To the Graduate Council:

I am submitting herewith a dissertation written by Mukta Lal Kar entitled "A Simplified Approach to Spectrum Estimation of Stochastic Processes." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Engineering Science.

L. J. Pinson

Lewis J. Pinson, Major Professor

E. C. Huebschmann

E. C. Huebschmann, Major Professor

We have read this dissertation
and recommend its acceptance:

Kenneth E. Harwell

R L Young

Roy D Joseph

V C Reddy

W. M. Farmer

Accepted for the Council:

L. Evans Peik

Vice Chancellor
Graduate Studies and Research

A SIMPLIFIED APPROACH TO SPECTRUM ESTIMATION OF STOCHASTIC PROCESSES

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Mukta Lal Kar

June 1981

3051904

DEDICATION

This work is dedicated to my mother and father.

ACKNOWLEDGMENTS

The work reported herein was supported in part by the U. S. Army Research Office (Research Triangle Park, North Carolina), Grant DAAG29-77-G-0138.

The author wishes to express his appreciation to many people at The University of Tennessee Space Institute (UTSI) who aided in the completion of this work. His thanks go to the chairmen of the committee, Dr. L. J. Pinson, Associate Professor of Electrical Engineering, and Dr. E. C. Huebschmann, Professor of Electrical Engineering, for their guidance and suggestions at different stages of the work.

He also thanks the other committee members, Dr. R. D. Joseph, Associate Professor of Electrical Engineering; Dr. K. C. Reddy, Professor of Mathematics; and Dr. R. L. Young, Professor of Mechanical Engineering, for their constructive suggestions and valuable discussions related to this work. In particular, he wishes to thank his supervisor, Dr. K. E. Harwell, Professor and Director of the Gas Diagnostic Division, for his support and continuous encouragement throughout this study. Special thanks are also due to Dr. W. M. Farmer, Associate Professor and Manager of the Gas Diagnostic Laboratory.

He also wishes to thank Mr. J. O. Hornkohl, Senior Research Engineer at the Gas Diagnostic Laboratory,

for valuable assistance during the progress of this work. Thanks are also due to the other personnel of the Gas Diagnostic Laboratory for their direct or indirect help.

The author wishes to thank Mrs. Helen Mason, Mrs. Mary Lo, and Mrs. Marjorie Joseph for their help in acquiring the necessary information for the literature survey.

He remembers the cheerful company of Dr. and Mrs. K. Sarkar during their three years in Tullahoma and thanks them for sharing some of their time.

The author is deeply indebted to his wife, father-in-law, mother-in-law, and brothers for their patience, understanding and encouragement. He wishes to express love to his son, Partha, who missed so many glorious moments while his father was so involved with this investigation.

Sincere thanks are also due to Messrs. James F. Taylor, Larry Janzen, and Kurt Ingenthron of the Intergraph Corporation, Huntsville, Alabama, for their interest in this work.

Finally, the author wishes to express sincere appreciation to Mrs. Jeannie Harper for her devotion and excellence in preparing the manuscript.

ABSTRACT

Application of the popular Fast Fourier Transform (FFT) algorithm is limited to equi-spaced data. A literature survey of spectrum analyses using randomly sampled records shows that most are not suitable for real time applications and require considerable high speed storage. For these reasons they are not suitable for small system applications.

A simple approach to spectrum estimation for randomly sampled data was investigated. This approach is based on the Fourier series and the method of modified least squares fit to estimate the Fourier coefficients. The estimator is algebraically very simple and produces an alias-free spectrum if a sufficient number of data points are available. This is true even if the mean sampling rate is less than the Nyquist rate required for spectral computation when using equi-spaced data samples.

Detailed statistical and error analyses are presented for the estimator. A statistical analysis shows the estimator to be asymptotically unbiased and consistent. Also, a number of simulation experiments were conducted using both computer generated data and the data acquired in practical laboratory setups. For some cases the same data were input to other algorithms for comparison. The

experimental results verify the analytic solutions in the majority of the cases.

Finally, stepwise procedures to use this algorithm are stated and the implementation in small computer systems is discussed.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
II.	LITERATURE SURVEY	4
	Discrete Signals and Laser Doppler Velocimetry	4
	Random Sampling and Laser Doppler Velocimeter Measurements	7
	Spectral Estimators for Randomly Sampled Records	10
	Discretized Lag Product Method (Mayo's Approach)	19
	Other Techniques	23
III.	STATEMENT OF THE PROBLEM	25
IV.	DEVELOPMENT OF THE ALGORITHM	27
	Statistical Analysis of the Estimator	34
	Error Analysis of the Estimator	40
	Window Selection for the Present Estimator	46
V.	EXPERIMENTS WITH THE UNIVERSITY OF TENNESSEE SPACE INSTITUTE ALGORITHM AND THEIR RESULTS	48
	Study with Simulated Band-Limited White Noise	49
	Noise Rejection Study	53

CHAPTER PAGE

Aliasing Study with The University of Tennessee Space Institute Algorithm. 56

Tests with Data Obtained from Laboratory Setup 62

VI. CONCLUSIONS AND RECOMMENDATIONS. 69

BIBLIOGRAPHY 73

APPENDICES 78

A. LEAST SQUARE ERROR FIT METHOD. 79

B. SOME USEFUL TIME STATISTICS AND THEIR ESTIMATORS 82

C. BIAS ERROR, LEAKAGE, AND ALIASING. 88

D. THE TECHNIQUE OF BRICK-WALL FILTERING. 96

VITA 100

LIST OF FIGURES

FIGURE		PAGE
1.	Pictorial Representation of Major Sampling Schemes	5
2.	Block Schematic of the Band Limited White Noise Experiment through Simulation	50
3.	Low Pass Filter Transfer Function Reconstruction Using The University of Tennessee Space Institute and Mayo Algorithms	51
4.	Bandpass Filter Transfer Function Reconstruction Using The University of Tennessee Space Institute and Mayo Algorithms.	52
5.	Block Schematic of the Noise Rejection Test of The University of Tennessee Space Institute and Mayo Algorithms through Simulated Data.	54
6.	Block Schematic of Noise Rejection Test of FFT Algorithm with Simulated Data	54
7.	Part of Time History of Single Frequency Signal Plus Noise with S/N = 0.1.	55
8.	Part of Autocovariance of Time History of Fig. 7.	55

FIGURE PAGE

9. Spectra of Time History Shown in Fig. 7 with The University of Tennessee Space Institute Approach and Mayo's Approach. . . . 57

10. Spectra of Time History Shown in Fig. 7 Using FFT through Periodogram and Autocovariance Approaches, Respectively . . . 58

11. Spectra for Aliasing Comparison Using The University of Tennessee Space Institute Algorithm 60

12. Typical Experimental Setup. 63

13. Typical Plot of One of the Time Histories Acquired in the Experiment Shown in Fig. 12 64

14. Spectra of the Record Shown in Fig. 13 Using Three Different Algorithms. 65

15. Block Schematic of the Experimental Setup to Acquire Data for Band Limited White Noise Experiment. 67

16. Spectra of a Record Obtained in an Experiment Shown in Fig. 15 Using Three Different Algorithms. 68

C- 1. Pictorial Representation of Eq. (C-1) 89

C- 2. Spectral Window for Boxcar Weighting Function. 90

C- 3. Hanning Spectral Window Function. 93

C- 4. Pictorial Presentation of Aliasing. 94

FIGURE PAGE

D- 1. Process of Filtering in the Frequency Domain. 97

D- 2. Real and Imaginary Parts of Fourier Transformed Signal. 98

D- 3. Brick-Wall Filtered Frequency Response of Fig. D-2. 98

CHAPTER I

INTRODUCTION

The concept of period and frequency is at least as old as solar cycles, lunar cycles, and so forth. However, frequency domain analysis of signals did not start until the time of the late Joseph P. Fourier. His famous postulates made possible the consideration of signals from one domain in the other (e.g., from time to frequency and vice versa).

Noise is inherent in any physical system although it is usually unwanted. It is unavoidable in any measurement process. Its most undesirable property is that it often obscures recognition of time domain signals. Even in the absence of noise, some compound signals are not identifiable in the time domain. In such cases, it is often desirable to consider the signal in the frequency plane and then to examine its frequency components and their relative strengths. Spectrum analysis is a mathematical tool which transfers a time domain signal to its frequency plane resulting in the representation of a time series into its various frequency components.

The works of Nyquist [1][1] and Shannon [2] initiated the treatment of continuous signals in terms of only sampled data. Large scale processing of signals through sampled data methods started with the introduction of fast digital computers. Thus, such spectral analysis via sampled data became a reality. The use of a regular Fourier series to compute a spectrum is very time consuming, however, so a real time application cannot be achieved. A modification to the normal computation of Fourier series (by Cooley and Tukey [3]) leads to a significant saving in computational time (by a factor of approximately $N/\text{Log}_2 N$ for an N-point Fourier transform). This modified Fourier transform algorithm is popularly known as the Fast Fourier Transform (FFT).

Since its introduction, the FFT has become the most widely used tool for frequency domain analysis. Although the FFT is a well-established algorithm for computation of discrete Fourier transforms, at least in its most efficient form, it is limited to samples taken at equally spaced time intervals.

While considerable effort has produced realizable theories in the area of equi-spaced sampling, much remains to be done in the domain of random sampling. Randomly sampled signals or random discrete signals arise in diverse

---

[1]Numbers in brackets refer to similarly numbered references in the Bibliography.

fields of engineering. The amplitudes of such signals are in general also random. In such cases, there is no efficient tool like the FFT (as in regular sampling cases) for frequency domain analysis. The purpose of this work is to develop a faster (relative to existing ones) algorithm to estimate the spectra of randomly sampled data.

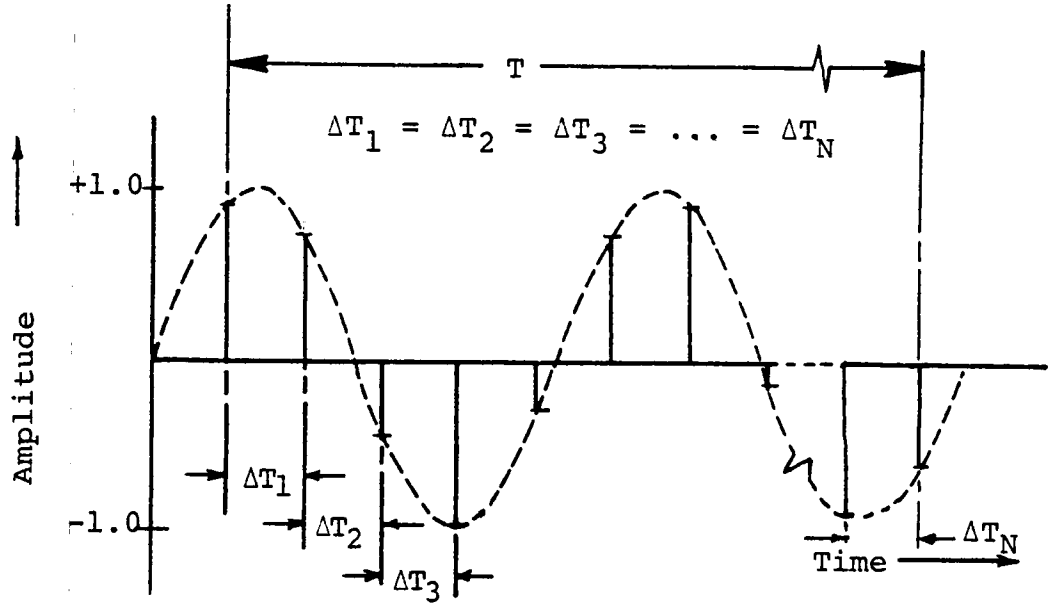
CHAPTER II

LITERATURE SURVEY

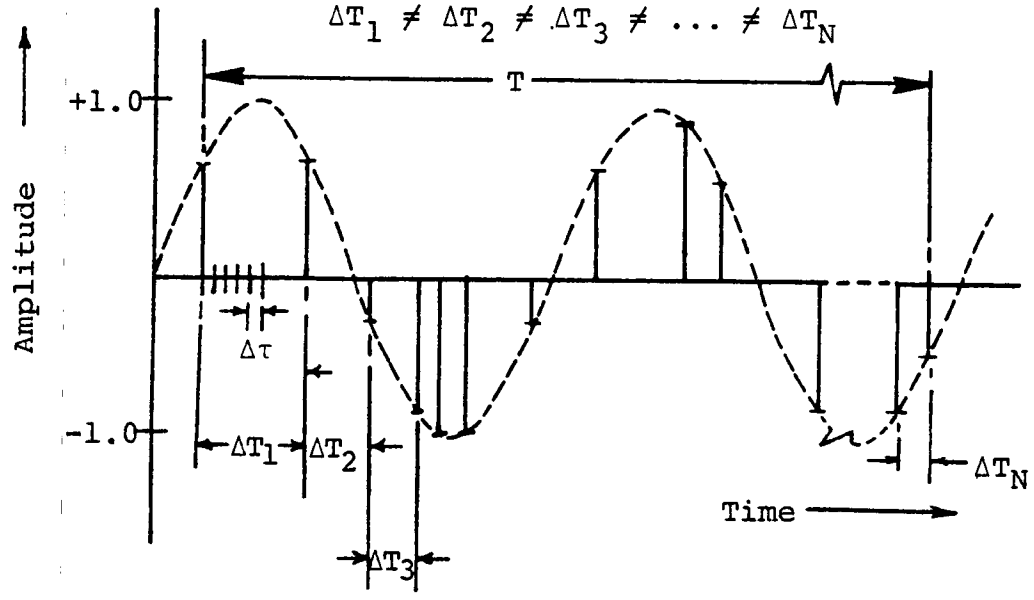
Discrete Signals and Laser Doppler Velocimetry

With the introduction of Fourier series and the Fourier transform, the era of frequency domain analysis began. Spectral estimation through analog techniques has become inadequate, mainly due to the finer frequency resolution and accuracy needed for present-day signal and noise studies. The use of digital computers has simplified such problems, permitting the rather simple computation of spectra from discrete samples. Broadly speaking, the samples may be taken in two different ways, with respect to a time grid (see Fig. 1). In the first, the sampling intervals are chosen equal, while in the second they are unequal and/or random.

For the case of equi-spaced samples, estimation of power spectra may be carried out using procedures described by Bendat and Piersol [4]. A vast amount of literature is available on the computation of the discrete Fourier transform for equi-spaced samples using Fast Fourier Transform (FFT) techniques--Refs. [5, 6, 7, 8] are but a few examples. However, the FFT method cannot be applied efficiently in the case of random sampling.



a. Equi-spaced Sampling



b. Randomly Spaced Sampling

Figure 1. Pictorial representation of major sampling schemes.

Random occurrence of samples is inherent in many diverse fields of engineering such as control and communication engineering, astronomy, meteorology, oceanography, and economics. Aircraft control guidance systems, radio receivers, radar echoes from distant planets, fluctuations in wind velocities and atmospheric temperatures, and vibration effects due to the tides on a ship are other examples of random phenomena which are subject to stochastic excitations.

An important instrument which produces data with random time samples is the Laser Doppler Velocimeter (LDV), a widely used noninterfering flow measuring instrument. Unlike the conventional methods using probes, it measures the flow parameters with satisfactory accuracy without reacting with the flow field. In this measurement, the experimentalist has no control over when the particles associated with the fluid will cross the laser beam and provide a data sample. So, the arrival times of data samples, obtained from scattering of the particles, are random and short-lived (often lasting less than a microsecond). If the observations are made in terms of velocities, the fluid velocity is known and the turbulence power spectrum can be computed for the position where the LDV is focused.

The objective of this study is to compute such power spectra either in pseudo-real time or a posteriori. An algorithm for computing power spectra is developed for

application to the LDV data. Parameters such as mean sampling rate which affect this computation are examined in detail.

Random Sampling and Laser Doppler Velocimeter Measurements

A large body of literature is available in the area of equi-spaced sampling schemes and comprehensive theories on power spectra calculations have resulted. However, the literature on random sampling is limited and much effort remains before comprehensive theories can be developed. In the case of equi-spaced sampling, aliasing[1] error is unavoidable unless care is taken to ensure that the highest signal frequency and the available sampling rate meet the Nyquist criterion. Blackman and Tukey [9] suggested that aliasing may be avoided, possibly by random sampling. This possibility was first investigated by Shapiro and Silverman [10], and later, the critical works of Beutler and Leneman [11, 12] and Masry [13] confirmed this valuable result. Their analyses show that random sampling is, in general, alias-free, even for a non-bandlimited signal. Also, they [11, 13] carried out a detailed analysis for Poisson sampling, an important case of random sampling, which is of particular interest here. Beutler has also shown that recovery of most of the spectrum is possible if a Poisson

[1] Aliasing is duscussed in more detail in Appendix C.

sampling is applied over the positive (or negative) half-axis in the time domain.

The Poisson probability law--where events occur independent of one another and which have a mean and variance both equal to μ--has the probability density function,

$$p(k,\mu) = \frac{\mu^k e^{-\mu}}{k!} . \tag{1}$$

Experimental investigation shows that the interval between detection of two particles in the laser probe volume is independent of all other intervals in the record. Although it is pointed out by Papoulis [14] that no natural phenomena exactly follow Eq. (1), the sampling law of the LDV is assumed closer to Poisson statistics than any other. Mayo, Shay, and Riter [15] have carried out investigations using these statistics. Accordingly, Shay [16] modeled the probability of having k-points in an interval of length τ as

$$\frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!} , \tag{2}$$

which is obviously a Poisson probability law with parameter λ, where λ is the mean sampling rate. The density function of their sum t_i is given by a gamma density function

$$f_i(\tau) = \frac{\lambda^i}{(i-1)!} \tau^{i-1} \exp(-\lambda\tau) u(\tau) , \tag{3}$$

with mean $E[\tau] = \frac{i}{\lambda}$ and variance $\sigma^2 = \frac{i}{\lambda^2}$. The function $u(\tau)$ is a unit step function. It is their conclusion [15] that LDV data follow Poisson sampling best when the data rate is low.

Scott [17] independently presented a model of Poisson sampling with dead time. He has shown that such instrument dead time has no harmful effect, and helps reduce the "sampling noise." Another similar investigation and modeling has been reported by Erdmann [18]. Considerable time is required to calculate velocity information from the laser velocimeter data. The time required to process a single velocity is usually referred to as instrument dead time and is dependent on the computer, interface logic circuits, and various signal acceptance tests. For a micro-computer based LDV (differential scatter type) system, this dead time is usually greater than or equal to 75 microseconds. A number of particles may pass unaccounted for through the probe volume. Erdmann used the LDV apparatus not to measure velocity, but to detect the arrival times of particles in a selected sampling volume. Thus, the dead time of this instrument was reduced by an order of magnitude and provided better information on particle distribution in the fluid. He conducted this experiment for both laminar and turbulent flows. Experimental results show that in laminar flow, the arrival rate distribution is Poisson for small mean and tends to be Gaussian as the mean

gets larger; in the case of turbulence it is more or less Gaussian.

In summary, equi-spaced sampling techniques provide error-free recovery if the sampling rate meets the Nyquist criterion (i.e., sampling rate $\geq \frac{\omega}{\pi}$, where $\omega = 2\pi f$, and f is the highest signal frequency). Violation of this criterion causes aliasing error which prevents accurate spectral calculation. Aliasing error may be avoided by random sampling of a signal so that improved spectral estimation is possible, particularly in the case of Poisson sampling. The data intervals in LDV measurements closely follow Poisson sampling statistics.

Spectral Estimators for Randomly Sampled Records

For equi-spaced samples, two basic methods have evolved to compute spectra using digital computers. In one method, correlation or covariance is computed first and then the transform is applied. In the other, the Fast Fourier Transform (FFT) is applied directly to the sampled data and the square of the absolute value of the complex Fourier coefficients is the raw periodogram or power spectrum.

Most of the computational methods for estimating spectra of randomly sampled data have been reported in the last decade. Thompson [19] in 1971 showed an approach to spectrum analysis based on complex demodulation. His

analysis was limited to only sinusoids rather than wide band signals. In the same year, Adegbola [20] reported another estimator based on the indirect approach, which is discussed in the following paragraphs.

The estimators reported to date may be broadly classified into two categories--indirect approach and direct approach. In the indirect approach, an equi-spaced discrete autocovariance is estimated from a randomly sampled record and then its transform is computed to obtain the spectral density. In the direct approach, the Fourier transform is applied to the data to estimate the power spectrum. These methods are described below.

Indirect Approach

The spectrum is expressed as (continuous time approximation of Eq. (B-16) in Appendix B):

$$S(\omega) = \int_{-\tau_m}^{\tau_m} w(\tau)\hat{C}(\tau,T)\cos(\omega\tau)\,d\tau \ , \qquad (4)$$

where τ_m represents maximum lag used to compute the autocovariance estimate and $\hat{C}(\tau,T)$ is the autocovariance estimate for data duration T:

$$\hat{C}(\tau,T) = \frac{1}{T-\tau_m} \int_{-T/2}^{\frac{T}{2}-\tau_m} \tilde{x}(t)\tilde{x}(t+\tau)\,dt$$

$$0 \leq \tau << T \ , \qquad (5)$$

and $w(\tau)$ is a suitable data window to reduce "leakage."[2] The term $\tilde{x}(t)$ in Eq. (5) is given by

$$\tilde{x}(t) = x(t) - \bar{x} , \tag{6}$$

where $\bar{x}$ is the mean value of $x(t)$ given by

$$\bar{x} = \frac{1}{T} \int_{-T/2}^{T/2} x(t)dt . \tag{7}$$

The mean is subtracted to avoid leakage from the zero frequency or d.c. value, particulary if it is large compared to the other frequency terms.

The discrete forms of Eqs. (5) and (7) are written as

$$\hat{C}(k,N) = \frac{1}{N-k} \sum_{n=1}^{N-k} \tilde{x}(n)\tilde{x}(n+k) \tag{8}$$

and

$$\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x(n) , \tag{9}$$

respectively. It is easy to see from Eq. (8), for each shift value, k, of the autocovariance, N-k lag products are calculated and summed. For this reason, the indirect approach is sometimes termed as the lagged product approach.

---

[2]The phenomenon "leakage" is explained in Appendix C.

Substituting Eq. (5) into Eq. (4) gives a double integral in the variables time t and lag τ:

$$\hat{S}(\omega) = \int_0^{\tau_m} \frac{1}{T-\tau_m} \int_{-T/2}^{\frac{T}{2}-\tau_m} \tilde{x}(t)\tilde{x}(t+\tau)w(\tau)\cos(\omega\tau)dt\,d\tau \qquad 0 \leq \tau << T \ . \quad (10)$$

The discretized form of Eq. (10) may be written as

$$\hat{S}_1(\omega) = \frac{1}{\pi\lambda^2 T} \sum_j^N \sum_k^N w(t_k-t_j)\tilde{x}(t_j)\tilde{x}(t_k) \cos \omega(t_k-t_j) \ , \quad (11)$$

where T is the record length and λ is the mean data rate of the random sampling scheme. This particular estimator is worked out by Gaster and Roberts [21]. Estimators similar to the one in Eq. (11) are found in Adegbola [20] and Masry and Lui [22, 23], with slight differences in multiplying factors and in limits of summation. This type of estimator is usually very slow computationally, as it is obvious that almost N^2 cosine values must be computed. Adegbola [20] presented a comprehensive analysis of his estimator, but failed to produce adequate experimental results.

Gaster and Roberts [21] tried an approach based on the theory in Ref. [10], estimating spectra without the use of sampling instants (as in the case of equally spaced samples) and found that such an approach produces unstable

estimates for a relatively large number of samples. They also developed the estimator in Eq. (11) and derived a theoretical model for the variance of the power spectrum estimator and compared their result with some experimentation. Their primary conclusion was that it produces stable estimates of power spectrum from reasonable amounts of data. Masry and Lui [22, 23] have also shown that this type of estimator is asymptotically unbiased. It is also alias-free, even if the mean data rate is less than the Nyquist rate, provided the sampling scheme follows a Poisson-type probability law.

Jones [24, 25, 26] tried several different estimators but found that one of them, termed the "discretizing the lag" method, produced a reasonably faster algorithm. In this procedure, an equi-spaced covariance is estimated using lag-products at discrete lags and then a transform is applied to the autocovariance to estimate the power spectrum. Jones, however, did not perform a detailed analysis. Later, Mayo, Shay, and Riter [15, 16] analyzed the estimator in detail and concluded that it is one of the best algorithms available to date. They termed this estimator the "discretized lagged product" method. This particular method, along with works of Fredrick [17] will be discussed later.

Regarding the lagged product method cited in Eq. (11), in brief, it may be said that a number of researchers proposed and analyzed such an estimator

independently. They came to the conclusion that the estimator produces a reasonably stable and alias-free spectrum from a considerable amount of data. Its disadvantage is that it is computationally slow.

Direct Approach

In the direct approach the power spectrum $S_d(\omega)$ is computed by applying a Fourier transform directly to the data. This method uses the definition of the raw periodogram (Eq. (B-11) in Appendix B) for the case of equi-spaced intervals, but with the following modification. To introduce the Poisson sampling process in place of regular time intervals, Eq. (B-11) in Appendix B may be written as

$$S(\omega) = \frac{1}{T}\left| \int_0^T x(t)\,\delta(t-t_n)\,e^{-j\omega t}\,dt \right|^2 , \qquad (12)$$

where

$$\delta(t-t_n) = 1 \quad ; \; t = t_n$$

$$= 0 \quad ; \text{ elsewhere} ,$$

and also

$$\int \delta(t-t_n)\,dt = 1 .$$

The discrete form of Eq. (12) is

$$\hat{S}(\omega) = \frac{1}{T}\left| \sum_{n=1}^{N} \tilde{x}_n e^{-j\omega t_n} \right|^2 , \tag{13}$$

where $\tilde{x}_n = x(t_n) - \bar{x}$, and $\bar{x}$ is the mean of process $x(t_n)$.

Again, if λ is the mean sample rate and $R(0)$ is the autocovariance at zero lag, the direct estimator for the power spectrum $\hat{S}_d(\omega)$ is written in terms of $\hat{S}(\omega)$ as

$$\hat{S}_d(\omega) = \frac{1}{\lambda^2}\hat{S}(\omega) - \frac{R(0)}{\lambda} . \tag{14}$$

$R(0)$ and λ are estimated as

$$\hat{\lambda} = \frac{N}{T} \tag{15}$$

and

$$\hat{R}(0) = \frac{1}{N}\sum_{n=1}^{N} \tilde{x}_n^2 . \tag{16}$$

After insertion of $\hat{\lambda}$, $\hat{R}(0)$, and $\hat{S}(\omega)$ in Eq. (14) and rearranging, the direct estimator takes the form

$$\hat{S}_d(\omega) = \frac{T}{N^2}\left[\left| \sum_{n=1}^{N} \tilde{x}_n e^{-j\omega t_n} \right|^2 - \sum_{n=1}^{N} \tilde{x}_n^2 \right] . \tag{17}$$

A suitable data window $w(t_n)$ may be added to Eq. (14) to avoid leakage, then $\hat{S}_d(\omega)$ looks like

$$\hat{S}_d(\omega) = \frac{T}{N^2}\left[\left|\sum_{n=1}^{N} \tilde{x}_n w(t_n) e^{-j\omega t_n}\right|^2 - \sum_{n=1}^{N} \tilde{x}_n^2\right] . \quad (18)$$

Such an estimator was studied by Gaster and Roberts [27]. They made a comparative study of the two approaches--indirect (through autocovariance and transform) and direct, based on computing time and variability error. They found that the indirect approach has less variability error than the direct approach; whereas, the latter method is faster than the former for the same amount of data.

The procedure for comparisons as given in Ref. [27] is as follows: Let N be the total number of samples taken over time T and λ be the mean sampling rate. The parameters N, T, and λ are the same for both methods. Assuming trigonometric functions require the major part of computing time and each data point needs one sine and one cosine computation, the total time taken by N data point in the direct method is

$$t_{dir} \simeq 2N . \quad (19a)$$

To have the same frequency resolution using the indirect method the maximum allowable lag value is $\tau_m = T/2$. Thus, for Eq. (11),

$$t_{indir} = N \cdot \tau_m \cdot \lambda \simeq N \cdot \frac{T}{2} \cdot \frac{N}{T} \simeq \frac{N^2}{2} . \quad (19b)$$

Again, to hold the variability error to the same level, the direct method needs two and two-thirds times the number of samples needed in the indirect method. When this condition is applied, $t_{dir} = \frac{16}{3} N$. Now, the direct method will be faster than the indirect method if

$$\frac{16N}{3} < \frac{N^2}{2} , \tag{19c}$$

inequality (Eq. (19c)) holds for

$$N > 11 , \tag{19d}$$

or (using Eq. (19b)),

$$T\lambda > 11 . \tag{19e}$$

The above constraint is met in many cases. A point to note is that Gaster and Roberts [27] took $t_{dir} = N$ (for one data point sine and cosine computations are probably assumed as one trigonometric computation). Thus, the above constraint comes out to be $T\lambda > 5$.

The above comparison is an encouraging factor for taking up a direct approach to develop a simple and computationally fast algorithm. The development of the algorithm is discussed later in detail.

In summary, it is found that the direct approach is faster, but takes a larger amount of data to produce the same quality of spectrum obtained using the indirect method.

This requirement is not stringent at all and in many cases the required data length is achieved easily.

Discretized Lag Product Method (Mayo's Approach)

This particular estimator is analyzed below. Later, the results of this estimator will be compared with that of the estimator developed in this investigation.

In this method, the time intervals ΔT_i between pairs of samples x_i are quantized with respect to a unit of time $\Delta\tau$, as shown in Fig. 1b, page 5. The term $\Delta\tau$ is selected such that the expected maximum signal frequency $f_{max} \leq \frac{1}{2\Delta\tau}$, as required by the Nyquist criterion [1].

The sample mean is subtracted first and a limit is imposed on the maximum time interval ΔT_{max} between pairs of samples. If the time interval is not greater than ΔT_{max}, then a product is performed using two samples corresponding to the interval and the product is accumulated in an array SUM(k) and a corresponding count is maintained in a histogram array H(k). The value of k is determined as

$$k = \frac{(\Delta T)_i}{\Delta\tau} ; i = 1,2,3...,(N-1) ,$$

where

k ranges from 0 to (M-1) ,

N = total number of samples taken over a time T ,

$\Delta\tau$ = minimum time spacing .

If enough data are taken it is reasonable to expect that the spectrum will be computed for M frequencies.

The sum of lag products, SUM(k) at each integer multiple of lags 0, $\Delta\tau$, $2\Delta\tau$,...,$(M-1)\Delta\tau$, is to be performed. Also, $\Delta T_{max} = (M-1)\Delta\tau$; SUM(k) is expressed as

$$SUM(k) = \sum \tilde{x}_i \, \tilde{x}_j$$

for all i,j, where

$$\left[\frac{t_i - t_j}{\Delta\tau} \right] = k$$

and

$$|i - j| \leq M-1 \ .$$

To be more understandable, with respect to Fig. 1 (page 5), let the value of i be held to 1, and j is varied from 1 to N. Then, i is incremented to 2, j = 2,N or j = i,N-- it continues until i = N. In this way, two arrays,

$$SUM(k) \quad ; \ k = 0,1,\ldots,(M-1)$$

$$H(k) \quad ; \ k = 0,1,\ldots,(M-1)$$

are computed. The autocovariance is obtained as

$$\hat{C}(k) = \frac{SUM(k)}{H(k)} \quad ; \ k = 0,1,\ldots,(M-1) \ . \qquad (20a)$$

Now, if $M = 2^p$, where p is an integer, an FFT algorithm can be applied as an M-point autocovariance estimate. This

gives

$$\hat{S}(i\Delta\omega) = F\,[\,\hat{C}(k)\,] \quad ; \quad i = 0,1,\ldots,M-1 \; . \tag{20b}$$

If M is not an integer multiple of 2, a cosine transform may be applied as shown below:

$$\hat{S}(\ell\Delta\omega) = \sum_{k=-(M-1)}^{M-1} \hat{C}(k)\, \exp \frac{-j2\pi\ell k}{2M} \; , \tag{21}$$

where

$$\ell = 1,2,\ldots,M$$

$$\Delta\omega = 1/(2M\Delta\tau) \; .$$

Equation (21) may be written as

$$\hat{S}(\ell\Delta\omega) = \sum_{k=-(M-1)}^{M-1} \hat{C}(k)\, \exp \frac{-j\pi\ell k}{M} + \hat{C}(0) + \sum_{k=1}^{M-1} \hat{C}(k)\, \exp \frac{-j\pi\ell k}{M}$$

$$= \sum_{k=M-1}^{1} \hat{C}(k)\, \exp \frac{j\pi\ell k}{M} + \hat{C}(0) + \sum_{k=1}^{M-1} \hat{C}(k)\, \exp \frac{-j\pi\ell k}{M} \; .$$

Now, $\hat{C}(k)$ is an even function, $\hat{C}(k) = \hat{C}(-k)$; then,

$$\hat{S}(\ell\Delta\omega) = \hat{C}(0) + \sum_{k=1}^{M-1} \hat{C}(k) \cdot 2 \left[\frac{1}{2} \left(\exp \frac{j\pi\ell k}{M} + \exp \frac{-j\pi\ell k}{M} \right) \right]$$

or

$$\hat{S}(\ell\Delta\omega) = \hat{C}(0) + 2 \cdot \sum_{k=1}^{M-1} \hat{C}(k)\, \cos \left(\frac{\pi\ell k}{M} \right) \; . \tag{22}$$

Note that if the FFT is applied, the frequency information available is $f_{max} = \frac{1}{2\Delta\tau}$, or $\frac{M}{2}$ frequency terms with resolution $\frac{1}{2M\Delta\tau}$. Whereas, using the cosine transform, for an M-point autocovariance estimate, M points frequency values are produced. It is of interest to determine the number of products necessary to calculate the discretized autocovariance. If i = 1, and j varies from 1 to N, the number of products = N; for i = 2, j = 2,...,N, the number of products = (N-1); likewise, i = N, j = N, the number of products = 1. Thus, the total number of products = $N + (N-1) + \ldots + 1 = \frac{N(N+1)}{2}$. To compute an M-point autocovariance, $\frac{N(N+1)}{2}$ multiplications and M divisions are required (neglecting the addition operations).

When data are randomly sampled, covariances are estimated at regularly spaced intervals to take advantage of the FFT. In so doing, the limitations due to aliasing are again introduced. The lag product method quantizes intervals between pairs of samples. In addition to the usual errors generated in computing power spectra, quantization error, inherent in analog-digital conversion techniques, is introduced. Sufficient experimental results are not available to accept it as a practical estimator. Smith and Meadows [29] reported the use of this algorithm and presented some results, but they used a large number of data points.

The estimator of Scott [17], for all practical aspects, is similar to Mayo's estimator. Here also, a regularly spaced autocovariance estimate is generated out of a random record with a little different computer logic. The lag time τ_i is given a tolerance $\pm\frac{\Delta\tau}{2}$. This may be considered as the quantizing unit of time $\Delta\tau$ chosen in Mayo's approach. Nothing is mentioned about the computational time or speed of his algorithm.

It may be concluded that the discretized lagged product technique seems to be a more convincing candidate than the lagged product method written in Eq. (11). However, it is not as fast as one might expect because a large number of lagged products are to be performed as a pre-transform requirement. The regularization process (forcing time to be multiples of $\Delta\tau$) introduces quantization error and aliasing into the resultant spectrum.

Other Techniques

Beuther [28] talks about an autoregressive technique estimation. This technique is particularly useful when available data length is unusually short and one needs to extract maximum information out of it. It may be applied for long data length, but will require increased computing time. In the literature, one finds its use very limited in this area.

In the direct method, as discussed earlier, a significant amount of time is spent in computing

trigonometric functions like sines and cosines. Roberts and Gaster [30] proposed two schemes to approximate sine and cosine functions through equivalent periodic rectangular waveforms. This will enhance the speed of computation of spectra by the direct method expressed in Eq. (18). However, in using the above schemes individually or in a combined fashion, one should be careful with the limitations imposed by the assumptions. The saving of time is not discussed in quantitative figures.

Many approaches based on linear regression techniques have been tried by Norsworthy [31] to estimate spectra. One of his approaches is somewhat similar to the work reported here (to be discussed later in detail). He concluded that regression techniques produce stable spectra if the number of coefficients is much less than the number of data points. Mayo [32] reported similar conclusions. Neither Norsworthy nor Mayo elaborated on their findings.

There is evidence that a linear regression technique does produce stable estimates for the coefficients if the number of data points is much larger than that of the coefficients to be computed. It should be noted that the theory of least square error fit is based on such an assumption. An approach based on the Fourier series and a least square fit technique is the subject of this investigation. The algorithm developed is discussed and analyzed in the following chapters.

CHAPTER III

STATEMENT OF THE PROBLEM

It was found that aliasing, inherent in equi-spaced sampling processes, can be eliminated by random sampling [11, 12]; Poisson sampling, in particular, is alias-free. A number of different algorithms for estimating spectra of randomly occurring discrete signals have been developed which seek to take advantage of this important concept. Some of them are slow computationally. The faster ones convert randomly spaced signals to an equi-spaced covariance function and then take advantage of the speed of an FFT algorithm. In so doing, not only the originality of the random data is lost in the process of quasi-regularization, but aliasing errors are also introduced. On the other hand, Gaster and Roberts [27, 30] have shown that a direct method is computationally faster than the indirect ones and provides stable spectra from the relatively long record which is usually available.

A new computationally fast algorithm, which may be used for both pseudo-real or non-real time applications, is the subject of this study. The algorithm takes advantage of the random character of the data acquisition to prevent aliasing. This study shows that, with a sufficient number of data points, the sampling rate can be much lower than the highest frequency to be measured.

The method of least squares is well known and provides a reliable method for estimating system parameters. Also, it is used to estimate statistical errors in experimental data. The present algorithm for unequally spaced sampling is based on the least square estimation technique and also upon the simple observation that the mean value of a randomly sampled sine or cosine is zero. In addition to retaining the fully random property of the data, this algorithm provides a reliable error estimate for each of the spectral coefficients. Since the algorithm is relatively simple, it can be applied easily in applications utilizing micro-computers and mini-computers.

The algorithm has been tested with both simulated and experimental data. Through use of a computer, band limited Gaussian noise has been generated in discrete form and used as input data. Noise (white noise with a large mean) rejection ability has also been tested with the help of computers. Single frequency as well as multiple frequency (band limited noise) characteristics have been determined experimentally. A discussion of the results is included and a comprehensive error analysis has been performed. Additionally, a brief discussion regarding the hardware implementation of the process is included for completeness.

CHAPTER IV

DEVELOPMENT OF THE ALGORITHM

The system of equations relating dependent y and independent variables x_i is written as

$$\begin{aligned}
y^{(1)} &= \beta_0 + \beta_1 x_1^{(1)} + \beta_2 x_2^{(1)} + \ldots + \beta_M x_M^{(1)} \\
y^{(2)} &= \beta_0 + \beta_1 x_1^{(2)} + \beta_2 x_2^{(2)} + \ldots + \beta_M x_M^{(2)} \\
&\vdots \\
y^{(N)} &= \beta_0 + \beta_1 x_1^{(N)} + \beta_2 x_2^{(N)} + \ldots + \beta_M x_M^{(N)}
\end{aligned} \quad , \tag{23}$$

where a subscript implies the coefficient number, and a superscript, the observation number. The vector-matrix form for Eq. (23) is given by

$$\underline{Y} = X\underline{\beta} \ , \tag{24}$$

where

$\underline{Y}$ is a column vector (N,1) ,

$\underline{\beta}$ is also a column vector (M,1) ,

X is a rectangular matrix (N,M) .

Based on the technique of the least squares (Appendix A), Eq. (24) has the solution

$$\underline{\beta} = (X'X)^{-1} X'\underline{Y} \ , \tag{25}$$

where X' is the transpose of the X matrix, and $(X'X)^{-1}$ is the inverse of (X'X). Note that (X'X) is a square matrix (M,M), and $X'\underline{Y}$ is a column vector (M,1).

As in the system of Eq. (23), a velocity observation at any instant of time t_k, say, $V(t_k)$, can be expressed in terms of a finite length Fourier series:

$$\begin{aligned}
V(t_1) &= a_0 + a_1\cos(\omega t_1) + a_2\cos(2\omega t_1) + \ldots + a_M\cos(M\omega t_1) \\
&\quad + b_1\sin(\omega t_1) + b_2\sin(2\omega t_1) + \ldots + b_M\sin(M\omega t_1) \\
V(t_2) &= a_0 + a_1\cos(\omega t_1) + a_2\cos(2\omega t_2) + \ldots + a_M\cos(M\omega t_2) \\
&\quad + b_1\sin(\omega t_1) + b_2\sin(2\omega t_2) + \ldots + b_M\sin(M\omega t_2) \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
V(t_N) &= a_0 + a_1\cos(\omega t_N) + a_2\cos(2\omega t_N) + \ldots + a_M\cos(M\omega t_N) \\
&\quad + b_1\sin(\omega t_N) + b_2\sin(2\omega t_N) + \ldots + b_M\sin(M\omega t_N)
\end{aligned} \qquad (26)$$

In the system of Eq. (26), a_0 is the average or d.c. value of the series and is the expected average velocity. Once the average is computed and subtracted, Eq. (26) may be written as

$$V(t_k) - a_0 = \sum_{i=1}^{M} \{ a_i \cos(i\omega t_k) + b_i \sin(i\omega t_k) \} \qquad (27)$$

or

$$\tilde{v}(t_k) = \sum_{i=1}^{M} \{ a_i \cos(i\omega t_k) + b_i \sin(i\omega t_k) \} . \qquad (28)$$

The term a_0 is estimated by

$$a_0 \simeq \frac{1}{N} \sum_{k=1}^{N} V(t_k) \ , \tag{29}$$

and $\tilde{v}(t)$ is the fluctuating part of the observations. The vector-matrix representation of Eq. (28) is given by

$$\underline{\tilde{V}} = T \, \underline{C} \ , \tag{30}$$

where

$$\underline{\tilde{V}}' = \left[\tilde{v}(t_1) \ , \ \tilde{v}(t_2), \ldots, \tilde{v}(t_N) \right] \ ,$$

$$\underline{C}' = \left[a_1 \ , \ a_2 , \ldots , a_M , b_1 , \ldots , b_M \right] ,$$

and T is a rectangular matrix whose elements are the trigonometric cosine or sine values. In Eq. (30), $\underline{C}$ is the unknown coefficient vector. The vector $\underline{C}$ has the least square error solution,

$$\underline{C} = (T'T)^{-1} \ (T'\underline{\tilde{V}}) \ . \tag{31}$$

In Eq. (31) (T'T) is a symmetric square matrix (2M,2M). It is difficult to find the inverse of the (T'T) matrix, particularly if most of the elements of (T'T) are non-zero. Let (T'T) be partitioned as

$$(T'T) = \left[\begin{array}{c:c} P1 & P2 \\ \hdashline P3 & P4 \end{array} \right] \tag{32}$$

such that each of the submatrices P1, P2, P3, and P4 is also a square matrix (M,M). As (T'T) is symmetric, P3 can be obtained by transposing P2 and vice versa. The elements of the above submatrices have the forms:

$$p1_{ij} = \sum_{k=1}^{N} \cos(i\omega t_k) \cos(j\omega t_k) , \qquad (33)$$

$$p2_{ij} = \sum_{k=1}^{N} \cos(i\omega t_k) \sin(j\omega t_k) , \qquad (34)$$

$$p4_{ij} = \sum_{k=1}^{N} \sin(i\omega t_k) \sin(j\omega t_k) . \qquad (35)$$

In the above three equations, i and j have the range,

$$1 \leq i \leq M$$

$$1 \leq j \leq M .$$

The (T'T) matrix is real and symmetric. Also, the magnitudes of the off-diagonal elements are no more than one-fifth of the diagonal elements. If the off-diagonal elements of (T'T) are set to zero, then the matrix becomes diagonal and real, as shown below:

$$(T'T) = \begin{bmatrix} \sum_{k=1}^{N} \cos^2(\omega t_k) & & & 0 \\ & \ddots & & \\ & & \sum_{k=1}^{N} \cos^2(M\omega t_k) & \\ & & & \sum_{k=1}^{N} \sin^2(\omega t_k) \\ & & & \ddots \\ 0 & & & \sum_{k=1}^{N} \sin^2(M\omega t_k) \end{bmatrix}, \quad (36)$$

and its inverse is

$$(T'T)^{-1} = \begin{bmatrix} 1\Big/\sum_{k=1}^{N} \cos^2(\omega t_k) & & & 0 \\ & \ddots & & \\ & & 1\Big/\sum_{k=1}^{N} \cos^2(M\omega t_k) & \\ & & & 1\Big/\sum_{k=1}^{N} \sin^2(\omega t_k) \\ & & & \ddots \\ 0 & & & 1\Big/\sum_{k=1}^{N} \sin^2(M\omega t_k) \end{bmatrix}. \quad (37)$$

Note that the $(T'T)^{-1}$ computation is greatly simplified since the diagonal elements of $(T'T)^{-1}$ are the reciprocal of the corresponding elements of $(T'T)$. In fact, once the above assumption is applied, the system of equations need not be computed as a matrix, but rather, as a simple one-dimensional array. Usually, in matrix computations, all of the element values must first either be supplied or computed before any operation (such as inverse, etc.) is performed. But in this simplified method, computation of the coefficients does not require a priori computation of all the elements of $(T'T)$. The effects of setting off-diagonal elements to zero will be discussed later in this chapter.

Using Eq. (31), expressions for the Fourier coefficients become

$$\hat{a}_i = \frac{\sum_{k=1}^{N} \cos(i\omega t_k)\tilde{v}(t_k)}{\sum_{k=1}^{N} \cos^2(i\omega t_k)} \tag{38a}$$

and

$$\hat{b}_i = \frac{\sum_{k=1}^{N} \sin(i\omega t_k)\tilde{v}(t_k)}{\sum_{k=1}^{N} \sin^2(i\omega t_k)} \; . \tag{38b}$$

The spectral components are (see Appendix B)

$$\hat{S}(i\omega) = \hat{a}_i^2 + \hat{b}_i^2$$
$$i = 1,2,\ldots,M \; . \tag{39}$$

Since the spectral analysis measurement is concerned more with relative frequency content than absolute magnitude, in the forthcoming sections on statistical analysis and error analysis it is shown that no appreciable information is lost by setting the off-diagonal elements to zero. However, an order of magnitude is gained in computing time and required memory space. These factors make the algorithm comparatively fast and simple.

In the above development it should be noted that the new estimator does not exactly give a least square fit, rather, it yields a "modified" least square fit. The coefficients, using least square criterion, give

$$\hat{\underline{C}} = (T'T)^{-1} T'\tilde{\underline{Y}} , \tag{40}$$

whereas the new estimator takes the form

$$\tilde{\underline{C}} = D^{-1} T'\tilde{\underline{Y}} . \tag{41}$$

Here D, the diagonal matrix, is derived from (T'T) = D + F, where D contains the diagonal terms, and F, the off-diagonal ones of (T'T). The important characteristic in developing the equation is the assumption that the elements of F are equal to zero. This not only helps to speed up the computation but also greatly simplifies the analytical expressions for $\underline{C}$. Also, it might be noted that the analytical expressions for coefficients are different from those of the direct method in Eq. (12), but similar.

Statistical Analysis of the Estimator

Any good statistical estimator should satisfy the following three inequalities:

$$\text{Unbiased:} \quad E[\hat{\Phi}] = \Phi \ ,$$

$$\text{Consistent:} \quad \lim_{N\to\infty} E\left[\left(\hat{\Phi} - \Phi \right)^2 \right] = 0 \ ,$$

and

$$\text{Efficient:} \quad E\left[\left(\hat{\Phi} - \Phi \right)^2 \right] - E\left[\left(\hat{\Phi}_i - \Phi \right)^2 \right]$$

for any other estimate $\hat{\Phi}_i$, where Φ is the actual value of the parameter, and $\hat{\Phi}$ is the statistical estimate of the parameter. Here, E[ ] is the statistical operator known as the expected value. These inequalities are "measures of goodness" for estimates. In the case of a power spectrum estimator, it is customary to show the bias error and consistency. Analysis for efficiency is usually complicated, and for all practical purposes is impossible, since all estimators cannot be considered. The following analysis will deal with the bias error and consistency of the power spectrum estimator developed earlier in this chapter.

The bias error of an estimator $\hat{\Phi}$ of the parameter Φ is defined as

$$B[\hat{\Phi}] = E[\hat{\Phi}] - \Phi$$

If the bias error is zero, i.e., $E[\hat{\Phi}] = \Phi$, then it is known that the estimator is unbiased. The estimator under discussion is expressed as

$$\tilde{\underline{C}} = D^{-1}\, T'\tilde{V} \ , \tag{43}$$

where the true model is

$$\underline{V} = T\,\underline{C} + \underline{n} \ . \tag{44}$$

Here, $\underline{n}$ is the noise vector with the mean equal to zero. The diagonal matrix D comes from $T'T = D + F$, where D is the diagonal, and F is the matrix whose diagonal elements are zero,

$$E[\tilde{\underline{C}}] = E\,[\,D^{-1}\,T'\,(T\underline{C} + \underline{n})\,]$$

$$= E\,[\,D^{-1}\,T'T\,\underline{C}\,] + E\,[\,D^{-1}\,T'\,]\cdot E[\underline{n}]$$

$$= E\,[\,D^{-1}\,(D + F)\,\underline{C}\,]$$

or

$$E[\tilde{\underline{C}}] = E[\underline{C}] + E\,[\,D^{-1}\,F\,\underline{C}\,] \ . \tag{45}$$

In derivation of Eq. (45) it is assumed that $(D^{-1}\,T')$ and $(\underline{n})$ are statistically independent. Now the estimator will be unbiased if

$$E[\tilde{\underline{C}}] = E[\underline{C}] \ ,$$

from Eq. (45) $E[D^{-1}F\,\underline{C}]$ is the bias error.

Let $D^{-1}\,F = R$. The element r_{ij} may take one of the following forms:

$$\frac{\sum_{k=1}^{N} \cos(m\omega t_k)\cos(n\omega t_k)}{\sum_{k=1}^{N} \cos^2(\ell\omega t_k)}, \tag{46}$$

$$\frac{\sum_{k=1}^{N} \cos(m\omega t_k)\sin(n\omega t_k)}{\sum_{k=1}^{N} \cos^2(\ell\omega t_k)}, \tag{47}$$

$$\frac{\sum_{k=1}^{N} \cos(n\omega t_k)\sin(m\omega t_k)}{\sum_{k=1}^{N} \sin^2(\ell\omega t_k)}, \tag{48}$$

$$\frac{\sum_{k=1}^{N} \sin(m\omega t_k)\sin(n\omega t_k)}{\sum_{k=1}^{N} \sin^2(\ell\omega t_k)}. \tag{49}$$

Note that R is a square matrix (2M,2M), and ℓ, m, and n are integers. The integers ℓ, m, and n have a maximum range up to M. Also, $r_{ij} = 0$ for $i = j$.

The numerator of Eq. (46) may be written as

$$\sum_{k=1}^{N} \cos(m\omega t_k)\cos(n\omega t_k)$$

$$= \frac{1}{2} \sum_{k=1}^{N} \left(\cos\left((m+n)\omega t_k\right) + \cos\left((m-n)\omega t_k\right) \right) .$$

Equation (46) is written as

$$\frac{\frac{1}{2}\sum_{k=1}^{N}\cos\left((m+n)\omega t_k\right)}{\sum_{k=1}^{N}\cos^2(\ell\omega t_k)} - \frac{\frac{1}{2}\sum_{k=1}^{N}\cos\left((m-n)\omega t_k\right)}{\sum_{k=1}^{N}\cos^2(\ell\omega t_k)} .$$

Based on statistical analyses it can be shown that

$$E\left[\lim_{N\to\infty}\sum_{k=1}^{N}\cos^2(\ell\omega t_k)\right] = \frac{N}{2} ,$$

where ℓ is an integer, and

$$E\left[\lim_{N\to\infty}\frac{1}{2}\sum_{k=1}^{N}\cos((m+n)\omega t_k)\right] = 0$$

$$E\left[\lim_{N\to\infty}\frac{1}{2}\sum_{k=1}^{N}\cos((m-n)\omega t_k)\right] = 0 .$$

Thus, expected value of Eq. (46) goes to zero in the limit. In the same way, expected value of Eqs. (47), (48), and (49) may be shown to approach zero as N tends to infinity. In general, r_{ij} of the R matrix behaves the same way as N approaches infinity.

Let

$$D^{-1}\,F\underline{C} = R\underline{C} = \underline{S}$$

$$s_i = \sum_{j=1}^{2M} r_{ij}c_j , \qquad (50)$$

where

$$\underline{C}' = [\, c_1 , c_2 , \ldots , c_{2M} \,] \ .$$

In Eq. (44), $\underline{C}$ is the characteristic, or true, solution of the system. The vector $\tilde{\underline{C}}$ is a random variable, but $\underline{C}$ is not. Thus, $E[s_i]$ becomes zero as $E[r_{ij}]$ approaches zero in the limit as N approaches infinity. Thus, for large N, bias error approaches zero. Hence, the estimator under study is asymptotically unbiased.

Variance of the estimate is written as

$$\text{Var}\ [\tilde{\underline{C}}] = E\ [\ (\tilde{\underline{C}} - \underline{C})(\tilde{\underline{C}} - \underline{C})'\ ] \ . \qquad (51)$$

With the help of Eqs. (43) and (44) the difference between $\tilde{\underline{C}}$ and $\underline{C}$ is found to be

$$\tilde{\underline{C}} - \underline{C} = D^{-1}\, F\underline{C} + D^{-1}\, T'\underline{n} \ . \qquad (52)$$

Then,

$$\begin{aligned}(\tilde{\underline{C}} - C)\ (\tilde{\underline{C}} - \underline{C})' &= (D^{-1}\, F\underline{C} + D^{-1}\, T'\underline{n})(\underline{C}'F'D^{-1} + \underline{n}'TD^{-1}) \\ &= D^{-1}\,(F\underline{C} + T'\underline{n})\ (\underline{C}'F' + \underline{n}'T)\, D^{-1} \\ &= D^{-1}\,(F\underline{C}\ \underline{C}'\ \underline{F} + 2T'\underline{n}\underline{C}'F' + T'\underline{n}\underline{n}'T)\, D^{-1} \ . \qquad (53)\end{aligned}$$

Thus,

$$\begin{aligned}\text{Var}\ [\tilde{\underline{C}}] &= E\ [\, D^{-1}\, F\underline{C}\ \underline{C}'\ F'D^{-1}\,] \\ &\quad + E\ [\, 2D^{-1}\, T'\underline{n}\ \underline{C}'\ F'D^{-1}\,] \\ &\quad + E\ [\, D^{-1}\, T'\underline{n}\ \underline{n}'T\ D^{-1}\,] \ , \qquad (54)\end{aligned}$$

or

$$\text{Var}\,[\tilde{\underline{C}}] = E[\underline{S}\,\underline{S}'] + 2E[D^{-1}T']E[\underline{n}]\,E[\underline{S}'] + E[D^{-1}T']E[\underline{n}\,\underline{n}']E[TD^{-1}],$$

where $S = D^{-1}F\,\underline{C}$, as in Eq. (50), or

$$\text{Var}\,[\tilde{\underline{C}}] = E[\underline{S}\,\underline{S}'] + E[D^{-1}T']E[\underline{n}\,\underline{n}']E[TD^{-1}] .$$

Using the procedure for finding the limit of bias error, E[SS'] approaches zero as N becomes large. From Eq. (44) the elements of T or T' may be found. Now the denominators of elements of $D^{-1}T$ or TD^{-1} matrices will be either

$$\sum_{k=1}^{N} \cos^2(i\omega t_k)$$

or

$$\sum_{k=1}^{N} \sin^2(i\omega t_k) .$$

For large N,

$$E\left[\sum_{k=1}^{N} \cos^2(i\omega t_k)\right] = E\left[\sum_{k=1}^{N} \sin^2(i\omega t_k)\right] = \frac{N}{2} .$$

Now, $D^{-1}T'$ or TD^{-1} elements look like

$$\frac{\cos(i\omega t_k)}{N/2} \quad \text{or} \quad \frac{\sin(i\omega t_k)}{N/2} \quad ; \quad \begin{matrix} i = 1,2,\ldots,M \\ k = 1,2,\ldots,N \end{matrix} ,$$

and in the limit $N \to \infty$, those elements go to zero. Hence, the estimator is consistent for a large sample size.

Error Analysis of the Estimator

In solving the system of linear equations (Eq. (26)) using the least square error fit method, the off-diagonal elements of (T'T) are neglected to simplify matrix storage and inversion problems. The matrix (T'T) is studied to further understand the effects of this approximation. The matrix (T'T) becomes diagonal for the equispaced sampling case with spectral resolution = 1/T, where T is the total time covered by the data set. For random sampling (T'T) is no longer a diagonal matrix. The random sampling coupled with flexible frequency resolution (provided that $M \cdot f_{res} \leq \lambda$ is satisfied), the case with the present algorithm, makes (T'T) deviate from a diagonal matrix. It is not even always diagonally dominant. A matrix P is diagonally dominant, by definition of Ref. [33], if it satisfies for a row or column;

$$\sum_{\substack{j \\ j \neq i}} | p_{ij} | \leq | p_{ii} | , \tag{55}$$

where p_{ij} is the i-j^{th} element of P. Let (T'T) = P. A careful study of some typical computations of the elements of (T'T) indicates that

$$\sum_{i \neq j} |p_{ij}| \leq \frac{1}{5} |p_{ii}| . \qquad (56)$$

Also, it has been observed that when (T'T) is diagonally dominant, the new approach provided a solution closer to the true one (obtained using Gaussian elimination) and the iteration process in Eq. (60) becomes convergent. An attempt has been made to get an idea of the amount of loss in information due to neglecting the off-diagonal elements of (T'T), by computing the ℓ_1 norms of (T'T), D, and F, where (T'T) = D+F, D is the diagonal matrix consisting of diagonal elements of (T'T), and F contains only the off-diagonal elements. Matrix norms have been computed using the procedures mentioned in Ref. [34]. The signal used in this study is

$$x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t) + \sin(2\pi f_3 t) + \sin(2\pi f_4 t) , \qquad (57)$$

where

$$f_1 = 1 , \; f_2 = 4 , \; f_3 = 7 , \; f_4 = 9.5 .$$

The signal in Eq. (57) is selected so signal frequencies are distributed over the observation band; sampling is at a sub-Nyquist rate. The parameters for this typical case are:

1. Data points, N = 200.
2. Number of coefficients, 2M = 40.
3. Mean sampling rate $\lambda \simeq 10$ samples/sec.
4. Signal, $f_{max} = 9.5$ Hz.
5. Frequency resolution, $f_{res} = 0.5$ Hz.

Matrix norms are:

1. ℓ_1 matrix norm for (T'T) is 300.19.
2. ℓ_1 matrix norm for D is 117.49.
3. ℓ_1 matrix norm for F is 192.28.

When the sample size is increased to N = 600, the matrix becomes diagonally dominant and the matrix norms are as follows:

1. ℓ_1 matrix norm for (T'T) is 615.97.
2. ℓ_1 matrix norm for D is 321.53.
3. ℓ_1 matrix norm for F is 318.42.

The matrix also becomes diagonally dominant when λ is increased from 10 to 35 (approximately), keeping N = 200:

1. ℓ_1 matrix norm for (T'T) is 212.24.
2. ℓ_1 matrix norm for D is 106.13.
3. ℓ_1 matrix norm for F is 106.10.

This study shows that (T'T) can be made diagonally dominant by increasing either N or λ, or both. In this way the error in the solution by the UTSI method with respect to the one using Gaussian elimination may be reduced.

The error due to neglecting off-diagonal elements can be reduced, and in effect, the estimate may be improved

by an iterative scheme (given below). This is particularly desirable when available data length is small.

$$\tilde{\underline{V}} = T\,\tilde{\underline{C}} \ ,$$

or $$T\,\tilde{\underline{C}} = \tilde{\underline{V}} \ ,$$

or $$T'T\,\tilde{\underline{C}} = T'\tilde{\underline{V}} \ ,$$

or $$(D+F)\tilde{\underline{C}} = T'\tilde{\underline{V}} = \underline{Y} \ ,$$

or $$D\,\tilde{\underline{C}} = \underline{Y} - F\,\tilde{\underline{C}} \ ,$$

or $$\tilde{\underline{C}} = D^{-1}\underline{Y} - D^{-1}\,F\,\tilde{\underline{C}} \ . \qquad (58)$$

Now in Eq. (58), let $D^{-1}\underline{Y}$ be the first estimate of $\tilde{\underline{C}}$ and be denoted by $\tilde{\underline{C}}^{(1)}$. The term $(D^{-1}\,F)$ is the first-order error matrix and $(D^{-1}\,F)\tilde{\underline{C}}$ is the error vector. The ℓ_1 norm of $D^{-1}\,F$ and ℓ_2 norm of $D^{-1}\,F\,\tilde{\underline{C}}$ are called NORM1 and NORM2, respectively. The second estimate $\tilde{\underline{C}}^{(2)}$ is then written as (using Eq. (58)):

$$\tilde{\underline{C}}^{(2)} = \tilde{\underline{C}}^{(1)} - D^{-1}\,F\,\tilde{\underline{C}}^{(1)} \ , \qquad (59)$$

and the third becomes

$$\tilde{\underline{C}}^{(3)} = \tilde{\underline{C}}^{(1)} - D^{-1}\,F\,\tilde{\underline{C}}^{(2)} \ .$$

The iterative process takes the general form

$$\tilde{\underline{C}}^{(i+1)} = \tilde{\underline{C}}^{(1)} - D^{-1}\,F\,\tilde{\underline{C}}^{(i)} \ . \qquad (60)$$

The iteration in Eq. (60) always converges if the ℓ_1 matrix norm of $D^{-1}\,F$ is less than one. In actual calculations, iteration converges even if the ℓ_1 norm of $D^{-1}\,F$

is slightly above one. The overall convergence studies indicate that the iteration converges very quickly for NORM1 < 1, and relatively slowly for 1 < NORM1 < 2. The convergence is slower for 2 < NORM1 < 3. For NORM1 < 3, it usually diverges and NORM2 gets larger and larger with increasing iteration. The two typical cases of convergence are presented in Table I.

In Table I, it may be noticed that increased sample size reduces the ℓ_1 norm (NORM1), which in turn helps the iteration to converge faster than the case with sample size = 100. When the iteration converges, the above iteration procedure provides results as good as the Gaussian elimination solution, using less memory than that required in the latter method. Computational time may be of the same order. To guarantee fast convergence, one of the following two conditions should be satisfied:

1. $N \geq 10\ M$ while $\lambda \geq f_{max}$.
2. $N \geq 5\ M$ while $\lambda \geq 3\ f_{max}$.

These are not difficult conditions to satisfy in practice. If, for short record length, one desires to extract maximum information, the Gaussian elimination solution or the iteration techniques are recommended, provided larger storage and computing powers are available.

Table I. Two Typical Cases of Convergence

Iteration Step (i)	Sample Size = 100		Sample Size = 200	
	ℓ_1 Matrix Norm of $D^{-1} F$ (NORM1)	ℓ_2 Matrix Norm of $D^{-1} F \tilde{C}(i)$ (NORM2)	ℓ_1 Norm of $D^{-1} F$ (NORM1)	ℓ_2 Norm of $D^{-1} F \tilde{C}(i)$ (NORM2)
1	2.7271	1.1624	1.9435	0.811
2		0.708		0.346
3		0.489		0.179
4		0.381		0.1044
5		0.326		6.0664
6		0.308		0.0445

Note: The signal used here is the same as given in Eq. (57); $2M = 40$, $\lambda \simeq 10$ samples/sec; $f_{res} = 0.5$ Hz, signal $f_{max} = 9.5$ Hz.

Window Selection for the Present Estimator

Leakage is an unavoidable problem in the estimation of spectra from a finite amount of data. Using no window is equivalent to the use of a boxcar type of window, as explained in Appendix C, and causes considerable leakage through its side lobes. The amount of leakage can always be reduced by proper selection of a window (either in time domain or in frequency domain).

The model in Eq. (29) is rewritten without and with a time window function as

$$\tilde{\underline{V}} = T\ \underline{C} \tag{61}$$

and

$$W\ \tilde{\underline{V}} = T\ \underline{C}_w\ , \tag{62}$$

respectively. Here, W is the discrete data-window and may be considered as the diagonal matrix (N,N) whose elements are $w_1, w_2, \ldots, w_N$, and $0 \leq w_i \leq 1$.

The new estimator without a data window as developed earlier, looks like

$$\tilde{\underline{C}} = D^{-1}\ T'\ \tilde{\underline{V}}\ , \tag{63}$$

where (T'T) = D + F, and F is neglected. Similarly, the above estimator with a data window may be obtained from Eq. (62) as

$$\tilde{\underline{C}}_w = D^{-1}\ T'\ W\ \tilde{\underline{V}}\ . \tag{64}$$

Comparing the estimators $\tilde{\underline{C}}$ and $\tilde{\underline{C}}_W$, one may notice that in Eq. (64), the observations (components of $\tilde{\underline{V}}$) are weighted by a weighting matrix W. In the case of a boxcar window, each element of W is equal to one.

Selection of a particular window depends on an individual's requirement. A trade-off is usually made between the width of the main lobe (of the window function in the frequency domain) and leakage through the side lobes. Typically, a window whose transform has very small side lobes will have a wider main lobe and vice versa. However, one can find some optimized window functions in the literature. These optimized window functions possess more suitable features, but they are usually complicated algebraically, and hence, need considerable time for computation. More information on the selection of a window and a comparison of window features may be found in Refs. [9, 35, 36].

The new estimator developed in this chapter and represented in Eq. (41) will be referred to as the UTSI estimator in this study.

CHAPTER V

EXPERIMENTS WITH THE UNIVERSITY OF TENNESSEE SPACE INSTITUTE ALGORITHM AND THEIR RESULTS

To test the performance of The University of Tennessee Space Institute (UTSI) algorithm, a number of experiments have been conducted. The important ones will be described here and their results will be presented. Wherever possible, the same data have been input to the Discretized Lag Product (Mayo's) algorithm. The results of both algorithms are presented side by side for comparison. In some experiments, spectra computed using the Fast Fourier Transform (FFT) are also included for comparison. For testing the algorithms, input data were generated in two ways: using (1) digital simulations and (2) actual laboratory measurements.

Three computer simulations were performed to evaluate the algorithm under controlled conditions. The first was for discrete sinusoidal components at selected frequencies. The second demonstrated algorithm performance for detection of a signal imbedded in noise (white). The third is an aliasing study using discrete sinusoids. A description of the simulations, corresponding results, and discussions follow.

Study with Simulated Band-Limited White Noise

As shown in Fig. 2, using the IBM scientific subroutine RANDU, random numbers are generated. Each Gaussian random variable has been generated by summing 12 random numbers. The FFT is applied to the Gaussian random variables and white noise in the frequency domain is obtained. Now a "brick-wall"[1] filter is used to make it band-limited. The band-limited white noise is transformed to the time domain through an inverse FFT. But this time domain signal is equi-spaced. Using a random number generator along with an approximation taken in Ref. [37], Poisson sampling instants are generated. Through linear interpolation the sample values at Poisson sampling instants are found from the equi-spaced ones. These randomly spaced samples are the input to the algorithms. The final spectral estimate is the average over 30 such sets of data. It is to be noted that for each run, the starting seed number of the pseudorandom number generator should be different from the others. The output spectra obtained from the algorithm under discussion and from the discretized lag method are plotted in Figs. 3 and 4 for low pass and bandpass data, respectively. Figure 3 is the case where a brick-wall low pass (0 to 100 Hz) filter was used.

---

[1]The technique of brick-wall filtering is discussed in Appendix D.

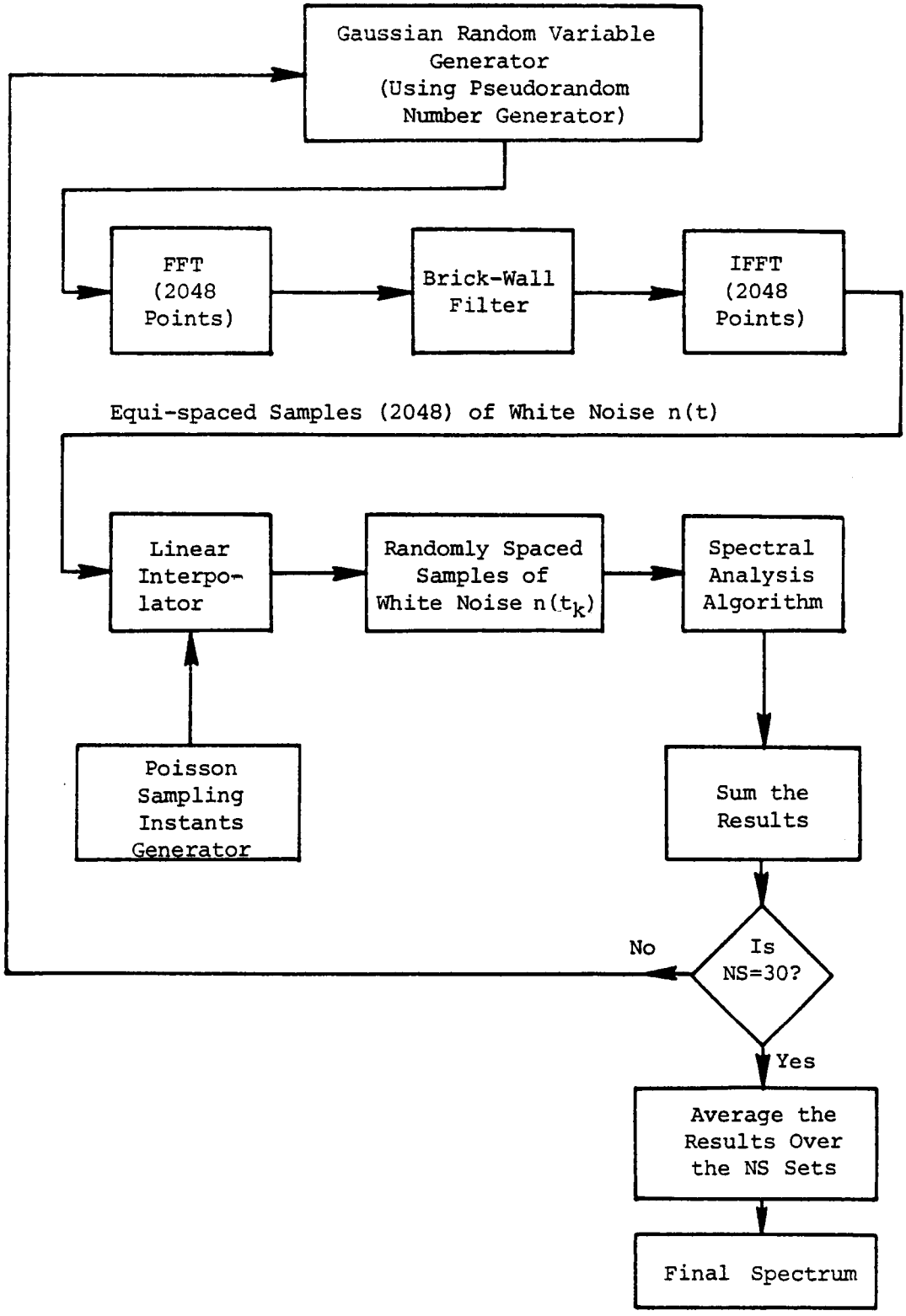


Figure 2. Block schematic of the band limited white noise experiment through simulation.

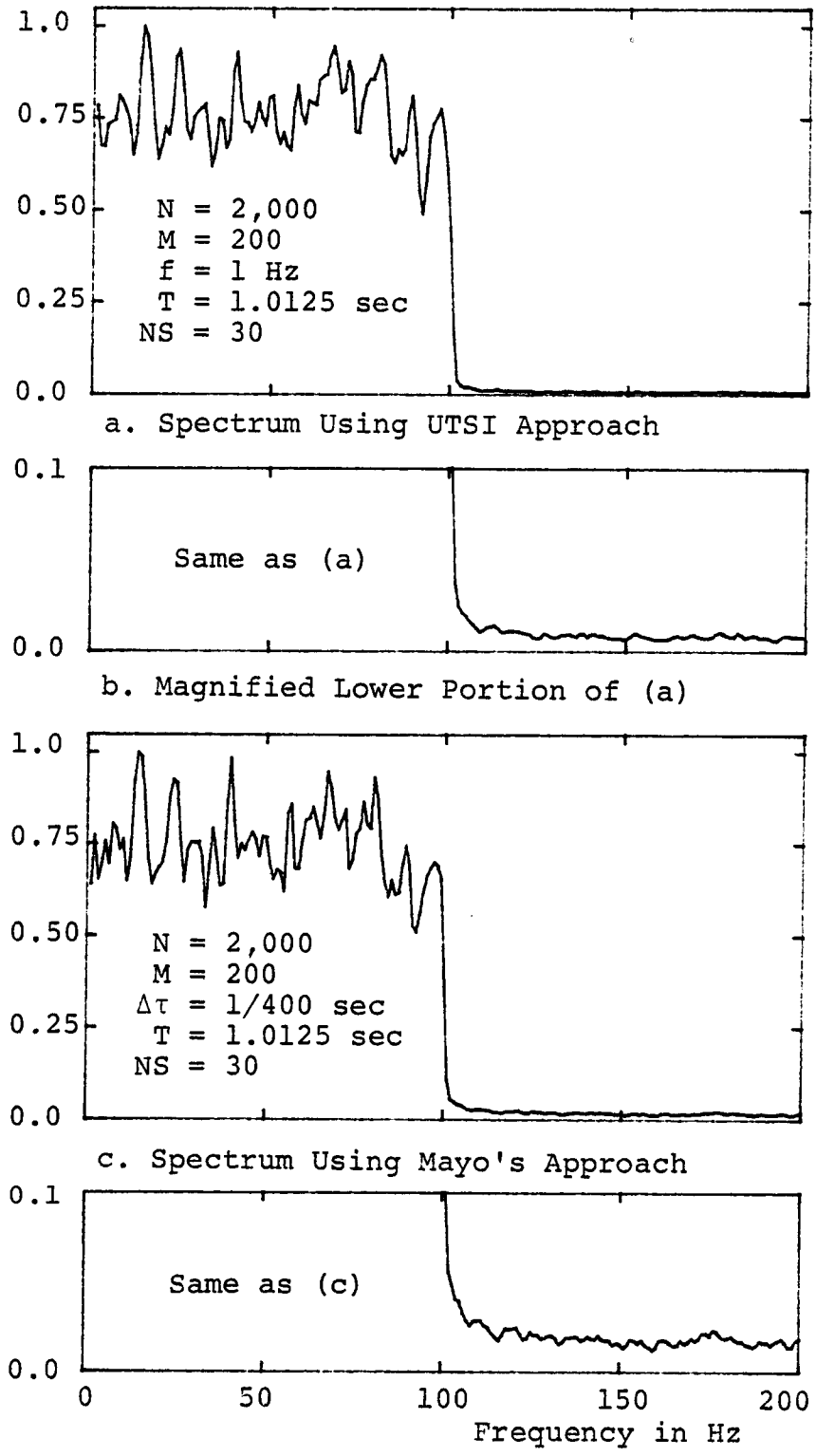


Figure 3. Low pass filter transfer function reconstruction using The University of Tennessee Space Institute and Mayo algorithms.

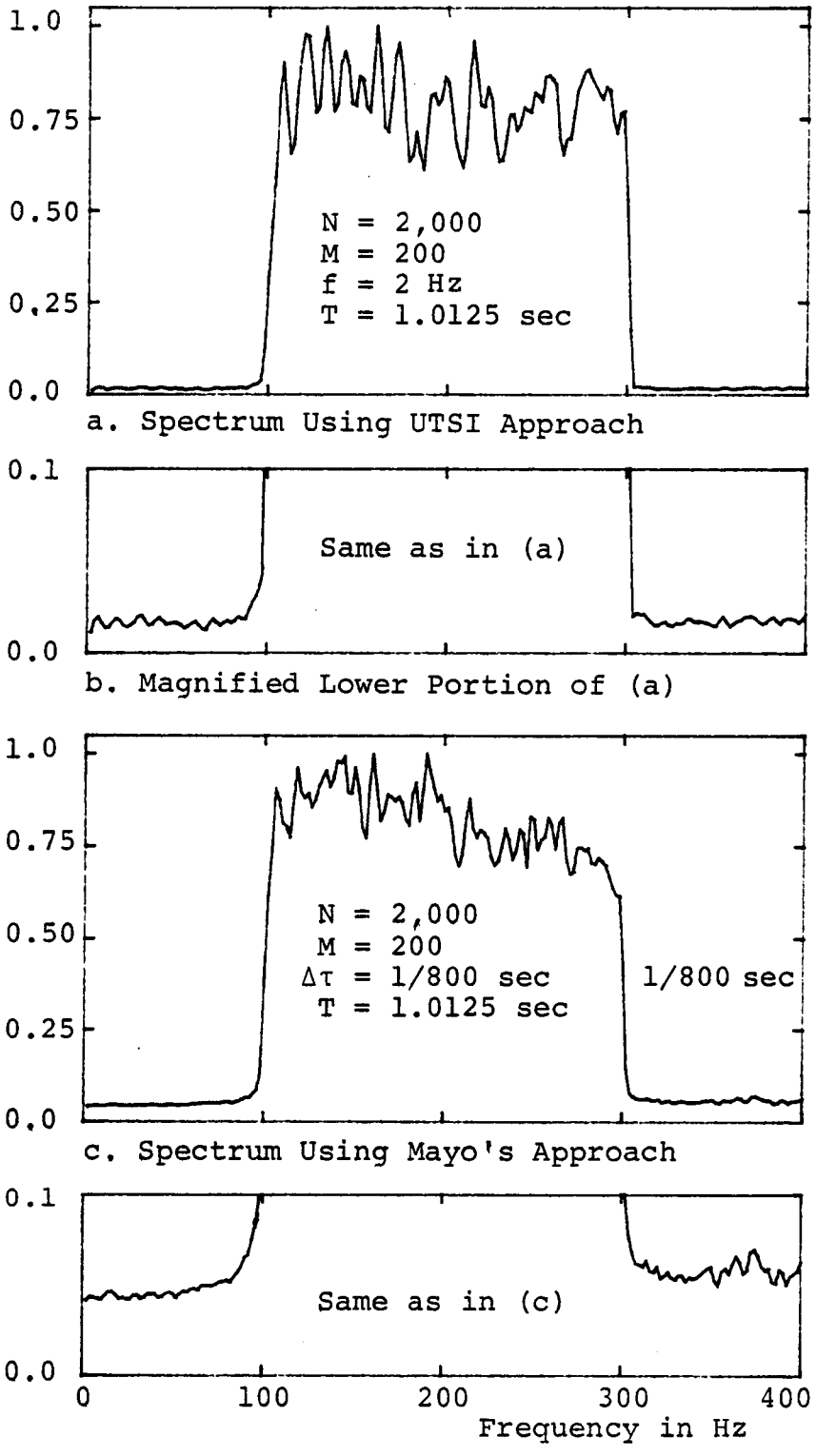


Figure 4. Bandpass filter transfer function reconstruction using The University of Tennessee Space Institute and Mayo algorithms.

The filter was a bandpass (pass band of 200 Hz with a center frequency f_c = 200 Hz) type for Fig. 4. For the UTSI algorithm, a Hanning spectral window is used, and in the case of the discretized lag product method, a triangular data window is used on the autocovariance estimate for smoothing.

The above experiment shows the response of the algorithm for band-limited white noise. The response of the two algorithms is similar in the pass band. The stop band response of the UTSI method indicates that it does not generate significant leakage. The two algorithms produce comparable results, and considering stop band rejection, the UTSI method is even better.

Noise Rejection Study

In this experiment the noise rejection capability of the algorithms has been tested through computer simulation. As shown in Fig. 5, a single frequency signal and white noise are generated independently and then added to obtain a randomly spaced noisy signal with a specified signal-to-noise ratio (S/N), but a S/N = 0.1 was chosen for the simulation. A similar equi-spaced noisy signal with S/N = 0.1 was generated as shown in Fig. 6. Part of this equi-spaced time signal (f = 10 Hz) is plotted in Fig. 7. The normalized autocovariance function of the above equi-spaced record is computed and part of it is plotted in Fig. 8. The randomly sampled signal with record length of

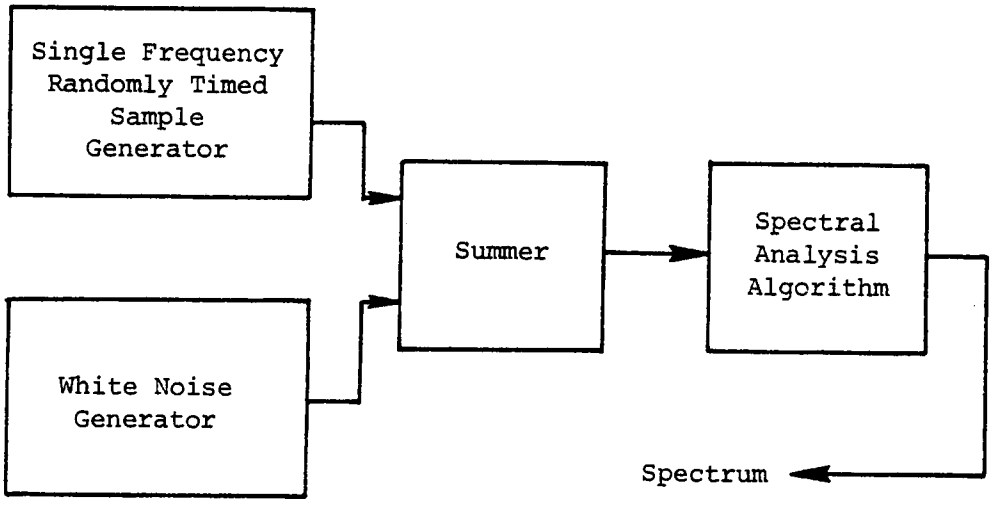


Figure 5. Block schematic of the noise rejection test of The University of Tennessee Space Institute and Mayo algorithms through simulated data.

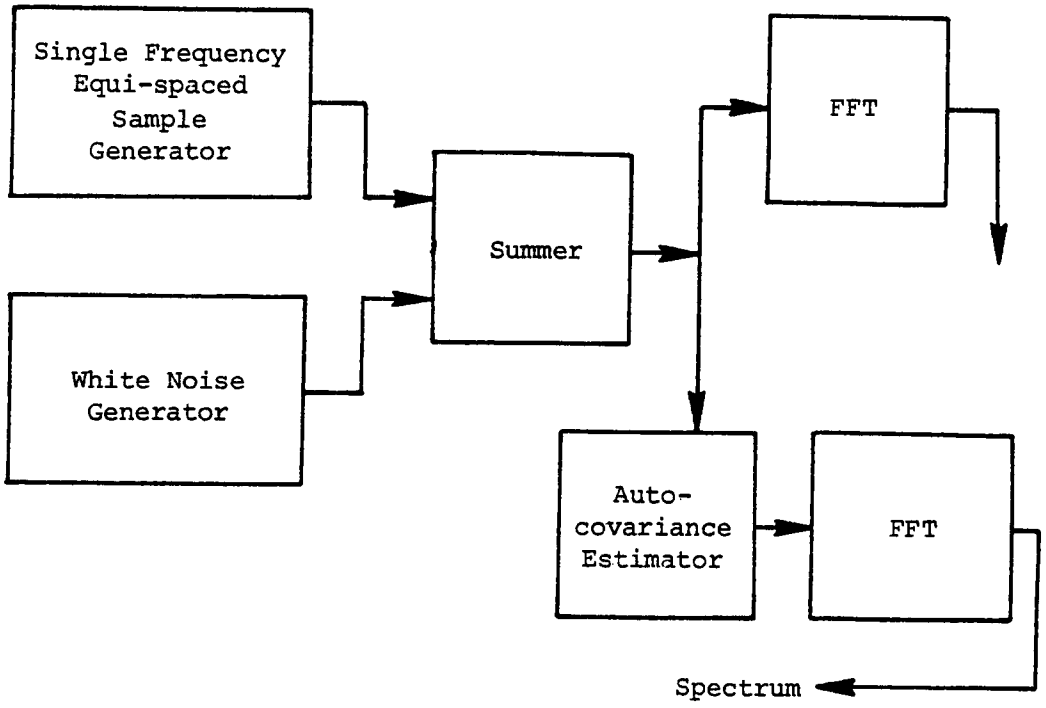


Figure 6. Block schematic of noise rejection test of FFT algorithm with simulated data.

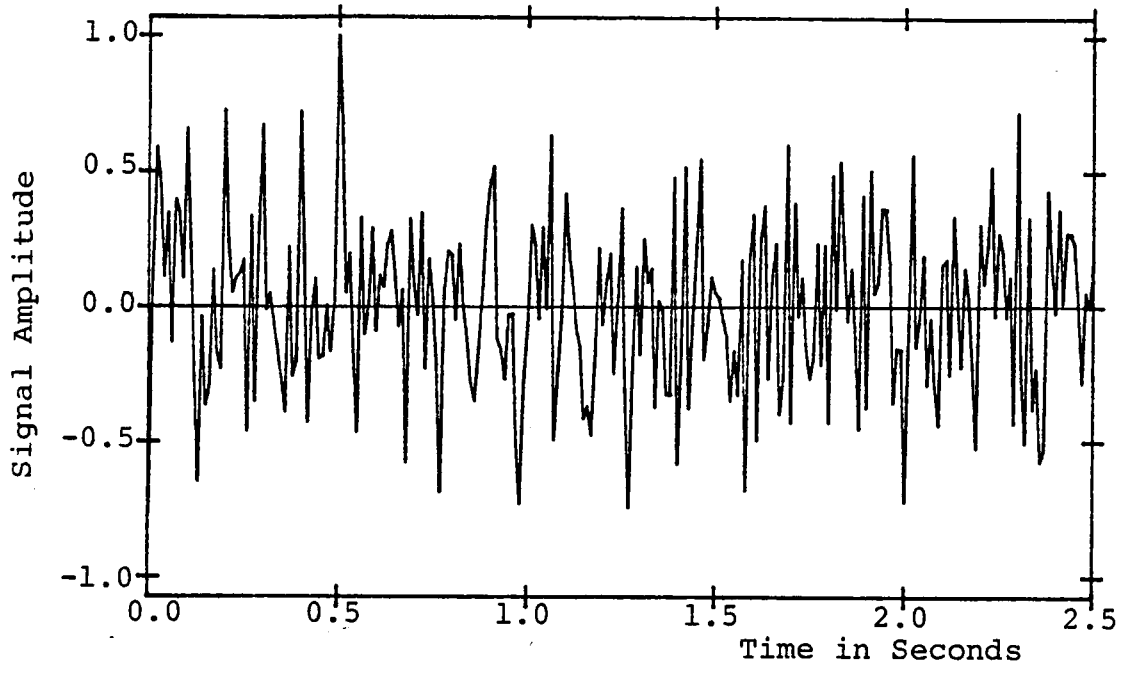


Figure 7. Part of time history of single frequency signal plus noise with S/N = 0.1

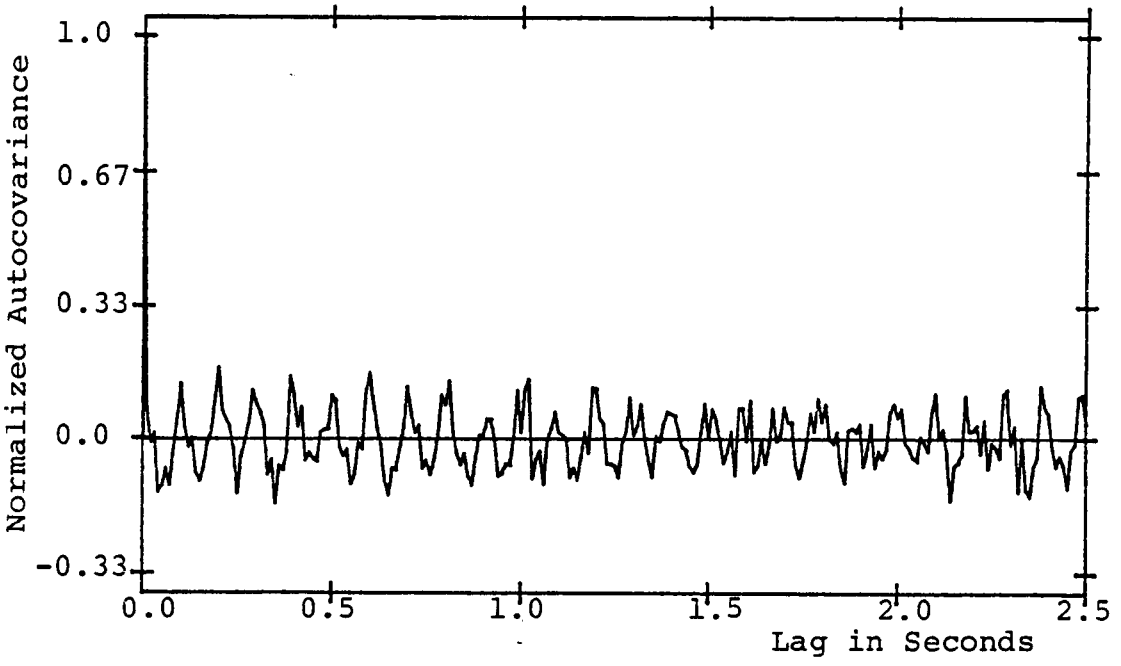


Figure 8. Part of autocovariance of time history of Fig. 7.

N = 500 and mean sampling rate of λ = 100 samples/sec was input to the UTSI algorithm and Mayo's algorithm. The results are plotted in Fig. 9. The spectrum of the equi-spaced data with N = 512 and sampling interval = 0.01 is computed using the FFT. Figures 10a and 10b are the resulting periodograms. The power spectrum is also computed from the FFT of the autocovariance function for the equi-spaced noisy signal. The result is shown in Figs. 10c and 10d.

Comparing the noise rejection of the three algorithms, the UTSI algorithm is as good as the FFT, and is better than Mayo's indirect approach.

Aliasing Study with The University of Tennessee Space Institute Algorithm

An experimental study using computer simulation was performed to determine the aliasing properties of the algorithm. A randomly sampled record was generated from the following signal:

$$x(t) = \sin(2\pi f_1 t + \pi/4) + \sin(2\pi f_2 t + \pi/4) + 2\sin(2\pi f_3 t + \pi/4) , \quad (65)$$

where

$$f_1 = 10 \text{ Hz}, \; f_2 = 20 \text{ Hz}, \text{ and } f_3 = 40 \text{ Hz} .$$

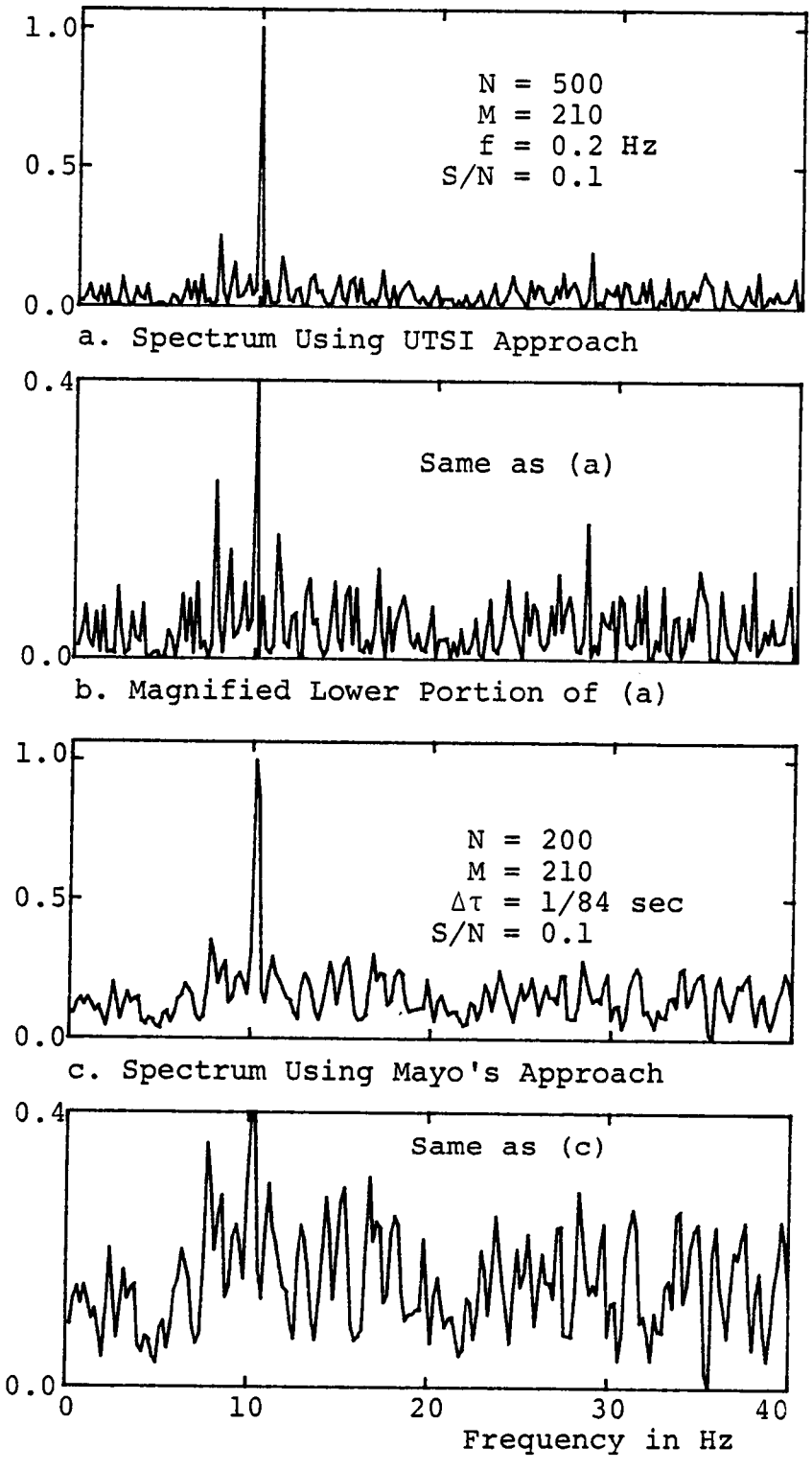


Figure 9. Spectrums of time history shown in Fig. 7 with The University of Tennessee Space Institute approach and Mayo's approach.

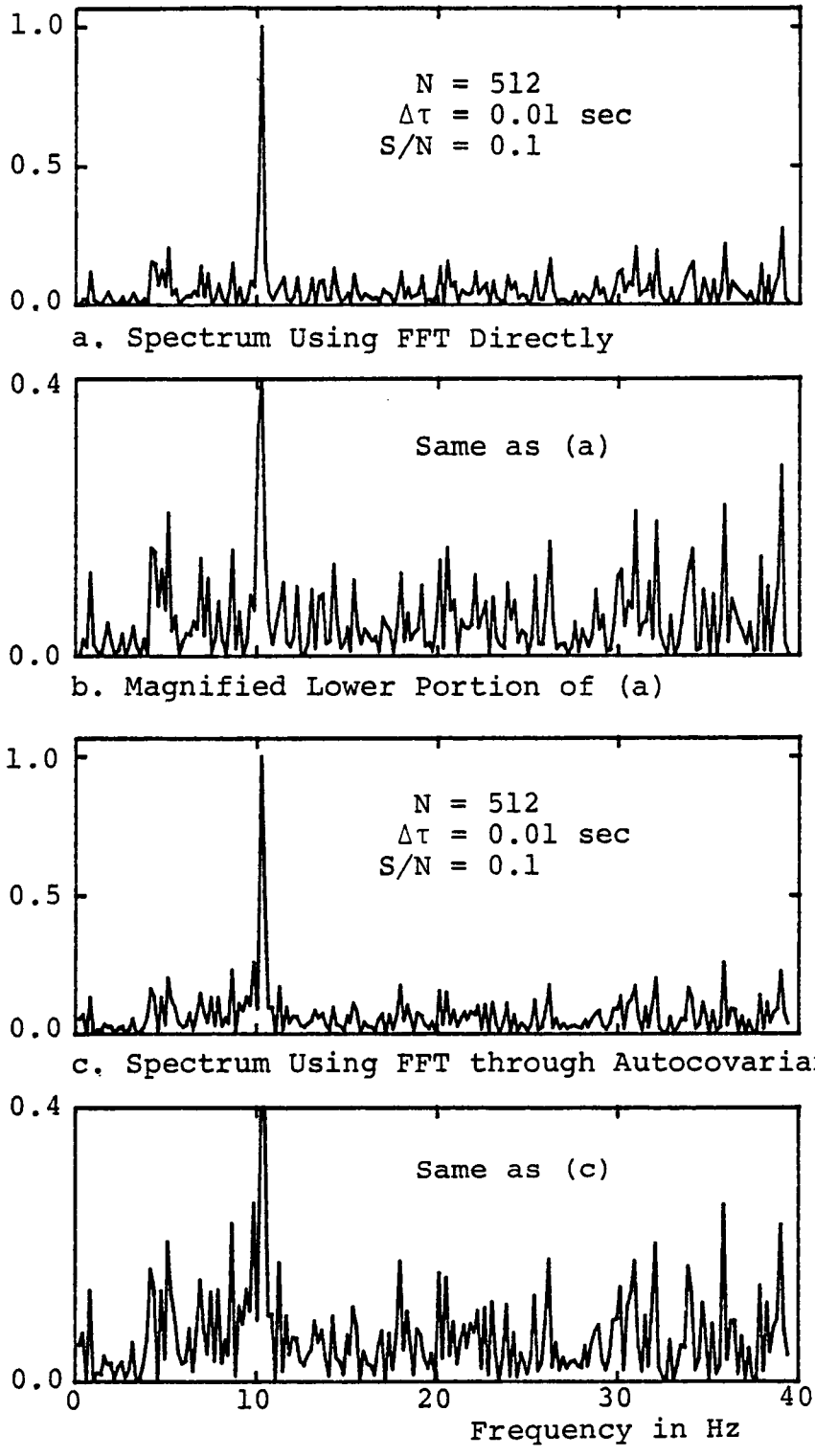


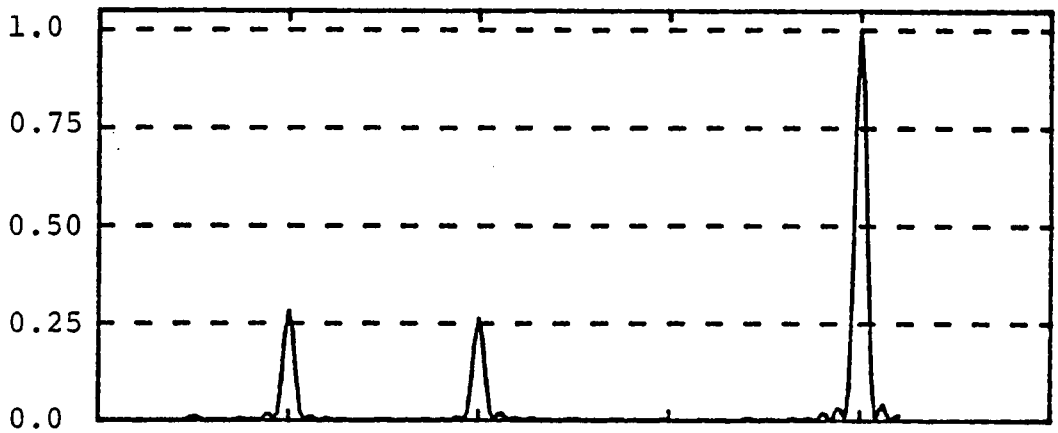
Figure 10. Spectra of time history shown in Fig. 7 using FFT through periodogram and autocovariance approaches, respectively.

The Nyquist rate or Nyquist sampling frequency f_N for the signal x(t) is equal to 80 samples/sec (considering f_{max} = 40 Hz). The sampling rate in the case of random sampling is known as the mean sampling rate λ ($\lambda \simeq \frac{N}{T}$, where N is the total number of samples taken over time T sec). Six typical cases are given below:

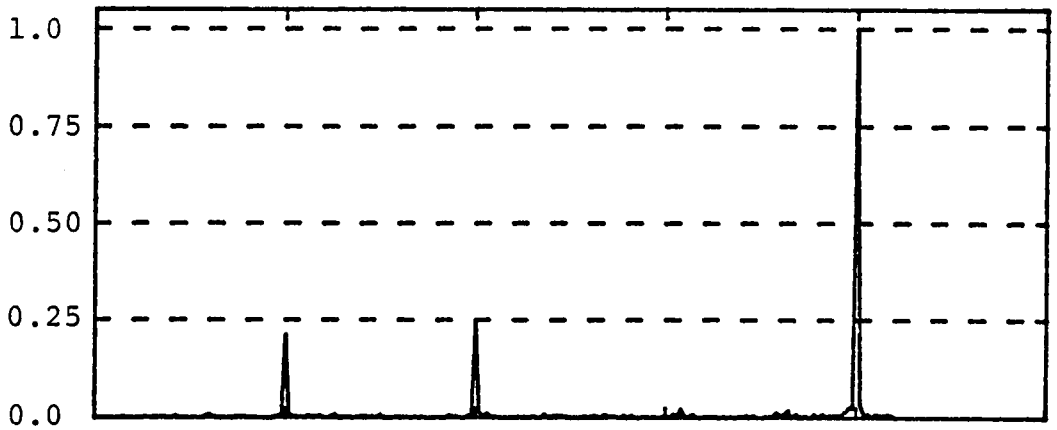
1. N = 500 samples ; λ = 390 samples/sec.
2. N = 500 samples ; λ = 79 samples/sec.
3. N = 500 samples ; λ = 7.89 samples/sec.
4. N = 500 samples ; λ = 5.81 samples/sec.
5. N = 2,000 samples ; λ = 7.89 samples/sec.
6. N = 5,000 samples ; λ = 7.89 samples/sec.

Characteristics of aliasing are studied by decreasing λ while holding the sample size constant in the first four cases. Comparing cases (3), (5), and (6), the effect of increased sample size on the aliasing case is shown. The spectra of the above six cases are plotted in Fig. 11. It is to be noted from Eq. (65) that the amplitudes of the first two frequencies are the same and are equal to 1.0, while the amplitude of the third one is 2.0, so the power spectral density at f_3 will be four times the value at f_1 and f_2.

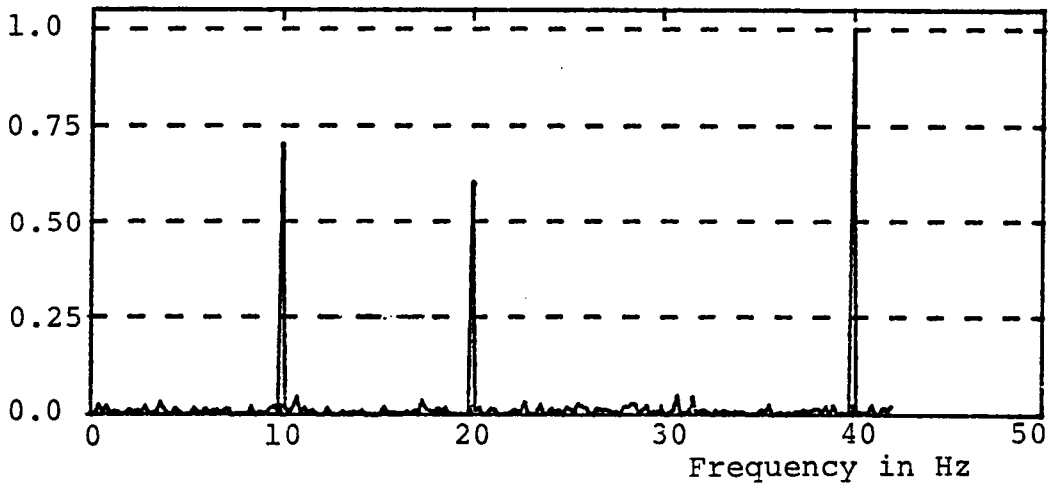
These results show that the close relation between Nyquist rate sampling and an alias-free spectrum in the case of equi-spaced data may be violated in the random sampling scheme. The UTSI approach will provide alias-free



a. N = 600, M = 210, $\lambda = 5\ f_N$

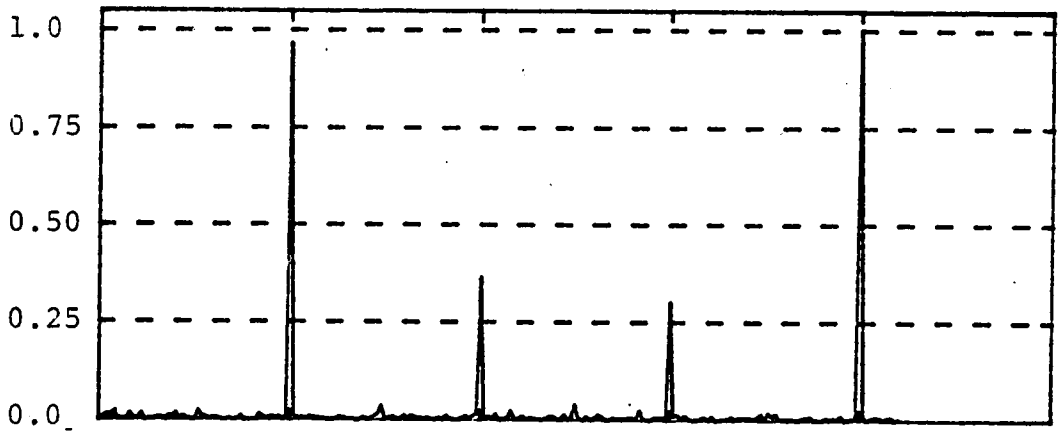


b. N = 600, M = 210, $\lambda = f_N$

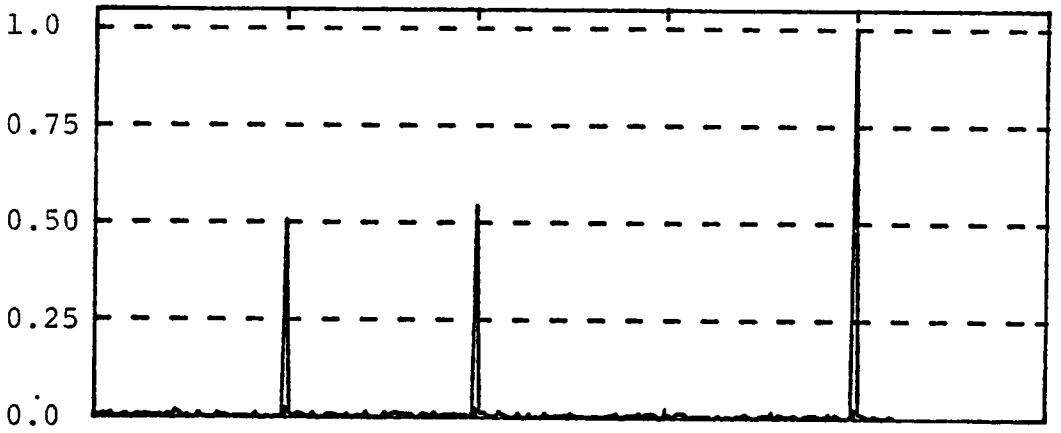


c. N = 600, M = 210, $\lambda = 0.1\ f_N$

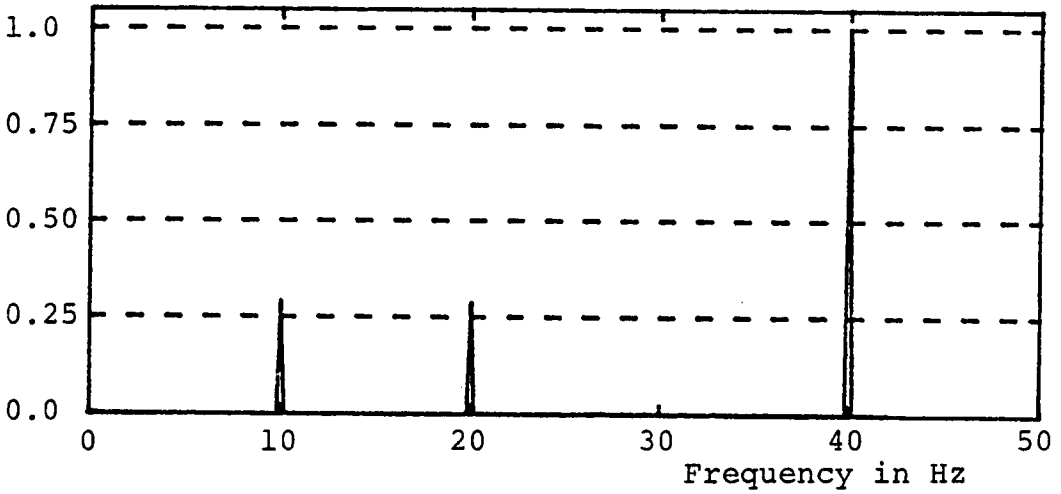
Figure 11. Spectra for aliasing comparison using The University of Tennessee Space Institute algorithm.



d. N = 600, M = 210, $\lambda = 0.063\ f_N$



e. N = 2,000, M = 210, $\lambda = 0.1\ f_N$



f. N = 5,000, M = 210, $\lambda = 0.1\ f_N$

Figure 11. (Continued)

spectra at sub-Nyquist rate if the record length is long enough.

Tests with Data Obtained from Laboratory Setup

Study with Single Frequency Signal

In this test, two kinds of data are generated: a single frequency sinusoidal signal and band-limited white noise. The experimental setup is shown in Fig. 12. A loudspeaker is kept near the fringe area of the differential scatter laser velocimeter. The loudspeaker is excited with a sine wave signal frequency of 100 Hz, so the data obtained are fluctuating velocity data with almost zero mean velocity. One of the typical velocity time histories is plotted in Fig. 13. Note that the data were originally discrete in nature. The data were plotted with pen down wherever a bust of observations was available. To find the spectrum of this record, the UTSI and Mayo algorithms have both been applied to the data. The results are plotted in Figs. 14a and 14b, respectively. The same data were converted to equi-spaced samples through linear interpolation and were input to the FFT algorithm. The corresponding result is shown in Fig. 14c.

The results with the data set obtained from the experimental setup in Fig. 12 verify the ability of the UTSI algorithm to predict frequency spectra in a laboratory

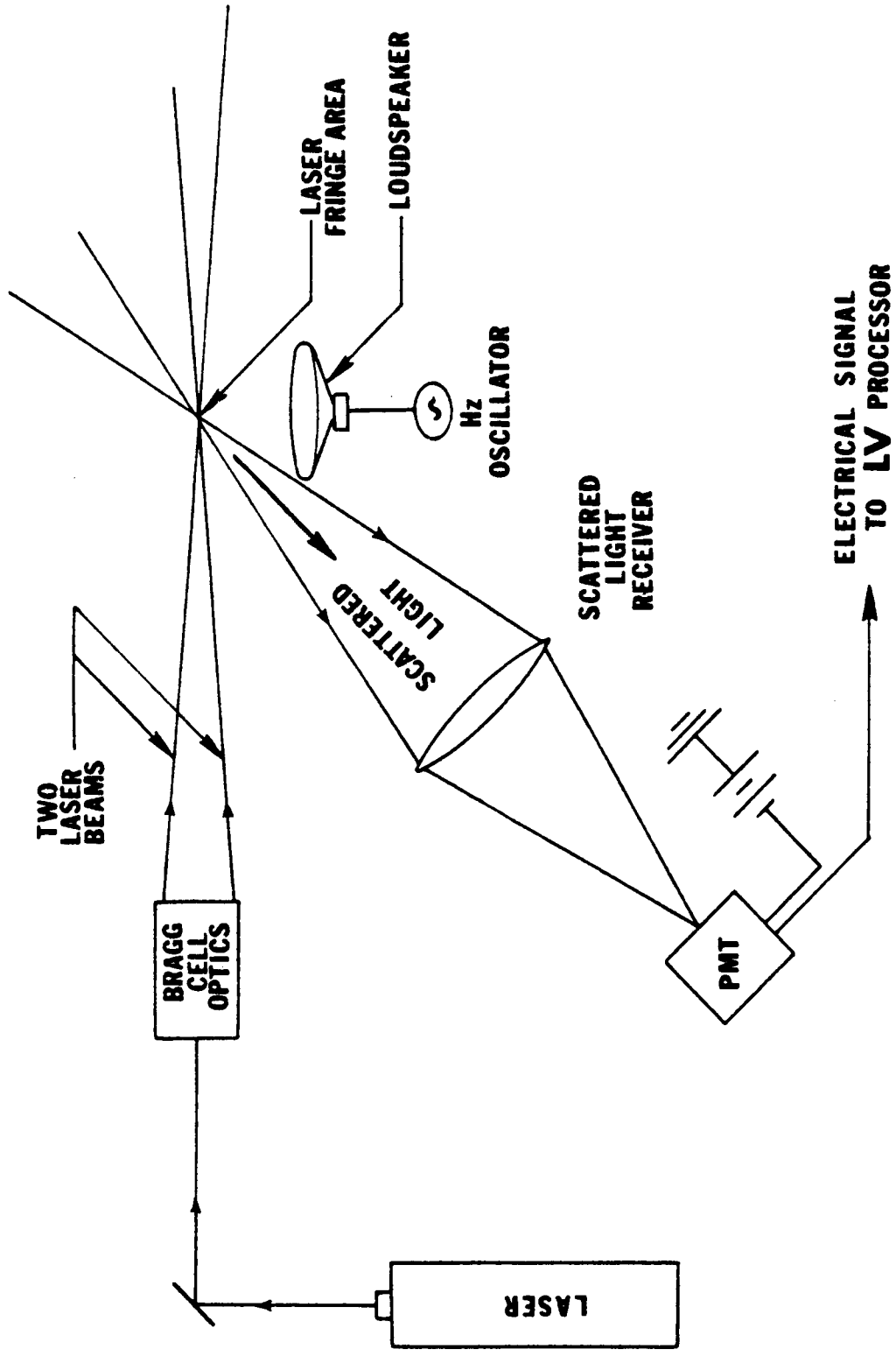


Figure 12. Typical experimental setup.

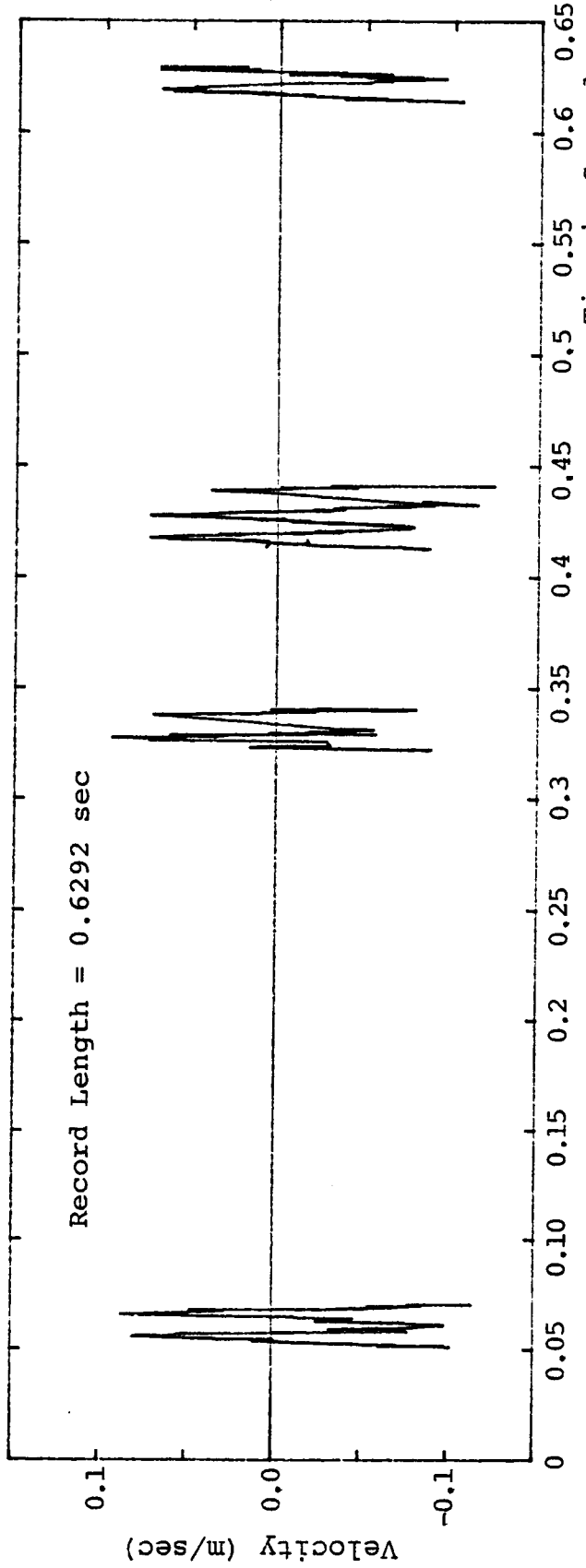
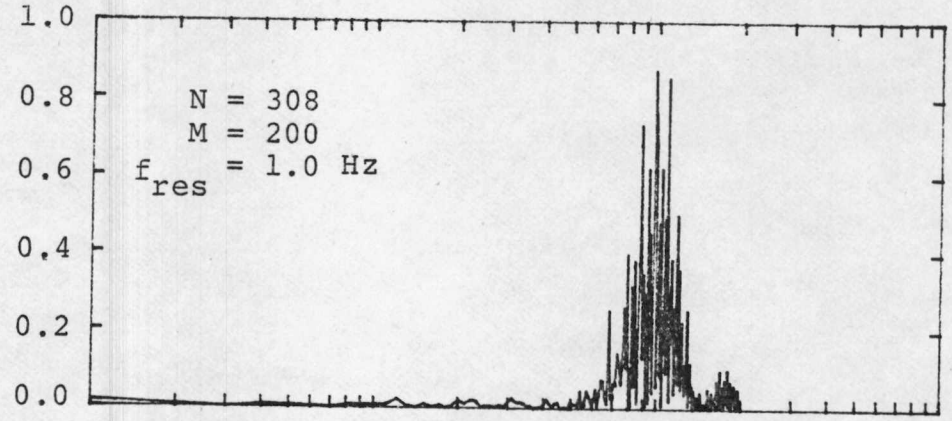
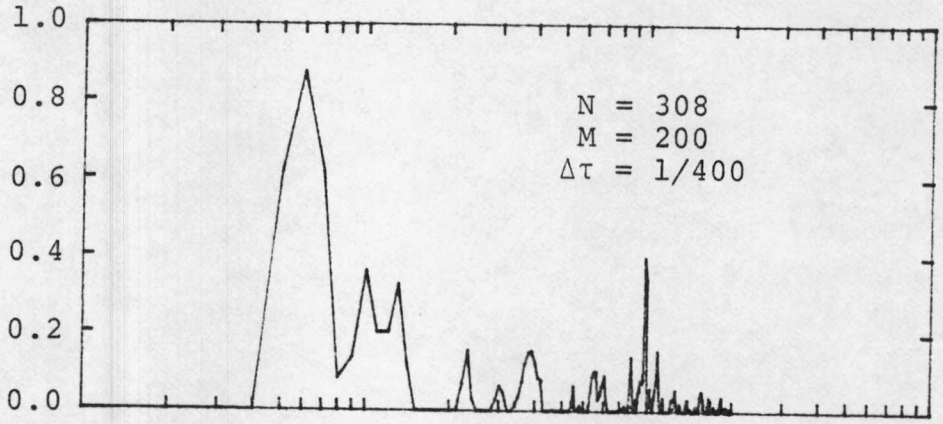


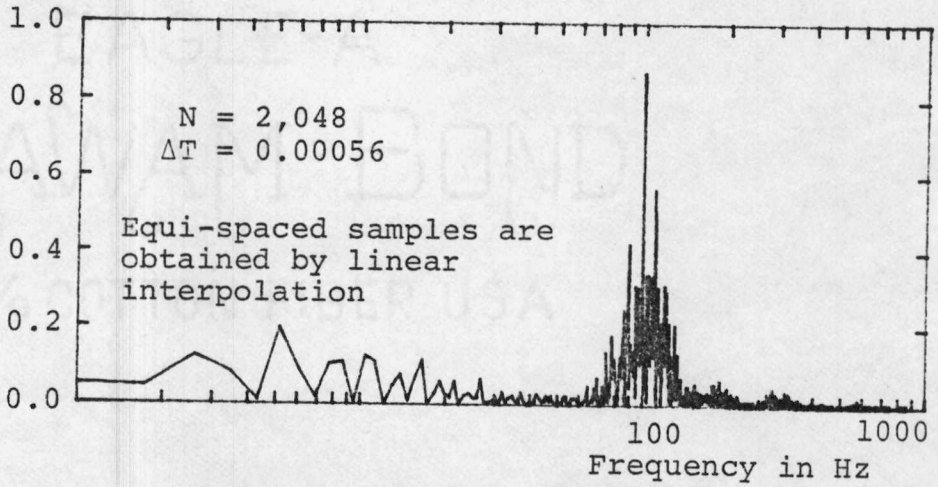
Figure 13. Typical plot of one of the time histories acquired in the experiment shown in Fig. 12.



a. Spectrum Using UTSI Algorithm



b. Spectrum Using Mayo's Algorithm



c. Spectrum Using FFT Algorithm

Figure 14. Spectra of the record shown in Fig. 13 using three different algorithms.

environment. While the output from the FFT algorithm confirms the result obtained using the UTSI algorithm, Mayo's algorithm produces a different result. It generates some artificial low frequencies in the spectrum which are not expected from the data.

Study with Band-Limited White Noise Using Experimental Data

To acquire data containing band-limited white noise, the output of a standard laboratory noise generator is passed through a Butterworth low-pass active filter, as shown in Fig. 15. Random sampling instants are generated with the computer. The band-limited noise is sampled with a mean sampling rate of λ = 5,000 samples/sec and digitized by a 12-bit analog-digital converter. The digitized data are used as input to the UTSI and Mayo algorithms and the corresponding results are presented in Figs. 16a and 16b. The same noise was also sampled in equi-spaced fashion keeping the sampling frequency f_s = 5 KHz and was input to the FFT; the corresponding result is plotted in Fig. 16c. Again, with the UTSI and FFT methods, a Hanning spectral window was used for smoothing, whereas, for Mayo's algorithm, a triangular data window was used.

The spectra obtained from the band-limited white noise data set are comparable for all three methods.

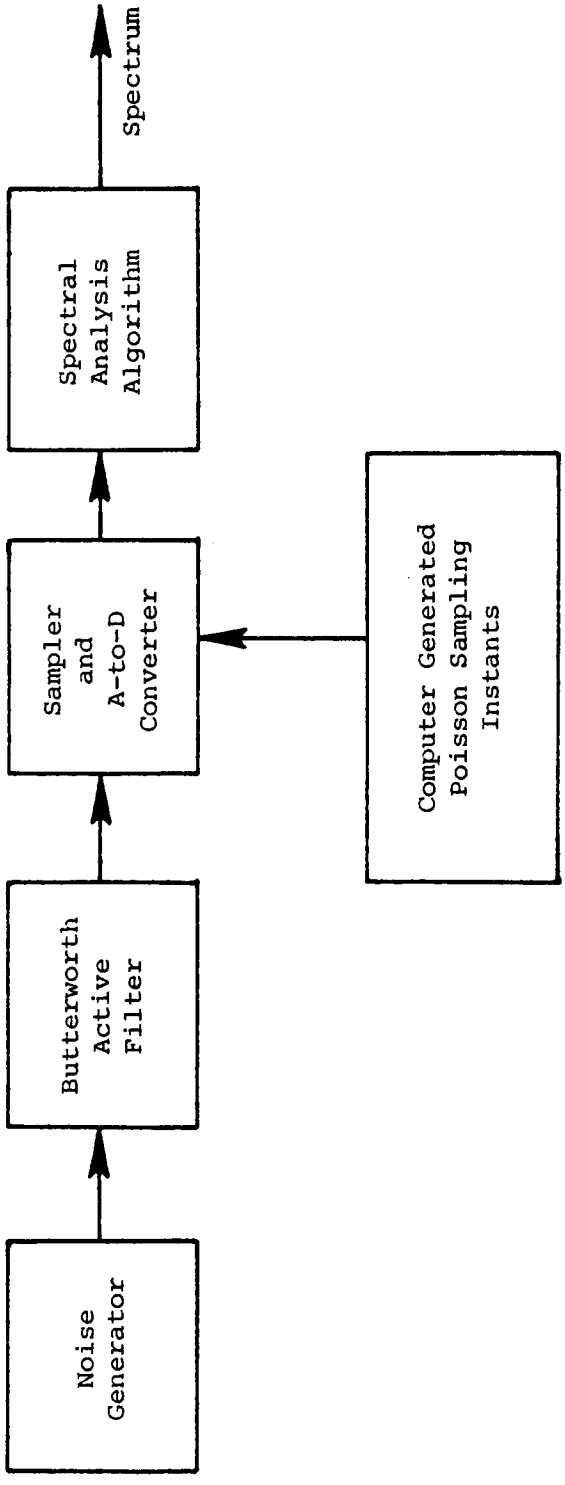
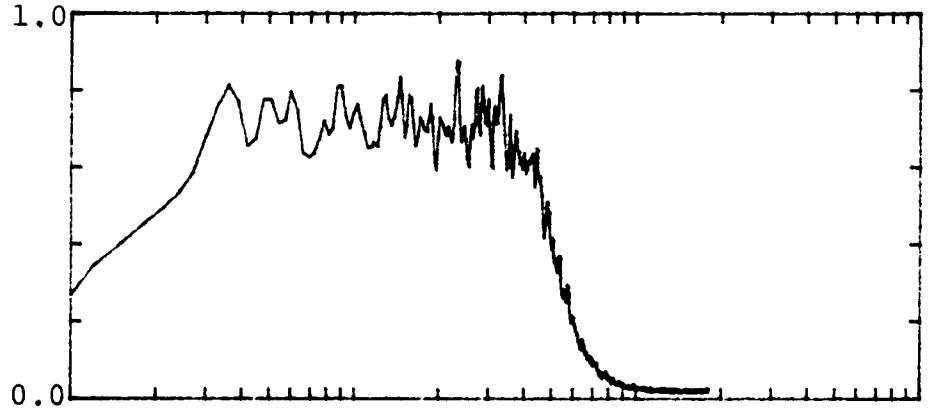
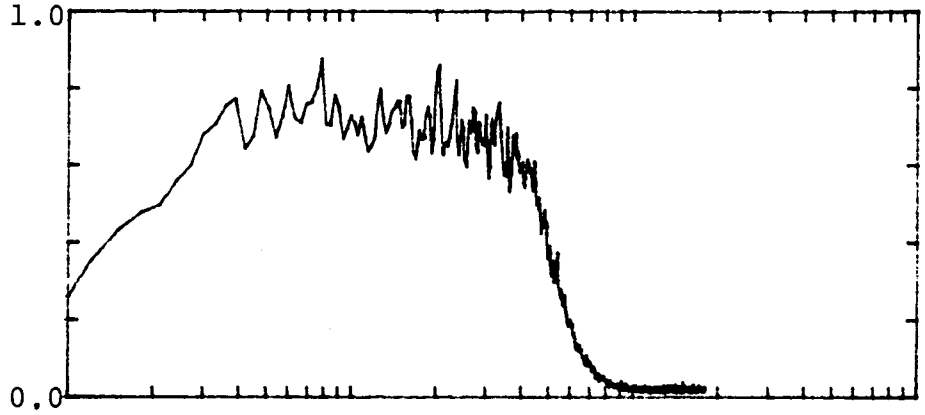


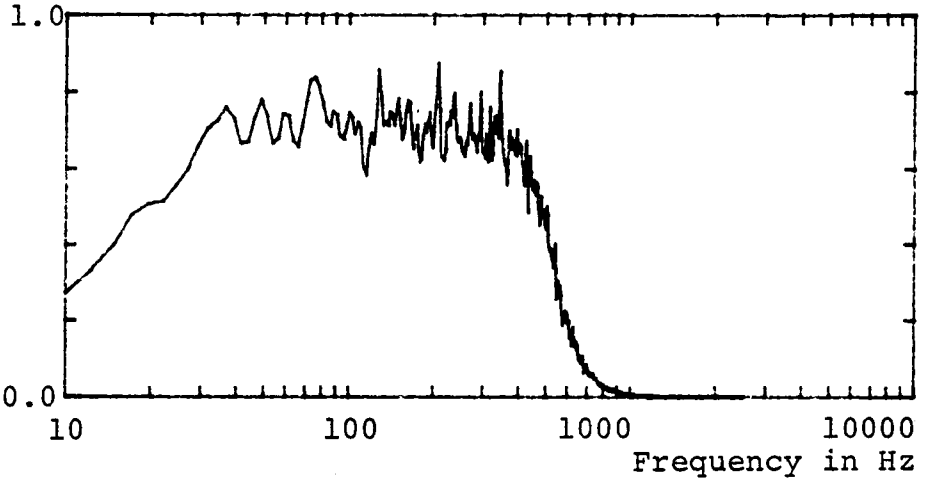
Figure 15. Block schematic of the experimental setup to acquire data for band limited white noise experiment.



a. Spectrum Using UTSI Algorithm



b. Spectrum Using Mayo's Algorithm



c. Spectrum Using FFT Algorithm

Figure 16. Spectra of a record obtained in an experiment shown in Fig. 15 using three different algorithms.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

The experiments reported in Chapter V and their results demonstrate the potential of the UTSI algorithm as a practical spectrum estimator of stochastic processes. Its noise rejection ability and nonaliasing characteristic are superior to other algorithms available in this area of random sampling. By a proper choice of parameters (e.g., data length), it neither generates artificial frequencies nor distorts existing frequencies in a signal. Computationally, it is as fast as the other algorithms; additionally, its algebraic simplicity makes it easily implementable in micro-processor based systems.

The above discussion shows the specific features of the UTSI algorithm. One might ask about the procedure to be taken to estimate the spectrum of an unknown process using the UTSI algorithm. The following procedural steps are recommended.

Two cases may arise: (1) Case A, data set (time versus signal amplitude) already available, and (2) Case B, an experiment to be performed to find the spectrum of an unknown process.

1. Case A:
    a. Find the mean sampling rate $\lambda \simeq \frac{N}{T}$, where T is the record length, and N is the number of data points in the record.
    b. Select f_{max} such that $f_{max} \leq \lambda$, where f_{max} is the high end of the frequency spectrum.
    c. Choose the number of spectral components, $M \leq \frac{1}{5} N$.
    d. Frequency resolution, $f_{res} = f_{max/M}$.

In this case, f_{max} may be increased provided increased data length ($N > 5 M$) is available. For example, $f_{max} \simeq 2 \lambda$ may be taken if at least $N \geq 10 M$ is available. These tentative figures are the result of the aliasing study.

2. Case B:
    a. Expect or select f_{max} for the spectrum.
    b. Choose $\lambda \geq f_{max}$.
    c. Select desired f_{res}.
    d. Get the number of spectral components, $M \simeq \frac{f_{max}}{f_{res}}$.
    e. Find N such that $N \geq 5 M$.

The instrument dead time of the data acquisition system may impose a limitation on the selection of a high speed data rate to measure high frequency. This limitation may be practically circumvented by collecting increased data length accordingly.

Based on the experimental results and discussions presented in Chapter V, a number of conclusions may be drawn:

1. The UTSI approach is a potentially valuable method for spectrum analysis of randomly sampled data.
2. Further analytical study of aliasing with the UTSI method is needed to determine the optimal relationship between f_{max} and N. This would make it possible to exploit the alias-free characteristics of the algorithm.
3. Instrument dead time is a barrier to the acceptance of high sampling rate data for turbulent spectral measurements with an LDV. Using the UTSI approach, f_{max} can be taken greater than λ if a larger data length is available.
4. For a short record length a Gaussian elimination solution is recommended to extract the maximum information.
5. In conducting the above investigation it was the author's experience that noise spikes in the input data cause problems with spectral analysis for any algorithm. The noise spikes become easily detectable when the time history is plotted. However, such a procedure is time consuming for a large data set. These spikes

should be suppressed at their origins as much as possible. In some cases, large amplitude spikes may be detected and discarded using appropriate preprocessing algorithms prior to spectral analysis.

6. It is noted that computation of the sine and cosine functions accounts for approximately three-fourths of the computing time needed by the UTSI algorithm. A method of replacing the sine and cosine functions by the equivalent periodic rectangular waveforms, mentioned in Ref. [30], might be implemented to speed up the computation.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Nyquist, H. "Certain Topics in Telegraph Transmission Theory," Trans. AIEE, 47:617-644, April, 1928.

2. Shannon, C. E. "Communication in the Presence of Noise," Proc. IRE, 37:10-21, January, 1949.

3. Cooley, J. W., and J. W. Tukey. "An Algorithm for Machine Calculation of Complex Fourier Series," Math Comp. 19:297-301, April, 1965.

4. Bendat, J. S., and A. G. Piersol. Random Data Analysis and Measurement Procedures. New York: John Wiley and Sons, Inc., 1971.

5. Brigham, E. O. The Fast Fourier Transform. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1974.

6. Bracewell, Ronald N. The Fourier Transform and Its Applications. New York: McGraw-Hill Book Co., Inc., 1973.

7. Childers, D., and A. Durling. Digital Filtering and Signal Processing. New York: West Publishing Co., Inc., 1975.

8. Oppenheim, Alan V., and R. Schafer. Digital Signal Processing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.

9. Blackman, R. B., and J. W. Tukey. The Measurement of Power Spectra from the Point of View of Communications Engineering. New York: Dover Publications, Inc., 1953.

10. Shapiro, H. S., and R. A. Silverman. "Alias-Free Sampling of Random Noise," J. Soc. Indust. Appl. Math., 8(No. 2):225-248, June, 1960.

11. Beutler, F. J., and O. A. Z. Leneman. "The Theory of Point Process," Acta Mathematica, Budapest, Hungary, 159-197, 1966.

12. Beutler, F. J. "Alias-Free Randomly Timed Sampling of Stochastic Processes," IEEE Trans. Information Theory, IT-16:147-152, March, 1970.

13. Masry, E. "Random Sampling and Reconstruction of Spectra," Information and Control, 19:275-288, 1971.

14. Papoulis, A. Probability, Random Variables and Stochastic Processes. New York: McGraw-Hill Book Co., Inc., 1965.

15. Mayo, W. T., M. T. Shay, and S. Riter. "Digital Estimation of Turbulence Power Spectra from Burst Counters," Proceedings of the Second International Workshop on Laser Velocimetry. W. Lafayette, Indiana: Purdue University, 1974. Pp. 16-24.

16. Shay, M. T. "Digital Estimation of Auto-Covariance Functions and Power Spectra from Randomly Sampled Data Using a Lag Product Technique." Ph.D. dissertation, Texas A & M University, College Station, 1976. (Available from University Microfilm International, Ann Arbor, Michigan, No. 77-2666.)

17. Scott, P. F. "Distortion and Estimation of the Auto-Correlation Function and Spectrum of a Randomly Sampled Signal." Ph.D. dissertation, Rensselaer Polytechnic Institute, Troy, New York, 1976. (Available from University Microfilm International, Ann Arbor, Michigan, No. 76-27209.)

18. Erdmann, J. C. Lecture notes and theoretical supplement presented at a short course on "Theory and Application of Laser Velocimetry," The University of Tennessee Space Institute, Tullahoma, Tennessee, June, 1980.

19. Thompson, Rory O. R. Y. "Spectral Estimation from Irregularly Spaced Data," IEEE Trans. on Geoscience Electronics, GE-9(No. 2):107-110, April, 1971.

20. Adegbola, M. O. "Alias-Free Spectral Estimation of Stochastic Processes." Ph.D. dissertation, California Institute of Technology, Pasadena, 1971. (Available from University Microfilm International, Ann Arbor, Michigan, No. 72-467.)

21. Gaster, M., and J. B. Roberts. "Spectral Analysis of Randomly Sampled Signals," J. Inst. Math. Appl., 15:195-216, 1975.

22. Masry, E., and M. C. Lui. "A Consistent Estimate of the Spectrum by Random Sampling of the Time Series," SIAM J. Appl. Math., 28(No. 4):793-810, June, 1975.

23. Masry, E., and M. C. Lui. "Discrete-Time Spectral Estimation of Continuous Parameter Processes--A New Consistent Estimate," IEEE Trans. Information Theory, IT-22(No. 3):298-311, May, 1976.

24. Jones, R. H. "Spectral Analysis with Regularly Missed Observations," Annals of Mathematical Statistics, 33:455-461, June, 1962.

25. Jones, R. H. "Spectrum Estimation with Unequally Spaced Observations." Paper No. AFOSR-70-2747TR presented at the Proceedings of the Kyoto International Conference on Circuit and System Theory, Kyoto, Japan, September 9-11, 1970.

26. Jones, R. H. "Spectrum Estimation with Missing Observations," Annals of Institute of Statistical Mathematics, 23:287-298, 1971.

27. Gaster, M., and J. B. Roberts. "The Spectral Analysis of Randomly Sampled Records by a Direct Transform," Proc. R. Soc. London, 354:27-58, 1977.

28. Beuther, Paul D. "Autoregressive Spectral Estimation," Proceedings of the Dynamic Flow Conference on Dynamic Measurements in Unsteady Flows, 1978. Skovlunde, Denmark: Dynamic Flow Conference, 1978. Pp. 794-796.

29. Smith, D. M., and D. M. Meadows. "Power Spectra from Random-Time Samples for Turbulence Measurements with a Laser Velocimeter," Proceedings of the Second International Workshop on Laser Velocimetry. W. Lafayette, Indiana: Purdue University, 1974. Pp. 27-44.

30. Roberts, J. B., and M. Gaster. "Rapid Estimation of Spectra from Irregularly Sampled Records," Proc. IEEE, 125(No. 2):92-96, February, 1978.

31. Norsworthy, K. H. "Fourier Transformation and Spectrum Analysis of Sparsely Sampled Signals," Laser Velocimetry and Particle Sizing. Washington, D.C.: Hemisphere Publishing Corp., 1979. Pp. 314-327.

32. Mayo, W. T., Jr. "Spectrum Measurements with Laser Velocimeters," Proceedings of the Dynamic Flow Conference on Dynamic Measurements in Unsteady Flows, 1978. Skovlunde, Denmark: Dynamic Flow Conference, 1978. Pp. 851-868.

33. Ortega, James M. Numerical Analysis, a Second Course. New York: Academic Press, Inc., 1971.

34. Dongarra, J. J., C. B. Moler, J. R. Bunch, and A. B. Stewart. Linpack User's Guide. Philadelphia: Society of Industrial and Applied Mathematics, 1979.

35. Jenkins, G. M., and D. G. Watts. Spectral Analysis and Its Applications. San Francisco: Holden Day, Inc., 1968.

36. Harris, F. J. "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," Proc. IEEE, 66(No. 1):51-83, January, 1978.

37. Carnahan, Brice, H. A. Luther, and J. O. Wilkes. Applied Numerical Methods. New York: John Wiley and Sons, Inc., 1969.

38. Draper, N. R., and H. Smith. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1966.

APPENDICES

APPENDIX A

LEAST SQUARE ERROR FIT METHOD

The vector matrix form of a set of equations is

$$\underline{Y} = X\underline{\beta} + \underline{e} \ , \qquad \text{(A-1)}$$

where,

$\underline{Y}$ is a (n×1) matrix which is known ,

X is a (n×m) matrix which is also known ,

$\underline{\beta}$ is a (m×1) matrix which is to be found ,

$\underline{e}$ is a (n×1) matrix which is the error vector .

The error vector may be written,

$$\underline{e} = \underline{Y} - X\underline{\beta} \ . \qquad \text{(A-2)}$$

Let S be a scalar equal to the quadratic form $\underline{e}'\underline{e}$, which is the sum of the squared error,

$$S = \underline{e}'\underline{e} = (\underline{Y} - X\hat{\underline{\beta}})' \ (\underline{Y} - X\hat{\underline{\beta}}) \ . \qquad \text{(A-3)}$$

Now the problem is to find $\hat{\underline{\beta}}$ so that S is minimized ($\hat{\underline{\beta}}$ = estimate of $\underline{\beta}$). One can write S after expanding Eq. (A-3),

$$S = \underline{Y}'\underline{Y} - \hat{\underline{\beta}}' \ X'\underline{Y} - \underline{Y}'X \ \hat{\underline{\beta}} + \hat{\underline{\beta}} \ X'X \ \hat{\underline{\beta}} \ . \qquad \text{(A-4)}$$

Let

$$P = X'X$$

$$\underline{Q} = X'\underline{Y} \ .$$

Now S becomes

$$S = \underline{Y}'\underline{Y} - \hat{\underline{\beta}}'\underline{Q} - \underline{Q}'\hat{\underline{\beta}} + \hat{\underline{\beta}}P\hat{\underline{\beta}} \ , \qquad \text{(A-5)}$$

or

$$S = (\hat{\underline{\beta}}' - \underline{Q}'P^{-1}) P (\hat{\underline{\beta}} - P^{-1}\underline{Q}) + \underline{Y}'\underline{Y} - \underline{Q}'P^{-1}\underline{Q}$$

(by adding and subtracting $\underline{Q}'P^{-1}\underline{Q}$ on the right-hand side of Eq. (A-5),

or

$$S = \{ (\hat{\underline{\beta}}' - \underline{Q}'P^{-1}) X' \} \{ X (\hat{\underline{\beta}} - P^{-1}\underline{Q}) \} + \underline{Y}'\underline{Y} - \underline{Y}'XP^{-1}X'\underline{Y} \ ,$$

or

$$S = \{ \hat{\underline{\beta}}'X' - \underline{Q}'P^{-1}X' \} \{ X\hat{\underline{\beta}} - XP^{-1}\underline{Q} \} + \underline{Y}'\underline{Y} - \underline{Y}'XP^{-1}X'\underline{Y} \ ,$$

or

$$S = (X\hat{\underline{\beta}} - XP^{-1}\underline{Q})' (X\hat{\underline{\beta}} - XP^{-1}\underline{Q}) + \underline{Y}'\underline{Y} - \underline{Y}'XP^{-1}X'\underline{Y} \ . \qquad \text{(A-6)}$$

Now, to minimize S, the terms $(\underline{Y}'\underline{Y})$ and $(\underline{Y}'XP^{-1}\ X'\underline{Y})$ cannot be zero. So

$$(X\hat{\underline{\beta}} - XP^{-1}\underline{Q})' \ (X\hat{\underline{\beta}} - XP^{-1}\underline{Q}) = 0 \ ,$$

or

$$(X\hat{\underline{\beta}} - XP^{-1}\underline{Q}) = 0 \ ,$$

or

$$\hat{\underline{\beta}} = P^{-1}\underline{Q} \ ,$$

or

$$\hat{\underline{\beta}} = (X'X)^{-1}X'\underline{Y} \ . \qquad \text{(A-7)}$$

A detailed treatment on the least square error fit may be found in Ref. [38], which is one of the many good references on the subject.

APPENDIX B

SOME USEFUL TIME STATISTICS AND THEIR ESTIMATORS

Among the many time statistics available in the literature, the following four are very useful ones in the processing of random signals. Their brief definitions and discrete estimators are included here and the reader will find them useful in going through the earlier material. In defining the time statistics, the important assumption is made that the process is stationary and ergodic. Detailed discussions and related assumptions may be found in Refs. [4, 7, 8, 9, 35].

1. Mean.
2. Variance.
3. Autocorrelation and autocovariance.
4. Spectral density.

Let $x(t)$ be a time statistic taken over a large time interval $(0,T)$. Usually, the random physical data record has two components--the static or time invariant part and the dynamic or fluctuating part. The static component is known as the mean value or average value. The mean value μ_x is written as

$$\mu_x = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)dt$$

$$= \langle x(t) \rangle \ , \qquad \text{(B-1)}$$

where < > is denoted as the infinite time average process. The dynamic component has different measures. A significant measure of the dynamic component is the variance which gives the spread of the signal about the mean value. The expression for variance is

$$\text{var}(x(t)) = \sigma_x^2 = \lim_{T\to\infty} \frac{1}{T} \int_0^T (x(t) - \mu_x)^2 dt \ . \qquad \text{(B-2)}$$

After expansion and algebraic manipulation the expression for σ_x^2 yields

$$\sigma_x^2 = \Psi_x^2 - \mu_x^2 \ , \qquad \text{(B-3)}$$

where Ψ_x^2 is the mean square value and is given by

$$\Psi_x^2 = \lim_{T\to\infty} \frac{1}{T} \int_0^T x^2(t)dt \ , \qquad \text{(B-4)}$$

and μ_x^2 is the square of the mean value. The positive square root of variance is known as the standard deviation.

The autocorrelation of random data shows the general dependence of the values of the data at one time on the values at another time. For a large T, the autocorrelation function is written as

$$R_x(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt$$

$$= \langle x(t)x(t+\tau) \rangle \ , \qquad \text{(B-5)}$$

where $R(\tau)$ is a real and even function of τ. The auto-covariance $C_x(\tau)$ is given by

$$C_x(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T (x(t) - \mu_x)(x(t+\tau) - \mu_x)dt \ . \qquad \text{(B-6)}$$

The two-sided power spectral density is defined as

$$S(\omega) = P \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau \ , \qquad \text{(B-7)}$$

which is the Fourier transform of $R(\tau)$. The term P is a constant whose value for different definitions is 1 or $1/2\pi$ or $1/\sqrt{2\pi}$. The difference is only a scaling factor and does not affect the spectrum shape. Since $R(\tau)$ is an even function and with P = 1, $S(\omega)$ is given by

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)\cos(\omega\tau)d\tau$$

$$= 2 \int_0^{\infty} R(\tau)\cos(\omega\tau)d\tau \ . \qquad \text{(B-8)}$$

To avoid leakage caused by a large zero frequency term, $C(\tau)$ is typically used in place of $R(\tau)$. Then $S(\omega)$ is given by

$$S(\omega) = \int_{-\infty}^{\infty} C(\tau)\cos(\omega\tau)d\tau \qquad \text{(B-9)}$$

and

$$C(\tau) = \int_{-\infty}^{\infty} S(\omega)\cos(\omega\tau)d\tau \ . \qquad \text{(B-10)}$$

Equations (B-9) and (B-10) form a cosine transform pair. A second definition of spectral density, an estimator of $S(\omega)$ (known as a periodogram), is computed from the Fourier transform of the data samples directly

$$S_T(\omega) = \frac{1}{T}\left| \int_0^T x(t)e^{-j\omega t}\, dt \right|^2 , \qquad \text{(B-11)}$$

and Papoulis [14] has shown that for large T,

$$E\,[\,S_T(\omega)\,] = S(\omega) \ .$$

The estimators, based on discrete equi-spaced data for the above defined continuous time statistics, are given by

1. Mean:

$$= \frac{1}{N}\sum_{n=1}^{N} x_n \ . \qquad \text{(B-12)}$$

2. Variance:

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \hat{\mu})^2$$

$$= \frac{1}{N-1} \left(\sum_{n=1}^{N} x_n^2 \right) - \frac{1}{N(N-1)} \left(\sum_{n=1}^{N} x_n \right)^2 . \qquad \text{(B-13)}$$

3. Autocorrelation:

$$R_r = \frac{1}{N-r} \sum_{n=1}^{N-r} x_n \, x_{n+r}$$

$$r_{max} << N . \qquad \text{(B-14)}$$

4. Autocovariance:

$$C_r = \frac{1}{N-r} \sum_{n=1}^{N-r} (x_n - \hat{\mu}) \, (x_{n+r} - \hat{\mu}) . \qquad \text{(B-15)}$$

In all the above cases, N = total number of samples of the random process x_n, and r = lag number. The power spectrum estimate may be written as

$$S(f_k) = \frac{2\Delta T}{N} \left| X_k \right|^2 , \qquad \text{(B-16)}$$

where,

$$\left| X_k \right|^2 = XR_k^2 + XI_k^2 ,$$

where,

XR = real part of Fourier transform of x_n ,

XI = imaginary part of Fourier transform of x_n ,

f_k = kf and f = spectral resolution = $\frac{1}{N\Delta T}$,

ΔT = time spacing between two consecutive data points for a regularly spaced record x_n ,

k = 1, 2, ... , N .

APPENDIX C

BIAS ERROR, LEAKAGE, AND ALIASING

In Fourier or frequency analysis, the finite record length influences the estimated spectrum. The estimated spectrum $S_X(f)$ is calculated for a finite length data set $s_X(x)$, which may be mathematically written as

$$s_X(x) = s(x) \cdot w_X(x) \ , \qquad \text{(C-1)}$$

where $w_X(x)$ is a rectangular window. It is also known as the "boxcar" window function and is given by

$$w_X(x) = 1 \ ; \ |x| \leq \frac{X}{2}$$

$$= 0 \ ; \ |x| > \frac{X}{2} \ . \qquad \text{(C-2)}$$

Equation (C-1) is shown pictorially in Fig. C-1.

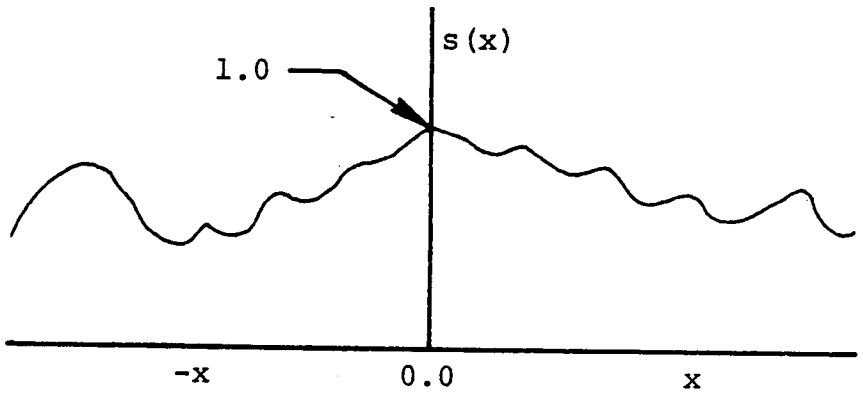
The Fourier transform of $w_X(x)$ gives

$$W_X(f) = \int_{-\infty}^{\infty} w_X(x) \exp(-j2\pi fx)\,dx$$

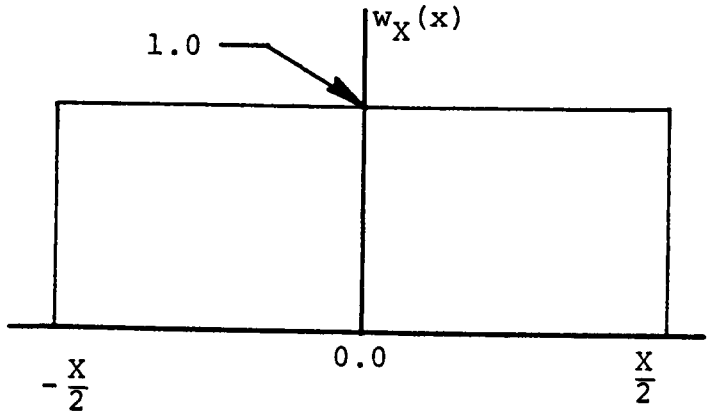
$$= X \left| \frac{\sin(\pi fX)}{\pi fX} \right| . \qquad \text{(C-3)}$$

$W_X(f)$ is known as the spectral window function and is plotted as shown in Fig. C-2.

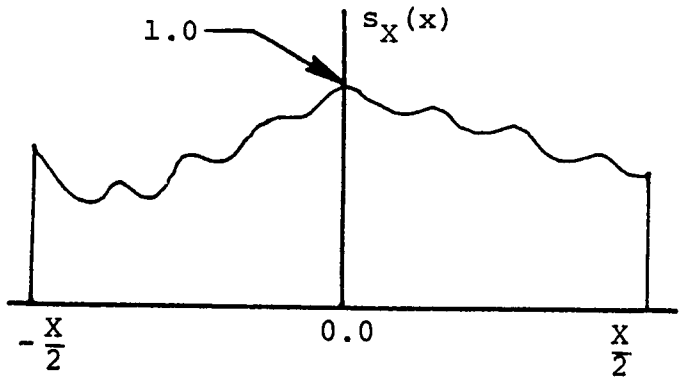
If two functions are multiplied as in Eq. (C-1), they are related equivalently in the frequency domain by



a. s(x) as a Function of x



b. Boxcar Weighting Function $w_X(x)$



c. s(x) Weighted by $w_X(x)$

Figure C-1. Pictorial representation of Eq. (C-1).

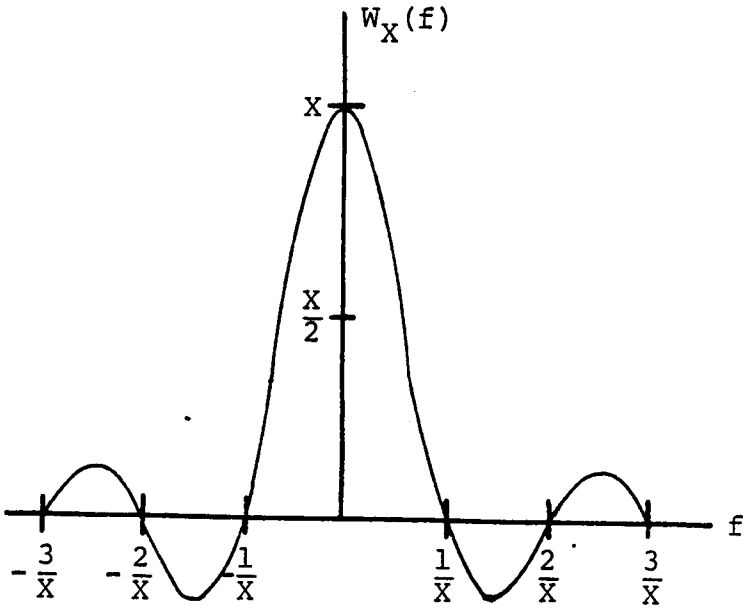


Figure C-2. Spectral window for boxcar weighting function.

the convolution integral,

$$\hat{S}_X(f) = \int_{-\infty}^{\infty} S_X(f_1) W_X(f-f_1) df_1$$

$$= \int_{-\infty}^{\infty} S_X(f_1) X \left(\frac{\sin(\pi(f-f_1)X)}{\pi(f-f_1)X} \right) df_1 , \qquad \text{(C-4)}$$

where,

$\hat{S}_X(f)$ = estimated spectrum ,

$S_X(f)$ = true spectrum .

The difference between $\hat{S}_X(f)$ and $S_X(f)$ is known as bias error. The strongest objection to the boxcar or rectangular window function is the "leakage" through the side lobes. Side lobes are those portions of the spectral window between 1/X and 2/X, between 2/X and 3/X, and so forth. Also note that half of the side lobes have negative amplitudes.

Leakage is reduced by using either a very long record or a window which tapers the ends of the data set. A long data set, even if available, is time consuming and expensive to process. So, the second scheme is often used. This method helps to reduce energy in the side lobes at the expense of a wider main lobe (reduced resolution). Tapering can be achieved in a large number of ways. Each method yields a distinct spectral window with different side lobe and main lobe characteristics. A number of

variations on window functions are described in the literature (for a comparison, see Ref. [36]). One of the popular windows is the Hanning window, which is defined as

$$\begin{aligned} h_X(x) &= 0 && ; \; x < -X/2 \\ &= \frac{1}{2}\left[1 + \cos(2\,x/X)\right] && ; \; |x| \leq X/2 \\ &= 0 && ; \; x > X/2 \; . \end{aligned} \qquad \text{(C-5)}$$

Then the Fourier transform of $h_X(x)$ gives the Hanning spectral window $H_X(f)$, where

$$H_X(f) = \frac{1}{4} W_X\,(f - f_o) + \frac{1}{2} W_X(f) + \frac{1}{4} W_X(f + f_o) \; , \qquad \text{(C-6)}$$

and f_o is given by

$$f_o = 1/X \; .$$

The Hanning spectral window function is shown in Fig. C-3.

For equally spaced data, two frequencies are aliases of one another if sinusoids of the corresponding frequencies cannot be distinguished by their equally spaced values [9]. An example of aliasing is shown in Fig. C-4a. The data sampling rate is not high enough to accurately represent the higher frequency; additionally, in calculating the spectrum, the energy of the higher frequency term contributes to the energy of the lower frequency term. Figure C-4b illustrates an alias-free spectrum. An

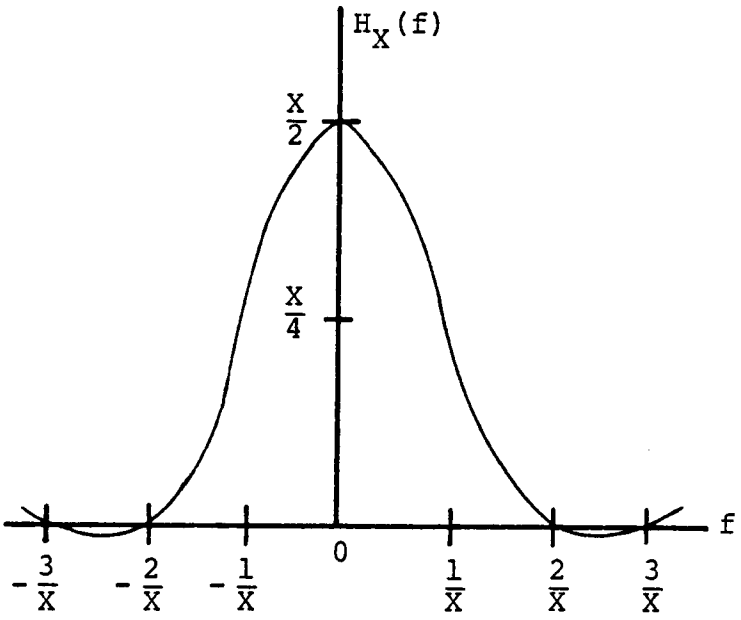
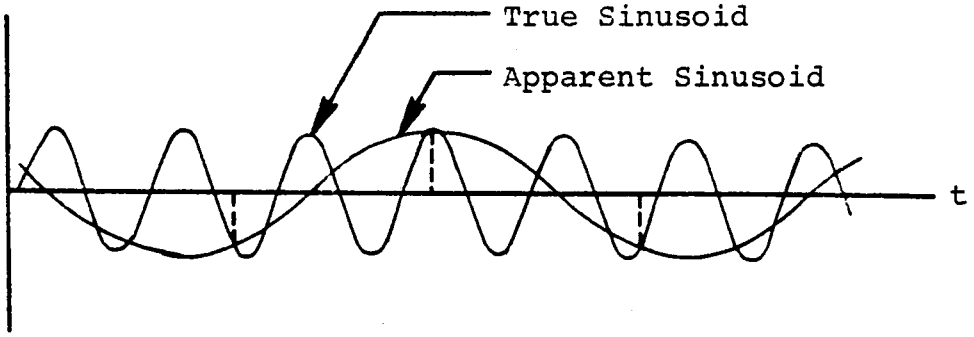
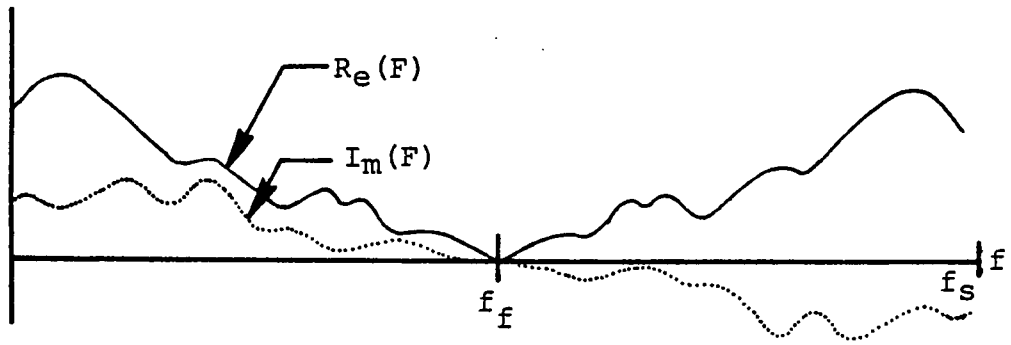


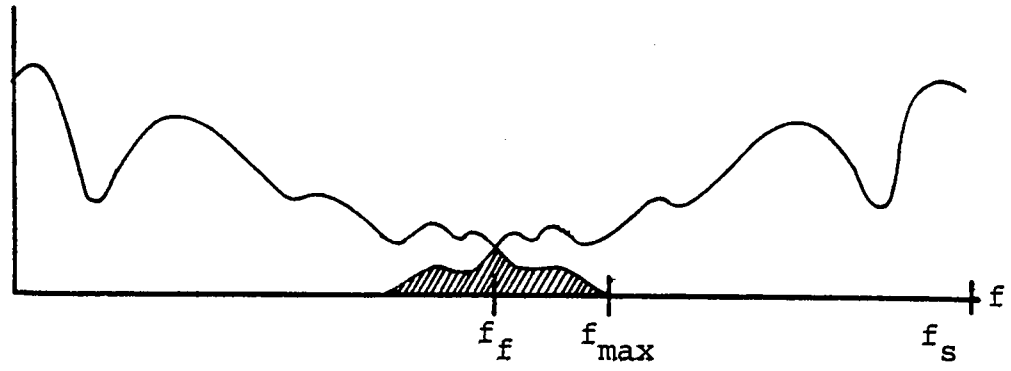
Figure C-3. Hanning spectral window function.



a. High Frequency Signal Impersonating Lower Frequency Due to Improper Sampling Rate



b. Frequency Spectrum "Folded" Around Folding Frequency $f_f = f_s/2$ (f_s = Sampling Frequency)



c. Aliasing Due to $f_{max} > f_f$ Causing Overlap in the Frequency Spectrum

Figure C-4. Pictorial presentation of aliasing.

aliased spectrum is shown in Fig. C-4c. Aliasing may be eliminated by (1) higher sampling rate, or (2) filtering out the higher frequencies before sampling. The course of action usually is dictated by the nature of the data, type of instrumentation and its limitations, data recording procedures, etc.

APPENDIX D

THE TECHNIQUE OF BRICK-WALL FILTERING

Filtering a signal can be accomplished by multiplying the signal transform with a filter transfer function. The product is then inverse transformed into a time domain signal. Pictorially, the process is demonstrated in Figs. D-1a, D-1b, and D-1c. The process is carried out via computer routine after all the data are collected.

In the method of brick-wall filtering, unwanted frequency components in the signal are set to zero and then an inverse transform is applied to get the filtered signal. An example to produce band-limited white noise using the FFT algorithm is given here to illustrate the procedure.

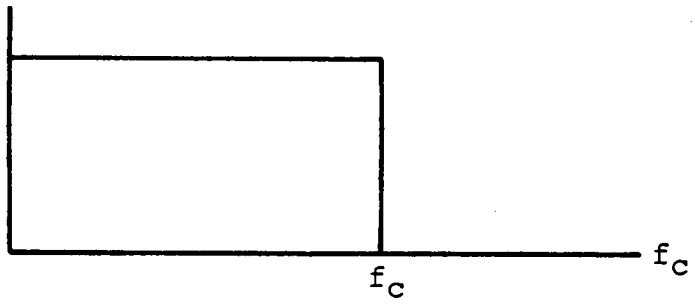
A real Gaussian random variable is generated and transformed yielding the result in Fig. D-2.

If the signal is band limited to the frequency N/4, the frequency components for $N/4 < n \leq N/2$ are set to zeros (both real and imaginary parts). The real and imaginary parts for frequency components between (N/2+1) to (N-1) are obtained from the symmetry of the FFT (even for real and odd for imaginary) about N/2. The band-limited spectrum is shown in Fig. D-3.

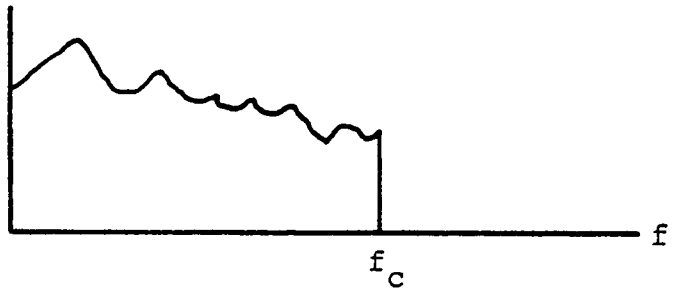
Now the inverse FFT is applied. The real part of the inverse Fourier transform provides the desired band-limited discrete (N equi-spaced samples) Gaussian noise in



a. Frequency Spectrum of Signal



b. Ideal Filter Transfer Function



c. Spectrum of Filtered Signal

Figure D-1. Process of filtering in the frequency domain.

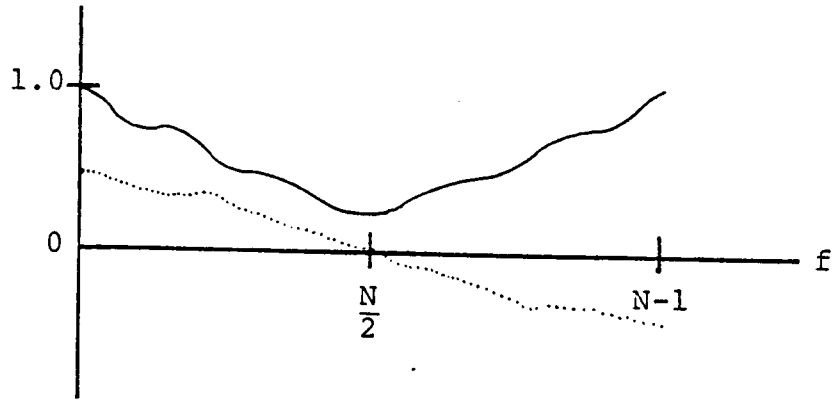


Figure D-2. Real and imaginary parts of Fourier transformed signal.

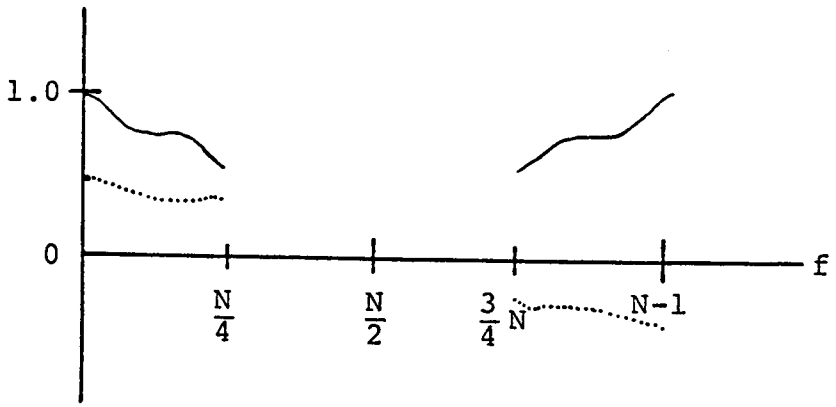


Figure D-3. Brick-wall filtered frequency response of Fig. D-2.

the time domain and the imaginary parts (whose magnitudes are very small compared to real parts) are neglected.

The above brick-wall filtering technique may easily be applied for bandpass and band reject filtering also.

VITA

Mukta Lal Kar was born on January 20, 1949 in Howrah, India, where he received his elementary and high school education. In 1963 he joined the Presidency College, Calcutta University, India, to study Pre-University in Science. In 1964 he entered the School of Engineering, Jadavpore University, Calcutta, India, to study for the five-year integrated engineering degree. He was awarded the degree of Bachelor in Electronics and Tele-Communication Engineering in August 1969.

Mr. Kar was employed by the Electronics and Radar Development Establishment (LRDE), Bangalore, India, from August 1969 to December 1969. That same year he joined the Vikram Sarabhai Space Center, Trivandrum, India, where he was involved in the research, design, and development project of the digital telemetry system for space vehicles until August 1974. That same year he received an offer of Graduate Research Assistantship from The University of Tennessee, Knoxville, and in the Fall of 1974 he began graduate study at The University of Tennessee Space Institute (UTSI). In December 1976 he received the Master of Science degree in Electrical Engineering. Mr. Kar continued graduate study, receiving his Ph.D. degree in Engineering Science in June 1981.

Since December 1980 Mr. Kar has been employed by the Intergraph Corporation, Huntsville, Alabama.

Mr. Kar is married to the former Manjula Biswas of Taki, India, and they have a six-year-old son, Partha.