Estimation of Sinusoids from Incomplete Time Series

P. S. NAIDU AND BINA PARAMASIVAIAH

Abstract—**The Gerchberg-Papoulis extrapolation algorithm was applied to the problem of recovering missing observations from incomplete time series. It has been experimentally shown that with as few as 10 percent of data points and *a priori* information on the bandwidth, it is possible to recover all the missing points in the time series. The technique is illustrated by means of a pair of sinusoids sampled randomly. It has been possible to estimate the entire time series with mean-square error less than 0.0035 after 30 iterations. The parameters of the sinusoids could be estimated with high accuracy.**

I. Introduction

The problem of estimating frequency of sinusoids from a finite time series has been attacked by several authors [1], [3]-[5]. In many practical situations the time series is not complete on account of either measurements at certain sampling time instants not being made due to the malfunction of the hardware or the signal-to-noise ratio being so low that the data point had to be discarded. Such situations often arise in processing of array data. Some of the sensors do not function at all, or they are extremely noisy and hence the data from such sensors is not acceptable. Further, in some experiments it may not be possible to take measurement at regular intervals. One such situation is in measurement of motion of fine particles by laser anemometer [6]. The spectral analysis of such incomplete time series was attempted by Nuttall [7], who used maximum entropy and linear prediction techniques. He showed that a reasonable estimate can be obtained from data with up to 30 percent missing observations. The missing data may be recovered by means of one of the standard mathematical interpolation schemes, for example, Lagrange interpolation. The interpolation error can be made small by requiring the average sampling rate to be higher than the Nyquist rate [8].

In this paper we investigate the possibility of applying the Gerchberg-Papoulis algorithm, originally meant for extrapolation of band-limited signals to interpolation of the missing observations. Once the missing observations are recovered, the spectrum is easily obtained by Fourier transformation. It is shown that in a noise free environment, two closely spaced sine waves can be resolved with as few as ten percent of total data, i.e., 90 percent of the data points being missing. The missing points can be recovered with a mean-square error of less than 0.0035 after 30 iterations.

Manuscript received August 14, 1983; revised November 22, 1983. This work was supported by the Department of Electronics, Government of India.

P. S. Naidu is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India.

B. Paramasivaiah was with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India. She is now with the University of Connecticut, Storrs, CT.

II. Gerchberg-Papoulis Algorithm

Consider a function $X(t)$ with the Fourier transform $\bar{X}(w)$, such that

$$\bar{X}(w) = 0 \qquad |w| > w_0. \tag{1}$$

The observed time series is obtained by randomly sampling $X(t)$ over the entire duration of the time series. Let

$$X_0(t) = D(t)\, X(t) \qquad t = 0, 1, 2, \cdots, T-1 \tag{2}$$

where $D(t)$ is a random sequence having value 1 or 0 with a prescribed probability p or $1-p$.

$$\text{prob}\,\{D(t) = 1\} = p \qquad \text{for} \quad t = 0, 1, 2, \cdots, T-1.$$

Note that whenever $D(t) = 0$, the data point is lost. The procedure to determine $X(t)$ for $t = 0, 1, 2, \cdots, T-1$ given $X_0(t)$, $t = 0, 1, 2, \cdots, T-1$ consists of the following steps.

Step One: Compute the Fourier transform of $X_0(t)$.

Step Two: Band limit the Fourier transform $\bar{X}_0(w)$ to $\pm w_0$, i.e., let

$$\bar{P}(w) = \bar{X}(w) \qquad |w| \leqslant w_0$$
$$= 0 \qquad \text{otherwise.}$$

Step Three: Compute the inverse Fourier transform of $\bar{P}(w)$. The new sequence is combined with the observed sequence as follows:

$$X_1(t) = [1 - D(t)]\; P(t) + X_0(t) \tag{3}$$

where $P(t)$ is the inverse Fourier transform of $\bar{P}(w)$. $X_1(t)$ gives us the first estimate of the interpolated time series. To get the next improved estimate we go back to step one and go through all the three steps.

This procedure is continued until the relative mean-square difference is less than a prescribed number

$$\sum_t |X_n(t) - X_{n-1}(t)|^2 \leqslant \delta. \tag{4}$$

The numerical experiments to be described in the next section have shown that the mean square difference goes to zero exponentially (Fig. 4). Further information on the above algorithm and its comparison with other extrapolation techniques may be found in an excellent tutorial paper by Fitzgerald and Byrne [9].

III. Numerical Simulation

We have considered as an example of a bandlimited signal a sum of two sinusoids

$$X(t) = a \cos(2\pi f_1 t + \varphi_1) + b \cos(2\pi f_2 t + \varphi_2). \tag{5}$$

0096-3518/84/0600-0559$01.00 ©1984 IEEE

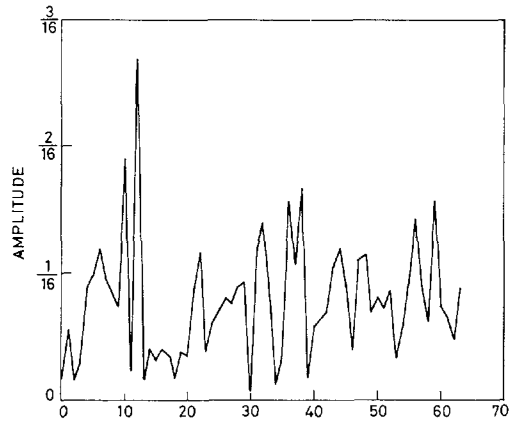


Fig. 1. Amplitude spectrum of incomplete time series obtained by randomly selecting 26 data points from a time series of 256 points.

We have used the following numerical values for a, b, f_1, f_2, φ_1, and φ_2:

$$a = 1.25, \quad b = 1.5, \quad f_1 = 10 \text{ Hz}, \quad f_2 = 12 \text{ Hz}$$

$$\varphi_1 = \pi/2, \quad \varphi_2 = \pi/3.$$

The signal was sampled at the rate of 256 samples/s. The signal duration was one second. The random sequence $D(t)$ was generated with the help of a pseudorandom number generator. For every t, a random number was generated having a uniform distribution between 0 and 1. $D(t)$ is set to 1 if the pseudorandom number is less than or equal to p, otherwise it is set to 0. The simulated observed sequence $X_0(t) \cdot t = 0, 1, 2, \cdots, T - 1$ is subjected to Fourier analysis. The amplitude spectrum is shown in Fig. 1. Note that the peaks corresponding to 10 Hz and 12 Hz stand above the erratic sidelobes. It is remarkable that the two peaks could be resolved with just 26 data points (10 percent) which if taken contiguously could not have resolved the peaks. The simulated sequence was subjected to the interpolation algorithm described in the last section. The Fourier spectrum was band limited to 4 and 24 Hz. In addition to band limiting, the amplitude spectrum was subjected to thresholding, i.e.,

$$\bar{P}(w) = 0 \quad \text{if} \quad |\bar{P}(w)| \leqslant \epsilon$$

where ϵ is a predetermined constant. This technique is effective when the spectrum is made of sharp peaks [1]. After nearly 30 iterations we obtained interpolated time series as shown in Fig. 2 and its spectrum in Fig. 3.

The parameters of the sinusoids as estimated from the reconstructed time series are shown in Table I. The frequency and phase measurements are quite accurate, however; amplitude measurement is in error by less than 8 percent. In Fig. 4 we have plotted the mean-square difference as defined in (4) as a function of number of iterations. Note that the mean-square difference seems to decay exponentially.

IV. Discussion of Results

The reconstructed waveform after 30 iterations is compared with the actual waveform in Fig. 2 where we have shown the actual data points by means of filled circles. The mean-square error is 0.00327. The mean-square error decreases as the number of iterations is increased (Fig. 5). From this it is clear that

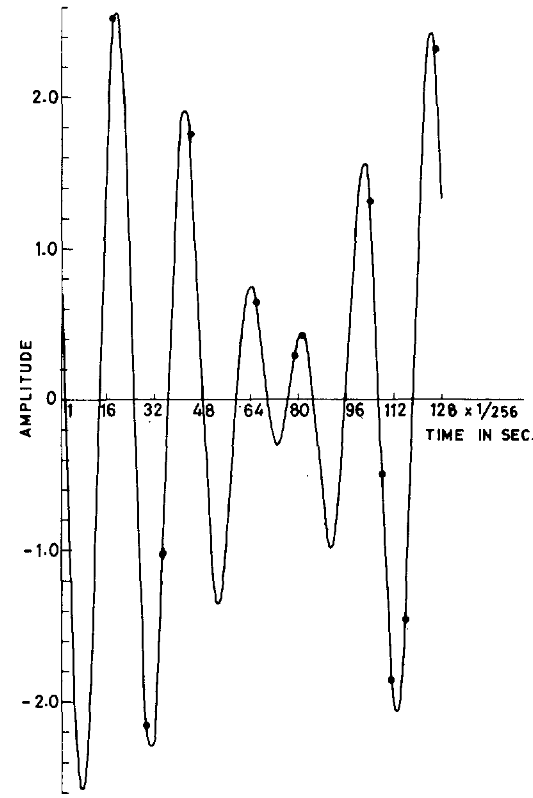


Fig. 2. Interpolated time series after 30 iterations. The actual data points are also shown in the figure by means of filled circles.

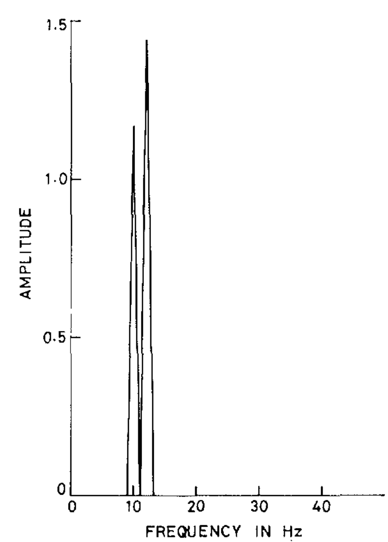


Fig. 3. Spectrum of interpolated time series shown in Fig. 2. Note that the peaks appear at exact frequencies.

TABLE I

Parameters	Actual Values	Observed Values
Frequencies	i) 10 Hz ii) 12 Hz	i) 10 Hz ii) 12 Hz
Amplitudes	i) 1.25 ii) 1.50	i) 1.17 ii) 1.44
Phases	i) 1.5714 ii) 1.0476	i) 1.5709 ii) 1.0545

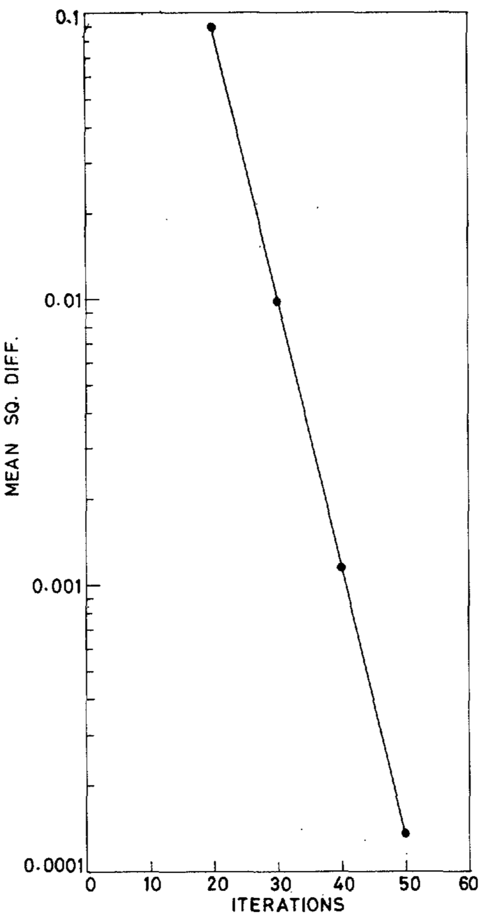


Fig. 4. Mean-square difference between the interpolated time series after every ten iterations.

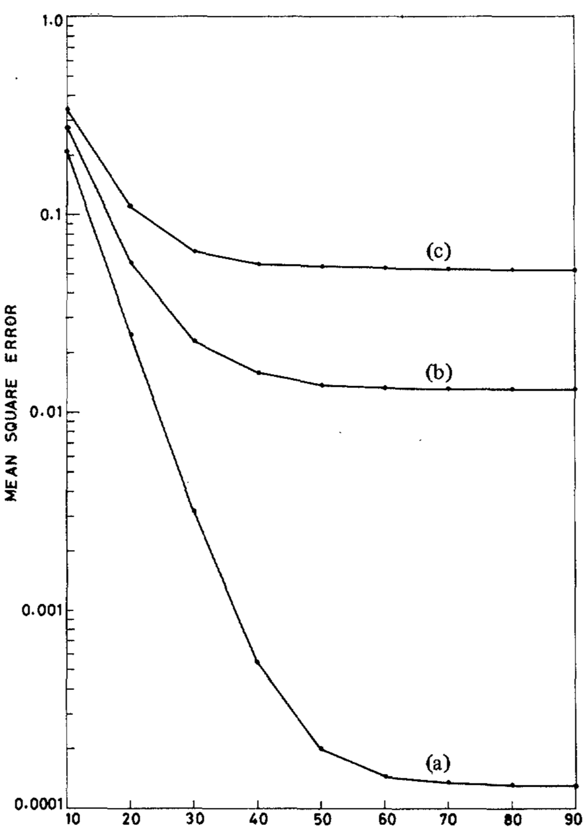


Fig. 5. Mean-square error as a function of number of iterations. The effect of background noise is also shown. (a) Zero noise. (b) SNR = 22.8 dB. (c) SNR = 16.8 dB.

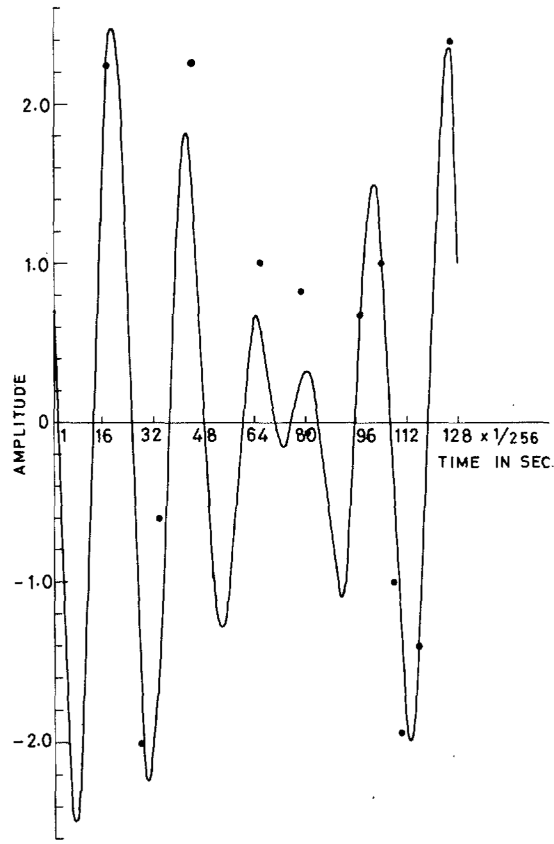


Fig. 6. Incomplete data with noise (filled circles) and reconstructed waveform after 40 iterations (continuous curve). SNR = 22.8 dB.

the missing points can be recovered from incomplete time series. In addition we can estimate the frequency, amplitude, and phase of the sinusoids. The Fourier spectrum was band limited between 4 and 24 Hz. This choice of band limiting frequencies was not found to be critical except in presence of background noise. In the presence of noise, however, tighter bounds on the cutoff frequency helped to remove the out-of-band noise and thus improve the spectrum estimation and waveform reconstruction. Any noise inside the band cannot be removed or suppressed by the present technique. We have encountered situations when the background noise peaks got enhanced along with the signal peaks. A proper selection of threshold can help to overcome such a situation. Indeed, proper selection of threshold plays an important role. As a rule of thumb the threshold was selected in such a manner that the background noise spectrum was suppressed. We have carried out a few experiments of waveform reconstruction by adding zero mean white noise of prescribed variance to the incomplete time series previously generated. It was found that the reconstruction mean-square error is largely controlled by in-band noise. The error cannot be reduced beyond a limit by increasing the number of iterations indefinitely (Fig. 5). The reconstructed waveform after 40 iterations is shown in Fig. 6.

References

[1] A. Papoulis and C. Chamzas, "Detection of hidden periodicities by adaptive extrapolation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 492–500, Oct. 1979.
[2] A. Papoulis, "A new algorithm in spectral analysis and band limited extrapolation," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 735–742, Sept. 1975.
[3] M. S. Sabri and W. Steenaart, "An approach to band limited signal

extrapolation: The extrapolation matrix," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 74–78, Feb. 1978.

[4] A. K. Jain and S. Ranganath, "Extrapolation algorithm for discrete signals with application in spectral estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 830–845, Aug. 1981.

[5] D. P. Kolba and T. W. Parks, "Extrapolation and spectral estimation for band limited signals," in *Proc. ICASSP*, 1978, pp. 372–374.

[6] M. Gaster and J. B. Roberts, "The spectral analysis of randomly sampled records by a direct transform," *Proc. Roy. Soc. London A*, vol. 354, pp. 27–58, 1977.

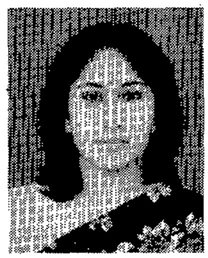
[7] A. H. Nuttall, "Spectral analysis of a univariate process with bad data points, via maximum entropy and linear predictive techniques," NUSC Tech. Rep. 5303, Naval Underwater Systems Center, New London, CT 06320, Mar. 1976.

[8] R. Radzyner and P. T. Bason, "An error bound for Lagrange interpolation of low pass function," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 669–671, Sept. 1972.

[9] R. Fitzgerald and C. Byrne, "Extrapolation of band limited signals: A tutorial," in *Proc. EUSIPCO-80*, M. Kunt and F. deCoulon, Eds., North-Holland, 1980, pp. 175–179.

P. S. Naidu was born in India in 1937. He received the B.Sc. (Honors) and M. Tech. degrees from the Indian Institute of Technology, Kharagpur, India, in 1960 and 1962, respectively, and the Ph.D. degree from the University of British Columbia, Vancouver, B.C., Canada, in 1965.

He worked for the Geological Survey of Canada and Dalhousie University, Halifax, N.S., Canada, as a Postdoctoral Fellow. In 1969 he joined the Indian Institute of Science, Bangalore, where he is now a Professor in the Department of Electrical Communication Engineering. His research interests are in geophysical signal processing, coherent optical computers, and underwater signal processing.



Bina Paramasivaiah received the Bachelor's degree in electronics engineering from Bangalore University, Bangalore, India. She is currently working towards a graduate degree at the University of Connecticut, Storrs.

Until 1983 she was a Project Assistant with the Indian Institute of Science, Bangalore.