Reconstruction of Nonperiodic Two-Dimensional Signals from Zero Crossings

SUSAN R. CURTIS, SHLOMO SHITZ, AND ALAN V. OPPENHEIM

Abstract—In this correspondence, we present new results on the reconstruction of two-dimensional signals from zero crossing or threshold crossing information. Specifically, we develop new theoretical results which state conditions under which two-dimensional bandlimited signals are uniquely specified to within a scale factor with this information. Unlike previous results in this area, our new results do not constrain the signals to be periodic or bandpass.

I. INTRODUCTION

A significant amount of research has been devoted to the problem of reconstruction of signals from zero crossings [1]. Although historically most of this work has been in the field of communication theory and has concentrated on one-dimensional signals, more recently, results have been developed on the reconstruction of multidimensional periodic signals from zero crossings [2], [3]. Alternative multidimensional results have been derived by directly extending one-dimensional results [4]–[6], but these results fail to take advantage of the fundamental differences between one- and two-dimensional signals, and thus require one-dimensional signals derived from the two-dimensional signal to be bandpass or to have zeros in particular locations. In this correspondence, we present new results on the reconstruction of arbitrary bandlimited two-dimensional signals from zero crossings.

In the next section, we shall define terminology and review the mathematics necessary for the remainder of this work. In Section II, we present our basic result on the unique representation of arbitrary bandlimited two-dimensional signals with zero crossings and a number of extensions to it.

II. BACKGROUND

In this section, we define the notation and terminology to be used in the remainder of the correspondence. We also review some properties of functions and their zeros which will be of importance in the following section.

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. In this work, we shall be primarily concerned with entire functions of exponential type (EFET) (see [1] for a review of the properties of EFET's in one variable, and [7] and [8] for EFET's in several variables). These functions are constrained to have at most an exponential growth rate in any direction in complex space. As is well known [9], [10], 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 band limitation are possible by using the Fourier-Stieltjes transform or the so-called 2-transform (see [1] for definitions and for other possible definitions of band limitation). It is also possible to use a more general definition of band limitation derived from the theory of generalized functions or distributions. This definition requires the spectral distribution as defined in [9] to have compact support. Unless otherwise noted, the results presented in this correspondence apply to this more general type of bandlimitness, 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 [8] or [11] for a precise characterization). For either polynomials or entire functions, the representation 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 complex zeros (for example, the function $e^w$ is non-zero for all real or complex values of $w$). 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 EFET's 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 have finite energy. A more general class of one-dimensional bandlimited functions is characterized precisely in [1] by developing a subset of EFET's referred to as $B$ functions. The class of $B$ functions includes the set of bandlimited signals under a number of common definitions of band limitation (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 [1] and can be used to eliminate the possibility of exponential factors. For functions of several variables, the Paley–Wiener–Schoenberg theorem [10], 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 considering the factorization of polynomials. Specifically, let $V_f$ and $V_r$ 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 $V_r = \{(s, w): g(s, w) = 0\}$. 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 $V_f$ and $V_r$ are both nonempty sets. (This definition is also used in [14].) Note that if $h$ is an entire function which never vanishes (such as $e^w$), 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 [7], [8]. For example, $V_f$ and $V_r$ as defined above are analytic sets, as is $V_f \cap V_r$. An irreducible analytic set is an analytic set which cannot

* Globally irreducible, in the terminology of [11]–[13] and others.
be expressed as the union of two distinct 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 [7].

III. UNIQUE SPECIFICATION WITH ZERO CROSSINGS

In this section, we develop our new results on reconstruction of two-dimensional signals from zero crossings by applying known results on analytic sets. The main result is developed in Section III-A and a number of extensions are presented in Section III-B.

A. Basic Result

The problem of establishing conditions under which a signal is uniquely specified with zero crossings can be solved with results from the theory of intersection of analytic sets. Since our result depends primarily on a result in this area which is readily available in the mathematics literature, we shall first state this result and then show how it applies to the problem of unique specification with zero crossings:

**Theorem 1** [7]: 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 allows us to develop our result on reconstruction from zero crossings in a straightforward 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 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 the Appendix for proof) the following.

**Theorem 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 $D \subset \mathbb{R}^2$ and sign $f(x, y) = \text{sign } g(x, y)$ for all values of $(x, y)$ in $D$, then $f(x, y) = g(x, y)$ 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 have finite length and are thus not strictly bandlimited. If the finite-length signal represents a finite segment of some bandlimited function, then we can apply Theorem 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 $\hat{g}(x, y)$, sign $f(x, y) = \text{sign } g(x, y)$, and $f(x, y)$ contains sign changes, then $f(x, y) = c \hat{g}(x, y)$. This is similar to a result presented in [2] 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.

At this point, it is worthwhile to consider the likelihood that a two-dimensional bandlimited function will satisfy the irreducibility constraints of Theorem 2. Although the only one-dimensional EFET's which are irreducible are of the form $f(w) = e^{iw} + (w - c)$, 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 a large nontrivial class of two-dimensional EFET's are irreducible [13]-[16]. We can precisely characterize this likelihood for polynomials, a special class of EFET's. Specifically, 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 $m > 1$) [17], and that this set is an algebraic set [18]. However, even if it could be shown that in some statistical sense, "almost all" EFET's 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. These functions will, in general, contain an infinite number of factors. 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 2 to include factorable signals.

B. Extensions

Although Theorem 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 conceptually similar to those presented in [2] for the case of periodic signals, although the mathematics and some of the conclusions are different.

One problem with Theorem 2 is that it requires two separate functions to be irreducible. If in some application, it is known that a particular signal is irreducible and satisfies the constraints of Theorem 2, then this information alone is not sufficient to guarantee that the signal is uniquely specified by its zero crossings. In particular, this theorem only guarantees that no other irreducible function can have the same zero crossings; it does not guarantee that there are no other reducible functions with those zero crossings. In the case of periodic signals represented as a Fourier series polynomial [2], [3], it was possible to eliminate this possibility when the exact bandwidth of the signal was known, since multiplying by an additional factor raises the degree of the Fourier series polynomial and thus increases the bandwidth of the signal. With the representation we are using for arbitrary signals, a similar extension is not necessarily valid since multiplication of an arbitrary signal by a polynomial will not necessarily increase its bandwidth. (A similar problem is encountered by Sanz [19] in discussing the reconstruction of multidimensional signals from algebraic sampling contours.) We can, however, establish the following result (see the Appendix for proof).

**Theorem 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 $D \subset \mathbb{R}^2$ and sign $f(x, y) = g(x, y)$ for all values of $(x, y)$ in $D$, then $g(x, y) = p(x, y) f(x, y)$ where $p(x, y)$ is a polynomial with no real zeros in the region $D$, except possibly where $f(x, y) = g(x, y) = 0$.

While this result does not guarantee uniqueness in the case where only one signal is guaranteed to be irreducible, it is possible that this theorem could be used in conjunction with the constraints of a particular application to guarantee uniqueness in that particular case.

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 similar result which allows the signals to be specified by crossings of an arbitrary threshold rather than simply zero crossings. This is important in applications such as image processing where signals represent energy or intensity and thus are constrained to be positive. These signals contain no zero crossings, but may contain points (contours) where the signal crosses a particular threshold. More generally, it is possible to allow crossings of an arbitrary bandlimited function. In particular, let us state (see the Appendix for proof) the following.

**Theorem 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 $D \subset \mathbb{R}^2$ and 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 2 which we will develop here is to allow reducible signals. This extension is important since, as mentioned earlier, we cannot precisely state the likelihood that a signal is irreducible. The reasoning used to develop this result is similar to the reasoning used in [2] 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 only consider the case where 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 the Appendix for proof).

**Theorem 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 bounded and closed region \(D\), then \(f(x, y) = cg(x, y)\) for some positive real constant \(c\).

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 [2] are not a special case of the results presented here. This is because the results presented in [2] consider the possible factorization of a signal in terms of a polynomial in \(e^{j\theta}\) and \(e^{j\lambda}\), whereas the results presented in this correspondence 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 \(e^{j\theta}\) and \(e^{j\lambda}\), but reducible as an entire function in \(s, w\), as in the case, for example, with the function \(f(s, w) = 1 - e^{j\theta}e^{j\lambda} = (1 - e^{j(\theta/2)})(1 + e^{j(\theta/2)}e^{j(\lambda/2)})\). A similar problem is mentioned by Sanz and Huang [14] when comparing their work on the reconstruction of signals from magnitude or phase to the work of Hayes [20]. In this case, it was found that the discrete-time problem considered by [20] is not a special case of the continuous-time problem considered by [14]. As is discussed in [14], 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 [2], [3], and [20] 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 work and in [14]. The interested reader should consult [14] for further details.

**IV. Conclusions**

In this correspondence, we have presented new results on the unique specification of arbitrary (nonperiodic) two-dimensional signals with zero crossing or threshold crossing information. Our primary result established that two-dimensional bandlimited signals which are irreducible as entire functions are uniquely specified to within a scale factor by their zero crossing contours. We also extended this result to permit crossings of an arbitrary threshold and to permit factorable signals with sign changes in each factor. Since previous results on unique specification of two-dimensional signals with zero crossings have required that the signal be bandpass or periodic or that a sine wave be added to the original signal, the results in this work represent an important generalization and extension of previous results. These results suggest practical applications in multidimensional signal processing, image processing, and vision, as well as the possibility for use as an analytical tool in areas such as communications and sampling theory.

**Proof of Theorem 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 [9, 10] that these functions can be extended to \(C^2\) as entire functions of an exponential type (EFET) denoted as \(f(s, w)\) and \(g(s, w)\), which are also EFET's in each variable separately, and have at most polynomial growth in the real plane. If \(f(x, y)\) and \(g(x, y)\) have finite energy (in the real plane), then their Fourier transforms exist and this result is known as the Polya–Plancherel theorem [8].

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) \(f(x, y) = 0\) [12]. The same is true for \(g(x, y)\). If \(f(x, y)\) 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 [7, proof of Theorem 4.11, p. 72]) which is contained in \(D\). The set \(V_f\) is an analytic set [8, p. 217] and is also an analytic surface \([7, p. 71]\) over a (complex) closed, bounded domain \(E: D \subset C \subset C^2\) since the analytic set is irreducible (because the related EFET is irreducible [11]). 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 1, the sets coincide 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 [14].

**Theorem A1:** Let \(f, g: C^2 \rightarrow C\) be entire functions such that \(V_f = V_g\). If \(g\) is irreducible, then there exists an entire function \(h: C^2 \rightarrow C\) that satisfies \(f = gh\).

Thus, we now have

\[
(f(s, w) = h(s, w)g(s, w) (A1)
\]

where \((h(s, w)\) is entire and nonzero everywhere in \(C^2\). Using growth arguments as in [14, p. 1448] or by applying [21, 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 \(e^{j\alpha x + j\beta y}\). Since \(f(s, w)\) and \(g(s, w)\) must be real for all real values of \(s\) and \(w\), then \(\alpha, \beta \) must all be real. We can also show that \(\alpha \) and \(\beta \) 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 [10]. Thus, \(h(s, w) = e^{j\gamma} = c(f(x, y)) = cg(x, y)\).

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 contours.

**Proof of Theorem 3**

Proceeding as in the proof of Theorem 2, we have

\[
g(s, w) = f(s, w)k(s, w) (A2)
\]

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 relationship among the bandwidths of \(f, g, k\) by applying known properties of the so-called \(P\) indicators.

**Theorem A2** [8, Theorem 3.4]: Let \(f(z)\) and \(k(z)\) be EFET's such that \(f(x)\) is real, then \(1 + \lambda h(z)\) has completely regular growth in the variable \(w \in C^2\). Then the \(P\) indicator of the function \(g(z) = f(z)k(z)\) is the sum of the \(P\) indicators of \(f(z)\) and \(k(z)\). The requirement that \(f, g\) and \(k\) be strictly bandlimited (and thus any one-dimensional slice of \(f\) or \(g\) is bandlimited) guarantees that they have completely regular growth along any slice [1]. The \(P\) indicator corresponds to the smallest convex domain completely containing the region of support of the spectrum (see [8] for precise definitions). Roughly speaking, Theorem A2 states that the bandwidth of \(g\) is the sum of the bandwidths of \(f\) and \(k\).
However, in Theorem 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 [8]. Thus, $k$ must be a polynomial of finite degree [19]. This polynomial cannot have real zeros in the region $D$, except possibly where $f(x, y) = g(x, y) = 0$, and the theorem is proven.

**Proof of Theorem 4**

The proof is straightforward due to Theorem 2. The functions $f_j(x, y) = f(x, y) - h(x, y)$ and $g_j(x, y) = g(x, y) - h(x, y)$ satisfy the constraints of Theorem 2. Note that if the functions $f$, $g$, $h$ are bandlimited in the general sense, then so are $f_j$ and $g_j$.

**Proof of Theorem 5**

The proof of this theorem is similar to the proof of Theorem 2 applied to each factor separately. According to Osgood's product [12,13,11], any EFET in $n$ dimensions can be represented uniquely by the product of either a finite or an infinite set of irreducible EFET's. We will 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 factor point where $f(s, w) = 0$ and $g(s, w) = 0$, at least one of the factors $f_j(s, w)$ must be zero and at least one of the factors $g_j(s, w)$ must be zero. For each zero contour of $f(s, w)$ and $g(s, w)$ corresponding to the irreducible factors $f_j(s, w)$ and $g_j(s, w)$, we can use Theorem 2 to show that $f_j(s, w) = c g_j(s, w)$. 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 $f_j(s, w)$, is complex, then $f_j(s, w)$ will also be a factor of $f(s, w)$ and thus $f(s, w)$ will contain a real factor $f_j(s, w) f_0(s, w)$ which is nonnegative everywhere, violating the constraints of the theorem. Thus, the factor $f_j(s, w)$ 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 $f(x, y) = c g(x, y)$.

**Acknowledgment**

Our sincere thanks to an anonymous reviewer who carefully reviewed the results of this paper and pointed out an error in our original statement of Theorem 3.

**References**


