Signal Reconstruction from Phase or Magnitude

MONSON H. HAYES, JAE S. LIM, AND ALAN V. OPPENHEIM, FELLOW, IEEE

Abstract-In this paper, we develop a set of conditions under which a sequence is uniquely specified by the phase or samples of the phase of its Fourier transform, and a similar set of conditions under which a sequence is uniquely specified by the magnitude of its Fourier transform. These conditions are distinctly different from the minimum or maximum phase conditions, and are applicable to both one-dimensional and multidimensional sequences. Under the specified conditions, we also develop several algorithms which may be used to reconstruct a sequence from its phase or magnitude.

I. INTRODUCTION

FOR both continuous-time and discrete-time signals, the magnitude and phase of the Fourier transform are, in general, independent functions, i.e., the signal cannot be recovered from knowledge of either alone. Under certain conditions, however, relationships exist between these components. For example, when the signal is minimum phase or maximum phase, the log magnitude and phase are related through the Hilbert transform. This relationship has been exploited in a variety of ways in many fields including network theory, communications, and signal processing [1]-[3].

In this paper we develop a set of conditions under which a discrete-time sequence is completely specified to within a scale factor by the phase of its Fourier transform, without the restriction of minimum or maximum phase, and propose several algorithms for implementing the reconstruction of a signal from the phase of its Fourier transform. In Section II we consider the case in which the phase is specified at all frequencies, and in Section III the case in which the phase is specified at a discrete set of frequencies. Algorithms for implementing the reconstruction are developed in Section V. In Section IV, we develop a different set of conditions, again without the restriction of minimum or maximum phase, in which a discrete-time sequence is completely specified by the magnitude of its Fourier transform. In Section VI, we extend the results of Sections II, III, IV, and V to the case of multidimensional sequences.

II. UNIQUENESS OF A SEQUENCE WITH A PHASE FUNCTION SPECIFIED AT ALL FREQUENCIES

The sequences that we consider are real with rational z-transforms. Since we are interested in conditions under which the

Manuscript received September 14, 1979; revised May 30, 1980. This work was sponsored by the Department of the Air Force. The views and conclusions contained in this document are those of the contractor and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government.

M. H. Hayes and J. S. Lim are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

A. V. Oppenheim is with the M.I.T. Lincoln Laboratory, Lexington, MA 02173, and also with the Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, MA 02139.

sequence can be uniquely specified by the phase of its Fourier transform, the Fourier transform is assumed to converge, i.e., the region of convergence of the z-transform includes the unit circle.

For such sequences, we first show that a finite length sequence is uniquely specified by the phase of its Fourier transform if its z-transform has no zeros in reciprocal pairs or on the unit circle.¹ More specifically, denoting the phase of the Fourier transform of x[n] and y[n] by $\theta_x(\omega)$ and $\theta_y(\omega)$, respectively, we have the following.

Theorem 1: Let x[n] and y[n] be two finite length sequences whose z-transforms have no zeros in reciprocal pairs or on the unit circle. If $\theta_x(\omega) = \theta_y(\omega)$ for all ω , then $x[n] = \beta y[n]$ for some positive constant β . If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$ for all ω , then $x[n] = \beta y[n]$ for some real constant β .

To demonstrate the validity of Theorem 1, we note first of all that if a finite-length sequence g[n] with z-transform G(z)has a phase which is zero or π for all ω , then g[n] is an even sequence, i.e., g[n] = g[-n]. Consequently, G(z) has zeros in conjugate reciprocal pairs so that if G(z) has a zero at $z = z_0$, then G(z) must also have a zero at $z = 1/z_0^*$. Now assume that x[n] and y[n] both satisfy the conditions of Theorem 1 and define g[n] as

$$g[n] = x[n] * y[-n]$$
(1)

so that

$$G(z) = X(z) Y(1/z).$$
 (2)

If $\theta_x(\omega) = \theta_y(\omega)$ or if $\tan \theta_x(\omega) = \tan \theta_y(\omega)$, then the phase of g[n] is zero or π . Therefore, g[n] is an even sequence and thus the zeros of G(z) occur in conjugate reciprocal pairs. Since the zeros of G(z) are collectively the zeros of X(z) and Y(1/z), if $X(z_0) = 0$, then either $X(1/z_0) = 0$ or $Y(z_0) = 0$. However, because the conditions of Theorem 1 exclude reciprocal zeros or zeros on the unit circle, $X(z_0)$ and $X(1/z_0)$ cannot both be zero. Thus, if $X(z_0) = 0$, then $Y(z_0) = 0$ and vice versa, i.e., the zeros of X(z) and Y(z) are identical. Consequently, since g[n] is an all zero sequence,²

$$X(z) = \beta Y(z) \tag{3a}$$

or

$$x[n] = \beta y[n]. \tag{3b}$$

¹Since we are considering only sequences which are real, zeros occur in complex conjugate pairs. In the following discussions, this symmetry is implicitly assumed, particularly in reference to reciprocal zeros.

²When we refer to a sequence as an all-zero (all-pole) sequence, this should be interpreted to mean that the z-transform has only zeros (poles) except possibly at z = 0 or $z = \infty$.

Combining (1) and (3b), we have

$$g[n] = \beta x[n] * x[-n].$$
 (4)

Since the phase of x[n] * x[-n] is always zero, if $\theta_x(\omega) = \theta_y(\omega)$, then the phase of g[n] is zero so β must be a positive constant. If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$, then the phase of g[n] is zero or π so β must be real.

An interpretation of Theorem 1 is suggested by the observation that for a rational z-transform, in general, a zero at $z = z_0$ and a pole at $z = 1/z_0^*$ contribute the same phase but different magnitude to the Fourier transform. Thus, with phase information alone, there is an inherent ambiguity in the z-transform in the sense that a zero (pole) at $z = z_0$ associated with the original sequence can potentially only be identified from the given phase as either a zero (pole) at $z = z_0$ or a pole (zero) at $z = 1/z_0^*$ and this ambiguity cannot be further resolved without additional information. The finite length condition in Theorem 1 resolves this ambiguity by restricting the z-transform to have only zeros except possibly at z = 0 or at $z = \infty$. The additional condition that the z-transform has no zeros in reciprocal pairs eliminates the possibility of zero phase components in the z-transfrom which, of course, could never be recovered from phase information alone. The conditions in Theorem 1 also eliminate the possibility of zeros on the unit circle. While the theorem can be modified to allow for the possibility of zeros on the unit circle, the result becomes somewhat more complicated and we have chosen not to include this additional generality.

Although Theorem 1 requires that x[n] be an all-zero sequence, a dual to Theorem 1 can be formulated for an allpole sequence. Specifically, let $\tilde{x}[n]$ denote the convolutional inverse of a sequence x[n], i.e.,

$$x[n] * \widetilde{x}[n] = \delta[n].$$
⁽⁵⁾

Then we have the following.

Theorem 2: Let x[n] and y[n] be two sequences whose z-transforms have no poles in reciprocal pairs, and which have finite duration convolutional inverses. If $\theta_x(\omega) = \theta_y(\omega)$ for all ω , then $x[n] = \beta y[n]$ for some positive constant β . If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$ for all ω , then $x[n] = \beta y[n]$ for some real constant β .

Theorem 2 follows directly from Theorem 1. Since the phase of the Fourier transform of $\tilde{x}[n]$ is the negative of the phase of the Fourier transform of x[n], $\tilde{x}[n]$ is uniquely specified to within a positive scale factor by the phase of the Fourier transform of x[n], by virtue of Theorem 1. Since x[n] is uniquely specified by $\tilde{x}[n]$, Theorem 2 follows.

In Section IV we will consider a number of numerical algorithms which can be implemented on a digital computer for reconstructing a sequence from its phase under the conditions of Theorem 1 or Theorem 2. At this point, however, we discuss a conceptual algorithm which may potentially have a practical implementation but which, more importantly, serves to lend insight into Theorems 1 and 2. We outline the algorithm under the conditions of Theorem 1 since it is easily modified for the conditions of Theorem 2.

Let $\theta_x(\omega)$ denote the specified phase function from which the sequence is to be reconstructed and $\hat{\theta}_x(\omega)$ the associated unwrapped phase [3]. From the conditions of Theorem 1, X(z) is restricted to be of the form

$$X(z) = C z^{n_0} \prod_{k=1}^{N_1} (1 - a_k z^{-1}) \prod_{k=1}^{N_2} (1 - b_k z)$$
(6)

with C real, n_0 an integer, $|a_k| < 1$, $|b_k| < 1$ for all k, and $a_k \neq b_l^*$ for any k and l.

Step 1: The algebraic sign of C is obtained from $\theta_x(\omega)$ using the fact that $\theta_x(0)$ is zero if and only if C is positive [3]. The value of n_0 in (6) is obtained from the unwrapped phase as

$$n_{0} = \frac{1}{\pi} [\hat{\theta}_{x}(\pi) - \hat{\theta}_{x}(0)].$$
⁽⁷⁾

Step 2: From the unwrapped phase function and the value of n_0 obtained in Step 1, a new phase function is specified as

$$\varphi_{\mathbf{x}}(\omega) \triangleq \widehat{\theta}_{\mathbf{x}}(\omega) - n_0 \omega - \widehat{\theta}_{\mathbf{x}}(0). \tag{8}$$

Using the Hilbert transform, a minimum phase sequence $x_{\min}[n]$ can be specified which has the phase $\varphi_x(\omega)$. The z-transform $X_{\min}(z)$ of $x_{\min}[n]$ is given by [3], [4]

$$X_{\min}(z) = \frac{\prod_{k=1}^{N_1} (1 - a_k z^{-1})}{\prod_{k=1}^{N_2} (1 - b_k^* z^{-1})}$$
(9)

where the coefficients a_k and b_k are identical to those in (6). Since pole-zero cancellations cannot occur in (9) by virtue of the condition in Theorem 1 which implies that $a_k \neq b_l^*$ for any k or l, the coefficients a_k in (6) can be obtained from the zeros of $X_{\min}(z)$ and the coefficients b_k^* , and thus b_k in (6) can be obtained from the poles of $X_{\min}(z)$.

The condition in Theorem 1 that there are no zeros in reciprocal pairs ensures that there are no pole-zero cancellations in (9). If the original sequence has reciprocal zeros, then the algorithm above may still be applied to recover all but those zeros in X(z) which are in reciprocal pairs.

III. UNIQUENESS OF A SEQUENCE WITH A PHASE FUNCTION SPECIFIED AT DISCRETE FREQUENCIES

In Theorems 1 and 2, we assumed that the phase function was specified at all frequencies. A similar set of theorems can be stated if the phase is specified at a sufficient number of discrete frequencies. As in Section II, we assume the sequences are real with rational z-transforms with a region of convergence that includes the unit circle. Then, an extension of Theorem 1 to discrete phase samples is given by the following.

Theorem 3: Let x[n] and y[n] be two finite length sequences which are zero outside the interval³ $0 \le n \le N-1$ with z-transforms which have no zeros in reciprocal pairs or on the unit circle. If $\theta_x(\omega) = \theta_y(\omega)$ at (N-1) distinct frequencies in the interval $0 \le \omega \le \pi$, then $x[n] = \beta y[n]$ for

³More generally, x[n] need only be zero outside any finite interval of length N. This added generality, however, is not considered in order to simplify the following discussions.

some positive constant β . If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some real constant β .

The validity of Theorem 3 follows in an almost identical manner to that of Theorem 1. Specifically, consider a finite length sequence g[n] for which g[n] = 0 outside the interval $(-N+1) \le n \le (N-1)$. Let $G(\omega) = |G(\omega)| e^{j\theta g(\omega)}$ denote the Fourier transform of g[n] with $\theta_g(\omega)$ zero or π at N-1 distinct frequencies $\omega_1, \omega_2, \cdots, \omega_{N-1}$ between zero and π , i.e.,

$$\theta_g(\omega_k) = 0 \quad \text{or} \quad \pi \quad \text{for} \quad k = 1, \cdots, N-1$$
 (10a)

with

$$\omega_k \neq \omega_l \quad \text{for } k \neq l \tag{10b}$$

and

$$0 < \omega_k < \pi \quad \text{for} \quad k = 1, \cdots, N-1. \tag{10c}$$

Then, $G(\omega_k)$ is real and

Im
$$[G(\omega_k)] = \sum_{n=-N+1}^{N-1} g[n] \sin n\omega_k = 0.$$
 (11)

Or, equivalently

$$\sum_{n=1}^{N-1} \{g[n] - g[-n]\} \sin n\omega_k = 0 \quad k = 1, 2, \cdots, N-1.$$
(12)

Equation (12) implies that [5]

$$\{g[n] - g[-n]\} = 0 \qquad n = 1, \cdots, N-1, \tag{13}$$

i.e., g[n] is an even sequence. Now, consider two sequences x[n] and y[n] satisfying the conditions of Theorem 3 and having the same phase at (N-1) distinct frequencies between zero and π . As with Theorem 1, we form the sequence

$$g[n] = x[n] * y[-n].$$
(14)

Since g[n] = 0 outside the interval $(-N+1) \le n \le (N-1)$ and $\theta_g(\omega_k)$ satisfies (10), g[n] is an even sequence. For reasons identical to those used in justifying Theorem 1, it then follows that

$$x[n] = \beta y[n] \tag{15}$$

where β is a positive constant if the phase samples of x[n] and y[n] are equal and a real constant if tangents of the phase samples are equal.

Although Theorem 3 requires that x[n] be an all-zero sequence, a dual to Theorem 3 for an all-pole sequence is easily formulated in terms of the convolutional inverse. Specifically, we have the following.

Theorem 4: Let x[n] and y[n] be two sequences whose z-transforms have no poles in reciprocal pairs, and which have convolutional inverses that are zero outside the interval $0 \le n \le N-1$. If $\theta_x(\omega) = \theta_y(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some positive constant β . If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some positive constant β . If $\tan \theta_x(\omega) = \tan \theta_y(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some real constant β .

Theorem 4 follows from Theorem 3 in the same manner that Theorem 2 follows from Theorem 1.

It should be noted that Theorem 1 implies that any finite duration sequence, x[n], which has a z-transform with no zeros on the unit circle or in conjugate reciprocal pairs is uniquely specified to within a scale factor by its phase, $\theta_x(\omega)$. Although there are many finite duration sequences with phase $\theta_x(\omega)$ which are not related to x[n] by a scale factor, x[n] is the only sequence which satisfies the z-transform constraints of Theorem 1. A similar statement, of course, can be made in the context of Theorem 3.

In Section V, numerical algorithms are developed for reconstructing a sequence x[n] from its phase when the sequence satisfies the constraints of Theorem 1. Although x[n] is uniquely specified to within a scale factor by its phase, some additional knowledge of x[n] is assumed to guarantee that the sequence obtained from the algorithm satisfies the constraints of Theorem 1 and thus that it is the correct sequence. The additional information which is assumed is the location of the first nonzero point of x[n]. In this case, as we next show, if x[n] satisfies the constraints of Theorem 1 and if it is known that x[n] = 0 outside the interval $0 \le n \le N - 1$ with $x[0] \ne 0$, then scaled versions of x[n] are the only sequences which are zero outside the given interval and have the same phase (or phase samples) as x[n]. Specifically, we have the following.

Theorem 5: Let x[n] be a sequence which is zero outside the interval $0 \le n \le N-1$ with $x[0] \ne 0$ and which has a z-transform with no zeros in reciprocal pairs or on the unit circle. Let y[n] be any sequence which is zero outside the interval $0 \le n \le N-1$. If $\theta_y(\omega) = \theta_x(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $y[n] = \beta x[n]$ for some positive constant β . If $\tan \theta_y(\omega) = \tan \theta_x(\omega)$ at (N-1)distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some real constant β .

Note that, in contrast to Theorem 3, there are no constraints on the location of the zeros of y[n]. Thus, y[n] may be any finite duration sequence which is zero outside the interval $0 \le n \le N-1$.

To demonstrate the validity of Theorem 5, we first form the sequence g[n] = x[n] * y[-n]. As discussed previously, since g[n] = 0 outside the interval $(-N+1) \le n \le (N-1)$ and the phase of its Fourier transform satisfies (10), g[n] is an even sequence. Now let $N_1 - 1$ represent the location of the last nonzero point in x[n], i.e., x[n] = 0 for $n \ge N_1$ and $x[N_1 - 1] \ne 0$. Then

$$G(z) = X(z) Y(z^{-1}) = \sum_{n=0}^{N_1 - 1} x[n] z^{-n} \sum_{n=0}^{N-1} y[n] z^n.$$
(16)

Since g[n] is even and $x[0] \neq 0$, y[n] = 0 for $n \ge N_1$ so that the number of zeros of y[n] is less than or equal to the number of zeros of x[n]. Now, for reasons identical to those used in justifying Theorem 1, if g[n] is an even sequence and if x[n] has no zeros in reciprocal pairs, then for each zero of x[n], y[n] must also have the same zero. Even though y[n] is not restricted to the class of sequences with no zeros in reciprocal pairs or on the unit circle, from our previous result, y[n]cannot have more zeros than x[n], and therefore, $y[n] = \beta x[n]$. For reasons identical to those used in justifying Theorem 1, β is a positive constant if the phase samples of x[n] and y[n] are equal whereas β is a real constant if the tangents of the phase samples are equal.

Although Theorem 5 requires that the sequence be an allzero sequence, a dual to Theorem 5 for an all-pole sequence is easily formulated in terms of the convolutional inverse. Specifically, we have the following.

Theorem 6: Let x[n] be a sequence whose z-transform has no poles in reciprocal pairs or on the unit circle, and whose convolutional inverse is zero outside the interval $0 \le n \le N-1$ and nonzero at n = 0. Let y[n] be any sequence whose convolutional inverse is zero outside the interval $0 \le n \le N-1$. If $\theta_y(\omega) = \theta_x(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some positive constant β . If $\tan \theta_y(\omega) = \tan \theta_x(\omega)$ at (N-1) distinct frequencies in the interval $0 < \omega < \pi$, then $x[n] = \beta y[n]$ for some real constant β .

Theorem 6 follows from Theorem 5 in the same manner that Theorem 2 follows from Theorem 1. Again note that, in contrast to Theorem 4, there are no constraints on the location of the poles of y[n].

IV. Uniqueness of a Sequence with a Specified Magnitude Function

In Section II, several sets of conditions are presented which establish a uniqueness between a sequence and its phase function. Unlike the case for minimum or maximum phase sequences, there is no dual statement of uniqueness between a sequence and its magnitude function under the same set of conditions. However, under a different set of conditions a sequence is uniquely specified to within a sign and a time shift by the magnitude of the Fourier transform. The conditions are embodied in the following theorems.

Theorem 7: Let x[n] and y[n] be two sequences whose z-transforms contain no reciprocal pole-zero pairs and which have all poles, not at $z = \infty$, inside the unit circle and all zeros, not at z = 0, outside the unit circle. If the magnitudes of the Fourier transforms of x[n] and y[n] are equal, then $x[n] = \pm y[n+m]$ for some integer m.

A dual to this theorem is the following.

Theorem 8: Let x[n] and y[n] be two sequences whose z-transforms contain no reciprocal pole-zero pairs and which have all poles, not at z = 0, outside the unit circle and all zeros, not at $z = \infty$, inside the unit circle. If the magnitudes of the Fourier transforms of x[n] and y[n] are equal, then $x[n] = \pm y[n+m]$ for some integer m.

Since the justification of Theorem 7 is almost identical to that of Theorem 8, we will focus only on the first. The validity of Theorem 7 is suggested by noting that a zero (pole) at $z = z_0$ and a zero (pole) at $z = 1/z_0^*$ contribute the same magnitude to the Fourier transform. Therefore, with magnitude information alone, there is an inherent ambiguity in the specification of the sequence in that a zero (pole) of the original sequence can potentially only be identified from the magnitude as either a zero (pole) at $z = z_0$ or at $z = 1/z_0^*$. In Theorem 7, this ambiguity is resolved by restricting the poles to lie inside the unit circle and the zeros to lie outside while in Theorem 8, the condition is the reverse. The additional condition that there are no conjugate reciprocal pole-zero pairs eliminates the possibility of all-pass terms which contribute only to the phase and not the magnitude.

To more formally demonstrate Theorem 7, let x[n] and y[n] be two sequences which satisfy the conditions of Theorem 7. The z-transforms of x[n] and y[n] may therefore be expressed in the form

$$P(z) = Cz^{n_0} \frac{\prod_{k=1}^{N_1} (1 - a_k z)}{\prod_{k=1}^{P_1} (1 - b_k z^{-1})}$$
(17)

with $a_k \neq b_l^*$ for any k and l, and where C is a real constant, n_0 is an integer, and $|a_k| < 1$, $|b_k| < 1$ for all k. Denoting the z-transforms of x[n] and y[n] by X(z) and Y(z), respectively, we wish to show that if X(z) and Y(z) both have the same magnitude on the unit circle then $x[n] = \pm y[n + m]$ for some integer m. Consider G(z) defined as the ratio X(z)/Y(z). Since X(z) and Y(z) both have the same magnitude on the unit circle, G(z) must be entirely all-pass with unity magnitude, i.e., for a zero at $z = z_0$ there must be a pole at $z = 1/z_0^*$ and vice versa. Therefore, G(z) consists only of poles and/or zeros at z = 0 or at $z = \infty$ and conjugate reciprocal pole-zero pairs. Because of the conditions in Theorem 7, this in turn requires that for any zero (or pole) of X(z) at $z = z_0$ there must be a zero (or pole) of Y(z) at $z = 1/z_0^*$ which, for $z_0 \neq 0$ or ∞ , violates the conditions in Theorem 7 since one will always be inside and the other outside the unit circle. Thus, G(z) must be of the form

$$G(z) = \pm z^m \tag{18}$$

or, equivalently,

$$x[n] = \pm y[n+m] \tag{19}$$

for some integer m.

A conceptual algorithm similar to that considered in Section II can be developed for the reconstruction of a sequence to within an algebraic sign and a delay from the magnitude of its Fourier transform under the conditions of Theorem 7 or Theorem 8. We outline the procedure below under the conditions of Theorem 7. It is easily modified for the conditions of Theorem 8.

Let $|X(\omega)|$ denote the specified magnitude function. Using the Hilbert transform, a minimum phase sequence $x_{\min}[n]$ can be specified which has the same magnitude function. The z-transform $X_{\min}(z)$ of $x_{\min}[n]$ is given by

$$X_{\min}(z) = |C| \frac{\prod_{k=1}^{N_1} (1 - a_k^* z^{-1})}{\prod_{k=1}^{P_1} (1 - b_k z^{-1})},$$
(20)

i.e., it has the same poles as X(z) and the zeros are reflected inside the unit circle. Since the conditions of Theorem 7 exclude the possibility of pole-zero cancellation, the coefficients a_k^* , and thus a_k in (17), can be obtained from the zeros of $X_{\min}(z)$ and the coefficients b_k in (17) can be obtained from the poles of $X_{\min}(z)$.

V. NUMERICAL ALGORITHMS FOR RECONSTRUCTION FROM SAMPLES OF A PHASE FUNCTION

In Section II, we presented two sets of conditions, embodied in Theorems 1 and 2, under which a sequence is uniquely specified to within a positive scale factor by the phase of its Fourier transform. In this section, we describe two numerical algorithms which can be used to reconstruct a sequence satisfying the requirements of Theorem 1 from samples of its phase function when the location of the first nonzero point of x[n]and the interval outside of which x[n] is zero are known. Although these algorithms will only be discussed in terms of reconstructing sequences satisfying the conditions of Theorem 1, the reconstruction of sequences meeting the requirements of Theorem 2 may be accomplished by simply reconstructing the finite length sequence $\tilde{x}[n]$ defined in (5) using the negative of the specified phase samples and then computing the convolutional inverse sequence.

The first algorithm presented below is an iterative technique in which the estimate of x[n] is improved in each iteration. This algorithm is similar to the iterative algorithms developed by Gerchberg and Saxton [6] and Fienup [7] for reconstructing a signal from magnitude information and to the iterative algorithm developed by Quatieri [8] for reconstructing a signal from its phase under the assumption that the signal is minimum phase. The second algorithm is a closed form solution which is obtained by solving a set of linear equations. Under the conditions specified in Theorem 1, this algorithm provides the desired sequence x[n] to within a scale factor when the location of the first nonzero point of x[n] and the interval outside of which x[n] is zero are known.

In the discussions which follow, x[n] is used to denote a sequence which satisfies the conditions of Theorem 1 and is zero outside the interval $0 \le n \le N-1$ with $x[0] \ne 0$. In the more general case (see footnote 3), a linear phase term may be added to the given phase to accomplish this.

A. Iterative Algorithm

The *M*-point discrete Fourier transform (DFT) of x[n] will be denoted as

$$X(k) = |X(k)| e^{j\theta_X(k)}$$
(21)

where it is assumed that $M \ge 2N$. Then, an iterative technique to reconstruct the sequence x[n] from the *M* samples of its phase $\theta_x(k)$, $k = 0, 1, \dots, M-1$, as illustrated in Fig. 1 and may be described as follows.

Step 1: We begin with $|X_0(k)|$, an initial guess of the unknown DFT magnitude and form the first estimate, $X_1(k)$, of X(k) using the specified phase function, i.e.,

$$X_1(k) = |X_0(k)| e^{j\theta_{\mathcal{X}}(k)}.$$
 (22)

Computing the inverse DFT of $X_1(k)$ provides the first estimate, $x_1[n]$, of x[n]. Since an *M*-point DFT is used, $x_1[n]$ is an *M*-point sequence which is, in general, nonzero for $N \le n \le M - 1$.

Step 2: From $x_1[n]$, another sequence, $y_1[n]$, is defined by

$$y_1[n] = \begin{cases} x_1[n] & 0 \le n \le N - 1 \\ 0 & N \le n \le M - 1. \end{cases}$$
(23)

Step 3: The magnitude $|Y_1(k)|$ of the *M*-point DFT of $y_1[n]$ is then considered as a new estimate of |X(k)| and a



Fig. 1. Block diagram of the iterative algorithm for reconstructing a signal from its phase.

new estimate of X(k) is formed by

$$X_{2}(k) = |Y_{1}(k)| e^{j\theta_{X}(k)}.$$
(24)

From this, a new estimate $x_2[n]$ is obtained from the inverse DFT of $X_2(k)$. Repetitive application of Steps 2 and 3 defines the iteration.

In this iterative procedure, the total squared error between x[n] and its estimate is nonincreasing with each iteration. To see this, let $x_p[n]$ denote the estimate after the *p*th iteration and define the error E_p as

$$E_p = \sum_{n=0}^{M-1} \{x[n] - x_p[n]\}^2.$$
 (25)

From Parseval's theorem,

$$E_p = \frac{1}{M} \sum_{k=0}^{M-1} |X(k) - X_p(k)|^2.$$
⁽²⁶⁾

Since X(k) and $X_p(k)$ have the same phase, then

$$E_{p} = \frac{1}{M} \sum_{k=0}^{M-1} \left[\left| X(k) \right| - \left| X_{p}(k) \right| \right]^{2}$$
$$= \frac{1}{M} \sum_{k=0}^{M-1} \left[\left| X(k) \right| - \left| Y_{p-1}(k) \right| \right]^{2}.$$
(27)

FFT Length	Number of Iterations	x[0]	x[1]	x[2]	x[3]	x[4]	x[5]	x[6]	x[7]	Total Squared Error
. 16 ,	10	4.000	1.892	-9.935	3.331	5.113	7.115	13.886	4.961	11.961
	100	4.000	2.022	-10.335	3.615	4.845	6.679	14.123	5,379	7.050
	500	4.000	2.040	-10.804	4.456	4.272	5.577	14.642	-5,854	8.925.10 ⁻¹
	1000	4.000	2.012	-10.947	4.849	4.075	5.159	14.901	-5.961	6.792·10 ⁻²
128	10	4.000	1.647	-10.245	4.875	5.015	6.563	15.286	-4.640	6.117
	100	4.000	1.855	-10.671	4.902	4.459	5.723	15.104	-5,412	1.229
	500	4.000	1.996	-10.991	4.997	4.013	5.019	15.003	5.984	9.109·10 ⁻²
	1000	4.000	2.000	-11.000	5.000	4.000	5.000	15.000	-6.000	4.118.10 ⁻⁵
Original Sequence		4.0	2.0	-11.0	5.0	4.0	5.0	15.0	-6.0	

 TABLE I

 Iterative Reconstruction of a Sequence from Its Phase

Using the triangle inequality for vector differences, (27) becomes

$$E_{p} \leq \frac{1}{M} \sum_{k=0}^{M-1} |X(k) - Y_{p-1}(k)|^{2}$$
⁽²⁸⁾

with equality holding if $\theta_x(k) = \theta_{p-1}(k)$ where $\theta_{p-1}(k)$ is the phase of $Y_{p-1}(k)$. Again using Parseval's theorem, (28) becomes

$$E_{p} \leq \sum_{n=0}^{M-1} \{x[n] - y_{p-1}[n]\}^{2}.$$
(29)

Since

$$y_{p-1}[n] = x_{p-1}[n],$$
 for $0 \le n \le N - 1$
 $y_{p-1}[n] = x[n] = 0,$ for $N \le n \le M - 1$

then

$$\sum_{n=0}^{M-1} \{x[n] - y_{p-1}[n]\}^2 \leq \sum_{n=0}^{M-1} \{x[n] - x_{p-1}[n]\}^2 = E_{p-1}$$
(30)

with equality if and only if $y_{p-1}[n] = x_{p-1}[n]$. Therefore,

$$E_p \leqslant E_{p-1}.\tag{31}$$

Although (31) is not sufficient to guarantee the convergence of the algorithm, it has recently been proved [9] that this algorithm always converges if $M \ge 2N$ and if x[n] satisfies the constraints of Theorem 5. Consistent with this theoretical result, in all the examples that we have considered so far, we have empirically observed that the algorithm converges to the correct solution when $M \ge 2N$ even though the number of iterations required to achieve a small total squared error is, in general, quite large. We have also observed that increasing Mmay increase the rate of convergence of the algorithm, but such an increase obviously results in an increase in the number of computations required for each iteration. The development of techniques to increase the rate of convergence are currently under investigation and some preliminary results are reported in [10]. Two examples of the iterative procedure applied to a mixed phase sequence x[n] of length 8 are shown in Table I. In the first example, an FFT of length 16 was used. In the second example, the FFT length was extended to 128 points. In both cases, the initial guess of the unknown magnitude was chosen to be a constant, and the scaling factor β was chosen so that the resulting sequences have the same value at the origin as x[n]. The results after 10, 100, 500, and 1000 iterations are presented along with the values of the total squared error.

B. Closed Form Solution

A closed form solution⁴ for reconstructing a sequence x[n] from samples of its phase, $\theta_x(\omega)$, follows from the definition of $\theta_x(\omega)$. With $\theta_x(\omega)$ defined so that $-\pi \leq \theta_x(\omega) < \pi$, we have

$$\tan \theta_{\mathbf{x}}(\omega) = \frac{-\sum_{n=0}^{N-1} \mathbf{x}[n] \sin n\omega}{\sum_{n=0}^{N-1} \mathbf{x}[n] \cos n\omega}.$$
(32)

For the case in which $\theta_x(\omega) = \pm \pi/2$ so that $\tan \theta_x(\omega) = \pm \infty$, (32) is equivalent to

$$\sum_{n=0}^{N-1} x[n] \cos n\omega = 0.$$
 (33)

Sampling $\tan \theta_x(\omega)$ at N-1 distinct frequencies $\omega_1, \omega_2, \cdots, \omega_{N-1}$ with $0 < \omega_k < \pi$ for $k = 1, 2, \cdots, N-1$ and using a standard trigonometric identity, (32) and (33) can be written as

$$\sum_{n=1}^{N-1} x[n] \sin \left[\theta_x(\omega_k) + n\omega_k\right]$$
$$= -x[0] \sin \theta_x(\omega_k), \quad \text{if } \theta_x(\omega_k) \neq \pm \frac{\pi}{2}$$
(34a)

⁴A closed form solution similar to the one presented in this section can be obtained by expressing the real and imaginary parts of $X(\omega)$ in terms of the given phase and then relating the real and imaginary parts through the discrete Hilbert transform relations for causal sequences.

$$\sum_{n=1}^{N-1} x[n] \cos n\omega_k = -x[0], \quad \text{if } \theta_x(\omega_k) = \pm_2^{\pi} \qquad (34b)$$

for $k = 1, 2, \dots, N-1$. Equation (34) represents (N-1) linear equations in the (N-1) unknowns of x[n] which, when augmented by the equation x[0] = x[0], can be expressed in matrix form as

$$Sx = x[0] b \tag{35}$$

where x represents the vector of elements of x[n]. Any solution s[n] to (35) has the property that s[n] is zero outside the interval $0 \le n \le N-1$, has the correct value at n = 0, and has the same tangent of the phase as x[n] for N-1 distinct frequencies between zero and π . Thus, from Theorem 5, we conclude that s[n] = x[n]. Therefore, there is a unique solution to (35). This implies that S^{-1} , the inverse of the matrix S, exists and that x[n] is given by

$$\boldsymbol{x} = \boldsymbol{x}[0] \ \boldsymbol{S}^{-1}\boldsymbol{b}. \tag{36}$$

Since x[0] is not known, in general, (36) specifies x[n] to within an unknown scale factor x[0]. To specify x[n] to within a positive scale factor, we first assume that x[0] > 0 and determine the phase of x in (36). If the resulting phase does not differ from $\theta_x(\omega)$, then x[0] > 0; otherwise x[0] < 0.

In reconstructing x[n] from $\theta_x(\omega)$ using (36), it should be noted that we have some control over the matrix S. Since the elements of the matrix S are functions of the samples of $\theta_x(\omega)$, S can be changed by choosing a different set of frequency samples. This control over S may be useful, for example, in avoiding potential numerical instabilities in computing the inverse matrix S^{-1} .

Compared with the iterative algorithm, the closed form solution presented above has the advantage that the desired sequence is guaranteed to be the solution to (35) and, in addition, no iterations are required in order to reach a solution. On the other hand, (35) requires computing the inverse of an $(N-1) \times (N-1)$ matrix which may lead to numerical problems and severe roundoff errors, particularly as N becomes large.

The algorithm discussed above has been applied to a variety of different examples. Consistent with our theoretical results, in all examples we have considered, the desired solutions have been obtained. Specifically, for the sequence shown in Table I, the closed form solution was used to reconstruct the sequence from its phase. The phase samples used to define the matrix S were first chosen to be equally spaced between zero and π , and then were randomly selected between zero and π . Within the limits of finite precision arithmetic, in both cases the sequence was reconstructed exactly when the scaling factor was chosen so that the solutions obtained had the correct value at the origin.

Finally, it should be noted that if the first nonzero point of x[n] is not at n = 0 but rather at $n = n_0 > 0$, then the matrix S will be singular. This follows from the observation that x is a solution to the equation Sx = 0. However, it is straightforward to show that the location of the first nonzero point of x[n] is given by

$$n_0 = N - \operatorname{rank}(S). \tag{37}$$

This equation, therefore, suggests an algorithm which may be used to determine, from phase samples, the interval over which a sequence is nonzero when the sequence satisfies the constraints of Theorem 1.

VI. EXTENSION TO MULTIDIMENSIONAL SEQUENCES

In this section, we extend the results of Sections II-V to the case of multidimensional sequences. This extension is achieved by mapping a multidimensional sequence into a 1-D (one-dimensional) sequence and then applying the results for 1-D sequences. Since the extension of the 2-D case to sequences of higher dimension is straightforward, our discussions in this section will concentrate on the 2-D case. Again, we consider only sequences which are real, have rational z-transforms, and have Fourier transforms that converge.

Let $x[n_1, n_2]$ represent a 2-D sequence which has a rational z-transform with $X(\omega_1, \omega_2)$ given by

$$X(\omega_1, \omega_2) = \frac{A(\omega_1, \omega_2)}{B(\omega_1, \omega_2)}$$
(38)

where $A(\omega_1, \omega_2)$ is a 2-D polynomial of degree M_1 in exp $[j\omega_1]$ and M_2 in exp $[j\omega_2]$ and where $B(\omega_1, \omega_2)$ is a 2-D polynomial of degree N_1 in exp $[j\omega_1]$ and N_2 in exp $[j\omega_2]$. Suppose we form a 1-D sequence $\hat{x}_1[n]$ or $\hat{x}_2[n]$ by

$$\hat{x}_1[n] = \sum_{m=-\infty}^{\infty} x[m, n-Nm]$$
 (39a)

or

$$\hat{x}_{2}[n] = \sum_{m=-\infty}^{\infty} x[n - Mm, m]$$
 (39b)

where $M \ge \max(M_1, N_1)$ and $N \ge \max(M_2, N_2)$. Then it can be shown [11], [12] that the transformation in (39) is invertible and that

$$\hat{X}_{1}(\omega) = X(\omega_{1}, \omega_{2}) \Big|_{\omega_{1} = \omega N, \omega_{2} = \omega}$$

$$\tag{40a}$$

and

$$\hat{X}_{2}(\omega) = X(\omega_{1}, \omega_{2}) \big|_{\omega_{1} = \omega, \, \omega_{2} = \omega M}.$$
(40b)

From (40), it is clear that the phase of $\hat{x}_1[n]$ or $\hat{x}_2[n]$ is specified by the phase of $x[n_1, n_2]$ and the magnitude of $\hat{x}_1[n]$ or $\hat{x}_2[n]$ is specified by the magnitude of $x[n_1, n_2]$. Therefore, all the theorems and numerical algorithms developed in Sections II-V for 1-D sequences may be extended to 2-D sequences by first transforming them into 1-D sequences using (39) and then applying the 1-D results to the resulting 1-D sequences. Thus, for example, Theorem 1 may be extended to 2-D sequences as follows. Let $x[n_1, n_2]$ and $y[n_1, n_2]$ be two 2-D sequences such that the two 1-D sequences x[n] and y[n] obtained from transforming $x[n_1, n_2]$ and $y[n_1, n_2]$ using (39) are finite in length and have no zeros in reciprocal pairs or on the unit circle. If $\theta_x(\omega_1, \omega_2) =$ $\theta_y(\omega_1, \omega_2)$ for all ω_1 and ω_2 , then $x[n_1, n_2] = \beta y[n_1, n_2]$ for some positive constant β . If $\tan \theta_x(\omega_1, \omega_2) = \tan \theta_y(\omega_1, \omega_2)$ for all ω_1 and ω_2 , then $x[n_1, n_2] = \beta y[n_1, n_2]$ for some real constant β .

If $x[n_1, n_2]$ is a 2-D sequence with finite support, then the transformation given by (39) can be reduced to a simpler



Fig. 2. Image reconstruction from phase information.

form. Specifically, let $x[n_1, n_2]$ be zero outside the region $0 \le n_1 \le N_1 - 1$ and $0 \le n_2 \le N_2 - 1$. Then (39) can be rewritten as

$$\hat{x}_1[n_1N+n_2] = x[n_1, n_2],$$

with $N \ge N_1$ and $0 \le n_2 \le N-1$
(41a)

or

$$\hat{x}_2[n_1 + n_2M] = x[n_1, n_2],$$

with $M \ge N_2$ and $0 \le n_1 \le M - 1.$
(41b)

Clearly, the transformation in (41) is invertible and it can be easily shown [12] that $\hat{X}_1(\omega)$ and $\hat{X}_2(\omega)$ are given by (40).

As an illustration of the results of this section, a 2-D sequence representing the intensity of an image $x[n_1, n_2]$ was created which is zero outside the region $0 \le n_1 \le 11$ and $0 \le n_2 \le 11$. From the phase of $\hat{X}_1(\omega)$ defined by (40a) with N = 12, the closed form solution was used to reconstruct $x[n_1, n_2]$. With the scale factor chosen so that the reconstructed image had the same value at the origin as $x[n_1, n_2]$; the result, shown in Fig. 2, is indistinguishable from the original. For illustration, the image shown has been enlarged by means of a zero-order hold.

Finally, it should be noted that this approach of transforming n-D sequences into their 1-D projections provides only a partial solution to the multidimensional phase-only problem since it circumvents the fundamental issues involved in multidimensional phase-only signal reconstruction. Specifically, this approach imposes constraints on a 1-D projection of an n-D sequence rather than directly on the n-D sequence. In addition, although it may not be possible to perform a phaseonly reconstruction of an n-D sequence from a particular projection, this does not preclude the possibility that there exists another projection for which the reconstruction is possible. Therefore, with this approach, it is difficult to determine which multidimensional sequences may be reconstructed from their phase. It is possible, however, to generalize Theorem 1 to n-D sequences [13]. However, due to the absence of a fundamental theorem of algebra for polynomials in more than one variable, the proof of this theorem [14] is more abstract than that required in the one-dimensional case.

VII. CONCLUDING REMARKS

In this paper, we have developed a set of conditions under which a sequence is uniquely specified by the phase or samples of the phase of its Fourier transform, and a similar set of conditions under which a sequence is uniquely specified by the magnitude of its Fourier transform. Under the specified conditions, we have also developed several algorithms which may be used to reconstruct a sequence from its phase or magnitude.

Even though the results reported in this paper seem to answer some of the important questions on the general problem of signal reconstruction from its phase or magnitude, there are a variety of issues that remain to be investigated. One such issue pertains to the conditions under which a sequence can be recovered from its magnitude. The conditions that we have developed under which a sequence can be reconstructed from its phase are quite general and most finite extent sequences of practical interest satisfy these conditions. However, the conditions we have developed under which a sequence can be reconstructed from its magnitude appear to be very restrictive. For example, few sequences of practical interest have all the poles within the unit circle and all the zeros outside the unit circle. It may be possible, however, to relax these conditions by imposing other constraints which do not exclude sequences of practical interest. Another issue that requires further investigation is an understanding of the sensitivity of the reconstructed sequence to inaccurate information about the original (unknown) sequence. For example, in most practical problems of interest, the phase may not be known exactly due to errors such as measurement noise and it is important to understand the effects of these errors on the reconstructed sequence. These and other important issues are currently under investigation.

The results reported in this paper have the potential to be applicable to a variety of practical problems. For example, consider an image which is blurred by an optical system whose transfer function is not known and suppose that we wish to reduce the effect of the blurring. If the blurring function is symmetrical so that the tangent of the phase is unaffected by the blurring and if the original image satisfies the conditions developed in this paper for its unique reconstruction from the tangent of its phase, then the results of this paper are potentially applicable. As another example, consider the signal coding problem. In signal coding by Fourier transform techniques, both the magnitude and phase are typically coded and transmitted. Since the results of this paper show that most finite extent sequences can be recovered from samples of their phase, it is reasonable to attempt to code only the phase and then reconstruct the signal from the coded phase. These and other potential application areas are also currently under investigation.

REFERENCES

- [1] W. D. Gregg, Analog and Digital Communication. New York: Wiley, 1977.
- [2] E. Guillemin, Theory of Linear Physical System. New York: Wiley, 1963.
- [3] A. V. Oppenheim and R. W. Schafer, Digital Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1975.
- [4] R. W. Schafer, "Echo removal by discrete generalized linear filtering," Ph.D. dissertation, Dep. Elec. Eng., M.I.T., 1968.
- [5] S. Karlin and W. J. Studden, Tchebycheff Systems: With Applications in Analysis and Statistics. New York: Interscience-Wiley, 1966.
- [6] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik, vol. 35, pp. 237-246, 1972.
- [7] J. R. Fienup, "Reconstruction of an image from the modulus of its Fourier transform," Opt. Lett., vol. 3, pp. 27-29, July 1978. [8] T. Quatieri, "Phase estimation with applications to speech analy-
- sis-synthesis," Sc.D. dissertation, Dep. Elec. Eng., M.I.T., 1979.
- [9] V. T. Tom, T. F. Quatieri, M. H. Hayes, and J. H. McClellan, "Convergence of iterative nonexpansive signal reconstruction algorithms," Lincoln Lab., M.I.T., Tech. Note 1980-21, 1980.
- [10] A. V. Oppenheim, M. H. Hayes, and J. S. Lim, "Iterative procedures for signal reconstruction from phase," presented at the SPIE Conf., Washington, DC, Apr. 1980.
- [11] R. M. Mersereau, "The digital reconstruction of multi-dimen-sional signals from their projections," Sc.D. dissertation, Dep. Elec. Eng., M.I.T., 1973.

- [12] R. M. Mersereau and D. E. Dudgeon, "The representation of twodimensional sequences as one-dimensional sequences," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-22, 1974.
- M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Phase-only signal [13] reconstruction," presented at ICASSP Conf., Denver, CO, Apr. 1980.
- [14] M. H. Hayes, "On the uniqueness of a multi-dimensional sequence in terms of its phase or magnitude," Lincoln Lab., M.I.T., Tech. Note 1980-22, Feb. 26, 1980.



processing.

Monson H. Hayes was born in Washington, DC, on October 27, 1949. He received the B.S. degree in physics from the University of California, Berkeley, in 1971, and the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1978.

He worked as a Systems Engineer in IR systems technology at Aerojet Electrosystems from 1972-1974. He is currently pursuing the Ph.D. degree at M.I.T. His research interests

include digital signal processing and its applications to digital image

Jae S. Lim, for a photograph and biography, see p. 197 of the April 1980 issue of this TRANSACTIONS.



Alan V. Oppenheim (S'57-M'65-SM'71-F'77) received the S.B. and S.M. degrees in 1961 and the Sc.D. degree in 1964, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge.

From 1961 to 1964 he was a member of the M.I.T. Research Laboratory of Electronics and an Instructor in the Department of Electrical Engineering. During this period his research activities involved the application of modern algebra to the characterization of nonlinear

systems and the development of homomorphic filtering as an approach to some nonlinear filtering problems. In 1964 he joined the faculty in the Department of Electrical Engineering and Computer Science and became a Staff Member of the M.I.T. Research Laboratory of Electronics. From 1967 to 1969, he was a staff member of the M.I.T. Lincoln Laboratory, working in the areas of digital signal processing and speech communication. In 1969, he was appointed as an Associate Professor and, in 1976, was appointed as a full Professor, both in the Department of Electrical Engineering and Computer Science at M.I.T. During the academic year 1972-1973, on sabbatical leave, he was a Guggenheim Fellow at the University of Grenoble, Grenoble, France. From 1976 to 1978, he held the Cecil H. Green Distinguished Chair in Electrical Engineering and Computer Science. He is currently Professor of Electrical Engineering and Computer Science at M.I.T. and Associate Head of the Data Systems Division at M.I.T. Lincoln Laboratory. His research interests are in the area of digital signal processing and its applications to speech, image, and seismic data processing. He is coauthor of the text Digital Signal Processing and Editor of the book Applications of Digital Signal Processing. He is also coauthor of the text Signals and Systems, to be published.

Dr. Oppenheim is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. In 1969 he received the IEEE G-AE Senior Award for "authoring a paper of exceptional merit," and in 1976 he received the IEEE Acoustics, Speech, and Signal Processing Society Technical Achievement Award "in recognition of fundamental contributions to the development of homomorphic processing techniques for speech and other signals."