#2

Reconstruction of Multidimensional Signals from Zero Crossings

Susan R. Curtis

Loan Copy Only

Technical Report 509

June 1985

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

Reconstruction of Multidimensional Signals from Zero Crossings

Susan R. Curtis

Technical Report 509

June 1985

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Research Laboratory of Electronics Massachusetts Institute of Technology		Office of Naval Research Mathematical and Information Scien. Div.

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
77 Massachusetts Avenue Cambridge, MA 02139	800 North Quincy Street Arlington, Virginia 22217

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Advanced Research Projects Agency		N00014-81-K-0742

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.			
1400 Wilson Boulevard Arlington, Virginia 22217	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
11. TITLE (Include Security Classification) Reconstruction of Multidimensional Signals from Zero Crossings			NR 049-506	

12. PERSONAL AUTHOR(S) Susan R. Curtis

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	June 1985	96

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

    This thesis addresses the problem of reconstruction of multidimensional signals from zero crossing or threshold crossing information. The basic theoretical result shows that most two-dimensional, periodic, band-limited signals are uniquely specified with zero crossings or with crossings of an arbitrary threshold. This result also applies to finite length signals and to situations where only a discrete set of zero crossing points are available. Additional results show that signals of higher dimensions and nonperiodic signals are also uniquely specified with zero crossings or threshold crossings. By applying the duality of the Fourier transform in a straightforward way, it is also shown that finite-length signals are uniquely specified with zero crossings (or sign information) in the Fourier domain under analogous sets of conditions.
                                                                    (cont.)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Kyra M. Hall RLE Contract Reports	(617) 253-2569	

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

19. Abstract continued

A problem distinct from that of uniquely specifying signals with zero crossings is that of developing specific algorithms for recovering signals from zero crossing information once it is known that the signals satisfy the appropriate constraints. Two algorithms for recovering signals from zero crossings or threshold crossings are presented and one is successfully used to recover example images from threshold crossings. The thesis concludes with suggestions for future research including possible applications of these results.

Reconstruction of Multidimensional Signals

from Zero Crossings

by

Susan R. Curtis

Submitted to the Department of Electrical Engineering
and Computer Science on May 16, 1985
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Abstract

This thesis addresses the problem of reconstruction of multidimensional signals from zero crossing or threshold crossing information. The basic theoretical result shows that most two-dimensional, periodic, bandlimited signals are uniquely specified with zero crossings or with crossings of an arbitrary threshold. This result also applies to finite length signals and to situations where only a discrete set of zero crossing points are available. Additional results show that signals of higher dimensions and nonperiodic signals are also uniquely specified with zero crossings or threshold crossings. By applying the duality of the Fourier transform in a straightforward way, it is also shown that finite-length signals are uniquely specified with zero crossings (or sign information) in the Fourier domain under analogous sets of conditions.

A problem distinct from that of uniquely specifying signals with zero crossings is that of developing specific algorithms for recovering signals from zero crossing information once it is known that the signals satisfy the appropriate constraints. Two algorithms for recovering signals from zero crossings or threshold crossings are presented and one is successfully used to recover example images from threshold crossings. The thesis concludes with suggestions for future research including possible applications of these results.

Thesis Supervisor: Prof. Alan V. Oppenheim
Title: Professor of Electrical Engineering

ACKNOWLEDGMENTS

Table of Contents

CHAPTER 1

INTRODUCTION

The importance of zero crossing locations in determining the nature of both one- and two-dimensional signals has been recognized for some time. Experiments in speech processing have shown that speech with only the zero crossing information preserved (hard-clipped speech) retains much of the intelligibility of the original speech [1]. Also, a wide variety of papers in image processing and vision stress the importance of the information contained in the edges of objects and one theory of human vision relies primarily on edge detection as the mechanism by which humans process visual information [2].

There are also a variety of other types of applications in which the zero crossings or threshold crossings are available and it is desired to recover the original signal. One possible application occurs when an image is clipped or otherwise distorted in such a way as to preserve zero crossing or level crossing information and it is desired to recover the original signal from this information. This might happen if an image is recorded on a high-contrast film or any film with an unknown characteristic. If it is possible to recover the original signal from its threshold crossings, then it is possible, at least in principle, to recover the original signal from its distorted version and to determine the type of nonlinearity present. In addition, it is not necessary for the nonlinearity to be monotonic; it is only necessary that it preserve the threshold crossing information. This could be useful in an application

such as medical archiving, where intensity levels of images recorded on film are likely to become distorted over time, but threshold crossing information could be preserved. In some archiving applications it is unlikely that any particular image may need to be retrieved, but it is important to be able to recover the image if necessary even if the process is expensive or time-consuming. Another possible type of application of results on reconstruction from zero crossing information is in a variety of design problems such as filter design [3] and antenna design [4]. In these cases, one could potentially specify the zero crossing points or null points of the filter response or antenna pattern and then use these points to derive the remainder of the response.

One might also consider the possibility of exploiting the information in zero crossings for signal coding and data compression. However, in representing a two-dimensional signal with zero crossings or threshold crossings, it is important to recognize that the amplitude information in the original signal is embedded in the exact location of the zero crossings. Consequently, it is not unreasonable that while the original signal can be sampled at the Nyquist rate, the zero crossing representation may require a considerably higher, possibly infinite, sampling rate to adequately preserve the zero crossing locations. Thus the total number of bits or bandwidth required in the zero crossing representation might well be higher than that required by sampling and quantizing the original signal. For this reason, we expect results on signal reconstruction from zero crossings to be more useful in applications where the exact zero crossing points are available. It is possible, how-

ever, to view the representation of signals with zero crossings as a potential trade-off between the bandwidth and the dynamic range necessary for transmitting a signal. If the available bandwidth is sufficient to accurately preserve the zero crossing locations, then the dynamic range requirements might be greatly reduced if the signal could be recovered from the zero crossing locations.

The possibility of recovering signals from zero crossings has been considered in a number of different papers in the field of communication theory (see [5] for references), although the great majority of this work has concentrated on one-dimensional signals. In this thesis, we study the problem of recovering multidimensional signals from zero crossings. As will be discussed in more detail later, there is a fundamental difference between one-dimensional zero-crossing problems and multidimensional problems since in one dimension, zero crossings represent a discrete set of points and in two dimensions or higher, zero crossings are contours or surfaces. Thus it is possible to develop results for the multidimensional problem which have no direct analog in the one-dimensional case. These results, which shall be developed in this thesis, appear to be less restrictive and more broadly applicable than two-dimensional extensions of one-dimensional results previously reported.

We shall begin this thesis by reviewing related research on reconstruction of signals from zero crossings. In developing our new theoretical results on unique specification with zero crossings, we shall begin by considering two-dimensional periodic signals since these signals can be represented as polynomials with a Fourier series expansion, allowing us to apply known properties of the zeros of polynomials.

These results will be presented in chapter 3, along with the extension to signals of higher dimensions. Most of the results developed for periodic signals can be generalized to nonperiodic signals; this generalization will be presented in chapter 4. In chapter 5, we will use the results of chapters 3 and 4 to develop conditions under which signals are uniquely specified with zero crossings in the Fourier domain by applying the duality of the Fourier transform in a straight-forward way.

A problem distinct from that of uniquely specifying signals with zero crossings is the problem of developing effective techniques for recovering a signal from zero crossing information once it is known that the signal satisfies the appropriate uniqueness constraints. Two possible algorithms are presented and evaluated in chapter 6. We include experimental results illustrating images we have successfully recovered from zero crossing information. We conclude the thesis with suggestions for future research including possible applications of our results.

CHAPTER 2

RELATED RESEARCH

A significant amount of research in the field of communication theory has been devoted to the problem of unique representation of signals with zero crossings, primarily for one-dimensional signals. In this chapter, we review this research to provide background for the results to be developed later in this thesis. A more detailed review of results in this area can be found in reference [5].

A wide variety of papers have dealt with the importance of zero crossings and the possibility of recovering a one-dimensional signal from zero crossing information alone. These results are primarily based upon the theory of entire functions since it is known that a bandlimited function can be uniquely extended to the complex plane as an entire function. Most of these results have come from the field of communication theory, where it is often of considerable value to know what type of information is sufficient to uniquely specify a signal. In addition, the theory of zeros of entire functions has been used for comparison and analysis of different types of modulation systems and for studying properties of modulated signals. For example, Voelcker [6] shows how modulation processes can be considered to be methods of manipulating or extracting the zeros of a signal, and that different systems can be analyzed as to how they effect the zeros of a signal. Additional applications for this work occur in a variety of design problems where one might want to specify a system response or other signal in terms of zero crossings and derive the

remainder of the response from these. Systems are often designed and analyzed in terms of poles and zeros, which in the case of lumped systems, involves the study of zeros of polynomials, and in the case of distributed systems, involves the study of zeros of non-polynomial entire functions (an example is given in [7]).

Most of the results on the unique specification of one-dimensional signals are based upon the fact that a bandlimited function is entire (analytic everywhere) and is thus uniquely specified by its zeros (real and complex) to within a constant and an exponential factor. An arbitrary bandlimited function is uniquely specified by its (real) zero crossings if all its zeros are guaranteed to be real. Thus, a number of previous research efforts concentrated on identifying conditions under which signals have only real zeros and developing methods for modifying a signal so that all of its zeros become real. One result in this area is that a one-dimensional complex signal with no energy for negative frequencies is uniquely specified by the zero crossings of its real part if the complex signal has zeros only in the upper half-plane [5, 8]. (A more general form of this result is given in [9].) One method of modifying signals so that all of their zeros become real is to add a sinusoid of sufficient amplitude at a frequency corresponding to the band edge [10]; another is to repeatedly differentiate the signal [8]. Some modulation schemes have also been shown to produce signals with only real zeros [6].

Additional work has involved identifying signals which are uniquely specified by their (real) zero crossings despite the fact that they also contain complex zeros. This is possible if the zero crossing rate is in some sense higher than the informa-

tion rate or bandwidth of a signal. Voelcker [11] has shown that for angle-modulated systems, knowledge of the zero crossing locations of the modulated signal is equivalent to knowledge of samples of the signal before modulation; thus if the original signal is bandlimited, then the zero crossings of the modulated signal are sufficient to uniquely specify it. Fairly recently, in response to experimental results presented by Voelcker and Requicha [12], Logan [9] developed a new class of bandpass signals which are uniquely specified by their zero crossings. Specifically, Logan showed that a signal with a bandwidth of less than one octave is uniquely specified by its zero crossings if it has no zeros in common with its Hilbert transform other than real simple zeros. This means that almost all bandpass signals of bandwidth less than one octave are uniquely specified by their zero crossings.

It is also possible to interpret results on unique specification of signals with zero crossings as a type of sampling, where the "samples" consist of the set of points (times) corresponding to zero crossings, as opposed to the amplitude of the signal at particular fixed instants [5]. Using this point of view, "sampling" might consist of adding a sine wave at the appropriate frequency and recording those instants when the resulting signal crosses zero, or equivalently, recording those instants where the original signal crosses a sinusoid. "Interpolation" would then consist of generating a signal with sine-wave crossings at the specified instants. Sampling and interpolation systems using this approach have been designed, implemented, and found to produce good results [13].

Despite the number of results on the unique specification of signals with zero crossings, most one-dimensional bandlimited signals encountered in practice do not satisfy the constraints associated with any of the above results and are not uniquely specified by their zero crossings unless they satisfy some additional constraints which effectively guarantee that they contain a sufficient number of zero crossings. In fact, it has been shown [14] that almost all sample functions of a bandlimited Gaussian random process are not uniquely specified by zero crossings.

Although a considerable amount of theoretical work has been devoted to the problem of reconstruction of one-dimensional signals from zero crossings, much less work has been devoted to the corresponding two-dimensional problem. Logan's result has been extended to two dimensions [2, 15] by requiring a one-dimensional signal derived from the original two-dimensional signal to satisfy the constraints of Logan's theorem. In addition, one-dimensional results on reconstruction from sine-wave crossings have been extended to two-dimensional problems [16]. However, as mentioned earlier, the two-dimensional problem is fundamentally different from the one-dimensional problem since in two dimensions, the "zero crossings" are actually zero crossing contours and not isolated points as in the one-dimensional case. It is possible to derive results on the unique specification of multidimensional signals with zero crossings which are based directly on the properties of multidimensional signals and are not derived from similar results for one-dimensional signals. These results will be presented in this thesis.

CHAPTER 3

UNIQUE SPECIFICATION OF PERIODIC SIGNALS WITH ZERO CROSSINGS

The representation of a signal in terms of zero crossings can be thought of as a form of nonuniform sampling, with each zero crossing representing one sample [5]. Most one-dimensional bandlimited signals are not uniquely determined by zero crossings since the average rate of zero crossings is not guaranteed to be sufficiently high [14]. Logan's condition [9] requires signals to be bandpass with a bandwidth of less than one octave so that the number of zero crossings (or the rate of zero crossings) is in some sense consistent with the amount of information or bandwidth in the signal. In two dimensions, in contrast to one dimension, the "zero crossings" (sign changes) of a signal are contours and thus each zero crossing contour corresponds to an infinite number of samples of the signal. Thus it is reasonable to suggest that a two-dimensional signal may be specified with zero crossings under more general conditions than those required for a one-dimensional signal.

This is in fact true, and in this chapter, we shall present new results on the unique specification of bandlimited, periodic, two-dimensional signals from zero crossings. The results are simpler to develop for periodic signals than for arbitrary signals since we can represent these signals as polynomials in a Fourier series representation and apply well-known results on polynomials from algebraic geometry. (The extension to nonperiodic signals will be presented in the next chapter.) In the first section, we shall first define this representation and present

the necessary results on polynomials. Following that, we shall develop a number of

different sets of conditions under which a signal is uniquely specified with zero

crossings. These results can be stated in a number of different forms since it is pos-

sible to uniquely specify a signal with zero crossings under a number of different

sets of constraints. We shall present the primary result in section 2 and then

develop a number of extensions for different types of signals in section 3. In sec-

tion 4, we consider the problem of sampling the zero crossing contours, that is, we

develop conditions under which a signal is uniquely specified with a finite set of

discrete points chosen from the zero crossing contours. In section 5, we extend

these results to signals of dimension higher than two.

3.1. Background

Consider a real, bandlimited, continuous-time, periodic signal $f(x,y)$ with

periods T_1 and T_2 in the x- and y- directions, respectively. We can express $f(x,y)$

as a polynomial using the Fourier series representation:

$$f(x,y) = \sum_{n_1} \sum_{n_2} F(n_1,n_2) \, W_1^{n_1} \, W_2^{n_2} \qquad (1)$$

where

$$W_1 = e^{j\frac{2\pi x}{T_1}}$$

$$W_2 = e^{j\frac{2\pi y}{T_2}}$$

The coefficients $F(n_1,n_2)$ are the Fourier series coefficients and represent the spec-

trum of $f(x,y)$. Since we are assuming $f(x,y)$ to be bandlimited, the sums in

equation (1) must be finite. The set of points (n_1,n_2) where $F(n_1,n_2)$ is not

constrained to be zero is referred to as the region of support of the spectrum. Since the Fourier coefficients $F(n_1, n_2)$ are nonzero for positive and negative values of n_1 and n_2, strictly speaking, $f(x, y)$ is not a polynomial in the variables W_1 and W_2 since it contains negative powers of n_1 and n_2. However, since $f(x, y)$ is bandlimited, $F(n_1, n_2)$ is nonzero only over some finite region in the (n_1, n_2) plane, and it is easy to modify equation (1) to relate $f(x, y)$ to a polynomial in W_1 and W_2. Assume $F(n_1, n_2) = 0$ outside the region $-N_1 \leq n_1 \leq N_1$, $-N_2 \leq n_2 \leq N_2$. Then we can write:

$$f'(x, y) = W_1^{N_1} W_2^{N_2} f(x, y) = \sum_{n_1=0}^{n_1=2N_1} \sum_{n_2=0}^{n_2=2N_2} F(n_1-N_1, n_2-N_2) W_1^{n_1} W_2^{n_2} \quad (2)$$

Although in the discussion that follows we shall refer to the representation of $f(x, y)$ as a Fourier series polynomial, it should be kept in mind that, strictly speaking, we are referring to the representation of the modulated signal $f'(x, y)$ in equation (2) as a polynomial.

Now that we have established a method of representing a signal as a polynomial, we will state a well-established result on polynomials in two variables which we will later use to develop our results on the unique specification of signals with zero crossings. We will state this result here without proof; the detailed proof is available in references [17] and [18], as well as a number of other texts on algebraic geometry.

Theorem 3.1. If $p(x,y)$ and $q(x,y)$ are two-dimensional polynomials of degrees r and s with no common factors, then there are at most rs distinct pairs (x,y) where:

$$p(x,y) = 0 \tag{3}$$

and

$$q(x,y) = 0$$

In this theorem, the degree of a polynomial in two variables is defined in terms of the sum of the degrees in each variable (for each term), that is, the degree of a two-dimensional polynomial $p(x,y)$ is equivalent to the degree of the one-dimensional polynomial $p(x,x)$. The rs distinct pairs (x,y) described in this theorem consist of rs points anywhere in the complex (x,y)-plane. Essentially, Theorem 3.1 places an upper bound on the number of points where two two-dimensional polynomials can both be zero if they do not have a common factor.

A stronger form of Theorem 3.1 is available which guarantees that the zero-sets of the polynomials intersect in exactly rs points, rather than simply stating an upper bound. This stronger result, referred to as Bezout's Theorem in algebraic geometry, requires including the multiplicity of intersections as well as points which lie "at infinity" (e.g., two parallel lines are considered to intersect in one point at infinity). Bezout's Theorem can be thought of as a generalization of the Fundamental Theorem of Algebra, which guarantees that an nth degree polynomial has exactly n roots provided multiplicity is included. To see the connection, consider applying Theorem 3.1 with $f(x,y) = y - h(x)$, where $h(x)$ is an nth-degree polynomial in x and let $g(x,y) = y$. The set of points where $f(x,y) = 0$ is a curve in the (x,y)-plane, and the set of points where $g(x,y) = 0$ is the x-axis. The points

of intersection (where $f(x,y) = g(x,y) = 0$) correspond to the n zeros of the polynomial $h(x)$. Theorem 3.1 guarantees that there will be at most $rs = n$ such points of intersection and Bezout's Theorem tells us there will in fact be exactly $rs = n$ points of intersection.

One interesting aspect of this theorem is that the number of points of intersection is in general greater than the number of coefficients or degrees of freedom in the polynomials. For example, a polynomial of degree N in each variable will have $(N+1)^2$ coefficients and one might expect the polynomial to be uniquely specified to within a scale factor with $(N+1)^2-1$ distinct points where it is zero. However, this is not the case since these polynomials will have an overall degree of $2N$ and two such polynomials will have $4N^2$ points of intersection, so that $(N+1)^2-1$ points cannot be guaranteed to uniquely specify the polynomial.

3.2. Primary Result

In this section, we will use the representation of signals as polynomials and the result on intersection of zero-sets of polynomials presented in the preceding section to establish our primary result cn the reconstruction of periodic signals from zero crossings. A number of extensions to this result will be presented in the following section.

To see how the results presented in the preceding section apply to the problem of unique specification of two-dimensional signals with zero crossings, consider a real, bandlimited, periodic signal $f(x,y)$ expressed as a polynomial in the Fourier

series representation in equation (2). We assume that there are some regions where $f(x,y)$ is positive and some regions where $f(x,y)$ is negative. These regions are separated from each other by a contour where $f(x,y) = 0$. If another signal $g(x,y)$ has the same zero crossing contours as $f(x,y)$, then there are an infinite number of points where both $f(x,y)$ and $g(x,y)$ are zero. We can then use Theorem 3.1 to show that $f(x,y)$ and $g(x,y)$ must have a common factor. If furthermore we know that $f(x,y)$ and $g(x,y)$ are irreducible when expressed as polynomials as in equation (2), then they must be equal to within a scale factor. Our result can be stated as follows:

Theorem 3.2. Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, doubly-periodic, bandlimited functions with sign $f(x,y)$ = sign $g(x,y)$, where $f(x,y)$ takes on both positive and negative values. If $f(x,y)$ and $g(x,y)$ are nonfactorable when expressed as polynomials in the Fourier series representation (2), then $f(x,y) = cg(x,y)$.

Proof: We will prove this result by starting with two signals $f(x,y)$ and $g(x,y)$ which satisfy the constraints of the theorem and showing that they must be equal to within a scale factor. Since we know that $f(x,y)$ takes on positive and negative values, there must be some region of the (x,y) plane where $f(x,y) > 0$ and another region where $f(x,y) < 0$. Since $f(x,y)$ is bandlimited and therefore continuous, the boundary between these regions is a contour where $f(x,y) = 0$. Since sign $f(x,y)$ = sign $g(x,y)$ for all (x,y), the same arguments also hold for $g(x,y)$. Thus, we have contours in the x,y plane where:

$$f(x,y) = g(x,y) = 0 \tag{4}$$

Also, if N_1 and N_2 are defined as in equation (2), we have:

$$W_1^{N_1} W_2^{N_2} f(x,y) = 0 \tag{5}$$

$$W_1^{N_1} W_2^{N_2} g(x,y) = 0$$

over these contours. Thus we have an infinite set of points where two polynomials in the variables W_1, W_2 are known to be zero. Thus, by Theorem 3.1, $f(x,y)$ and $g(x,y)$ must have a common factor. If furthermore, we assume that $f(x,y)$ and $g(x,y)$ are nonfactorable when expressed as polynomials in equation (2), then $f(x,y) = cg(x,y)$.

Note that in order to satisfy Theorem 3.1, it is not necessary to know the location of all the zero crossing contours; it is only necessary to know the location of a sufficient number of points along these contours. Thus any zero contour in the (x,y) plane is sufficient to uniquely specify the signal (since it contains an infinite number of points) even if the region where $f(x,y) < 0$ is very small. It is also possible to sample the zero crossing contours, i.e., to uniquely specify the signal with only a finite set of discrete points from the zero crossing contours. This possibility will be explored in more detail in section 3.4.

The fact that knowledge of all the zero contours in the (x,y) plane is not necessary to uniquely specify the signal allows us to extend this result to signals which are not periodic but are finite length. This extension is important since most signals encountered in practice are finite length. Consider the case where $f(x,y)$

- 20 -

is a finite segment of a periodic signal satisfying the constraints of Theorem 3.2. For example, if $f(x,y)$ represents one period of a bandlimited periodic function $\bar{f}(x,y)$:

$$\bar{f}(x,y) = \sum_{n_1} \sum_{n_2} f(x + n_1T_1, y + n_2T_2) \tag{6}$$

then it is possible to recover $f(x,y)$ from its zero crossings provided that $\bar{f}(x,y)$ satisfies the constraints of Theorem 3.2, even though $f(x,y)$ itself is not bandlimited. More generally, it is not necessary for the duration of $f(x,y)$ to be equal to one period of the corresponding periodic function. Thus, $f(x,y)$ can represent a finite segment of a variety of different periodic functions. In order for $f(x,y)$ to be uniquely specified by its zero crossings, we only need one periodic function containing $f(x,y)$ to be bandlimited. Specifically, let us state:

> **Theorem 3.3.** Let $f(x,y)$ and $g(x,y)$ be two-dimensional continuous functions defined over the same known region R of finite extent with sign $f(x,y) \equiv$ sign $g(x,y)$, where $f(x,y)$ takes on both positive and negative values. If $f(x,y) = f_p(x,y)$ in R and $g(x,y) = g_p(x,y)$ in R for any periodic, bandlimited functions $f_p(x,y)$ and $g_p(x,y)$ which are nonfactorable when expressed as polynomials in the form (2), then $f(x,y) = cg(x,y)$ for some positive constant c.

Proof: First we note that if $f_p(x,y)$ and $g_p(x,y)$ are not simply periodic replications of $f(x,y)$ and $g(x,y)$, then $f_p(x,y)$ and $g_p(x,y)$ may contain zero crossings that cannot be obtained from the zero crossings of $f(x,y)$ and $g(x,y)$. Thus, Theorem 3.3 does not quite follow directly from Theorem 3.2. However, unique specification in terms of zero crossings does not require knowledge of all the zero crossing contours; it requires only a specific number of points from these contours.

In this case, as long as $f(x,y)$ contains both positive and negative values, it will contain at least one zero crossing contour with an infinite number of points. Thus $f_p(x,y) = g_p(x,y) = 0$ at an infinite number of points, and using arguments taken from the proof of Theorem 3.2, $f_p(x,y) = cg_p(x,y)$. Since $f(x,y)$ and $g(x,y)$ are both known to be defined over the same region R and over R, $f(x,y) = f_p(x,y)$ and $g(x,y) = g_p(x,y)$, then $f(x,y) = cg(x,y)$.

Having established a set of conditions which guarantee that a signal is uniquely specified by some partial information, it is worthwhile to determine whether or not these conditions are likely to apply to a typical signal encountered in practice. First of all, we note that the irreducibility constraint is satisfied with probability one, since it has been shown that the set of reducible m-dimensional polynomials forms a set of measure zero in the set of all m-dimensional polynomials (for $m > 1$) [19] and that this set is an algebraic set [20]. Furthermore, we shall develop an extension to Theorem 3.2 in the next section which does not require the signals to be irreducible as Fourier series polynomials. The more restrictive constraint is that of requiring the signal to strictly bandlimited. Although signals encountered in practice are generally not strictly bandlimited, in many applications signals are commonly assumed to be bandlimited, and furthermore it is common to low-pass filter signals when necessary for particular processing techniques. Another conceivable difficulty with this result is that in some applications such as image processing, the signals are constrained to be positive and thus will not contain zero crossings. This

problem will be eliminated in the following section when we extend this result to include crossings of an arbitrary threshold instead of just zero crossings.

3.3. Extensions

Although Theorem 3.2 states a particular set of constraints under which a signal is uniquely specified with zero crossings, a number of different sets of constraints are possible. In this section, we shall introduce four extensions to Theorem 3.2. The first extension imposes most of its constraints on only one signal, as opposed to imposing a number of constraints on both signals $f(x,y)$ and $g(x,y)$. This result will be convenient to use when discussing algorithms for reconstruction from the zero crossings of $f(x,y)$. The second result generalizes the concept of zero crossings to include crossings of an arbitrary threshold. The third result extends Theorem 3.2 to include factorable signals by replacing the nonfactorability constraint with a constraint on each factor. The final result extends Theorem 3.2 to complex signals with spectral components constrained to one half-plane in the frequency domain. In this case, the signals are uniquely specified with the zero crossings of the real part of the signal.

Let us start by considering a case where it would be convenient to have a result which imposes most of its constraints on only one signal. Suppose a signal $f(x,y)$ is known to satisfy the constraints of Theorem 3.2, and we would like to develop an algorithm to recover $f(x,y)$ from its zero crossings. It would be convenient to have a set of constraints which guarantee that there are no other signals $g(x,y)$ with sign $f(x,y)$ = sign $g(x,y)$ whether or not $g(x,y)$ satisfies the

constraints of Theorem 3.2. While it is simple enough to guarantee that a recovered signal $g(x,y)$ is bandlimited, it is extremely difficult to guarantee that $g(x,y)$ is nonfactorable. Without this restriction, $g(x,y)$ might contain a positive factor, and thus we could have sign $f(x,y)$ = sign $g(x,y)$ but $f(x,y) \neq cg(x,y)$. To avoid this problem, we note that since multiplication by an additional factor in the spatial domain increases the degree of the Fourier series polynomial, the signal $g(x,y)$ would then have a larger bandwidth than $f(x,y)$. Thus, if the exact size of the bandwidth of $f(x,y)$ is known, then this information, together with the zero crossings of $f(x,y)$, is sufficient to uniquely specify $f(x,y)$. Specifically:

Theorem 3.4. Let $f(x,y)$ and $g(x,y)$ be real two-dimensional signals with $F(n_1,n_2) = 0$ and $G(n_1,n_2) = 0$ outside the region $-N_1 \leq n_1 \leq N_1$, $-N_2 \leq n_2 \leq N_2$, with sign $f(x,y) \equiv$ sign $g(x,y)$. If $F[N_1, N_2] \neq 0$, $f(x,y)$ takes on both positive and negative values, and $f(x,y)$ is nonfactorable when expressed as a polynomial in equation (2), then $f(x,y) = cg(x,y)$ for some positive constant c.

Proof: Following the proof of Theorem 3.2, we know that since sign $f(x,y)$ = sign $g(x,y)$, $f(x,y)$ and $g(x,y)$ must contain a common factor. Since $f(x,y)$ is nonfactorable, then if $f(x,y) \neq cg(x,y)$, $g(x,y) = f(x,y) h(x,y)$ for some real periodic bandlimited function $h(x,y)$. If $h(x,y)$ is not a constant, then $H(k,l) \neq 0$ for some (k,l) such that either $k>0$ or $l>0$. Since $F(N_1,N_2) \neq 0$, we know that $G(N_1+k,N_2+l) \neq 0$, violating the constraints of the theorem. Thus $h(x,y)$ must be a constant and $g(x,y) = cf(x,y)$.

Thus, if a signal $f(x,y)$ satisfies the constraints of Theorem 3.4, then we know that there are no other signals $g(x,y)$ with the same bandwidth and the same zero crossings. Therefore, in a somewhat more general sense than was possible with Theorem 3.2, we can say that a signal satisfying the constraints of Theorem 3.4 is uniquely specified to within a scale factor with its zero crossings and the known bandwidth.

It is also possible to generalize Theorems 3.2 and 3.4 to allow unique specification with crossings of an arbitrary threshold rather than just zero crossings. The possibility of developing such an extension provides an important distinction between our results and earlier results such as Logan's theorem which cannot be directly extended to include crossings of an arbitrary threshold. In fact, we can permit a more generalized form of threshold crossings by allowing crossings of an arbitrary periodic function. This form of threshold crossings can be considered as a generalization of sine-wave crossings as used in other work such as [16]. also generalizes previous work with sine-wave crossings mentioned earlier. Using this generalized form, we can extend Theorem 3.2 as follows:

Theorem 3.5. Let $f(x,y)$, $g(x,y)$, $h(x,y)$ be real, two-dimensional, doubly-periodic, bandlimited functions with sign $(f(x,y)-h(x,y))$ = sign $(g(x,y)-h(x,y))$, where $f(x,y)-h(x,y)$ takes on both positive and negative values. If $f(x,y)-h(x,y)$ and $g(x,y)-h(x,y)$ are nonfactorable when expressed as polynomials in the Fourier series representation (2), then $f(x,y)-h(x,y)$ = $c(g(x,y)-h(x,y))$.

Proof: Since $f(x,y)$, $g(x,y)$, and $h(x,y)$ are periodic and bandlimited, the functions $f'(x,y) = f(x,y) - h(x,y)$ and $g'(x,y) = g(x,y) - h(x,y)$ are also periodic and bandlimited, and according to the theorem statement, will satisfy the constraints of Theorem 3.2. Thus $f'(x,y) = cg'(x,y)$ and $f(x,y) - h(x,y) = c(g(x,y) - h(x,y))$.

A further extension to Theorem 3.2 involves replacing the nonfactorability constraint with a constraint on each factor. Let us express $f(x,y)$ as a product of real factors $f_i(x,y)$ (factors which are real for real values of x,y). Observe that if $f_i(x,y) = 0$ for any i, then $f(x,y) = 0$; similarly, if $f(x,y) = 0$, then at least one of the factors $f_i(x,y)$ must be zero. Thus, if each factor contributes a set of zero crossing contours, each factor will be uniquely specified by its own zero crossing contours, and thus we can develop a set of conditions under which $f(x,y)$ will be uniquely specified by the complete set of zero crossing contours. Specifically, we state:

Theorem 3.6. Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, doubly-periodic, bandlimited functions with sign $f(x,y) =$ sign $g(x,y)$. Consider the (polynomial) factorization of $f(x,y)$ and $g(x,y)$ into real factors which are irreducible over the set of real factors. If each of these factors has multiplicity one and takes on both positive and negative values, then $f(x,y) = cg(x,y)$ for some positive constant c.

Proof: Recall from the proof of Theorem 3.2 that if $f(x,y)$ and $g(x,y)$ have common zero contours then they must have a common factor. We will assume that

$f(x,y) \neq cg(x,y)$ and attempt to reach a contradiction. For convenience, let us assume that there is some irreducible factor of $f(x,y)$ which is not a factor of $g(x,y)$. First of all, note that if this factor, denoted $f_i(x,y)$, is complex, then $f_i^*(x,y)$ will also be a factor of $f(x,y)$ and thus $f(x,y)$ will contain a real factor $f_i(x,y)f_i^*(x,y)$ which is nonnegative everywhere, violating the constraints of the theorem. Thus, the factor $f_i(x,y)$ must be real, and since according to the theorem hypothesis, it has both positive and negative values and has multiplicity one, then we must have sign $f(x,y) \neq$ sign $g(x,y)$ for some values of (x,y), and we have reached a contradiction. Thus, there cannot be any factor of $f(x,y)$ which is not a factor of $g(x,y)$ and thus, $f(x,y) = cg(x,y)$.

This theorem could also be stated in a slightly different manner by considering all possible factorizations of $f(x,y)$ and $g(x,y)$ rather than one particular factorization. In this case, the requirement would be that every possible factor of $f(x,y)$ and $g(x,y)$ must take on both positive and negative values. The multiplicity constraint is then unnecessary since if a factor $f_i(x,y)$ occurs with multiplicity two (or higher), then there will also be a factor $f_i^2(x,y)$ which is nonnegative for real values of x,y and violates the constraints of the theorem.

Theorem 3.6 can be easily modified to permit finite length signals or to permit crossings of an arbitrary threshold. To modify this result for finite length signals, we simply require that the finite length signals be a finite segment of a periodic signal satisfying the constraints of Theorem 3.6. To modify this result for crossings of

an arbitrary threshold α, we simply require that the signals $f(x,y) - \alpha$ and $g(x,y) - \alpha$ satisfy the constraints of Theorem 3.6.

As our final extension, we will modify Theorem 3.2 for complex signals. We will assume that the zero crossings of the real parts of the signals are available and that the signal is constrained so that the real part of the signal is sufficient to specify the imaginary part. This is possible if the spectrum has a region of support over a nonsymmetric half-plane, i.e., if $F[n_1,n_2] \neq 0$, then $F[-n_1,-n_2] = 0$ except at $n_1 = n_2 = 0$, where we require $F[0,0]$ to be real. Our result can be stated as follows:

> **Theorem 3.7.** Let $f(x,y)$ and $g(x,y)$ be complex two-dimensional, doubly-periodic, bandlimited functions with sign Re $\{f(x,y)\}$ = sign Re $\{g(x,y)\}$, where Re $\{f(x,y)\}$ takes on both positive and negative values. Assume that $F[n_1,n_2]$ and $G[n_1,n_2]$ have support over the same nonsymmetric half plane (as stated above). If Re $\{f(x,y)\}$ and Re $\{g(x,y)\}$ are nonfactorable when expressed as polynomials in the Fourier series representation (2), then $f(x,y) = cg(x,y)$.

Proof: Since Re $\{f(x,y)\}$ has Fourier series coefficients

$$\frac{F[n_1,n_2] + F^*[-n_1,-n_2]}{2},$$

then since $f(x,y)$ is bandlimited, Re $\{f(x,y)\}$ is bandlimited, and similarly for $g(x,y)$. Since Re $\{f(x,y)\}$ and Re $\{g(x,y)\}$ satisfy the constraints of Theorem 3.2, Re $\{f(x,y)\}$ = c Re $\{g(x,y)\}$. Thus we have

$$\frac{F[n_1,n_2] + F^*[-n_1,-n_2]}{2} = c \frac{G[n_1,n_2] + G^*[-n_1,-n_2]}{2}.$$ Since we know that if $F[n_1,n_2] \neq 0$, then $F[-n_1,-n_2] = 0$ and $G[-n_1,-n_2] = 0$, we must have $F[n_1,n_2] = cG[n_1,n_2]$ and $f(x,y) = cg(x,y)$.

3.4. Sampling of Zero Crossing Contours

As mentioned earlier, it is possible to state Theorem 3.2 in a slightly different way so that it is possible to uniquely specify a signal with a finite set of discrete zero crossing points, essentially allowing us to sample the zero crossing contours. This result is important since any practical algorithm for recovering signals from zero crossing information can only make use of a finite number of zero crossing points. In this section, we develop this result and discuss a number of related issues.

Before proceeding, let us first emphasize that we are referring to sampling the zero crossing locations *along a zero crossing contour* and we are not referring to sampling of the sign of the original signal, that is, recording the sign of the signal at each point on a predetermined grid. This is distinct from the type of sampling used in many signal processing problems where signals are specified with samples over a particular grid. The difficulty with sampling the sign information is that the information necessary to apply our results to uniquely specify a signal is contained in the exact location of the zero crossings and this information is lost when (sign $f(x,y)$) is sampled. From another point of view, we can say that a finite set of samples of (sign $f(x,y)$) contains a finite number of *bits* of information and thus cannot be expected to uniquely specify a signal to infinite precision. This is distinctly different from typical sampling problems where each sample is of (theoretically) infinite precision and thus does not contain a finite number of bits of information. Note, however, that we are strictly referring to theoretical sampling problems; in

practical applications, of course, signals are generally represented with a finite number of bits, and it may be possible for a signal to be represented to sufficient accuracy with a finite set of samples of (sign $f(x,y)$).

Since Theorem 3.1 specifies the number of points where two two-dimensional polynomials can both be zero, we can use this theorem to establish that a particular number of arbitrarily-chosen zero crossing points is guaranteed to be sufficient for unique specification. We shall also show that this number of points may not be *necessary* for unique specification; in particular, if the zero crossing points are not chosen arbitrarily but are chosen in some particular way, a smaller set of zero crossing points can be sufficient for unique specification.

The exact number of zero crossing points sufficient for unique specification depends on the size and shape of the spectrum of the signal. We will state our results in terms of rectangular spectra since these shapes are common in applications and are straight-forward to understand. The result could be easily modified for spectra of different shapes or could be applied directly to a problem involving a different spectrum by simply assuming a rectangular region large enough to enclose the actual region. If reference to a region of support $R(N)$ specifies that the spectrum is zero outside the region $-N \leq n_1, n_2 \leq N$, then we can state:

Theorem 3.8. Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, doubly-periodic, bandlimited functions with a spectrum with region of support $R(N)$. If $f(x,y)$ and $g(x,y)$ are nonfactorable when expressed as polynomials in the Fourier series representation (2), and $f(x,y) = g(x,y) = 0$ at more than $16N^2$ distinct points in one period, then $f(x,y) = cg(x,y)$ for some real constant c.

Proof: Recall that the proof of Theorem 3.2 requires stating that two polynomials $W_1^N W_2^N f(x,y)$ and $W_1^N W_2^N g(x,y)$ are equal to within a scale factor given that they are both zero at an infinite number of points. (Let $N = N_1 = N_2$ in the proof of Theorem 3.2.) In the case of Theorem 3.8, we know that $W_1^N W_2^N f(x,y) = W_1^N W_2^N g(x,y) = 0$ at more than $16N^2$ points in one period, that is, at more than $16N^2$ distinct values of the variables (W_1, W_2). These polynomials are of degree $4N$ and thus, by Theorem 3.1, can have at most $16N^2$ common zeros. Thus, $W_1^N W_2^N f(x,y) = cW_1^N W_2^N g(x,y)$ and the theorem follows.

This result is important since any practical algorithm designed to recover signals from zero crossing information can only make use of a finite number of zero crossing points. In addition, as discussed in the development of Theorem 3.4, it is necessary to have a result which only requires one signal to be irreducible. We can now combine both of these ideas into one result which precisely states the conditions we will need to impose later:

Theorem 3.9. Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, doubly-periodic, bandlimited functions with bandwidth with region of support $R(N)$. If $f(x,y)$ is nonfactorable, when expressed as polynomials in the Fourier series representation (2), $F(N,N) \neq 0$, and $f(x,y) = g(x,y) = 0$ at more than $16N^2$ distinct points, then $f(x,y) = cg(x,y)$ for some real constant c.

Proof: The proof is identical to the proof of Theorem 3.8, with the references to Theorem 3.2 replaced by references to Theorem 3.4.

Another extension one might consider is to permit sampling of zero crossing contours for factorable signals. However, if a signal is factorable and satisfies the constraints of Theorem 3.6 (but not the constraints of Theorem 3.2), then a finite set of zero crossings is not guaranteed to uniquely specify the signal since these zero crossings may all correspond to the same factor. However, a result similar to Theorem 3.8 is easily developed for factorable signals by constraining the set of zero crossings to include N_i zero crossings of the ith factor, where each N_i is chosen by applying Theorem 3.8 to each factor individually. It would also be necessary to know which zero crossings correspond to the same factor in order to successfully recover a signal under these conditions. These constraints would be difficult to verify in an actual signal reconstruction problem, but would be easier to verify in a situation such as a filter design procedure [3] which involves first specifying the zeros of the response and then generating the remainder of the response with signal reconstruction techniques.

Note that the number of arbitrarily-chosen zero-crossing points sufficient to uniquely specify a signal is somewhat greater than the number of unknowns in the

signal (i.e., the number of spectral components). Specifically, if a real signal $f(x,y)$ has a spectrum with support over $R(N)$, then by the symmetry properties of the Fourier transform, $F[n_1,n_2]$ has $2N^2 + 2N + 1$ independent points, which in general, are complex (except $F[0,0]$). Since one complex point corresponds to two real unknowns (the real part and the imaginary part), $f(x,y)$ contains $4N^2 + 4N + 1$ real unknowns or degrees of freedom. According to Theorem 3.8, $f(x,y)$ is uniquely specified with p zero crossing points if $p > 16N^2$, approximately four times the number of unknowns.

Although Theorem 3.8 states that a particular number of zero crossing points is sufficient for unique specification, it does not state that this number of points is *necessary* for unique specification. In particular, as we show next, it is possible to specify a signal consisting of s spectral components with $s-1$ zero crossing points if the zero crossing points are not chosen arbitrarily but are specifically chosen so that they uniquely specify the signal. To establish this result, note that we can write a set of linear equations of the form:

$$\sum_{n_1} \sum_{n_2} F[n_1,n_2]\, e^{j\frac{2\pi x_i n_1}{T_1}}\, e^{j\frac{2\pi y_i n_2}{T_2}} = 0 \qquad (7)$$

where each equation uses a different pair of points (x_i, y_i) on a zero crossing contour, i.e., for which the equality is known to hold. If we assume that $f(x,y)$ satisfies the constraints of Theorem 3.9, then $F[N,N] \neq 0$, so we can substitute $F[N,N] = 1$ and obtain a non-zero solution. Thus, if $F[n_1,n_2]$ consists of s points, these equations contain $s-1$ (complex) unknowns. Although we have not

shown that $s - 1$ equations of this form are guaranteed to have a unique solution, we know from Theorem 3.9 that if a sufficient number of equations is used (say, p equations, where $p > 16N^2$) then these equations are guaranteed to have a unique solution. If we have p equations in $s - 1$ unknowns and $p > s - 1$, then some of the equations must be dependent and can be eliminated. Thus, it is possible (theoretically) to find $s - 1$ independent equations from the set of p equations, and thus the corresponding $s - 1$ zero crossing points are sufficient to uniquely specify $f(x,y)$. This result, however, does not suggest a practical algorithm for choosing the $s - 1$ zero crossing points so that these points uniquely specify the signal.

3.5. Higher-Dimensional Signals

Although up to this point we have been primarily concerned with the reconstruction of two-dimensional signals from zero crossings, it is also possible to recover signals of dimension higher than two from this information. One approach to this problem is to take two-dimensional slices of the higher-dimensional signal and recover a two-dimensional signal whose spectrum is a projection of the spectrum of the original signal and apply the results developed earlier. This is analogous to one method used by Hayes [21] to develop results on the unique specification of two-dimensional signals with Fourier phase information. Another approach, the one we shall follow here, is to extend the results developed earlier directly to m-dimensional problems by applying an m-dimensional extension of Theorem 3.1. Such an extension is available, although the result is not quite as straight-forward as Theorem 3.1. In general, it is not possible to state that two polynomials in an

arbitrary number of variables have common zeros at a finite number of points. However, it is possible to characterize the intersection of two surfaces, each described by a polynomial equation, as another surface with a specified dimension and degree.

In the terminology used in algebraic geometry, a *variety* is defined as the intersection of the zero-sets of one or more polynomials, corresponding to our intuitive notion of surfaces. The dimension of a variety is the number of independent variables present. In a three-dimensional space, a plane (or a sphere, paraboloid, etc.) is defined by one equation and thus has two independent variables and dimension two. A line is defined by two independent equations (the intersection of the zero-sets of two polynomials) and thus has one independent variable and dimension one. For a variety defined by one polynomial equation, the degree is equivalent to the degree of the polynomial defining it. A finite set of points is considered to have dimension zero, and in this case, the degree is defined as the number of points in the set.

In order to develop results on the unique specification of higher-dimensional signals with zero crossings, we need to characterize the intersection of varieties of arbitrary dimension. We begin by stating a result on the dimension of the intersection as given in [22]:

Theorem 3.10. Let $f(x)$ be an irreducible polynomial in the variables $x_1, x_2, \ldots, x_m$ and let V_f denote the $(m-1)$-dimensional irreducible variety $\{x: f(x) = 0\}$ in the space C^m. Let V_g denote any variety of dimension r in C^m. If V_g is irreducible and $V_g \not\subset V_f$, then every component of the intersection $V_f \cap V_g$ has dimension $r-1$.

In fact, it is possible to characterize the intersection of two varieties in more detail. In particular, under the conditions stated in Theorem 3.10 it is also possible to show that $V_f \cap V_g$ will have dimension $r-1$ and degree $(\deg V_f) \times (\deg V_g)$, provided that when counting the degree of intersections we consider the multiplicity of intersections and also include intersections which lie "at infinity". (In this case, two parallel planes can be considered to intersect in a line at infinity.) This stronger result, generally referred to as Bezout's Theorem in m dimensions, is given in [22,23,24], but will not be stated precisely here since a precise statement would require the introduction of a number of additional concepts from algebraic geometry and Theorem 3.10 is sufficient to develop our results.

Before proceeding further, it is worthwhile to consider a few examples in order to understand these results on the intersection of varieties. First of all, consider the case $m = 2$. Curves in a plane have dimension 1 and their intersection has dimension 0, that is, their intersection is a finite set of points. The number of points in this set (the degree of the intersection) is equal to the product of the degrees of the curves, as we found in Theorem 3.1. The case $m = 3$ corresponds to the intersection of surfaces in three-dimensional space. For example, two ellipsoids each have dimension 2 and degree 2. Their intersection consists of two ellipses, a one-

dimensional curve of degree 4 which is reducible into two curves of degree 2.

Theorem 3.10 can be applied to the problem of unique specification with zero crossings in the same way Theorem 3.1 was used to develop Theorem 3.2. We will show that if two signals have the same sign for all (x,y), then their zero crossing surfaces must intersect in a surface of dimension $m-1$; thus, by Theorem 3.1, the two surfaces must be identical and the two signals must be equal to within a scale factor. Specifically, we state the following theorem:

Theorem 3.11. Let $f(x)$ and $g(x)$ be real, m-dimensional, periodic, bandlimited functions with sign $f(x)$ = sign $g(x)$, where $f(x)$ takes on both positive and negative values. If $f(x)$ and $g(x)$ are nonfactorable when expressed as polynomials in the Fourier series representation (2), then $f(x) = cg(x)$.

Proof: Following the proof of Theorem 3.2 but extending it to arbitrary dimensions, from the fact that sign $f(x)$ = sign $g(x)$, we have a set of points where:

$$W_1^{N_1} W_2^{N_2} \cdots W_M^{N_M} f(x) = 0 \qquad (8)$$

and

$$W_1^{N_1} W_2^{N_2} \cdots W_M^{N_M} g(x) = 0,$$

that is, a set of points where two polynomials in the variables $W = W_1, W_2, \cdots W_M$ are known to be zero. The sets of all points (real and complex) $V_f = \{(W):f(W) = 0\}$ and $V_g = \{(W):g(W) = 0\}$ are $(m-1)$-dimensional varieties. By Theorem 3.10, if V_f and V_g are not identical, then $V_f \cap V_g$ must have dimension $m-2$. However, we know $V_f \cap V_g$ includes surfaces separating

regions where $f(x) > 0$ from regions where $f(x) < 0$. Consider two-dimensional slices of the signals $f(x)$ and $g(x)$ defined by $x_1 = \alpha_1$, $x_2 = \alpha_2$, ... $x_{m-2} = \alpha_{m-2}$, where the α_k are chosen so that each successive slice contains positive and negative values of $f(x)$ and $g(x)$ and thus contains some but not all of the zero crossings of $f(x)$ and $g(x)$. (If necessary reorder the x_k to make this possible.) This two-dimensional slice contains positive and negative values and thus contains a zero crossing contour consisting of an infinite number of points. From the point of view of varieties, taking a slice of a signal corresponds to forming the intersection of the original zeros with series of hyperplanes P_k: $W_K = e^{j\frac{2\pi \alpha_k}{T_k}} = \beta_k$. By Theorem 3.10, forming the intersection of $V_f \cap V_g$ with each hyperplane reduces the dimension by one. (In general the intersection may be reducible, but we can apply Theorem 3.10 to each irreducible component as long as no irreducible component is completely contained in the next slice.) We thus find that the resulting variety $V_f \cap V_g \cap P_1 \cap P_2 \cdots \cap P_{m-2}$ has dimension zero, and thus can only consist of a finite number of points. Thus, we have reached a contradiction, and $V_f \cap V_g$ must have dimension $m - 1$, $V_f = V_g$, and $f(x) = cg(x)$.

It is also possible to develop a number of extensions to Theorem 3.11 as we did for Theorem 3.2. In particular, Theorems 3.3, 3.4, 3.5, and 3.6 can be extended to the m-dimensional case. These results allow us to generalize Theorem 3.11 to permit finite length signals, constaints on one signal, threshold crossings,

- 38 -

and factorable signals. Specifically, these results are identical to Theorems 3.3 through 3.6 except that signals of any dimension $m \geq 2$ are permitted. These extensions can be derived by using the proof of the corresponding extension to Theorem 3.2 and replacing all references to Theorems 3.1 and 3.2 with references to Theorems 3.10 and 3.11.

A somewhat more difficult problem arises when we attempt to extend Theorem 3.11 to the case where only a finite set of zero crossing points are available. While in the two-dimensional case, a result is available which allows us to state that two curves may intersect in only a finite number of points, our results in the m-dimensional case inherently involve surfaces with an infinite number of points. Thus it is not possible to say that N zero crossing points are always sufficient to specify an m-dimensional signal for any finite value of N. However, it is possible to say that when the chosen points are independent in the sense that they do not lie on an $(m-2)$-dimensional contour of the appropriate degree, then these points are sufficient to uniquely specify the signal.

CHAPTER 4

UNIQUE SPECIFICATION OF NONPERIODIC SIGNALS

In this chapter, we extend the results developed earlier for periodic signals to the case of nonperiodic signals.* This problem is more difficult mathematically since it is in general not possible to express an arbitrary signal as a polynomial in a Fourier series representation. Nevertheless, it is possible to find applicable results on the intersection of zero sets of functions, this time from the theory of analytic sets. We will begin by developing the necessary mathematical background and then present the main result and a number of extensions.

4.1. Background

In this section, we define the notation and terminology to be used in the remainder of this chapter. In addition, mathematical results to be used later are presented here.

A two-dimensional complex-valued function (denoted $f(s,w)$) is said to be holomorphic if it is holomorphic (or analytic) in each variable separately. A function holomorphic for all finite values of s and w is called entire. We shall be primarily concerned with entire functions of exponential type (EFET) (see [5] for a review of the properties of EFETs in one variable, and [26,27] for EFETs in several variables). These functions are constrained to have at most an exponential

---

*The work presented in this chapter was performed in collaboration with S. Shitz and also appears in reference [25].

growth rate in any direction in complex space. As is well known [28,29], any bandlimited function of real variables can be uniquely extended to complex space as an EFET. (We will use the notation $f(x,y)$ to denote a function of real variables and $f(s,w)$ to denote its extension to complex variables.) This statement applies for a wide variety of common definitions of bandlimited functions. For finite energy signals, the Fourier transform will exist and any bandlimited signal will have a Fourier transform with a compact region of support. For bounded signals (with possibly infinite energy), alternate definitions of bandlimitation are possible by using the Fourier-Stieltjes transform or the so-called 2-transform (see [5] for definitions and for other possible definitions of bandlimitation). It is also possible to use a more general definition of bandlimitation derived from the theory of generalized functions or distributions. This definition requires the spectral distribution as defined in [28] to have compact support. Unless otherwise noted, the results presented in this chapter apply to this more general type of bandlimitedness, although in most practical applications the usual Fourier or Fourier-Stieltjes definition will apply.

Entire functions can be characterized in terms of their complex zeros much like polynomials (see [27] for a precise characterization). For either polynomials or entire functions, the represention of a function in terms of zeros requires both the real and complex zeros, not just the real zeros (zero crossings). However, there are some important differences between polynomials and entire functions since it is possible to have entire functions which are not constant yet still have no real or com-

plex zeros (for example, the function e^s is nonzero for all real or complex values of s). Thus if the set of complex zeros of an entire function is known, then the entire function may not be known even to within a constant due to the possibility of positive factors. However, it is known that the only positive EFETs are exponentials, and these can be eliminated by placing restrictions on the growth rate of the function. Such restrictions are often implicit in the definition of a bandlimited function. For example, if the Fourier transform definition of a bandlimited signal is used, then the signals are assumed to be finite energy. A more general class of one-dimensional bandlimited functions is characterized precisely in [5] by developing a subset of EFETs referred to as B-functions. The class of B-functions includes the set of bandlimited signals under a number of common definitions of bandlimitation (e.g., Fourier or Fourier-Stieltjes), as well as including a class of other signals with similar properties but which do not possess a Fourier (or similar) transform. B-functions are known to satisfy a number of different growth restrictions on the real axis which are given in [5] and can be used to eliminate the possibility of exponential factors. For functions of several variables, the Paley-Wiener-Schwartz theorem [29], which states that a function with a spectral distribution with compact support has at most polynomial growth in any direction in the real plane (or space), can be used to eliminate the possibility of exponential factors.

Because of the possibility of nonconstant positive factors in entire functions, it is common to exclude such factors when considering the factorization of entire functions into irreducible factors in the same way constants are excluded when con-

sidering the factorization of polynomials. Specifically, let V_f and V_g denote the set of real or complex zeros of $f(s,w)$ and $g(s,w)$, respectively, i.e., $V_f = \{(s,w): f(s,w) = 0\}$ and $V_g = \{(s,w): g(s,w) = 0\}$. The function $f(s,w)$ will then be referred to as irreducible if it cannot be expressed as $f(s,w) = g(s,w)h(s,w)$ where $g(s,w)$ and $h(s,w)$ are entire functions and V_g and V_h are both nonempty sets. (This definition is also used in [30].) Note that if h is an entire function which never vanishes (such as e^s) then f can still be irreducible in the sense defined above although $f = g \times h$.

We shall also use the term analytic set, defined as the intersection of the zero sets of one or more holomorphic functions [26, 27]. For example, V_f and V_g as defined above are analytic sets, as is $V_f \cap V_g$. An irreducible analytic set is an analytic set which cannot be expressed as the union of two distinct analytic sets. For example, if $f(s,w)$ is irreducible (as defined above), then V_f is an irreducible analytic set. If $f(s,w)$ is reducible, it can be expressed as $f(s,w) = g(s,w)h(s,w)$ and V_f can be expressed as $V_f = V_g \cup V_h$. An irreducible analytic set is also referred to as an analytic surface [26].

In this chapter, analytic sets play the same role as the zero sets of polynomials in the preceeding chapter. Thus, to develop results on reconstruction of signals from zero crossings, we will need a result from the theory of intersection of analytic sets:

Theorem 4.1. [26] Two surfaces analytic over a closed bounded region D intersect in at most a finite number of points in D. Two surfaces analytic over all of C^2 coincide if they have in common some sequence of points along with their limit point.

This theorem is similar to Theorem 3.1 but it only states that the number of points of intersection in a closed region must be finite; it does not specify the number of points of intersection. This is because in general EFETs can be thought of as "infinite-order polynomials" and thus will intersect in an infinite number of points. Nevertheless, it is still possible to constrain the number of points of intersection to be finite over any region. This fact allows us to apply the theory of analytic sets to the problem of unique specification with zero crossings.

4.2. Primary Result

In this section, we will apply the mathematical results stated in the preceeding section to the problem of unique specification with zero crossings. Theorem 4.1 allows us to develop our result on reconstruction from zero crossings in a straight-forward way. Note that if two irreducible signals $f(x,y)$ and $g(x,y)$ have identical zero crossing contours then the sets V_f and V_g must intersect in curves (at the zero crossing contours). Since these curves contain an uncountably infinite number of points in a finite region, by applying Theorem 4.1 we can show that the sets V_f and V_g must be identical. Then we know that f and g must be equal to within multiplication by an EFET which never vanishes, that is, by an exponential factor. This possibility can be eliminated by placing restrictions on the rate of growth of the function, as mentioned earlier. Specifically, let us state (see section 4.4 for

- 44 -

proof):

> **Theorem 4.2.** Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, bandlimited signals whose complex extensions are irreducible as entire functions in the sense defined in the previous section. If $f(x,y)$ takes on positive and negative values in a closed bounded region $D \subset R^2$ and $\text{sign} f(x,y) = \text{sign} g(x,y)$ for all values of (x,y) in D, then $f(x,y) = cg(x,y)$ for some real positive constant c.

Note that in this theorem it is not necessary for the zero crossings of $f(x,y)$ and $g(x,y)$ to be identical for all values of (x,y); it is sufficient for the signals to have one zero crossing contour in common. This fact allows us to apply this theorem to signals which are finite length are thus not strictly bandlimited. If the finite length signal represents a finite segment of some bandlimited function, then we can apply Theorem 4.2 by considering the region D to be the region of support of the function. Specifically, if $f(x,y)$ and $g(x,y)$ are finite length segments of the bandlimited signals $\hat{f}(x,y)$ and $\hat{g}(x,y)$, $\text{sign} f(x,y) = \text{sign} g(x,y)$, and $f(x,y)$ contains sign changes, then $f(x,y) = cg(x,y)$. This is similar to a result presented in chapter 3 which allows finite length signals to be uniquely specified by zero crossings if their periodic replications satisfy appropriate constraints. The result presented here is less restrictive since it does not require the underlying bandlimited function to be periodic.

While in chapter 3 we argued that the set of reducible polynomials has measure zero, the same statement does not apply to reducible EFET's. It is possible, however, to make a number of comments concerning the reducibility of EFET's. Although the only one-dimensional EFET's which are irreducible are of the form

$f(s) = e^{as+b}(s-c)$, this is not the case in two dimensions. Although we cannot precisely characterize the likelihood that a two-dimensional EFET is irreducible, it is commonly assumed that loosely speaking, most two-dimensional EFETs are irreducible [30, 31, 32]. However, even if it could be shown that in some statistical sense "almost all" EFETs are irreducible, there are some important examples of functions which are reducible. One example occurs if the two-dimensional bandlimited function can be expressed as a bandlimited function of only one variable, as is the case for circularly symmetric functions. Another example occurs if the function is separable and can thus be expressed as a product of two bandlimited functions, one in each variable. In the next section, we shall extend Theorem 4.2 to include factorable signals.

4.3. Extensions

Although Theorem 4.2 stated a number of conditions under which a signal is uniquely specified with its zero crossings, it is also possible to develop a number of variations or extensions of this result. These extensions are similar to those presented in chapter 3 for the case of periodic signals, but sufficiently different mathematically so that it is worthwhile to present these results in detail.

We begin by developing a result similar to Theorem 3.4 which allows us to impose the irreducibility constraint on only one signal. In Theorem 3.4, we imposed this constraint by requiring the signal to have a non-zero spectral component at the band edge. For the case of nonperiodic signals, we will require that the signal be bandlimited to a particular region B but no smaller region.

Specifically, (see section 4.4 for proof):

> **Theorem 4.3.** Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, signals and bandlimited to a region B but no smaller region. Let $f(s,w)$ be irreducible in the sense defined in the previous section. If $f(x,y)$ takes on positive and negative values in a closed bounded region $D \subset R^2$ and $\text{sign} f(x,y) = \text{sign} g(x,y)$ for all values of (x,y) in D, then $f(x,y) = cg(x,y)$ for some positive real constant c.

It is also possible to generalize the results presented above to a broader definition of zero crossings. In particular, it is possible to develop a result similar to Theorem 4.2 which allows the signals to be specified by crossings of an arbitrary threshold rather than simply zero crossings. As we found in chapter 3, it is also possible to allow the threshold to vary across the signal. While for the case of periodic signals we assumed that the threshold could be described as a periodic function, the result to be developed here allows the threshold to be specified by an arbitrary bandlimited function. In particular, let us state (see section 4.4 for proof):

> **Theorem 4.4.** Let $f(x,y)$, $g(x,y)$, and $h(x,y)$ be real, two-dimensional, bandlimited signals, where $f(s,w)-h(s,w)$ and $g(s,w)-h(s,w)$ are irreducible in the sense defined in the previous section. If $f(x,y) - h(x,y)$ takes on positive and negative values in a closed bounded region $D \subset R^2$ and $\text{sign}[f(x,y) - h(x,y)] = \text{sign}[g(x,y) - h(x,y)]$ for all values of (x,y) in D, then $f(x,y)-h(x,y) = c(g(x,y)-h(x,y))$ for some positive real constant c.

Another extension to Theorem 4.2 which we will develop here is to allow reducible signals. This extension is important since, as mentioned earlier, we can-

not precisely state the likelihood that a signal is irreducible. The reasoning used to develop this result is similar to the reasoning used in chapter 3 to develop a similar result for periodic signals. Let $f(s,w)$ denote the complex extension of a signal $f(x,y)$ and consider the factorization of $f(s,w)$ into real factors $f_i(s,w)$ (factors which are real for real values of s,w) which are irreducible over the set of real factors. We will assume that the number of such factors is finite. Observe that if $f_i(s,w) = 0$ for any i, then $f(s,w) = 0$; similarly, if $f(s,w) = 0$, then at least one of the factors $f_i(s,w)$ must be zero. Thus if each factor contributes a set of zero crossing contours, each factor will be uniquely specified by its own zero crossing contours, and thus we can develop a set of conditions under which $f(x,y)$ will be uniquely specified by its complete set of zero crossing contours. These conditions can be stated as follows (see section 4.4 for proof):

Theorem 4.5. Let $f(x,y)$ and $g(x,y)$ be real, two-dimensional, bandlimited signals. If $f(s,w)$ and $g(s,w)$ can be factored into a finite number of real irreducible factors (as described above), and if each factor of $f(s,w)$ and $g(s,w)$ has multiplicity one and takes on positive and negative values in a closed bounded region $D \subset R^2$ and $\text{sign} f(x,y) = \text{sign} g(x,y)$ for all values of (x,y) in D, then $f(x,y) = cg(x,y)$ for some positive real constant c.

This result allows us to argue that a broad class of bandlimited two-dimensional signals are uniquely specified with zero crossings without relying on assumptions of nonfactorability. In particular, it can be shown [25] that almost all signals in a class of bandlimited Gaussian random fields will contain zero crossings. Thus, if each factor is chosen in a random way, the resulting signal is quite likely to

satisfy the constraints of Theorem 4.5.

Although the results we have presented so far are limited to real signals, these results can be extended to permit signals which are complex but which have a spectral distribution limited to a half-plane. In these cases, it is possible to recover the real part of the signal from its zero crossings and then recover the imaginary part of the signal from the real part. This extension is very similar to Theorem 3.7 and is given in detail in [25].

At this point we should also point out that although the results presented here apply to periodic signals as well as nonperiodic signals, the results presented in chapter 3 are not a special case of the results presented here. This is because the results presented in chapter 3 consider the possible factorization of a signal in terms of a polynomial in e^{js} and e^{jw} whereas the results presented in this chapter consider the possible factorization of a signal in terms of an entire function in s and w. It is possible for a signal to be irreducible as a polynomial in e^{js}, e^{jw} but reducible as an entire function in s, w, as is the case, for example, with the function

$f(s,w) = 1 - e^{js}e^{jw} = (1 - e^{j\frac{s}{2}}e^{j\frac{w}{2}})(1 + e^{j\frac{s}{2}}e^{j\frac{w}{2}})$. A similar problem is mentioned by Sanz and Huang [30] when comparing their work on the reconstruction of signals from magnitude or phase to the work of Hayes [33]. In this case, it was found that the discrete-time problem considered by [33] is not a special case of the continuous-time problem considered by [30] As is discussed in [30], this problem can also be viewed in terms of different methods of extending the real signal to complex variables. In the case of periodic signals, the approach taken in chapter 3

and in [33,34,35] was effectively to map the periodic signal onto the unit surface in the complex space, as opposed to mapping the original signal onto the real plane in complex space as is done in this paper and in [30]. This problem is discussed in more detail in [30].

4.4. Proofs

In this section, we include the proofs for the results developed in this chapter.

Proof of Theorem 4.2.

As was mentioned earlier, if $f(x,y)$ and $g(x,y)$ are real, bandlimited (in the broad sense) functions, then it is well known [28,29] that these functions can be extended to C^2 as entire functions of exponential type (EFET) denoted as $f(s,w)$ and $g(s,w)$, which are also EFETs in each variable separately, and have at most polynomial growth in the real plane. If $f(x,y)$ and $g(x,y)$ are finite energy (in the real plane), then their Fourier transforms exist and this result is known as the Polya-Planckerel Theorem [27].

If $f(x,y)$ takes on positive and negative values in the closed, bounded region D, then since $f(x,y)$ is continuous (since it is entire), there must exist a contour (an uncountable number of points) where $f(x,y) = 0$. The same is true for $g(x,y)$. If there exists at least an infinite number of points $(x,y) \in D$ where $f(x,y) = 0$ and $g(x,y) = 0$ then there exists a limit point (see [26], proof of Theorem 4.11, p.72) which is contained in D. The set V_f is an analytic set [27, p.217] and is also an analytic surface [26, p.71] over a (complex) closed, bounded

domain $E : D \subset E \subset C^2$ since the analytic set is irreducible by assumption. In any

bounded and closed set $E \subset C^2$ if two distinct analytic surfaces have a sequence of

points in common along with their limit point then by Theorem 4.1, the sets coin-

cide not only in E but in C^2 so $V_f = V_g$ in all of C^2. Then we can make use of a

theorem stated precisely by Sanz and Huang [30]:

Theorem 4.6. Let $f, g : C^n \to C$ be entire functions such that $V_f = V_g$. If g is irreducible, then there exists an entire function $h : C^n \to C$ that satisfies $f = g \times h$.

Thus we now have:

$$f(s,w) = h(s,w)g(s,w) \tag{9}$$

where $h(s,w)$ is entire and nonzero everywhere in C^2. Using growth arguments as

in [30, p.1448] or by applying [36, Theorem 12] to any one-dimensional slice of

f, g, and h, we can also show that $h(s,w)$ is an EFET. It is well known that the

only EFET which is nonzero in all of C^2 is $e^{\alpha s + \beta w + \gamma}$. Since $f(s,w)$ and $g(s,w)$

must be real for all real values of s and w, then α, β, γ must all be real. We can

also show that α and β must be zero, since otherwise $f(s,w)$ or $g(s,w)$ would

have exponential growth in the real plane, and thus would not be bandlimited [29].

Thus $h(s,w) = e^{\gamma} = c$ and $f(x,y) = cg(x,y)$.

Note that in this proof we have only used the fact that there are an infinite

number of points where $f(x,y) = 0$ and $g(x,y) = 0$. Thus it is only necessary to

know a countably-infinite set of points on a zero crossing contour (e.g., a discrete

sequence of points); it is not necessary to know the complete zero crossing con-

tours.

Proof of Theorem 4.3.

Proceeding as in the proof of Theorem 4.2, we have:

$$g(s,w) = f(s,w)k(s,w) \tag{10}$$

where $k(s,w)$ is an EFET in C^2. We cannot assume that $k(s,w)$ is nonzero since

we have not assumed that $g(s,w)$ is irreducible. Instead, we can establish a rela-

tionship between the bandwidths of f,g,k by applying known properties of the so-

called P-indicators:

> **Theorem 4.7.** [27, Thm. 3.4.4] Let $f(z)$ and $k(z)$ be EFETs
> such that for almost all $x \in R^N$ the function $f(x+\lambda w), \lambda \in R^N$, has
> completely regular growth in the variable $w \in C^1$. Then the P-indicator
> of the function $g(z) = f(z)k(z)$ is the sum of the P-indicators of $f(z)$
> and $k(z)$: $h_g(\lambda) = h_f(\lambda) + h_k(\lambda)$.

The requirement that f,g (and therefore k) be strictly bandlimited (and thus any

one-dimensional slice of f or g is bandlimited) guarantees that they have com-

pletely regular growth along any slice [5]. The P-indicator corresponds to the smal-

lest convex domain completely containing the region of support of the spectrum

(see [27] for precise definitions). Roughly speaking, Theorem 4.7 states that the

bandwidth of g is the sum of the bandwidths of f and k. However, in Theorem

4.3, we have constrained g and f to have the same known bandwidth and thus k

must have "zero bandwidth", i.e., $h_k(\lambda) = 0$ in the notation of [27]. Thus, k must

be constant, and the theorem is proven.

Proof of Theorem 4.4.

The proof is straight-forward due to Theorem 4.2. The functions $f_1(x,y) = f(x,y) - h(x,y)$ and $g_1(x,y) = g(x,y) - h(x,y)$ satisfy the constraints of Theorem 4.2. Note that if the functions f, g, h are bandlimited in the general sense then so are f_1 and g_1.

Proof of Theorem 4.5.

The proof of this theorem is similar to the proof of Theorem 4.2 applied to each factor separately. Consider the factorization of $f(s,w)$ and $g(s,w)$ into a finite set of real factors, irreducible over the set of real factors, as described earlier. For each point where $f(s,w) = 0$ and $g(s,w) = 0$, at least one of the factors $f_i(s,w)$ must be zero and at least one of the factors $g_i(s,w)$ must be zero. For each zero contour of $f(s,w)$ and $g(s,w)$ corresponding to the irreducible factors $f_i(s,w)$ and $g_i(s,w)$, we can use Theorem 4.2 to show that $f_i(s,w) = cg_i(s,w)$. Let us assume that $f(x,y) \neq cg(x,y)$ and attempt to reach a contradiction. For convenience, let us assume that there is some irreducible factor of $f(s,w)$ which is not a factor of $g(s,w)$. First of all, note that if this factor, denoted $f_i(s,w)$, is complex, then $f_i^*(s,w)$ will also be a factor of $f(s,w)$ and thus $f(s,w)$ will contain a real factor $f_i(s,w)f_i^*(s,w)$ which is nonnegative everywhere, violating the constraints of the theorem. Thus, the factor $f_i(s,w)$ must be real, and since according to the theorem hypothesis, it has both positive and negative values and has multiplicity one, then we must have sign $f(x,y) \neq$ sign $g(x,y)$ for some values

of (x,y), and we have reached a contradiction. Thus, there cannot be any factor

of $f(s,w)$ which is not a factor of $g(s,w)$ and thus, $f(x,y) = cg(x,y)$.

CHAPTER 5

UNIQUE SPECIFICATION WITH FOURIER SIGN INFORMATION

The results developed so far on the unique specification of signals with zero crossings can be applied to the problem of unique specification of finite-length discrete-time sequences with sign information in the Fourier domain. In particular, since the Fourier transform of a finite length sequence is itself a periodic, bandlimited function, the results from chapter 3 can be applied directly. Similarly, the results of chapter 4 can be applied to the problem of uniquely specifying finite-length continuous-time signals with sign information in the Fourier domain. In this chapter, we will primarily discuss the unique specification of discrete-time sequences, although it should be kept in mind that by using the results of chapter 4, similar results can be developed for continuous-time signals.

A significant amount of research has been devoted to the problem of recovering finite length sequences from various forms of partial information in the Fourier domain [37], particularly the Fourier magnitude or phase [21,33,38]. In this chapter, we will begin by reviewing this previous research. We will then state some notation in section 2, and present our results in this area in section 3.

5.1. Related Research

Over the past few years, a considerable amount of research has been devoted to problems related to reconstructing a signal from various forms of partial informa-

tion in the Fourier domain, particularly the Fourier transform magnitude or phase. This work has involved developing conditions under which a signal is uniquely specified by its FT magnitude or phase and developing signal reconstruction algorithms.

In a variety of practical applications, only the magnitude or phase of a signal is available and it is desired either to reconstruct the original signal from this information or to synthesize a signal which retains many of the properties or features of the original signal. Reconstruction from magnitude (or "phase-retrieval") problems occur in areas such as electron microscopy [39], x-ray crystallography [40], and optical astronomy [41], where the magnitude or intensity of a diffraction or interference pattern is recorded and it is hoped that more complete information can be recovered. Reconstruction from phase (or "magnitude-retrieval") problems arise in cases where the "direction" of a field or signal is available and the magnitude is degraded in some unknown way; applications of these results have been suggested in image deblurring [21], arrival time estimation [42], and in the field of paleomagnetism [37].

In a number of different applications, it has been observed that many of the features of a signal are preserved when FT phase information is preserved and magnitude information is discarded but not in the opposite situation [43]. These results suggest that much of the intelligibility information resides in the phase, and that perhaps under some conditions, a signal could be reconstructed from the phase information alone. This is in fact true, and the restrictions are sufficiently mild to

allow almost all signals in one or more dimensions to be uniquely specified with FT phase. Specifically, it has been shown that if a one-dimensional discrete-time signal is finite length and has a z-transform with no zeros on the unit circle or in conjugate reciprocal pairs, then phase information alone is sufficient to uniquely specify the signal to within a scale factor [21,38]. A similar result has been developed for multidimensional signals which requires that the z-transform have no nontrivial symmetric factors [21,33]. These results have also been extended to the situation where the phase is known only at a set of sample frequencies such as when the phase of the DFT is available.

Unfortunately, a similar set of conditions has not been developed for the problem of reconstruction from Fourier transform magnitude. One-dimensional signals are not in general uniquely specified by FT magnitude since in the z-transform domain, zeros can be flipped inside or outside the unit circle without changing the magnitude on the unit circle. One-dimensional signals are uniquely specified by FT magnitude when all zeros are known to be either inside or outside the unit circle (the minimum or maximum phase conditions). In two or more dimensions, however, it has been shown that a finite-length discrete-time signal is uniquely specified to within a translation, reflection with respect to the origin, and a sign, by samples of Fourier transform magnitude [21,33] when its z-transform is not factorable, which is satisfied in most cases of practical interest. Similar results have been developed for reconstruction of a complex multidimensional bandlimited signal from either the magnitude or phase of the signal itself [30]. This is equivalent to

reconstructing a continuous-time signal from the magnitude or phase of its Fourier Transform.

The lack of satisfactory results on reconstruction of one-dimensional signals from FT magnitude prompted researchers to study the problem of unique specification with "signed-FT magnitude" (magnitude and one bit of phase). With the additional sign information, it is usually possible to reconstruct a one-dimensional signal. In particular, it has been shown that if a signal is real, causal, and its z-transform has no zeros on the unit circle it is uniquely specified by its signed-FT magnitude [37, 44, 45].

A variety of algorithms have been developed for reconstructing signals from the FT phase, magnitude, or signed magnitude. One algorithm for recovering a sequence from FT phase involves solving a set of linear equations. However, this method is practical only for small signals and applies only to reconstruction from phase. Another algorithm, used for solving a variety of reconstruction problems, is an iterative procedure which alternately imposes the finite length constraint in the time domain and the known Fourier domain information in the frequency domain [21, 33, 46]. A number of variations of these algorithms have been developed using different methods of imposing the constraints and different techniques for accelerating the convergence rate.

For reconstruction from phase, several forms of the iterative algorithm have been shown to converge to the correct sequence [47], and to yield good results in practice. In addition, Musicus [48, 49] has shown that by introducing two

extrapolation steps at each iteration, this algorithm can be converted to a conjugate gradient algorithm, guaranteed to converge in a finite number of iterations. The problem of reconstruction from magnitude appears to be more difficult, since the iterative algorithm will often generate a sequence which does not contain any recognizable features of the original image. A number of modifications have been proposed, with different authors claiming different degrees of success [21,33,50]. In general, the quality of the reconstruction depends strongly on the initial estimate and on the degree of complexity of the image. In fact, if an initial estimate is chosen which has the correct signed-magnitude, then the algorithm will converge to the desired result.

From the results discussed above, it is apparent that in many cases, use of either the phase or the signed-magnitude of the Fourier transform leads to good results, but that magnitude alone is not sufficient. This is true in the problem of direct Fourier synthesis, the development of theoretical results for one-dimensional signals, and in the development of reconstruction algorithms for signals of any dimension. These results suggest that important information about the signal is contained in the most significant bit of phase, the only information which is contained in both the signed-magnitude and the phase. In fact, it was shown in [37] that images synthesized with the most significant bit of phase (sign of the real part), unity magnitude, and zero or random numbers for the remainder of phase were found to be quite intelligible. As in previous work, this observation suggests that perhaps under some conditions, it may be possible to recover a signal from the one

bit of phase information, that is, from sign information in the Fourier domain. By applying the duality of the Fourier transform, we can use results on reconstruction from zero crossings to develop new results on reconstruction from sign information in the Fourier domain.

5.2. Notation

Before proceeding to develop results on the unique specification with Fourier sign information, let us define our notation for dealing with sequences. First of all, we will use $x[\underline{n}]$ to denote a discrete-time sequence of arbitrary dimension, and we will use $x[n]$ or $x[n_1,n_2]$ when the discussion is restricted to sequences of one or two dimensions, respectively. Similarly, we will use $X(z)$ to denote the z-transform, and $X(\omega)$ to denote the Fourier transform, i.e.:

$$X(z) = \sum_{\underline{n}} x[\underline{n}] z^{-\underline{n}} \qquad (11)$$

$$X(\omega) = \sum_{\underline{n}} x[\underline{n}] e^{-j\omega\underline{n}}$$

The Fourier transform sign information, or one bit of phase, will be defined as:

$$S_x(\omega) = \begin{cases} 1 & \text{if Re } \{X(\omega)\} \geq 0 \\ -1 & \text{otherwise} \end{cases} \qquad (12)$$

We will also refer to the even (symmetric) and odd (antisymmetric) components of a signal, defined as:

$$x_e[\underline{n}] = \frac{x[\underline{n}] + x[-\underline{n}]}{2} \qquad (13)$$

$$x_o[\underline{n}] = \frac{x[\underline{n}] - x[-\underline{n}]}{2}$$

Similarly, $X_e(z)$ will denote the z-transform of $x_e[\underline{n}]$. We will refer to z-transforms

as symmetric if they correspond to symmetric sequences, that is, $X(z)$ is symmetric if $X(z) = X(z^{-1})$. A factor of $X(z)$ will be said to be a real symmetric factor if it is symmetric as defined above and if all of its coefficients are real. This does not imply that $X(z)$ will take on only real values but does imply that $X(z)$ will be real on the unit surface $|z_i| = 1$ for all i. Furthermore, the set of real symmetric factors of a z-transform includes all possible factors which satisfy the definition above and is not limited to irreducible factors.

We shall also refer to two-dimensional signals with a region of support over a nonsymmetric half-plane (NSHP), defined to mean that if (n_1, n_2) is in the region of support, then $(-n_1, -n_2)$ is not in the region of support unless $n_1 = n_2 = 0$. Note that if a signal has NSHP support, then it is uniquely specified by its even component, since by equation (13) we have:

$$x[n_1, n_2] = \begin{cases} x_e[n_1, n_2] & n_1 = n_2 = 0 \\ 2x_e[n_1, n_2] & \text{otherwise} \end{cases} \qquad (14)$$

for values of (n_1, n_2) in the specified region of support. Also, note that since $x_e[n_1, n_2] \Longleftrightarrow \text{Re}\{X(\omega_1, \omega_2)\}$, if $x[n_1, n_2]$ has NSHP support, then it is uniquely specified with $\text{Re}\{X(\omega_1, \omega_2)\}$.

5.3. Results

We will next apply our results on unique specification of signals with zero crossings to the problem of unique specification of sequences with the Fourier transform sign information $S_x(\omega_1, \omega_2)$. These results can be stated in a number of different forms since a number of different results were developed for the dual

problem. Since many extensions follow directly from earlier results by duality, we will simply present the main result here and extend it for different definitions of $S_x(\omega_1,\omega_2)$. Details of other possible extensions of these results can be found in [34].

First of all, we note that since the real part of the Fourier transform only contains information about the even component of the sequence, we must require that $x[n_1,n_2]$ be even or be defined only over a nonsymmetric half-plane so it can be recovered from its even part. Also, as was the case for the unique specification with zero crossings, we note that if Re $\{X(\omega_1, \omega_2)\} > 0$ for all (ω_1,ω_2), then we could not expect sign (Re $\{X(\omega_1, \omega_2)\}$) to be sufficient to reconstruct the original signal. Thus, we will also assume that Re $\{X(\omega_1, \omega_2)\}$ and Re $\{Y(\omega_1,\omega_2)\}$ are positive in some regions of the (ω_1,ω_2) plane and negative in other regions.

Specifically, we state the following theorem:

Theorem 5.1. Let $x[n_1,n_2]$ and $y[n_1,n_2]$ be real two-dimensional sequences with support over a finite non-symmetric half-plane, with $S_x(\omega_1,\omega_2) = S_y(\omega_1,\omega_2)$. If Re $\{X(\omega_1, \omega_2)\}$ takes on both positive and negative values and $X_e(z_1,z_2)$ and $Y_e(z_1,z_2)$ are nonfactorable, then $x[n_1,n_2] = cy[n_1,n_2]$ for some positive constant c.

Proof: Since $x[n_1,n_2]$ and $y[n_1,n_2]$ are finite length sequences, $x_e[n_1,n_2]$ and $y_e[n_1,n_2]$ are finite length sequences, and Re $\{X(\omega_1, \omega_2)\}$ and Re $\{Y(\omega_1,\omega_2)\}$ are periodic, bandlimited signals. Since $X_e(z_1,z_2)$ and $Y_e(z_1,z_2)$ are nonfactorable, Re $\{X(\omega_1, \omega_2)\}$ and Re $\{Y(\omega_1,\omega_2)\}$ are nonfactorable as polynomials. Since

Re $\{X(\omega_1, \omega_2)\}$ and Re $\{Y(\omega_1,\omega_2)\}$ contain sign changes, we know that they must satisfy the constraints of Theorem 3.2, so Re $\{X(\omega_1, \omega_2)\} = c$ Re $\{Y(\omega_1,\omega_2)\}$, $x_e[n_1,n_2] = cy_e[n_1,n_2]$, and $x[n_1,n_2] = cy[n_1,n_2]$.

It is also possible to generalize Theorem 5.1 to allow unique specification with broader classes of sign information than the $S_x(\omega_1,\omega_2)$ as defined in equation (12). Since $S_x(\omega_1,\omega_2)$ can be viewed as one bit of Fourier transform phase, we can generalize $S_x(\omega_1,\omega_2)$ to allow quantizing the phase in different ways. Specifically, we can define:

$$S_x^\alpha(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } \alpha - \frac{\pi}{2} \le \phi_x(\omega_1,\omega_2) \le \alpha + \frac{\pi}{2} \\ -1 & \text{otherwise} \end{cases} \qquad (15)$$

or equivalently,

$$S_x^\alpha(\omega_1,\omega_2) = \text{sign} (\text{Re} \{X(\omega_1,\omega_2) e^{j\alpha}\}) \qquad (16)$$

The case $\alpha = 0$ corresponds to the definition of $S_x(\omega_1,\omega_2)$ given in equation (12). Alternatively, since $S_x(\omega_1,\omega_2)$ can be viewed as the zero crossings of the real part of the Fourier transform, we can generalize $S_x(\omega_1,\omega_2)$ to allow crossings of an arbitrary threshold as follows as we did in Theorem 3.5:

$$S_x^\beta(\omega_1,\omega_2) = \text{sign}(\text{Re} \{X(\omega_1, \omega_2)\}-\beta) \qquad (17)$$

To develop a result on unique specification with "generalized one bit of phase", we will combine these two ideas and define:

$$S_x^{\alpha,\beta}(\omega_1,\omega_2) = \text{sign}(\text{Re} \{X(\omega_1,\omega_2)e^{j\alpha}\} - \beta) \qquad (18)$$

We can then develop a result similar to Theorem 5.1 for this definition of sign

- 63 -

information:

Theorem 5.2. Let $x[n_1,n_2]$ and $y[n_1,n_2]$ be real two-dimensional sequences with support over a finite non-symmetric half-plane, with $S_x^{\alpha,\beta}(\omega_1,\omega_2) \equiv S_y^{\alpha,\beta}(\omega_1,\omega_2)$ for any α and β such that $\text{Re}\,\{X(\omega_1,\omega_2)e^{j\alpha}\} - \beta$ takes on both positive and negative values. Also, let:

$$\hat{x}[n_1,n_2] = \frac{x[n_1,n_2]\,e^{j\alpha} + x^*[-n_1,-n_2]\,e^{-j\alpha}}{2} - \beta\,\delta[n_1,n_2] \quad (19)$$

$$\hat{y}[n_1,n_2] = \frac{y[n_1,n_2]\,e^{j\alpha} + y^*[-n_1,-n_2]\,e^{-j\alpha}}{2} - \beta\,\delta[n_1,n_2]$$

where

$$\delta[n_1,n_2] = \begin{cases} 1 & \text{if } (n_1,n_2) = (0,0) \\ 0 & \text{otherwise} \end{cases}$$

If $\hat{X}(z_1,z_2)$ and $\hat{Y}(z_1,z_2)$ are nonfactorable, then $x[n_1,n_2] = cy[n_1,n_2]$ for $(n_1,n_2) \neq (0,0)$, and $x[0,0]\cos\alpha - \beta = c(y[0,0]\cos\alpha - \beta)$ for some positive constant c.

Proof: Since $x[n_1,n_2]$ and $y[n_1,n_2]$ are finite length sequences, $\hat{x}[n_1,n_2]$ and $\hat{y}[n_1,n_2]$ are finite length sequences, and $\hat{X}(\omega_1,\omega_2)$ and $\hat{Y}(\omega_1,\omega_2)$ are real, periodic, bandlimited signals. Since $\hat{X}(z_1,z_2)$ and $\hat{Y}(z_1,z_2)$ are nonfactorable, $\hat{X}(\omega_1,\omega_2)$ and $\hat{Y}(\omega_1,\omega_2)$ are nonfactorable as polynomials. Note that with the definitions given in the theorem statement,

$$\hat{X}(\omega_1,\omega_2) = \text{Re}\,\{X(\omega_1,\omega_2)e^{j\alpha}\} - \beta \quad (20)$$

$$\hat{Y}(\omega_1,\omega_2) = \text{Re}\,\{Y(\omega_1,\omega_2)e^{j\alpha}\} - \beta$$

Then, we note that $S_x^{\alpha,\beta}(\omega_1,\omega_2) = S_{\hat{x}}(\omega_1,\omega_2)$, and since $S_x^{\alpha,\beta}(\omega_1,\omega_2) = S_y^{\alpha,\beta}(\omega_1,\omega_2)$,

$S_{\hat{x}}(\omega_1,\omega_2) = S_{\hat{y}}(\omega_1,\omega_2)$, and $\hat{X}(\omega_1,\omega_2)$ contains sign changes. Since $\hat{X}(\omega_1,\omega_2)$ and

$\hat{Y}(\omega_1,\omega_2)$ satisfy the constraints of Theorem 3.2, $\hat{X}(\omega_1,\omega_2) = c\hat{Y}(\omega_1,\omega_2)$, and

$\dot{x}[n_1,n_2] = c\dot{y}[n_1,n_2]$. Then, with some algebra, $x[\dot{n}_1,n_2] = cy[n_1,n_2]$ for $(n_1,n_2) \neq (0,0)$, and $x[0,0]\cos\alpha - \beta = c(y[0,0]\cos\alpha - \beta)$ for some positive constant c.

The ambiguity at $(n_1,n_2) = (0,0)$ is not just a scale factor; it is a scaling of the value with respect to a threshold $\dfrac{\beta}{\cos\alpha}$. Note that if $\alpha = \dfrac{k\pi}{2}$ for some odd integer k, then $S_x^{\alpha,\beta}(\omega_1,\omega_2)$ contains no information about $x[0,0]$, and even if $\beta = 0$, $x[0,0]$ cannot be recovered.

The definition of Fourier sign information can also be modified to include complex values of (ω_1,ω_2). In particular, from a z-domain viewpoint, the real values of (ω_1,ω_2) correspond to values of (z_1,z_2) on the unit surface $|z_1| = |z_2| = 1$; we might also be interested in sign information on the surface $|z_1| = r_1$, $|z_2| = r_2$. In this case, if the sequence $y[n_1,n_2] = x[n_1,n_2]\, r_1^{-n_1}\, r_2^{-n_2}$ satisfies the constraints of the results developed earlier, then it is uniquely specified with the zero crossings of $\text{Re}\,\{Y(\omega_1,\omega_2)\}$, or equivalently, with the zero crossings of:

$$\text{Re}\,\{Y(z_1,z_2)\}\big|_{|z_1|=1,\ |z_2|=1} = \text{Re}\,\{X(r_1 z_1,\ r_2 z_2)\}\big|_{|z_1|=1,\ |z_2|=1} \qquad (21)$$

Then, since we can recover $x[n_1,n_2]$ from $y[n_1,n_2]$, $x[n_1,n_2]$ is uniquely specified with sign information over the surface $|z_1| = r_1$, $|z_2| = r_2$.

CHAPTER 6

RECONSTRUCTION ALGORITHMS

Having established that particular classes of signals are uniquely specified by some partial information, it is of interest to develop algorithms for recovering the original signal from this information. One common approach to developing algorithms for reconstruction from various forms of partial information is to develop an iterative algorithm which alternately imposes constraints in the space and frequency domains (similar to the Gerchberg-Saxton algorithm [51]). Another approach is to express the solution as a set of simultaneous linear equations. In this section, we will discuss each of these methods and present experimental results obtained with each method.

6.1. Iterative Method

The class of iterative algorithms mentioned above can be applied to the problem of reconstruction from zero crossings by imposing the correct sign of the signal (perhaps with respect to some threshold) in the space domain and the correct bandwidth in the Fourier domain. A similar algorithm can be developed for the problem of reconstruction from Fourier sign information as discussed in chapter 5.

Since knowledge of the exact points of discontinuity is necessary for the signal to be uniquely specified, the convergence of an iterative algorithm to the correct solution necessarily depends upon the use of the exact zero crossing points. Thus,

an algorithm for reconstruction from a sampled version of the sign of $f(x,y)$ cannot be guaranteed to converge to the correct solution. However, it can be shown that the continuous-space version of the algorithm (that is, a similar algorithm imposing the correct sign for all values of (x,y) and thus using an actual Fourier transform and not a DFT) will converge to the correct signal. It is also possible to show that the sampled version of the algorithm will converge to a signal which satisfies both the space and frequency domain constraints (provided such a solution exists), although the solution is not unique. These results can be developed within the theory of projections onto convex sets, as was used in [52] to establish the convergence of a number of different signal reconstruction algorithms. Specifically, the results developed in [52] apply directly to this problem provided the constraints in each domain are imposed in such a way as to be projections onto convex sets; the details will be presented in Appendix 1.

Once the theoretical properties of a reconstruction algorithm have been determined, it is important to empirically determine the effectiveness of the algorithm in recovering an actual signal. In particular, it is worthwhile to determine if a practical sampling rate limits the set of solutions to a sufficiently small set and if convergence (or a good approximation) can be obtained with a practical number of iterations. It is also worthwhile to investigate the effect of using different initial estimates in the iteration.

Experimentally, we have found that different types of signals require different sampling rates to limit the set of possible solutions to a reasonable size. Specifically,

- 67 -

signals with most of their energy concentrated in low frequencies inherently contain fewer zero crossing contours in the Fourier domain and thus each zero crossing point must be known to a greater accuracy, requiring the sign function to be significantly oversampled. When recovering practical signals from zero crossings, this is a serious problem since most signals which are bandlimited do not have their energy uniformly distributed throughout the passband. For these signals, the energy at frequencies near zero is often several orders of magnitude greater than the energy at frequencies near the band edges. For the dual problem of recovering sequences from Fourier sign information, there is no reason to believe that sequences would tend to have most of their energy in a particular region of the image. In fact, typical images tend to have many more zero crossing contours in the real part of their Fourier transform than threshold crossing contours (of any threshold) in the original image. Thus the problem of recovering signals from Fourier sign information tends to be more stable numerically (when typical pictures are used) and is less susceptible to inaccuracies in zero crossing locations caused by sampling the sign information.

We have had very little success with the iterative technique in recovering images from zero crossings due to the problems mentioned above. When recovering signals from Fourier sign information, we have found that it is necessary to use a DFT size at least 4 or 8 times the signal size, and that it is extremely desirable to start with an initial estimate which in some sense resembles the original signal. Specifically, we have had some success when using an initial estimate formed from

the correct one bit of Fourier transform phase and a Fourier transform magnitude which is the average of a number of unrelated images. This is because images often have most of their energy at low frequencies and thus the Fourier transform magnitudes corresponding to different images are often quite similar. In fact, as was shown in [37] an image synthesized from the correct one bit of Fourier transform phase and an average Fourier transform magnitude bears a strong resemblance to the original image. When the initial estimate has a flat Fourier transform magnitude (such as if random noise is used), then the recovered signal appears to be a noisy version of the original signal.

An example of an image reconstructed with this technique is included in Figure 1, where we show the original image (a) and the image reconstructed from Fourier sign information (b). In this example, the original image is 64×64 points, 256×256 DFTs were used, and the results shown were obtained with 25 iterations, accelerated with a method similar to that used in [21]. The initial estimate used here was an average Fourier transform magnitude combined with the correct one bit of phase or Fourier sign information. Although a large number of iterations are required for the algorithm to converge to a sequence satisfying the space and frequency domain constraints, the improvement of image quality after the first 20 iterations or so is somewhat negligible even if the image at this stage does not satisfy the frequency domain constraints at every point.

Figure 1. Reconstruction from Fourier Sign Information

(a) original image

(b) recovered image

6.2. Linear Equation Method

As mentioned earlier, it is also possible to express the solution to the problem of reconstruction from zero crossings as a set of linear equations. Thus, another possible reconstruction algorithm would involve solving the set of equations:

$$\sum_{n_1} \sum_{n_2} F[n_1,n_2] \, e^{j\frac{2\pi x_i n_1}{T_1}} \, e^{j\frac{2\pi y_i n_2}{T_2}} = 0 \tag{22}$$

where each equation uses a different pair of points (x_i, y_i) for which the equality is known to hold. This approach has the advantage that the exact zero crossing locations are used as opposed to sampling the sign information. Since these equations are written in terms of complex unknowns we can modify them for real unknowns and to take advantage of the fact that $f(x,y)$ is real as follows:

$$\sum_{n_1} \sum_{n_2} F_R[n_1,n_2] \cos(2\pi x_i n_1 + 2\pi y_i n_1) - F_I[n_1,n_2] \sin(2\pi x_i n_1 + 2\pi y_i n_1) = 0 \tag{23}$$

where $F_R[n_1,n_2]$ and $F_I[n_1,n_2]$ denote the real and imaginary parts of $F[n_1,n_2]$ respectively.

As mentioned earlier, if we assume that $f(x,y)$ satisfies the constraints of Theorem 3.9 and the number of equations used (i.e., the number of zero crossing points used) is also consistent with the constraints of Theorem 3.9, then these equations have a unique solution once the scale factor is specified by setting one point to its known value. We have found it simplest to set $F_R[0,0]$ to the known mean value of the signal (and $F_I[0,0] = 0$ since the image is real). We have also found that the number of equations specified by Theorem 3.9 is significantly higher than the number of equations usually required in practice. Our experience is that with a

small number of equations (signals with a narrow bandwidth), it is often possible to use the same number of equations as unknowns. As the number of unknowns (spectral components) increases, it is desirable to use more equations than unknowns and to compute a least squares solution in order to reduce the numerical difficulties associated with solving linear equations.

Examples of two images recovered with this method are shown in Figures 2 and 3,* which show (a) the original image, (b) an image consisting of the threshold crossings (i.e., contours showing where the original image crosses a particular threshold, and (c) the image reconstructed by solving the linear equations and then taking the inverse transform. The original image was obtained from a very similar image by low-pass filtering the image and removing some low-amplitude Fourier transform points so that it would be practical to solve a set of linear equations for the remaining points. The exact size and shape of the spectrum of the resulting image (i.e., the region of support of the Fourier transform) was then assumed to be known. Precise values of the zero crossing locations were found (to approximately 16 digit accuracy) by expressing the image as a polynomial as in equation (2) and using a numerical technique to find the zeros of this polynomial. The equations were then solved by using a QR decomposition [54] and double-precision arithmetic. In the case of Figure 2, the image contains 228 independent spectral components, a total of 600 equations in 454 unknowns were used (here almost all points are complex and contribute two unknowns), and the normalized rms error

---

*I would like to gratefully acknowledge the help of M. Steinberg in generating these figures and the ones that follow. Additional examples can be found in [53].

Figure 2. Reconstruction from Zero Crossings

(a) original image   (b) threshold crossings of (a)
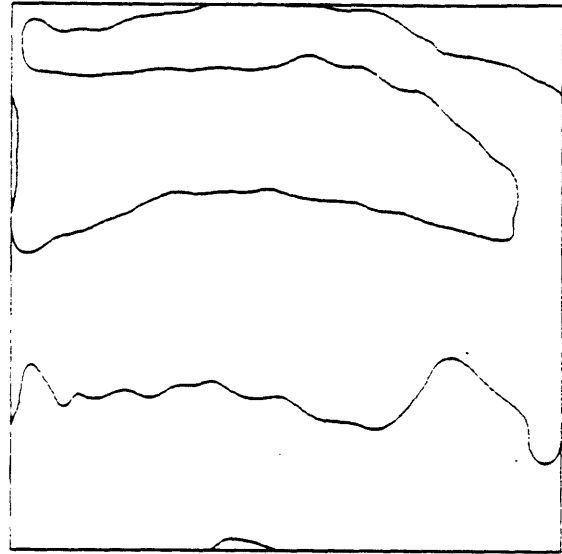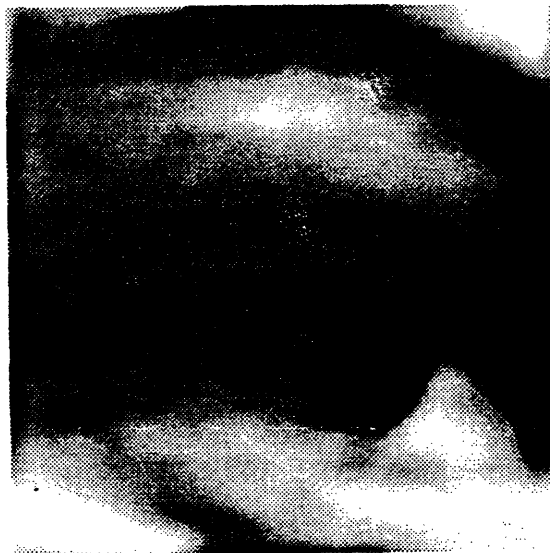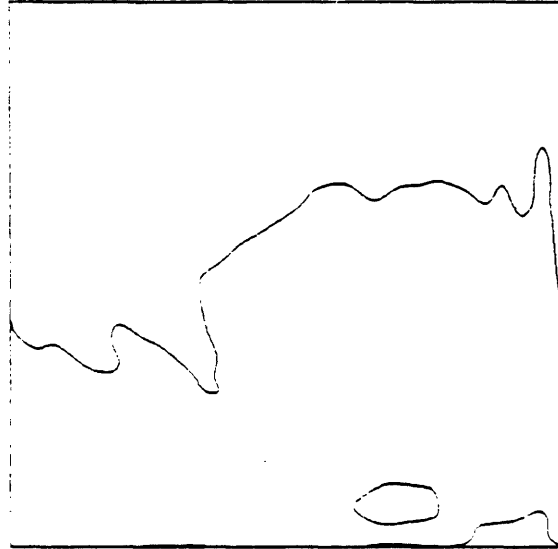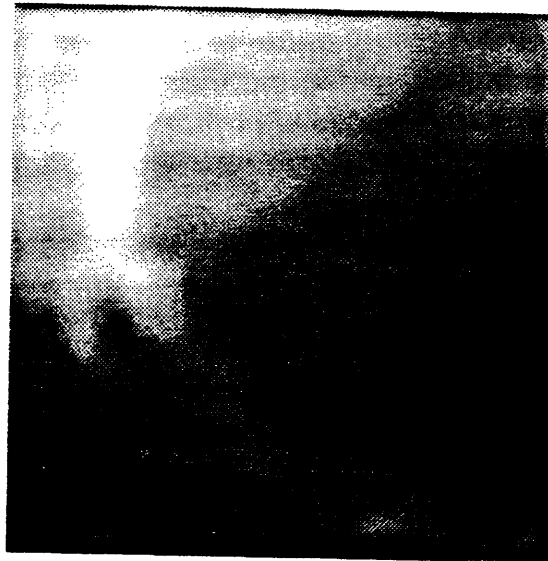

(c) recovered image

Figure 3. Reconstruction of X-Ray
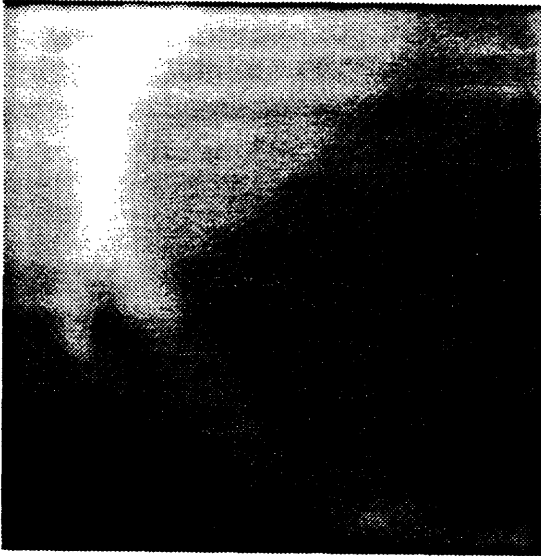
(a) original image

(b) threshold crossings of (a)

(c) recovered image

(rms error/rms signal) is approximately 0.000065. In the case of Figure 3, the image has 178 independent spectral components, a total of 600 equations in 354 unknowns were used, and the normalized rms error is approximately 0.027.

The success of this method depends on a number of different factors. The significant factors appear to be the accuracy of the zero crossing points and the degree to which the zero crossing points are spread out evenly throughout the picture. The locations of zero crossing points depend on the type of image as well as the threshold used. For example, we note that the reconstruction of the image in Figure 2 was more accurate than that of Figure 3 (despite the fact that the image of Figure 3 contains fewer spectral components and the same number of equations were used) and that Figure 2 contains more contours spread out throughout the picture. For both of these examples, the threshold chosen was somewhere near the mean value of the image. The mean value is not necessarily the best threshold to use but the best threshold is likely to be fairly close to the mean in most images. As the threshold is increased or decreased away from the mean, there will be fewer picture elements on one side of the threshold than the other, and furthermore these picture elements will tend to be concentrated in small areas of the picture. This means that the threshold crossing contours will be less evenly distributed throughout the picture, and the reconstruction process will be less stable. In particular, errors occurring during reconstruction will be greatest in areas which are the farthest from any zero crossing contour. For most images, there is a range of thresholds for which the reconstruction works well, and outside this range, significant

errors occur which increase as the threshold varies further from this range.

Examples of reconstruction showing the effects of different thresholds are shown in Figures 4 and 5. Figure 4 shows a range of different thresholds for the "eye" picture shown in Figure 1, and Figure 5 shows a range of different thresholds for the x-ray picture shown in Figure 2. The images are on a scale of 0 to 1, with mean values close to 0.5. Figure 4 (a) shows the threshold crossing contours obtained with a rather small threshold (.27), such that most of the picture elements are brighter than this threshold and the threshold crossing contours are concentrated in the center of the picture. Figure 4 (b) shows the image recovered from these contours. Notice that there are significant errors in the corners of the image, areas which are far from any threshold crossing contours. Figure 4 (c,d) shows another example with a slightly larger threshold (.30); in this case the image was recovered fairly accurately. Figure 4 (e,f) shows a different threshold (.62) somewhat above than the mean, where the signal was recovered accurately, and Figure 4 (g,h) shows a slightly larger threshold (.64) where again we begin to see distortions in areas far from the threshold crossing contours. Thus, the reconstruction appears to be most successful in the range of thresholds between .30 (Figure 4(c)) and .62 (Figure 4(e)).

A similar set of examples is shown in Figure 5 for the x-ray picture. Figure 5 (a,b) illustrates reconstruction using a threshold of .40. The distortions in the recovered image appear as a horizontal stripe rather than in a small area as was the case in the previous example. In the Fourier domain, the errors in these images
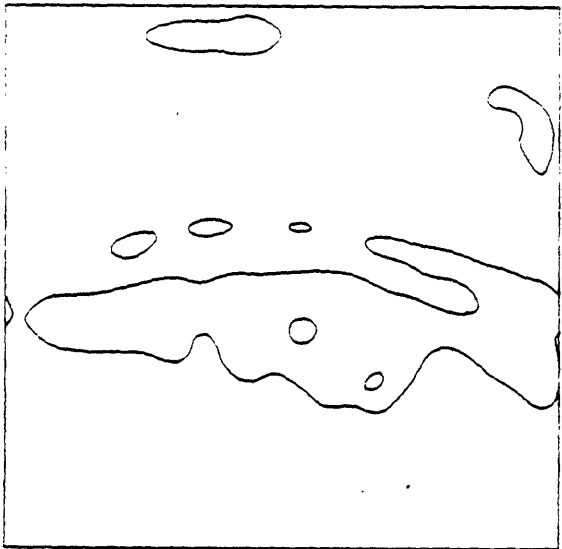
Figure 4. Effect of Different Thresholds
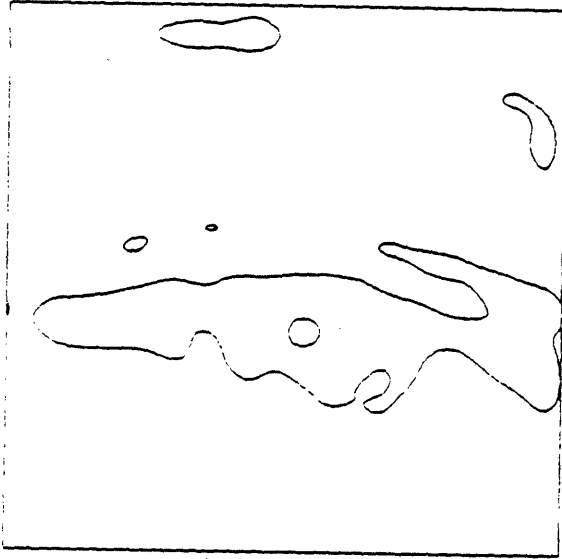
(a) threshold = 0.27
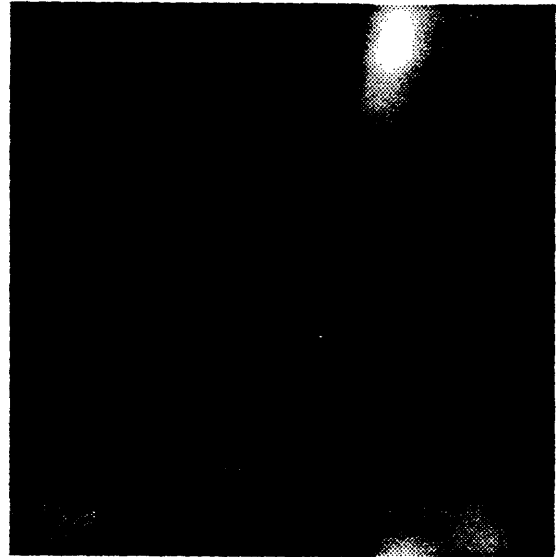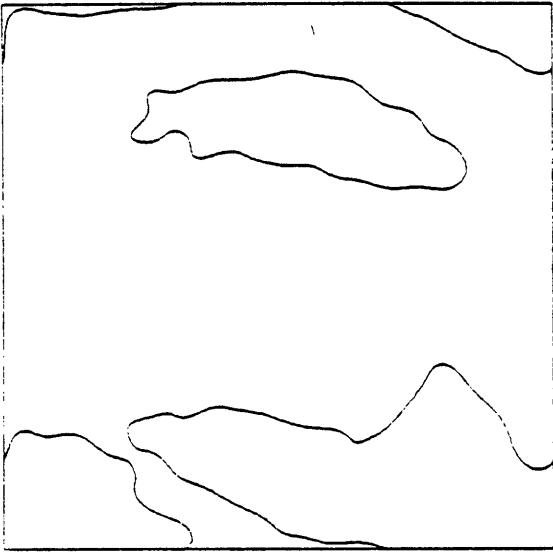

(b) image recovered from (a)

(c) threshold = 0.30
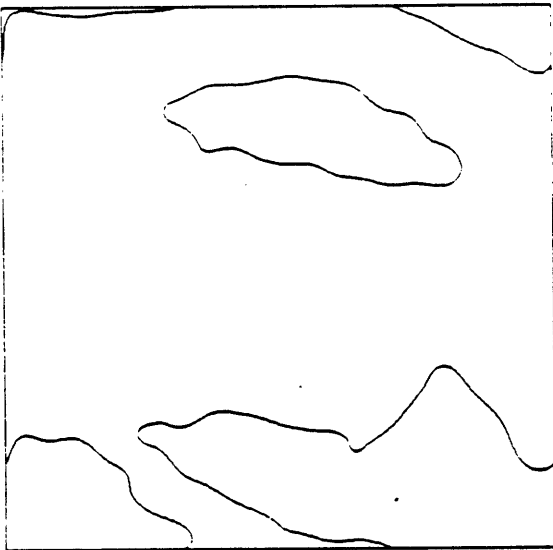

(d) image recovered from (c)

Figure 4 (cont'd). Effect of Different Thresholds

(e) threshold #3 = 0.62

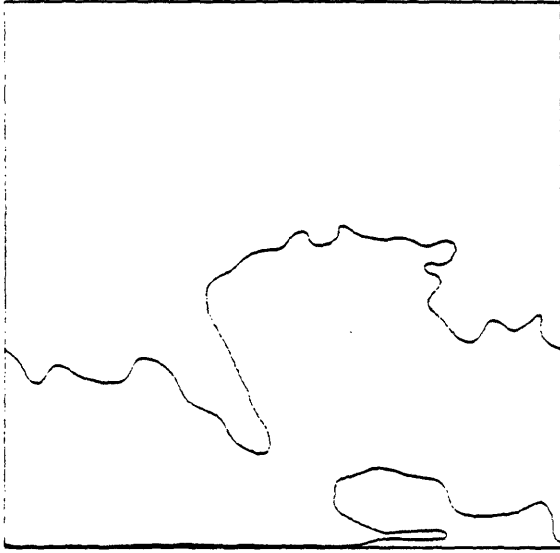(f) image recovered from (e)

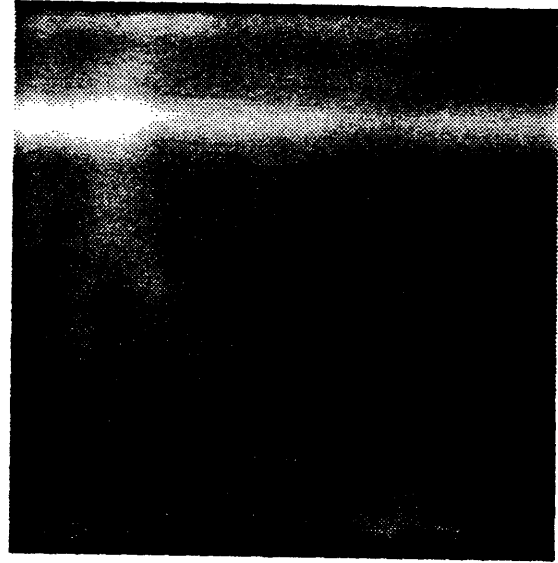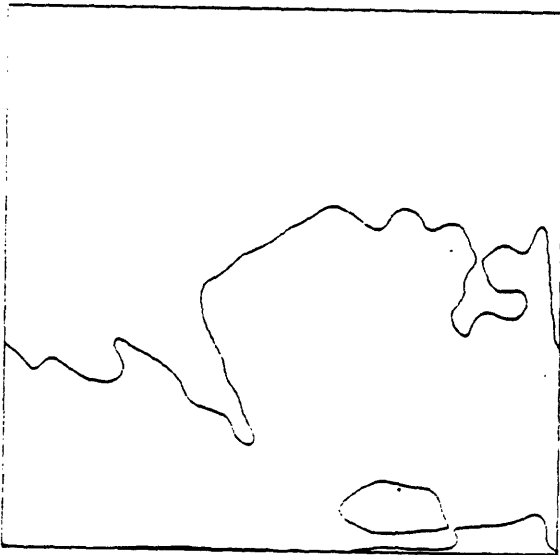(g) threshold #4 = 0.64

(h) image recovered from (g)

Figure 5. Effect of Different Thresholds on X-Ray
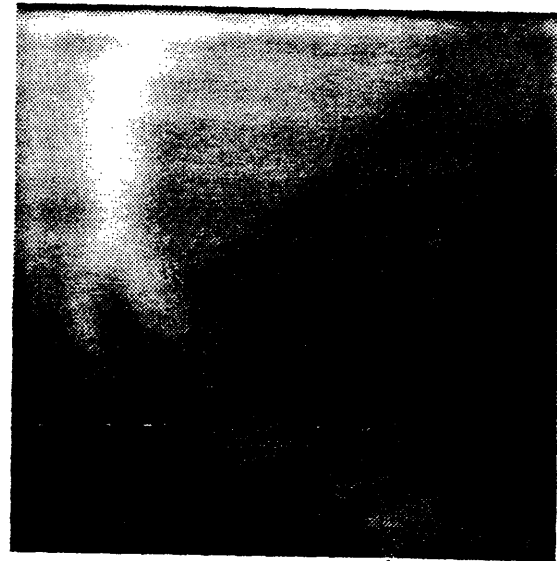

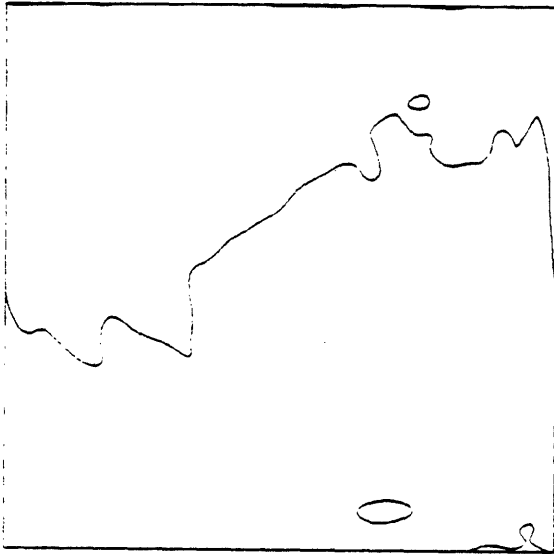(a) threshold = .40

(b) image recovered from (a)

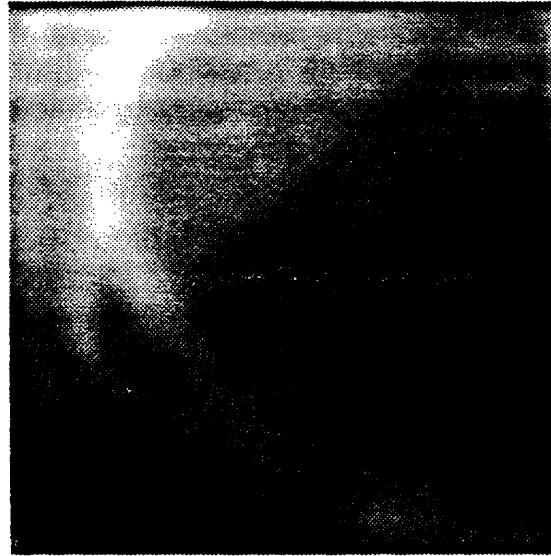(c) threshold = .43
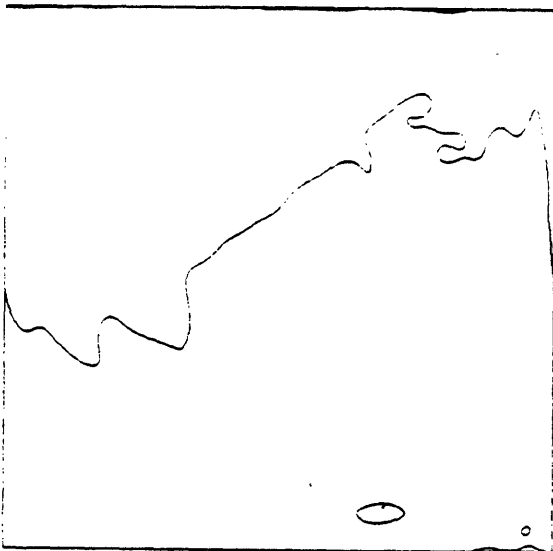

(d) image recovered from (c)

Figure 5 (cont'd). Effect of Different Thresholds on X-Ray
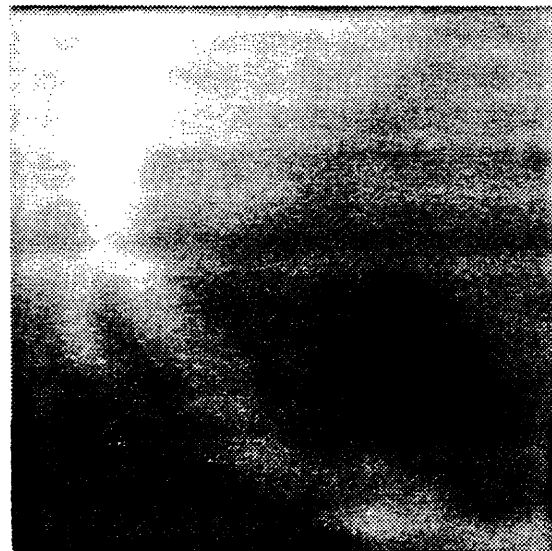

(e) threshold = .52

(f) image recovered from (e)
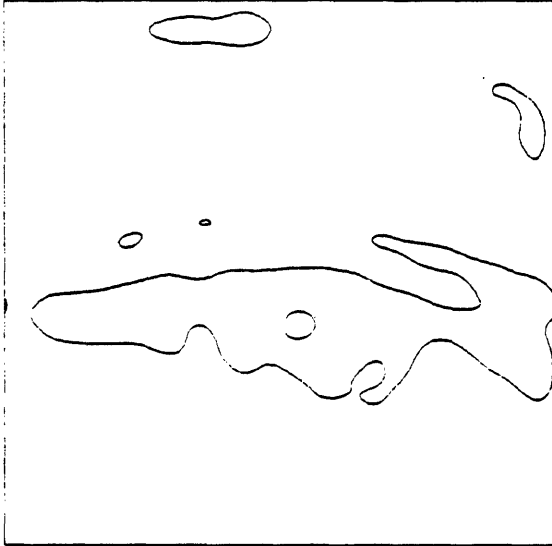

(g) threshold = .53


(h) image recovered from (g)

show up in the spectral components which have zero horizontal frequency and fairly low but nonzero vertical frequency. These components are under-constrained by the threshold crossing information since the threshold crossing contours are very nearly horizontal in this example. As seen in Figure 5 (c,d), when the threshold is increased to .43, the reconstruction is improved. Figure 5 (e,f) shows a reconstruction obtained with a threshold of 0.52, just slightly above the mean value. While this example was effectively recovered, when the threshold is increased slightly to 0.53, noticeable errors occur in the reconstruction (Figure 5 (g,h)). Note that in this image, the range of thresholds which result in effective signal reconstruction is much smaller than in the previous example and furthermore, this range is not centered about the mean value of the image (.49).

The stability of the reconstruction process can be improved by increasing the number of equations. An example is shown in Figure 6. Figure 6(a) shows the threshold crossings of the "eye" image used in Figure 1, with a threshold of 0.27. When 600 equations (454 unknowns) are used, the resulting image has significant distortion near the corners, which are far from the threshold crossing contours. When 750 equations are used, the resulting image has improved but the distortion is still noticeable. When 900 equations are used, the recovered image appears very similar to the original.

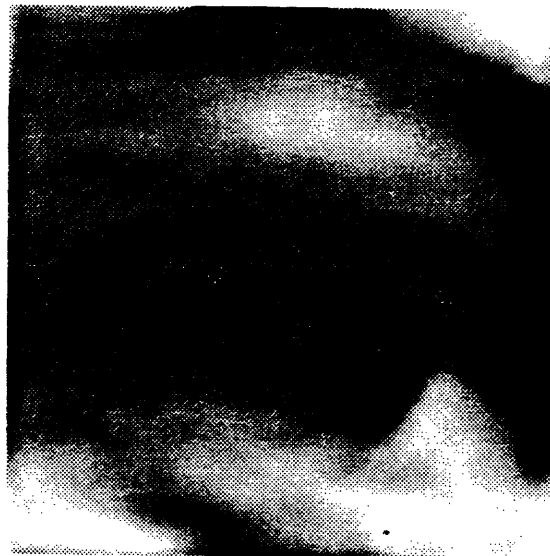Figure 6. Effect of Additional Equations

(a) threshold crossings
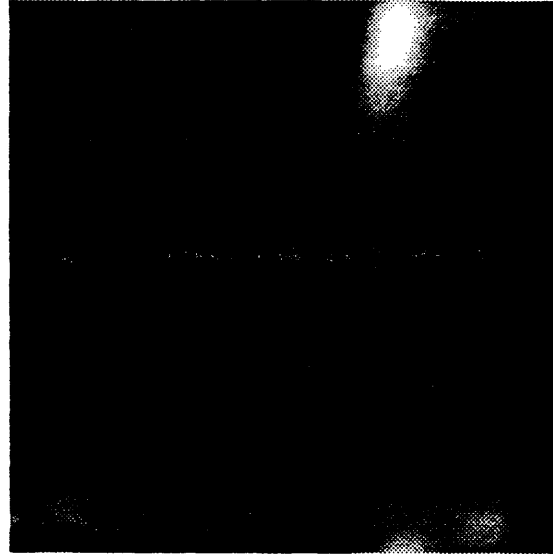
(b) 600 equations

(c) 750 equations

(d) 900 equations

CHAPTER 7

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

In this thesis, we have developed a number of new theoretical results on the unique specification of multidimensional signals with zero crossing or threshold crossing information. We began by considering the unique specification of two-dimensional periodic signals since these signals can be represented as a polynomial using the Fourier series. Our basic result established that a two-dimensional, periodic, bandlimited signal is uniquely specified to within a scale factor by its zero crossings if the signal is nonfactorable when expressed as a polynomial. We extended this result to include finite-length signals, crossings of an arbitrary threshold, sampling of the zero crossing contours, and factorable signals. We also developed similar results for signals of higher dimensions and for nonperiodic signals. We applied these results to the problem of unique specification of signals with zero crossings or sign information in the Fourier domain by using the duality of the Fourier transform in a straight-forward way.

We also discussed two algorithms for reconstruction from zero crossing information once it is known that a signal satisfies the uniqueness results. An iterative algorithm which imposes the known sign information in the space domain and the known bandwidth in the frequency domain was shown to converge to a signal satisfying the given constraints. This method was found not to be successful in practice because the constraints involved sampling the sign information and did not involve

imposing the exact zero crossing locations. An alternative method, involving the solution of linear equations where each equation uses one exact zero crossing point, was shown to have a unique solution if enough equations are used. This method was found to give results which are indistinguishable from the original signal provided that the original signal has a limited bandwidth and that the zero crossing locations are known very accurately.

A wide variety of different directions are possible for expanding or extending the results of this thesis. First of all, there are a number of theoretical issues which we have not investigated. One problem we have not addressed is the problem of determining when it is possible to synthesize an image with an arbitrary set of zero crossing points. In the results we have developed, we have always assumed that the zero crossing points did correspond to an image which we were trying to recover. A different problem involves assuming that the zero crossing points are not known apriori to correspond to an original image, but it is desired to develop constraints on the zero crossing locations such that it is possible to synthesize an image with zero crossings at these points. This problem has been investigated experimentally in [55] and has applications in areas such as microlithography where it is desired to create an image which will have the correct zero crossings after passing through some type of bandlimited or diffraction-limited system.

Another problem in the theoretical area is to develop a technique or criterion for choosing as few zero crossing points as possible to uniquely specify a signal. Although we have shown that it is possible to uniquely specify an image with a fin-

ite number of arbitrarily-chosen zero crossing points and furthermore it is possible to choose the zero crossing points in such a way that p zero crossing points will uniquely specify a signal with p degrees of freedom, a practical criterion for choosing the p zero crossing points so that this is possible could be beneficial in applications of this work.

A wide variety of experimental or algorithmic work is also possible to extend this thesis. One obvious possibility is to develop new reconstruction algorithms which might be more robust or more efficient than the method of solving linear equations. One possible method is to develop an iterative algorithm which uses the exact zero crossing points as opposed to sampling the sign of the signal, perhaps by using a conjugate-gradient or other numerical optimization routine [56]. Another possibility involves the use of regularization or stabilization techniques used in problems such as bandlimited extrapolation, which are known to be ill-conditioned. Another direction worth investigating involves using multiple thresholds or a threshold which varies across the picture. The thresholds can then be chosen in such a way that there are more zero crossing contours distributed more evenly throughout the picture, making the reconstruction process more stable. It would also be worthwhile to analyze the stability of the reconstruction process theoretically both with respect to errors in zero crossing locations and with respect to the assumption of bandlimitedness.

Another area for possible extension of the results in this thesis involves developing applications of these results. One possible application occurs in image

restoration. If an image is recorded with a high-contrast recording medium, then the recorded image consists primarily of two distinct levels of intensity, with a boundary between these two levels. In this case, it may be possible to determine the exact threshold crossing points even if it is not possible to determine the amplitude of the original signal at most points. By using the results presented in this thesis, it should be possible to recover the original image before recording provided that the image satisfies the constraints of any of the results developed earlier. A number of variations on this problem are possible. For example, if an image is over- or under- exposed unintentionally such that intensity values in a particular range are reproduced accurately but values are clipped outside this range, then again it should be possible to recover an undistorted version of the image.

Another method of applying the results developed in this area is to use the theoretical properties in developing additional theoretical results in areas such as communication theory as well as in signal processing. One possibility along these lines is to use the results to develop new results on multidimensional sampling theory. Although results on sampling of one-dimensional bandlimited functions are well established, sampling theorems for two-dimensional functions are generally limited to sampling on rectangular or other periodic grids. One area where it would be useful to develop new results pertains to sampling on polar grids.

To show how the results on zero crossings apply to sampling, let us define the phrase "sampling set" to specify a set of points chosen such that the samples of the signal at those points are sufficient to uniquely specify the signal. It can be easily

shown [5] that a set of points is a sampling set if and only if there is no signal which has zero crossings at these points (and has the required bandwidth). Specifically, note that if there are two signals $f(x,y)$ and $g(x,y)$ of the same bandwidth with the same sample values at a particular set of points S, then $h(x,y) = f(x,y)-g(x,y)$ is zero on the set of points in S. Thus if there is no signal $h(x,y)$ with zero crossings which include all the points in S, the set S will be a sampling set. Thus if we have a particular set of points and we want to know if these points form a sampling set, we can attempt to recover a signal which has zero crossings at all points in the set. If such a signal does not exist, we know that the points form a sampling set. Furthermore, we may be able to create interpolation functions by recovering signals which are unity at one point in the set and zero at the other $N-1$ points in the set. It may also be possible to use the results of this thesis to develop new theoretical results pertaining to other aspects of multidimensional sampling theory or to related fields such as communication theory.

Appendix 1. Convergence of Iterative Reconstruction Algorithm

In this appendix, we establish the convergence of our algorithm for reconstruction from zero crossings by using the theory of projection onto convex sets, using the approach presented in [52] for establishing the convergence of a variety of signal reconstruction algorithms. Specifically, the result we shall be using is as follows*:

> **Theorem A1.** Let H be a Hilbert space, G be a composition of projection operators onto closed convex sets, at least one of which is finite-dimensional, and G^* denote the intersection of these sets. If G^* is nonempty, then for all $x \in H$, the sequence $G^n x$ converges to a point in G^*.

The results developed in [52] include showing that a wide variety of constraints often imposed in signal reconstruction algorithms can be imposed in such a way that the transformations will satisfy the constraints of Theorem A1. In fact, the constraints used in our algorithm for reconstruction from zero crossings differ from some of those discussed in [52] in only trivial ways. The basic approach shall be repeated here although the mathematical details shall be omitted.

To show that Theorem A1 applies to our iterative algorithm, we must first carefully define the transformations applied in the time (or space) and frequency domains at each iteration so that they can be characterized as projection operators onto closed convex sets. First let us note that although the continuous-space

---

*Theorem A1 is a weak form of the results in [52, 57] but it is sufficient for our purposes.

variables (x,y) will be used throughout the discussion below, the properties of the transformations in the space and frequency domains apply equally well if the samples (x_k,y_k) are used. The only difference is that in the continuous-space case, the set G^* will contain exactly one sequence (provided the proper constraints are satisfied), whereas in the discrete-space case, the set G^* will contain an infinite number of signals.

Let $h(x,y)$ denote the signal we are trying to recover which is known to satisfy the constraints of Theorem 3.9 and is thus uniquely specified by sign $h(x,y)$. Let W^* denote the set of sequences which satisfy the frequency domain constraints:

$$F[n_1,n_2] = 0 \quad \text{for } n_1 \text{ or } n_2 \notin [-N,N] \quad (A1)$$

$$F[N,N] = H[N,N]$$

and T^* denote the set of sequences which satisfy the spatial domain constraints:

$$\begin{aligned} f(x,y) &\geq 0 \quad \text{if sign } h(x,y) = 1 \\ f(x,y) &\leq 0 \quad \text{if sign } h(x,y) = -1 \end{aligned} \quad (A2)$$

Note that the set W^* is finite-dimensional even if the space H includes infinite-length signals. Also note that the definition of the set T^* may not be precisely the same as stating sign $f(x,y) = $ sign $h(x,y)$, since at a zero crossing point where $f(x,y) = 0$ it is considered to satisfy either constraint sign $h(x,y) = 1$ or sign $h(x,y) = -1$. It is necessary to use the definition of T^* given in (A2) in order for T^* to be a closed set (a set which includes its limit points.)

Next we will define the operators T and W to be projections onto the sets T^* and W^*. For T and W to be projections, we need:

$$||Tx - x|| \le ||y - x|| \quad \text{for all } y \in T^* \tag{A3}$$

and

$$||Wx - x|| \le ||y - x|| \quad \text{for all } y \in W^* \tag{A4}$$

Thus the operator W which imposes frequency domain constraints is:

$$W[F[n_1,n_2]] = \begin{cases} F(n_1,n_2) & 0 \le n_1,n_2 \le N \quad (n_1,n_2) \ne (N,N) \\ H(N,N) & (n_1,n_2) = (N,N) \\ 0 & n_1,n_2 > N \end{cases} \tag{A5}$$

or in words, simply substituting the known values of $H[n_1,n_2]$. The operator T

which imposes spatial domain constraints is:

$$T[f(x,y)] = \begin{cases} f(x,y) & \text{if sign } f(x,y) = \text{sign } h(x,y) \\ 0 & \text{otherwise} \end{cases} \tag{A6}$$

or in words, setting the signal to zero at points where its sign is incorrect.

Next, we express our iterative algorithm in the form

$$x_{k+1} = Gx_k \tag{A7}$$

where $G = TW$ is a composition of projection operators. Then, by Theorem A1,

the sequence in equation (A7) will converge to a point in G^*, that is, a sequence

which satisfies the time and frequency domain constraints. Thus, if $h(x,y)$ satisfies

the constraints of Theorem 3.9, and the iteration imposes the correct sign $h(x,y)$

for all points (continuous-space, not sampled) then G^* contains exactly one

sequence, and the iteration must converge to that sequence. If sign $h(x,y)$ is sam-

pled, then the iteration must converge to a sequence in G^*, i.e., a sequence which

satisfies the time and frequency domain constraints, although this solution is not

unique and the solution actually obtained depends on the initial estimate.

References:

[1] J. C. R. Licklider and I. Pollack, "Effects of differentiation, integration and infinite peak clipping upon the intelligibility of speech," *J. Acoust. Soc. Amer.*, vol. 20, pp. 42-51, Jan. 1948.

[2] D. Marr, S. Ullman, and T. Poggio, "Bandpass Channels, Zero Crossings, and Early Visual Information Processing," *J. Opt. Soc. Am.*, vol. 69, no. 6, pp. 914-916, June 1979.

[3] A. A. G. Requicha and H. B. Voelcker, "Design of Nonrecursive Filters by Specification of Frequency Domain Zeros," *IEEE Trans. Electroacoust.*, vol. AU-18, pp. 464-470, Dec. 1970.

[4] J. Clarke, "Steering of zeros in the directional pattern of a linear array," *IEEE Trans. Antennas Propagat.*, vol. AP-16, pp. 267-268, March 1968.

[5] A. A. G. Requicha, "Zeros of Entire Functions: Theory and Engineering Application," *Proc. IEEE*, vol. 68, no. 3, pp. 308-328, March 1980.

[6] H. B. Voelcker, "Toward a Unified Theory of Modulation," *Proc. IEEE*, vol. 54, pp. 340-353, 735-755, 1966.

[7] R. F. Hyneman, "A Technique for the Synthesis of Line-Source Antenna Patterns having Specified Sidelobe Behavior," *IEEE Trans. Antennas Propagat.*, vol. AP-16, pp. 430-435, July 1968.

[8] S. J. Haavik, "The Conversion of the Zeros of Noise," M.Sc. thesis, Univ. Rochester, Rochester, NY, 1966.

[9] B. F. Logan, Jr., "Information in the Zero Crossings of Bandpass Signals," *Bell Syst. Tech. J.*, vol. 56, pp. 487-510, April 1977.

[10] I. Bar-David, "An implicit sampling theorem for bounded bandlimited functions," *Inform. Contr.*, vol. 24, pp. 36-44, Jan. 1974.

[11] H. B. Voelcker, "Zero Crossing Properties of Angle-Modulated Signals," *IEEE Trans. Communic.*, vol. COM-20, pp. 307-315, June 1972.

[12] H. B. Voelcker and A. A. G. Requicha, "Clipping and Signal Determinism: Two Algorithms Requiring Validation," *IEEE Trans. Commun.*, vol. COM-21, pp. 738-744, June 1973.

[13] A. Sekey, "A Computer Simulation Study of Real Zero Interpolation," *IEEE Trans. Electroacoust.*, vol. AU-18, pp. 43-54, March 1970.

[14] S. Shitz, "Nonuniformly Spaced Sampling for Bandlimited Functions and Stochastic Processes," Technion Israel Inst. of Tech., Elec. Eng. Pub. no. 505, Oct. 1984. Also submitted for publication in *IEEE Trans. Inform. Theory*.

[15] Y. Y. Zeevi and D. Rotem, "Information in the zero crossings of images," Technion Israel Inst. of Tech., Elec. Eng. Vision Res. Lab. Internal Report, 1984.

[16] B. F. Logan, Jr., "Signals Designed for Recovery after Clipping - I. Localization of Infinite Products, II. Fourier Transform Theory of Recovery," *AT&T Bell Syst. Tech. J.*, vol. 63, no. 2, pp. 281-306, Feb. 1984.

[17] A. Mostowski and M. Stark, *Introduction to Higher Algebra*. New York: Mac-Millan Co., 1964.

[18] R. J. Walker, *Algebraic Curves*. New York: Springer-Verlag, 1978.

[19] M. H. Hayes and J. H. McClellan, "Reducible Polynomials in More than One Variable," *Proc. IEEE*, vol. 70, no. 2, pp. 197-198, Feb. 1982..

[20] J. L. C. Sanz, T. S. Huang, and F. Cukierman, "Stability of unique Fourier-transform phase reconstruction," *J. Opt. Soc. Am.*, vol. 73, no. 11, pp. 1442-1445, Nov. 1983.

[21] M. H. Hayes, "Signal Reconstruction from Phase or Magnitude," Sc.D. Thesis, M.I.T., Dept. of EECS, June 1981.

[22] S. Lang, *Introduction to Algebraic Geometry*. New York: Interscience Publishers, Inc., 1958.

[23] I. R. Shafarevich, *Basic Algebraic Geometry*. New York: Springer-Verlag, 1974.

[24] R. Hartshorne, *Algebraic Geometry*. New York: Springer-Verlag, 1977.

[25] S. R. Curtis, S. Shitz, and A. V. Oppenheim, "Reconstruction of Nonperiodic Two-Dimensional Signals from Zero Crossings," *IEEE Trans. Acoust., Speech, Signal Proc.*, to be submitted.

[26] B. Fuks, *Introduction to the Theory of Analytic Functions in Several Complex Variables*. Providence: Amer. Mat. Soc. Transl., 1963.

[27] L. Ronkin, *Introduction to the Theory of Entire Functions in Several Variables*. Providence: Amer. Mat. Soc. Transl., 1974.

[28] L. Hormander, *Linear Differential Operators*. New York: Springer-Verlag, 1963.

[29] J. Barros-Neto, *An Introduction to the Theory of Distributions*. New York: Marcel Dekker, Inc., 1973.

[30] J. L. C. Sanz and T. S. Huang, "Unique Reconstruction of a bandlimited multidimensional signal from phase or magnitude," *J. Opt. Soc. Am.*, vol. 73, no. 11, pp. 1446-1450, Nov. 1983.

[31] J. C. Dainty and M. A. Fiddy, "The essential role of prior knowledge in phase retrieval," *Optica Acta*, vol. 31, no. 3, pp. 325-330, Mar. 1984.

[32] R. Barakat and G. Newsam, "Necessary conditions for a Unique Solution of Two-Dimensional Phase Recovery," *J. Math. Phys.*, vol. 25, no. 11, pp. 3190-3193, Nov. 1984.

[33] M. H. Hayes, "The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of its Fourier Transform," *IEEE Trans. on Acoust.*,

Speech, Signal Processing, vol. ASSP-30, no. 2, pp. 140-154, April 1982.

[34] S. R. Curtis, A. V. Oppenheim, and J. S. Lim, "Signal Reconstruction from Fourier Transform Sign Information," *IEEE Trans. Acoust., Speech, Signal Proc.*, to be published June 1985. Also in MIT Research Laboratory of Electronics Technical Report no. 500, Cambridge, MA, May 1984.

[35] S. R. Curtis, J. S. Lim, and A. V. Oppenheim, "Signal Reconstruction from One Bit of Fourier Transform Phase," *Int'l Conf. on Acoustics, Speech, Signal Proc.*, March 1984.

[36] B. Ja. Levin, *Distribution of Zeros of Entire Functions*. Providence: Amer. Mat. Soc. Transl., 1980.

[37] A. V. Oppenheim, J. S. Lim, and S. R. Curtis, "Signal Reconstruction from Partial Fourier Domain Information," *J. Optical Society of America*, vol. 73, no. 11, Nov. 1983.

[38] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal Reconstruction from Phase or Magnitude," *IEEE Trans. on Acoust., Speech, Signal Proc.*, vol. ASSP-28, no. 6, pp. 672-680, Dec. 1980.

[39] W. O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy*. New York: Academic Press, 1978.

[40] G. N. Ramachandran and R. Srinivasan, *Fourier Methods in Crystallography*. New York: Wiley-Interscience, 1970.

[41] J. R. Fienup, "Space Object Imaging through the Turbulent Atmosphere," *Optical Engineering*, vol. 18, pp. 529-534, Sept. 1979.

[42] Y. T. Li and A. L. Kurkjian, "Arrival Time Determination Using Iterative Signal Reconstruction from the Phase of the Cross Spectrum," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 2, pp. 502-504, April 1983.

[43] A. V. Oppenheim and J. S. Lim, "The Importance of Phase in Signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529-541, May 1981.

[44] P. L. Van Hove, "Signal Reconstruction from Fourier Transform Amplitude," S.M. thesis, M.I.T., Dept. of EECS, Sept. 1982.

[45] P. L. Van Hove, M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal Reconstruction from Signed Fourier Transform Magnitude," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Oct. 1983.

[46] A. V. Oppenheim, M. H. Hayes, and J. S. Lim, "Iterative Procedures for Signal Reconstruction from Phase," *Optical Engineering*, vol. 21, no. 1, pp. 122-127, Jan./Feb. 1982.

[47] V. Tom, T. Quatieri, M. Hayes, and J. McClellan, "Convergence of Iterative Nonexpansive Signal Reconstruction Algorithms," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. ASSP-29, Oct. 1981.

[48] B. R. Musicus, "Iterative Algorithms for Optimal Signal Reconstruction and Parameter Identification given Noisy and Incomplete Data," Ph.D. thesis, M.I.T. Dept of EECS, Cambridge, MA, 1982. See also M.I.T. Research Lab of Electronics Tech. Rep. no. 496.

[49] B. R. Musicus, "Iterative Algorithms for Optimal Signal Reconstruction and Parameter Identification given Noisy and Incomplete Data," *Int'l Conf. on Acoustics, Speech, Signal Proc.*, April 1983.

[50] J. R. Fienup, "Reconstruction from the Modulus of the Fourier Transform," *Optics Letters*, vol. 3, no. 1, pp. 27-29, July 1978.

[51] 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.

[52] D. C. Youla and H. Webb, "Image Restoration by the Method of Convex Projections: Part I - Theory," *IEEE Trans. Med. Imag.*, vol. MI-1, no. 2, pp. 81-94, Oct. 1982.

[53] M. Steinberg, "Experimental Aspects of Image Reconstruction from Threshold Crossings," S.B. Thesis, M.I.T., Dept. of EECS, June 1985.

[54] G. Strang, *Linear Algebra and Its Applications*. New York: Academic Press, 1980.

[55] S. I. Sayegh, B. E. A. Saleh, and K. M. Nashold, "Image Design: Generation of a Prescribed Image Through a Diffraction-Limited System with High-Contrast Recording," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-33, no. 2, pp. 460-465, April 1985.

[56] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, 1969.

[57] Z. Opial, "Weak Convergence of the Sequence of Successive Approximations for Nonexpansive Mappings," *Bull. Amer. Math. Soc.*, vol. 73, pp. 591-597, 1967.

DISTRIBUTION LIST

DODAAD Code

Director   HX1241   (1)
Defense Advanced Research Project Agency
1400 Wilson Boulevard
Arlington, Virginia 22209
Attn: Program Management

Group Leader Information Sciences   N00014   (1)
Associate Director for Engineering Sciences
Office of Naval Research
800 North Quincy Street
Arlington, Virginia 22217

Administrative Contracting Officer   N66017   (1)
E19-628
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Director   N00173   (6)
Naval Research Laboratory
Attn: Code 2627
Washington, D.C. 20375

Defense Technical Information Center   S47031   (12)
Bldg. 5, Cameron Station
Alexandria, Virginia 22314

Dr. Judith Daly   (1)
DARPA / TTO
1400 Wilson Boulevard
Arlington, Virginia 22209