Signal Reconstruction from Signed Fourier Transform Magnitude

PATRICK L. VAN HOVE, MONSON H. HAYES, MEMBER, IEEE, JAE S. LIM, MEMBER, IEEE, AND ALAN V. OPPENHEIM, FELLOW, IEEE

Abstract—**In this paper, we show that a one-dimensional or multidimensional sequence is uniquely specified under mild restrictions by its signed Fourier transform magnitude (magnitude and 1 bit of phase information). In addition, we develop a numerical algorithm to reconstruct a one-dimensional or multidimensional sequence from its Fourier transform magnitude. Reconstruction examples obtained using this algorithm are also provided.**

I. Introduction

In a variety of contexts, such as electron microscopy [1], X-ray crystallography [2], optics [3], and Fourier transform signal coding [4], it is desirable to reconstruct a sequence from partial Fourier domain information. As a consequence, considerable attention has been paid to this area, and some significant results have been developed. It has been previously established [5]–[7] that under very mild restrictions a finite extent one-dimensional (1-D) or multidimensional (MD) sequence is uniquely specified to within a scale factor by the tangent of its Fourier transform (FT) phase, and algorithms for implementing the reconstruction have been developed. It is well known that, in contrast, the FT magnitude does not uniquely specify a 1-D sequence. For MD sequences, the FT magnitude specifies a sequence to within a translation, sign, and a central symmetry [7], [8], and reconstruction algorithms developed so far have been successful [7] for only a very restricted class of MD sequences.

From the above results, on the question of unique specification of a sequence, there appear to be significant differences between 1-D and MD sequences, and between the tangent of the FT phase and the FT magnitude. In addition, the tangent of the phase and the magnitude of a complex number, which have been considered in previous studies, do not completely specify the complex number. In this paper, we show that if the signed FT magnitude (magnitude and one bit of phase information) is considered rather than the FT magnitude, there are only minor differences on the question of unique specification of a sequence, between 1-D and MD sequences, and between the tangent of the FT phase and the signed FT magnitude. In particular, it is shown that under very mild restrictions, the signed FT magnitude is sufficient to uniquely specify a 1-D or MD sequence. We note that the tangent of the phase and the signed magnitude of a complex number completely specify the complex number.

In Section II of this paper, the basic theory is presented. In Section III an algorithm for implementing the reconstruction is discussed, and Section IV illustrates several examples.

II. Theory

In this section, we discuss the unique specification of a sequence by its FT magnitude and 1 bit of phase. We initially consider the one-dimensional (1-D) case and then extend the 1-D result to the multidimensional (MD) case. Before we present the theoretical results, we define the notation that will be used throughout the paper.

Let $x(n)$ denote a 1-D sequence which is *causal* and *finite extent* so that $x(n)$ is zero outside $0 \leqslant n \leqslant L-1$. Furthermore, we restrict $x(n)$ to be real-valued. Let $X(z)$ and $X(\omega)$ represent the z transform and Fourier transform of $x(n)$, so that

$$X(z) = \sum_{n=0}^{L-1} x(n) z^{-n} \tag{1}$$

$$X(\omega) = X(z)\Big|_{z=e^{j\omega}} = \sum_{n=0}^{L-1} x(n) e^{-j\omega n}. \tag{2}$$

The Fourier transform $X(\omega)$ can be represented in terms of its real part $X_R(\omega)$ and imaginary part $X_I(\omega)$, or in terms of its magnitude $|X(\omega)|$ and phase $\theta_x(\omega)$ as follows:

$$X(\omega) = X_R(\omega) + jX_I(\omega) = |X(\omega)| e^{j\theta_x(\omega)}. \tag{3}$$

To ensure that $\theta_x(\omega)$ is well defined at all ω, we assume that $X(z)$ has no zeros on the unit circle. The phase function $\theta_x(\omega)$ in (3) represents the principal value of the phase so that

$$-\pi < \theta_x(\omega) \leqslant \pi. \tag{4}$$

The 1-bit FT phase information will be represented by the function $S_x^\alpha(\omega)$ defined as

$$S_x^\alpha(\omega) = \begin{cases} +1 & \alpha - \pi \leqslant \theta_x(\omega) \leqslant \alpha \\ -1 & \text{otherwise} \end{cases} \tag{5}$$

Manuscript received August 24, 1982; revised March 23, 1983. This work was supported by the Advanced Research Projects Agency monitored by ONR under Contract N00014-81-K-0742 NR-049-506 and in part by the National Science Foundation under Grants ECS80-07102 and ECS82-04793.

P. L. Van Hove, J. S. Lim, and A. V. Oppenheim are with the Department of Electrical Engineering and Computer Science, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139.

M. H. Hayes was with the Department of Electrical Engineering and Computer Science, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139. He is now with the School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

0096-3518/83/1000-1286$01.00 © 1983 IEEE

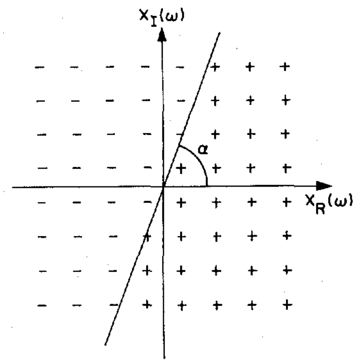


Fig. 1. Mapping of the 1-bit phase function.

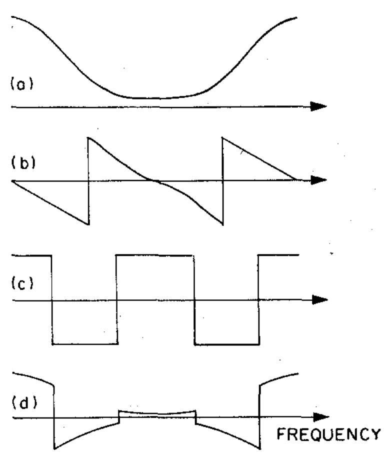


Fig. 2. Fourier transform magnitude, phase, 1-bit phase, and signed magnitude of the sequence $X(z) = 1 + 3z^{-1} + 5z^{-2} + 2z^{-3}$.

where α is a known constant in the range of $0 < \alpha \leqslant \pi$. Thus, the complex plane is divided into two regions separated by a straight line passing through the origin and at an angle α with the real axis, as shown in Fig. 1. For example, for $\alpha = \pi/2$, $S_x^{\pi/2}(\omega)$ represents the algebraic sign of $\text{Re}\{X(\omega)\}$. More generally, $S_x^{\alpha}(\omega)$ is the algebraic sign of $\text{Re}\{e^{j(\pi/2-\alpha)}X(\omega)\}$. The algebraic sign of zero is assumed to be positive.

The function $G_x^{\alpha}(\omega)$ is defined as

$$G_x^{\alpha}(\omega) = S_x^{\alpha}(\omega)|X(\omega)| \tag{6}$$

and will be referred to as the signed Fourier transform magnitude since it contains both magnitude and sign information. An example of $|X(\omega)|$, $\theta_x(\omega)$, $S_x^{\alpha}(\omega)$, and $G_x^{\alpha}(\omega)$ when $\alpha = \pi/2$ and $X(z) = 1 + 3z^{-1} + 5z^{-2} + 2z^{-3}$ is shown in Fig. 2.

Finally, given a positive integer N, we define a constant P and an interval R as

$$P = \frac{N-1}{2} \text{ and } R = (0, \pi) \text{ for } N \text{ odd}$$

$$P = \frac{N}{2} \text{ and } R = (0, \pi] \text{ for } N \text{ even.} \tag{7}$$

The uniqueness of a 1-D sequence when the signed Fourier transform magnitude $G_x^{\alpha}(\omega)$ is specified is based on the following statements. The proof of these statements is given in the Appendix.

Statement A1: Let $x(n)$ and $y(n)$ be two real, causal, and finite extent sequences. If $|X(\omega)| = |Y(\omega)|$, $x(n)$ and $y(n)$ can always be expressed as

$$x(n) = b(n) * a(n)$$

and

$$y(n) = \epsilon b(n) * a(N-1-n)$$

where $\epsilon = +1$ or -1 and $a(n)$ and $b(n)$ are real, causal, and finite extent sequences with N corresponding to the length of $a(n)$, i.e., $a(n) = 0$ outside $0 \leqslant n \leqslant N-1$.

Statement A2: Let $b(n)$ be a real, causal, and finite extent sequence. For any positive integer N, the equation

$$\text{Re}\{B(z)\, z^{-(N-1)/2}\,|_{z=e^{j\omega}}\} = 0$$

is satisfied for at least P distinct values of ω in the interval R, where P and R are as defined in (7).

Statement A3: Let $a(n)$ be a real sequence which is zero outside $0 \leqslant n \leqslant N-1$. If the equation

$$\text{Im}\{A(z)\, z^{(N-1)}\,|_{z=e^{j\omega}}\} = 0$$

is satisfied for at least P distinct values of ω in the interval R, then it is identically equal to zero and $a(n) = a(N-1-n)$.

We use the above three statements, whose proofs are shown in the Appendix, to demonstrate the following theorem:

Theorem 1: Let $x(n)$ and $y(n)$ be two real, causal, and finite extent sequences with z transforms which have no zeros on the unit circle. If $G_x^{\pi/2}(\omega) = G_y^{\pi/2}(\omega)$ for all ω, then $x(n) = y(n)$.

To show Theorem 1, we note from (5) and (6) that the condition $G_x^{\pi/2}(\omega) = G_y^{\pi/2}(\omega)$ is equivalent to

$$\text{sign}\{X_R(\omega)\}\,|X(\omega)| = \text{sign}\{Y_R(\omega)\}\,|Y(\omega)| \tag{8}$$

which in turn implies that $|X(\omega)| = |Y(\omega)|$, and therefore that

$$\text{sign}\{X_R(\omega)\} = \text{sign}\{Y_R(\omega)\}. \tag{9}$$

From Statement A1, then, $x(n)$ and $y(n)$ can be expressed as

$$x(n) = b(n) * a(n)$$

$$y(n) = \epsilon\, b(n) * a(N-1-n) \tag{10}$$

where $\epsilon = \pm 1$. Fourier transforming (10) we obtain

$$X(\omega) = A(\omega)\, B(\omega)$$

$$Y(\omega) = \epsilon\, e^{-j\omega(N-1)}\, A(-\omega)\, B(\omega). \tag{11}$$

To show that $\epsilon = 1$ in (11), we evaluate (9) at $\omega = 0$ and recognize that $X_R(0) = A(0)\,B(0)$ and $Y_R(0) = \epsilon A(0)\,B(0)$, so that

$$\text{sign}\{A(0)\,B(0)\} = \text{sign}\{\epsilon A(0)\,B(0)\}. \tag{12}$$

Since $X(\omega)$ is not zero at $\omega = 0$, (12) requires that $\epsilon = +1$.

Since $\epsilon = 1$, from (10), showing that $x(n) = y(n)$ is equivalent to showing that $a(n) = a(N-1-n)$. Toward this end, we consider the sum

$$X_R(\omega) + Y_R(\omega).$$

From (11) with $\epsilon = 1$, it can be shown that

$$X_R(\omega) + Y_R(\omega) = 2\,\text{Re}[A(\omega)e^{j\omega(N-1)/2}] \cdot \text{Re}[B(\omega)e^{-j\omega(N-1)/2}]. \tag{13}$$

From Statement A2, there are at least P distinct values of ω in the interval R which we denote as ω_i, $i = 1, 2, \cdots, P$ for which

$$\mathrm{Re}[B(\omega_i)e^{-j\omega_i(N-1)/2}] = 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (14)$$

From (13) and (14),

$$X_R(\omega_i) + Y_R(\omega_i) = 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (15)$$

From (9), both terms of the left-hand side of (15) have the same sign for all ω. Since a sum of two terms having the same sign can be zero only when both terms are zero, we have

$$X_R(\omega_i) = Y_R(\omega_i) = 0$$

and therefore also,

$$X_R(\omega_i) - Y_R(\omega_i) = 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (16)$$

From (11) and the fact that $\epsilon = 1$, it can be shown that (16) can be expressed as

$$X_R(\omega_i) - Y_R(\omega_i) = -2\,\mathrm{Im}[A(\omega_i)e^{j\omega_i(N-1)/2}] \cdot \mathrm{Im}[B(\omega_i)e^{-j\omega_i(N-1)/2}] = 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (17)$$

Since $B(\omega)$ is not zero for any ω, it follows from (14) that the second factor in (17) satisfies the property

$$\mathrm{Im}[B(\omega_i)e^{-j\omega_i(N-1)/2}] \neq 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (18)$$

From (17) and (18),

$$\mathrm{Im}[A(\omega_i)e^{j\omega_i(N-1)/2}] = 0, \quad i = 1, 2, \cdots, P, \omega_i \in R. \quad (19)$$

From (19) and Statement A3, $a(n) = a(N - 1 - n)$ so that $x(n) = y(n)$, thus demonstrating Theorem 1.

The result in Theorem 1 can be generalized in various ways. Specifically, in Theorem 1, we have assumed that $\alpha = \pi/2$, which is a specific representation of the 1-bit phase information. It can be shown that the statement is true for other choices of $0 < \alpha < \pi$. When $\alpha = \pi$ so that $S_x^\pi(\omega) = \mathrm{sign}[\theta_x(\omega)]$, a sequence is uniquely specified by $G_x^\pi(\omega)$ when $x(0) = 0$. Theorem 1 can also be extended to anticausal (left-sided) sequences. The proofs of these extensions can be found in [9]. When the above extensions are incorporated in Theorem 1, we have the following general theorem:

Theorem 2: Let $x(n)$ and $y(n)$ be two real, causal (or anticausal), and finite extent sequences, with z transforms which have no zeros on the unit circle. If $G_x^\alpha(\omega) = G_y^\alpha(\omega)$ for all ω and $0 < \alpha < \pi$, then $x(n) = y(n)$. When $\alpha = \pi$, if $G_x^\pi(\omega) = G_y^\pi(\omega)$ and $x(0) = y(0) = 0$, then $x(n) = y(n)$.

Theorems 1 and 2 explicitly require that the sequences be real-values and causal (or anticausal). The necessity of these conditions can be illustrated through counterexamples. Consider first the condition that the sequences be real, and let $y(n)$ equal $e^{j(\alpha-\pi)}x(n)$ where $x(n)$ is real. In this case, it is straightforward to show that $G_x^\pi(\omega) = G_y^\alpha(\omega)$. Since $G_x^\pi(\omega)$ does not uniquely specify $x(n)$, $G_y^\alpha(\omega)$ does not uniquely specify $y(n)$. To indicate the necessity of the causality (or anticausality) condition, consider as one counterexample the two-sided sequences $x(n)$ and $y(n)$ for which the z transforms are

$$X(z) = -z^2 + 6 - z^{-2} = (z + 2 - z^{-1})(-z + 2 + z^{-1})$$

$$Y(z) = z^2 + 4z + 2 - 4z^{-1} + z^{-2} = (z + 2 - z^{-1})^2. \quad (20)$$

For these two sequences it can be easily shown that $|X(\omega)| = |Y(\omega)|$ and $S_x^{\pi/2}(\omega) = S_y^{\pi/2}(\omega)$. In this case, then, $x(n)$ and $y(n)$ are different sequences, but they have the same signed FT magnitude.

In Theorems 1 and 2, uniqueness results were presented assuming that the signed spectral magnitude of a finite length sequence is known for all frequencies in the interval $(0, 2\pi)$. In the case of FT phase, it is possible to generalize the uniqueness results to the case in which the FT phase is known only for a finite number of distinct frequencies. Specifically, it has been shown [6] that for a finite length sequence of length N which has no symmetric (zero-phase) factors in its z transform, any $(N - 1)$ samples of the FT phase are sufficient to uniquely define the sequence to within a scale factor. Therefore, since the FT phase need not be known for all ω, such a result has been useful [6] in the development of practical algorithms for recontructing a finite length sequence from its FT phase samples. Unfortunately, however, a fixed finite set of signed magnitude samples is not always sufficient to uniquely specify a real, causal, and finite length sequence. For example, consider the following two causal sequences of length $N = 3$.

$$x(n) = 1.0\,\delta(n) + 2.6\,\delta(n-1) + 1.2\,\delta(n-2) \quad (21)$$

$$y(n) = 1.2\,\delta(n) + 2.6\,\delta(n-1) + 1.0\,\delta(n-2). \quad (22)$$

Since $y(n)$ is obtained from $x(n)$ by flipping both of the zeros of $X(z)$ about the unit circle, both $x(n)$ and $y(n)$ have the same spectral magnitude. Furthermore, in the interval $(0, \pi)$ the real part of the Fourier transform of $x(n)$ is equal to zero at only one frequency, $\omega = 0.477023\pi$ and the real part of the Fourier transform of $y(n)$ is equal to zero only at $\omega = 0.526166\pi$. Therefore, the signed magnitude of $X(\omega)$ is equal to the signed magnitude of $Y(\omega)$ for all ω outside the intervals $(0.477023\pi, 0.526166\pi)$ and $(-0.526166\pi, -0.477023\pi)$. Consequently, an arbitrary number of signed magnitude samples within this region is not sufficient to distinguish $x(n)$ from $y(n)$.

Even though a real, causal, finite extent sequence is not uniquely specified by samples of its signed FT magnitude at a finite number of arbitrary frequencies, it is specified by samples of its signed FT magnitude at a finite number of properly chosen frequencies which are different for different sequences. Specifically, for $x(n)$ which is zero outside $0 \leqslant n \leqslant N - 1$, the FT magnitude $|X(\omega)|$ is completely specified by $(N - 1)$ discrete Fourier transform (DFT) samples in the interval $(0, \pi)$. The 1 bit of FT phase $S_x^\alpha(\omega)$ is completely specified by the positions of its discontinuities and by its value at $\omega = 0$. Since the function $S_x^\alpha(\omega)$ has at most $2N$ discontinuities in $(-\pi, +\pi)$, $G_x^\alpha(\omega)$ is completely specified by a maximum of $3N$ samples at properly chosen frequencies.

In the above discussion, we considered only 1-D sequences. We now extent Theorem 2 to MD sequences. Let $x(\boldsymbol{n})$ denote an MD sequence $x(n_1, n_2, \cdots, n_M)$, and let $G_x^\alpha(\boldsymbol{\omega})$ denote the signed FT magnitude of $x(\boldsymbol{n})$, where $G_x^\alpha(\boldsymbol{\omega})$ represents $G_x^\alpha(\omega_1, \omega_2, \cdots, \omega_M)$ and is given by $S_x^\alpha(\boldsymbol{\omega})|X(\boldsymbol{\omega})|$. We define an MD sequence $x(\boldsymbol{n})$ to have a one-sided region of support in the M-dimensional space $n_1, n_2, \cdots, n_M$ if it has nonzero values for only one polarity of each n_i. For example, for a two-dimensional sequence there are four possible regions of support which are consistent with the sequence being one sided, corresponding to the four quadrants. Theorem 3, which follows, represents a generalization of Theorem 2 to encompass MD sequences.

Theorem 3: Let $x(\boldsymbol{n})$ and $y(\boldsymbol{n})$ be two real finite extent sequences with one-sided support and with z transforms which have no zeros at $|z_1| = |z_2| = \cdots = |z_M| = 1$. If $G_x^\alpha(\boldsymbol{\omega}) = G_y^\alpha(\boldsymbol{\omega})$ for all $\boldsymbol{\omega}$ and $0 < \alpha < \pi$, then $x(\boldsymbol{n}) = y(\boldsymbol{n})$. When $\alpha = \pi$, if $G_x^\pi(\boldsymbol{\omega}) = G_y^\pi(\boldsymbol{\omega})$ and $x(0) = y(0) = 0$, then $x(\boldsymbol{n}) = y(\boldsymbol{n})$.

We demonstrate the validity of Theorem 3 for a 2-D sequence which has the first-quadrant support size $M_1 \times M_2$ so that

$$x(n_1, n_2) = y(n_1, n_2) = 0 \text{ outside } 0 \leqslant n_1 \leqslant M_1 - 1 \text{ and } 0 \leqslant n_2 \leqslant M_2 - 1.$$

The proof for a higher dimension and for a different quadrant support is analogous to the 2-D case with the first-quadrant support. To demonstrate Theorem 3, we map the 2-D sequences $x(n_1, n_2)$ and $y(n_1, n_2)$ into two 1-D sequences $\hat{x}(n)$ and $\hat{y}(n)$ by the following transformation:

$$\begin{aligned} \hat{x}(n_1 \cdot M_2 + n_2) &= x(n_1, n_2) \\ \hat{y}(n_1 \cdot M_2 + n_2) &= y(n_1, n_2). \end{aligned} \tag{23}$$

In essence, the transformation in (23) corresponds to mapping a 2-D sequence to a 1-D sequence by concatenating the columns of the 2-D sequence. Clearly, $\hat{x}(n)$ and $\hat{y}(n)$ given by (23) are real, causal, and finite extent sequences. From (23), it is clear that the transformation is invertible. Furthermore, it can be shown [10] that

$$\hat{X}(\omega) = X(\omega_1, \omega_2) \Big|_{\omega_1 = \omega \cdot M_2, \quad \omega_2 = \omega}$$

and

$$\hat{Y}(\omega) = Y(\omega_1, \omega_2) \Big|_{\omega_1 = \omega \cdot M_2, \quad \omega_2 = \omega}. \tag{24}$$

From (24), it follows that the signed FT magnitudes of $\hat{x}(n)$ and $\hat{y}(n)$ are specified by the signed FT magnitudes of $x(n_1, n_2)$ and $y(n_1, n_2)$. Therefore, if $G_x^\alpha(\omega_1, \omega_2) = G_y^\alpha(\omega_1, \omega_2)$, then $G_{\hat{x}}^\alpha(\omega) = G_{\hat{y}}^\alpha(\omega)$. In addition, since $X(z_1, z_2)$ and $Y(z_1, z_2)$ have no zeros at $|z_1| = |z_2| = 1$, from (24), $\hat{X}(z)$ and $\hat{Y}(z)$ have no zeros on the unit circle. Since $\hat{x}(n)$ and $\hat{y}(n)$ satisfy all the conditions in Theorem 2, it follows from Theorem 2 that $\hat{x}(n) = \hat{y}(n)$. Since the transformation (23) is invertible, $x(n_1, n_2) = y(n_1, n_2)$ as required by Theorem 3.

The condition that $X(\boldsymbol{\omega}) \neq 0$ at any $\boldsymbol{\omega}$ is much more restrictive for 2-D sequences than for 1-D sequences, since $X(z) = 0$ represents surfaces in the (z_1, z_2) plane for 2-D sequences and points in the z plane for 1-D sequences. From the proof of Theorem 3 described above, however, it is not necessary to require $X(\boldsymbol{\omega}) \neq 0$ at any $\boldsymbol{\omega}$. We only need to require that $X(\boldsymbol{\omega}) \neq 0$ at the slices of $\boldsymbol{\omega}$ needed to form $\hat{X}(\boldsymbol{\omega})$ in (24). This is a much less restrictive condition than the condition in Theorem 3.

The theoretical result in Theorem 3 differs from that by Hayes [5] in several respects. In the result by Hayes [5], only samples of the FT magnitude are required, but the sequence is restricted to have a nonfactorizable z transform and the unique specification of the sequence is only to within a sign, a translation, and a central symmetry. In Theorem 3, the signed FT magnitude is required, but the sequence may have a factorizable z transform and is uniquely specified in the strict sense.

III. Algorithm

In Section II, we showed that under certain conditions a sequence is uniquely specified by its signed FT magnitude. In this section, we discuss an algorithm to implement the reconstruction of a sequence $x(\boldsymbol{n})$ from its signed FT magnitude. The sequence $x(\boldsymbol{n})$ is assumed to satisfy the conditions of Theorem 3. In addition, its signed FT magnitude $G_x^\alpha(\boldsymbol{\omega})$ is assumed known.

The algorithm that we have developed is an iterative procedure which is similar in style to other iterative procedures studied by Gerchberg-Saxton [11] and Fienup [12]. In the iterative algorithm, the "time" domain constraint that $x(\boldsymbol{n})$ is real and finite extent with a one-sided region of support, and the frequency domain constraint that the signed FT magnitude of $x(\boldsymbol{n})$ is given by $G_x^\alpha(\boldsymbol{\omega})$, are imposed separately in each iteration. Specifically, let $X_p(\boldsymbol{\omega})$ denote the estimate of $X(\boldsymbol{\omega})$ at the pth iteration. The estimate $X_p(\boldsymbol{\omega})$ is inverse Fourier transformed to the time domain to obtain $x_p'(\boldsymbol{n})$

$$x_p'(\boldsymbol{n}) = F^{-1}[X_p(\boldsymbol{\omega})]. \tag{25}$$

From $x_p'(\boldsymbol{n})$, we generate an estimate $x_p''(\boldsymbol{n})$ which satisfies the time domain constraints

$$x_p''(\boldsymbol{n}) = \begin{cases} \mathrm{Re}[x_p'(\boldsymbol{n})] & \text{for } \boldsymbol{n} \in A \\ 0 & \text{for } \boldsymbol{n} \notin A \end{cases} \tag{26}$$

where A represents the known support region of $x(\boldsymbol{n})$.

The sequence $x_p''(\boldsymbol{n})$ is then Fourier transformed back to the frequency domain to obtain $X_p''(\boldsymbol{\omega})$ as follows:

$$X_p''(\boldsymbol{\omega}) = F[x_p''(\boldsymbol{n})]. \tag{27}$$

The new frequency domain estimate $X_{p+1}(\boldsymbol{\omega})$ is then obtained by enforcing the constraint that $G_{x_{p+1}}^\alpha(\boldsymbol{\omega}) = G_x^\alpha(\boldsymbol{\omega})$ as follows:

$$X_{p+1}(\boldsymbol{\omega}) = \begin{cases} |X(\boldsymbol{\omega})| e^{j\theta_{x_p''}(\boldsymbol{\omega})} & \text{if } S_{x_p''}^\alpha(\boldsymbol{\omega}) = S_x^\alpha(\boldsymbol{\omega}) \\ |X(\boldsymbol{\omega})| e^{j(2\alpha - \theta_{x_p''}(\boldsymbol{\omega}))} & \text{if } S_{x_p''}^\alpha(\boldsymbol{\omega}) = -S_x^\alpha(\boldsymbol{\omega}). \end{cases} \tag{28}$$

Specifically, the correct magnitude is substituted for the estimated magnitude. If $S_{x_p''}^\alpha(\boldsymbol{\omega}) = S_x^\alpha(\boldsymbol{\omega})$, then the phase of the estimate is retained. Otherwise, the estimate is reflected about a line that passes through the origin with angle α to correct the sign of $S_{x_p''}^\alpha(\boldsymbol{\omega})$. This completes one iteration. The initial estimate $X_0(\boldsymbol{\omega})$ we have used is given by

$$X_0(\boldsymbol{\omega}) = |X(\boldsymbol{\omega})| e^{j\theta_{x_0}(\boldsymbol{\omega})} \tag{29}$$

where $\theta_{x_0}(\boldsymbol{\omega})$ is given by

$$\theta_{x_0}(\boldsymbol{\omega}) = \begin{cases} \alpha - \dfrac{\pi}{2} & \text{for } S_x^\alpha(\boldsymbol{\omega}) = +1 \\ \alpha + \dfrac{\pi}{2} & \text{for } S_x^\alpha(\boldsymbol{\omega}) = -1. \end{cases} \tag{30}$$

The iterative algorithm discussed above is illustrated in Fig. 3.

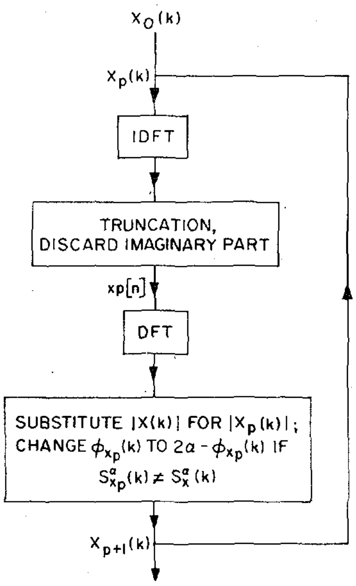


Fig. 3. Block diagram of the iterative algorithm.

The asymptotic behavior of the algorithm in Fig. 3 has not yet been studied theoretically. We have observed experimentally that a stable estimate of the sequence to be retrieved is always attained after a large number of iterations.

To implement the algorithm in Fig. 3, the Fourier and inverse Fourier transform operations are approximated by discrete Fourier transfrom (DFT) and inverse DFT (IDFT) operations. Although the uniqueness is not guaranteed in terms of the signed FT magnitude samples, we have empirically observed that the algorithm reconstructs the desired sequence provided that the signed FT magnitude is densely sampled in the frequency domain, so that the FT magnitude is completely specified and the discontinuities of $S_x^{\alpha}(\boldsymbol{\omega})$ are individually resolved by the samples of $S_x^{\alpha}(\boldsymbol{\omega})$. The FT magnitude $|X(\boldsymbol{\omega})|$ is completely specified by samples of $|X(\boldsymbol{\omega})|$ when the DFT size is twice the size of the known support of $x(n)$ in each dimension.

IV. Examples

The algorithm discussed in Section III has been used to reconstruct a variety of different 1-D and 2-D sequences from their signed FT magnitudes. In this section, we present some of these examples.

Fig. 4 illustrates one example in which a 1-D sequence is reconstructed from its signed FT magnitude. In Fig. 4(a) is shown a 47-point sequence obtained by sampling female speech at a 10 kHz rate. In Fig. 4(b) is shown the sequence reconstructed by using the iterative algorithm with the DFT size of 1024 after 50 iterations. In addition to the above example, a number of other examples have been considered. In all cases, we observed that the algorithm reconstructs the desired sequence.

Fig. 5 illustrates an example in which a 2-D sequence is reconstructed from its signed FT magnitude. In Fig. 5(a) is shown an image of size 256 X 256 pixels. In Fig. 5(b) is shown the image reconstructed by using the iterative algorithm using the DFT size of 512 X 512 after 10 iterations.

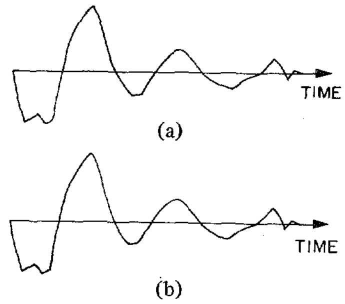


Fig. 4. Speech segment sampled at 47 points. (a) Original sequence. (b) Reconstructed sequence after 50 iterations.

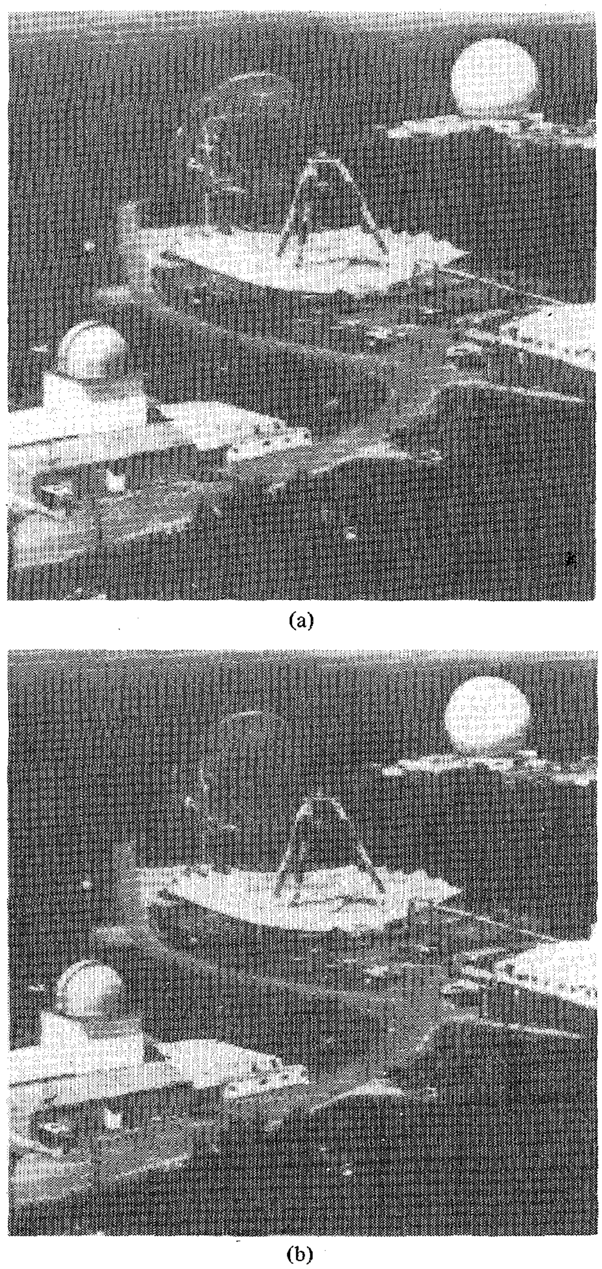


Fig. 5. Image of size 256 X 256 pixels. (a) Original image. (b) Reconstructed image after 10 iterations.

In addition to the examples shown in this section, we have studied a number of other examples. From these examples, we have made the following observations about the iterative algorithm. First, for sequences satisfying the uniqueness constraints, if a DFT size below some threshold value is used, the algorithm does not lead to the desired sequence. The threshold value is different for different sequences, and we have not yet found a simple way to determine the threshold value for a given sequence. In practice, therefore, the DFT size is typically much larger than the threshold value to reconstruct a sequence from its signed FT magnitude. Second, the DFT size required is typically much larger (by more than a factor of 10 typically) than the size of the data for 1-D signals. For 2-D signals, we have observed that the DFT size of $2N \times 2N$ when the data size is $N \times N$ is sufficient for all examples we considered. This difference is in part due to the fact that the magnitude of $2N \times 2N$ DFT when the data size is $N \times N$ uniquely specifies a 2-D sequence within a sign factor, a translation, and a central symmetry, and therefore the ambiguity that needs to be resolved by 1 bit of phase information is much less for 2-D signals than for 1-D signals. Third, the threshold DFT length is approximately the same for different choices of α, as long as α is not too close to 0 or π. As α approaches 0 or π, the thresold length is significantly increased. The choice of $\alpha = \pi/2$ permits the use of FFT routines specific to real sequences, and therefore, uses less computation time and less storage space. Fourth, the convergence rate of the iterative algorithm is rapid initially and becomes slow as the number of iterations is increased. Fifth, we have observed that the mean square error between the original and reconstructed sequences decreases monotonically as the number of iterations increases. Sixth, the convergence rate of the algorithm can be significantly improved by using an acceleration procedure similar to that used by Oppenheim *et al.* [13]. Further details on the behavior of the iterative algorithm can be found in Van Hove [9].

V. Conclusions

In this paper, we have shown that a 1-D or MD sequence is uniquely specified under mild restrictions by its signed FT magnitude. In addition, we have developed an iterative algorithm to reconstruct a 1-D or MD sequence from its signed FT magnitude. When this result is combined with the previous result [5] on the problem of reconstructing a 1-D or MD sequence from its FT phase, we obtain a very general result that a 1-D or MD sequence is uniquely specified by its FT phase or its signed FT magnitude. In addition, under mild restrictions, an iterative algorithm which is similar in style can be used to reconstruct a 1-D or MD sequence from its FT phase or signed magnitude.

Appendix

Statement A1: Let $x(n)$ and $y(n)$ be two real, causal, and finite extent sequences. If $|X(\omega)| = |Y(\omega)|$, $x(n)$ and $y(n)$ can always be expressed as

$$x(n) = b(n) * a(n)$$
$$y(n) = \epsilon b(n) * a(N - 1 - n)$$

where $\epsilon = +1$ or -1 and $a(n)$ and $b(n)$ are real, causal, and finite extent with N corresponding to the length of $a(n)$, i.e., $a(n) = 0$ outside $0 \leqslant n \leqslant N - 1$.

Proof: A general expression of the z transform $X(z)$ of a sequence $x(n)$ which is causal and has a finite support is given by

$$X(z) = z^{-n_1} x_0 \prod_{i=1}^{Q} (1 - z_i z^{-1}) \tag{A1.1}$$

where z_i, $i = 1, 2, \cdots, Q$, are the zeros of $X(z)$, x_0 is the first nonzero sample, and n_1 is the positive initial delay in $x(n)$. It is well known that the FT magnitude of a finite extent 1-D sequence remains unchanged only when the sequence is subject to linear shifts, sign inversions, and/or zero "flipping." The z transform $Y(z)$ may therefore be written as

$$Y(z) = \pm z^{-n_2} x_0 \prod_{i \in \{u\}} (1 - z_i z^{-1}) \prod_{i \in \{r\}} (-z_i + z^{-1}) \tag{A1.2}$$

where n_2 is the positive initial delay in $y(n)$, $\{r\}$ is the set of indexes of the R zeros of $Y(z)$ which are zeros of $X(z)$ reflected across the unit circle, and $\{u\}$ is the set of indexes of zeros which are unchanged from $X(z)$ to $Y(z)$. We may also write (A1.1) and (A1.2) as

$$X(z) = A(z) \cdot B(z)$$
$$Y(z) = \pm C(z) \cdot B(z)$$

or

$$x(n) = a(n) * b(n)$$
$$y(n) = \pm c(n) * b(n) \tag{A1.3}$$

where

$$A(z) = z^{-(n_1 - n_2)} \prod_{i \in \{r\}} (1 - z_i z^{-1})$$
$$B(z) = z^{-n_2} x_0 \prod_{i \in \{u\}} (1 - z_i z^{-1})$$
$$C(z) = \prod_{i \in \{r\}} (-z_i + z^{-1}). \tag{A1.4}$$

We now show that $c(n)$ is $a(n)$ time reversed, represented by $a'(n)$. The length of the sequence $a'(n)$ is $N = n_1 - n_2 + R + 1$, if we include the leading zeros. Therefore,

$$a'(n) = a(N - 1 - n)$$
$$A'(z) = A(z^{-1}) z^{-(N-1)} = z^{-R} \prod_{i \in \{r\}} (1 - z_i z) = C(z)$$

so that $c(n) = a(N - 1 - n)$. From (A1.3), the sequences $x(n)$ and $y(n)$ are expressed in the adequate form. To characterize $a(n)$ and $b(n)$, we examine their z transforms. Since $B(z)$ contains only a finite number of negative powers of z, the sequence $b(n)$ has a finite causal support. Since $A(z)$ and $A'(z) = C(z)$ contain only negative powers of z, it follows that $a(n)$ and $a(N - 1 - n)$ are causal so that $a(n)$ is zero outside $0 \leqslant n \leqslant N - 1$. If the z transform $X(z)$ contains a pair of complex conjugate zeros, than they must both belong to $\{u\}$ or both to $\{r\}$ for $y(n)$ to be real-valued. The z transforms $A(z)$ and $B(z)$ may therefore contain complex zeros only in conjugate pairs so that $a(n)$ and $b(n)$ are real. In the case $n_2 > n_1$, we simply exchange the roles of $x(n)$ and $y(n)$. This completes the proof of Statement A1.

Statement A2: Let $b(n)$ be a real, causal, and finite extent sequence. For any positive integer N, the equation

$$\text{Re}\,\{B(z)\,z^{-(N-1)/2}\,|_{z=e^{j\omega}}\} = 0$$

is satisfied for at least P distinct values of ω in the interval R where P and R are as defined in (7) of the text.

To prove this statement, we introduce the notion of unwrapped phase. Given a Fourier transform $M(\omega)$ which has no zeros, we define its unwrapped phase $\phi_M(\omega)$ as the unique continuous function of ω which satisfies

$$M(\omega) = |M(\omega)|\,e^{j\phi_M(\omega)} \tag{A2.1}$$

for all ω and which takes the value of 0 or $-\pi$ at $\omega = 0$. The unwrapped phase has the following properties. If we define the function $F(\omega)$ as

$$F(\omega) = D(\omega)\,B(\omega) \tag{A2.2}$$

then if follows that

$$\phi_F(\omega) = \phi_D(\omega) + \phi_B(\omega) + 2\alpha\pi$$

where

$$\alpha = \begin{cases} 1 & \text{if } \phi_D(0) = \phi_B(0) = -\pi \\ 0 & \text{otherwise.} \end{cases} \tag{A2.3}$$

The unwrapped FT phase $\phi_B(\omega)$ of a causal sequence $b(n)$ satisfies

$$\phi_B(0) \geqslant \phi_B(\pi). \tag{A2.4}$$

The unwrapped phase of the function

$$D(\omega) = e^{-j\omega(N-1)/2} \tag{A2.5}$$

is

$$\phi_D(\omega) = -\omega\,\frac{N-1}{2}. \tag{A2.6}$$

We now proceed to the proof of statement A2. We consider the unwrapped phase $\phi_F(\omega)$ of the function

$$F(\omega) = B(\omega)e^{-j\omega(N-1)/2}.$$

The equation $\text{Re}(F(\omega)) = 0$ has the same roots as the equation

$$\phi_F(\omega) = \frac{\pi}{2} + k\pi, \text{ with } k \text{ an integer,}$$

since $F(\omega)$ has no zeros. From our previous discussion, we have

$$\phi_F(\pi) - \phi_F(0) = \phi_B(\pi) - \phi_B(0) + \phi_D(\pi) - \phi_D(0) \leqslant -\left(\frac{N-1}{2}\right)\pi.$$

Since the continuous function $\phi_F(\omega)$ decreases at least by $(N-1)/2\ \pi$ on the interval R, it follows that the graph of $\phi_F(\omega)$ crosses at least $N/2$ lines of phase $\pi/2 + k\pi$ in $(0, \pi]$ if N is even and at least $(N-1)/2$ such lines in $(0, \pi)$ if N is odd. Fig. 6 shows $\phi_F(\omega)$ when $b(n) = \delta(n)$, for the cases $N = 4$ and $N = 5$.

Statement A3: Let $a(n)$ be a real valued sequenced which is zero outside $0 \leqslant n \leqslant N-1$. If the equation

$$\text{Im}\,\{A(z)\,z^{(N-1)/2}\,|_{z=e^{j\omega}}\} = 0$$

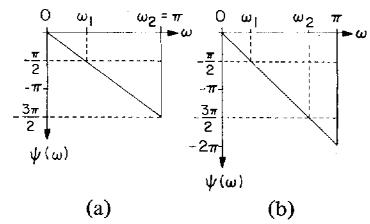


Fig. 6. Unwrapped phase of the function $F(\omega)$ for $b(n) = \delta(n)$. (a) $N = 4$. (b) $N = 5$.

is satisfied for at least P distinct values of ω in the interval R, then it is identically equal to zero and $a(n) = a(N-1-n)$. P and R are defined as in (7) in the text as

$$P = \frac{N-1}{2} \text{ and } R = (0, \pi) \qquad \text{for } N \text{ odd}$$

$$P = \frac{N}{2} \text{ and } R = (0, \pi] \qquad \text{for } N \text{ even}$$

Proof for N Odd: With the use of trigonometric formulas, we obtain

$$G(\omega) = \text{Im}(A(\omega)e^{j\omega(N-1)/2}) = \sum_{n=0}^{N-1} a(n)\sin\left(\frac{N-1}{2} - n\right)\omega \tag{A3.1}$$

$$G(\omega) = \sum_{n=1}^{(N-1)/2}\left\{a\left(\frac{N-1}{2} - n\right) - a\left(\frac{N-1}{2} + n\right)\right\}\sin n\omega. \tag{A3.2}$$

Since the set of the $(N-1)/2$ functions $\sin\omega, \sin 2\omega, \cdots, \sin (N-1)\omega/2$ is a Chebyshev set on the interval $(0, \pi)$ as is shown in [9] and since $G(\omega)$ has at least $(N-1)/2$ distinct roots in the interval $(0, \pi)$, it follows that the coefficients of the expansion in the right-hand side of (A3.2) must vanish

$$a\left(\frac{N-1}{2} - n\right) = a\left(\frac{N-1}{2} + n\right) = 0;$$

$$n = 1, 2, \cdots, \frac{N-1}{2}$$

or

$$a(n) = a(N-1-n); \qquad n = 0, 1, \cdots, N-1.$$

When N is even, the expansion of $G(\omega)$ is

$$G(\omega) = \sum_{n=0}^{(N/2)-1}\left\{a\left(\frac{N}{2} - 1 - n\right) - a\left(\frac{N}{2} + n\right)\right\} \cdot \sin\left(n + \frac{1}{2}\right)\omega.$$

Since the functions $\sin\omega/2, \sin 3\omega/2, \cdots, \sin N - 1/2\ \omega$ form a Chebyshev set on the interval $(0, \pi]$ as is shown in [9], it follows that

$$a\left(\frac{N}{2} - 1 - n\right) - a\left(\frac{N}{2} + n\right) = 0; \qquad n = 0, 1, \cdots, \frac{N}{2} - 1$$

or

$$a(n) = a(N - 1 - n); \qquad n = 0, 1, \cdots, N - 1.$$

This completes the proof of Statement A3.

REFERENCES

[1] W. O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy.* New York: Academic, 1978.
[2] G. N. Ramachandran and R. Srinivasan, *Fourier Methods in Crystallography.* New York: Wiley-Interscience, 1970.
[3] D. L. Misell, "An examination of an iterative method for the solution of the phase problem in optics and electron optics: I and II," *J. Phys. D. Appl. Phys.*, vol. 6, pp. 2200-2225, 1973.
[4] W. K. Pratt, *Digital Image Processing.* New York: Wiley, 1978.
[5] M. H. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 140-154, Apr. 1982.
[6] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 672-680, Dec. 1980.
[7] M. H. Hayes, "Signal reconstruction from phase or magnitude," Sc.D. dissertation, Dep. Elec. Eng. Comput. Sci., M.I.T., June 1981.
[8] Y. H. Bruck and L. G. Sodin, "On the ambiguity of the image reconstruction problem," *Opt. Commun.*, pp. 304-308, Sept. 1980.
[9] P. L. Van Hove, "Signal reconstruction from Fourier transform amplitude," S.M. thesis, Dep. Elec. Eng. Comput. Sci., M.I.T., Sept. 1982.
[10] R. M. Mersereau and D. E. Dudgeon, "The representation of two-dimensional sequences as one-dimensional sequences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 320-325, Oct. 1974.
[11] 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.
[12] J. R. Fienup, "Space object imaging through the turbulent atmosphere," *Opt. Eng.*, vol. 18, pp. 529-534, Sept. 1979.
[13] A. V. Oppenheim, M. H. Hayes, and J. S. Lim, "Iterative procedure for signal reconstruction from Fourier transform phase," *Opt. Eng.*, vol. 21, pp. 122-127, Jan.-Feb. 1982.

Patrick L. Van Hove was born in Brussels, Belguim, on March 17, 1956. He received the Ingenieur Civil Electricien degree from the Universite Libre de Bruxelles (ULB), Brussels, Belgium, in 1978 and the M.S. in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1982.

From 1978 to 1981 he was an assistant at the ULB where he worked primarily on coherent optics and on acoustics. From September 1981 to August 1982, he was studying at M.I.T. under a fellowship from the Belgian American Educational Foundation, and he completed his thesis with the digital processing group. After completing his military duties in Belgium, he is now a Research Assistant with the Digital Signal Processing Group at M.I.T. His interests include theoretical aspects of signal processing, its application to image and speech processing, and optical holography.

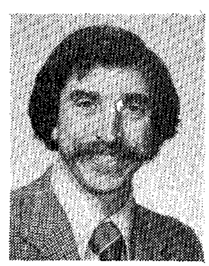
Monson H. Hayes (S'76-M'80-S'80-M'81) was born in Washington, DC, on October 27, 1949. He received the B.S. degree in physics from the University of California, Berkeley, in 1971, and the S.M., E.E., and Sc.D. degrees, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge, in 1978, 1978, and 1981, respectively.

He worked as a Systems Engineer in IR systems technology at Aerojet Electrosystems from 1972 to 1974. From 1975 to 1979 he was a Teaching Assistant in the Department of Electrical Engineering at M.I.T., and from 1979 to 1981 was a Research Assistant at M.I.T. Lincoln Laboratory, Lexington, in the Multidimensional Digital Signal Processing Group. Since 1981 he has been with the Georgia Institute of Technology, Atlanta, where he is currently an Assistant Professor of Electrical Engineering. His research interests include digital signal processing and multidimensional signal processing and its applications to image processing.

Jae S. Lim (S'76-M'78) was born on December 2, 1950. He received the S.B., S.M., E.E., and Sc.D. degrees in electrical engineering and computer Science from the Massachusetts Institute of Technology Cambridge, in 1974, 1975, 1978, respectively.

He joined the M.I.T. faculty in 1978 as an Assistant Professor, and is currently Associate Professor in the Department of Electrical Engineering and Computer Science. His research interests include digital signal processing and its applications to image and speech processing. He has contributed more than 60 articles to journals and conference proceeedings, and is the editor of a reprint book, *Speech Enhancement* (Englewood Cliffs, NJ: Prentice-Hall, 1982). He is the winner of two prize paper awards, one from the Boston Chapter of the Acoustical Society of America in December 1976, and one from the IEEE Acoustics, Speech, and Signal Processing Society in April 1979.

Dr. Lim is a member of Eta Kappa Nu, Sigma Xi, and Chairman of the IEEE ASSP Society Technical Committee on Digital Signal Processing.



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

From 1961 to 1964 he was a member of the Massachusetts Institute of Technology 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 homorphic filtering as an approach to some nonlinear filtering problems. In 1964 he joined the faculty in the Department of Electrical Engineering and Computer Science, where he is currently full professor. He has been a Guggenheim Fellow at the University of Grenoble, Grenoble, France, he has held the Cecil H. Green Distinguished Chair in Electrical Engineering and Computer Science at M.I.T. and for a two year period he served as Associate Head of the Data Systems Division at M.I.T. Lincoln Laboratory. He has received a number of IEEE awards and the M.I.T. Graduate Student Council Teaching Award. His research interests are in the area of digital signal processing and its applications to speech, image, and seismic data processing, in the areas of system theory and knowledge-based signal processing. He is coauthor and editor of several texts in the area of signal processing, and is editor of a quarterly publication, *Trends & Perspectives in Signal Processing.*

Dr. Oppenheim is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.