The Cepstrum: A Guide to Processing

DONALD G. CHILDERS, FELLOW, IEEE, DAVID P. SKINNER, MEMBER, IEEE,
AND ROBERT C. KEMERAIT, MEMBER, IEEE

Abstract—This paper is a pragmatic tutorial review of the cepstrum literature focusing on data processing. The power, complex, and phase cepstra are shown to be easily related to one another. Problems associated with phase unwrapping, linear phase components, spectrum notching, aliasing, oversampling, and extending the data sequence with zeros are discussed. The advantages and disadvantages of windowing the sampled data sequence, the log spectrum, and the complex cepstrum are presented. The influence of noise upon the data processing procedures is discussed throughout the paper, but is not thoroughly analyzed. The effects of various forms of liftering the cepstrum are described. The results obtained by applying whitening and trend removal techniques to the spectrum prior to the calculation of the cepstrum are discussed.

We have attempted to synthesize the results, procedures, and information peculiar to the many fields that are finding cepstrum analysis useful. In particular we discuss the interpretation and processing of data in such areas as speech, seismology, and hydroacoustics. But we must caution the reader that the paper is heavily influenced by our own experiences; specific procedures that have been found useful in one field should not be considered as totally general to other fields.

It is hoped that this review will be of value to those familiar with the field and reduce the time required for those wishing to become so.

I. INTRODUCTION

THIS PAPER has two objectives: first, to present a guide to the cepstrum literature which is becoming increasing diverse; and second, to survey cepstrum signal processing procedures which have found application in the analysis of data. While a review section is included to establish a notational reference, no attempt is made to provide an extensive rederivation of previous results. Instead, key sources are referenced.

We recognize that for some readers our summary of various cepstrum processing procedures which have been and continue to be applied to the analysis of data may be too concise. It is our hope that this weakness is counterbalanced by our discussion of the pitfalls, advantages, and disadvantages of such procedures. While the majority of these results appear elsewhere, sometimes mentioned only briefly, we do include some new findings. A good deal of what follows is based upon our own experiences but we have tried to incorporate and synthesize the results and experiences of the numerous other investigators in this field as well. Perhaps the single most important observation that has been made by us and others is that the results obtained from cepstrum analysis are highly *data dependent*. Because of this, few generalities, which cut across application areas, can be formulated from empirical results from one application area alone. Nonetheless, investigators appear to be obtaining satisfactory results from the processing of their data by cepstrum techniques. Thus it is our hope that the reader will gain from this paper some new insights, ideas, and

perspectives; but, hopefully, these will be tempered by a healthy skepticism.

Fundamentally, cepstrum techniques are suited for the analysis of data that contain echoes (wavelets) or reverberations of a fundamental wavelet (sometimes called a signature) whose shape need not be known *a priori*. The power cepstrum is usually used to determine the arrival times of the fundamental wavelet and its echoes and their relative amplitudes; the processing of the complex cepstrum can determine the wavelet waveform.

The application areas encompass radar and sonar [6]-[8], [29], [40]-[41], [43], [61], [62] where cepstrum processing can be used to advantage to reduce reflection interference, speech [13], [19], [21], [42], [46]-[51], [53], [55], [57], [61]-[66], [71], where speaker fundamental frequency (pitch) is estimated and spectrum envelopes are calculated, marine and earth seismology, seismic exploration and detection [1], [31], [34]-[38], [45], [58]-[59], [80] where source depth determinations are made and the ocean bottom is mapped, and the electroencephalogram (EEG) or brain waves [4], [9], [11], [12], [22]-[24], [40], [54], [64] where correlates of electrophysiological events are derived. Other areas of interest include the deconvolution of probability density functions [52]. More recent applications are two-dimensional functions [5], [20] and aeroacoustics or noise pollution [44]. Some additional exciting work is the restoration of old recordings [30] and image processing [19], [30], [61], [63]. Throughout the paper we attempt to synthesize the observations from these many fields.

The paper's outline is as follows. First, we provide a brief historical review and the fundamental mathematical formulation for processing discrete data sequences by cepstrum techniques. This is followed by a section on phase problems encountered in the complex cepstrum, e.g., linear phase terms, spectral notching, oversampling, phase unwrapping, and noise. Next comes a section on other problems, e.g., aliasing, oversampling, the addition of zeros. The following section is on windowing the original data sequence, the log spectrum sequence, and the cepstrum sequence. In the next to the last section we discuss data processing for the three areas of speech, seismology, and hydroacoustics. Finally, we conclude with a summary, recapitulation, and some general observations. The paper is organized in this manner in an attempt to localize, as well as possible, the terminology that is special to each of the diverse fields. Thus, the next to the last section on data processing employs many terms peculiar to each field.

It is our hope that this guide will prove useful to those just becoming involved in cepstrum analysis as well as to those with previous experience.

II. THE CEPSTRA

Historically, the cepstrum has its roots in the general problem of the deconvolution of two or more signals. This literature is rich and varied and encompasses linear prediction,

predictive deconvolution, inverse filtering, and deconvolution. In the interest of brevity we have provided only a selected list of references and books in this area [66]–[86].

In what follows it will be seen that the (complex) cepstrum is also concerned with the deconvolution of two signals, namely, a basic or fundamental wavelet and a train of impulses. The phase cepstrum is defined and it is shown how the power, complex, and phase cepstra are related via a form which simplifies computation.

A. The Power Cepstrum

The power cepstrum was first described by Bogert *et al.* [1] in 1963 as a heuristic technique for finding echo arrival times in a composite signal. Basically, these authors defined the cepstrum (which we term the power cepstrum [9], [12] to avoid confusion with the complex cepstrum) of a function as the power spectrum of the logarithm of the power spectrum of that function.

These authors quickly showed (as we repeat below) that the effect of a delayed echo will manifest itself as a ripple in the log spectrum. The "frequency" of this ripple is easily determined by calculating the spectrum of the log spectrum wherein this "frequency" will appear as a peak. However, the units of "frequency" of this ripple in the log spectrum are in units of time; thus, the independent variable (abscissa) in the spectrum of the log spectrum is time. Other parameters were also observed to undergo similar transformations of units. To avoid confusion, Bogert *et al.* [1] introduced the following now-classical paraphrased terms according to a syllabic interchange rule:

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

along with others. Today the two most prevalent terms are cepstrum and quefrency, e.g., filtering in the cepstrum domain is usually called just that and not "liftering" as suggested by Bogert *et al.* [1], but this can and often does lead to confusion.

In practice the power cepstrum is effective if the wavelet and the impulse train, whose convolution comprise the composite data, occupy different quefrency ranges. In actuality, the power cepstrum does not exist for most signals; it is meaningful only when defined in a sampled data sense (as is the complex cepstrum) although attempts to extend it exist [3]. Thus the following definition is offered: the power cepstrum of a data sequence is the square of the inverse z-transform of the logarithm of the magnitude squared of the z-transform of the data sequence. When this definition is evaluated on the unit circle, the result (except for the normalization factors associated with the power spectrum) is the same as that obtained in [1]. Thus we may write the power cepstrum as

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

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

where $X(z)$ is the z-transform of the data sequence $x(nT)$. Alternately, the definition could be changed to use the for-

ward z-transform and/or the final squaring could be changed to magnitude squared. In actuality the final squaring operation in (1) is unnecessary and is frequently omitted for several reasons, but it has been used here to provide historical continuity with [1].

Therefore, if we have the convolution of two sequences, then

$$x(nT) = f(nT) * g(nT) \quad (2)$$

or

$$|X(z)|^2 = |F(z)|^2 \cdot |G(z)|^2 \quad (3)$$

or

$$\log |X(z)|^2 = \log |F(z)|^2 + \log |G(z)|^2. \quad (4)$$

If we apply (1), then

$$x_{pc}(nT) = f_{pc}(nT) + g_{pc}(nT) + \text{a cross-product term.} \quad (5a)$$

If the power cepstra of f and g occupy different quefrency ranges, then (5a) can be reduced to

$$x_{pc}(nT) = f_{pc}(nT) + g_{pc}(nT). \quad (5b)$$

(This is the result that would be obtained if the final squaring operation in (1) were not included in the definition.) Under this condition the individual contributions of each power cepstrum can be separated by liftering (filtering) in the quefrency domain.

For the case of a composite signal consisting of the basic wavelet and a single echo, then

$$g(nT) = \delta(nT) + a\delta(nT - n_0 T) \quad (6)$$

where $\delta(nT)$ denotes the unit *pulse* function in a sampled data sequence.

Equation (3) then becomes

$$|X(z)|^2 = |F(z)|^2 |(1 + az^{-n_0})|^2 \quad (7)$$

and if we evaluate (4) on the unit circle ($z = e^{j\omega T}$), then

$$\log |X(e^{j\omega T})|^2 = \log |F(e^{j\omega T})|^2$$
$$+ \log (1 + a^2 + 2a \cos (\omega n_0 T))$$
$$= \log |F(e^{j\omega T})|^2 + \log (1 + a^2)$$
$$+ \log \left(1 + \frac{2a}{1 + a^2} \cos (\omega n_0 T) \right). \quad (8)$$

We may now expand the third term on the right of (8) in a power series (except for the point values $a = \pm 1$ and $\cos (\omega n_0 T) = \pm 1$ [6]) to obtain

$$\log (1 + a_0 \cos (\omega n_0 T)) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} (a_0 \cos (\omega n_0 T))^m \quad (9)$$

where $a_0 = (2a)/(1 + a^2)$.

We see that the logarithm of the magnitude squared of the z-transform of the composite signal will contain cosinusoidal ripples (sometimes referred to as spectral modulation) whose gamnitude (amplitude) and quefrency (i.e., the frequency of the ripples) are related to the echo amplitude (a) and delay ($n_0 T$), respectively.

Using (9) in (8), we can take the inverse z-transform of (8) to obtain the term within the brackets in (1) which will have

peaks at quefrencies of $(n_0 T)$ and multiples thereof. (We perform this task in detail later for the complex cepstrum.) These peaks will be detectable provided the $\log |F(e^{j\omega T})|^2$ is approximately quefrency limited to less than $(n_0 T)$, i.e., the ripples in $\log |F(e^{j\omega T})|^2$ should not have a repoid (period) greater than $(n_0 T)^{-1}$. In other words (5b) holds or at least approximately so. It is apparent that the echo arrival time can be estimated by simply noting the time of occurrence of the first peak in the power cepstrum. Further, it is possible using (9) to estimate the echo amplitude. This will be discussed more fully for the complex cepstrum. The presence of multiple echoes can and does confuse the interpretation procedure because of the nonlinearity introduced by the log function. In addition, aliasing causes problems [32]. These points are given further consideration later. It should also be apparent that if (5b) does not hold, then the cross-product term in (5a) will introduce further extraneous peaks.

It should be noted that the peaks in the power cepstrum can be removed by notch liftering (filtering) to yield an estimate of the power cepstrum of the basic wavelet. Further, if the final squaring operation in the calculation of the power cepstrum in (1) is not performed, then the peaks still appear and again can be removed by notch liftering, but now the operations performed to calculate this modified power cepstrum can be reversed to obtain an estimate of the log power spectrum and with exponentiation to yield the power spectrum of the basic wavelet itself. But the waveform of the basic wavelet cannot be recovered by processing the power cepstrum since the phase information is discarded. This latter situation is corrected with the complex cepstrum which we discuss in the next subsection along with the inversion process.

The power cepstrum has been applied to seismic data [1], [36], sonar [43], speech [13], [46]–[49], and the electroencephalogram (EEG) [12], [22], [24]. Its statistical properties have also been examined [2], [7].

It is hopefully beneficial to point out that alternate viewpoints and, thus, subsequent terminologies have arisen since the original paper by Bogert *et al.* [1]. These viewpoints have led to what might well be considered two lines of investigation: (1) the use of varying degrees of spectral whitening; and (2) the attempts to devise methods for obtaining the phase relations of the wavelet with respect to the reference signal [36].

We have seen that the occurrence of an echo in the time domain signal leads to what amounts to a spectral modulation (or ripple) in the frequency domain.

The spectral whitening approach to echo detection considers the application of the logarithm a severe spectral whitener (rather than a mechanism to transform the product of two functions into the sum of the logarithm of the two functions as Bogert *et al.* intended). Cohen [36] and his co-workers consider the logarithm too severe a spectral whitener for some applications, e.g., when the signal is narrow bandwidth and the signal-to-noise ratio (SNR) is low, when multiple echoes are present, and when the echo amplitude is large, $|a| > 0.5$. They consider separating the echo and the original wavelet in three ways: 1) the unwhitened version, which calculates the power spectrum of the data where some form of mean removal or lifter in the quefrency domain is always applied to the spectrum before the second transformation; 2) the moderately whitened version, which calculates the power spectrum of the square root of the power spectrum; and 3) the severe whitened version, which calculates the power spectrum of the logarithm of the power spectrum.

Cohen [36], [37] discusses the echo polarity determination problem. In some experimental situations the echo may undergo 180° phase reversals at certain boundary reflection interfaces, i.e., the reflection coefficient, a, may be negative. This knowledge can be useful in data interpretation. For example, in (8) if a is negative this leads to spectral nulls at $f = (m)/(n_0 T)$, $m = 0, 1, 2, \cdots$. By measuring the frequency spacing between two successive spectral nulls one can then determine the delay time $n_0 T$. The source depth can in turn be estimated if the average signal velocity is known.

Cohen [36] also points out that the cepstrum may contain many peaks which can confuse the analyst. He suggests doing pseudo-autocorrelation (defined as the inverse Fourier transform of the liftered power spectrum [1], [36]) analysis simultaneously. If a negative reflection has occurred, then the cepstrum will have a positive peak while the pseudo-autocorrelation will show a negative peak. The dot product (i.e., keeping track of the phase) of the cepstrum and the pseudo-autocorrelation can even be determined. Here the cepstrum can be any one of those obtained by transforming one of the three whitened spectra discussed above. This procedure can apparently help separate multiple echoes (or events) from a single echo (or event) such as might occur with multipath, i.e., multiple reflections.

B. The Complex Cepstrum and Wavelet Recovery

The complex cepstrum is an outgrowth of homomorphic system theory developed by Oppenheim [14]–[17]. In fact, the power cepstrum is also a specific application of homomorphic system theory. The complex cepstrum has been investigated extensively [9], [12], [13], [19], [21]–[23], [26]–[32], [61]–[64].

Since the complex cepstrum retains the phase information of the composite data, it can be used not only for echo detection but also wavelet recovery; this process is also known as homomorphic deconvolution or homomorphic filtering and has since been applied to seismic data [31], [34]–[38], [45], [57]–[59], speech [19], [21], [30], [50], [51], [53], [57], [62], [63], [65], [66], image processing [19], [30], [61], [63], and EEG analysis [9], [11], [12], [22]–[24].

Formally, we define the complex cepstrum of a data sequence as the inverse z-transform of the complex logarithm of the z-transform of the data sequence [21], [63], i.e.,

$$\hat{x}(nT) = \frac{1}{2\pi j} \oint_c \log(X(z)) z^{n-1} \, dz \qquad (10)$$

where $\hat{x}(0) = \log[x(0)]$ and $X(z)$ is the z-transform of the data sequence $x(nT)$. Frequently, $\hat{X}(z)$ is used to denote the $\log X(z)$; then $\hat{x}(nT)$, the complex cepstrum, is the inverse z-transform of $\hat{X}(z)$. The contour of integration lies within an annular region in which $\hat{X}(z)$ has been defined as single valued and analytic. If we have the convolution of two sequences, then

$$x(nT) = f(nT) \bullet g(nT) \qquad (11)$$

or

$$X(z) = F(z) G(z) \qquad (12)$$

and

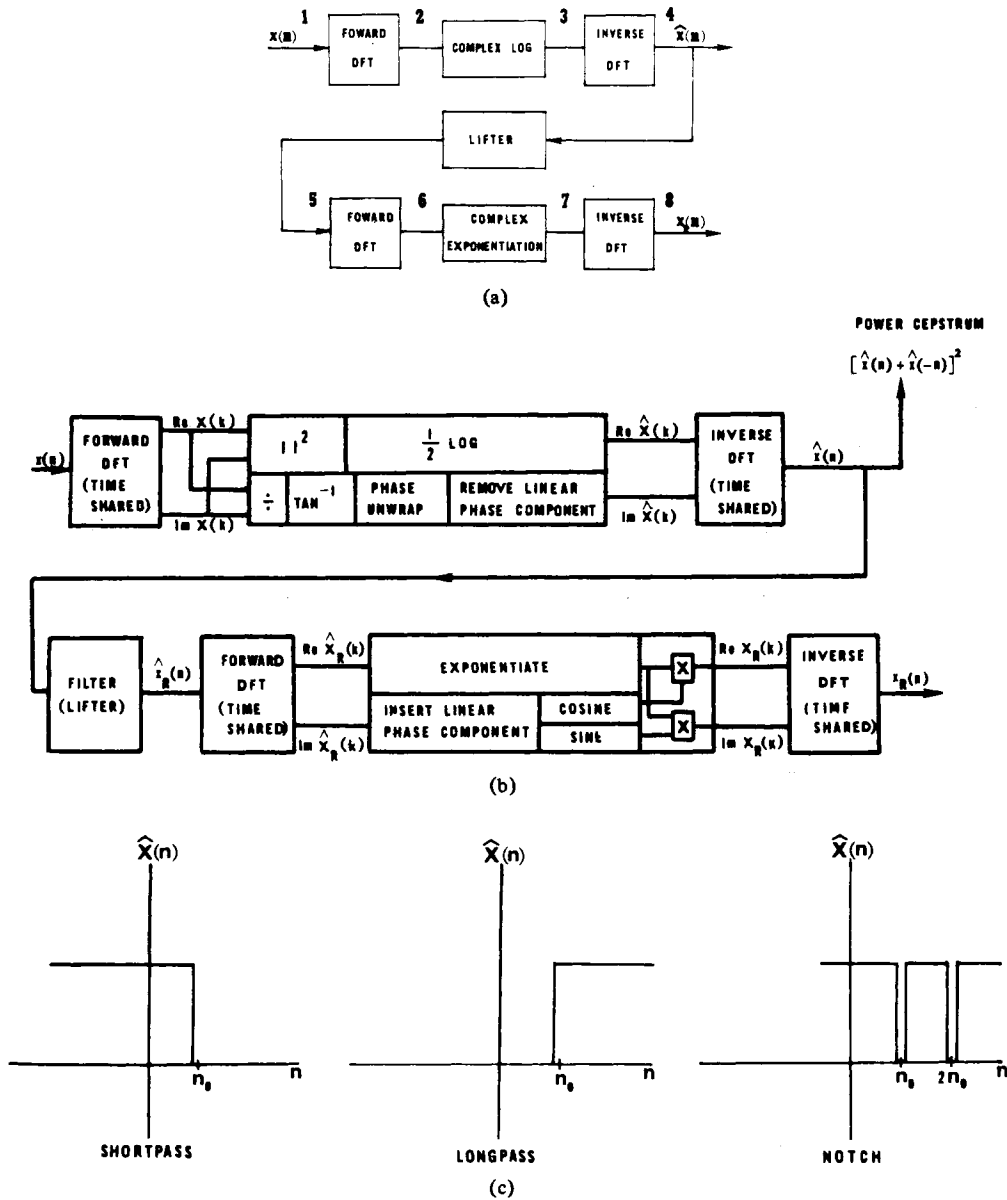$$\hat{X}(z) = \log X(z) = \log F(z) + \log G(z) \qquad (13)$$

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

or

$$\hat{x}(nT) = \hat{f}(nT) + \hat{g}(nT). \qquad (14)$$

Further, if $\hat{f}$ and $\hat{g}$ occupy different quefrency ranges, then the complex cepstrum can be liftered (filtered) to remove one or the other of the convolved sequences. Since the phase information is retained, the complex cepstrum is invertible. Thus if $\hat{g}(nT)$ is rejected from $\hat{x}(nT)$ by liftering, then $\hat{x} = \hat{f}$ and we may then z-transform, exponentiate, and inverse z-transform to obtain the sequence $f(nT)$, i.e., f and g have been deconvolved.

Fig. 1 illustrates an overall wavelet recovery or homomorphic deconvolution (filtering) system which is not only functional for off-line computations but can be implemented in real-time [26]-[28]. Examples of long pass, short pass, and notch lifters (filters) are also shown. These lifters are analogous to high pass, low pass, and notch filters, respectively, in the frequency do-

main. The notch lifter has been frequently referred to as a comb lifter in the literature.

1) Phase Unwrapping: The computation of the complex cepstrum is complicated by the fact that the complex logarithm is multivalued. If the imaginary part of the logarithm is computed module 2π, i.e., evaluated as its principal value, then discontinuities appear in the phase curve. This is not allowed since $\log(X(z))$ is the z-transform of $\hat{x}(nT)$ and thus must be analytic in some annular region of the z-plane. This problem may be rectified by making the following observations:

1) The imaginary part of $\log(X(z))$ must be a continuous and periodic (evaluated on the unit circle) function of ω with period $(2\pi/T)$ since it is the z-transform of $\hat{x}(nT)$.
2) Since it is required that the complex cepstrum of a real function be real, it follows that the imaginary part of $\log(X(z))$ must be an odd function of ω.
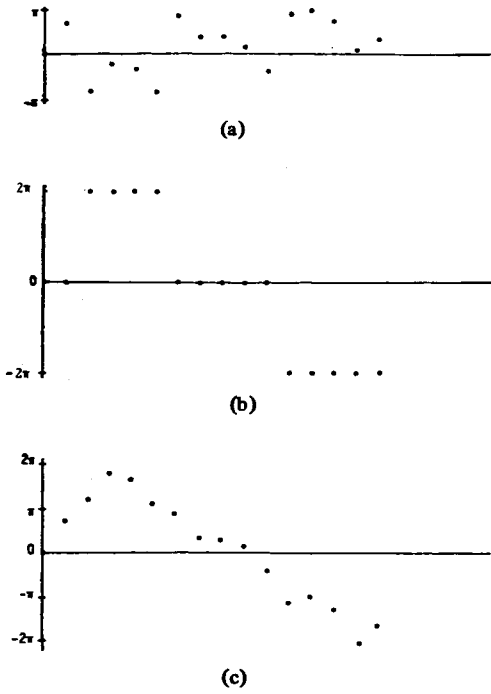
Fig. 2. Phase unwrapping. (a) Phase modulo 2π. (b) $C(k)$, the correction sequence. (c) Unwrapped phase.

Subject to these conditions we may compute the unwrapped phase curve as follows [21], [22] (provided the phase is sampled at a rate sufficiently great to assure that it never changes by more than π between samples [26], [27]): a correction sequence $C(k)$ is added to the modulo 2π phase sequence $P(k)$ where $C(k)$ is

$$C(o) = 0$$
$$C(k) = \begin{cases} C(k-1) - 2\pi, & \text{if } P(k) - P(k-1) > \pi \\ C(k-1) + 2\pi, & \text{if } P(k-1) - P(k) > \pi \\ C(k-1), & \text{otherwise.} \end{cases} \quad (15)$$

This is illustrated in Fig. 2.

Alternately, the phase may be unwrapped by computing the relative phase between adjacent samples of the spectrum. These phases may be added to achieve a cumulative (unwrapped) phase for each point. Both methods have the drawback that the computation must be done sequentially. It is also noted that if the phase never changes by more than $\pi/2$ between samples, the phase modulo π could be computed and unwrapped with algorithms similar to the above. This is interesting since it is slightly easier to calculate the phase modulo π than the phase modulo 2π (the arctangent algorithm is simpler) and many signals have this property (though noise generally does not).

Several other phase unwrapping procedures have been discussed, e.g., integrating the phase derivative [21], [31], an adaptive numerical integration procedure [56], and a recursive procedure to remove the linear phase [31].

Phase unwrapping is unnecessary for the class of minimum phase signals, i.e., a sequence whose z-transform has no poles or zeros outside the unit circle, which implies that $\hat{x}(nT) = 0$ for $n < 0$ [21], [61], [63], [64]. The complex cepstrum of such a sequence is zero at negative quefrencies. Further, for $n > 0$ the complex cepstrum is identical to the power cepstrum (except for a factor of 2 and the squaring operation); for $n = 0$ the two cepstra are identical (except for the squaring opera-

tion). Analogously, a maximum phase sequence may be defined (the z-transform has no poles or zeros inside the unit circle). The complex cepstrum for such a sequence is zero for positive quefrencies. Maximum and minimum phase sequences are discussed more extensively in [18], [33], [63], [84]. This is a very important topic since the signals of general interest are frequently of mixed phase. It is difficult to properly process such signals especially in the presence of noise. It has recently been shown that a procedure called homomorphic prediction, which is a combination of homomorphic deconvolution and linear prediction, is quite helpful in processing such mixed phase signals [18], [42], [57]. We mention this again later.

We now show that the impulses that appear in the complex cepstrum can be caused by the presence of a single additive echo. These impulses are nonzero on only one side of the origin and are therefore referred to as minimum or maximum phase impulse trains.

In (11) let $g(nT) = \delta(nT) + a\delta(nT - n_0 T)$, then

$$x(nT) = f(nT) + af(nT - n_0 T). \quad (16)$$

Taking the z-transform and evaluating it on the unit circle, we have

$$X(e^{j\omega T}) = F(e^{j\omega T})(1 + ae^{-j\omega n_0 T}). \quad (17)$$

Taking the log of both sides, we obtain

$$\hat{X}(e^{j\omega T}) = \log(F(e^{j\omega T})) + \log(1 + ae^{-j\omega n_0 T}). \quad (18a)$$

If $a < 1$ (corresponding to a minimum phase sequence), then we may expand the right most term in (18) in a power series, then

$$\hat{X}(e^{j\omega T}) = \log(F(e^{j\omega T})) + ae^{-j\omega n_0 T} - \frac{a^2}{2} e^{-2j\omega n_0 T}$$
$$+ \frac{a^3}{3} e^{-3j\omega n_0 T} \cdots. \quad (19a)$$

Inverse z-transforming, we have the complex cepstrum

$$\hat{x}(nT) = \hat{f}(nT) + a\delta(nT - n_0 T) - \frac{a^2}{2} \delta(nT - 2n_0 T)$$
$$+ \frac{a^3}{3} \delta(nT - 3n_0 T) \cdots. \quad (20a)$$

Thus the complex cepstrum of the composite signal consists of the complex cepstrum of the basic wavelet $\hat{f}$ plus a train of δ functions located at positive quefrencies at the echo delay (and its multiples) whose amplitudes are directly related to the echo amplitude. Notch liftering and interpolation (smoothing) can be performed to remove these δ functions [9], [11], [12], [19], [21]–[24], [50], [51], [63], [64]. The basic wavelet can then be recovered by inverting the operations used to compute the complex cepstrum (see Fig. 1). If the complex cepstra of the basic wavelet and the impulse train are sufficiently separated in quefrency, then short-pass liftering can be used to recover the basic wavelet. Analogously, the impulse train g, can be recovered by using long-pass liftering.

If the echo amplitude is greater than unity $a \geqslant 1$, (corresponding to a maximum phase sequence) then (18) can be rewritten as

$$\hat{X}(e^{j\omega T}) = \log(ae^{-j\omega n_0 T} F(e^{j\omega T})) + \log\left(1 + \frac{1}{a} e^{j\omega n_0 T}\right)$$
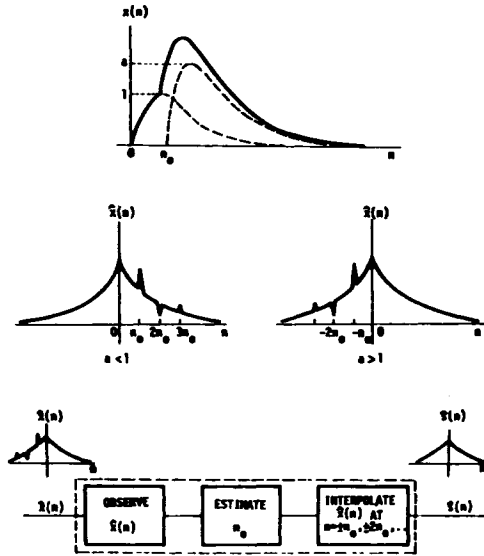
$$(18b)$$

Fig. 3. The superposition of two wavelets to form $x(nT)$; the complex cepstra for $a < 1$ and $a > 1$; and the liftering of $\hat{x}(nT)$ to eliminate the echo pulse train. This process is analogous to what is commonly called notch filtering in the frequency domain.

which may be expanded as

$$\hat{X}(e^{j\omega T}) = \log\left(ae^{-j\omega n_0 T} F(e^{j\omega T})\right) + \frac{1}{a} e^{j\omega n_0 T}$$

$$- \frac{1}{2a^2} e^{j2\omega T} \cdots$$

$$= \log\left(F(e^{j\omega T})\right) - j\omega n_0 T + \log a + \frac{1}{a} e^{j\omega n_0 T}$$

$$- \frac{1}{2a^2} e^{j2\omega n_0 T} \cdots . \tag{19b}$$

Upon removal of the linear phase term, $-j\omega n_0 T$, the complex cepstrum becomes

$$\hat{x}(nT) = \hat{f}(nT) + (\log a)\,\delta(nT) + \frac{1}{a}\,\delta(nT + n_0 T)$$

$$- \frac{1}{2a^2}\,\delta(nT + 2n_0 T) \cdots . \tag{20b}$$

Thus the complex cepstrum again has peaks at the echo delay (and its multiples), but these peaks now occur at negative rather than positive quefrencies and their gamnitudes (amplitudes) are related to $(1/a)$ rather than a. If these peaks are removed by liftering and the wavelet recovery procedure is followed including the reinsertion of the linear phase term, then the echo is recovered rather than the basic wavelet. The effect of liftering on the complex cepstrum is schematized in Fig. 3 for $a > 1$ and $a < 1$ within the context of Fig. 1.

It will be noted that the peaks in the complex cepstrum due to the impulse train may never have an amplitude greater than unity regardless of the value of a. Further, note that multiplying the original composite signal by a scale factor only changes the coefficient of the $\delta(nT)$ term in the complex cepstrum, since the scale factor appears as a shift in the mean of the log spectrum. Therefore, the complex cepstrum does not depend on the composite signal scale factor, but does depend on the SNR as well as on the ratio of the signal-to-noise bandwidths [2], [7], [30].

Another interesting, and perhaps, more representative example is the one with an infinite series of decaying echoes. Here

$$g(nT) = \delta(nT) + \alpha_1^{n_0}\,\delta(nT - n_0 T)$$

$$+ \alpha_1^{2n_0}\,\delta(nT - 2n_0 T) + \cdots \quad \text{where } 0 \leqslant \alpha_1 < 1.$$

Or with $\alpha = \alpha_1^{n_0}$, then

$$g(nT) = \delta(nT) + \alpha\delta(nT - n_0 T) + \alpha^2\,\delta(nT - 2n_0 T) + \cdots$$

$$= \sum_{m=0}^{\infty} \alpha^m\,\delta(nT - mn_0 T) \tag{21}$$

which when convolved with $f(nT)$ will give us a minimum phase sequence. Then

$$G(z) = 1 + \alpha z^{-n_0} + \alpha^2 z^{-2n_0} + \cdots$$

$$= \frac{1}{1 - \alpha z^{-n_0}}. \tag{22}$$

Thus the equation corresponding to (20a) is

$$\hat{x}(nT) = \hat{f}(nT) + \alpha\delta(nT - n_0 T) + \frac{\alpha^2}{2}\,\delta(nT - 2n_0 T) + \cdots . \tag{23}$$

This complex cepstrum is minimum phase and is nearly identical to that in (20a) except that the signs of the train of pulse functions are all positive rather than alternating in sign. The remarks following (20a) apply here as well. When α is near unity this example might be considered more representative of speech data for the situation of a sustained vowel phonation such as /i/.

In the general multiple echo case the delays become "mixed" via the series expansion of the logarithm. This greatly complicates the proper estimation of the true echo delay times [9], [9a], [12], [31], [35], [38]–[40]. The estimation is even further complicated if aliasing is severe [12], [32].

C. The Relationship Between the Complex and Power Cepstra

Clearly the complex and power cepstra are closely related. The simple formal relationship can be obtained from (1) as follows:

$$x_{pc}(nT) = (Z^{-1}\,(\log(X(z) \cdot X^*(z))))^2$$

$$= (Z^{-1}\,(\log X(z) + \log X^*(z)))^2. \tag{24}$$

Assuming $x(nT)$ is real and evaluating its z-transform on the unit circle, we find $X^*(z) = X(z^{-1})$, thus we may write

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

$$\left. + \frac{1}{2\pi j} \oint \log X(z^{-1})\, z^{n-1}\, dz\right)^2. \tag{25}$$

Letting $z' = z^{-1}$, we obtain

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

$$\left. + \frac{1}{2\pi j} \oint \log X(z')\, z'^{-n-1}\, dz'\right)^2. \tag{26}$$

Then by the definition of the complex cepstrum in (10) we have

$$x_{pc}(nT) = (\hat{x}(nT) + \hat{x}(-nT))^2. \qquad (27)$$

Thus the power cepstrum is four times the square of the even part of the complex cepstrum. This also follows from the fact that the power cepstrum is the square of the inverse transform of twice the real part of the log spectrum; and, as was noted earlier, the power cepstrum contains no phase information.

Equation (27) is of value since the power cepstrum is often superior to the complex cepstrum for echo arrival time estimation [12], [24], [25]. This is apparently due to the fact that the linear phase contribution (to be discussed below) of the imaginary part of the logarithm tends to mask the echo delay. There are probably other phase unwrapping errors as well as noise errors which contribute to this observation. Complex exponential weighting [25], which we discuss later, appears to be a method which can assist the investigator in the determination of echo delay times from the complex cepstrum.

A wavelet recovery (homomorphic filtering) system can easily compute both the power and complex cepstra as shown in Fig. 1.

Finally, as was noted earlier, if the squaring operation in (27) is not performed, then the system in Fig. 1 can be used to obtain an estimate of the log power spectrum and in turn the power spectrum of the basic wavelet. Note that if this is one's objective (and not wavelet recovery), then the problems associated with phase unwrapping are not encountered.

D. The Phase Cepstrum

The inverse transform of the phase of the complex logarithm yields peaks at multiples of the echo arrival time in much the same way that the inverse transform of the log magnitude does. This can be shown as follows for the single additive echo case:

$$\hat{X}(e^{j\omega T}) = \log (F(e^{j\omega T})) + \log (1 + ae^{-j\omega n_0 T})$$
$$= \log |F(e^{j\omega T})| + j \text{ phase } (F(e^{j\omega T}))$$
$$+ \frac{1}{2} \log (1 + a^2 + 2a \cos \omega n_0 T)$$
$$+ j \tan^{-1}\left(-\frac{a \sin \omega n_0 T}{1 + a \cos \omega n_0 T}\right). \qquad (28)$$

The fourth term on the right produces ripples in the phase, just as the third term produces ripples in the log magnitude. Since $\hat{X}(e^{j\omega T})$ is obtained from the transform of a real sequence, its real part (magnitude of the transform of the real sequence) is an even function of ω, and its imaginary part (phase of the transform of the real sequence) is an odd function of ω. Thus the inverse transform of Re $(\hat{X}(e^{j\omega T}))$ will yield the even portion of the complex cepstrum and the inverse transform of j Im $(\hat{X}(e^{j\omega T}))$ will produce the odd portion of the complex cepstrum. Since the inverse transform of the term $\log (1 + ae^{-j\omega n_0 T})$ produces peaks on one side of the origin only, the peaks produced by its real and imaginary parts must be equal in magnitude and opposite in sign on one side of the origin but of the same sign on the other side of the origin (depending upon whether the echo amplitude a is less or greater than unity).

From these observations we formally define the phase cepstrum of a data sequence as the square of the inverse z-transform of twice the phase (the imaginary part of the

logarithm) of the z-transform of the data sequence. This may be written as

$$x_L(nT) = (Z^{-1}(2 \log X(z) - 2 \log |X(z)|))^2 \qquad (29)$$

where the factor of 2 has been introduced to eliminate any normalization factors in the relation between the phase and complex cepstra and $x_L(0) = 0$. From (10), (24), and (27) the phase cepstrum can be easily shown to be

$$x_L(nT) = (\hat{x}(-nT) - \hat{x}(-nT))^2. \qquad (30)$$

Thus the phase cepstrum is to the phase as the power cepstrum is to the log magnitude. Once again the final squaring operation could be changed to magnitude squared or eliminated.

Empirically, it has been determined that the phase cepstrum is less useful than the power cepstrum in the determination of echo arrival times [26], [27]. This is apparently due to the phase unwrapping errors produced by additive noise and linear phase terms. The phase cepstrum is as difficult to compute as the complex cepstrum, since both require phase unwrapping. However, the phase cepstrum has proven valuable in evaluating the effects of noise on the signal phase [26]. Significant differences in the appearance of the phase and power cepstra can be indicative of phase unwrapping problems which might otherwise go unnoticed. This has proven to be the case in some of the work by one of the authors (DPS) on echoes generated by chirp signals.

III. Phase Perplexities

Many problems arise in the computation of the phase sequence for the complex cepstrum. Here we address several of these problems along with their alleviation.

A. Linear Phase Components

The presence of a linear component in the phase sequence introduces rapidly decaying oscillations in the complex cepstrum, e.g., let the spectrum of such a signal be represented as $X(e^{j\omega T}) = e^{-jr\omega} X'(e^{j\omega T})$ or $X(z) = z^{-r/T} X'(z)$. Then the cepstrum of the linear phase term alone is

$$\hat{x}_{LP}(nT) = \begin{cases} 0, & n = 0 \\ \dfrac{-r}{nT} \cos n\pi = \dfrac{-r}{nT}(-1)^n, & n \neq 0. \end{cases} \qquad (31)$$

This term is added to the cepstrum of the remaining portion of the data being analyzed. Note that it changes sign at each sample and although it does decay, it may be quite large depending upon r. Such a term may mask echo peaks in the complex cepstrum, and should be removed by subtraction from the composite signal phase. Several procedures for doing this appear in the literature [9], [31]. Basically, this is just trend removal, which is standard practice for improving spectral estimates. The removed linear phase term can be recorded and then reinserted during the inversion process if necessary.

The presence of a linear phase term may influence the choice of liftering to be applied to the complex cepstrum. If the echo is to be removed and the basic wavelet is to be recovered, then the echo peaks should not be notch liftered (removed) by simply replacing them with the average of their adjacent points, since these adjacent points have contributions from the linear phase component (if it has not been completely removed) which are opposite in sign to the contribution of the echo

point to be removed. Instead, if the echo is located at n_0 in the complex cepstrum then this point should be replaced with the average of the $(n_0 + 2)$ and $(n_0 - 2)$ points. This form of liftering results in a smaller mean-square error (MSE) in the recovered wavelet than when the average of the points adjacent to the echo peak is used. This has been found to be the case even when the linear phase component has been completely removed [26]. This liftering procedure is not claimed to be optimum. In fact the liftering procedure is undoubtedly signal and noise dependent and would in general involve averaging more than just two points in the complex cepstrum.

A serious problem in phase unwrapping is encountered when discontinuities in the phase occur in calculating the phase modulo 2π via the arctangent routine. The phase unwrapping algorithm previously described removes these discontinuities provided the phase changes by less than π between samples. Recently, it has been pointed out that a linear phase component with a large slope will cause errors in this unwrapping procedure [26], [31]. If the phase changes between samples are greater than π due to the presence of a linear phase term, then this unwrapping problem can be alleviated by increasing the record length with the addition of zeros [26]. This is equivalent to sampling the z-transform more frequently. If one is unsure whether the phase change between samples is less than π, then one can check such an hypothesis with the above procedure by comparing the unwrapped phase before and after the record length has been appended with zeros. Others suggest that an iterative approach to phase unwrapping is helpful [31].

One example of where the linear phase component gives problems is when $x(nT) = f(nT - n_0 T)$, $[0, N - 1]$, zero otherwise, then $X(e^{j\omega T}) = e^{-j\omega n_0 T} F(e^{j\omega T})$. As expected the phase of x is the sum of a linear phase component and the phase of f. If ω is the minimum rate ($\omega = (n2\pi)/(NT)$) then $X(e^{jn(2\pi/N)}) = e^{j2\pi n(n_0/N)} F(e^{jn(2\pi/N)})$. If $n_0 > N/2$ the linear phase component will change by more than π between samples and unless the phase of f counteracts this change, the phase unwrapping algorithm will yield erroneous results. This has been observed in computer experiments when the composite signal is delayed by more than half the record length. As expected this not only reduces the echo detectability in both the phase and complex cepstra, but also severely distorts the recovered wavelet.

B. Spectrum Notching

It should also be noted that zeros near the unit circle in the z-transform of the echo sequence result in notches in the spectrum sequence wherein additive noise may dominate. We have seen earlier that one phase unwrapping algorithm requires that the changes in phase between samples must be less than $\pm\pi$, i.e., the derivative of the phase with respect to frequency must be less than $\pm\pi$.

Consider the z-transform of the data sequence evaluated on the unit circle, then

$$X(e^{j\omega T}) = |X(e^{j\omega T})| \, e^{j\angle X(e^{j\omega T})} = X_{Re}(e^{j\omega T}) + jX_{Im}(e^{j\omega T})$$

or

Thus the change in phase is inversely proportional to the magnitude squared of the spectrum. If a notch occurs in the spectrum, then the change in the phase may be quite large, and, therefore, proper phase unwrapping may be difficult to achieve. Further, the phase may change sign rapidly in these spectrum notches. This represents a serious problem even in the absence of noise as the above example illustrates. Therefore, it is quite possible for the unwrapped phase curve to contain discontinuities (jumps or steps) in the vicinity of a spectrum notch.

As was pointed out earlier spectral nulls can be caused physically by $180°$ phase reversals in reflections at boundary interfaces [36], [37]. Nulls in the spectrum may be an important aid to data interpretation. The investigator needs to understand the physical situation under which the data are collected and to model it well [37].

C. Other Results

Finally, our results have indicated that when cepstrum analysis is performed on a bandpass function phase unwrapping outside the signal band is of little value and may actually be detrimental to wavelet recovery since the spectrum outside the signal band is dominated by noise. Similar considerations lead us to avoid oversampling since this leads to large segments of the log spectrum being dominated by noise. (See the next section.)

IV. OTHER PROBLEMS

A. Aliasing [9], [26], [32]

Aliasing of the cepstrum is of course an ever present problem since the nonlinear complex logarithm introduces harmonics into $\hat{X}(z)$. The appending of zeros to the input data sequence reduces aliasing as will selecting the data record length NT to be as large as possible. This latter choice is subject to the constraints imposed by the investigator on the number of points that can be analyzed and the minimum sampling rate. If the total data record length exceeds the duration of the composite signal contained within the record, then it is questionable if the total data record length should be further extended with still more "data." The reason for this doubt is that the spectral samples will increasingly reflect the effect of the noise rather than the signal as the total data record length surpasses that of the composite signal duration.

B. Oversampling

Oversampling of the data record when noise is present is also a problem. Outside the signal band noise dominates the spectrum. This usually presents no problem in ordinary spectrum analysis since these components frequently contain little power but this may not be the case for the cepstrum. Because of the nonlinear logarithmic operation, the regions of low power in the spectrum may contribute as much or more to the cepstrum as the regions which contain the signal in the spectrum. When this occurs it affects both echo detectability and wavelet recovery. Oversampling also aggravates phase unwrapping and aliasing since it shortens the data record (if the total number of data points or samples is fixed), which in turn

$$\frac{dX(e^{j\omega T})}{df} = \frac{X_{Re}(e^{j\omega T})[dX_{Im}(e^{j\omega T})/df] - X_{Im}(e^{j\omega T})[dX_{Re}(e^{j\omega T})/df]}{|X(e^{j\omega T})|^2} . \tag{32}$$

implies that the samples of the log spectrum are spaced farther apart.

C. Appending Zeros

It is well known that appending zeros to a data sequence increases the sampling "rate" of its discrete Fourier transform. This benefits the computation of the cepstrum in two ways. First, the increased sampling "rate" in the frequency domain reduces aliasing of the cepstrum. Second, increasing the fineness with which the phase curve is sampled reduces the number of phase unwrapping errors (which result from jumps greater than π between samples). Our results have indicated that extending the record length with zeros results in a modest improvement in the recovered wavelet even when aliasing and phase unwrapping errors do not appear to be a problem. It should be noted that unless the record length is extended with zeros, then aliasing causes an ambiguity in the determination of the echo epoch (arrival time) and amplitude. This is due to the fact that there is no way to distinguish between an echo of relative amplitude a and delay n_0 and one with amplitude $1/a$ and delay $(N - n_0)$ where N is the total number of samples.

Mathematically, these statements are verified as follows: consider the z-transform of the sequence $x(nT)$ where $x(nT) = 0$ outside $[0, N - 1]$

$$X(z) = \sum_{n=0}^{N-1} x(nT) z^{-n} \tag{33a}$$

which when evaluated on the unit circle gives

$$X(e^{j\omega T}) = \sum_{n=0}^{N-1} x(nT) e^{-j\omega nT}. \tag{33b}$$

If we sample at uniformly spaced intervals around the unit circle, we obtain

$$X_s(e^{j(2\pi m/N)}) = \sum_{n=0}^{N-1} x(nT) e^{-j(2\pi mn/N)} \tag{34}$$

which is just the DFT of $x(nT)$. It follows that

$$\hat{X}_s(e^{j(2\pi m/N)}) = \log X_s(e^{j(2\pi m/N)}). \tag{35}$$

Since the logarithm (which is a zero memory nonlinearity) of a sampled function is equivalent to sampling the logarithm of the function, then with a little additional effort it follows that the complex cepstrum of the DFT of $x(nT)$ (or the z-transform of $x(nT)$ sampled on the unit circle) is just the periodic extension of the complex cepstrum of the original data sequence. We see that the effect of appending zeros is to increase N. This implies we sample the log spectrum at smaller intervals, since the spacing between these samples is proportional to $1/N$. As described above, the errors introduced by a linear phase component or aliasing are reduced by increasing N through the appendage of zeros.

V. WINDOWING

A. The Composite Data

Echo detection and extraction are degraded by applying to the data record a window ordinarily used to reduce leakage, e.g., Hamming, Hanning, Tapering (Tukey window), unless the window is relatively constant (flat) over that portion of the data record containing the composite signal. In speech pro-cessing this is not the case. Here the data are highly non-stationary. And the investigator is frequently interested in analyzing the speech signal over one pitch period (or at most three pitch periods). In this case windowing is of considerable benefit.

One can see for the single echo case that windowing the input data record normally prevents the logarithmic operation from fully separating the basic wavelet and the echo series as follows:

$$x(nT) = [f(nT) + af(nT - n_0 T)] w(nT) \tag{36}$$

or

$$X(z) = [F(z)(1 + az^{-n_0})] * W(z). \tag{37}$$

For arbitrary $W(z)$, the contributions of the basic wavelet and the echo cannot generally be separated by taking the logarithm of (37) since the term in brackets is convolved with $W(z)$. Fortunately, as will be discussed more fully below, in practice the cepstrum procedure can still be applied with effectiveness even though there is some error.

Schafer [21] suggested a window which does preserve the separability of the basic wavelet and echo series and which has proven extremely useful in cepstrum analysis. This window denoted as

$$w(nT) = \begin{cases} \alpha^{nT}, & 0 \leqslant n \leqslant N - 1, \quad 0 < \alpha < 1 \\ 0, & \text{otherwise} \end{cases} \tag{38}$$

was first proposed to reduce the error associated with truncating the echo when it extended beyond the end of the record [21]. Our results and those of others [26], [31], [58] have indicated that this window is quite useful because it reduces the aliasing of the echo impulse train in the complex cepstrum by imposing an $(\alpha^{n_0 T})^n$ weighting on the impulses. This follows directly from a calculation of the z-transform of (36) with (38) used for $w(nT)$, i.e., for this case

$$X(z) = F(\alpha^{-T} z)(1 + a\alpha^{n_0 T} z^{-n_0}) \tag{39}$$

provided no truncation error is present and that the basic wavelet begins at $n = 0$.

From (20) we see that when no window is used and a is near unity and the echo delay is a substantial portion of the record length (NT), the higher order peaks may not decrease rapidly enough to avoid aliasing. This problem can be overcome with the window under consideration. Our results indicate that the choice of α is data dependent and α should be chosen as close to unity as possible, consistent with the desired reduction in aliasing. The closer the data sequence is to a maximum phase sequence, the more one can reduce α, e.g., from 0.99 to 0.98 or 0.96. The choice of α is also dependent on the echo delay time which is discussed more fully later.

The exponential window can introduce some distortion into the recovered wavelet even if the data are unweighted by the inverse window in the recovery process [10], [26]. This is primarily due to the distortions introduced into the data that extend beyond the duration of the wavelet of interest.

In summary the exponential window performs nearly as well as the rectangular window when no noise is present but does introduce some distortion as noted above. Further, the echo arrival time can be determined even when wavelet recovery cannot be effected. Also if rectangular windowing is judiciously applied, then the cepstrum can be used to detect

similar but not necessarily identical wavelets. We suspect that if the exponential window is used to make the composite signal minimum phase, then the echo, most probably, will be lost.

Finally, it should be noted that the exponential window may be used to alter the SNR of a data record more effectively than the rectangular window. This can be effected when the composite signal occupies only a portion of the total record. In this case the window may weight the signal more or less heavily than those portions of the record containing the noise. However, caution should be exercised in echo detection and extraction when the signal (wavelet) of interest occurs near the end of the data record and thus will be greatly reduced by the exponential window.

We wish to emphasize that the comments made in the last three paragraphs have been greatly influenced by our analyses of echo type data. It is our opinion (as well as that of others) that such observations are and will be highly data dependent.

Recently it has been proposed that the exponential window be generalized to include complex exponential weighting, i.e., $\alpha^{nT} e^{j\phi nT}$ [25]. It may at first appear that this phase factor will have no significant effect on the complex cepstrum, i.e., it will introduce only a phase shift. However, the procedure can be used to change the phase relation of the echo (multipath reflection) by 180° [25]. This may make it easier to detect a peak in the cepstrum. The complex exponential factor ϕ can be varied in a prescribed fashion so that it may be used as a hypothesis tester. Thus trial sweeps of the complex weight can be generated to confirm or deny *a priori* estimates of the echo delay [25]. It appears that this technique may prove to be a powerful investigative tool to assist the researcher in interpreting his data.

B. The Log Spectrum

One might be motivated to window the log spectrum in order to reduce leakage in the complex cepstrum which could be falsely interpreted as peaks due to echoes. Windowing of the log spectrum will, of course, introduce some loss in time resolution in the cepstrum domain. Then, if the echo contributions can be liftered from the complex cepstrum and if the recovered log spectrum can be corrected (by multiplying by the inverse of the windowing series), we should be able to recover the basic wavelet. Our results have, however, indicated that such windowing of the log spectrum raises the echo detection threshold by around 12 dB and severely distorts the recovered wavelet when additive noise is present [26]. This is apparently due to the fact that windowing the log spectrum is equivalent to smoothing the complex cepstrum. Thus, it appears that windowing the log spectrum may smooth out the very peaks one wishes to detect in the complex cepstrum. The distortion introduced into the recovered wavelet is undoubtedly due to this windowing of the log spectrum (or smoothing of the complex cepstrum).

C. The Complex Cepstrum

Since noise is usually interspersed throughout the data record and the composite signal may occupy only a portion of the record, it seems reasonable that the high quefrency components of the complex cepstrum may frequently contain more noise than signal information. Our results have shown that by judiciously zeroing the high quefrency components of the complex cepstrum we may significantly improve the

fidelity of the recovered wavelet in a noisy environment [26]. At low SNR the MSE can be reduced by a factor of 2 by a judicious rectangular windowing of the complex cepstrum. This is essentially short pass liftering [9], [12], [19], [21] in which the aim is not to eliminate the echo peaks (which are generally notch filtered prior to the windowing) but rather to eliminate the high quefrency noise dominated sections of the complex cepstrum. This concurs with the results of [12] in which it is reported that a Hanning smoothing of the log spectrum (which is equivalent to Hanning windowing of the complex cepstrum) improves wavelet recovery. It appears that there is little to choose between the rectangular or Hanning window of the complex cepstrum to improve the fidelity (MSE) of the recovered wavelet. We mention once again that these observations are probably data dependent and are influenced by the duration of the window as well.

D. Sequence Truncation

As mentioned above in subsection A errors may be introduced by truncating the echo if it extends beyond the end of the record. In addition, aliasing of the echo impulse train in the complex cepstrum may occur. These errors can be reduced appreciably by exponentially windowing the sampled data sequence. But we suspect that if the exponential window is too severe, then the echo may be lost.

VI. DATA PROCESSING—SPEECH, SEISMIC, AND HYDROACOUSTIC

Three application areas which appear to be using cepstrum analysis quite frequently are speech research, seismology, and hydroacoustics. In the three subsections that follow we have tried to itemize the major data processing procedures that investigators in each of these areas tend to use. For some situations these lists may be simplistic but we feel that they are nonetheless indicative of the typical steps considered by some if not all investigators in these areas.

A. Speech [18], [19], [21], [30], [42], [46]–[51], [53], [55], [61]–[65], [66], [71]

Within speech research there are at least three problems to which cepstrum analysis is applied. The first is perhaps the most difficult. This problem seeks to achieve the deconvolution of three signals which form the basis for a model of the speech process. This simple model considers voiced sounds to be produced by quasi-periodic pulses of air which in turn cause the vocal cords to vibrate producing glottal pulses which excite the vocal tract to finally produce speech. For nonnasal sounds the vocal tract is modeled as an all pole filter over short time intervals. The glottal source is modeled with zeros in the z-domain again over short time intervals. The vocalized speech signal is, therefore, modeled as the three fold convolution of an impulse train, the glottal impulse response, and the vocal track impulse response. These three signals are to be deconvolved. This is difficult to achieve without considerable additional information or assumptions. It is less difficult to achieve the deconvolution of the pulse (impulse) train with the composite convolution of the glottal impulse response and vocal track impulse response since these two time sequences occupy different quefrency ranges in the cepstrum domain.

Another related problem previously mentioned is to estimate the envelope of the speech spectrum. The speech spectrum is generally quite scalloped, i.e., it looks like an undulating picket fence. The scalloping or spectral modulation is

due to the speaker fundamental frequency (pitch) or periodic pulse train. The pulse train can be liftered from the cepstrum by a shortpass lifter. The inverse process shown in Fig. 1 is then followed to the point where the spectrum is obtained. This yields an estimate for the speech spectrum envelope. However, this estimate is not as good as that obtained by linear prediction (predictive deconvolution) [66], [71]. We illustrate this later with an example.

The third common problem is to achieve an estimate of the pitch period or the interval between the pulses in the excitation pulse train. This is easily accomplished by longpass liftering the cepstrum and then following the inverse process outlined in Fig. 1. The pitch period can also be measured directly from the cepstrum by measuring the time interval from the origin to the first peak.

Very recently it has been suggested that homomorphic filtering in the form of cepstrum analysis be combined with linear prediction to effect a useful pole-zero modeling and inverse filtering procedure for mixed phase signals [18], [42], [57]. This procedure has been applied to speech [42] and seismic data [57]. It is sure to see increased application within the near future.

The specific information and procedures commonly used in processing speech data are the following:

1) The speech record is usually windowed with a Hanning or Hamming window. The total window duration is often on the order of three pitch periods or less, i.e., approximately 24 ms. This will, of course, vary with the speech signal being analyzed and the individual investigator.

2) Zeros may be appended to the windowed speech signal to increase its effective record length; sometimes this is on the order of a factor of 10 but is usually less. A typical sampling rate is 10 kHz.

3) The speech bandpass is typically in the range of 50 Hz to 4 or 5 kHz with resonances appearing in the spectrum at the natural frequencies (formants) of the vocal tract. While the location of these formants varies with the phonation or articulation a great deal is known about their typical location and bandwidth. This information helps considerably in the analysis of speech data.

4) Aliasing is not a serious problem in the analysis of speech data.

5) The composite convolution of the glottal impulse response and the vocal tract impulse response is generally less than 5 ms in the cepstrum domain. The first peak in the cepstrum due to the pulse train is in the vicinity of 8 ms. This information is used to design the shortpass and longpass lifters.

Example—Speech: In this simple example we compare the speech spectral envelopes obtained by both linear prediction and cepstrum processing.

The steps followed for cepstrum processing are those shown in Fig. 1(a) [or 1(b)]. The resultant envelope is obtained at step (6). Included in the complex logarithm step is the calculation of the squared magnitude of the DFT. All DFT calculations are performed by an FFT algorithm.

The speech record analyzed appears in Fig. 4(a) and was sampled at 10 kHz to give 256 data points with a 512 point FFT being used. The phonation was a sustained /i/. The record was then Hanning windowed (256 points) prior to cepstrum analysis. No zeros were appended for this example. The "cepstrum" (inverse DFT of log $\|\|^2$) is shown is Fig. 4(b). The peak at 6.7 ms corresponds to the pitch period which can also be estimated from Fig. 4(a). Next, shortpass liftering was applied. The lifter was constant from zero to 1.5 ms and then
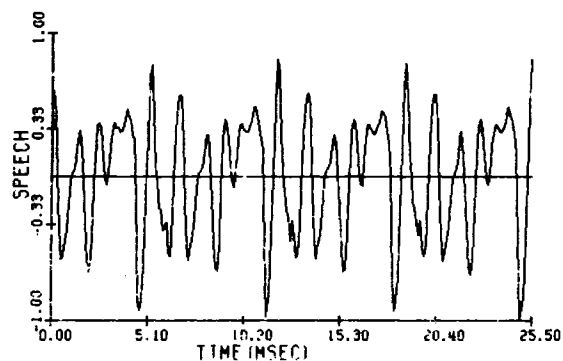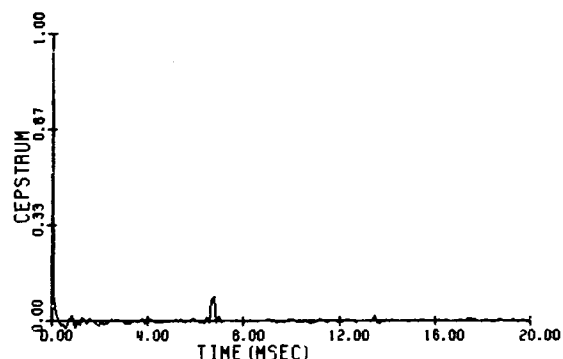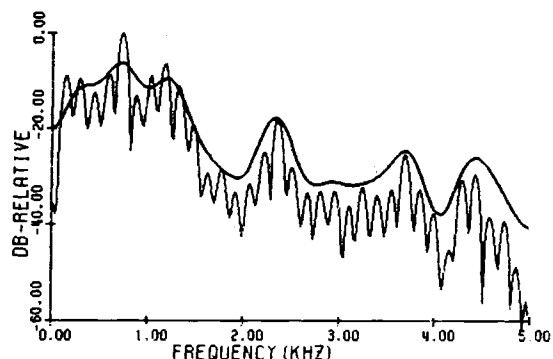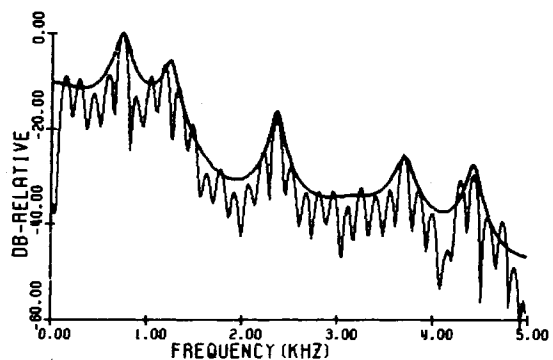

Fig. 4. (a) Speech signal. (b) The corresponding cepstrum.

Fig. 5. (a) Speech spectrum of Fig. 4(a) and envelope derived by liftering the cepstrum in Fig. 4(b) (the superimposition is less than perfect). (b) Speech spectrum of Fig. 4(a) and envelope derived by linear prediction.

had a cosine taper from 1.5 ms to 2.5 ms and was zero beyond 2.5 ms. The final step was a forward FFT to yield the speech spectrum envelope shown in Fig. 5(a) which is superimposed albeit imperfectly upon the power spectrum of windowed speech signal. The lifter selected gave the "best" spectral envelope fit in the opinion of the authors. However, the results are sensitive to the type of lifter used.

The peaks shown correspond to the formants, but it is possible in some cases for the cepstrum procedure described to yield false formants in between the actual formants.

The linear prediction procedure calculates the coefficients for an all pole filter

$$\frac{1}{1 - \sum_{k=1}^{P} a_k z^{-k}}$$

from the windowed speech signal by an autocorrelation technique [71]. For this example the windowed data was 20 ms with $p = 16$. The envelope of the spectrum was determined by finding the FFT of the sequence $\{1, -a_1, -a_2, \cdots, -a_p\}$ and then calculating the reciprocal of the FFT. Zeros were appended prior to calculating the 512 point FFT. The results appear in Fig. 5(b) superimposed upon the speech power spectrum. It will be noted that linear prediction provides a "better" spectral envelope and yields smaller formant bandwidths than cepstrum processing.

B. Seismic Data Processing [1], [18], [34], [36], [37], [57]-[59], [80]

For more than a decade cepstrum analysis has been applied to seismic data. 1) to determine the focal depth of a seismic event; 2) to remove spectral modulations caused by multipath reflection; and 3) to determine the slapdown phase which results from spallation of the earth's surface near ground zero, and other situations. Knowledge of the depth of the seismic event can be used to help discriminate whether the event is an earthquake or a man-made explosion. The elimination of multipath reflections assists the data interpreter in the determination of source (event) parameters. Similar remarks apply to spalling as well.

Usually, the power cepstrum, or a variation thereof (as discussed earlier), is used to estimate the $(P - pP)$ time difference (see Fig. 6) which is the most realistic indicator of the depth of the event. The complex cepstrum is used to separate (deconvolve) the basic P phase wavelet from the impulse train caused by the echo or echoes. Epoch timing information can on occasion be obtained by longpass liftering the complex cepstrum. But this may require that the event be deep so that the P phase cepstral information and the echo information are adequately separated in quefrency.

Good success in processing seismic data is, however, apparently dependent upon a good SNR and wide bandwidth data [34].

For many years predictive deconvolution (or linear prediction) has been used in the analysis of seismic data [66]-[86], (see [80] for a review). This procedure is parametric and proposes a model for the basic wavelet. As such it does not work particularly well if the wavelet to be removed is mixed phase. Cepstrum analysis (homomorphic deconvolution) is a more general method for deconvolution, and is effective when the cepstra of the signals to be deconvolved occupy different quefrency ranges. A method, called homomorphic prediction,
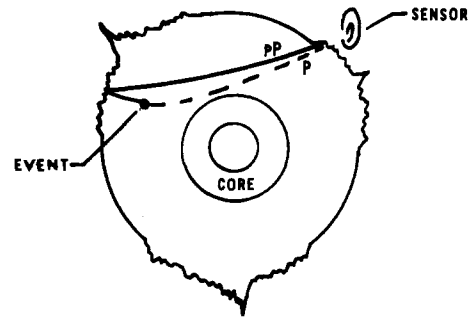

Fig. 6. Simple representation of earth and two seismic waves resulting from a seismic event. The P wave is the direct, longitudinal wave. The pP wave is the single reflected wave. Neither passes through the inner or outer cores, but rather propagate through the mantle.

has just been proposed which combines linear prediction and homomorphic deconvolution to more effectively analyze mixed phase data [18], [42], [57].

The specific information and procedures commonly used in processing seismic data are the following:

1) A short time series window (usually starting about 1 s before the onset of the signal and lasting from 3 to 7 s into the data record) is used. Some form of tapering is almost always applied, e.g., for a 7 s window, a linear taper may be applied for 1 s before the signal onset as well as to the last second of the record with the window being constant for the 5 s in between the two tapered ends. Longer windows are not generally used since these would include more of the coda (i.e., tails of the data) which contain too many multipath reflection signals.

2) A weighting of the time series may also be used, independent of whether a window is applied. This weighting is generally in the form α^{nT}, where $0.96 \leqslant \alpha < 1.0$. This procedure tends to make the P phase more minimum phase.

3) Zeros are appended to the time series to increase its length, sometimes by as much as a factor of 10. A typical sampling rate is 20 samples per s. The windowed data record may then be 5 s to give 100 data samples. This record is then extended with zeros to a duration of 1024 samples.

4) The seismic bandpass for body waves is generally considered to be in the range 0.1 to 2-5 Hz depending on various factors [36]. Echo delay (epoch) times for seismic events are in the range $0.1 \leqslant n_0 T \leqslant$ several seconds. The lower end of this range requires 10 Hz bandwidth which is not always available. At 20 samples per s, the data is over sampled so aliasing is of no concern.

5) Since seismic data has a narrow passband, the spectrum is nonwhite. An elastic absorption by the earth of teleseismic signals above 2 Hz means that the cepstrum is dominated by the bandpass characteristic of the earth [36]. This effect is aggravated by the frequency response of the sensing instrument, but this can be corrected by inverse filtering [36]. Pseudo-heterodyning, a translation in quefrency (a method analogous to conventional heterodyning), may be helpful for such data [9(a)].

6) Cepstra are also calculated using the coda and then averaged, taking into account the predictable travel time differences. Spectra are also similarly averaged to enhance spectral nulls [36], [37].

Example—Seismic Data: This example may be considered a simulation of seismic data with the echo being negative at the air-earth interface. In Fig. 7 we present four groupings of three graphs each. The first graph in the first group is the
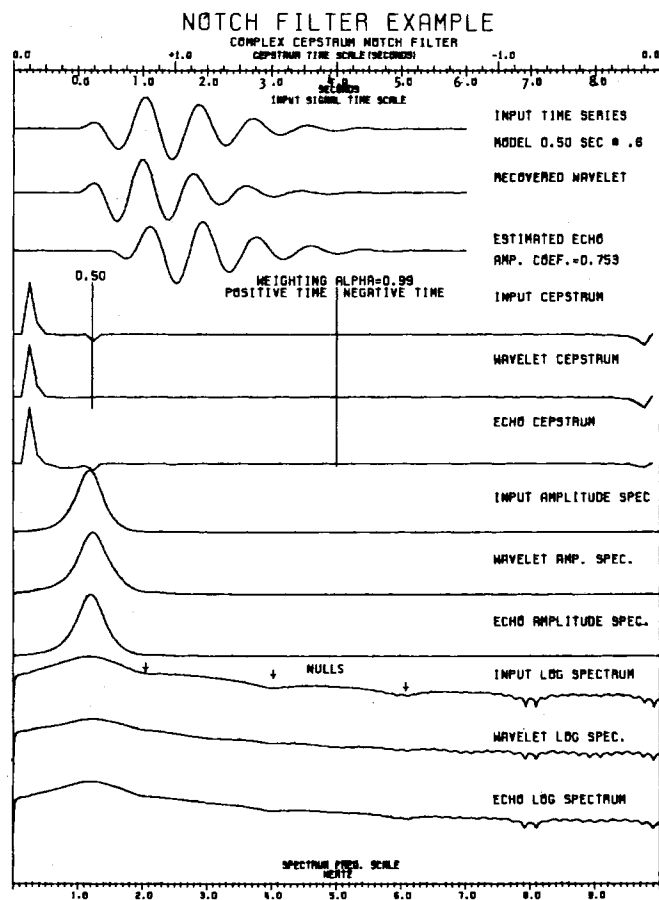
NOTCH FILTER EXAMPLE

Fig. 7. Simulated seismic data example. The normalized amplitude of the wavelet plus a negative echo is 1000 units. The echo occurs at 0.5 s and is 0.6 that of the wavelet. An exponential window ($\alpha = 0.99$) was applied to the composite signal prior to calculating the cepstrum. The four groupings of graphs show the time domain data and the corresponding cepstra, spectra and log spectra, respectively.

composite signal, i.e., the wavelet plus the negative echo. The overall normalized peak amplitude of the composite signal is 1000 units. The echo occurs 0.5 s after the onset of the basic wavelet. The echo amplitude is 0.6 that of the original wavelet. The second graph in this group is the wavelet recovered by homomorphic filtering, i.e., notch liftering (rejecting) only the first peak in the cepstrum. The third graph is an estimate of the echo obtained by calculating the difference between the first and second graphs. At the top of this figure are the time scales for both the input data and the cepstrum. The spectrum frequency scale appears at the bottom of the figure.

An exponential window ($\alpha = 0.99$) was applied to the composite signal prior to calculating the cepstrum. This made the cepstrum of the composite signal nearly minimum phase.

The second grouping of three graphs presents the cepstra for the previous three waveforms respectively. The first two points at the extreme left in each cepstra were zeroed. The vertical line at 0.5 s shows the peak that was notch liftered. Ideally, the second and third cepstra should be identical when normalized, but due to the fact that only one cepstrum peak was notch liftered, these waveforms are in fact different.

The third grouping of three waveforms presents the respective amplitude spectra. Finally, we have in the last grouping the respective log spectra. The spectral nulls due to the echo can barely be seen due to the narrow bandwidth simulated in the first graph. The second graph is the log spectrum after

notch liftering. Here it can be seen that the spectral nulls have not been completely removed. And finally we have the log spectrum of the estimate of the echo wavelet.

C. Hydroacoustic Data Processing [18], [31], [35], [38], [45], [57]

The power cepstrum has been used to estimate the source depth of a known explosive charge by analyzing the data received at long ranges. This is accomplished by measuring the time period of the bubble pulse modulation on the spectrum which results in a peak in the power cepstrum [45].

The cepstrum has also been used to investigate multipath conditions in shallow water as well [31], [35], [38]. But the echoes are not all identical in waveshape as commonly assumed. The cepstrum is also apparently affected to a considerable degree by fluctuations in the transmission media, bottom reflections, and surface scattering [38].

As mentioned under the subsections on speech and seismic data processing, it has been proposed that linear prediction (predictive deconvolution) and homomorphic deconvolution be combined to more effectively analyze mixed phase signals [18], [42], [57]. This suggestion has been tested on marine seismic data with apparently good success [57].

The specific information and procedures commonly used in processing hydroacoustic data are the following:

1) The sampling rates used are dependent on specific applications as well as the computational resources available, but typically they are in the range of 100 to 1000 samples per s.

2) The effects of windowing are not considered to be as important as for seismic data. Windows commonly used include the rectangular (boxcar), Hamming, Hanning (cosine), and linear taper.

3) Rather than extend the data record with zeros, a longer time window is frequently used. It is commonly assumed that hydroacoustic data is stationary, thus time averaging can be employed. Typical window durations are 1024 or 2048 samples, or longer depending on the computational capacity available.

4) The spectrum is whitened prior to computing the power cepstrum. The whitening may be achieved with either the square root or the logarithmic operation.

5) The spectrum is often highpass liftered to remove low-quefrency components prior to computing the power cepstrum. This is a form of trend removal to reduce leakage in the cepstrum domain.

6) The expected echo delay (epoch) times are $0.02 \leqslant n_0 T \leqslant 2$ or 3 s. The bandwidth is at least 50 Hz.

VII. CONCLUDING REMARKS

A. The Effects of Noise

The effects of noise are discussed in a qualitative way at various points throughout the paper along with two recommended procedures for alleviating these effects for wavelet extraction and echo detection, namely, windowing the complex cepstrum and reducing errors due to aliasing and phase unwrapping by appending zeros to the sampled data sequence. Noise analysis is presented in a more quantitative and extensive manner in [2], [7], [9], [12], [26], [30]. We point out in particular that it has been recently shown that SNR alone is an insufficient measure for determining cepstrum performance and that the relative bandwidths of the signal and noise are also needed [7].

B. Summary

We have attempted to provide the reader with a unified approach to the power, complex, and phase cepstra, namely if $\hat{x}(nT)$ denotes the complex cepstrum, then the power cepstrum is $x_{pc}(nT) = (\hat{x}(nT) + \hat{x}(-nT))^2$ and the phase cepstrum is $x_\angle(nT) = (\hat{x}(nT) - \hat{x}(-nT))^2$.

The problems associated with phase unwrapping, linear phase components, and spectrum notching have been described along with those of aliasing and over sampling. The extension of the sampled data sequence by appending zeros was shown to provide computational benefits, namely, the reduction of aliasing in the cepstrum and the reduction of phase unwrapping errors.

As a recapitulation we offer the following comments relative to the procedures followed by investigators in the various fields at each step in the cepstrum process shown in Fig. 1. First, the data are usually windowed or weighted in some manner. But the type of window used is data dependent. This is true even for echo type data. Here the exponential window may be useful or no window at all (i.e., the rectangular window). For speech data the Hamming or Hanning windows are frequently used. After windowing zeros are usually appended to extend the data record. The first DFT is then performed with an FFT algorithm. The magnitude (or magnitude squared) of the DFT is then usually found, even if the complex cepstrum is to be calculated. At this point a number of procedures may be followed. The square root of the magnitude of the DFT may be calculated in order to whiten the spectrum. The logarithm performs a similar function but may be considered too severe a whitener. Trend removal may also be performed on the spectrum to prevent leakage in the cepstrum. Phase unwrapping is also done during this stage if the complex cepstrum is to be calculated. It appears that windowing the log spectrum should not be performed for the reasons already given. And apparently zeros are rarely appended to the log spectrum. At the next step the forward or inverse FFT may be calculated. The forward FFT is usually calculated if the power cepstrum is desired. The inverse FFT of the liftered (usually long pass) power spectrum gives the pseudo-autocorrelation which contains sign information concerning echo reflections. The inverse FFT of the log magnitude spectrum is also useful for obtaining estimates of the spectrum envelope. If the inverse FFT of the log magnitude spectrum is squared, then the power cepstrum is obtained. The inverse FFT of the whitened spectrum is also frequently found as well. The complex cepstrum is the inverse FFT of the complex logarithmic spectrum, i.e., keeping track of phase. The "cepstrum," i.e., any of the above forms, is then either short-pass, longpass, or notch (comb) liftered. Included in this liftering operation may be a procedure to simply zero the cepstrum at various points or the cepstrum may be windowed. After liftering the inverse steps are usually followed in a conventional manner. The investigator may terminate the process at any step depending upon his or her goal.

The complex cepstrum is basically a method for deconvolving a train of impulses from a basic wavelet. Thus the form of the wavelet or the echo impulse train or both can be recovered. But this procedure clearly has its limitations. This is why we have included a limited list of references to the literature on linear prediction, predictive deconvolution, inverse filtering, and general deconvolution [66]-[80] including some books on computational seismology [81]-[86]. For those who may wish to learn more about homomorphic systems and the cepstrum including applications we recommend [61]-[65] in addition to the specific papers already cited. Aspects of real-time computation of the complex cepstrum are discussed in [26], [28]. And those interested in homomorphic prediction to process mixed phase signals should consult the recent work in [18], [42], [57].

We wish to point out that the principles of homomorphic deconvolution have recently been applied to effect a transformation of the independent variable t rather than to effect a deconvolution of two signals [60]. This results in a new type of nonlinear filter which because of its signal dependency can filter out superimposed noise on a signal, leaving large peaks of the signal unattenuated. This filter is signal dependent; it is therefore apparently not a true homomorphic technique. The filter can be realized in real time. One of the authors (DGC) has successfully simulated these results on a large computer.

We anticipate that there will be other interesting results in the future and we hope that this paper will both stimulate and assist others to this end.

REFERENCES

Theory—Homomorphic Systems and Cepstra (with Applications)

[1] B. P. Bogert, M. J. Healy, and J. W. Tukey, "The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking," in *Time Series Analysis*, M. Rosenblatt, Ed. New York: Wiley, 1963, Chap. 15, pp. 209-243.

[2] B. P. Bogert and J. F. Ossanna, "The heuristics of cepstrum analysis of a stationary complex echoed Gaussian signal in stationary Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 373-380, July 1966.

[3] J. F. Bohme, "The cepstrum as a generalized function," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 650-653, Sept., 1974.

[4] D. Childers, "Composite signal decomposition techniques," in *Int. Conf. on Comm.* (ICC), Seattle, WA, June 11-13, 1973, pp. 1-6.

[5] D. E. Dudgeon, "Existence of cepstra for two-dimensional rational polynomials," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 242-243, Apr. 1975.

[6] J. C. Hassab, "On the convergence interval of the power cepstrum," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 111-112, Jan. 1974.

[7] J. C. Hassab and R. Boucher, "A probabilistic analysis of time delay extraction by the cepstrum in stationary Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 444-454, July 1976.

[8] ——, "Analysis of signal extraction, echo detection and removal by complex cepstrum," *J. Sound and Vibration*, vol. 40, pp. 321-335, June 1975.

[9] R. C. Kemerait, "Signal detection and extraction by cepstrum techniques," *Ph.D. Dissertation*, University of Florida, 1971.

[9a] ——, "Pseudo-heterodyning in the cepstral domain," in *8th Ann. Southeastern Symp. System Theory*, Apr. 26-27, 1976, pp. 37-41.

[10] R. C. Kemerait and L. Balceda, "Signal detection and extraction by weighted cepstrum techniques," in *1976 Southeastcon*, Clemson University, IEEE Catalog No. 76 CH1059-5 Reg. 3, pp. 3B-1-3B-3, 1976.

[11] R. Kemerait and D. G. Childers, "Composite signal decomposition by cepstrum techniques," in *IEEE Region 3 Convention Record*, pp. K1-1-K1-4, 1972.

[12] ——, "Signal detection and extraction by cepstrum techniques," *IEEE Trans. on Inform. Theory*, vol. IT-18, pp. 745-759, Nov. 1972.

[13] A. Noll, "The cepstrum and some close relatives," in *Signal Processing*, J. W. R. Griffiths, P. L. Stocklin, and C. Van Schooneveld, Eds. London: Academic Press, 1973, pp. 11-22.

[14] A. Oppenheim, "Superposition in a class of nonlinear systems," M.I.T. Res. Lab. of Electronics, Cambridge, MA, Tech. Rep. 432 (Ph.D. dissertation), Mar. 31, 1965.

[15] ——, "Optimum homomorphic filters," M.I.T. Res. Lab. of Electronics, Quarterly Progress Rep. no. 77, vol. XIII, Statistical Communication Theory, AD 615324, Apr. 15, 1965, pp. 248-260.

[16] ——, "Nonlinear filtering of convolved signals," M.I.T. Res. Lab. of Electronics, Quarterly Progress Rep. no. 80, 1966.

[17] ——, "Generalized superposition," *Inform. Contr.*, vol. 11, pp. 528-536, Nov.-Dec., 1967.

[18] A. V. Oppenheim, G. E. Kopec, and J. M. Tribolet, "Signal analy-

sis by homomorphic prediction," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-24, pp. 327–332, Aug. 1976.

[19] A. V. Oppenheim, R. W. Schafer, and T. G. Stockham, Jr., "Nonlinear filtering of multiplied and convolved signals," *Proc. IEEE*, vol. 56, pp. 1264–1291, Aug. 1968.

[20] R. Rom, "On the cepstrum of two-dimensional functions," *IEEE Trans. on Inform. Theory*, pp. 214–217, Mar. 1975.

[21] R. W. Schafer, "Echo removal by discrete generalized linear filtering," Ph.D. Dissertation, M.I.T., Cambridge, MA, 1968.

[22] S. Senmoto, "Adaptive decomposition of composite signals in noise," Ph.D. Dissertation, University of Florida, Gainesville, FL, 1971.

[23] S. Senmoto and D. G. Childers, "Analysis of a composite signal by complex cepstrum and adaptive filter," *Trans. Inst. Elec. Comm. Engrs.*, (Japan), pt. A, pp. 9–16, 1972.

[24] ——, "Adaptive decomposition of a composite signal of identical unknown wavelets in noise," *IEEE Trans. on Syst., Man, and Cybern.*, vol. SMC-2, pp. 59–66, Jan. 1972.

[25] M. J. Shensa, "Complex exponential weighting applied to homomorphic deconvolution," *Geophys. J. Roy. Astron. Soc.*, vol. 44, pp. 379–387, 1976.

[26] D. P. Skinner, "Real-time composite signal decomposition," Ph.D. Dissertation, University of Florida, Gainesville, FL, 1974.

[27] D. P. Skinner and D. G. Childers, "The power, complex, and phase cepstra," presented at 1975 Proc. Nat. Telecommun. Conf., Dec. 1975.

[28] ——, "Real-time composite signal decomposition," *IEEE Trans. on Acoust., Speech, and Signal Processing*, vol. ASSP-24, pp. 267–270, June 1976.

[29] R. G. Smith, "Cepstrum descrimination function," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 332–334, May, 1975.

[30] T. G. Stockham, Jr., T. M. Cannon, R. B. Ingebretsen, "Blind deconvolution through digital signal processing," *Proc. IEEE*, vol. 63, pp. 678–692, Apr. 1975. (Refer to for other related reports and Master's theses.)

[31] P. L. Stoffa, P. Buhl, and G. M. Bryan, "The application of homomorphic deconvolution to shallow-water marine seismology —Part I: Models," *Geophysics*, vol. 39, no. 4, pp. 401–416, Aug. 1974.

[32] ——, "Cepstrum aliasing and the calculation of the Hilbert transform," *Geophysics*, vol. 39, no. 4, pp. 543–544, Aug. 1974.

Theory-Related Material

[33] A. J. Berkhout, "On the minimum phase criterion of sampled signals," *IEEE Trans. Geosci. Electron*, vol. GE-11, pp. 186–198, Oct. 1973.

Applications–Homomorphic Filtering and Cepstra (with Theory)

[34] W. H. Bakun and L. R. Johnson, "The deconvolution of teleseismic P waves from explosions Milrow and Cannikin," *Geophys. J. Roy. Astron. Soc.*, vol. 34, pp. 321–342, 1973.

[35] P. Buhl, P. L. Stoffa, and G. M. Bryan, "The application of homomorphic deconvolution to shallow-water marine seismology —Part II: Real data," *Geophysics*, vol. 39, no. 4, pp. 417–426, Aug. 1974.

[36] T. Cohen, "Source-depth determinations using spectral, pseudo-autocorrelation and cepstral analysis," *Geophys. J. Roy. Astron. Soc.*, vol. 20, pp. 223–231, 1970.

[37] T. J. Cohen, "P_s and pP phases from seven Pahute Mesa events," *Bull. Seismological Soc. Amer.*, vol. 65, no. 4, pp. 1029–1032, Aug. 1975.

[38] P. O. Fjell, "Use of the cepstrum method for arrival times extraction of overlapping signals due to multipath conditions in shallow water," *J. Acoust. Soc. Amer.*, vol. 59, no. 1, pp. 209–211, Jan. 1976.

[39] O. S. Halpeny, "Epoch detection by cepstrum analysis," Master's Thesis, University of Florida, Gainesville, FL, 1970.

[40] R. Kemerait and D. G. Childers, "Detection of multiple echoes immersed in noise," in *Proc. 15th Midwest Symp. Circuit Theory*, May 4–5, 1972 pp. 1–10.

[41] ——, "Decomposition of pulse-type data by cepstrum techniques," *1973 IEEE Southeast Conf.*, Apr. 30, May 1–2, 1973, pp. F-4-1–F-4-4.

[42] G. E. Kopec, A. V. Oppenheim, and J. M. Tribolet, "Speech analysis by homomorphic prediction," submitted to *IEEE Trans. Acoust., Speech and Signal Processing*. (This paper is apparently an earlier version of [18].)

[43] L. R. LeBlanc, "Narrow-band sampled-data techniques for detection via the underwater acoustic communication channel," *IEEE Trans. Commun. Technol.*, vol. COM-17, pp. 481–488, Aug. 1969.

[44] J. H. Miles, G. H. Stevens, and G. G. Leininger, "Analysis and correction of ground reflection effects in measured narrowband

sound spectra using cepstral techniques," presented at 90th Meeting of Acoust. Soc. Amer., Nov. 4–7, 1975.

[45] S. K. Mitchell and N. R. Bedford, "Long range sensing of explosive source depths using cepstrum," paper presented at 90th Meeting Acoust. Soc. Amer., Nov. 1975.

[46] A. Noll, "Short-time spectrum and 'cepstrum' techniques for vocal-pitch detection," *J. Acoust. Soc. Amer.*, vol. 36, pp. 296–302, Feb. 1964.

[47] ——, "Cepstrum pitch determination," *J. Acoust. Soc. Am.*, vol. 41, no. 2, pp. 293–309, Feb. 1967.

[48] ——, "Clipstrum pitch determination," *J. Acoust. Soc. Am.*, vol. 44, no. 6, pp. 1585–1591, Dec. 1968.

[49] ——, "Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and a maximum likelihood estimate," in *Computer Processing in Communications*, J. Fox, Ed. Brooklyn, NY: Polytechnic, 1970, pp. 779–797.

[50] A. Oppenheim, "Speech analysis-synthesis system based on homomorphic filtering," *J. Acoust. Soc. Amer.*, vol. 45, no. 2, pp. 458–465, Feb. 1969.

[51] A. Oppenheim and R. W. Schafer, "Homomorphic analysis of speech," *IEEE Trans. Audio Electroacoust.*, vol. AU-16, pp. 221–226, June 1968.

[52] J. Prabhakar and S. C. Gupta, "Separation of Rayleigh and Poisson density functions through homomorphic filtering," *Nat. Elec. Conf.*, pp. 605–610, Dec. 1970.

[53] R. Schafer and L. R. Rabiner, "System for automatic formant analysis of voiced speech," *J. Acoust. Soc. Amer.*, vol. 47, (Pt. 2), pp. 634–648, Feb. 1970.

[54] S. Senmoto and D. G. Childers, "Decomposition of a composite signal of unknown wavelets in noise," in *Int. Conf. Comm.* (ICC), 71C 28-COM, Montreal, Canada, 1971, pp. 5-14–5-19.

[55] J. Tenold, D. H. Crowell, R. H. Jones, T. H. Daniel, D. F. McPherson, A. N. Popper, "Cepstral and stationarity analysis of full-term and premature infants' cries," *J. Acoust. Soc. Amer.*, vol. 56, no. 3, pp. 975–980, Sept. 1974.

[56] J. M. Tribolet, "A new phase unwrapping algorithm," submitted to *IEEE Trans. Acoust., Speech*, and *Signal Processing*.

[57] J. M. Tribolet, A. V. Oppenheim, and G. E. Kopec, "Deconvolution by homomorphic prediction," submitted to *Geophysics*.

[58] T. Ulrych, "Application of homomorphic deconvolution to seismology," *Geophysics*, vol. 36, no. 4, pp. 650–660, Aug. 1971.

[59] T. Ulrych, O. G. Jensen, R. M. Ellis, and P. G. Sommerville, "Homomorphic deconvolution of some teleseismic events," *Bull. Seismological Soc. Amer.*, vol. 62, no. 5, pp. 1253–1265, Mar. 1972.

Extensions

[60] D. J. H. Moore and D. J. Parker, "On nonlinear filters involving transformation of the time variable," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 415–422, July 1973.

Books–Homomorphic Systems, Cepstrum Theory, and Applications

[61] B. Gold and C. M. Rader, *Digital Processing of Signals*. New York: McGraw-Hill, 1969.

[62] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[63] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[64] D. G. Childers and A. E. Durling, *Digital Filtering and Signal Processing*. St. Paul, MN: West Publishing, 1975.

[65] J. L. Flanagan, *Speech Analysis, Synthesis, and Perception*, 2nd Ed. New York: Springer-Verlag, 1972.

Linear Prediction, Predictive Deconvolution, Inverse Filtering, and Deconvolution

[66] B. S. Atal, "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification," *J. Acoust. Soc. Amer*, vol. 55, pp. 1304–1312, June 1974.

[67] D. G. Childers, R. S. Varga, and N. W. Perry, Jr., "Composite signal decomposition," *IEEE Trans. Audio Electroacoust*, vol. AU-18, pp. 471–477, Dec. 1970.

[68] M. P. Ekstrom, "A spectral characterization of the ill-conditioning in numerical deconvolution," *IEEE Trans. Audio Electroacoust.* vol. AU-21, pp. 344–348, Aug. 1973.

[69] B. R. Hunt, "A theorem on the difficulty of numerical deconvolution," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 94–95, Mar. 1972.

[70] ——, "Deconvolution of linear systems by constrained regression and its relationship to the Wiener theory," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 703–705, Oct. 1972.

[71] J. Makhoul, "Linear prediction: a tutorial review," *Proc. IEEE*,

vol. 63, pp. 561-580, Apr. 1975.

[72] B. Mitchell, M. Landisman, and Z. A. Der, "Predictive deconvolution applied to long range seismic refraction observations," *Pure Appl. Geophysics*, vol. 96, pp. 127-133, 1972.

[73] K. L. Peacock and S. Treitel, "Predictive deconvolution: theory and practice," *Geophysics*, vol. 34, pp. 155-169, Apr. 1969.

[74] E. A. Robinson, "Predictive decomposition of seismic traces," *Geophysics*, vol. 22, pp. 767-778, 1957.

[75] ——, "Properties of the Wold decomposition of stationary stochastic processes," *Theory Prob. Appl.*, vol. 8, pp. 187-194, 1963.

[76] ——, "Mathematical development of discrete filters for the detection of nuclear explosions," *J. Geophysics Res.*, vol. 68, pp. 5559-5567, 1963.

[77] ——, "Multichannel z-transforms and minimum delay," *Geophysics*, vol. 31, pp. 482-500, June 1966.

[78] ——, "Predictive decomposition of time series with application to seismic exploration," *Geophysics*, vol. 32, pp. 418-484, June 1967.

[79] J. Schell, "Dereverberation by linear systems techniques," *IEEE Trans. Geosci. Electron.*, vol. GE-9, no. 1, pp. 28-34, Jan. 1971.

[80] L. C. Wood and S. Treitel, "Seismic signal processing," *Proc. IEEE*, vol. 63, pp. 649-661, Apr. 1975.

Books—Computational Seismology

[81] M. Bath, *Mathematical Aspects of Seismology*. Amsterdam: Elsevier, 1968.

[82] *Computational Seismology*, V. I. Keilis-Borok, Ed. New York: Consultants Bureau, 1972.

[83] E. A. Robinson, *Multichannel Time Series Analysis with Digital Computer Programs*. San Francisco: Holden-Day, 1967.

[84] ——, *Statistical Communication and Detection with Special Reference to Digital Data Processing of Radar and Seismic Signals*. New York: Hafner, 1967.

[85] E. A. Robinson and S. Treitel, *Robinson-Treitel Reader*, 3rd Ed. Tulsa, OK: Seismographic Service Corp., 1973.

[86] *Seismic Filtering*, R. Van Nostrand, Ed. Tulsa, OK: Society of Exploration Geophysicists, 1971.

FILSYN—A General Purpose Filter Synthesis Program

GEORGE SZENTIRMAI, FELLOW, IEEE

Abstract—A very general computer program is described that can be used for the synthesis of passive *LC*, active *RC*, and (infinite impulse response) digital filters. Although it operates in both batch and interactive modes, this discussion deals exclusively with the interactive mode, which is somewhat more general and very easy to use. Apart from offering superior accuracy and flexibility, this program offers many firsts, including the passive realization of complex quadruplets of transmission zeros, the simultaneous realization of two transmission zero pairs, the active *RC* leapfrog realization, and many others.

I. INTRODUCTION

VERY CLOSE to ten years ago the author had reported on a computer program package that was developed for the synthesis of passive *LC* ladder filters [1]. This paper is about a vastly expanded and completely rewritten program for the interactive synthesis of filters of all kinds: passive *LC*, active *RC*, or digital.

One must justify the amount of space devoted to such a project, not to mention the effort expended in developing it, but this is easily done in this instance.

First of all, ten years is more than the average lifetime of computer programs and our previous effort is no exception. Although it has been converted to Fortran IV and is still in use, it has serious shortcomings. Some of these come from the fact that our philosophy of computer usage has changed tremendously during the past ten years. At the time we were fighting the reluctance of design engineers to use the computer and therefore were making a very strenuous effort to make the use of the program as easy and painless as possible. We in fact,

were attempting to present a package of foolproof programs that could be used without any knowledge of filter design theory and still yield meaningful results. Great emphasis was hence placed on the computer second guessing the user and supplying answers to questions the designer should have answered in the first place. This philosophy necessarily led to considerable restrictions.

Today we should perhaps try to keep people off the computers (that should at least improve turnaround time) and the author himself came around to the point of view that no matter how sophisticated a computer program is, users must at least have a reasonable idea as to what is going on in the design process to be able to utilize the program to its fullest. Such users would be more knowledgeable, but they would also appreciate the greater flexibility available, coupled with the possibility of correcting wrong choices that is available in an interactive operation. In keeping with this philosophy, the new version to be described below needs more understanding of the theory behind the program and hence a user who can make decisions only a designer can make, but in return it allows him a much greater flexibility in filter types, configurations, and other options. At the same time, to help the inexperienced user to become an experienced one, we retained most of the computer aids and added some new ones to help in this decision making process and sometimes take it over completely.

Another drawback of the old programs was found to be still numerical in the case of very high degree filters. The new version combines the two methods [2], [3] previously suggested as remedies and uses the transformed variable in the product form for most of the computations. Numerical problems still exist, but only where they do no harm.

One may still ask: why worry about passive *LC* filters? Of course they are still around (and will remain so for some time