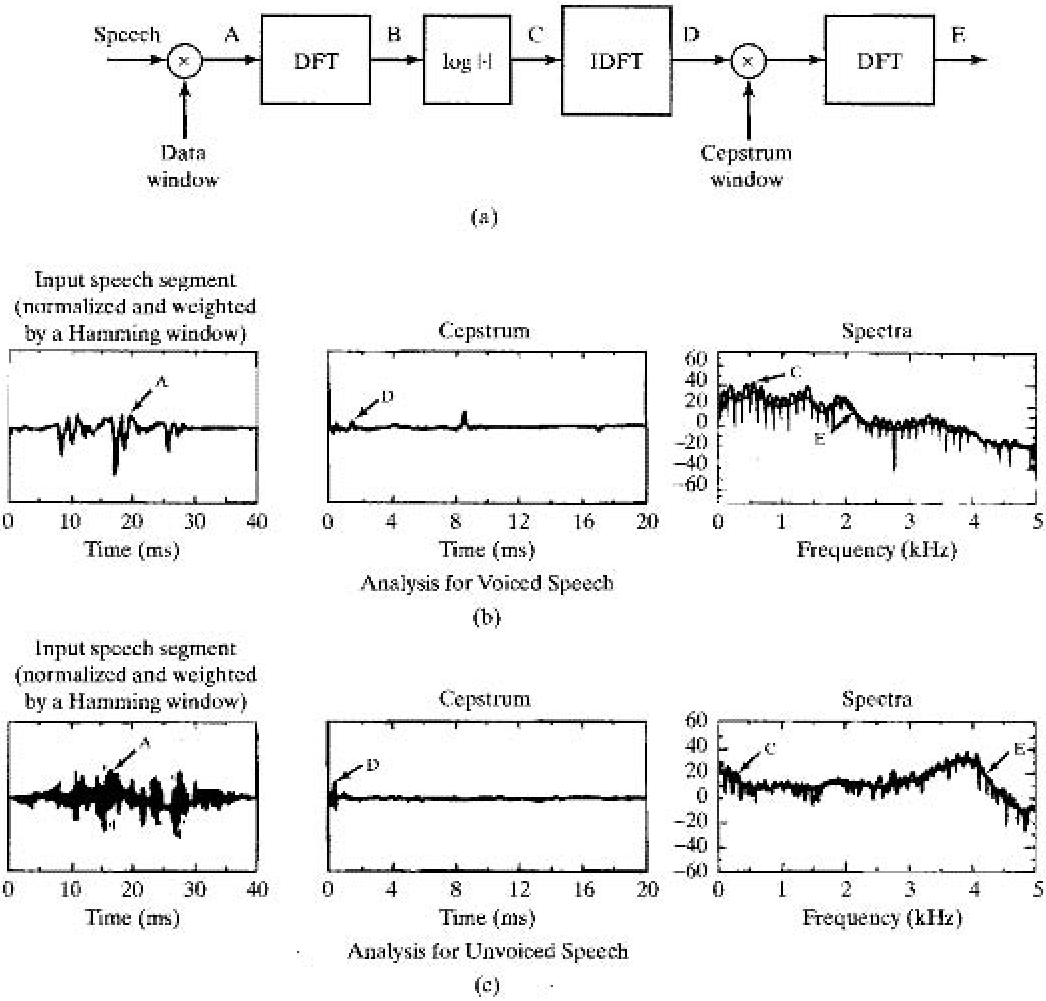


Figure 13.26 (a) System for cepstrum analysis of speech signals. (b) Analysis for voice speech. (c) Analysis for unvoiced speech.

complex cepstrum, our previous discussion suggests that the low-time portion of $c_x[n]$ should correspond to the slowly varying components of the log magnitude of the Fourier transform of the speech segment, and for voiced speech, the cepstrum should contain impulses at multiples of the pitch period. An example is shown in Figure 13.26.

Figure 13.26(a) shows the operations involved in estimating the speech parameters using the cepstrum. Figure 13.26(b) shows a typical result for voiced speech. The windowed speech signal is labeled A, $\log|X[k]|$ is labeled C, and the cepstrum $c_x[n]$ is labeled D. The peak in the cepstrum at about 8 ms indicates that this segment of speech is voiced with that period. The smoothed spectrum, or *spectrum envelope*, obtained by frequency-invariant lowpass filtering with cutoff below 8 ms is labeled E and is superimposed on C. The situation for unvoiced speech, shown in Figure 13.26(c), is similar, except that the random nature of the excitation component of the input speech segment causes a rapidly varying random component in $\log|X[k]|$ instead of a periodic component. Thus, in the cepstrum the low-time components correspond as before to the vocal tract system function; however, since the rapid variations in $\log|X[k]|$ are not periodic, no strong peak appears in the cepstrum. Therefore, the presence or absence of a peak in the cepstrum in the normal pitch period range serves as a very good voiced/unvoiced detector and pitch period estimator. The result of lowpass frequency-invariant filtering in the unvoiced case is similar to that in the voiced case. A smoothed spectrum envelope estimate is obtained as in E.

In speech analysis applications, the operations of Figure 13.26(a) are applied repeatedly to sequential segments of the speech waveform. The length of the segments must be carefully selected. If the segments are too long, the properties of the speech signal will change too much across the segment. If the segments are too short, there will not be enough of the signal to obtain a strong indication of periodicity. Usually the segment length is set at about three to four times the average pitch period of the speech signal. Figure 13.27 shows an example of how the cepstrum can be used for pitch detection and for estimation of the vocal tract resonance frequencies. Figure 13.27(a) shows a sequence of cepstra computed for speech waveform segments selected at 20-ms intervals. The existence of a prominent peak throughout the sequence of speech segments indicates that the speech was voiced throughout. The location of the cepstrum peak indicates the value of the pitch period in each corresponding time interval. Figure 13.27(b) shows the log magnitude with the corresponding smoothed spectra superimposed. The lines connect estimates of the vocal tract resonances obtained by a heuristic peak-picking algorithm. (See Schafer and Rabiner, 1970.)

13.10.4 Applications

As indicated previously, cepstrum analysis methods have found widespread application in speech processing problems. One of the most successful applications is in pitch detection (Noll, 1967). They also have been used successfully in speech analysis/synthesis systems for low bit-rate coding of the speech signal (Oppenheim, 1969b; Schafer and Rabiner, 1970).

Cepstrum representations of speech have also been used with considerable success in pattern recognition problems associated with speech processing such as speaker identification (Atal, 1976), speaker verification (Furui, 1981) and speech recognition (Davis and Mermelstein, 1980). Although the technique of linear predictive analysis of Chapter 11 is the most widely used method of obtaining a representation of the vocal tract component of the speech model, the linear predictive model representation is often transformed to a cepstrum representation for use in pattern recognition problems (Schroeder, 1981; Juang, Rabiner and Wilpon 1987). This transformation is explored in Problem 13.30.

13.11 SUMMARY

In this chapter, we discussed the technique of cepstrum analysis and homomorphic deconvolution. We focused primarily on definitions and properties of the complex cepstrum and on the practical problems in the computation of the complex cepstrum. An

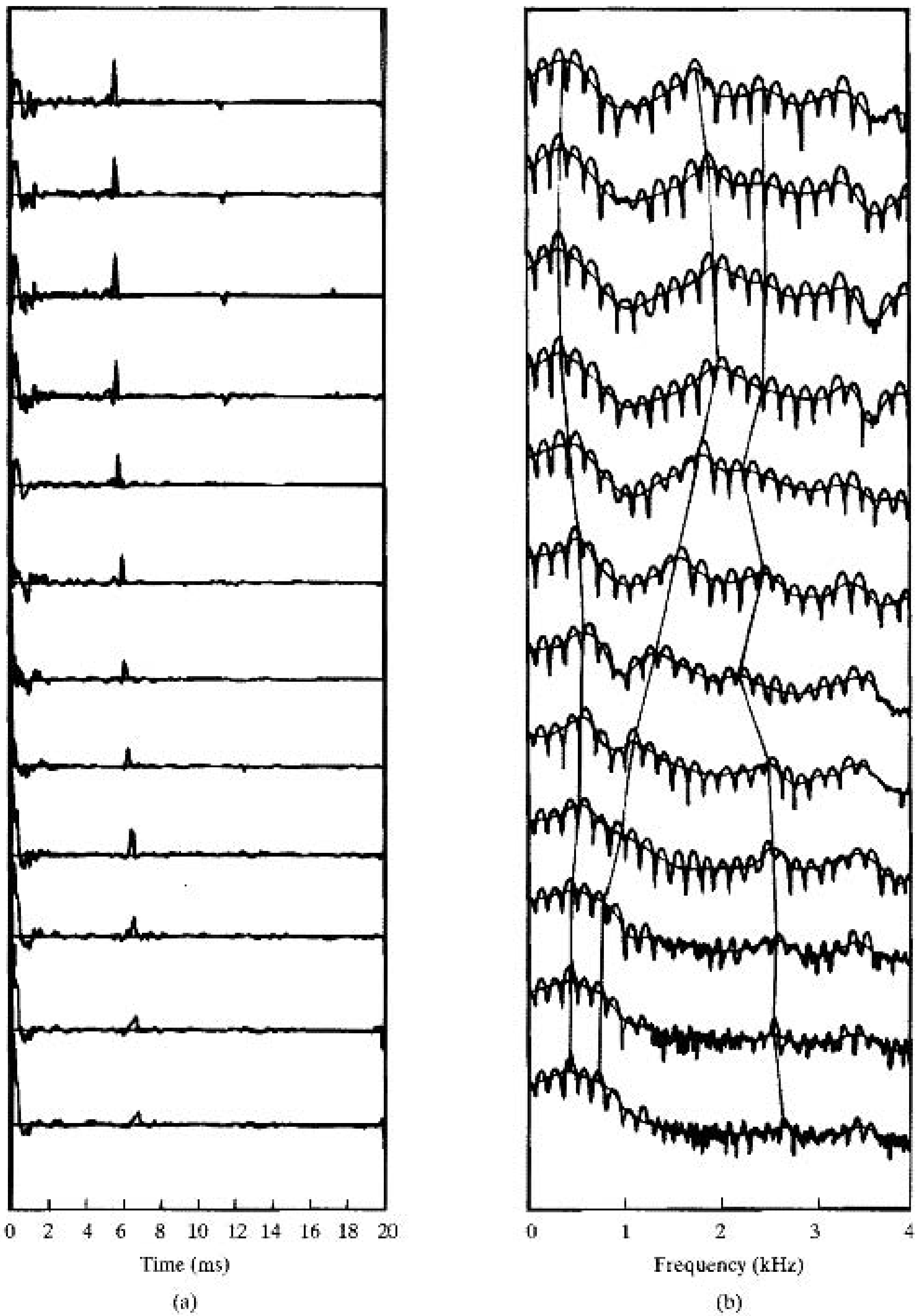


Figure 13.27 (a) Cepstra and (b) log spectra for sequential segments of voiced speech.

idealized example was discussed to illustrate the use of cepstrum analysis and homomorphic deconvolution for separating components of a convolution. The application of cepstrum analysis techniques to speech processing problems was discussed in some detail as an illustration of their use in a real application.

Problems

Basic Problems

13.1. **(a)** Consider a discrete-time system that is linear in the conventional sense. If $y[n] = T\{x[n]\}$ is the output when the input is $x[n]$, then the *zero signal* $\mathbf{0}[n]$ is the signal that can be added to $x[n]$ such that $T\{x[n] + \mathbf{0}[n]\} = y[n] + T\{\mathbf{0}[n]\} = y[n]$. What is the zero signal for conventional linear systems?

(b) Consider a discrete-time system $y[n] = T\{x[n]\}$ that is homomorphic, with convolution as the operation for combining signals at both the input and the output. What is the zero signal for such a system; i.e., what is the signal $\mathbf{0}[n]$ such that $T\{x[n] * \mathbf{0}[n]\} = y[n] * T\{\mathbf{0}[n]\} = y[n]$?

(c) Consider a discrete-time system $y[n] = T\{x[n]\}$ that is homomorphic, with convolution as the operation for combining signals at both the input and the output. What is the zero signal for such a system; i.e., what is the signal $\mathbf{0}[n]$ such that $T\{x[n] * \mathbf{0}[n]\} = y[n] * T\{\mathbf{0}[n]\} = y[n]$?

13.2. Let $x_1[n]$ and $x_2[n]$ denote two sequences and $\hat{x}_1[n]$ and $\hat{x}_2[n]$ their corresponding complex cepstra. If $x_1[n] * x_2[n] = \delta[n]$, determine the relationship between $\hat{x}_1[n]$ and $\hat{x}_2[n]$.

13.3. In considering the implementation of homomorphic systems for convolution, we restricted our attention to input signals with rational z-transforms of the form of Eq. (13.32). If an input sequence $x[n]$ has a rational z-transform but has either a negative gain constant or an amount of delay not represented by Eq. (13.32), then we can obtain a z-transform of the form of Eq. (13.32) by shifting $x[n]$ appropriately and multiplying by -1. The complex cepstrum may then be computed using Eq. (13.33).

Suppose that $x[n] = \delta[n] - 2\delta[n-1]$, and define $y[n] = \alpha x[n-r]$, where $\alpha = \pm 1$ and r is an integer. Find α and r such that $Y(z)$ is in the form of Eq. (13.32), and then find $\hat{y}[n]$.

13.4. In Section 13.5.1, we stated that linear-phase contributions should be removed from the unwrapped phase curve before computation of the complex cepstrum. This problem is concerned with the effect of not removing the linear-phase component due to the factor z^r in Eq. (13.29).

Specifically, assume that the input to the characteristic system for convolution is $x[n] = \delta[n+r]$. Show that formal application of the Fourier transform definition

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log[X(e^{j\omega})]e^{j\omega n}\,d\omega \tag{P13.4-1}$$

leads to

$$\hat{x}[n] = \begin{cases} r\dfrac{\cos(\pi n)}{n}, & n \neq 0, \\ 0, & n = 0. \end{cases}$$

The advantage of removing the linear-phase component of the phase is clear from this result, since for large r such a component would dominate the complex cepstrum.

13.5. Suppose that the z-transform of $s[n]$ is

$$S(z) = \frac{(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{4}z)}{(1 - \frac{1}{3}z^{-1})(1 - \frac{1}{5}z)}.$$

Determine the pole locations of the z-transform of $n\hat{s}[n]$, other than poles at $|z| = 0$ or ∞.

13.6. Suppose that the complex cepstrum of $y[n]$ is $\hat{y}[n] = \hat{s}[n] + 2\delta[n]$. Determine $y[n]$ in terms of $s[n]$.

13.7. Determine the complex cepstrum of $x[n] = 2\delta[n] - 2\delta[n-1] + 0.5\delta[n-2]$, shifting $x[n]$ or changing its sign, if necessary.

13.8. Suppose that the z-transform of a stable sequence $x[n]$ is given by

$$X(z) = \frac{1 - \frac{1}{2}z^{-1}}{1 + \frac{1}{2}z},$$

and that a stable sequence $y[n]$ has complex cepstrum $\hat{y}[n] = \hat{x}[-n]$, where $\hat{x}[n]$ is the complex cepstrum of $x[n]$. Determine $y[n]$.

13.9. Equations (13.65) and (13.68) are recursive relationships that can be used to compute the complex cepstrum $\hat{x}[n]$ when the input sequence $x[n]$ is minimum phase and maximum phase, respectively.

(a) Use Eq. (13.65) to compute recursively the complex cepstrum of the sequence $x[n] = a^n u[n]$, where $|a| < 1$.

(b) Use Eq. (13.68) to compute recursively the complex cepstrum of the sequence $x[n] = \delta[n] - a\delta[n+1]$, where $|a| < 1$.

13.10. $\text{ARG}\{X(e^{j\omega})\}$ represents the principal value of the phase of $X(e^{j\omega})$, and $\arg\{X(e^{j\omega})\}$ represents the continuous phase of $X(e^{j\omega})$. Suppose that $\text{ARG}\{X(e^{j\omega})\}$ has been sampled at frequencies $\omega_k = 2\pi k/N$ to obtain $\text{ARG}\{X[k]\} = \text{ARG}\{X(e^{j(2\pi/N)k})\}$ as shown in Figure P13.10. Assuming that $|\arg\{X[k]\} - \arg\{X[k-1]\}| < \pi$ for all k, determine and plot the sequence $r[k]$ as in Eq. (13.49) and $\arg\{X[k]\}$ for $0 \le k \le 10$.

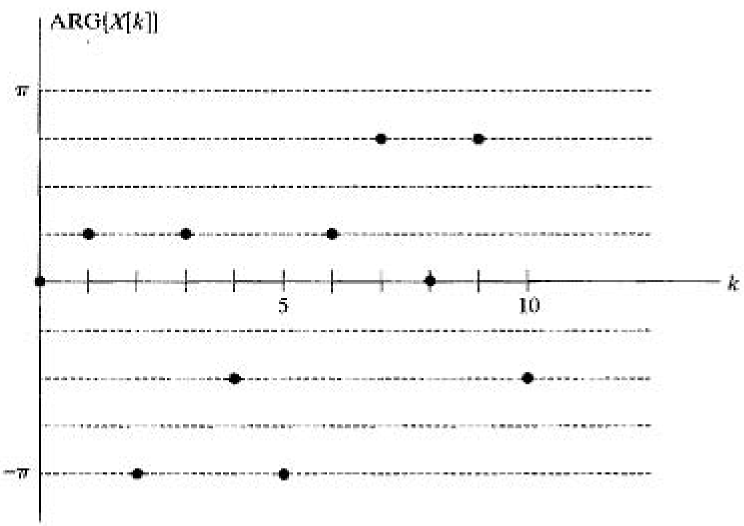


Figure P13.10

13.11. Let $\hat{x}[n]$ be the complex cepstrum of a real-valued sequence $x[n]$. Specify whether each of the following statements is true or false. Give brief justifications for your answers.

Statement 1: If $x_1[n] = x[-n]$ then $\hat{x}_1[n] = \hat{x}[-n]$.

Statement 2: Since $x[n]$ is real-valued, the complex cepstrum $\hat{x}[n]$ must also be real-valued.

Advanced Problems

13.12. Consider the system depicted in Figure P13.12, where S_1 is an LTI system with impulse response $h_1[n]$ and S_2 is a homomorphic system with convolution as the input and output operations; i.e., the transformation $T_2\{\cdot\}$ satisfies

$$T_2\{w_1[n] * w_2[n]\} = T_2\{w_1[n]\} * T_2\{w_2[n]\}.$$

Suppose that the complex cepstrum of the input $x[n]$ is $\hat{x}[n] = \delta[n] + \delta[n-1]$. Find a closed-form expression for $h_1[n]$ such that the output is $y[n] = \delta[n]$.

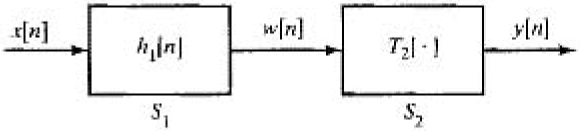


Figure P13.12

13.13. The complex cepstrum of a *finite length* signal $x[n]$ is computed as shown in Figure P13.13-1. Suppose we know that $x[n]$ is minimum phase (all poles and zeros are inside the unit circle). We use the system shown in Figure P13.13-2 to find the real cepstrum of $x[n]$. Explain how to construct $\hat{x}[n]$ from $c_x[n]$.

Figure P13.13-1

Figure P13.13-2

13.14. Consider the class of sequences that are real and stable and whose z-transforms are of the form

$$X(z) = |A| \frac{\prod_{k=1}^{M_i}(1 - a_k z^{-1}) \prod_{k=1}^{M_o}(1 - b_k z)}{\prod_{k=1}^{N_i}(1 - c_k z^{-1}) \prod_{k=1}^{N_o}(1 - d_k z)},$$

where $|a_k|, |b_k|, |c_k|, |d_k| < 1$. Let $\hat{x}[n]$ denote the complex cepstrum of $x[n]$.

(a) Let $y[n] = x[-n]$. Determine $\hat{y}[n]$ in terms of $\hat{x}[n]$.

(b) If $x[n]$ is causal, is it also minimum phase? Explain.

(c) Suppose that $x[n]$ is a finite-duration sequence such that

$$X(z) = |A| \prod_{k=1}^{M_i} (1 - a_k z^{-1}) \prod_{k=1}^{M_o} (1 - b_k z),$$

with $|a_k| < 1$ and $|b_k| < 1$. The function $X(z)$ has zeros inside and outside the unit circle. Suppose that we wish to determine $y[n]$ such that $|Y(e^{j\omega})| = |X(e^{j\omega})|$ and $Y(z)$ has no zeros outside the unit circle. One approach that achieves this objective is depicted in Figure P13.14. Determine the required sequence $\ell[n]$. A possible application of the system in Figure P13.14 is to stabilize an unstable system by applying the transformation of Figure P13.14 to the sequence of coefficients of the denominator of the system function.

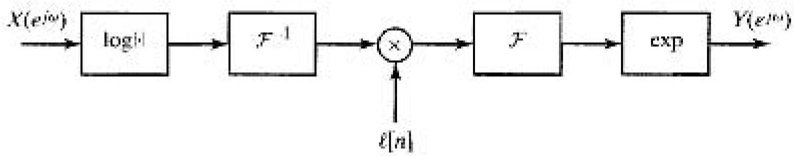


Figure P13.14

13.15. It can be shown (see Problem 3.50) that if $x[n] = 0$ for $n < 0$, then

$$x[0] = \lim_{z \to \infty} X(z).$$

This result was called the *initial value theorem for right-sided sequences*.

(a) Prove a similar result for *left-sided sequences*, i.e., for sequences such that $x[n] = 0$ for $n > 0$.

(b) Use the initial value theorems to prove that $\hat{x}[0] = \log(x[0])$ if $x[n]$ is a minimum-phase sequence.

(c) Use the initial value theorems to prove that $\hat{x}[0] = \log(x[0])$ if $x[n]$ is a maximum-phase sequence.

(d) Use the initial value theorems to prove that $\hat{x}[0] = \log|A|$ when $X(z)$ is given by Eq. (13.32). Is this result consistent with the results of parts (b) and (c)?

13.16. Consider a sequence $x[n]$ with complex cepstrum $\hat{x}[n]$, such that $\hat{x}[n] = -\hat{x}[-n]$. Determine the quantity

$$E = \sum_{n=-\infty}^{\infty} x^2[n].$$

13.17. Consider a real, stable, even, two-sided sequence $h[n]$. The Fourier transform of $h[n]$ is positive for all ω, i.e.,

$$H(e^{j\omega}) > 0, \qquad -\pi < \omega \le \pi.$$

Assume that the z-transform of $h[n]$ exists. Do not assume that $H(z)$ is rational.

(a) Show that there exists a minimum-phase signal $g[n]$, such that

$$H(z) = G(z)G(z^{-1}),$$

where $G(z)$ is the z-transform of a sequence $g[n]$, which has the property that $g[n] = 0$ for $n < 0$. State explicitly the relationship between $\hat{h}[n]$ and $\hat{g}[n]$, the complex cepstra of $h[n]$ and $g[n]$, respectively.

(b) Given a stable signal $s[n]$, with rational z-transform

$$S(z) = \frac{(1-2z^{-1})(1-\frac{1}{2}z^{-1})}{(1-4z^{-1})(1-\frac{1}{3}z^{-1})}.$$

Define $h[n] = s[n] * s[-n]$. Find $G(z)$ (as in part (a)) in terms of $S(z)$.

(c) Consider the system in Figure P13.17, where $\ell[n]$ is defined as

$$\ell[n] = u[n-1] + (-1)^n u[n-1].$$

Determine the *most general* conditions on $x[n]$ such that $y[n] = x[n]$ for all n.

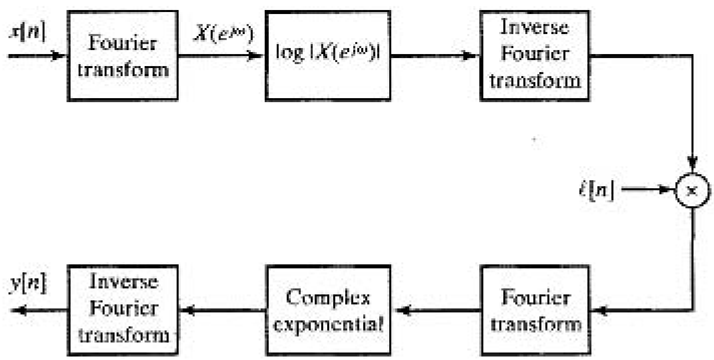


Figure P13.17

13.18. Consider a maximum-phase signal $x[n]$.

(a) Show that the complex cepstrum $\hat{x}[n]$ of a maximum-phase signal is related to its cepstrum $c_x[n]$ by

$$\hat{x}[n] = c_x[n]\ell_{max}[n],$$

where $\ell_{max}[n] = 2u[-n] - \delta[n]$.

(b) Using the relationships in part (a), show that

$$\arg\{X(e^{j\omega})\} = \frac{1}{2\pi}\mathcal{P}\int_{-\pi}^{\pi} \log|X(e^{j\theta})| \cot\left(\frac{\omega-\theta}{2}\right)d\theta.$$

(c) Also show that

$$\log|X(e^{j\omega})| = \hat{x}[0] - \frac{1}{2\pi}\mathcal{P}\int_{-\pi}^{\pi} \arg[X(e^{j\theta})] \cot\left(\frac{\omega-\theta}{2}\right)d\theta.$$

13.19. Consider a sequence $x[n]$ with Fourier transform $X(e^{j\omega})$ and complex cepstrum $\hat{x}[n]$. A new signal $y[n]$ is obtained by homomorphic filtering where

$$\hat{y}[n] = (\hat{x}[n] - \hat{x}[-n])u[n-1].$$

(a) Show that $y[n]$ is a minimum-phase sequence.

(b) What is the phase of $Y(e^{j\omega})$?

(c) Obtain a relationship between $\arg[Y(e^{j\omega})]$ and $\log|Y(e^{j\omega})|$.

(d) If $x[n]$ is minimum phase, how is $y[n]$ related to $x[n]$?

13.20. Equation (13.65) represents a recursive relationship between a sequence $x[n]$ and its complex cepstrum $\hat{x}[n]$. Show from Eq. (13.65) that the characteristic system $D_*[\cdot]$ behaves as a causal system for minimum-phase inputs; i.e., show that for minimum-phase inputs, $\hat{x}[n]$ is dependent only on $x[k]$ for $k \leq n$.

13.21. Describe a procedure for computing a causal sequence $x[n]$, for which

$$X(z) = -z^3 \frac{(1 - 0.95z^{-1})^{2/5}}{(1 - 0.9z^{-1})^{7/13}}.$$

13.22. The sequence

$$h[n] = \delta[n] + \alpha\delta[n - n_0]$$

is a simplified model for the impulse response of a system that introduces an echo.

(a) Determine the complex cepstrum $\hat{h}[n]$ for this sequence. Sketch the result.

(b) Determine and sketch the cepstrum $c_h[n]$.

(c) Suppose that an approximation to the complex cepstrum is computed using N-point DFTs as in Eqs. (13.46a) to (13.46c). Obtain a closed-form expression for the approximation $\hat{h}_p[n]$, $0 \leq n \leq N - 1$, for the case $n_0 = N/6$. Assume that phase unwrapping can be accurately done. What happens if N is not divisible by n_0?

(d) Repeat part (c) for the cepstrum approximation $c_{xp}[n]$, $0 \leq n \leq N - 1$, as computed using Eqs. (13.60a) and (13.60b).

(e) If the largest impulse in the cepstrum approximation $c_{xp}[n]$ is to be used to detect the value of the echo delay n_0, how large must N be to avoid ambiguity? Assume that accurate phase unwrapping can be achieved with this value of N.

13.23. Let $x[n]$ be a *finite-length* minimum-phase sequence with complex cepstrum $\hat{x}[n]$, and define $y[n]$ as

$$y[n] = \alpha^n x[n]$$

with complex cepstrum $\hat{y}[n]$.

(a) If $0 < \alpha < 1$, how is $\hat{y}[n]$ related to $\hat{x}[n]$?

(b) How should α be chosen so that $y[n]$ is no longer minimum phase?

(c) How should α be chosen so that if linear-phase terms are removed before computing the complex cepstrum, then $\hat{y}[n] = 0$ for $n > 0$?

13.24. Consider a minimum-phase sequence $x[n]$ with z-transform $X(z)$ and complex cepstrum $\hat{x}[n]$. A new complex cepstrum is defined by the relation

$$\hat{y}[n] = (\alpha^n - 1)\hat{x}[n].$$

Determine the z-transform $Y(z)$. Is the result also minimum phase?

13.25. Section 13.9.4 contains an example of how the complex cepstrum can be used to obtain two different decompositions involving convolution of a minimum-phase sequence with another sequence. In that example,

$$X(z) = \frac{(0.98 + z^{-1})(1 + 0.9z^{-15} + 0.81z^{-30})}{(1 - 0.9e^{j\pi/6}z^{-1})(1 - 0.9e^{-j\pi/6}z^{-1})}.$$

(a) In one decomposition, $X(z) = X_{min}(z)X_{ap}(z)$ where

$$X_{min}(z) = \frac{(1 + 0.98z^{-1})(1 + 0.9z^{-15} + 0.81z^{-30})}{(1 - 0.9e^{j\pi/6}z^{-1})(1 - 0.9e^{-j\pi/6}z^{-1})}$$

and

$$X_{ap}(z) = \frac{(0.98 + z^{-1})}{(1 + 0.98z^{-1})}.$$

Use the power series expansion of the logarithmic terms to find the complex cepstra $\hat{x}_{min}[n]$, $\hat{x}_{ap}[n]$, and $\hat{x}[n]$. Plot these sequences and compare your plots with those in Figure 13.19.

(b) In the second decomposition, $X(z) = X_{mn}(z)X_{mx}(z)$ where

$$X_{mn}(z) = \frac{z^{-1}(1 + 0.9z^{-15} + 0.81z^{-30})}{(1 - 0.9e^{j\pi/6}z^{-1})(1 - 0.9e^{-j\pi/6}z^{-1})}$$

and

$$X_{mx}(z) = (0.98z + 1).$$

Use the power series expansion of the logarithmic terms to find the complex cepstra and show that $\hat{x}_{mn}[n] \neq \hat{x}_{min}[n]$ but that $\hat{x}[n] = \hat{x}_{mn}[n] + \hat{x}_{mx}[n]$ is the same as in part (a). Note that

$$(1 + 0.9z^{-15} + 0.81z^{-30}) = \frac{(1 - (0.9)^3 z^{-45})}{(1 - 0.9z^{-15})}.$$

13.26. Suppose that $s[n] = h[n] * g[n] * p[n]$, where $h[n]$ is a minimum-phase sequence, $g[n]$ is a maximum-phase sequence, and $p[n]$ is

$$p[n] = \sum_{k=0}^{4} \alpha_k \delta[n - kn_0],$$

where α_k and n_0 are not known. Develop a method to separate $h[n]$ from $s[n]$.

Extension Problems

13.27. Let $x[n]$ be a sequence with z-transform $X(z)$ and complex cepstrum $\hat{x}[n]$. The magnitude-squared function for $X(z)$ is

$$V(z) = X(z)X^*(1/z^*).$$

Since $V(e^{j\omega}) = |X(e^{j\omega})|^2 \geq 0$, the complex cepstrum $\hat{v}[n]$ corresponding to $V(z)$ can be computed without phase unwrapping.

(a) Obtain a relationship between the complex cepstrum $\hat{v}[n]$ and the complex cepstrum $\hat{x}[n]$.

(b) Express the complex cepstrum $\hat{v}[n]$ in terms of the cepstrum $c_x[n]$.

(c) Determine the sequence $\ell[n]$ such that

$$\hat{x}_{min}[n] = \ell[n]\hat{v}[n]$$

is the complex cepstrum of a minimum-phase sequence $x_{min}[n]$ for which

$$|X_{min}(e^{j\omega})|^2 = V(e^{j\omega}).$$

(d) Suppose that $X(z)$ is as given by Eq. (13.32). Use the result of part (c) and Eqs. (13.36a), (13.36b), and (13.36c) to find the complex cepstrum of the minimum-phase sequence, and work backward to find $X_{min}(z)$.

The technique employed in part (d) may be used in general to obtain a minimum-phase factorization of a magnitude-squared function.

13.28. Let $\hat{x}[n]$ be the complex cepstrum of $x[n]$. Define a sequence $x_e[n]$ to be

$$x_e[n] = \begin{cases} x[n/N], & n = 0, \pm N, \pm 2N, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

Show that the complex cepstrum of $x_e[n]$ is given by

$$\hat{x}_e[n] = \begin{cases} \hat{x}[n/N], & n = 0, \pm N, \pm 2N, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

13.29. In speech analysis, synthesis, and coding, the speech signal is commonly modeled over a short time interval as the response of an LTI system excited by an excitation that switches between a train of equally spaced pulses for voiced sounds and a wideband random noise source for unvoiced sounds. To use homomorphic deconvolution to separate the components of the speech model, the speech signal $s[n] = v[n] * p[n]$ is multiplied by a window sequence $w[n]$ to obtain $x[n] = s[n]w[n]$. To simplify the analysis, $x[n]$ is approximated by

$$x[n] = (v[n] * p[n]) \cdot w[n] \simeq v[n] * (p[n] \cdot w[n]) = v[n] * p_w[n]$$

where $p_w[n] = p[n]w[n]$ as in Eq. (13.123).

(a) Give an example of $p[n]$, $v[n]$, and $w[n]$ for which the above assumption may be a poor approximation.

(b) One approach to estimating the excitation parameters (voiced/unvoiced decision and pulse spacing for voiced speech) is to compute the real cepstrum $c_x[n]$ of the windowed segment of speech $x[n]$ as depicted in Figure P13.29-1. For the model of Section 13.10.1, express $c_x[n]$ in terms of the complex cepstrum $\hat{x}[n]$. How would you use $c_x[n]$ to estimate the excitation parameters?

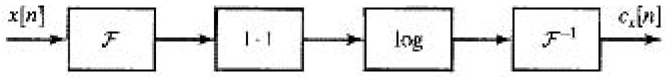


Figure P13.29-1

(c) Suppose that we replace the log operation in Figure P13.29-1 with the "squaring" operation so that the resulting system is as depicted in Figure P13.29-2. Can the new "cepstrum" $q_x[n]$ be used to estimate the excitation parameters? Explain.

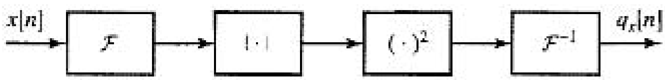


Figure P13.29-2

13.30. Consider a stable LTI system with impulse response $h[n]$ and all-pole system function

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

Such all-pole systems arise in linear-predictive analysis. It is of interest to compute the complex cepstrum directly from the coefficients of $H(z)$.

(a) Determine $\hat{h}[0]$.

(b) Show that

$$\hat{h}[n] = a_n + \sum_{k=1}^{n-1} \left(\frac{k}{n}\right) \hat{h}[k] a_{n-k}, \qquad n \geq 1.$$

With the relations in parts (a) and (b), the complex cepstrum can be computed without phase unwrapping and without solving for the roots of the denominator of $H(z)$.

13.31. A somewhat more general model for echo than the system in Problem 13.22 is the system depicted in Figure P13.31. The impulse response of this system is

$$h[n] = \delta[n] + \alpha g[n - n_0],$$

where $\alpha g[n]$ is the impulse response of the echo path.

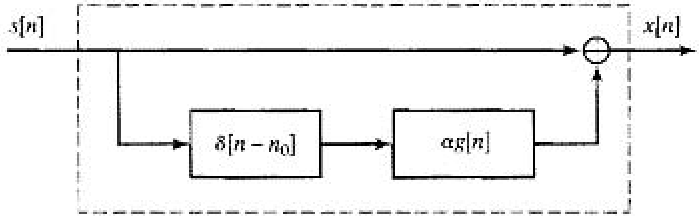


Figure P13.31

(a) Assuming that

$$\max_{-\pi<\omega<\pi} |\alpha G(e^{j\omega})| < 1,$$

show that the complex cepstrum $\hat{h}[n]$ has the form

$$\hat{h}[n] = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\alpha^k}{k} g_k[n - kn_0],$$

and determine an expression for $g_k[n]$ in terms of $g[n]$.

(b) For the conditions of part (a), determine and sketch the complex cepstrum $\hat{h}[n]$ when $g[n] = \delta[n]$.

(c) For the conditions of part (a), determine and sketch the complex cepstrum $\hat{h}[n]$ when $g[n] = a^n u[n]$. What condition must be satisfied by α and a so that the result of part (a) applies?

(d) For the conditions of part (a), determine and sketch the complex cepstrum $\hat{h}[n]$ when $g[n] = a_0\delta[n] + a_1\delta[n-n_1]$. What condition must be satisfied by α, a_0, a_1, and n_1 so that the result of part (a) applies?

13.32. An interesting use of exponential weighting is in computing the complex cepstrum without phase unwrapping. Assume that $X(z)$ has no poles and zeros on the unit circle. Then it is possible to find an exponential weighting factor α in the product $w[n] = \alpha^n x[n]$, such that none of the poles or zeros of $X(z)$ are shifted across the unit circle in forming $W(z) = X(\alpha^{-1}z)$.

(a) Assuming that no poles or zeros of $X(z)$ move across the unit circle, show that

$$\hat{w}[n] = \alpha^n \hat{x}[n]. \qquad \text{(P13.32-1)}$$

(b) Now suppose that instead of the complex cepstrum, we compute $c_x[n]$ and $c_w[n]$. Use the result of part (a) to obtain expressions for both $c_x[n]$ and $c_w[n]$ in terms of $\hat{x}[n]$.

(c) Now show that

$$\hat{x}[n] = \frac{2(c_x[n] - \alpha^n c_w[n])}{1 - \alpha^{2n}}, \qquad n \neq 0. \qquad \text{(P13.32-2)}$$

(d) Since $c_x[n]$ and $c_w[n]$ can be computed from $\log|X(e^{j\omega})|$ and $\log|W(e^{j\omega})|$, respectively, Eq. (P13.32-2) is the basis for computing the complex cepstrum without computing the phase of $X(e^{j\omega})$. Discuss some potential problems that might arise with this approach.