Interpolation and Extrapolation Using a High-Resolution Discrete Fourier Transform

Mauricio D. Sacchi, Tadeusz J. Ulrych, and Colin J. Walker

***Abstract*—We present an iterative nonparametric approach to spectral estimation that is particularly suitable for estimation of line spectra. This approach minimizes a cost function derived from Bayes' theorem. The method is suitable for line spectra since a "long tailed" distribution is used to model the prior distribution of spectral amplitudes. An important aspect of this method is that since the data themselves are used as constraints, phase information can also be recovered and used to extend the data outside the original window.**

The objective function is formulated in terms of hyperparameters that control the degree of fit and spectral resolution. Noise rejection can also be achieved by truncating the number of iterations. Spectral resolution and extrapolation length are controlled by a single parameter. When this parameter is large compared with the spectral powers, the algorithm leads to zero extrapolation of the data, and the estimated Fourier transform yields the periodogram.

When the data are sampled at a constant rate, the algorithm uses one Levinson recursion per iteration. For irregular sampling (unevenly sampled and/or gapped data), the algorithm uses one Cholesky decomposition per iteration.

The performance of the algorithm is illustrated with three different problems that frequently arise in geophysical data processing:

1) **harmonic retrieval from a time series contaminated with noise;**
2) **linear event detection from a finite aperture array of receivers [which, in fact, is an extension of 1)],**
3) **interpolation/extrapolation of gapped data.**

The performance of the algorithm as a spectral estimator is tested with the Kay and Marple data set. It is shown that the achieved resolution is comparable with parametric methods but with more accurate representation of the relative power in the spectral lines.

***Index Terms*—Bayes procedures, discrete Fourier transforms, interpolation, inverse problems, iterative methods, signal restoration, signal sampling/reconstruction, spectral analysis.**

I. Introduction

SPECTRAL analysis is a very active field of research. Possible applications are extremely diverse and have been particularly well detailed in an excellent review by Kay and Marple [4]. Basically, spectral analysis is an underdetermined, linear inverse problem. The goal is to determine a spectral estimate from the infinite number of estimates that are consistent with the data that represent, in general, a few samples of the autocorrelation function. Most of the techniques that have been proposed depend on the use of some particular norm that imposes a particular feature on the spectral estimate. In a similar fashion, the computation of the Fourier transform in a manner consistent with a given set of discrete observations can also be regarded as a linear inverse problem [9].

We show that Fourier transform estimation by inversion with prior information is quite similar to a bandlimited signal extrapolation problem. Different algorithms that achieve this objective have been devised [1], [10]. The underlying basis of these algorithms can be summarized as the minimization of a frequency weighted norm subject to data constraints. The weights are chosen to incorporate some *a priori* knowledge of the bandwidth and shape of the spectrum of the signal target. A byproduct of the algorithm is the Fourier transform of the extended time series and, therefore, a high-resolution spectral estimate as well.

The technique is also used to reconstruct a signal from a set of nonuniform samples. The algorithm computes the discrete Fourier transform (DFT) from the nonuniform sampled signal, and finally, an inverse DFT is used to obtain an evenly sampled signal. An algorithm for recovering signals from nonuniform samples has been proposed by Marvasti *et al.* [7], [8]. This algorithm recovers a bandlimited signal by iteratively applying bandlimiting and nonuniform sampling operators. The bandlimiting operator is used to constrain the signal to have a preassigned spectral support. Our technique is quite different in that we model the signal by assigning a prior distribution to the spectral samples of the DFT.

The paper is organized as follows. In Section II, we present an inverse procedure to compute the DFT. The inverse problem is regularized using two different criteria. First, we propose a zero-order quadratic regularization (damped least-squares) [5] that is derived from Bayes' theorem, assuming the DFT and the observational errors can be modeled with a Gaussian distribution. This regularization leads to the well-known formulas for the DFT and its inverse, which is equivalent to assuming that the original times series has been padded with zeros. Oldenburg [9] reached the same conclusion by minimizing the first Dirichlet criterion of the Backus and Gilbert formalism. In our second approach, we use the Cauchy distribution to model the samples of the DFT. This regularization leads to an iterative nonparametric procedure to estimate the DFT.

Manuscript received April 7, 1995; revised February 18, 1997. This work was supported by an NSERC grant to T. J. Ulrych. The associate editor coordinating the review of this paper and approving it for publication was Dr. Farokh Marvasti.

M. D. Sacchi was with the Department of Earth and Ocean Sciences, The University of British Columbia, Vancouver, B.C., Canada V6T 1Z4. He is now with the Department of Physics, Abadh Bhatia Physics Laboratory, University of Alberta, Edmonton, Alta., Canada T6G 2J1.

T. J. Ulrych is with the Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, B.C., Canada V6T 1Z4.

C. J. Walker is a private consultant in Vancouver, B.C., Canada V6K 1T1.

Publisher Item Identifier S 1053-587X(98)00529-7.

In Section III, we describe the algorithm that is utilized to retrieve the DFT. In Section IV, we discuss the estimation of the hyperparameters that control the resolution of the spectral estimate. Finally, Section V is devoted to numerical examples.

II. Theoretical Considerations

A. Estimation of the Fourier Transform by Linear Inversion

Consider an N-sample time series $x_0, x_1, x_2, \cdots, x_{N-1}$. The DFT of the discrete series is given by

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi nk/N} \qquad k = 0, \cdots, N-1 \tag{1}$$

and similarly, the inverse DFT is given by

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{i2\pi nk/N} \qquad n = 0, \cdots, N-1. \tag{2}$$

Let us suppose that we wish to estimate M spectral samples where $M > N$. A standard approach to solving this problem is zero padding. Defining a new time series consisting of the original series plus a zero extension for $n = N, \cdots, M-1$, we can estimate M spectral samples using the DFT. This procedure helps to remove ambiguities due to discretization of the Fourier transform, but as is well known, it does not reduce the sidelobes created by the temporal window. Let us therefore consider the estimation of M spectral samples but using (1) without zero padding. In other words, we want to estimate the DFT using only the available information. Moreover, in order to avoid biasing our results by the discretization, we also impose the condition $M > N$. Rewriting (2) as

$$x_n = \frac{1}{M} \sum_{k=0}^{M-1} X_k e^{i2\pi nk/M} \qquad n = 0, \cdots, N-1 \tag{3}$$

gives rise to a linear system of equations

$$\mathbf{x} = \mathbf{F}\mathbf{X} \tag{4}$$

where the vectors $\mathbf{x} \in \mathbf{R}^N$ and $\mathbf{X} \in \mathbf{C}^M$ denote the available information and the unknown DFT, respectively. The $N \times M$ matrix $\mathbf{F}$ contains the exponential terms $F_{n,k} = (1/M)\, e^{i2\pi nk/N}$. Equation (4) is a linear underdetermined problem that, as is well known, can be satisfied by many different solutions. Uniqueness is imposed by defining a regularized solution [12] $\hat{\mathbf{X}}$, which is obtained by minimizing the expression

$$J(\mathbf{X}) = \Phi(\mathbf{X}) + ||\mathbf{x} - \mathbf{F}\mathbf{X}||_2^2 \tag{5}$$

where $||.||_2^2$ stands for the l_2 norm. We will alternatively indicate the l_2 norm as $||\mathbf{X}||_2^2 = \mathbf{X}^H\mathbf{X} = \sum_k X_k^* X_k$, where H denotes Hermitian transpose. The regularizer $\Phi(\mathbf{X})$ serves to impose a particular feature on the solution.

B. Bayesian Approach to Regularization—The Gaussian–Gaussian (GG) Model

Throughout this paper, we will consider data contaminated by noise that is distributed as $N(0, \sigma_n^2)$. The conditional distribution of the data is given by

$$p(\mathbf{x}|\mathbf{X}, \sigma_n) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{(N-1)/2} e^{-(1/2\sigma_n^2)||\mathbf{x}-\mathbf{F}\mathbf{X}||_2^2}. \tag{6}$$

Consider a prior distribution for $p(\mathbf{X}|\sigma_X)$ conditional on a parameter σ_X. According to Bayes' rule (see, for example, [6]), the *a posteriori* distribution of the vector of parameters is given by

$$p(\mathbf{X}|\mathbf{x}, \sigma_X, \sigma_n) = \frac{p(\mathbf{X}|\sigma_\mathbf{X})\, p(\mathbf{x}|\mathbf{X}, \sigma_n)}{p(\mathbf{x}|\mathbf{X}, \sigma_X, \sigma_n)}. \tag{7}$$

The maximum *a posteriori* (MAP) estimator $\hat{\mathbf{X}}$ maximizes $p(\mathbf{X}|\mathbf{x}, \sigma_X, \sigma_n)$ for given σ_X and σ_n. Let us assume that the prior distribution of X_k is Gaussian and, since X_k is a complex variable, we will also assume that the real and imaginary parts are independent variables with Gaussian distributions. Then

$$\begin{aligned} p[\Re(\mathbf{X}), \Im(\mathbf{X})|\sigma_X] &= \left(\frac{1}{2\pi\sigma_X^2}\right)^{M-1} e^{-(1/2\sigma_X^2)[||\Re(\mathbf{X})||_2^2 + ||\Im(\mathbf{X})||_2^2]} \\ &= \left(\frac{1}{2\pi\sigma_X^2}\right)^{M-1} \\ &\quad \cdot \exp\left\{-\left(\frac{1}{2\sigma_X^2}\right) \sum_{k=0}^{M-1} [\Re(X_k)^2 + \Im(X_k)^2]\right\}. \end{aligned} \tag{8}$$

The last equation describes the joint distribution of $2(M-1)$ random variables that represent the joint distribution of $M-1$ complex variables [3]. Since $||\Re(\mathbf{X})||_2^2 + ||\Im(\mathbf{X})||_2^2 = \mathbf{X}^H\mathbf{X}$, (8) may be regarded as the joint distribution of the complex variable $\mathbf{X}$ conditional on σ_X, which we designate by $p(\mathbf{X}|\sigma_X)$. The MAP solution that maximizes the *posteriori* probability also minimizes the function

$$\begin{aligned} J_{gg}(\mathbf{X}) &= \frac{1}{2\sigma_X^2}||\mathbf{X}||_2^2 + \frac{1}{2\sigma_n^2}||\mathbf{x} - \mathbf{F}\mathbf{X}||_2^2 \\ &= \frac{1}{2\sigma_X^2}\,\mathbf{X}^H\mathbf{X} + \frac{1}{2\sigma_n^2}\,(\mathbf{x} - \mathbf{F}\mathbf{X})^H(\mathbf{x} - \mathbf{F}\mathbf{X}) \end{aligned} \tag{9}$$

where the subscript gg stands for the GG model. Taking derivatives and equating to zero gives the following estimator (see the Appendix):

$$\hat{\mathbf{X}} = (\mathbf{F}^H\mathbf{F} + \lambda\mathbf{I}_M)^{-1}\mathbf{F}^H\mathbf{x}. \tag{10}$$

The scalar $\lambda = \sigma_n^2/\sigma_X^2$ is also known as the tradeoff or ridge regression parameter. We can write (10) in another form using the identity

$$(\mathbf{F}^H\mathbf{F} + \lambda\mathbf{I}_M)^{-1}\mathbf{F}^H = \mathbf{F}^H(\mathbf{F}\mathbf{F}^H + \lambda\mathbf{I}_N)^{-1} \tag{11}$$

where $\mathbf{I}_M$ and $\mathbf{I}_N$ denote $M \times M$ and $N \times N$ identity matrices, respectively. Recalling that $\mathbf{F}\mathbf{F}^H = (1/M)\mathbf{I}_N$, we end up with the Fourier transform estimate for the GG model

$$\hat{\mathbf{X}} = \left(\frac{1}{M} + \lambda\right)^{-1}\mathbf{F}^H\mathbf{x}. \tag{12}$$

The result is nothing more than the DFT of x_n modified by a scale factor. When the additive noise is absent, $\lambda = 0$, and the solution expressed by (12) becomes

$$\hat{X}_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi nk/(M-1)} \tag{13}$$

which is the DFT of the windowed time series (1) and is equivalent to padding the data with zeros in the range $n = N, \cdots, M-1$.

It is clear that the GG model yields the DFT. Hence, the associated periodogram

$$P_k = \hat{X}_k \hat{X}_k^*, \qquad k = 0, \cdots, M-1 \tag{14}$$

will exhibit a resolution that is proportional to the length of the time series.

C. Regularization by the Cauchy–Gaussian (CG) Model

In problems that involve the estimation of line spectra, we have found that a sparse distribution of spectral amplitudes provides not only a manner to regularize the inversion of the DFT but also enhances the spectral resolution of the associated periodogram. We propose the following distribution to model the samples of the DFT:

$$p(X_k|\sigma_X) \propto \frac{1}{\left(1 + \dfrac{X_k X_k^*}{2\sigma_X^2}\right)} \tag{15}$$

which is a Cauchy distribution of the complex variables X_k [3]. The multidimensional distribution is given by $p(\mathbf{X}|\sigma_X) = \prod_k p(X_k|\sigma_X)$, and we have used this distribution as a prior for the complex vector $\mathbf{X}$. Combining the Cauchy prior with the data likelihood, we find the cost function for the CG model

$$J_{cg}(\mathbf{X}) = S(\mathbf{X}) + \frac{1}{2\sigma_n^2}(\mathbf{x} - \mathbf{F}\mathbf{X})^H(\mathbf{x} - \mathbf{F}\mathbf{X}) \tag{16}$$

where $S(\mathbf{X})$ is the regularizer imposed by the Cauchy distribution

$$S(\mathbf{X}) = \sum_{k=0}^{M-1} \ln\left(1 + \frac{X_k X_k^*}{2\sigma_X^2}\right) \tag{17}$$

and is a measure of the sparseness of the vector of powers $P_k = X_k X_k^*,\ k = 0, \cdots, M-1$. The constant σ_X controls the amount of sparseness that can be attained by the inversion, which will also depend on the noise level since σ_n may inhibit a reliable sparse solution. This issue will be examined in detail in the section devoted to the estimation of hyperparameters.

Taking derivatives of $J_{cg}(\mathbf{X})$ and equating to zero yields

$$\mathbf{X} = (\lambda \mathbf{Q}^{-1} + \mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{x} \tag{18}$$

where $\lambda = \sigma_n^2/\sigma_X^2$, and $\mathbf{Q}$ is a $M \times M$ diagonal matrix with elements

$$Q_{ii} = 1 + \frac{X_i X_i^*}{2\sigma_X^2}, \qquad i = 0, \cdots, M-1. \tag{19}$$

Equation (18) resembles the damped least squares solution, but the elements of the diagonal matrix $\mathbf{Q}$ are nonlinearly related to the Fourier coefficients X_k. Thus, an iterative procedure is needed to compute the estimator $\hat{\mathbf{X}}$ that solves (18). Expression (18) can be rewritten using the identity

$$\mathbf{F}^H(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H) = (\lambda \mathbf{Q}^{-1} + \mathbf{F}^H\mathbf{F})\mathbf{Q}\mathbf{F}^H. \tag{20}$$

The forms $(\lambda \mathbf{Q}^{-1} + \lambda \mathbf{F}^H\mathbf{F})$ and $(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)$ are positive definite leading to the identities

$$(\lambda \mathbf{Q}^{-1} + \mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H = \mathbf{Q}\mathbf{F}^H(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)^{-1}. \tag{21}$$

Using the identity expressed by (21), we can write (18) as

$$\mathbf{X} = \mathbf{Q}\mathbf{F}^H(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)^{-1}\mathbf{x}. \tag{22}$$

Since $\mathbf{Q}$ depends on $\mathbf{X}$, the last equation has to be solved by means of an iterative procedure. This is outlined in Section III.

The following observations can be made.

1) Equation (18) demands the inversion of a $N \times N$ matrix, whereas in (22), we perform the inversion of a $M \times M$ matrix.
2) The operator $(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)$ in (22) is Hermitian Toeplitz, provided that the time series is discretized at a constant rate. Therefore, a fast solver like Levinson's recursion can be used in the inversion. In the case of nonuniform discretization, a Cholesky decomposition is appropriate.

As we mentioned, the CG model leads to an algorithm that resembles the minimum norm solution of (16). This is particularly true when σ_X is large compared with the spectral amplitudes we are seeking. In this case, the functional $S(\mathbf{X}) \approx K + \mathbf{X}^H\mathbf{X}/2\sigma_X^2$, where K is a constant. Thus, minimizing $J_{cg}(\mathbf{X})$ is equivalent to minimizing $J_{gg}(\mathbf{X})$. In the contrary case, the algorithm will seek a DFT with a sparse distribution of spectral amplitudes, reducing windowing effects and enhancing the spectral peaks.

III. Iterative Nonlinear Estimation of the CG Model

We now describe the procedure that enables us to minimize the cost function $J_{cg}(\mathbf{X})$. We iteratively solve (22), which involves the solution of a $N \times N$ system of equations. It is straightforward to see the computational advantages of using (22) rather than (18). First, we will rewrite (22) as

$$\mathbf{X} = \mathbf{Q}\mathbf{F}^H(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)^{-1}\mathbf{x} = \mathbf{Q}\mathbf{F}^H\mathbf{b} \tag{23}$$

where the auxiliary vector $\mathbf{b} \in \mathbf{R}^N$ is obtained from the solution of the system

$$(\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}\mathbf{F}^H)\mathbf{b} = \mathbf{x}. \tag{24}$$

Equations (23) and (24) suggest an iterative scheme. The algorithm starts with the DFT of the finite length data set $\mathbf{X}^{(0)}$. The initial solution is also used to generate the matrix $\mathbf{Q}^{(0)}$. In each iteration, we compute

$$\mathbf{b}^{(\mu-1)} = [\lambda \mathbf{I}_N + \mathbf{F}\mathbf{Q}^{(\mu-1)}\mathbf{F}^H]^{-1}\mathbf{x} \tag{25}$$

which we subsequently use to update the DFT as

$$\mathbf{X}^{(\mu)} = \mathbf{Q}^{(\mu-1)}\mathbf{F}^H\mathbf{b}^{(\mu-1)} \tag{26}$$

where μ denotes iteration number. The procedure is stopped when the following criterion is satisfied:

$$\frac{|J_{cg}^{(\mu)} - J_{cg}^{(\mu-1)}|}{[|J_{cg}^{(\mu)}| + |J_{cg}^{(\mu-1)}|]/2} < \text{tolerance.} \tag{27}$$

The estimator $\hat{\mathbf{X}}$ is computed at the point where the last condition is satisfied. In general, a few iterations ($\leq$10) are needed to minimize the cost function J_{cg}.

IV. Estimation of the Hyperparameters

A. Data in the Absence of Noise

For the purpose of this discussion, we will concentrate on the CG model. In the case of accurate data, we set the parameter $\sigma_n = 0$. This is equivalent to seeking a perfect fit between the observed and the predicted data within the original window. Equations (23) and (24) reduce to the formulation we would have obtained by minimizing

$$J_{cg}(\mathbf{X}) = S(\mathbf{X}) + \mathbf{b}^H(\mathbf{x} - \mathbf{F}\mathbf{X}) \tag{28}$$

where $\mathbf{b}$ is the vector of Lagrange multipliers. Minimizing (28) leads to (22) with $\sigma_n^2 = 0$. In this case, several models will satisfy the data. The resolution of the model is controlled by σ_X, which is a parameter that invokes sparseness in the spectral estimate and therefore enhances resolution. The length of the extended time series is also controlled by σ_X.

When $\sigma_X \geq P$, where P is the maximum power encountered in the initial model, the algorithm leads to the periodogram. An explanation of this feature is that when σ_X is large compared with the spectral powers of the periodogram, the functional $S(\mathbf{X})$ reduces to $\sum X_k X_k^*/2\sigma_X^2$, and we are, therefore, estimating the minimum norm DFT of the windowed time series. Thus, the periodogram will coincide with the periodogram of the windowed time series. When σ_X is small compared with the spectral powers, the algorithm will enhance the spectral peaks.

B. Estimation of σ_X When σ_n Is Known

For harmonic signals with additive Gaussian noise of a given power σ_n, the parameter σ_X can be easily estimated according to a chi-squared criterion. In many geophysical applications, we can assess the noise level. For expensive computational procedures, i.e., 2-D problems, the degree of fitting can also be controlled by limiting the number of iterations.

Since we have assumed that the noise is normally distributed, the misfit $\Phi = (1/\sigma_n^2)\|\mathbf{x} - \mathbf{F}\mathbf{X}\|_2^2$ reduces to a chi-squared statistic. Therefore, $E[\Phi] = N$ [11], where the largest acceptable value at a 99% confidence limit is $\approx N + 3.3\sqrt{N}$.

V. Numerical Simulations

A. Harmonic Retrieval

The simulated example we illustrate consists of a 50-point time series composed of a pair of unit amplitude harmonics of 0.2 and 0.21 Hz contaminated with Gaussian noise with standard deviation $\sigma = 0.1$. The frequency spacing is 50% below the classic resolution limit given by the inverse of the time series length. The periodogram of the time series is shown in Fig. 1. The length of the time series is not sufficient to resolve both harmonics. The periodogram is dominated by sidelobes that mask the signals. In Fig. 2, we show the power spectrum using the CG model after nine iterations. The parameter σ_n was assumed to be known, and σ_X was obtained using the chi-squared criterion. Fig. 3 portrays the power spectrum versus the iteration number. The initial spectrum, with $\mu = 0$, corresponds to the periodogram of the finite length time series. In each iteration, sidelobes are progressively diminished, and the spectral peaks are enhanced until the harmonics are precisely recovered. The extended time series at each iteration are shown in Fig. 4. The first time series corresponds to the original data, and it is easy to see that the resolution is enhanced as a consequence of the extension we obtain.

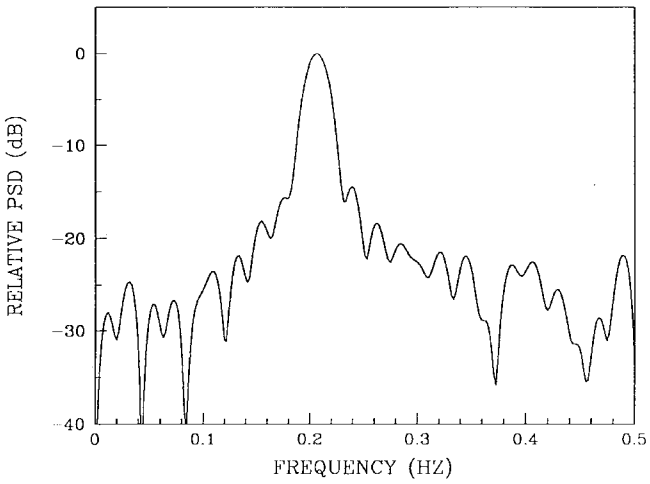

Fig. 1. Periodogram of a synthetic process that consists of a pair of unit amplitude harmonics of 0.2 and 0.21 Hz immersed in additive white noise ($\sigma_n = 0.2$). The sidelobes inhibit the correct identification of the harmonics.

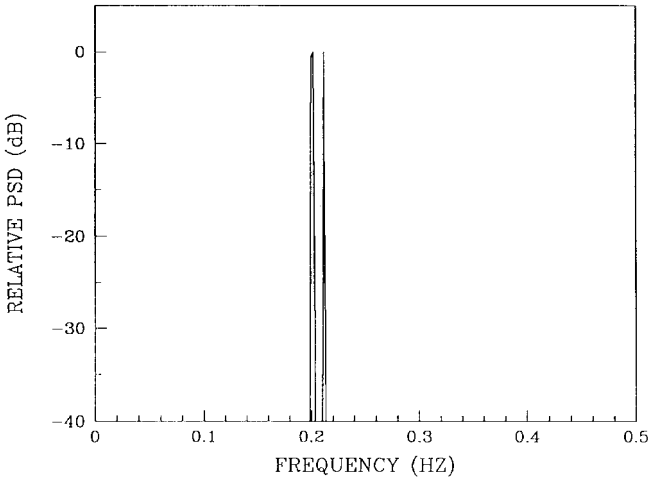

Fig. 2. Power spectrum computed with the Cauchy–Gauss model after nine iterations. The data correspond to the harmonic process used in Fig. 1. The sidelobes are suppressed, and the resolution is enhanced.

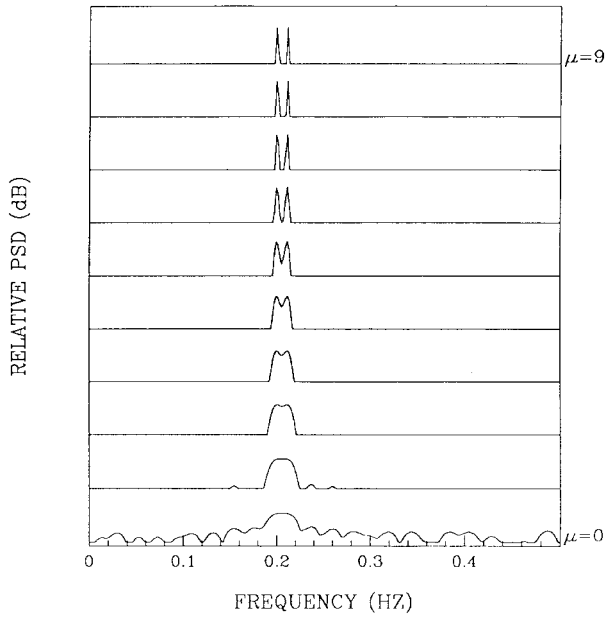

Fig. 3. Power spectra versus iteration number. The iteration $\mu = 0$ corresponds to the periodogram of the finite length time series, which is also plotted in Fig. 1. The final spectrum ($\mu = 9$) is also plotted in Fig. 2.

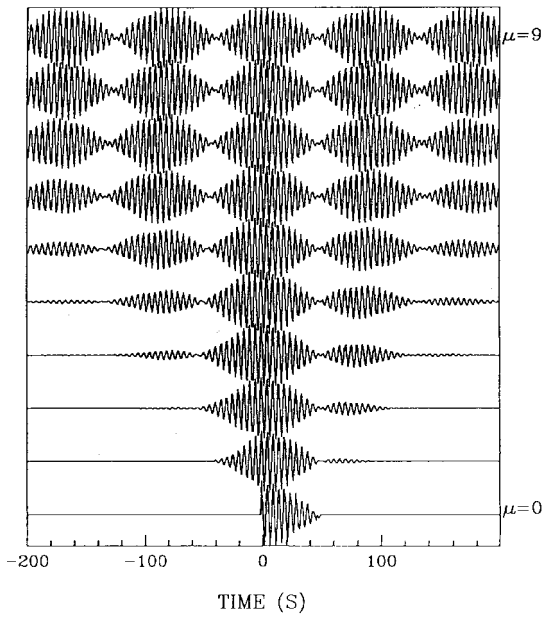

Fig. 4. Time series reconstructed from the DFT versus number of iterations. The first time series ($\mu = 0$) corresponds to the original data. The associated spectra are shown in Fig. 3.

B. The Kay and Marple (KM) Example [4]

We illustrate the performance of our algorithm by application to the well known KM data set. The results of 11 spectral estimators that have been applied to this data set by Kay and Marple are shown in [4, Fig. 16]. These data constitute an important benchmark in testing the performance of different spectral estimators. The data consist of 64 samples that describe a pair of unit amplitude harmonics with frequencies of 0.2 and 0.21 Hz, a third harmonic with an amplitude of 0.1 (20 dB down) at 0.1 Hz, and a continuous noise process

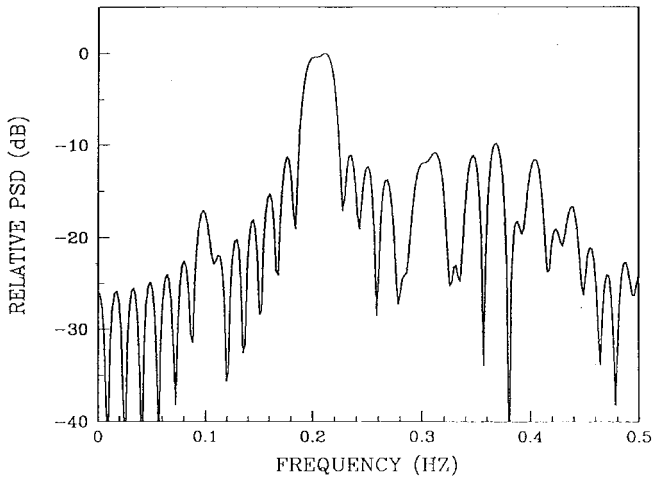

Fig. 5. Periodogram of the Kay and Marple [4] data set.

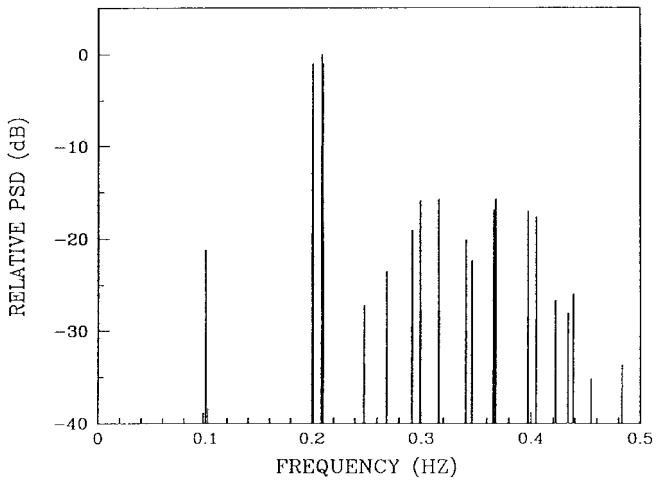

Fig. 6. High-resolution spectral estimate of the Kay and Marple [4] data set computed with the Cauchy–Gauss model.

in the upper band centered at about 0.35 Hz, where it rises to about 15 dB below the pair of unit amplitude signals. Fig. 5 illustrates the usual periodogram of the data, and Fig. 6 illustrates the PSD estimated with our algorithm using 2048 frequency samples. The parameter σ_n was set to zero since we are fitting the data exactly.

It is easy to see that the method enables us to perfectly resolve the harmonic components, but the broad part of the spectrum is incorrectly replaced by a number of distinct spectral lines (which do, however, outline the shape of the continuous noise spectrum). This problem is also encountered in the high-resolution parametric spectral estimators discussed by Kay and Marple [4]. This is not to say that our approach is doomed to failure, the phenomenon being simply a consequence of using a norm that mimics a sparse distribution of spectral amplitudes.

Of the estimators tested in [4], the Hildebrand–Prony method, which is a variant of Prony's approach for real undamped sinusoids in noise [2], provides the most accurate representation of the frequencies and power of the three harmonics. Like our approach, Prony's method does not require autocorrelation estimates. It does, however, require

the selection of the correct order of the input process, which can be half the data length, and may result in poor estimates if the order is estimated incorrectly. For the harmonic part of the process, our approach yields a resolution accuracy comparable with Prony's approach. If σ_n is increased, the broad part of the spectrum can be partly rejected. Unfortunately, this can lead to misleading results since small amplitudes might be rejected.

C. Spatio-Temporal Spectrum of Signals Received by a Passive Array

We now apply the algorithm to estimate the spatio-temporal spectrum of a signal received by a passive array of receivers. This problem frequently arises in radar, sonar, and seismic processing. The goal is to estimate the direction of arrival and the temporal spectral signature of a set of sources impinging from different angles on a uniform array of N receivers. In seismology, the problem has been intensively studied to detect plane-wave signals and estimate the slowness vector. In this paper, we present a narrowband example, but the problem can be easily extended to broadband signals. We are also assuming that the array of receivers is linear, but the method can be easily generalized to any real distribution of receivers.

Instead of developing a two-dimensional (2-D) version of our method, we prefer to present a hybrid procedure based on standard Fourier analysis in the temporal variable, whereas for the spatial variable, we invert the wavenumber using the CG regularization. Usually, the length of the temporal window is sufficient to achieve high resolution with simple methods based on standard Fourier analysis, whereas it is the aperture of the array that limits the spatial resolution. The 2-D algorithm works as follows.

1) Each record is transformed to the frequency offset domain using the FFT.
2) High-resolution analysis is performed at each frequency that comprises the signal.
3) The amplitude in the f–k space is plotted to identify the spatio-temporal structure of each source.
4) Alternatively, the data outside the original aperture may be extrapolated to simulate a longer array, and any 2-D spectral technique may be used in conjunction with the extended data set.

The method was tested with a simulated array of 10 equally spaced receivers. We modeled three sinusoids with unit amplitude and with normalized wavenumbers of 0.20, 0.25, and -0.25 units and normalized frequencies of 0.20, 0.20, and 0.35 units, respectively. The temporal extension of each channel is 100 samples. Gaussian noise with standard deviation $\sigma_n = 0.2$ was added to the composite record, and each channel was tapered with a Hamming window. The spatio-temporal spectrum computed using the periodogram is illustrated in Fig. 7. The contour lines correspond to normalized amplitudes ranging from 0 to -40 db with an interval of 5 db. The f–k plane is dominated by sidelobes due to truncation in space and time. This is more evident for the wavenumber since the aperture of the array is one order smaller that the length of the time series. The data set was then processed with the hybrid procedure based on the CG model. The result is portrayed in Fig. 8. There is a clear enhancement of the spatial resolution. In order to obtain such a resolution with a standard f–k analysis, the required aperture must be at least 10 times larger than the actual aperture.

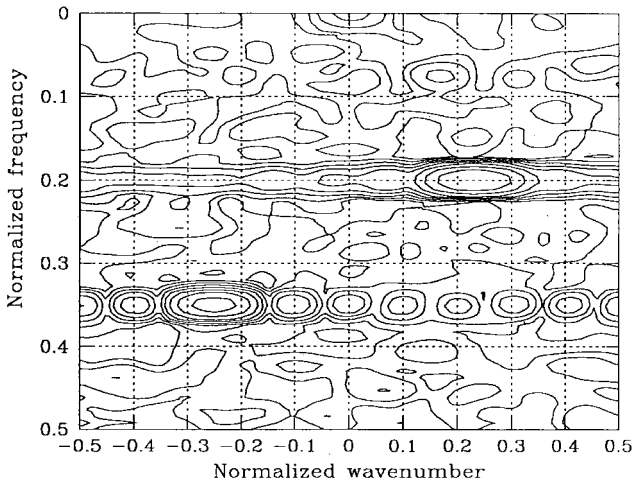

Fig. 7. Conventional f–k plane of three narrowband waves recorded by an array of ten receivers. The panel is dominated by sidelobes that mask the signals.

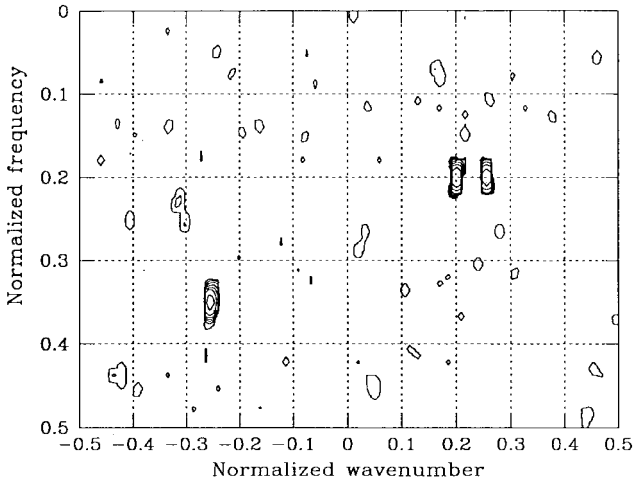

Fig. 8. High-resolution f–k panel estimated with the Cauchy–Gauss model. Sidelobes are diminished, and consequently, we can clearly recognize the spectral signature of each wave.

D. Application to Unevenly Sampled Data and to Gap Filling

Two important problems in time series analysis are spectral estimation of unevenly sampled data and gap filling. We show in this section that our algorithm may be used to cope with both problems. For irregular data, the vector of observation in (4) contains the amplitude of the time series at position t_j, $j = 0, \cdots, N-1$, and the Fourier kernel is written as

$$F_{j,k} = e^{i2\pi t_j f_k}, \qquad k = 0, \cdots, M-1. \tag{29}$$

The frequency axis is discretized at constant rate as in the previous analysis. Notice that the matrix $\mathbf{FQF}^H$ in (22) possesses Hermitian symmetry but is not Toeplitz. The resampling and/or gap filling is accomplished by the following scheme.

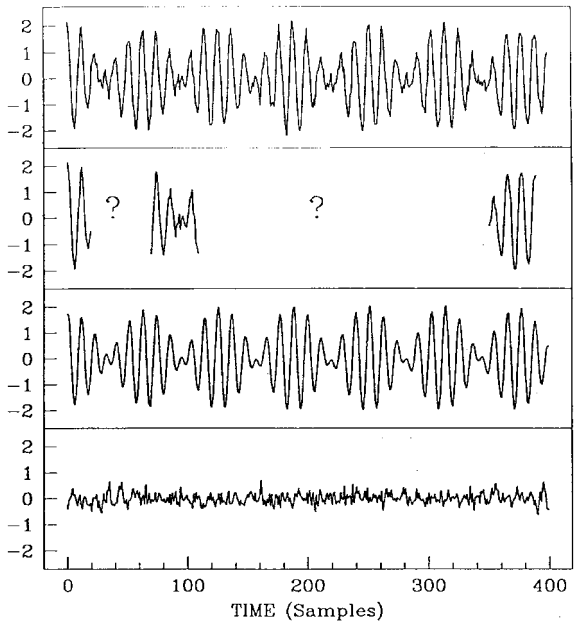

Fig. 9. (Top) Synthetic time series; the series contains two unit amplitude harmonics of frequencies 0.795 and 0.954 Hz, respectively; (second) Gapped time series; (third) Reconstructed time series using the Cauchy–Gauss DFT; (bottom) Error sequence obtained subtracting (third) from (top).

1) Select the proper Nyquist frequency, and define the number of samples of the frequency axis.
2) Form $\mathbf{d}$ and $\mathbf{F}$, allocating only the times corresponding to actual observations.
3) Iteratively retrieve the DFT using the Cauchy–Gauss regularization. At this stage, an estimate of the spectrum may also be computed.
4) Apply the inverse DFT transformation to compute a time series sampled at regular intervals and/or to fill the gaps.

The top of Fig. 9 portrays a time series that consists of 400 samples. There are two harmonics of unit amplitude, which are located at 0.795 and 0.954 Hz. The time series is contaminated with Gaussian noise with $\sigma_n = 0.2$. Three segments are extracted from the time series to produce the gapped data (see the second level of Fig. 9). The first segment has 20 samples, and the second and the third have 40 samples each. We recall that the gaps in the figure are not treated as zeros. They are simply not considered in the inversion. We use the CG DFT to estimate the power spectrum of the gapped data, which is shown in Fig. 10. The DFT is also used to reconstruct the time series. The gaps represent 75% of the total length of the reconstructed time series. The hyperparameters of the problem are chosen to yield $\chi^2 = N$. The third level of Fig. 9 shows the reconstruction after ten iterations of our algorithm. Finally, the bottom of Fig. 9 shows the difference between the original synthetic time series and the reconstruction using the CG DFT. This is also an estimate of the noise sequence used in the synthetic time series.

VI. Discussion and Concluding Remarks

The high-resolution technique for the estimation of the power spectrum presented in this paper and the prediction of the time series is based on the application of an algorithm that seeks a sparse solution to the ubiquitous problem of spectral estimation from a finite set of data. What makes the algorithm very attractive is that the sparseness measure is minimized subject to data constraints, and therefore, phase information is also recovered and is used in the extrapolation of the signal outside the original window, or aperture, depending on the problem.

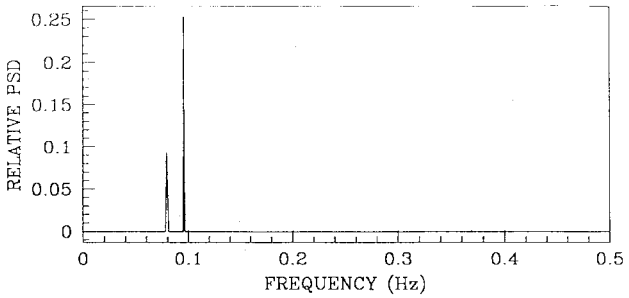

Fig. 10. Power spectrum of the gapped time series computed using the Cauchy–Gauss model.

The present algorithm is best suited to the analysis of undamped harmonic signals. It is a spectral line estimate; therefore, just as in the case of the Hildebrand–Prony estimator, the broadband process is less well modeled. However, we believe that the present approach obtains a result that achieves a good compromise between an accurate determination of the harmonic components and a measure of the power present in the stochastic component. Like the Hildebrand–Prony estimator, our algorithm does not require autocorrelation lag estimates but, unlike the Hildebrand–Prony approach, our method also does not require the *a priori* determination of the order of the process.

The hybrid technique developed to retrieve the spatio-temporal signature of spatial signals offers an efficient alternative to overcome aperture artifacts in short arrays. When the data are sampled at a constant rate, the algorithm is very fast since Levinson's recursion is applied.

Finally, we point out that the CG DFT can also be applied to retrieve harmonics from irregularly sampled time series. A byproduct of the technique is a time series resampled at constant rate that is obtained from the DFT.

Appendix
Proof of (10) and (18)

Using the rules for differentiation of a quadratic form by a complex vector, we have

$$\frac{\partial J_{gg}(\mathbf{X})}{\partial \mathbf{X}^*} = \frac{\partial}{\partial \mathbf{X}^*}\left[\frac{1}{2\sigma_X^2}\|\mathbf{X}\|_2^2 + \frac{1}{2\sigma_n^2}\|\mathbf{x}-\mathbf{F}\mathbf{X}\|_2^2\right] = \frac{1}{2\sigma_X^2}\,\mathbf{X} - \frac{1}{2\sigma_n^2}\,\mathbf{F}^H(\mathbf{x}-\mathbf{F}\mathbf{X}) \tag{A1}$$

which, after equating to zero, leads to

$$\lambda\hat{\mathbf{X}} + \mathbf{F}^H\mathbf{F}\hat{\mathbf{X}} = \mathbf{F}^H\mathbf{x} \tag{A2}$$

where $\lambda = \sigma_n^2/\sigma_X^2$. Finally, the solution is given by

$$\hat{\mathbf{X}} = (\mathbf{F}^H\mathbf{F} + \lambda\mathbf{I}_M)^{-1}\mathbf{F}^H\mathbf{x} \tag{A3}$$

which is (10).

When the Cauchy prior is adopted, we equate the derivatives of the cost function (16) to zero

$$\frac{\partial J_{cg}(\mathbf{X})}{\partial \mathbf{X}^*} = \frac{\partial}{\partial \mathbf{X}^*}\left[S(\mathbf{X}) + \frac{1}{2\sigma_n^2}\,||\mathbf{x}-\mathbf{F}\mathbf{X}||_2^2\right] = \frac{\partial S(\mathbf{X})}{\partial \mathbf{X}^*} + \frac{1}{2\sigma_n^2}\,\mathbf{F}^H(\mathbf{x}-\mathbf{F}\mathbf{X}). \tag{A4}$$

To evaluate the derivative of the regularization term $S(\mathbf{X})$, for simplicity, we compute the derivative with respect to the elements of the vector $\mathbf{X}^*$

$$\frac{\partial S(\mathbf{X})}{\partial X_j^*} = \frac{\partial}{\partial X_j^*}\sum_{k=1}^{M}\ln\left(1+\frac{X_k X_k^*}{2\sigma_X^2}\right) = \frac{1}{2\sigma_X^2}\left(1+\frac{X_k X_k^*}{2\sigma_X^2}\right)^{-1} X_k = \frac{1}{2\sigma_X^2}\,Q_{k,k}^{-1}X_k. \tag{A5}$$

Writing (A5) in matrix form, we have

$$\frac{\partial S(\mathbf{X})}{\partial \mathbf{X}^*} = \frac{1}{2\sigma_X^2}\,\mathbf{Q}^{-1}\mathbf{X}. \tag{A6}$$

Combining (A4) and (A6), we obtain, after equating (A4) to zero

$$\mathbf{X} = (\lambda\mathbf{Q}^{-1} + \mathbf{F}^H\mathbf{F})^{-1}\mathbf{F}^H\mathbf{x} \tag{A7}$$

which is equivalent to (18) and solved iteratively, as discussed in Section III.

REFERENCES

[1] S. D. Cabrera and T. W. Parks, "Extrapolation and spectral estimation with iterative weighted norm modification," *IEEE Trans. Signal Processing*, vol. 39, Apr. 1991.

[2] F. B. Hildebrand, *Introduction to Numerical Analysis.* New York: McGraw-Hill, 1956.

[3] N. L. Johnson and S. Kotz, *Distributions in Statistics: 1 and 2.* New York: Wiley, 1970.

[4] S. M. Kay and L. M. Marple, "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, Nov. 1981.

[5] L. R. Lines and S. Treitel, "Tutorial: A review of least-squares inversion and its applications to geophysical problems," *Geophysical Prospecting*, vol. 32, 1984.

[6] T. J. Loredo, "From Laplace to supernova SN 1987A: Bayesian inference in astrophysics," in *Maximum Entropy and Bayesian Methods*, P. Fougère, Ed. Dordrecht, The Netherlands: Kluwer, 1990.

[7] F. Marvasti, M. Analoui, and M. Gamshadzahi, "Recovery of signals from nonuniform samples using iterative methods," *IEEE Trans. Signal Processing*, vol. 39, Apr. 1991.

[8] M. Nafie and F. Marvasti, "Implementation of recovery of speech with missing samples on a DSP chip," *Electron. Lett.*, vol. 30, no. 1, Jan. 1994.

[9] D. W. Oldenburg, "Calculation of Fourier transforms by the Backus–Gilbert method," *Geophys. J. R. Astron. Soc.*, vol. 44, 1976.

[10] A. Papoulis and C. Chamzas, "Detection of hidden periodicities by adaptive extrapolation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, Oct. 1979.

[11] C. R. Rao, *Linear Statistical Inference and Its Applications.* New York: Wiley, 1973.

[12] A. H. Tikhonov and A. V. Goncharsky, *Ill-Posed Problems in the Natural Sciences.* Moscow, Russia: MIR, 1987.

Mauricio D. Sacchi was born in Colonel Brandsen, Argentina, in 1965. He received the diploma in geophysics from the National University of La Plata, La Plata, Argentina, and the Ph.D. degree in geophysics from The University of British Columbia (UBC), Vancouver, B.C., Canada, in 1988 and 1996, respectively.

After completing the Ph.D. degree, he joined the Consortium for the Development of Specialized Seismic Techniques, UBC, as a Post-Doctoral Fellow. He recently joined the Department of Physics, University of Alberta, Edmonton, Alta., Canada, as an Assistant Professor of Geophysics. His current interests are in the area of seismic data processing and seismic imaging and inversion.

Tadeusz J. Ulrych was born in Warsaw, Poland. He received the B.Sc. (Hons.) degree in electrical engineering from London University, London, Ont., Canada, in 1957 and the M.Sc. and Ph.D. degrees from the University of British Columbia (UBC), Vancouver, B.C., Canada, in geophysics in 1961 and 1963, respectively.

He joined the faculty at UBC in 1965, where he is presently Professor of Geophysics. He spent 1984 and 1987 to 1989 teaching and as Visiting Professor at PPPG/Universidade Federal de Bahia, Salvador, Brazil. From 1992 to 1993, he was a Visiting Scholar at JNOC, Tokyo, Japan. His research interests include information and inverse theory, time series and spectral analysis, and the processing of geophysical data. Of particular current interest is the joint application of Bayesian and maximum entropy principles to inverse problems.

Dr. Ulrych is an Associate Editor for the *Journal of Applied Geophysics* and for the *Journal of Seismic Exploration.*

Colin J. Walker was born in Fort William, Ont., Canada. He received the B.Math. degree from the University of Waterloo, Waterloo, Ont., in 1970.

He worked as a Research Scientist with the Department of Geophysics and Astronomy, University of British Columbia, Vancouver, B.C., Canada, from 1976 to 1987. His current interests are in time adaptive modeling and spectral analysis.