Ballena

systems corporation

UTR92-017

LA-SUB--93-283

Wideband Pulse Reconstruction **RECEIVED**
from Sparse Spectral-Amplitude Data **FFR 14 1997**

OSTI

Final Report

Submitted by

Ballena Systems Corporation

901 18th Street

Suite 113

Los Alamos, New Mexico 87544

Submitted to

Foreign Technology Center

Wright-Patterson Air Force Base

Ohio 45433

January 1993

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

Work funded through Los Alamos National Laboratory
Subcontract 9-XS1-U6573-1

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Wideband Pulse Reconstruction From Sparse Spectral-Amplitude Data

Kendall F. Casey
Brian A. Baertlein

Ballena Systems Corporation*

January 1993

Abstract

Methods are investigated for reconstructing a wideband time-domain pulse waveform from a sparse set of samples of its frequency-domain amplitude spectrum. Approaches are outlined which comprise various means of spectrum interpolation followed by phase retrieval. Methods for phase retrieval are reviewed, and it is concluded that useful results can only be obtained by assuming a minimum-phase solution. Two reconstruction algorithms are proposed. The first is based upon the use of Cauchy's technique for estimating the amplitude spectrum in the form of a ratio of polynomials. The second uses B-spline interpolation among the sampled values to reconstruct this spectrum. Reconstruction of the time-domain waveform via inverse Fourier transformation follows, based on the assumption of minimum phase. Representative numerical results are given.

*5820 Stoneridge Mall Road, Suite 205, Pleasanton, CA 94588. (510) 460-3740.

1 Introduction

In the fields of communication, remote sensing, and radio frequency (rf) weapon development there exists a trend toward the use of electromagnetic waveforms which have very broad bandwidths. This trend is motivated by several factors, including the potential for greater information content in broadband signals, the development of high-power pulsed electromagnetic sources, and the decreased potential for jamming and interception. The remote detection and analysis of such signals is of interest. For reasons which are both historical and implementation-related, however, many—if not most—existing sensors of electromagnetic waveforms have narrowband frequency responses. For such sensors it is usually only the incident waveform's power which is of interest; the phase information in these signals is generally not significant.

The problem addressed in this work is that of reconstructing a broadband time-domain pulse waveform from a relatively small number of samples of its frequency-domain amplitude spectrum. This problem arises not only because of the typically narrow bandwidths of radio receivers alluded to above, but also because of the limited number of such receivers which can be deployed for the detection of transient waveforms of interest. One scenario for such a situation is shown in Figure 1: a source of a transient electromagnetic field is located on or near the earth's surface and one or more receivers on satellite platforms are deployed to detect this field. The individual receiver bandwidths are narrow, and the receivers are tuned to different center frequencies. It is also assumed that dispersion (caused, for example, by transionospheric propagation)¹ gives rise to loss of phase information and that information regarding the relative times of arrival at the various receivers is not available. Thus only amplitude data can be used to draw inferences regarding the transient signal.

In this report we explore the possibility of combining the outputs of multiple, narrowband, phase-insensitive, electromagnetic sensors to synthesize a virtual sensor capable of recovering wideband time-domain waveforms. We will show that several existing mathematical techniques can be brought to bear on the problem, and that faithful reconstruction of a broadband waveform is possible under reasonably general conditions. The goals of the study include not only the development

¹Note that transionospheric propagation will also remove essentially all signal information at frequencies below the ionospheric cutoff frequency.

Figure 1: An example scenario for transient signal detection.

of reliable and robust means for reconstruction of the time-domain signal, but also the development of receiver deployment strategies (that is, selection of center frequencies) which enhance the reliability and robustness of the reconstruction.

The work is organized in five major sections. The electromagnetic field produced by a representative transient source is described in Section 2. In Section 3 we state the technical problem, we review the constraints imposed by electromagnetic theory on the received waveform, and we outline our solution approach. In Sections 4 and 5 respectively we review the fundamental processes of spectrum reconstruction and phase retrieval. Numerical examples of our solution procedures are presented in Section 6. Concluding remarks and suggestions for further investigation are given in Section 7. Listings of the FORTRAN programs which implement the solution procedures are given in the Appendix.

Figure 2: The VPD-II HEMP environment simulator.

2 A Transient Field Source: VPD-II

To motivate the investigation, we begin by considering a specific example source of interest, a vertically polarized high-altitude electromagnetic pulse (HEMP) environment simulator. Several such simulators exist worldwide; we consider the Vertically Polarized Dipole simulator VPD-II located at Kirtland Air Force Base, New Mexico. In what follows, we describe the simulator and develop a simple model to describe its radiated electromagnetic field. In particular, we show that the Fourier transform of this field can be approximately, but accurately, represented in terms of a ratio of polynomials in $i\omega$.

The VPD-II is a resistively loaded vertical monocone antenna about 40 meters in height, located over a conducting ground plane. It is fed at its base by a Marx generator. The fields radiated by this HEMP environment simulator, which is shown in Figure 2, have been studied in varying degrees of detail elsewhere [1, 2, 3]. We now consider an elementary model for the antenna and pulser.

A simple model for the VPD-II antenna is a nonuniformly loaded monopole of height h above a perfectly conducting ground plane of infinite extent, as shown in Figure 3. The loading is resistive; the resistance per unit length depends on the

Figure 3: Monopole model for VPD-II antenna. The ground plane at $z = 0$ is assumed to be perfectly conducting.

vertical position z as

$$R'(z) = \frac{2R_\infty}{h - z} \quad (0 \leq z < h) \quad (1)$$

where R_∞ denotes the characteristic impedance of the biconical transmission line formed by the antenna and the ground plane. This loading profile causes the frequency-domain antenna current $\tilde{I}(z)$ to behave as

$$\tilde{I}(z) = \tilde{I}(0)(1 - z/h)e^{-ikz} \quad (0 \leq z \leq h) \quad (2)$$

where $\tilde{I}(0)$ is the current at the feed point and $k = \omega/c$; c denotes the speed of light. The time dependence $\exp(i\omega t)$ is assumed and suppressed. The resistive loading causes the antenna current to behave as an outward-traveling wave whose amplitude smoothly decreases to zero at the end of the antenna. This behavior of the current reduces the amplitude of the field radiated from the upper end of the antenna, a desirable feature in HEMP environment simulation.

It is straightforward to calculate the frequency-domain radiated electric field

of this antenna. We find

$$\tilde{E}_\theta(r, \theta) = \frac{Z_0 \tilde{I}(0)}{2\pi r \sin \theta} e^{-ikr} g(kh; \theta) \quad (3)$$

where r and θ denote the usual spherical coordinates and Z_0 is the intrinsic impedance of free space. The function $g(kh, \theta)$ is given by

$$g(kh, \theta) = 1 - \frac{1 + \cos^2 \theta}{ikh \sin^2 \theta} + \frac{(1 + \cos \theta)^2}{2ikh \sin^2 \theta} e^{-ikh(1 - \cos \theta)} + \frac{(1 - \cos \theta)^2}{2ikh \sin^2 \theta} e^{-ikh(1 + \cos \theta)} \quad (4)$$

It is evident from eq. (4) that the radiated field comprises three distinct contributions: one emanating from the base of the antenna, one from the top of the antenna, and one from the image of the top of the antenna. We note that as the antenna height decreases toward zero, the radiated electric field reduces to

$$\tilde{E}_\theta = \frac{Z_0 \tilde{I}(0) ikh \sin \theta}{4\pi r} e^{-ikr} \quad (5)$$

so that in this limit the radiated electric field is proportional to the first derivative of the feed current; and as $h \rightarrow \infty$, we have

$$\tilde{E}_\theta = \frac{Z_0 \tilde{I}(0)}{2\pi r \sin \theta} e^{-ikr} \quad (6)$$

and the radiated electric field is proportional to the feed current.

An approximate expression for the radiated electric field in the frequency domain which reproduces this behavior in the two limiting cases is easily developed. We have

$$\tilde{E}_\theta \approx \frac{Z_0 \tilde{I}(0)}{4\pi r} e^{-ikr} \frac{ikh \sin \theta}{1 + (ikh/2) \sin^2 \theta} \quad (7)$$

The amplitudes of the exact and the approximate expressions for the normalized radiated electric field $\tilde{e}_\theta$, defined as

$$\tilde{e}_\theta = \frac{4\pi r}{Z_0 \tilde{I}(0)} \tilde{E}_\theta e^{ikr} \quad (8)$$

are compared in Figures 4 and 5. The differences in the amplitude spectra

Figure 4: Amplitudes of exact (dashed curve) and approximate (solid curve) expressions for the normalized radiated electric field vs. normalized frequency kh ; $\theta = 30^\circ$.

obviously depend upon the observation angle. These differences arise primarily because of interference among the fields radiated from the base, the top, and the image of the top of the antenna which is represented in the exact expression but not in the approximate expression for the radiated field in the frequency domain.

The input impedance Z_a of the nonuniformly loaded monopole is well known [2]. We have

$$Z_a = R_\infty + 1/(i\omega C_a) \quad (9)$$

where R_∞ (defined earlier) is the characteristic impedance of the biconical transmission line formed by the antenna and the ground plane and C_a is the capacitance of the antenna over the ground plane. Thus if the source which drives the antenna is expressed in terms of a Thévenin equivalent circuit with open-circuit voltage $\tilde{V}_{oc}$ and impedance Z_{eq} , we can express the antenna feed current $\tilde{I}(0)$ as

$$\tilde{I}(0) = \frac{\tilde{V}_{oc}}{Z_a + Z_{eq}} \quad (10)$$

The frequency-domain radiated electric field is thus expressed in approximate form

Figure 5: Amplitudes of exact (dashed curve) and approximate (solid curve) expressions for the normalized radiated electric field vs. normalized frequency kh ; $\theta = 60^\circ$.

as

$$\tilde{E}_\theta = \frac{Z_0 \tilde{V}_{oc}}{4\pi r(Z_a + Z_{eq})} \left[\frac{ikh \sin \theta}{1 + (ikh/2) \sin^2 \theta} \right] e^{-ikr} \quad (11)$$

The VPD-II antenna is fed by a Marx generator. An elementary circuit model for such a source is shown in Figure 6. The capacitance C_m represents the erect Marx capacitance and the voltage V_m is the Marx generator voltage prior to firing. The switch inductance is denoted L_s . The Thevenin equivalent source elements are simply $\tilde{V}_{oc} = V_m/(i\omega)$ and

$$Z_{eq} = i\omega L_s + 1/(i\omega C_m) \quad (12)$$

Making the appropriate substitutions in eq. (11), we obtain the following approximate expression for the radiated electric field in the frequency domain:

$$\tilde{E}_\theta = \frac{Z_0 V_m}{4\pi r(i\omega R_\infty + 1/C_t + (i\omega)^2 L_s)} \left[\frac{ikh \sin \theta}{1 + (ikh/2) \sin^2 \theta} \right] e^{-ikr} \quad (13)$$

where $1/C_t = 1/C_a + 1/C_m$. We observe that this expression (except, of course, for the propagation factor $\exp(-ikr)$) is a ratio of polynomials in the variable $i\omega$.

Figure 6: An equivalent circuit for the Marx generator.

The frequency-domain field vanishes at dc as it should, since the integral of the electric field over all time must vanish. The degree of the denominator polynomial exceeds that of the numerator by two, indicating that the early-time behavior of the field is a ramp function; the field is thus continuous at $t = 0$.

We now present some representative results. Approximate values of the relevant circuit parameters for VPD-II are as follows: $R_\infty = 60$ ohms, $C_a = 2.9$ nF, $C_m = 5.4$ nF, and $L_s = 0.25$ μ H. The antenna height h is taken to be 50 meters.² Normalized frequency-domain amplitude spectra of the radiated electric field are shown in Figures 7 and 8, and in Figures 9 through 12 are shown plots of the normalized time-domain electric field $e_\theta(\theta; t) = rE_\theta(r, \theta; t - r/c)/V_m$ as a function of time for $\theta = 30^\circ$ and $\theta = 60^\circ$. Both the exact and the approximate representations for the frequency-domain fields of the loaded vertical monopole antenna were used in the calculations. It will be noted that the differences between the curves are not great in either the time or frequency domains. The principal difference between the exact and the approximate results in the time domain is the smoothing, in the approximate results, of the slope discontinuities which occur in the exact results. These discontinuities, which are particularly evident in Figure

²This figure approximates the slant height (52.6 m) of the VPD-II antenna.

Figure 7: Normalized electric-field spectra for VPD-II models vs. normalized frequency kh ; $\theta = 30^\circ$. The dashed curve describes the exact model.

Figure 8: Normalized electric-field spectra for VPD-II models vs. normalized frequency kh ; $\theta = 60^\circ$. The dashed curve describes the exact model.

Figure 9: Normalized electric field vs. time for approximate VPD-II model; $\theta = 30^\circ$.

Figure 10: Normalized electric field vs. time for exact VPD-II model; $\theta = 30^\circ$.

Figure 11: Normalized electric field vs. time for approximate VPD-II model; $\theta = 60^\circ$.

Figure 12: Normalized electric field vs. time for exact VPD-II model; $\theta = 60^\circ$.

12, occur at the arrival times of the fields radiated from the upper end of the antenna and from the image of this point.

The radiated-field waveforms obtained in the foregoing are representative of those observed from VPD-II and other HEMP environment simulators. The electric field rises rapidly to a peak, then drops through zero to a less pronounced negative peak, and finally returns slowly to zero. The field has significant spectral content at frequencies up to several tens of megahertz.

A Remark on Ionospheric Propagation Effects

Propagation of such a pulse signal through the earth's ionosphere results in drastic changes in the waveform. These changes result from (1) loss of spectral content at frequencies below the ionospheric cutoff frequency (this frequency is approximately equal to $\omega_{pm} \sec \theta$, where ω_{pm} is the maximum (radian) plasma frequency ω_p encountered along the propagation path); and (2) phase dispersion above the cutoff frequency. The phase dispersion is represented approximately by a transfer function of the form

$$H(i\omega) = \exp(-\omega_0/i\omega) \quad (\omega > \omega_{pm} \sec \theta) \quad (14)$$

where the frequency ω_0 can be expressed as

$$\omega_0 = \frac{\omega_{pm}^2 L_e}{2c} \sec \theta \quad (15)$$

L_e denotes an effective ionospheric depth and is given by

$$L_e = \frac{1}{\omega_{pm}^2} \int_0^z \omega_p^2(z') dz' \quad (16)$$

The ionospherically propagated signal is stretched in time and behaves as a downwardly chirped, frequency modulated pulse.

The signal of interest for reconstruction purposes is more likely to be the original pulse than the ionospherically dispersed waveform. Thus the more important effect of the ionosphere on the signal is the loss of its low-frequency spectral content. Reconstruction of the original undispersed pulse must be based on the information contained in its high-frequency amplitude spectral content, which is relatively unaffected by propagation through the ionosphere.

3 Problem Statement; Solution Formulation

It is appropriate now to state in precise terms the mathematical problem to be considered and to discuss the relevant constraints on its solution.

Let $g(t)$ be an unknown, causal, time-domain signal, and let $f(t)$ be the signal formed by convolving g with a known, linear filter $h(t)$. We write

$$f(t) = (g \otimes h)(t) \quad (17)$$

We denote the Laplace and Fourier transforms of $f(t)$ by $F(s)$ and $F(i\omega)$ respectively. We wish to estimate the time-domain waveform $g(t)$ from a set of discrete spectral data of the form

$$\mathcal{D} = \{|F(i\omega_k)|, k = 1, 2, \dots, N\} \quad (18)$$

The physics of electromagnetic radiation lead to several constraints on the function g , most of which are also applicable to f . Specifically,

1. radiating electromagnetic fields have no dc component:

$$\lim_{\omega \rightarrow 0} F(i\omega) = 0 \quad (19)$$

2. the radiated waveform is real valued:

$$F(-i\omega) = F^*(i\omega) \quad (20)$$

3. the energy in the radiated waveform is finite:

$$\int_{-\infty}^{\infty} [f(t)]^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(i\omega)|^2 d\omega < \infty \quad (21)$$

4. the radiated waveform is a causal function of time $\Rightarrow$ all singularities of $F(s)$ occur in the half plane $\Re(s) < 0$.

In special situations there may exist additional constraints. Some of the more important cases are

- if the radiated waveform is time limited, then $F(i\omega)$ is an analytic function in the entire complex plane (holomorphic), and
- if the radiated waveform is smoothly varying with a continuous derivative of order n , then $F(i\omega) \sim O(\omega^{-n-2})$ for large ω .

In this work we discuss methods for solving the simpler problem of estimating $f(t)$ from the data $\{|F(i\omega_k)|\}$. Estimates of g follow from suitable filtering and deconvolution operations. These operations are, in general, nontrivial to perform, but they are adequately addressed elsewhere [4, 5] and will not be reviewed here.

Although several approaches to the problem can be envisioned, the authors have found that the most profitable approach comprises the following steps:

1. **Spectral Reconstruction:** Produce an amplitude spectrum estimate $|\hat{F}(i\omega)|$ from the available discrete data $\mathcal{D}$.
2. **Phase Recovery:** From the spectral amplitude estimate $|\hat{F}(i\omega)|$ construct a complex spectral response $\hat{F}(i\omega)$ which will satisfy the foregoing constraints. The time-domain waveform estimate $\hat{f}(t)$ follows immediately via an inverse Fourier transform.
3. **Quality Assessment:** Develop confidence estimates for the reconstructed function $\hat{f}$.

A few comments on this procedure are in order.

The first step, spectral reconstruction, is essentially a problem in interpolation of sparse data. Related problems arise in the construction of models from sparse experimental data (cf. [6, ch. 14]). Numerous approaches to spectral reconstruction can be identified, the more useful of which are reviewed in Section 4.

For the second step, phase retrieval, there appear to be a limited number of workable approaches. We find that only techniques which employ a minimum-phase assumption are feasible here. These methods are described in Section 5.

The final step, quality assessment, is an important but ill-defined component of problems of this type. Its importance can be made manifest by noting that the first two steps define a nearly automatic procedure for estimating $\hat{f}$ from $\mathcal{D}$. An estimate can always be produced, no matter what the quality of the

data. Without an assessment of a model's goodness-of-fit, users of the derived information will have no idea what confidence to place in their results. For the problem of interest here, there exists no well-defined theoretical basis for describing the relation between errors in the data set $\mathcal{D}$ and errors in the final waveform estimate $\hat{f}(t)$. One can employ numerical methods to study the results of errors in $\mathcal{D}$, but it is difficult to obtain insight in this manner. Although a systematic error analysis of our algorithms would be useful, such an analysis is beyond the scope of our investigation at this time. Limited further discussion of this topic appears in the Concluding Remarks.

4 Spectral Reconstruction

In this section we describe techniques with which one can reconstruct a continuous function from sampled data. These techniques are essentially interpolation procedures. In the present work it is found that parametric interpolation methods are most convenient since they permit one to easily enforce the waveform constraints of Section 3. In Subsections 4.1 and 4.2 we discuss rational-function interpolation and B-spline interpolation, both of which have been found useful in the present investigation.

It is relevant to note that the spectral data contained in $\mathcal{D}$ are not necessarily sampled at uniform increments in frequency, but the use of unequally spaced sample points ω_k does not pose a problem in principle since one can always construct a monotonic, continuously differentiable function (e.g., a quadratic spline) which maps the unequally spaced samples onto an equally spaced set of points. In fact, non-uniform spacing can be a very effective means of reducing the data requirements. Further discussion of this aspect of the problem appears in Subsection 4.3.

Finally, we note that the resulting spectrum estimate $|\hat{F}(i\omega)|$ is required to be non-negative. This condition is often difficult to enforce explicitly, and small regions of ω in which $|\hat{F}(i\omega)| < 0$ may arise in practice. Nonetheless, this phenomenon has not been found to pose an insurmountable problem in numerical experiments.

4.1 Rational Function Interpolation

In many cases the time-domain waveforms of interest are well approximated by sums of decaying exponentials, the Laplace transforms of which are ratios of polynomials in the transform variable.³ In this section we describe the use of rational-function interpolation for spectrum reconstruction. An extensive theoretical discussion of this topic has been presented by Stoer and Bulirsch [7, §2.2], whose work may be consulted for more details.

Using the rational-function model, we obtain the following expression for the

³Exponential-series models are of considerable interest in electromagnetic theory, since it is often possible to establish a physical basis for this behavior in terms of the decaying characteristic resonances of a radiating or scattering structure.

amplitude spectrum:

$$|\hat{F}(i\omega)|^2 = \frac{a_0 + a_1\omega^2 + \dots + a_m\omega^{2m}}{1 + b_1\omega^2 + \dots + b_n\omega^{2n}} \quad (22)$$

where the coefficients $\{a_k\}$ and $\{b_k\}$ and the model orders m and n are to be determined.

The constraints given in Section 3 require $a_0 = 0$ and $m \geq 1$ in order that the dc component vanish. In this formulation we will also impose the requirement that $\hat{f}$ be continuous at $t = 0$. In the Fourier domain we express the latter constraint as follows:

$$\lim_{\omega \rightarrow \infty} \omega^2 \hat{F}(i\omega) < \infty \quad (23)$$

which requires that the model orders satisfy $n \geq m + 2$. Incorporating these results we have

$$|\hat{F}(i\omega)|^2 = \frac{a_1\omega^2 + \dots + a_m\omega^{2m}}{1 + b_1\omega^2 + \dots + b_n\omega^{2n}} \quad (24)$$

It remains to determine the model orders (m, n) and coefficients $\{a_i, b_i\}$ which will most accurately interpolate the data.

A number of algorithms have been developed for the determination of the coefficients of an interpolating rational function (cf. [7]), but the constraints imposed above make it difficult to apply the majority of these algorithms. A useful approach to this problem was developed by Cauchy [8] who obtained the function coefficients as the solution of a least-squares problem. Some recent examples and extensions of this technique have been given by Kottapalli et al. [9]. Equating the model in equation (24) to the data points $|F(i\omega_k)|$ at the N discrete frequencies ω_k , we obtain the equations

$$|F(i\omega_k)|^2 = \sum_{j=1}^m a_j \omega_k^{2j} - |F(i\omega_k)|^2 \sum_{j=1}^n b_j \omega_k^{2j} \quad k = 1, 2, \dots, N \quad (25)$$

This set of simultaneous linear equations can be solved using standard methods for least-squares problems. We observe that the sum of the model orders $m + n$ must be no larger than N , while the situation $m + n < N$ constitutes a permissible least-squares problem.

4.2 B-Spline Interpolation

Polynomial splines are another possible approach to the problem of constrained interpolation.⁴ The discussion presented here emphasizes B-splines, so named because of their bell-shape. A recent, concise discussion of the theory of these functions has been presented by Unser and Aldroubi [10] and forms the basis for the exposition that follows.

B-splines are defined in a recursive manner. The B-spline of order 0, which we denote $\beta^0(x)$, is the characteristic function in the interval $[-1/2, 1/2)$, viz:

$$\beta^0(x) = \begin{cases} 1, & |x| < 1/2 \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

Higher-order splines are generated from β^0 as follows:

$$\beta^n(x) = [\beta^0 \otimes \beta^{n-1}](x) \quad (27)$$

where $\otimes$ signifies convolution. One can obtain the explicit formula

$$\beta^n(x) = \sum_{j=0}^{n+1} \frac{(-1)^j}{n!} \binom{n+1}{j} [x - j + (n+1)/2]_+^n \quad (28)$$

where

$$[x]_+^n = \max(0, x)^n \quad (29)$$

The B-splines of orders 1, 2, and 3 are given by

$$\beta^1(x) = \begin{cases} 1 - |x|, & |x| < 1 \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

$$\beta^2(x) = \begin{cases} [(3/2 - |x|)^2 - 3(1/2 - |x|)^2]/2!, & |x| < 1/2 \\ (3/2 - |x|)^2/2!, & 1/2 < |x| < 3/2 \\ 0, & \text{otherwise} \end{cases} \quad (31)$$

$$\beta^3(x) = \begin{cases} [(2 - |x|)^3 - 4(1 - |x|)^3]/3!, & |x| < 1 \\ (2 - |x|)^3/3!, & 1 < |x| < 2 \\ 0, & \text{otherwise} \end{cases} \quad (32)$$

⁴Entire-domain polynomial expansions are often used in applied work because of their simplicity and ease of analytical manipulation. For the problem under consideration here, however, it is difficult to construct a whole-domain polynomial interpolator which will both be non-negative and satisfy the constraints identified in Section 3.

The function β^3 is commonly known as a cubic spline and is widely used as a basis for interpolation.

It is not difficult to show that the first $n-1$ derivatives of $\beta^n(x)$ are continuous for all x . In the limit of large n , arguments which parallel the proof of the central limit theorem yield a Gaussian form, viz:

$$\beta^n(x) \sim \sqrt{\frac{6}{\pi(n+1)}} e^{-6x^2/(n+1)} \quad (n \rightarrow \infty) \quad (33)$$

For reference, we note that the Fourier transform of β^n is

$$\mathcal{F}[\beta^n] = \left[\frac{\sin(\omega/2)}{(\omega/2)} \right]^{n+1} \quad (34)$$

Several characteristics of B-splines make them useful for the problem under consideration. We note first that the B-splines are positive definite, which is appropriate for the interpolation of a spectrum. There is, however, no straightforward way to enforce the requirement that the interpolated function be positive definite.

Second, because splines also have compact support, it is possible to formulate the solution for the interpolation coefficients as a matrix equation wherein the nonzero terms of the matrix are confined to a limited band near the diagonal of the matrix. In practice, however, it is more convenient to derive the coefficients via a somewhat more efficient (but more complicated) algorithm which obviates the solution of this banded matrix equation. The cubic-spline form of this algorithm, which can be found in standard references [7, §2.4], enforces continuity of the interpolating function and its first two derivatives at all of the data points. For the purposes of our work here it is convenient to supplement the data set $\mathcal{D}$ with two additional data points, namely:

$$|F(0)| = 0 \quad (35)$$

$$|F(i\omega_{N+1})| = 0 \quad (36)$$

The first of these relations enforces the requirement of zero dc component. In the second constraint ω_{N+1} is an arbitrarily imposed upper limit on the spectral range of the waveform. The method also requires the values of the first derivative of the

interpolating function at the boundaries. In our work we have used the so-called natural boundary condition at dc, namely:

$$\frac{d}{d\omega}|F(0)| = \frac{|F(i\omega_1)|}{|i\omega_1|} \quad (37)$$

(i.e., an approximate, numerically derived, derivative) and we have forced the derivative at ω_{N+1} to vanish.

4.3 Other Methods; Sampling Considerations

A number of methods have been developed for interpolation of discrete data, but for the majority of these methods it is difficult to enforce the constraints described in Section 3. In spite of the limitations of these methods, some of them provide important theoretical insight. In this section we present a short review of other methods for interpolation, and we comment on their applicability to the question of optimal sampling.

It is well known that a uniformly sampled waveform can be reconstructed via interpolation of the sampled data with functions of the form $\sin(x)/x$. This result holds when the sampling rate exceeds the bandwidth of the data (cf. [11, §5.1]). When the spectral content of the signal f is non-zero only over the interval $|\omega| < \sigma$ we have

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\Delta_t) \frac{\sin \sigma(t - n\Delta_t)}{\sigma(t - n\Delta_t)} \quad (38)$$

where $\Delta_t = \pi/\sigma$ is the temporal sampling increment. By virtue of the time-frequency duality of the Fourier transform, it follows that we can interpolate a time-limited ($|t| < \tau$) function with the series

$$F(i\omega) = \sum_{n=-\infty}^{\infty} F(in\Delta_\omega) \frac{\sin \tau(\omega - n\Delta_\omega)}{\tau(\omega - n\Delta_\omega)} \quad (39)$$

where $\Delta_\omega = \pi/\tau$ is the radian frequency sampling increment.

When the samples are not uniformly spaced in time or frequency, one can still formulate relations similar to those given above by remapping the samples to a uniform spacing. To show this result, we consider a real function $f(t)$ with Fourier

transform $F(i\omega)$. Let $F(i\omega)$ be known at the discrete points ω_k , $k = 0, 1, \dots, N$. Define a differentiable, one-to-one function $T(\omega)$ such that

$$k\Delta_\Omega = T(\omega_k) \quad (40)$$

where Δ_Ω is a constant interval in the transformed radian frequency Ω . Finally, let $S(\Omega)$ be the inverse function for T such that

$$S[T(\omega)] = \omega \quad (41)$$

and define

$$G(i\Omega) = F[iS(\Omega)] \quad (42)$$

With the assumption that f is time-limited on an interval τ , we can approximate $G(i\Omega)$ via an interpolation similar to that in equation (39). We have

$$G(i\Omega) \approx \sum_{n=-N}^N G(in\Delta_\Omega) \frac{\sin \tau(\Omega - n\Delta_\Omega)}{\tau(\Omega - n\Delta_\Omega)} \quad (43)$$

where the values of G for negative frequency ($n < 0$) are related to the values at positive frequency via equation (20). Equation (43), like equation (39) above, becomes exact as the number of sample points N goes to infinity. It is clear that requirements for sample rates and number of samples in the Ω domain can be significantly different from those in the ω domain. Related results have been presented previously by Soumekh [12] and Levi [13]. The function $F(i\omega)$ is obtained by inverting the transformation between Ω and ω , and the desired time-domain waveform $f(t)$ is thus recovered.

The use of a suitable mapping T allows us to reduce the number of points required for interpolation. As an example, consider the so-called double-exponential pulse given by

$$f(t) = [e^{-\beta t} - e^{-\alpha t}] U(t) \quad (44)$$

where α and β are constants which control respectively the rise and fall time of the waveform, and $U(t)$ is the unit step function. For $\alpha \gg \beta$ this waveform has a rise time (from 10% to 90% of peak) of $\approx 2.2/\alpha$, and a fall time (decay to 10% of peak) of $\approx 2.3/\beta$. Hence, to capture the essential aspects of this waveform with uniform time-domain sampling, one must use a sampling interval $\Delta_t \sim 2/\alpha$ and a total number of points on the order of $(2/\beta)/\Delta_t = \alpha/\beta \gg 1$. A plot of this

Figure 13: The double-exponential pulse.

function for $\alpha = 100$, $\beta = 3$ is shown in Figure 13. A total of 256 time-domain samples are displayed in this figure.

The Fourier transform of this waveform is given by

$$F(i\omega) = \frac{\alpha - \beta}{(i\omega + \alpha)(i\omega + \beta)} \quad (45)$$

This frequency-domain dependence is illustrated in Figure 14 in an area-preserving presentation.⁵ For $\alpha \gg \beta$ one finds analytically that $F(i\omega)$ is largely independent of frequency for $\omega < \beta$, while a rolloff as ω^{-2} is noted for $\omega > \alpha$.

A general, effective method for sparse sampling of the frequency-domain data can be formulated as follows: Let $C_p(\omega)$ be the following normalized integral of the power-law remapped spectral content:

$$C_p(\omega) = \frac{\int_{-\omega}^{\omega} |F(i\omega')|^p d\omega'}{\int_{-\infty}^{\infty} |F(i\omega')|^p d\omega'} \quad (46)$$

⁵An area-preserving presentation of a function $f(x)$ plots $xf(x)$ versus $\ln(x)$. This display has the property that the area under the plot is proportional to the integral of the function.

Figure 14: An area-preserving plot of the spectral content of a double-exponential pulse.

For $p = 2$ this function gives the cumulative energy distribution of $F(i\omega)$, but in our work we have found $p = 1$ to be most useful. For any realizable function $f(t)$ the Paley-Wiener theorem requires that the quantity $F(i\omega)$ cannot vanish over any finite interval in ω . Hence, the integral C_p defines a differentiable, one-to-one mapping from the frequency variable ω to the interval $[0, 1]$.

A parsimonious sampling of $F(i\omega)$ is obtained by remapping a uniform sampling of $[0, 1]$ via C_p^{-1} . The virtue of such a sampling is that it places the highest density of points in regions where the integral of $|F(i\omega)|^p$ is growing most rapidly. As an example, consider the double-exponential pulse for the case $p = 1$. A plot of C_1 for the above pulse is shown in Figure 15. Dividing the interval $[0, 1]$ into eight equally spaced subintervals and remapping the subinterval endpoints to the frequency domain, we obtain the sample points listed in Table 1. As expected, these points lie in that region in Figure 14 where the integral of $|F(i\omega)|$ has its largest contributions, and the density of sample points increases near the peak of that curve.

Figure 16 compares inverse Fourier transforms obtained via (1) 128 complex samples of the exact spectral amplitude (at positive frequencies) and (2) simplistic

Figure 15: The normalized cumulative spectral content C_1 for the double-exponential pulse.

Table 1: Data points used in a sparse sampling of $F(i\omega)$.

f_k
0.301
0.681
1.24
2.29
4.34
8.63
19.6

Figure 16: A comparison of double exponential pulses obtained from Fourier domain sampling at uniform intervals (128 samples) and exponential intervals (8 samples) and simplistic interpolation.

(linear) interpolation of the eight complex samples defined above. We observe that these results are in good agreement despite the use of 16 times fewer data samples in the approximate result.

We conclude from this discussion that by using judiciously chosen sample points ω_k it is possible to reconstruct a wideband waveform from a data set that is substantially smaller than would be required from a straightforward application of the Nyquist sampling rate. Conversely, the use of a given set of sample points has obvious implications for the waveforms which can be reconstructed. The optimum sampling will employ higher sample densities in frequency regimes where large spectral contributions are expected. Knowledge of the expected spectral structure of signals of interest can thus be used to identify the optimum center frequencies for a suite of receivers deployed to capture and reconstruct those signals.

5 Phase Retrieval

Reconstruction of the phase of a complex signal from its magnitude, a procedure which is sometimes referred to as “factorization”, is a problem of interest in many fields, and a number of approaches to this problem have been developed. In this section we review techniques for phase retrieval. Our discussion concerns itself primarily with minimum-phase reconstruction, which we have found to be the most practical method for the problems considered in this work. An informative overview of phase retrieval is given by Goodman [14].

5.1 Minimum Phase Factorization

The Fourier transform $F(i\omega)$ of a real, causal function f has several important properties which we noted in Section 3. Recall that the Laplace transform $F(s)$ will be an analytic function for $\Re(s) > 0$. We can also show that the real and imaginary parts of $F(i\omega) = R(i\omega) + iX(i\omega)$ will form a Hilbert transform pair, viz:⁶

$$X(i\omega) = -\frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{R(i\omega')}{\omega - \omega'} d\omega' \quad (47)$$

$$R(i\omega) = \frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{X(i\omega')}{\omega - \omega'} d\omega' \quad (48)$$

where PV indicates the principal value of the integral. For the purposes of this work we define the Hilbert transform operator $\mathcal{H}$ as follows:

$$\mathcal{H}[F](\omega) = -\frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{F(i\omega')}{\omega - \omega'} d\omega' \quad (49)$$

It follows that $X(i\omega) = \mathcal{H}[R](\omega)$ and $R(i\omega) = \mathcal{H}^{-1}[X](\omega)$. It is evident from the foregoing results that the inverse of $\mathcal{H}$ is simply related to the forward transform. We have

$$\mathcal{H}^{-1}[F] = -\mathcal{H}[F] \quad (50)$$

⁶Note that the negative of $\mathcal{H}$ is also referred to as a Hilbert transform by some authors. The choice of sign can be related to the convention used for the exponent in the Fourier transform. The present work follows the convention of Papoulis [16] who uses $e^{-i\omega t}$ for the forward transform.

In addition, if f is a real causal signal, then the analytic signal $z(t)$ formed by an inverse Fourier transform over twice the positive frequencies, viz:

$$z(t) = \frac{1}{2\pi} \int_0^{\infty} 2F(i\omega') e^{i\omega' t} d\omega' \quad (51)$$

is given by

$$z(t) = f(t) + i\mathcal{H}[f] \quad (52)$$

Clearly,

$$Z(\omega) = 2F(i\omega)U(\omega) \quad (53)$$

The discrete-time analogs of these results are straightforward to obtain and are discussed extensively by Oppenheim and Schaffer [15, ch. 7].

A function $F(s)$ is said to be minimum phase if it is analytic in $\Re(s) \geq 0$ and $1/F(s)$ is also analytic in the same region. The following theorem will be important in this work:

Theorem 1 *If*

$$F(i\omega) = e^{-\alpha(\omega) - i\theta(\omega)} \quad (54)$$

is minimum phase, then

$$\theta(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\omega \alpha(\omega')}{\omega^2 - \omega'^2} d\omega' \quad (55)$$

$$\alpha(\omega) = \alpha(0) + \frac{\omega^2}{\pi} \int_{-\infty}^{\infty} \frac{\theta(\omega')}{\omega'(\omega^2 - \omega'^2)} d\omega' \quad (56)$$

The proof of this result is relatively involved (cf. [16, pp. 206-208]) and will be omitted.

Equation (55) implies that if $f(t)$ is a real causal function with the Fourier transform $F(i\omega) = e^{-\alpha(\omega) - i\theta(\omega)}$ then

$$\theta(\omega) = \mathcal{H}[\alpha] \quad (57)$$

To see this, note that since f is real it satisfies the symmetry relation $\alpha(-\omega) = \alpha(\omega)$. Rewriting the integrand as

$$\frac{\omega \alpha(\omega')}{\omega^2 - \omega'^2} = \frac{1}{2} \left[\frac{\alpha(\omega')}{\omega - \omega'} + \frac{\alpha(-\omega')}{\omega - (-\omega')} \right] \quad (58)$$

and using the definition of $\mathcal{H}$ we readily obtain the foregoing result. Hence, given the magnitude data $|F(i\omega)|$, we can readily compute the minimum phase factorization of $F(i\omega)$. It is interesting to note that a reconstruction of the magnitude is possible under much less restrictive conditions. Given the phase of $F(i\omega)$ we can recover the magnitude of a minimum-phase reconstruction via equation (56), but Hayes et al. [17] have shown that the minimum-phase assumption is not necessary. In fact, knowledge of the phase uniquely determines the magnitude of a discrete, finite-length sequence under very mild restrictions.

A waveform characteristic which is important in many applications is the energy density as a function of time. The energy density of a minimum phase waveform f_m occurs earlier in time than the density of any other waveform f_0 with the same spectral magnitude. We express this property as follows:

$$\int_0^t \{[f_m(t')]^2 - [f_0(t')]^2\} dt' \geq 0 \quad (59)$$

An important result of this property is that the rise time of a minimum-phase estimate tends to be at least as short as that of the true waveform.

An implementation of a minimum-phase reconstruction is straightforward and can be made quite efficient. We note that the interpolated function $|\hat{F}(i\omega)|$ is constrained to have a zero at dc. Because minimum phase reconstruction is only applicable to functions with no poles or zeros on the $i\omega$ axis, it is appropriate to work with a time integral of $\hat{F}$. To this end we define the function $\tilde{F}$ as follows:

$$|\tilde{F}(i\omega)| = \frac{|\hat{F}(i\omega)|}{\omega} \quad (60)$$

The log amplitude is readily obtained from $|\tilde{F}(\omega)|$

$$\alpha(\omega) = -\ln |\tilde{F}(\omega)| \quad (61)$$

We can then construct an analytic function $\zeta(\omega)$ as indicated in equation (51), viz:

$$\begin{aligned} \zeta(\omega) &= -\alpha(\omega) - i\theta(\omega) \\ &= -\alpha(\omega) - i\mathcal{H}[\alpha] \\ &= \ln |F| + i\mathcal{H}[\ln |F|] \end{aligned} \quad (62)$$

This construction is most easily performed via the Fourier transform:

$$\zeta(\omega) = \mathcal{F}^{-1} \{ 2U(\omega') \mathcal{F} [\ln |F|] (\omega') \} \quad (63)$$

(It is straightforward to show that the same function ζ is obtained if the roles of $\mathcal{F}$ and $\mathcal{F}^{-1}$ are interchanged. That formulation is used in many standard treatments of minimum-phase reconstruction.) The complex estimate $\hat{F}$ is then obtained as follows

$$\hat{F}(i\omega) = i\omega e^{\zeta(\omega)} \quad (64)$$

and the estimate $\hat{f}$ follows via an inverse Fourier transform. It is relevant to note that a discrete-time implementation of these equations results in an evaluation of the waveform's cepstrum [15, ch. 10], but this relationship will not be exploited in our formulation. Other aspects of the Hilbert transform and some implementation-related findings have recently been presented by Tesche [18].

5.2 Other Phase-Reconstruction Techniques

Researchers in areas which include astronomy, electron microscopy, and crystallography have developed techniques for phase retrieval which do not require the minimum-phase assumption. A number of these techniques have been reviewed and compared by Fienup [19]. Several of these approaches can be viewed as variations on the Gerchberg-Saxton algorithm [20], an iterative procedure in which (1) an estimate of the spectral-domain signal is obtained, (2) constraints are applied in the spectral domain, (3) the constrained estimate is inverse Fourier transformed and constraints in the original domain are applied, and (4) the resulting constrained signal is Fourier transformed to obtain a new spectral-domain estimate. The process begins with an arbitrary phase estimate and continues until the constraints in both domains are satisfied to some prescribed degree of accuracy.

The use of such algorithms in reconstructing the phases of wavefronts which produce an optical image exploits the complementary constraints of (1) the optical intensity measured in a diffraction (spectral) plane and (2) an optical intensity measurement in the object (image) domain or an estimate of the dimensions of the object (obtained from the autocorrelation function). In the reconstruction of wideband signals of unknown duration, it is difficult to identify analogous constraints for these complementary domains. The issue of causality is also difficult to

address for these nonminimum-phase algorithms.⁷ In general, the reconstruction of a nonminimum-phase function remains an unsolved research problem.

⁷The minimum-phase estimate is causal by definition.

6 Numerical Examples

In this section we present numerical results obtained from the algorithms discussed above. Two wideband pulse waveforms are considered. The first, a function which both satisfies the minimum-phase criterion and is not time limited, permits us to test the interpolation procedures without regard for the validity of our phase reconstruction method. The second waveform, a single cycle of a sinusoid, is nonminimum phase and permits an assessment of both the interpolation and phase reconstruction processes. These waveforms, which exhibit a considerable range of time- and frequency-domain behaviors, permit us to identify rough bounds on the performance of the algorithms.

6.1 A Minimum-Phase Function

A minimum-phase function which satisfies the constraints defined in Section 3 is as follows:

$$f_1(t) = \frac{a(b-c)e^{-at} + b(c-a)e^{-bt} + c(a-b)e^{-ct}}{(a-b)(b-c)(c-a)} U(t) \quad (65)$$

where a , b , and c are real constants. It is easily shown that f_1 has no dc component and is continuous at $t = 0$. The Laplace transform of this function is

$$F_1(s) = \frac{s}{(s+a)(s+b)(s+c)} \quad (66)$$

We see that the time integral of this waveform, given by the inverse transform of $F_1(s)/s$, is minimum phase whenever a , b , and c are strictly positive. The time- and frequency-domain behaviors of this function are shown in Figures 17 and 18, where we have used the parameters $a = 1$, $b = 1.25$ and $c = 3$. In the numerical results presented here and throughout this section the units of time and frequency are arbitrary but inversely related.

The data used in our demonstration of the waveform reconstruction algorithms will be the magnitude of $F_1(i\omega_n)$ at four frequencies ω_n . These data correspond to points near the maximum of $|F_1(i\omega)|$ and are presented in Table 2.

Figures 19 and 20 display the results of reconstructing f_1 via rational function and spline interpolation of the spectral magnitude data. The model orders for the

Figure 17: The function $f_1(t)$.

Figure 18: The magnitude of the function $F_1(i\omega)$.

Table 2: Data used to reconstruct the minimum-phase function.

$\omega_n/2\pi$	$ F_1(i\omega_n) $
0.05	0.0771
0.22	0.131
0.29	0.113
0.40	0.0855

Figure 19: Reconstruction of $f_1(t)$ based on rational-function interpolation of F_1 .

rational-function interpolation were taken to be $(m, n) = (1, 3)$, and the bandwidth ω_{N+1} for the spline interpolator was set at $2\omega_4$ (twice the highest frequency in the data). We observe that both algorithms permit one to identify features such as rise time and pulse length, but the rational-function interpolation produces much more accurate results. Given that F_1 is a rational function of s , this result is not surprising.

Figure 20: Reconstruction (dashed curve) of $f_1(t)$ based on cubic-spline interpolation of F_1 .

6.2 A Monocycle Sinusoid

A single cycle of a sinusoid is a waveform often used in remote sensing and communications systems. We express this time-limited waveform as follows:

$$f_2(t) = \sin(2\pi f_0 t) [U(t) - U(t - \tau)] \quad (67)$$

where τ is the period of the sinusoid and $f_0 = 1/\tau$ is the frequency of the sinusoid. The Fourier transform of the monocycle is given by

$$F_2(i\omega) = \frac{i\tau}{\pi} e^{-i\omega\tau/2} \frac{\sin(\omega\tau/2)}{1 - \omega^2\tau^2/(2\pi)^2} \quad (68)$$

with the limiting value

$$\lim_{\omega\tau \rightarrow 2\pi} F_2(i\omega) = -\frac{i\tau}{2} \quad (69)$$

Because this function has an infinite sequence of zeroes on the $i\omega$ axis, it fails to satisfy the minimum-phase criterion. An example of such a pulse for $\tau = 1$ is shown in Figures 21 and 22.

Figure 21: A single-cycle sinusoid.

Figure 22: The magnitude of the Fourier transform of a single-cycle sinusoid.

Table 3: Data used to reconstruct the monocycle.

$\omega_n/2\pi$	$ F_2(i\omega_n) $
0.5	0.424
1.0	0.50
1.5	0.255
2.0	0.0
2.5	0.0606
3.0	0.0
3.5	0.0283

The data used in reconstructing this waveform are presented in Table 3. We have sampled the waveform's spectrum at half-integer multiples of the dominant frequency. This sampling produces zeroes in the frequency response that accentuate the limitations of minimum-phase reconstruction.

Figures 23 and 24 display the results of reconstructing f_2 via rational-function and cubic-spline interpolation. In these results we have taken the rational-function interpolator model orders to be $(m, n) = (2, 5)$, and for the spline interpolator we have taken $\omega_{N+1} = 8\pi$. We find that the cubic-spline interpolated result is far superior to the rational-function interpolated result in this case; the rational-function interpolation fails completely. This finding may be attributed to the ability of the cubic-spline interpolation to better match the zeroes of the data.

Figure 23: Reconstruction (dashed curve) of $f_2(t)$ based on rational-function interpolation of F_2 .

Figure 24: Reconstruction (dashed curve) of $f_2(t)$ based on cubic-spline interpolation of F_2 .

7 Concluding Remarks

We have reviewed several methods for the reconstruction of a pulsed time-domain waveform from discrete samples of the waveform's spectrum. The problem has been decomposed into two fundamental tasks—spectrum reconstruction and phase retrieval.

Spectrum reconstruction from a sparse data set is equivalent to constrained interpolation, and several approaches to this problem have been discussed. We have found the techniques of rational-function interpolation and (cubic) B-spline interpolation to be useful. These methods allow many of the relevant frequency-domain constraints to be imposed explicitly, although the problem of non-negative spectra can still arise.

Retrieval of a waveform's phase from its magnitude data is an unsolved research question at this time. Our approach to this problem employs minimum-phase reconstruction. Although this method has been largely successful in our example calculation (even when the true waveform was nonminimum phase), the minimum-phase assumption imposes an artificial constraint on the problem. Several other methods of phase retrieval were considered, but none were found to be useful when only the magnitude data were available. If, however, the duration of the waveform were also known, then other previously developed phase reconstruction methods are applicable. Further investigation of wideband pulse reconstruction should include more detailed studies of phase retrieval techniques.

We have found that minimum-phase reconstruction of the spectra obtained by interpolating the spectral magnitude data with either rational functions or cubic B-splines can produce good results. Our results suggest that rational function interpolation is better suited for functions which are dominated by exponential behavior, while cubic-spline interpolation provides better results when the spectra contain zeroes. These two interpolation methods are complementary to a degree, and it is appropriate to examine the outputs produced by both methods when one is confronted with data sets of unknown origin.

We have also examined the question of optimum sampling for wideband waveforms. It was shown that by using a nonlinear transformation of the frequency variable one can capture in a sparse data set the important characteristics of the waveform. A transformation which was found to be both practical and theoretically meaningful was derived from a normalized integral of the expected,

power-law remapped spectrum. We demonstrated that by using this approach a particular wideband pulse could be faithfully reconstructed from a limited number of complex samples. The reduction in sample rate achieved by this technique was 16:1 with respect to a naive application of the sampling theorem.

An aspect of this work which bears further investigation is the problem of assessing the quality of a waveform reconstruction. Without such an estimate it is impossible to determine one's confidence in the validity of a given reconstruction. Techniques for model confidence assessment exist and have been useful in other fields. A relevant example arises in least-squares modeling where it is frequently assumed that errors in the input parameters are Gaussian distributed. This situation leads to χ^2 statistics for the model parameters. Although the interpolation procedures described above are equivalent to least-squares solutions of matrix equations, an error analysis of the entire reconstruction process must also account for the phase retrieval algorithms. Numerical simulation techniques will probably be necessary to obtain quantitative error estimates. The large number of variables which arise in the reconstruction process (i.e., waveform shape, sampling procedure, and measurement error distribution) suggests that meaningful confidence assessments will require extensive simulations.

Appendix: Program Listings

The two spectrum estimation and signal reconstruction methods described in the text have been implemented in FORTRAN programs. These programs, CAUCHY and BSPLINE, are listed in this Appendix.

```
=====
c
c      program cauchy
c
=====
```

```
c
c      Function:
```

```
c      Reconstructs a function from discrete samples of the magnitude of its
c      frequency response using Cauchy's method for interpolation of the sampled
c      data.
```

```
c
c      Algorithm:
```

```
c      Employs Cauchy's algorithm for interpolation via rational functions and
c      minimum phase reconstruction of the complex frequency response.
```

```
c      Ref: Papoulis, A. 'The Fourier Integral and Its Applications'
c           pp. 206-208.
```

```
c
c      Data Description:
```

```
c      Input file: Ascii file of two columns:
```

```
c           First column is frequency f (arbitrary units).
```

```
c           Second column is abs() of response at frequency f.
```

```
c           Length of data sequence is limited by
c           program variable MAXSAM
```

```
c      Output file: Ascii file of two columns:
```

```
c           First column is time (units are inverse of units of freq.).
```

```
c           Second column is estimated time-domain response.
```

```
c           Number of data points is limited by program
c           variable NPTSFFT.
```

```
c
c      Required External Subroutines (From 'Numerical Recipes')
```

```
c      SVDCMP      (Singular value decomposition of real matrix)
```

```
c      SVBKSB      (backsubstitution of SVD to solve linear systems)
```

```
c      REALFT      (FFT of real data)
```

```
c      FOUR1       (FFT of complex data)
```

```
c
c      Revision History
c
=====
```

```
c      30 July 1992 Creation                      B.A. Baertlein and K.F. Casey
```

```
c      09 March 1993 Correct bugs in model order selection and
c                      scale frequency variables.
```

```
c                      B.A.B
c
=====
```

```
c
c      implicit none
c
```

```
c      integer
```

```
c      +      maxsam,
c      +      unitin,
c      +      unitout,
c      +      unitdata,
c      +      nptsfft
```

```
c
c      real
```

```
c      +      svlimit,
c      +      twopi
```

```
c
c      parameter (
```

```
c      +      maxsam = 64,
c      +      unitin = 5,
c      +      unitout = 6,
c      +      unitdata = 1,
c      +      svlimit = 1.0e-5,
```

```
+      twopi = 6.2831853,
+      nptsfft = 256
+    )
c
c *** variables
c
      integer
+      J,
+      i,
+      rank,
+      k,
+      P,
+      Q,
+      npts
c
      real
+      freq(maxsam),
+      data(maxsam),
+      AB1(maxsam,maxsam),
+      W(maxsam),
+      V(maxsam,maxsam),
+      ampl(nptsfft),
+      coefs(maxsam),
+      omega,
+      fmax,
+      ratint,
+      fscale,
+      rxform(4*nptsfft)
c
      complex
+      xform(2*nptsfft)
c
      character
+      fname*80
c
      equivalence (xform,rxform)
c
c *** Begin executable code
c
10  continue
   write (unitout,*) 'Enter file name for input data'
   read (unitin,*) fname
c
   open ( unit=unitdata, file=fname, access='sequential',
+   form='formatted', status = 'old', err=9 )
   goto 20
9   continue
   write (unitout,*) 'Error opening file ',fname
   write (unitout,*) 'Please try again'
   goto 10
20  continue
c
c *** Read the data
c
   J = 1
30  continue
   read (unitdata,*,end=99) freq(J), data(J)
   J = J + 1
   if ( J .gt. MAXSAM ) then
       write (unitout,*) 'Too many input data points'
       write (unitout,*) 'Increase MAXSAM'
       stop
   endif
```

```

    goto 30
99  continue
    close(unitdata)
    J = J - 1
    write (unitout,*) 'Number of data points = ', J
    if ( J .lt. 4 ) then
        write (unitout,*) 'Too few points to proceed'
        stop
    endif
C
C *** Get initial estimates of input data order
C
40  continue
    write (unitout,*) 'Orders for numerator (m) and denominator (n)?'
    write (unitout,*) 'Sum must be less than or equal to ',J,
+    ' and must have n>=m+2'
    read (unitin,*) P, Q
    if ( P+Q .gt. J .or. P+2 .gt. Q ) then
        write (unitout,*) 'Invalid selection'
        write (unitout,*) 'Please try again'
        goto 40
    endif
C
C *** Get maximum frequency for use in scaling frequency data
C
    fscale = 0.
    do i = 1, J
        fscale = max(fscale,freq(i))
    enddo
C
C *** Fill the coefficient matrix
C
    do i = 1, J
        do k = 1, P
            AB1(i,k) = (twopi*freq(i)/fscale)**(2*k)
        enddo
        do k = 1, Q
            AB1(i,k+P) = -(data(i)**2)*(twopi*freq(i)/fscale)**(2*k)
        enddo
    enddo
C
C *** Find the singular values
C
    call svdcmp ( AB1, J, P+Q, maxsam, maxsam, W, V )
    write (unitout,*) 'singular values'
    do i = 1, P+Q
        write (unitout,*) ' ', W(i)
    enddo
C
C *** Determine the matrix rank by working though the singular values
C
    omega=0
    do i = 1, P+Q
        omega = max(omega,abs(W(i)))
    enddo
    rank = 0
    do i = 1, P+Q
        if ( abs(W(i)) / omega .gt. svlimit ) rank = rank+1
    enddo
C
    write (unitout,*) 'Rank of matrix = ', rank
    if ( rank .ne. P+Q ) then
        write (unitout,*) '*** Warning: model order is too large'
```

```

endif
c
c *** Solve for the coefficients
c
do i = 1, J
    ampl(i) = data(i)**2
enddo
call svbksb (AB1,W,V,J,P+Q,maxsam,maxsam,ampl,coefs)
write (unitout,*) 'rational approx coefs for normalized freqs'
write (unitout,*) 'order      numerator      denominator'
write (unitout,5) 0, 0., 1.
do i=1, max(P,Q)
    if ( i .le. P .and. i .le. Q ) then
        write (unitout,5) i, coefs(i), coefs(i+P)
    else if ( i .le. P ) then
        write (unitout,5) i, coefs(i), 0.
    else if ( i .le. Q ) then
        write (unitout,5) i, 0., coefs(i+P)
    endif
enddo
5 format ('omega**',i2,2x,e10.4,7x,e10.4)
c
c *** Decide how to sample the output spectrum.
c *** Reconstruction accuracy is limited by both aliasing of the
c *** output function and aliasing of the log-amplitude spectrum.
c *** As an estimate, divide the response of the integrated
c *** waveform at DC out of the response of the integrated
c *** waveform at high frequency.
c *** Set this ratio to some small number (10^{-10} here)
c *** for good sampling of the log-amplitude response.
c *** This is a crude approximation at best.
c
npts=nptsfft
fmax = (1.e-10*coefs(1)*coefs(P+Q)/coefs(P))**(0.5/(P-Q-1))/twopi
write (unitout,*) 'Recommended fmax = ', fmax*fscale
write (unitout,*)
+ '(Recommendation is unreliable if model order is too large)'
write (unitout,*) 'Enter fmax'
read (unitin,*) fmax
c
c *** Form a list of the log-amplitude spectrum for the integral of the
c *** function. Note that the log-amplitude is an even function.
c
do i = 2, npts+1
    omega = twopi*fmax*(i-1)/(npts)
    xform(i) = 0.5*
+ log(ratint(omega/fscale,coefs,P,Q,unitout)/omega**2)
enddo
c
c *** DC value is the first numerator coefficient.
c *** Combine with the real part of the nyquist part for FFT.
c
xform(1) = cmplx(0.5*log(abs(coefs(1))/fscale**2),
+ real(xform(npts+1)))
c
c *** compute the phase via the Hilbert transform.
c *** See Papoulis for details.
c
c *** In this transform, input is real and symmetric => ditto for output.
c
call realft(rxform,npts,-1)
c
c *** Zero the negative frequencies, double the positive frequencies

```

```

c *** (except DC and Nyquist).
c
  rxform(1) = rxform(1)/npts
  rxform(npts+1) = rxform(npts+1)/npts
  do i = 2, npts
    rxform(i) = rxform(i)*2/(npts)
    rxform(2*npts-i+2) = 0
  enddo

c
c *** Copy the double length sequence back into the complex work vector
c *** and FFT. Note that xform and rxform are equivalenced!
  do i = 2*npts, 1, -1
    xform(i) = rxform(i)
  enddo
  call four1(rxform,2*npts,1)

c
c *** At this point the variable 'xform' contains the analytic signal
c *** corresponding to the log of the time integral of the input.
c *** Time convention is d/dt <==> -i omega.
c *** Size of this vector is 2*npts complex values.
c *** The imaginary part of this signal is the Hilbert transform.
c *** Form the complex response, differentiate and inverse transform.
c
  do i = 1, npts+1
    omega = twopi*fmax*(i-1)/(npts)
    xform(i) = cmplx(0.,-omega) * exp(xform(i))
  enddo

c
c *** Move real part of Nyquist into imag part of DC.
c
  xform(1) = cmplx(real(xform(1)),real(xform(npts+1)))

c
c *** Inverse transform.
  call realft(rxform,npts,-1)

c
c *** write out the time-domain response,
c *** and don't forget the scaling for the FFT (df)=(fmax/N)
c
  write (unitout,*) 'Enter file name for output data'
  read (unitin,*) fname

c
  open ( unit=unitdata, file=fname, access='sequential',
+   form='formatted', status = 'new' )

c
  do i = 1, 2*npts
    write(unitdata,*) (i-1)/(2*fmax),
+   rxform(i)*2.*fmax/npts
  enddo

c
  close (unit=unitdata)

c
  stop
end

=====
c
  function ratint(omega,coefs,P,Q,unitout)
c
=====
c
  Evaluate the rational function approximation to the data
  at a given radian frequency (omega).
c

```



```
c      Coefficients are in coefs():
c      coefs(1:P) are numerator coefficients
c      (omega^0 coefficient vanishes)
c      coefs(P+1:P+Q) are denominator coefficients with
c      omega^0 coefficient (unity) suppressed.
c=====
c
c      implicit none
c
c *** global variables
c
c      integer
c      +      P,
c      +      Q,
c      +      unitout
c
c      real
c      +      ratint,
c      +      omega,
c      +      coefs(P+Q)
c
c *** local variables
c
c      integer
c      +      i
c
c      real
c      +      num,
c      +      den,
c      +      temp
c
c *** Begin executable code.
c
c      temp = omega*omega
c
c      num = 0
c      do i = 1, P
c          num = num + (temp**i)*coefs(i)
c      enddo
c
c      den = 1
c      do i = 1, Q
c          den = den + (temp**i)*coefs(i+P)
c      enddo
c
c      ratint = num/den
c      if ( ratint .lt. 0 ) then
c          write (unitout,*) 'spectrum is negative', ratint
c          ratint = - ratint
c      endif
c
c      return
c      end
```

```
=====
c
c      program bspline
c
c=====
c
c      Function:
c      Reconstructs a function from discrete samples of the magnitude of its
c      frequency response using B-spline interpolation of the sampled data.
c
c      Algorithm:
c      Employs B splines to interpolate the sample points of the magnitude
c      of the response and
c      minimum phase reconstruction of the complex frequency response.
c      Ref: Papoulis, A. 'The Fourier Integral and Its Applications'
c          pp. 206-208.
c
c      Data Description:
c      Input file: Ascii file of two columns:
c                  First column is frequency f (arbitrary units).
c                  Second column is abs() of spectral response at frequency f.
c                  Length of data sequence is limited by
c                  program variable MAXSAM
c
c      Output file: Ascii file of two columns:
c                  First column is time (units are inverse of units of freq.).
c                  Second column is estimated time-domain response.
c                  Number of data points is limited by program
c                  variable NPTSFFT.
c
c      Required External Subroutines (From 'Numerical Recipes')
c          SPLINE      (Evaluation of cubic B-spline coefficients.
c                      Note that this routine accepts at most 100 points.
c                      This limitation is tied to the value of MAXSAM.)
c          SPLINT      (Cubic B-spline interpolation)
c          REALFT      (FFT of real data)
c          FOUR1       (FFT of complex data)
c
c      Revision History
c      -----
c      25 Aug 1992   Creation                               B.A. Baertlein
c=====
c
c      implicit none
c
c      integer
c      +      maxsam,
c      +      unitin,
c      +      unitout,
c      +      unitdata,
c      +      nptsfft
c
c      real
c      +      twopi,
c      +      naturl,
c      +      logzero
c
c      parameter (
c      +      maxsam = 100,
c      +      unitin = 5,
c      +      unitout = 6,
c      +      unitdata = 1,
```

```
+      twopi = 6.2831853,
+      naturl = 1.0e35,
+      logzero = -20.0,
+      nptsfft = 256
+      )
C
C *** variables
C
      integer
+      J,
+      i,
+      npts
C
      real
+      freq(maxsam),
+      data(maxsam),
+      secdrv(maxsam),
+      omega,
+      f,
+      fmax,
+      intrpval,
+      rxform(4*nptsfft)
C
      complex
+      xform(2*nptsfft)
C
      character
+      fname*80
C
      equivalence (xform,rxform)
C
C *** Begin executable code
C
10  continue
   write (unitout,*) 'Enter file name for input data'
   read (unitin,*) fname
C
   open ( unit=unitdata, file=fname, access='sequential',
+   form='formatted', status = 'old', err=9 )
   goto 20
9   continue
   write (unitout,*) 'Error opening file ',fname
   write (unitout,*) 'Please try again'
   goto 10
20  continue
C
C *** Read the data, prepend the correct DC response
C
   freq(1) = 0.
   data(1) = 0.
   J = 2
30  continue
   read (unitdata,*,end=99) freq(J), data(J)
   J = J + 1
   if ( J.gt. MAXSAM-1 ) then
       write (unitout,*) 'Too many input data points'
       write (unitout,*) 'Increase MAXSAM'
       stop
   endif
   goto 30
99  continue
   close(unitdata)
   J = J - 1
```

```
      write (unitout,*) 'Number of data points = ', J-1
c
c *** Decide how to sample the output spectrum.
c *** Reconstruction accuracy is limited by both aliasing of the
c *** output function and aliasing of the log-amplitude spectrum.
c *** As an estimate, double the highest frequency present in the data.
c
      npts=nptsfft
      fmax = 2 * freq(J)
      write (unitout,*) 'Recommended fmax = ', fmax
      write (unitout,*) 'Enter fmax'
      read (unitin,*) fmax
c
c *** Append a zero to the data at this frequency.
c
      J = J +1
      freq(J) = fmax
      data(J) = 0.
c
c *** Form the interpolation coefficients, forcing the "natural"
c *** boundary conditions at DC and zero derivative at the highest frequency.
c
      call spline ( freq, data, J, naturl, 0., secdrv )
c
c *** Form a list of the log-amplitude spectrum for the integral of the
c *** function. Note that the log-amplitude is an even function.
c
      do i = 2, npts+1
        f = fmax*(i-1)/(npts)
        call splint(freq,data,secdrv,J,f,interpval)
c *** roundoff errors can lead to small negative function values
        if ( interpval .le. 0 ) then
          xform(i) = logzero
        else
          xform(i) = log(interpval/(twopi*f))
        endif
      enddo
c
c *** DC value is a simple linear estimate for the "natural"
c *** spline boundary condition at zero.
c *** Note:  $\lim_{\omega \rightarrow 0} F(\omega)/\omega = F'(\omega) = \text{data}(2)/\text{freq}(2)$ 
c *** Combine with the real part of the nyquist part for FFT.
c
      xform(1) = cmplx(log(data(2)/(freq(2))),
+      real(xform(npts+1)))
c
c *** Compute the phase via the Hilbert transform.
c *** See Papoulis for details.
c
c *** in this transform, input is real and symmetric => ditto for output
c
      call realft(rxform,npts,-1)
c
c *** Zero the negative frequencies, double the positive frequencies
c *** (except DC and Nyquist).
c
      rxform(1) = rxform(1)/npts
      rxform(npts+1) = rxform(npts+1)/npts
      do i = 2, npts
        rxform(i) = rxform(i)*2/(npts)
        rxform(2*npts-i+2) = 0
      enddo
c
```

```
c *** Copy the double length sequence back into the complex work vector
c *** and FFT. Note that xform and rxform are equivalenced!
c
  do i = 2*npts, 1, -1
    xform(i) = rxform(i)
  enddo
  call fourl(rxform,2*npts,1)
c
c *** At this point the variable 'xform' contains the analytic signal
c *** corresponding to the log of the time integral of the input.
c *** Time convention is d/dt <==> -i omega.
c *** Size of this vector is 2*npts complex values.
c *** The imaginary part of this signal is the Hilbert transform.
c *** Form the complex response, differentiate and inverse transform.
c
  do i = 1, npts+1
    omega = twopi*fmax*(i-1)/(npts)
    xform(i) = cmplx(0.,-omega) * exp(xform(i))
  enddo
c
c *** Move real part of Nyquist into imag part of DC.
c
  xform(1) = cmplx(real(xform(1)),real(xform(npts+1)))
c
c *** Inverse transform.
c
  call realft(rxform,npts,-1)
c
c *** Write out the time-domain response,
c *** and don't forget the scaling for the FFT (df)=(fmax/N)
c
  write (unitout,*) 'Enter file name for output data'
  read (unitin,*) fname
c
  open ( unit=unitdata, file=fname, access='sequential',
+   form='formatted', status = 'new' )
c
  do i = 1, 2*npts
    write(unitdata,*) (i-1)/(2*fmax), rxform(i)*2.0*fmax/npts
  enddo
c
  close (unit=unitdata)
c
  stop
end
```

References

- [1] Casey, K. F., "The external environment of VPD-II: space-wave field," *Sensor and Simulation Notes*, Note 323, April 1990.
- [2] Baum, C. E., "Resistively loaded radiating dipole based on a transmission-line model for the antenna," *Sensor and Simulation Notes*, Note 81, April 1969.
- [3] Kehrer, W. S. and C. E. Baum, "Electromagnetic design parameters for ATHAMAS II (VPD-II), *ATHAMAS Memos*, Memo 4, May 1975.
- [4] Mendel, J. M., *Maximum-Likelihood Deconvolution*, Springer-Verlag, 1990.
- [5] Jansson, P. A., *Deconvolution with Applications to Spectroscopy*, Academic, 1984.
- [6] Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, UK, 1986.
- [7] Stoer, J. and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, 1980.
- [8] Cauchy, A. L., "Sur la formule de Lagrange relative a l'interpolation", *Analyse Algebrique*, Paris, 1821.
- [9] Kottapalli, K., T. K. Sarkar, Y. Hua, E. K. Miller, and G. J. Burke, "Accurate computation of wide-band response of electromagnetic systems utilizing narrow-band information", *IEEE Trans. Microwave Theory Tech.*, MTT-39(4), pp. 682-686, April 1991.
- [10] Unser, M. and A. Aldroubi, "Polynomial splines and wavelets—a signal processing perspective", in C. K. Chui (ed.), *Wavelets—A Tutorial in Theory and Applications*, pp. 91-122, Academic Press, 1992.
- [11] Papoulis, A., *Signal Analysis*, McGraw-Hill, New York, 1977.
- [12] Soumekh, M., "Band-Limited interpolation from unevenly sampled data", *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-36(1), pp. 110-122, January 1988.

- [13] Levi, L., "Fitting a bandlimited signal to given points", *IEEE Trans. Information Theory*, pp. 372-376, July 1965.
- [14] Goodman, J. W., *Statistical Optics*, §7.4.4, Wiley, 1985.
- [15] Oppenheim, A. V. and R. W. Schaffer, *Digital Signal Processing*, Prentice-Hall, ch. 7, 1975.
- [16] Papoulis, A., *The Fourier Integral and Its Applications*, McGraw-Hill, New York, 1962.
- [17] Hayes, M. H., J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude", *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-28(6), pp. 672-680, December 1980.
- [18] Tesche, F. M., "On the use of the Hilbert transform for processing measured cw data", *IEEE Trans. Electromag. Compat.*, 34(3), pp. 259-266, 1992.
- [19] Fienup, J. R., "Phase retrieval algorithms: a comparison", *Appl. Opt.*, 21(15), pp. 2758-2769, 1 August 1982.
- [20] Gerchberg, R. W. and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures", *OPTIK*, 35(2), pp. 237-246, 1972.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.