MULTIPLE CHANNEL MAXIMUM ENTROPY SPECTRAL ESTIMATOR AND ITS APPLICATION

by

ALBERT TUNG-YIU NG

B.S., Physics, California Institute of Technology (1974)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE

DEGREE OF MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1977

Signature of Author _____
Department of Earth and Planetary Sciences

Certified by _____
Thesis Supervisor

Accepted by _____
Chairman, Departmental Committee on Graduate Students

MULTIPLE CHANNEL MAXIMUM ENTROPY SPECTRAL ESTIMATOR AND ITS APPLICATION

by

ALBERT TUNG-YIU NG

Submitted to the Department of Earth and Planetary Sciences on January 21, 1977, in partial fulfillment of the requirements for the degree of Master of Science in Geophysics.

ABSTRACT

The performance of Burg's multivariate prediction estimator used for computing maximum entropy (autoregressive) spectra was evaluated. Specifically, the accuracy, resolution, effects of noise, and estimation of autoregressive order were studied using synthesized data. Results of these empirical studies demonstrated the superiority of Burg's technique over the conventional windowing and transforming method. Spectral analysis was then carried out on three geological data sets from South Indian Ocean cores. Results of the analysis tend to support the conclusion of Hays et. al. (1976) that pleistocene ice ages have as one of their possible causes variation in the orbital motion of the Earth.

Thesis advisor: M. Nafi Toksoz, Professor of Geophysics

TABLE OF CONTENTS

INTRODUCTION

In geophysics, the effect of a given physical process can often show up in several types of observable quantities. In such cases, multi-channel time series spectral analysis can be used to establish correlations between the different data sets and thereby help to define the basic process controlling the observations.

Examples taken from the area of seismology include three component seismic data recorded at a single station, seismic array data, multi-station/multi-source seismic reflection data and pore-pressure, tilt, elastic velocity profile, and so on. Geophysical variables such as magnetic and gravity values sampled along a single profile, temperature, pressure, and wind speed measured at a single station, or chemical composition, density, and interval velocity collected as functions of depth in a well log constitute other examples of multiple channel data. Although each of the above quantities may be analysed independently, the interval mechanism of the system or process generating them can be intimately related. In some cases such as pore pressure and velocity or pressure and wind speed, one of the variables is the cause of the other. A knowledge of whether certain variables are correlated, and a measure of such a correlation is therefore vital to understanding the nature of the systems governing the observations.

Multi-variate spectral analysis provides one method with which correlation information existing between several variables can be uncovered and evaluated. For many types of data, the shape and amplitude of a spectrum is itself useful information, in particular when the form of the spectrum can be predicted by theory.

If the various channels are correlated, multi-variate spectral analysis can be applied to deduce the relative phase relationships between them. The success of spectral analysis in this respect has been well demonstrated by the use of the maximum likelihood method in wave number analysis of array processing (Capon, 1969).

Conventional methods for estimating cross spectral density of multi-channel time series are based on generalizations of the Blackman and Tukey (1959) and the periodogram techniques (Jones, 1965). In the Blackman and Tukey approach, autocorrelation matrix R is estimated from the data $X_i(n)$, i=1, ,m, n-1, ,T by

$$Rij(l) = \frac{1}{N-l} \sum_{n=1}^{N-1} X_i(n)\, X_j^*(n+l)$$

$$l=o,1, \quad ,k$$

where an asterisk (*) denotes, conjugate transpose of a column vector. Each auto-correlation matrix element is tapered with w(t) and then Fourier transformed to produce the spectral matrix Pij(w):

$$Pij(w) = \quad R_{ij}(t)e^{\,iwt}\, dt$$

For this reason, the method is also referred to as taper and transform.

The periodogram approach does not require an estimate of the autocorrelation sequence. Instead, the data sequence is first partitioned into smaller sequences of identical length. Each individual sequence is Fourier transformed and multiplied by the transform of its conjugate transpose to produce one matrix spectral sequence. The spectral sequences from all partitions are then averaged to result in an estimate of the matrix power spectral sequence.

The Blackman and Tukey and periodogram approaches as applied to single channel data have difficulties that do not disappear in their multi-variate extensions. Specifically, smearing and leakage of power due to their smoothing operations reduce the resolution of the resulting spectra. The multi-variate extension of the periodogram method likewise has the same problem.

Autoregressive spectral estimation techniques provide an alternative method. One estimator in particular has been studied and discussed extensively by geophysicists and electrical engineers (Lacoss, 1971; Ulrych, 1972; Ulrych and Bishop, 1975; Kavah and Cooper, 1976). This particular estimator was proposed by Burg (1967) from the point of view of maximizing the entropy per sample of a time series, subject to

constraints of known estimates of its autocorrelation samples. He, therefore, referred to it as the Maximum Entropy Method. An alternative approach of arriving at an estimator, based on fitting an autoregressive model to the time series has also been suggested by Parzen (1969). This autoregressive model acts as a linear filter which reduces the original data to a white process if the data is truly autoregressive of finite order. This white process is generally referred to as the innovation, or prediction error with equivalent connotation.

An algorithm, developed by Burg (1975), utilizes the Levinson-Durbin Recursion relation to estimate the autoregressive coefficients directly from the data. Specifically, at each iteration of the recursion, a new set of coefficients is computed by minimizing the sum of squares of the one step ahead forward and backward prediction errors. Alternative to Burg's method, the Yule-Walker method computes the model parameters directly from the autocorrelation of the data, which must be estimated in a windowed or non-windowed manner.

Superior accuracy, resolution, and stability of Maximum Entropy spectral estimates over B-T or periodogram for single channel data have been demonstrated by Lacoss (1971) and others for both the Yule-Walker and the Burg algorithms. It, therefore, seems logical that this method should be extended to spectral analysis of multiple time series.

Until recently, almost all research into Maximum Entropy Method of spectral estimation was concerned with single channel processing. Extension of the Yule-Walker algorithm to multiple time series is given in Burg (1975) and Parzen (1975). Applications of this algorithm in estimation of cross-power, coherence, and phase spectra have also been reported (Jones, 1974; Parzen, 1976).

A major problem encountered by the application of Yule-Walker algorithm is the need for estimating the sample autocorrelation sequence. Most often the estimated autocorrelation sequence may not necessarily be positive definite, resulting in negative values of power spectra at certain frequencies.

The Burg algorithm applied to single channel time series can be shown to guarantee non-negative spectral estimates. A basic requirement of extending this algorithm to spectral analysis of multiple time series is therefore the preservation of this distinctive property. Such an extension was done by Jones (1976), Nuttal (1976) and Lacoss (1976) independently.

The main objective of this thesis is to demonstrate the properties of Burg's Maximum Entropy Method, as extended to multiple time series as an estimator of cross power spectra. To the best knowledge of the author, no application of this method in multichannel processing of real data has been reported.

The organization of the present thesis is conceptually divided into three major chapters. Chapter I reviews the theoretical development of the Multivariate Burg Estimator. Algorithm derived from this development has been implemented into a system of Fortran programs with which the experimental studies reported in the other chapters were conducted.

Results of experiments conducted with synthesized data are given in Chapter II. Specifically, accuracy and resolution of estimated spectra will be evaluated. Several problems associated with applications of the Burg estimator are also discussed, which include the very important problem of choosing an optimal autoregressive order.

An application of the Multivariate Burg Spectral Estimator to real data is reported in Chapter III. This technique has been used to explore the correlation relation among three geological data sets to test the validity of the Astronomical Theory of ice ages. The results are compared to the findings of Hays and others (1976) who utilized the procedures of Blackman and Tukey.

Conclusions and future work are discussed in the final chapter. Owing to the empirical nature of this investigation, these conclusions reached are far from universal and unchangeable. It is hoped that a more theoretical approach in studying the method will be forthcoming.

Burg algorithm applied to single channel time series has been proven to yield non-negative spectral estimates. A basic requirement of extending this algorithm to spectral analysis of multiple time series is therefore the preservation of this distinctive property. Such an extension has been suggested by Jones (1976) and Nuttal (1976) independently.

The main objective of the investigation reported by the present thesis is to demonstrate the feasibility of Burg's Maximum Entropy Method, as extended by Jones and Nuttal to multiple time series processing, as an estimator of cross power spectra.

To the best knowledge of the author, no application of this method in multichannel processing of real data has been reported. Therefore, such an investigation is warranted for its contribution to understanding the various characteristics which distinguish the Multivariate Burg Spectral Estimators from the other estimators.

The organization of the present thesis is conceptually divided into three major chapters. Chapter I reviews the theoretical development of the Multivariate Burg Estimator. Algorithm derived from this development has been implemented into a system of Fortran programs with which the experimental studies reported in the other chapters were conducted.

Results of experiments conducted with synthesized data are given in Chapter II. Specifically, accuracy and resolution of estimated spectra will be evaluated. Several problems associated with applications of the Burg estimator are also discussed, which include the very important problem of choosing an optimal autoregressive order.

One application of the Multivariate Burg Spectral Estimator to real data is reported in Chapter III. This technique has been used to explore the correlation relation among three goelogical data in order to test

the validity of the Astronomical Theory of ice ages. The results are compared to the findings of Hays and others (1976) who utilized the procedures of Blackman and Tukey.

Conclusions and future works are discussed in the final chapter. Owing to the complexity of the field of spectral analysis, and the empirical nature of investigation as reported by the present thesis, these conclusions are far from universal and unchangeable. It is hoped that a more theoretical approach in studying the bias, error, consistency, and efficiency of the Multivariate Burg Spectral Estimator can be developed in the future.

I. Maximum Entropy Estimation of Multivariate Spectra

I.1 Multivariate Yule-Walker Equation

Let $x(n)$ represent an m dimensional, stationary vector random process with zero mean. Given that an order L is specified, a set of forward autoregression coefficients $A_\ell^{(L)}$, and a set of backward autoregression coefficients $B_\ell^{(L)}$ are defined by the relations

$$e_+^{(L)}(n) = x(n) + \sum_{\ell=1}^{L} A_\ell^{(L)}\, x(n-\ell) \qquad n=L+1,\ldots,N \tag{1.1}$$

and

$$e_-^{(L)}(n) = x(n) + \sum_{\ell=1}^{L} B_\ell^{(L)}\, x(n+\ell) \qquad n=1,\ldots,N-L \tag{1.2}$$

where $e_+^{(L)}(n)$ is the forward innovation, or forward prediction error sequence; similarly, $e_-^{(L)}(n)$ is the backward innovation, or backward prediction error sequence.

Let $\tilde{a}$ denote the conjugate transpose of a complex vector a, then a forward prediction error matrix and a backward prediction error matrix can be defined by

$$J_+ = \sum_{n=L+1}^{N} e_+^{(L)}(n)\, \tilde{e}_+^{(L)}(n) \tag{1.3}$$

and

$$J_- = \sum_{n=1}^{N-L} e_-^{(L)}(n)\, \tilde{e}_-^{(L)}(n) \tag{1.4}$$

substituting (1.1) in (1.3)

$$J_+ = \sum_{n=L+1}^{N} \{x(n)\tilde{x}(n) + x(n) \sum_{\ell=1}^{L} \tilde{x}(n-\ell)\tilde{A}_\ell^{(L)} + \sum_{\ell=1}^{L} A_\ell^{(L)} x(n-\ell)\tilde{x}(n)$$

$$+ \sum_{\ell=1}^{L} \sum_{m=1}^{L} A_\ell^{(L)} x(n-\ell)\tilde{x}(n-m)\tilde{A}_m^{(L)}\} \tag{1.5}$$

The sample autocorrelation matrix of lag k is defined as

$$\hat{R}(k) = \frac{1}{N-L} \sum_{n=L+1}^{N} x(n)\tilde{x}(n-k) \tag{1.6}$$

Using this definition in (1.5)

$$J_+ = (N-L)\ \{\hat{R}(o) + \sum_{\ell=1}^{L} \hat{R}(\ell)\tilde{A}_\ell^{(L)} + \sum_{\ell=1}^{L} A_\ell^{(L)}\hat{R}(-\ell)$$

$$+ \sum_{\ell=1}^{L} \sum_{m=1}^{L} A_\ell^{(L)}\hat{R}(m-\ell)\tilde{A}_m^{(L)}\} \tag{1.7}$$

The criterion for choosing the autoregressive coefficients is then to minimize the trace of this matrix. If a_{ij}^k denotes the ith row and jth column element of $A_k^{(L)}$, then

$$\frac{\partial}{\partial a_{ij}^k} \text{trace}\{J_+\} = (N-L) \sum_{\ell=1}^{L} \text{trace}\ \{\hat{R}(\ell)\ \frac{\partial}{\partial a_{ij}^k} \tilde{A}_\ell^{(L)}$$

$$+ \frac{\partial}{\partial a_{ij}^k} A_\ell^{(L)}\hat{R}(-\ell)\}$$

$$+ (N-L) \sum_{=1}^{L} \sum_{m=1}^{L} \text{trace}\{[A_\ell^{(L)}\hat{R}(m-\ell)]\frac{\partial}{\partial a_{ij}^k} \tilde{A}_m^{(L)}$$

$$+ \frac{\partial}{\partial a_{ij}^k} A_\ell^{(L)}[\hat{R}(m-\ell)A_m^{(L)}]\} \tag{1.8}$$

If I_{ij} denotes a matrix with a nonzero element of 1 only at the ith row and jth column, then the following relations are true.

(i) $\frac{\partial}{\partial a^k_{ij}} A^{(L)}_{\ell} = I_{ij}\ \delta_{k\ell}$

(ii) $\frac{\partial}{\partial a^k_{ij}} \tilde{A}^{(L)}_{m} = I_{ji}\ \delta_{km} = \tilde{I}_{ij}\ \delta_{km}$

(iii) trace $\{I_{ij}\ B\} = B_{ji} = \tilde{B}_{ij}$

(iv) trace $\{B\ \tilde{I}_{ij}\} = B_{ij}$

(v) trace $\{B\}$ = trace $\{\tilde{B}\}$

Using these relations, (1.8) can be reduced to

$$\frac{1}{N-L} \frac{\partial}{\partial a^k_{ij}} \text{trace}\ \{J_+\} + 2\hat{R}_{ij}(k) + 2 \sum_{m=1}^{L} [A^{(L)}_m \hat{R}(k-m)]_{ij}$$

$$k=1,2,\ldots,L \qquad (1.9)$$

To minimize the term on the L.H.S., the sum of terms on the R.H.S. are set equal to zero, resulting in the matrix equation

$$\hat{R}(k) + \sum_{m=1}^{L} A^{(L)}_m \hat{R}(k-m) = 0 \qquad (1.10)$$

The above equation can be rewritten as follows:

$$[A^{(L)}_1\ A^{(L)}_2\ \ldots\ A^{(L)}_L] \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \ldots\ldots \hat{R}(L-1) \\ \hat{R}(-1) & \hat{R}(0) \ldots\ldots \hat{R}(L-2) \\ \vdots & \vdots \\ \hat{R}(-L+1) & \hat{R}(-L+2) \ldots \hat{R}(0) \end{bmatrix} = [-\hat{R}(1)\ -\hat{R}(2)\ \ldots\ -\hat{R}(L)] \qquad (1.11)$$

If the row vector $[A_1^{(L)}\ A_2^{(L)}\ .\ .\ .A_L^{(L)}]$ is augmented on the right by adding the identity matrix I, equation (1.11) can be rearranged into

$$[I\ A_1^{(L)}\ \ldots\ A_L^{(L)}]\begin{bmatrix}\hat{R}(0) & \hat{R}(1)\ldots\ldots\hat{R}(L)\\ \hat{R}(-1) & \hat{R}(0) \quad\quad \hat{R}(L-1)\\ \vdots & \quad\quad\quad\quad \vdots\\ \hat{R}(-L) & \hat{R}(-L+1)\ldots\hat{R}(0)\end{bmatrix} = [U_+^{(L)}\ \ 0\ \ 0\ \ldots\ 0] \tag{1.12}$$

The matrix $U_+^{(L)}$, also called the forward prediction error matrix for reasons that will be given in the next section, is artificially created by adding an extra column to the augmented autocorrelation matrix in order to keep it in Toeplitz form. This matrix is given by

$$U_+^{L)} = \hat{R}(0) + \sum_{m=1}^{L} A_m^{(L)}\,\hat{R}(-m) \tag{1.13}$$

A similar derivation applied to the backward prediction error martix J_- results in the equation for the backward autoregressive coeffi-cients $B_\ell^{(L)}$ as follows:

$$[B_L^{(L)}\ B_{L-1}^{(L)}\ \ldots\ldots B_1^{(L)}\ \ I]\begin{bmatrix}\hat{R}(0) & \hat{R}(1)\ldots\ldots\ldots\hat{R}(1)\\ \hat{R}(-1) & \hat{R}(0) \quad\quad\quad \hat{R}(L-1)\\ \vdots & \quad\quad\quad\quad\quad\quad \vdots\\ \hat{R}(-L) & \hat{R}(-L+1)\ldots\ldots\hat{R}(0)\end{bmatrix} = [0\ \ldots\ldots 0\ U_-^{(L)}] \tag{1.14}$$

$$\text{with } U_-^{(L)} = \hat{R}(0) + \sum_{m=1}^{L} B_m^{(L)}\,\hat{R}(m) \tag{1.15}$$

Equations (1.12) and (1.14) are then combined to form a single equation, which is the multivariate equivalence of the Yule-Walker equation in univariate maximum entropy spectral estimation theory. This equation is

$$\begin{bmatrix} I & A_1^{(L)} \cdots A_{L-1}^{(L)} & A_L^{(L)} \\ B_L^{(L)} & B_{L-1} \cdots B_1^{(L)} & I \end{bmatrix} \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \cdots \cdots \hat{R}(L) \\ \hat{R}(-1) & \hat{R}(0) \cdots \cdots \hat{R}(L-1) \\ \vdots & \vdots \\ \hat{R}(-L) & \hat{R}(-L+1) \cdots \hat{R}(0) \end{bmatrix} = \begin{bmatrix} U_+^{(L)} & 0 \cdots 0 \; 0 \\ 0 & 0 \cdots 0 \; U_-^{(L)} \end{bmatrix} \tag{1.16}$$

I.2 Prediction error matrices

Two particularly important properties of $U_+^{(L)}$ and $U_-^{(L)}$ can be deduced. First, post-multiplying equation (1.14) by the column vector $[I \; \tilde{A}_1^{(L)} \; \tilde{A}_2^{(L)} \cdots \tilde{A}_L^{(L)}]$ results in

$$[I \; A_1^{(1)} \cdots A_L^{(L)}] \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \cdots \cdots \hat{R}(L) \\ \hat{R}(-1) & \hat{R}(0) \cdots \cdots \hat{R}(L-1) \\ \vdots & \vdots \\ \hat{R}(-L) & \hat{R}(-L+1) \cdots \hat{R}(0) \end{bmatrix} \begin{bmatrix} I \\ \tilde{A}_1^{(L)} \\ \vdots \\ \tilde{A}_L^{(L)} \end{bmatrix} = U_+^{(L)} \tag{1.17}$$

Taking the conjugate transpose of both sides of (1.19) and using the symmetry property of $\hat{R}(L)$,

$$[I\ A_1^{(L)} \ldots A_L^{(L)}] \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \ldots\ldots \hat{R}(L) \\ \hat{R}(-1) & \hat{R}(0) \quad\quad \hat{R}(L-1) \\ \vdots & \vdots \quad\quad\quad \vdots \\ \hat{R}(-L) & \hat{R}(-L+1) \quad \hat{R}(0) \end{bmatrix} \begin{bmatrix} I \\ \tilde{A}_1^{(L)} \\ \\ \\ \tilde{A}_L^{(L)} \end{bmatrix} = \tilde{U}_+^{(L)} \tag{1.18}$$

Comparing (1.17) and (1.18), it is obvious that the matrix $U_+^{(L)}$ as given by (1.15) is symmetric (or Hermitian for complex valued data), that is

$$U_+^{(L)} = \tilde{U}_+^{(L)} \tag{1.19}$$

It can be proven by similar derivation that

$$U_-^{(L)} = \tilde{U}_-^{(L)} \tag{1.20}$$

The second property of these matrices is now deduced by noting that the matrices J_+ and J_- whose traces are being minimized can be written in a form similar to that given by equation (1.17). Specifically, equation (1.7) becomes

$$\frac{1}{N-L} J_+ = [I\ A_1^{(L)}\ A_2^{(L)} \ldots A_L^{(L)}] \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \ldots\ldots \hat{R}(L) \\ \hat{R}(-1) & \hat{R}(0) \ldots\ldots \hat{R}(L-1) \\ \vdots & \vdots \quad\quad\quad \vdots \\ \hat{R}(-L) & \hat{R}(-L+1) \quad \hat{R}(0) \end{bmatrix} \begin{bmatrix} I \\ A_1^{(L)} \\ \vdots \\ A_L^{(L)} \end{bmatrix} \tag{1.21}$$

substituting (1.12) into (1.21) results in

$$\frac{1}{N-L} J_+ = U_+^{(L)} \tag{1.22}$$

Similarly, it can be proven that

$$\frac{1}{N-L} J_- = U_-^{(L)} \tag{1.23}$$

Therefore, the two matrices $U_+^{(L)}$ and $U_-^{(L)}$ are indeed the minimum trace prediction error matrices at the autoregressive order of L.

I.3 Multivariate Levinson-Durbin Recursion

Owing to the special form of a Toeplitz type matrix, the Yule-Walker equation (1.16) can be solved recursively. This recursion was first proposed by Levinson and improved by Durbin for application in single channel signal processing. Extension of the recursion scheme to multichannel signal processing has been given by Wiggins and Robinson (1965). An adaptation of this scheme for the solution of equation (1.16) is described below.

It would be assumed that forward and backward autoregressive coefficients of order L-1 has been known, together with their corresponding prediction error matrices as given by (1.13) and (1.15). The (L-1)th order Yule-Walker equation is then given by

$$\begin{bmatrix} I & A_1^{(L-1)} \cdots\cdots A_{L-1}^{(L-1)} \\ B_{L-1}^{(L-1)} \cdots\cdots\cdots B_1^{(L-1)} & L \end{bmatrix} \begin{bmatrix} \hat{R}(0) & \hat{R}(1) \cdots\cdots \hat{R}(L-1) \\ \hat{R}(-1) & \hat{R}(0) \qquad\quad \hat{R}(L-2) \\ \vdots & \vdots \qquad\quad \vdots \\ \hat{R}(-L+1) & \hat{R}(-L+2) \qquad \hat{R}(0) \end{bmatrix} = \begin{bmatrix} U_+^{(L-1)} \cdots 0 \\ 0 \cdots\cdots U_-^{(L-1)} \end{bmatrix} \tag{1.24}$$

Define matrix coefficients $C_+^{(L)}$, $C_-^{(L)}$ by the recursion relation

$$\begin{bmatrix} I & A_1^{(L)} \cdots\cdots A_L^{(L)} \\ B_L^{(L)} \cdots\cdots B_1^{(L)} & I \end{bmatrix} = \begin{bmatrix} I & A_1^{(L-1)} \cdots\cdots A_{L-1}^{(L-1)} & 0 \\ 0 & B_{L-1}^{(L-1)} \cdots\cdots B_1^{(L-1)} & I \end{bmatrix}$$

$$+ \begin{bmatrix} C_+^{(L)} & 0 \\ 0 & C_-^{(L)} \end{bmatrix} \begin{bmatrix} 0 & B_{L-1}^{(L-1)} \cdots B_1^{(L-1)} & I \\ I & A_1^{(L-1)} \cdots A_{L-1}^{(L-1)} & 0 \end{bmatrix} \tag{1.25}$$

Substituting this relation into equation (1.16) and using (1.24), one has

$$\begin{bmatrix} U_+^{(L)} \cdots 0 \cdots 0 \\ 0 \cdots\cdots 0 \cdots U_-^{(L)} \end{bmatrix} = \begin{bmatrix} U_+^{(L-1)} \cdots 0 & E_+^{(L)} \\ E_-^{(L)} & 0 \cdots\cdots U_-^{(L-1)} \end{bmatrix} + \begin{bmatrix} C_+^{(L)} & 0 \\ 0 & C_-^{(L)} \end{bmatrix} \begin{bmatrix} E_-^{(L)} & 0 \cdots\cdots U_-^{(L-1)} \\ U_+^{(L-1)} \cdots 0 & E_+^{(L)} \end{bmatrix} \tag{1.26}$$

where

$$E_+^{(L)} = \hat{R}(L) + \sum_{m=1}^{L-1} A_m^{(L-1)} \hat{R}(L-m) \tag{1.27}$$

and

$$E_-^{(L)} = \hat{R}(-L) + \sum_{m=1}^{L-1} B_m^{(L-1)} \hat{R}(-L+m) \tag{1.28}$$

Equating both sides of equation (1.26) column by column, equations for the matrices $C_+^{(L)}$ and $C_-^{(L)}$ can be obtained as follows:

$$E_+^{(L)} + C_+^{(L)} U_-^{(L-1)} = 0 \tag{1.29}$$

and

$$E_-^{(L)} + C_-^{(L)} U_+^{(L-1)} = 0 \tag{1.30}$$

from which

$$C_+^{(L)} = -E_+^{(L)} [U_-^{(L-1)}]^{-1} \tag{1.31}$$

and

$$C_-^{(L)} = -E_-^{(L)} [U_+^{(L-1)}]^{-1} \tag{1.32}$$

once $C_+^{(L)}$ and $C_-^{(L)}$ are computed, $U_+^{(L)}$ and $U_-^{(L)}$ can readily be expressed in terms of $U_+^{(L-1)}$, $U_-^{(L-1)}$ using equation (1.26), and $A_\ell^{(L)}$, $B_\ell^{(L)}$ be given by $A_\ell^{(L-1)}$, $B_\ell^{(L-1)}$ using equation (1.25). A summary of the recursion beginning with order zero is given below

(i)
$$A_0^{(0)} = I$$
$$B_0^{(0)} = I$$
$$U_+^{(0)} = \hat{R}(0)$$
$$U_-^{(0)} = \hat{R}(0)$$

(ii)
$$C_+^{(L)} = -\{\sum_{m=0}^{L-1} A_m^{(L-1)} \hat{R}(L-m)\} [U_-^{(L-1)}]^{-1}$$
$$A_0^{(L)} = I$$
$$A_\ell^{(L)} = A_\ell^{(L-1)} + C_+^{(L)} B_{L-\ell}^{(L-1)} \qquad \ell = 1,2,\ldots,L-1$$
$$A_L^{(L)} = C_+^{(L)}$$
$$U_+^{(L)} = [I - C_+^{(L)} C_-^{(L)}] U_+^{(L-1)}$$

(iii)
$$C_-^{(L)} = -\{\sum_{m=0}^{L-1} B_m^{(L-1)} \hat{R}(-L+m)\} [U_+^{(L-1)}]^{-1}$$
$$B_0^{(L)} = I$$
$$B_\ell^{(L)} = B_\ell^{(L-1)} + C_-^{(L)} A_{L-\ell}^{(L-1)} \qquad \ell = 1,2,\ldots,L-1$$
$$B_L^{(L)} = C_-^{(L)}$$
$$U_-^{(L)} = [I - C_-^{(L)} C_+^{(L)}] U_-^{(L-1)}$$

(1.33)

I.4 Multivariate Burg Spectral Estimator

The Yule-Walker algorithm computes the reflection coefficients $C_+^{(L)}$, $C_-^{(L)}$ using estimates of the sample autocorrelation sequence $\hat{R}(\ell)$ up to a maximum lag of L. Numerical experimentation indicates that this method does not necessarily give estimates of power spectra which are non-negative at all frequencies. To remedy this pitfall, Burg suggested estimating the reflection coefficient directly from the data by minimizing the sum of squares of the next step prediction error. In single channel signal processing, the forward reflection coefficient $C_+^{(L)}$ and the backward reflection coefficient $C_-^{(L)}$ are both scalars which are also conjugate of one another. Thus minimizing the forward prediction error and backward prediction error separately results in the same power spectral estimate. No such simple relation exists in multichannel signal processing, and the two separate spectral estimates may not be the same.

To assure a single multivariate spectral estimate by minimizing the sum of squares of both forward and backward prediction errors simultaneously, one single parameter in place of the two reflection coefficients has to be defined. Judging by the forms of $C_+^{(L)}$ and $C_-^{(L)}$ given in equations (1.31) and (1.32), Jones suggested a matrix $G^{(L)}$ defined by

$$C_+^{(L)} = -G^{(L)} \left[U_-^{(L-1)}\right]^{-1} \tag{1.34}$$

$$C_-^{(L)} = -\tilde{G}^{(L)} \left[U_+^{(L-1)}\right]^{-1} \tag{1.35}$$

where $\tilde{G}$ denotes the conjugate transpose of G.

The recursions in the forward and backward autoregressive coefficients now become

$$\left.\begin{aligned} A_0^{(L)} &= I \\ A_\ell^{(L)} &= A_\ell^{(L-1)} - G^{(L)} \, [U_-^{(L-1)}]^{-1} \, B_{\ell-1}^{(L-1)} \\ A_1^{(L)} &= -G^{(L)} \, [U_-^{(L-1)}]^{-1} \end{aligned}\right\} \qquad (1.36)$$

and

$$\left.\begin{aligned} B_0^{(L)} &= I \\ B_\ell^{(L)} &= B_\ell^{(L-1)} - \tilde{G}^{(L)} \, [U_+^{(L-1)}]^{-1} \\ B_L^{(L)} &= -\tilde{G}^{(L)} \, [U_+^{(L-1)}]^{-1} \end{aligned}\right\} \qquad (1.37)$$

If $e_+^{(L)}$, $e_-^{(L)}$ represent the Lth order forward and backward prediction errors respectively, that is,

$$e_+^{(L)}(n) = x(n) + \sum_{\ell=1}^{L} A_\ell^{(L)} \, x(n-\ell) \qquad n=L+1,\ldots,N \qquad (1.38)$$

and

$$e_-^{(L)}(n) = x(n) + \sum_{\ell=1}^{L} B_\ell^{(L)} \, x(n+\ell) \qquad n=1,2,\ldots,N \qquad (1.39)$$

then substituting (1.35), (1.36) in these relations results in

$$e_+^{(L)}(n) = e_+^{(L-1)}(n) - G^{(L)} \, [U_-^{(L-1)}]^{-1} \, e_-^{(L-1)}(n-L)$$

$$n=L+1,\ldots,N \qquad (1.40)$$

and

$$e_-^{(L)}(n) = e_-^{(L-1)}(n) - \tilde{G}^{(L)} [U_+^{(L-1)}]^{-1} \, e_+^{(L-1)}(n+L)$$

$$n=1,2,\ldots,N-L \qquad (1.41)$$

The following prediction error matrices are defined:

$$
\left.
\begin{aligned}
W_{++}^{(L)} &= \sum_{n=L+1}^{N} e_{+}^{(L)}(n)\ \tilde{e}_{+}^{(L)}(n) \\
W_{--}^{(L)} &= \sum_{n=1}^{N-L} e_{-}^{(L)}(n)\ \tilde{e}_{-}^{(L)}(n) \\
W_{+-}^{(L)} &= \sum_{n=L+1}^{N} e_{+}^{(L)}(n)\ \tilde{e}_{-}^{(L)}(n-L) \\
W_{-+}^{(L)} &= \sum_{n-1}^{N-L} e_{-}^{(L)}(n)\ \tilde{e}_{+}^{(L)}(n+L) = \tilde{W}_{+-}^{(L)}
\end{aligned}
\right\} \quad (1.42)
$$

It is shown by substituting (1.39), (1.40) in these that

$$
W_{++}^{(L)} = W_{++}^{(L-1)} = G^{(L)}\ [U_{-}^{(L-1)}]^{-1}\ W_{--}^{(L-1)}\ [\tilde{U}_{-}^{(L-1)}]^{-1}\ \tilde{G}^{(L)}
$$

$$
- W_{+-}^{(L-1)}\ [\tilde{U}_{-}^{(L-1)}]^{-1}\ \tilde{G}^{(L)} - G^{(L)}\ [U_{-}^{(L-1)}]^{-1}\ W_{-+}^{(L-1)} \qquad (1.43)
$$

and

$$
W_{--}^{(L)} = W_{--}^{(L-1)} + \tilde{G}^{(L)}[U_{+}^{(L-1)}]^{-1}\ W_{++}^{(L-1)}\ [\tilde{U}_{+}^{(L-1)}]^{-1}\ G^{(L)}
$$

$$
- W_{-+}^{(L-1)}\ [\tilde{U}_{+}^{(L-1)}]^{-1}\ G^{(L)} - \tilde{G}^{(L)}\ [U_{+}^{(L-1)}]^{-1}\ W_{+-}^{(L-1)} \qquad (1.44)
$$

The matrix $G^{(L)}$ is then estimated by minimizing the trace of the sum of $W_{++}^{(L)}$ and $W_{--}^{(L)}$. Upon some matrix algebra manipulation, an equation for $G^{(L)}$ is obtained as follows:

$$\{[U_+^{(L-1)}]^{-1} W_{++}^{(L-1)} [\tilde{U}_+^{(L-1)}]^{-1}\} G^{(L)} = G^{(L)} \{[U_-^{(L-1)}]^{-1} W_{--}^{(L-1)} [\tilde{U}_-^{(L-1)}]^{-1}\}$$

$$= [U_+^{(L-1)}]^{-1} W_{+-}^{(L-1)} + W_{+-}^{(L-1)} [\tilde{U}_-^{L-1)}]^{-1} \qquad (1.45)$$

Matrix equation of this form can be solved by a numerical procedure given by Lacoss and Shakal (1976).

The extended Burg spectral estimator for multiple channel signal processing will now be summarized in the following algorithm:

(i) Prediction errors and error matrices are initialized.

$$e_+^{(0)}(n) = x(n)$$

$$e_-^{(0)}(n) = x(n)$$

$$W_{++}^{(0)} = W_{--}^{(0)} = \hat{R}(0) \ , \ W_{+-}^{(0)} = \hat{R}(1)$$

$$U_+^{(0)} = U_-^{(0)} = \hat{R}(0)$$

(ii) $G^{(L)}$ is obtained by solving the equation

$$\{[U_+^{(L-1)}]^{-1} W_{++}^{(L-1)} [\tilde{U}_+^{(L-1)}]^{-1}\} G^{(L)} + G^{(L)} \{[U_-^{(L-1)}]^{-1} W_{--}^{(L-1)} [\tilde{U}_-^{(L-1)}]^{-1}\}$$

$$= [U_+^{(L-1)}]^{-1} W_{+-}^{(L-1)} + W_{+-}^{(L-1)} [\tilde{U}_-^{(L-1)}]^{-1}$$

(iii) Reflection coefficients are computed.

$$C_+^{(L)} = -G^{(L)} [U_-^{(L-1)}]^{-1}$$

$$C_-^{(L)} = -\tilde{G}^{(L)} [U_+^{(L-1)}]^{-1}$$

(iv) Forward and backward autoregressive coefficients are generated.

$$A_0^{(L)} = I \qquad B_0^{(L)} = I$$

$$A_\ell^{(L)} = A_\ell^{(L-1)} + C_+^{(L)} B_{L-\ell}^{(L-1)} \qquad B_\ell^{(L)} = B_\ell^{(L-1)} + C_-^{(L)} A_{L-\ell}^{(L-1)}$$

$$A_L^{(L)} = C_+^{(L)} \qquad B_L^{(L)} = C_-^{(L)}$$

$$\ell = 1,2,\ldots, L-1$$

(v) Prediction errors updated.

$$e_+^{(L)}(n) = e_+^{(L-1)}(n) - G^{(L)} [U_-^{(L-1)}]^{-1} e_-^{(L-1)}(n-L) \quad n = L+1,\ldots,N$$

$$e_-^{(L)}(n) = e_-^{(L-1)}(n) - \tilde{G}^{(L)} [U_+^{(L-1)}]^{-1} e_+^{(L-1)}(n+L) \quad n = 1,2,\ldots,N-L$$

(vi) Prediction error matrices are updated.

$$W_{++}^{(L)} = \sum_{n=L+1}^{N} e_+^{(L)}(n)\, \tilde{e}_+^{(L)}(n)$$

$$W_{--}^{(L)} = \sum_{n=1}^{N-L} e_-^{(L)}(n)\, \tilde{e}_-^{(L)}(n)$$

$$W_{+-}^{(L)} = \sum_{n=L+1}^{N} e_+^{(L)}(n)\, \tilde{e}_-^{(L)}(n-L)$$

$$U_+^{(L)} = [I - C_+^{(L)} C_-^{(L)}]\, U_+^{(L-1)}$$

$$U_-^{(L)} = [I - C_-^{(L)} C_+^{(L)}]\, U_-^{(L-1)}$$

and step (ii) is repeated for the next iteration.

I.5 Autoregressive Spectra

If the data to be processed is truly autoregressive of finite order, or an infinite number of autoregressive coefficients are estimated, the prediction error, or innovation, is by theory a white process. This white process can be regarded as the output of a linear, time-invariant system acting on the data. The impulse response of this system is then given by the estimated autoregressive coefficients. That is,

$$e(n) = A_n * x(n) \tag{1.46}$$

where * denotes convolution of two sequences. The z-transform of equation (1.46) is

$$E(z) = A\ (z)\ X(z) \tag{1.47}$$

The multivariate power spectra are defined by

$$W(e^{i\omega}) = E(e^{i\omega})\ \tilde{E}(e^{i\omega}) \tag{1.48}$$

and

$$S(e^{i\omega}) = X(e^{i\omega})\ \tilde{X}(e^{i\omega}) \tag{1.49}$$

where a tilda denotes the conjugate transpose of a vector or matrix.

Substituting these in (1.47) results in

$$W(e^{i\omega}) = A(e^{i\omega})\ S(e^{i\omega})\ \tilde{A}(e^{i\omega}) \tag{1.50}$$

Since e(A) is a white process, $W(e^{i\omega})$ = constant and (1.50) becomes

$$S(e^{i\omega}) = [A\ (e^{i\omega})]^{-1}\ W[\tilde{A}\ (e^{i\omega})]^{-1} \qquad (1.51)$$

The Lth order spectral estimate of $S(e^{i\omega})$ is therefore given by

$$S^{(L)}(e^{i\omega}) = [A^{(L)}(e^{i\omega})]^{-1}\ U_{+}^{(L)}\ [\tilde{A}^{(L)}(e^{i\omega})]^{-1} \qquad (1.52)$$

II. Empirical Study of Multivariate Burg Estimator

II.1 Introduction

Various aspects of Multivariate Burg Spectral Estimator will be investigated in this chapter. Specifically, the accuracy and resolution of estimated spectra will be evaluated. Several problems associated with applications of the Burg Estimator are also discussed. These include the specific form of the spectra to be analyzed, the determination of an optimal autoregressive order, and the effect of additive white noises on the correctness of an autoregressive model for the data.

Although theoretical study of the properties of autoregressive spectral estimator has been attempted by several authors (Kromer, 1970), a complete and satisfactory theory has not yet been developed. Therefore, the investigation undertaken in the present paper is strictly empirical. Owing to the large variety of data and their respective spectra that have to be processed and analyzed in all disciplines of science, no statement based on results of any empirical investigation can be conclusive. Also, the lack of a quantitative measure of performance imposes a serious handicap in evaluating, interpreting, and comparing results. Therefore, conclusions derived from these studies are unavoidably general and subjective.

Most often, interpretation of results can be affected by any apriori knowledge of the data processed. Consequently, studies based on certain idealized data types can be very beneficial. Other than illuminating various characteristics of the Multivariate Burg Estimator under a well controlled and adjustable experimental environment, these studies can

also provide insights and guidance in understanding and interpreting spectra of more complicated nature.

In general, a spectrum can be decomposed into a broad band component and a narrow band component. The boundary between these two catagories is nevertheless not too well defined. However, the extreme forms of these can be identified. On one side of the scale, there is the perfectly flat spectrum of a white process. And on the other end, one has the discrete line spectrum of signal composing of multiple sinusoids. These two ideal test data will be used extensively by the empirical studies described in this chapter.

II.2 Representations of cross spectrum

The first problem encountered by spectral analysis is necessarily the form of spectrum to be analyzed. In the analysis of power spectra, which do not have an imaginary part, the problem does not exist. This is no longer true in the analysis of cross spectra. A cross spectrum is in general a complex valued function of frequency. And it is well known that a non-zero complex variable can be represented by its real and imaginary components, or amplitude and phase components. It will be demonstrated later that amplitude and phase spectra possess distinctive characters intimately related to the original data, and are therefore used most often by spectral analysts.

Several other spectra can be computed from the amplitude and phase spectra. If a complex valued cross spectrum between channels i and j is represented by

$$P_{ij}(\omega) = A_{ij}(\omega) \exp[i \Phi_{ij}(\omega)]$$

where $A_{ij}(\omega)$ is the amplitude spectrum, and $\Phi_{ij}(\omega)$ is the phase spectrum, a coherence spectrum is defined by

$$C_{ij}(\omega) = P_{ij}(\omega)/[P_{ii}(\omega)P_{jj}(\omega)]^{\frac{1}{2}}$$

where $P_{ii}(\omega)$ and $P_{jj}(\omega)$ are the power spectra of channels i and j respectively. A phase delay spectrum is defined by

$$D_{ij}(\omega) = \Phi_{ij}(\omega)/\omega$$

and a group delay spectrum is defined by

$$G_{ij}(\omega) = \frac{d}{d\omega} \Phi_{ij}(\omega)$$

provided such a derivative exists. The significance of these spectra will be made clear in the next two sections.

II.3 Cross spectrum of shifted broad band signals

Consider the following model for two channels of data

$$x_1(t) = s(t) + n_1(t)$$

$$x_2(t) = as(t+d) + n_2(t)$$

where s(t) is a random process with non-zero power spectrum, $n_1(t)$ and $n_2(t)$ are uncorrelated white noises, a is a scaling factor, and d is the relative shift of the signals. The cross correlation function is then given by

$$\rho_{12}(\tau) = E\{x_1(t)x_2(t+\tau)\}$$

with $E\{\cdot\}$ denoting the expected value operator. If the autocorrelation function of the signal $s(t)$ is represented by $\rho_{ss}(t)$ and its power spectrum by $P_{ss}(\omega)$, then it is easily verified that

$$\rho_{12}(\tau) = a\rho_{ss}(\tau+d)$$

The cross spectrum is therefore given by

$$P_{12}(\omega) = \int_{-\infty}^{\infty} \rho_{12}(\tau)e^{-i\omega\tau}d\tau$$

$$= \int_{-\infty}^{\infty} a\rho_{ss}(\tau)e^{i\omega d}e^{-i\omega\tau}d\tau$$

$$= a\, P_{ss}(\omega)\, e^{i\omega d}$$

It is obvious that the amplitude spectrum in this case is just the power spectrum of the signal multiplied by the scaling factor. More important is the linear phase spectrum typical of this model. If the power of the white noises $n_1(t)$ and $n_2(t)$ are denoted by N_1 and N_2 respectively, then

$$P_{11}(\omega) = P_{ss}(\omega) + N_1$$

$$P_{22}(\omega) = a^2P_{ss}(\omega) + N_2$$

and the coherence is therefore

$$C_{12}(\omega) = \frac{aP_{ss}(\omega)}{(P_{ss}(\omega)+N_1)^{\frac{1}{2}}(a^2P_{ss}(\omega)+N_2)^{\frac{1}{2}}}$$

$$= \left(1 + \frac{N_1}{P_{ss}(\omega)}\right)^{-\frac{1}{2}} \left(1 + \frac{N_2}{a\,P_{ss}(\omega)}\right)^{-\frac{1}{2}}$$

The special case when a=1, and $N_1=N_2=N$ results in

$$C_{12}(\omega) = \frac{1}{1+N/P_{ss}(\omega)}$$

The phase delay is

$$D_{12}(\omega) = \omega d/\omega = d$$

and the group delay is

$$G_{12}(\omega) = \frac{d}{d\omega}(\omega d) = d$$

II.4 Cross spectrum of multiple sinusoids

Consider the following model for two channels of data

$$x_1(t) = \sum_{j=1}^{n} a_j \exp[i\nu_j(t+\alpha_j)] + n_1(t)$$

$$x_2(t) = \sum_{j=1}^{n} b_j \exp[1\nu_j(t+\beta_j)] + n_2(t)$$

with $\nu_{-j}=-\nu_j$, $a_{-j}=a_j$, $b_{-j}=b_j$, $\alpha_{-j}=\alpha_j$, $\beta_{-j}=\beta_j$, and $n_1(t)$, $n_2(t)$ are uncorrelated white noises with powers N_1 and N_2 respectively. If expected value is computed using time averaging, the cross correlation function is then given by

$$\rho_{12}(\tau) = \int_{-\infty}^{\infty} x_1(t)x_2(t+\tau)dt$$

$$= \sum_{i=-n}^{n} \sum_{j=-n}^{n} a_i b_j \exp[i(\nu_i\alpha_i+\nu_j\beta_j)] \exp(i\nu_j\tau)$$

$$\int_{-\infty}^{\infty} \exp[i(\nu_i+\nu_j)t]dt$$

$$= \sum_{j=-n}^{n} a_j b_j \exp[i\nu_j(\beta_j-\alpha_j)] \exp(i\nu_j\tau)$$

The cross spectrum is

$$P_{12}(\omega) = \int_{-\infty}^{\infty} \rho_{12}(\tau)\, e^{-i\omega\tau} d\tau$$

$$= \sum_{j=-n}^{n} a_j b_j \exp[i\nu_j(\beta_j-\alpha_j)]\, \delta(\omega-\nu_j)$$

where $\delta(\omega-\nu_j)$ is the Dirac delta function in the frequency ω. It can be seen that the cross spectrum is likewise discrete, with spectral lines at frequencies of the signal power spectrum. At these frequencies, the phase is the difference between the corresponding time delays of the same component in the two channels, multiplied by the frequency. The power spectra of these channels are

$$P_{11}(\omega) = \sum_{j=-n}^{n} a_j^2 \, \delta(\omega-\nu_j) + N_1$$

$$P_{22}(\omega) = \sum_{j=-n}^{n} b_j^2 \, \delta(\omega-\nu_j) + N_2$$

The coherence at frequency ν_j is therefore

$$C_{12}(\nu_j) = \frac{a_j b_j \, \delta(\nu-\omega_j)}{[a_j^2 \, \delta(\omega-\nu_j)+N_1]^{\frac{1}{2}}[b_j^2 \, \delta(\omega-\nu_j)+N_2]^{\frac{1}{2}}}$$

$$\frac{a_j b_j \, \delta(\omega-\nu_j)}{a_j b_j \, \delta(\omega-\nu_j)}$$

$$= 1.$$

The phase delay at frequency ν_j is

$$D_{12}(\nu_j) = \nu_j(\beta_j-\alpha_j)/\nu_j = \beta_j-\alpha_j$$

The group delay is obviously not defined for this type of data.

II.5 Accuracy and resolution of Multivariate Burg Estimator

The results derived in the previous sections will now be verified using a two dimensional Burg Estimator. The resulting cross spectral amplitude, coherence, phase, and phase delay estimates are compared with those predicted by theory. The ability of this estimator to resolve spectral peaks close to one another in a cross spectrum is also examined.

Test data for the experiments are generated from a system depicted in Figure 1a. In one experiment, the signal s(t) was 100 samples of a

white process with variance 100., and the time shift was chosen to be -3 samples. Uncorrelated white noises of unit variances were then added to produce the data as shown in Figure 2.

Cross spectra were generated using an autoregressive order of 3. The amplitude, coherence, phase, and phase delay spectra are shown in Figures 3 to 6 respectively. It can be observed that the maximum and minimum cross powers are within 3 dB of the true value of 100. Values of the estimated coherence at all frequencies are well above 0.97, with maxima close to 0.99, which is the true coherence for a signal to noise ratio of 100. Phase estimate is strikingly linear, with a slope of exactly -3 samples computed at a frequency of 0.5 cycles per sample. Phase delays computed at other frequencies vary from this value by less than 0.03 sample.

The high accuracy of the Burg estimator is well demonstrated. Indeed, the superiority of this estimator over the windowing method is clearly illustrated by Figures 7 and 8. In these figures, coherence and phase estimated by Burg's method are compared with those computed by taking the Fourier transform of a windowed estimate of the cross correlation sequence. Two coherence and phase spectra are shown, one with a six sample Bartlett window, and the other a 100 sample Bartlett window. The six sample Bartlett spectral estimates are evidently erroneous. Even the 100 sample Bartlett estimates cannot reach the same accuracy achieved by the Burg estimates.

The same procedure has also been applied on data composing of multiple sinusoids. Referring to Figure 1b, the signal s(t) now contains 100 samples of two sinusoids with frequencies 0.04 and 0.05 cycles per sample respectively. The peak amplitudes of the harmonic components are

10. and 5. in the first channel, 5. and 10. in the second channel; and time shifts at both frequencies of the second channel are both -3 samples. Uncorrelated white noises of unit variances are added to produce the data as shown in Figure 9.

Cross spectral amplitude, coherence, phase, and phase delay spectra computed from a Burg Estimator of order 10 are shown in Figures 10 through 13. The two spectral peaks are well resolved. Although the peak values at frequency of 0.04 is smaller than that at frequency of 0.05, the first spectral line has a larger width than the other. Consequently the area within each spectral line can be equal.

Coherences at frequencies 0.04 and 0.05 are practically equal to one. However, they are not as well resolved as the cross amplitude. Phases at these frequencies are -46° and -55°, which are equivalent to time shifts of -3.2 and -3.05 respectively. The errors are within one sample and are generally acceptable.

It has long been recognized that one important property of single channel Burg estimator is its ability to reproduce a sharply peaked spectral line in estimating power spectrum of a truncated sinusoid. Smearing of the spectral peak by the conventional method as a result of convolution with a window function is not present in a Burg estimate. That the same statement can be said about Burg estimate of cross spectrum of a truncated sinusoid is demonstrated in Figures 14, 15. Three estimates of cross spectrum between two truncated sinusoids with the same frequency of 0.1, but with phases differ by 108° at the frequency are shown. Of these, the Burg estimate of order 4 obviously resembles most of a line spectrum. The 100 sample Bartlett estimator shows a much smaller peak value, and a number of side lobes can be seen adjacent to the main lobe.

The 8 sample Bartlett estimate is too flat to be considered a peak. Phase estimate from the fourth order Burg estimator and the 100 sample Bartlett estimator generally agree within the frequency range where cross power is substantial. The phase estimate from the 8 sample Bartlett estimator is also approximately correct, except that it is 360° out of phase with the others.

II.6 Determination of optimal autoregressive order

A problem of paramount importance facing a user of either the univariate or multivariate Burg estimator is the choice of an autoregressive order. Since spectra can be computed at any order less than the number of data samples being processed, the compatibility of these spectra is of major concern to an analyst. Two situations should be considered separately.

The basic principle of Burg's method is to fit an autoregressive model of finite order to the data with minimum squared residual error. If the data to be processed is a true autoregressive process of finite order, then residuals computed from orders less than this true order would have gross error, while those computed from orders at or larger than this true order would be samll and the same. On the other hand, if the data is not autoregressive, all residuals computed from a fitted autoregressive model of any order are always decreasing. As a consequence, the conventional definition of an optimal spectral estimate is that one whose squared prediction error is minimum. Specifically, autoregression coefficients are computed recursively until the residual error, which is a monotonically decreasing function of regression order, stabilizes and ceases to decrease substantially.

However, there is a second contribution to the error in the spectral estimate. This contribution is the consequence of estimating the autoregressive coefficients from a finite number of data samples. The result is a statistical deviation of the estimated coefficients from their true values. As the autoregression order increases, more coefficients have to be estimated from the same number of data samples, and consequently the statistical deviation of those estimates are expected to increase.

One way to measure the combined effect of these two contributions is the use of a criterion called Final Prediction Error (Akaike, 1971). This criterion is

$$FPE(p) = |\hat{U}_p| \left(\frac{n+1+mp}{n-1-mp}\right)^m$$

where p is the order of the autoregression, m the dimensionality of the data, and n the number of data samples. $|\hat{U}_p|$ is the determinant of the residual error matrix as defined in Chapter II. An order of the autoregression is chosen to be that which minimizes this criterion.

A similar approach was also proposed by Akaike. In this scheme, it is assumed that there is a true order p such that the data is autoregressive. Then, at each successive order, a criterion called Akaike Information Criterion (Akaike, 1974) can be computed. Conceptually, this can be regarded as a measure of information available to the estimation of the autoregression coefficients. Specifically it is

$$AIC(p) = \log |\hat{U}_p| + \frac{2m^2p}{n}$$

A third approach was proposed by Parzen, This approach does not assume that the data is finite order autoregressive, or even autoregres-

sive at all, provided that the z-transform of the data model impulse response converges. A theorem was given (Parzen, 1976) which states that, for a stationary time series Y(t) with continuous positive spectral density function, there is always a frequency transfer function denoted

$$G_{\infty}(z) = I + A_1 z + A_2 z^2 + \ldots$$

and a white process $\varepsilon(t)$ such that

$$Y(t) + A_1 Y(t-1) + A_2 Y(t-2) + \ldots = \varepsilon(t)$$

An approximation to the true transfer function A(z) is usually obtained by terminating it at a finite order. This approximation is referred to as the Autoregressive Transfer Function Approximator Converging to the Truth (ARTFACT). A criterion, designated Criterion of Autoregressive Transfer Function (Parzen, 1975) is introduced which provides an estimate of the mean square overall error of ARTFACT. This is defined as

$$CAT(p) = \text{trace } \{\frac{m}{n} \sum_{j=1}^{P} \hat{\hat{U}}_j^{-1} - \hat{\hat{U}}_p^{-1}$$

when

$$\hat{\hat{U}}_j = \frac{n}{n-jm} \hat{U}_j \quad .$$

Aided by the above criterion, the problem of determining an optimal order reduces to choosing that one which minimizes the FPE, AIC, or CAT. In many experiments conducted with synthesized data, the orders thus chosen generally agree with one another. Nevertheless, exceptions to this do occur, and in these cases, CAT is found to be a more reliable indicator of an optimal order.

To illustrate the use of these three criterions, two channels of white processes generated as described in section III.5 were processed. The two white processes have a relative time shift of 10 samples. Residual error matrices were estimated recursively up to a maximum order of 45, and the three criterions mentioned above were computed at each order of the recursion. These are plotted in Figures 16 to 18.

The data used in this experiment is truly autoregressive of order 10. This statement is evidently supported by a study of the three criterions, which all show a significant reduction in value at the autoregressive order of 10. However, FPE continues to decrease gently until order 38, while both AIC and CAT start rallying upwards immediately after order 10. It is generally observed that FPE, and occasionally AIC, has a flat bottom ranging over a large number of orders with very small difference in value. This shows that FPE is a less relaible indicator of an optimal order than either AIC or CAT. The same phenomenon can be observed when data seemingly not autoregressive are analyzed and discussed in Chapter IV.

One widely acclaimed advantage of autoregressive estimator is its ability to provide good estimates with a relatively small order, thus requiring less computation. Recalling from Chapter I that an autoregressive model of order p is equivalent to using autocorrelation information of the data up to a maximum lag of p samples. A comparison of a pth order Burg's estimate with a Bartlett windowed transform of the autocorrelation sequence with maximum lag of p has been made for p ranging from 2 to 50, for the two shifted white processes as described earlier in this section. The coherence and phase estimate are shown in Figures 19 and 20.

It is clearly demonstrated in the figures how Burg's estimate rapidly converge to the true spectra at and beyond the autoregressive order of 10. Phase estimates from the other method are comparable to Burg's estiamtes only when autocorrelation lags of 20 and longer are used. Coherence estimates, nevertheless, seem incapable of achieving the same level of accuracy as Burg's estimates even when 50 samples of the autocorrelation sequence are used.

II.7 Effect of noise on error in autoregressive model

As mentioned above, one major source of error in estimating spectra using Burg's method is caused by errors in fitting the data with a finite order autoregressive model. For certain types of data, this model may not be appropriate. One example that occurs often in data processing is autoregressive processes contaminated by additive white noises. Indeed, it has been shown (Pagano, 1974) that this kind of data can be modelled as autoregressive moving average. One would expect that the larger the level of the noise, the larger the deviation from a true spectrum will be. Figure 21 shows coherence and phase estimates of a two channel shifted white signal with five levels of additive white noises. It is obvious from the figure that the accuracy of the estimates deteriorates as the noise power increases. The coherence estimate is particularly sensitive to the presence of noise. A comparison with estimates obtained from a 100 sample Bartlett estimator indicates that, even in the presence of noise, Burg's estimator can still provide more accurate coherence and phase estimates than the other estimator.

II.8 Conclusions

Empirical studies have been carried out on synthesized data to evaluate certain characteristics of the Multivariate Burg Estimator. High accuracy and good resolution of this estimator were verified by results of these studies. Superiority of this estimator over the convention windowed and transformed method was well demonstrated. Burg's estimates of spectrum were found to be stable in the presence of noise. The determination of an optimal autoregressive order was achieved utilizing the criterions of FPE, AIC, and CAT. It has also been shown that Burg Estimator requires a shorter lag of autocorrelation sequence than the windowed and transformed method.

III. Spectral Analysis of Ocean Core Geological Data

III.1 Introduction

One of the many explanations of Pleistocene ice ages attributes climatic changes to variations in the Earth's orbital geometry. One strategy of testing this orbital theory is to treat secular changes in the orbit as a forcing function of a system whose output is the geological record of climate (Hays, et al., 1976). With the assumption of constant sedimentation rate, spectral analysis can be applied to the geological data. The objective of the analysis is to measure any significant frequencies existing in the data. These significant frequencies are compared to the Earth's obliquity and precession frequencies. This comparison will furnish a test on the validity of the orbital theory. Phase relation between each of the geological records can also be estimated at these frequencies. These are also compared with values obtained from other methods. The purpose of this chapter is to apply the Multivariate Burg Estimator as a new spectral analysis technique in such a testing procedure.

III.2 Sources of data

Three climatically sensitive geological variables were measured from two deep-sea cores located in the Southern Indian Ocean. A continuous record of the last 450,000 years can be obtained. With an accumulation rate of approximately 3 cm/1,000 yrs, resolution of climatic fluctuations is well below 20,000 years. Identical samples were analyzed for the three variables at 10 cm intervals throughout both cores.

The first geological variable is an estimate of summer sea-surface temperatures at the core site, derived from a statistical analysis of radiolarian assemblages. Accuracy of these estimates is ±1.5°C.

The second geological variable is the oxygen isotopic composition of planktonic foraminifera. Down-core variations in $\delta\ O^{18}$ reflect changes in oceanic isotopic composition, caused by the waxing and waning of northern hemisphere ice sheets. It ties the core record into a global stratigraphic framework based on isotopic and geomagnetic variations. Therefore, although this variable was measured in subantarctic cores, it represents a record of the northern hemisphere climate.

The third geological variable is the percent of Cycladorphora davisiana relative to all other radiolaria. A high abundance of this species can be related to the low salinity of summer surface waters. For the analysis procedure described below, 157 samples of each variable were used, representing geological records from 3,000 to 471,000 years ago. The means of each record were estimated and subtracted from the data. Each record was also scaled independently so that the variance of all three records are the same and equal to unity.

III.3 Analysis procedures

The procedure with which power and cross spectra of the three geological data were estimated can be divided into three steps. First, reflection coefficients (or partial correlation coefficients), and residual (or prediction) error matrices were estimated at each order of the recursion up to a maximum order generally taken to be the number of data samples in each record divided by twice the dimension of the data, that is, twice the number of records being processed.

An autoregressive order analysis was then carried out for each of these reflection coefficients and residual error matrices. Specifically, the three criterions of FPE, AIC, and CAT were computed and plotted as a function of autoregressive order. The optimal order, and consequently the reflection coefficient and residual error matrix used to compute the autoregressive coefficients, was determined by choosing that which indicated sufficiently small values of all three criterions.

Finally, autoregressive coefficients were computed and spectra were generated. The power spectral density of each geological data was displayed and studied. This provides evidence of periodicities, or significant frequencies, existing in the data. The cross power spectral amplitude between two records then indicates those frequency components contributing to the correlation between the corresponding records. The degree of correlation is given by the coherence at those frequencies. A knowledge of the coherence is critical in assigning confidence to the phase information obtained from the phase spectrum at these frequencies.

III.4 Preliminary analysis

The normalized data constructed as described in the last section was processed by the procedure outlined also in the last section. The three geological records are shown in Figure 22. Plots of the AIC and CAT are shown in Figure 23. An abrupt reduction in the values of these criterions at a small order usually indicates the presence of broad band structure, although existence of small narrow band signals is not precluded.

The power spectral densities of the geological records estimated using an autoregressive order of 2 are shown in Figures 24. The spectral

densities are evidently dominated by low frequency components up to a period of 15,000 years. The most dominant frequency is approximately 0.015 cycles/1,000 yrs, or a period of 67,000 years.

The coherence and phase spectra are also shown in Figures 25 and 26. Coherence at this frequency is 0.75, which is generally considered large enough for a reliable estimate of phase informations. On the whole, coherences at frequencies less than 0.035 cycles/1,000 yrs are well above 0.50. The phase of the oxygen ratio record is found to lead the surface temperature record, with the maximum phase lead of 0.68 radians occurring at periods of 50,000 years to 67,000 years, equivalent to a time lead of 5,000 to 7,000 years. The phase of the salinity record also leads the surface temperature record by as much as 0.40 radians occurring at periods of 60,000 to 100,000 years, equivalent to a time lead of 4,000 to 6,000 years. While the salinity record lags the oxygen ratio record by 0.60 radians around a period of 20,000 years, equivalent to a time lag of 2,000 years.

III.5 Analysis of low-passed data

Although analysis at this stage does disclose some preliminary nature regarding the gross structure of the spectral decomposition and correlation of the original data, more detailed analysis is necessary to reveal local structure of the spectra. However, the presence of a broad band background conceals any fine structure otherwise observable in the criterions of FPE, AIC, and CAT. As a consequence, it is not permissible to locate an autoregressive order beyond the one indicative of the broad band background, which would lead to spectra containing more detailed structures than those obtained above.

A solution to the above problem is to destroy the broad band background by low passing the original data. Guided by results of the preliminary analysis, the three geological records were low-passed with a first order Butterworth filter. The cut-off frequency was chosen to be 0.055. As a result, power of components with period shorter than 18,000 years are suppressed, while spectral information of periods longer than 18,000 years remains intact. This low-passed data was then processed by the same procedure as described in section IV.3.

Results of autoregressive order analysis are shown in Figure 27. With the exception of FPE, the abrupt reductions in the other criterions are replaced by more gradually decreasing patterns. It can also be observed that a new optimal autoregressive order of 24 is indicated by both AIC and CAT. However, a more liberal judgement would agree that autoregressive orders ranging from 21 through 30 are acceptable with no substantial difference in values of either AIC of CAT.

In general, if a span of autoregressive orders is available, the smallest one should be chosen in order to eliminate unnecessary complexity in the spectra. Therefore, spectra are estimated using an autoregressive model of order 21. These spectra will be displayed and analysed in the following sections.

III.6 Power spectral densities

The power spectral densities of the three geological data are shown in Figure 28. Only powers at frequencies less than 0.06 cycles/1,000 yrs are plotted. The spectra evidently possess more complicated and finer structures from which several spectral groups can be identified.

Power spectral density of surface temperature record is characterised by five well separated clusters of peaks. The most prominant one occupies the low frequency region of the spectrum with periods ranging from 110 to 75 thousand years. The next three clusters, with dominant periods of approximately 43, 31 and 23 thousand years respectively, have more or less the same power. A smaller, broader peak with period of 18,000 years can be observed at the end of the spectrum.

Presence of the same clusters of spectral peaks is also evident in the power spectral density of oxygen isotope ratio record. The broadness and levels of these clusters are, however, quite different. The range of the first group is reduced to 90 to 75 thousand years. The one around 43,000 yrs is partially occulted by the enhanced broad band background. The spectral groups centered at 31,000 and 23,000 years still stand out very well, while the one with 18,000 yrs is almost invisible. Instead, a new one emerges at a period of approximately 25,000 yrs.

Power spectral density of the salinity record shows a much broader cluster of low frequency peaks. Typically, large power density extends from a period of 200,000 yrs to a period of 75,000 yrs. Strangely enough, though the long period bound of this prominant cluster varies from one geological variable to the others, the short period bound remains close to 75,000 yrs. A slightly smaller spectral group around a dominant period of 60,000 yrs can be seen between the long period band and the one at 43,000 yrs. The spectral group of 31,000 yrs is again present, followed by smaller group of 23,000 yrs, and an even smaller one of 18,000 yrs.

Besides the prominant low frequency cluster, the spectral groups with dominant periods of 43,000 yrs, 31,000 yrs, and 23,000 yrs are also common to all three geological measurements. To understand the significance of these periods, they are compared to the dominant periods of three orbital variables postulated to have affected the Pleistocene ice ages cycle. One of these is the eccentricity of the Earth's orbit, which exhibits an average period of 93,000 years. Obliquity, defined as the angle between the equatorial and ecliptic planes, is also an important variable. Its average period was found to be about 41,000 years. The third variable is called precessional index, which is approximately equal to the Earth-sun distance in June, expressed as a fraction of that distance in 1950. This index has an average period of 21,000 years.

Of the three orbital variables, the eccentricity affects only the total annual insolation (solar radiation received at the top of the atmospheres), while the others also affect the orientation of the Earth relative to the sun. Therefore, eccentricity effect is generally considered unimportant. More important, direct analysis of insolation data (Hays, et al., 1976) indicates that its spectrum is dominated by periods of 41,000 yrs, 23,000 yrs, and 19,000 yrs. This further strengthens the conviction that eccentricity effect is small, if not negligible.

As a consequence, the contribution to low frequency power present in spectra of the geological data is of some unknown origins. The dominant periods of 43,000 yrs and 23,000 yrs obtained from spectral analysis of the geological records agree well enough with the average periods of obliquity and precessional index, and with the insolation data, which is the combined consequence of these. The other dominant period of 31,000 yrs in the geological spectra does not have a counterpart in the orbital data, and remains unexplained in the present study.

III.7 Coherence and phase analysis

Coherence spectra as shown in Figure 29 furnish support for the conclusion that the three geological variables are strongly correlated at the periods of 43,000 yrs and 23,000 yrs. The spectral period of 31,000 yrs is also strongly correlated among the three variables. It is interesting, however, to note that salinity is not as strongly correlated with the other two variables, at the low frequency region, as that between surface temperature and oxygen isotope ratio.

Phase and phase delay spectra are shown in Figures 30 and 31 respectively. The phase shifts and time shifts between any two of the three geological variables can then be deduced. At the spectral period of 43,000 yrs, oxygen isotope ratio leads surface temperature by a phase of 8°, or equivalently, 1,000 yrs. Salinity also leads surface temperature by approximately 25°, or 3,000 yrs. At the spectral period of 23,000 yrs, oxygen isotope ratio again leads surface temperature by about 39°, or 2,500 yrs while salinity lags surface temperature by 16°, or 1,000 yrs.

This tends to imply that obliquity variation produces changes in salinity before it produces changes in the others. Likewise, oxygen isotope ratio is affected by the precessional variation more efficiently. However, these findings do not agree with those obtained by Hays who employed a time domain technique to estimate these phase relations. The discrepancy obviously results from the difference between the method employed in this study and the one used by Hays. Further investigations, probably with an entirely different technique, are needed to verify either of the results.

Also, such an investigation is needed to verify the significance of the apparently unaccountable period of 31,000 yrs. At this period, the oxygen isotopes ratio leads surface temperature by 42°, or 3,500 yrs. And salinity leads surface temperature by 17°, or 1,500 yrs. It seems that if certain physical parameters can be identified with an average period of 31,000 yrs, it would have an earlier effect on the oxygen isotopes ratio (and therefore on the northern hemisphere climate) than the other geological variables.

III.8 Conclusions

Spectral analysis of three gological data has been completed. The results tend to confirm the validity of the Astronomical Theory of Pleistocene ice ages as proposed by Hays and others. Discrepancies exist between phase estimates obtained from this study and that of Hays. These discrepancies, however, are within the errors of the estimation. Further works with better precision are needed to test the accuracy of either estimates.

IV. Conclusions and Discussion

An extension of Burg's method for estimating maximum entropy spectra to multiple channel signal processing has been developed. Implementation of this method into a system of Fortran programs was completed. Empirical studies were carried out on synthesized data to evaluate certain characteristics of the Multivariate Burg Spectral Estimator. Results from these studies indicated that high accuracy and good resolution of spectral estimates can be achieved. The level of accuracy and resolution evident in the Burg estimates are superior to those obtained from the Taper and Transform method using the same maximum lag of autocorrelation information.

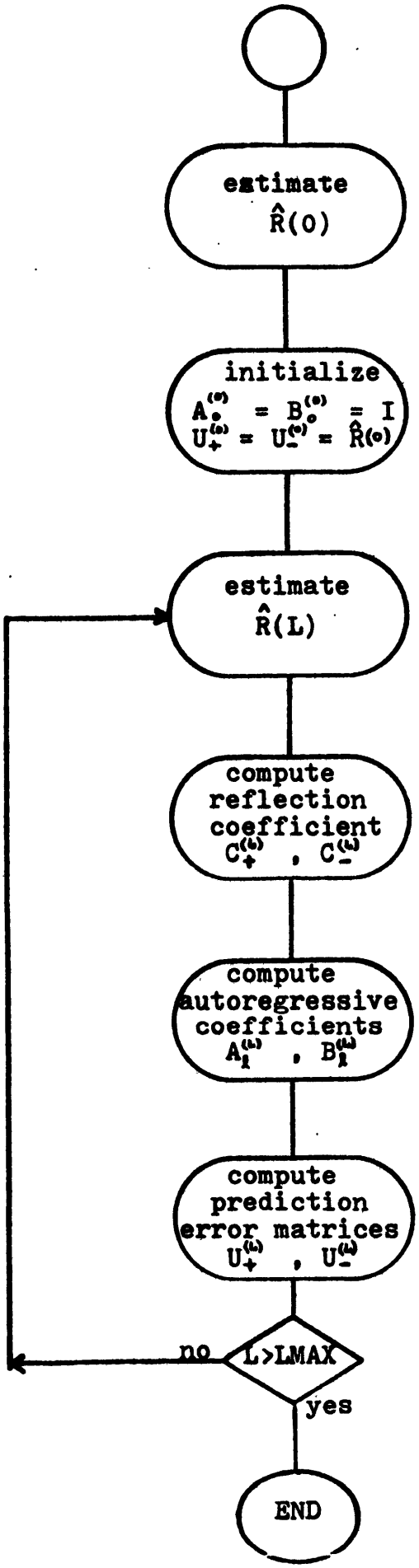
Determination of an optimal autoregressive order has also been discussed. The strategy is to choose an order at which the Final Prediction Error, Akaike's Information Criterion, and Criterion of Autoregressive Transfer Function have minimum values. It has been found that autoregressive orders indicated by the three criterions are in general consistent. Empirical studies seem to show that CAT is a more relaible indicator of autoregressive order in case discrepancy does arise.

Spectral analysis of three geological data utilizing the Multivariate Burg Estimator was conducted to test an Astronomical Theory of Pleistocene ice ages. Frequencies of spectral peaks estimated by this method agree very well with the findings of Hays and others, and therefore offer strong evidence of the orbital-climatic relation as an origin of ice ages. However, phase relations among the three geological variables do not agree with those obtained by Hays. In veiw of the small magnitudes of these estimates, the discrepancies may well be within the errors of the estimation. Although an accurate estimation of these phase relations

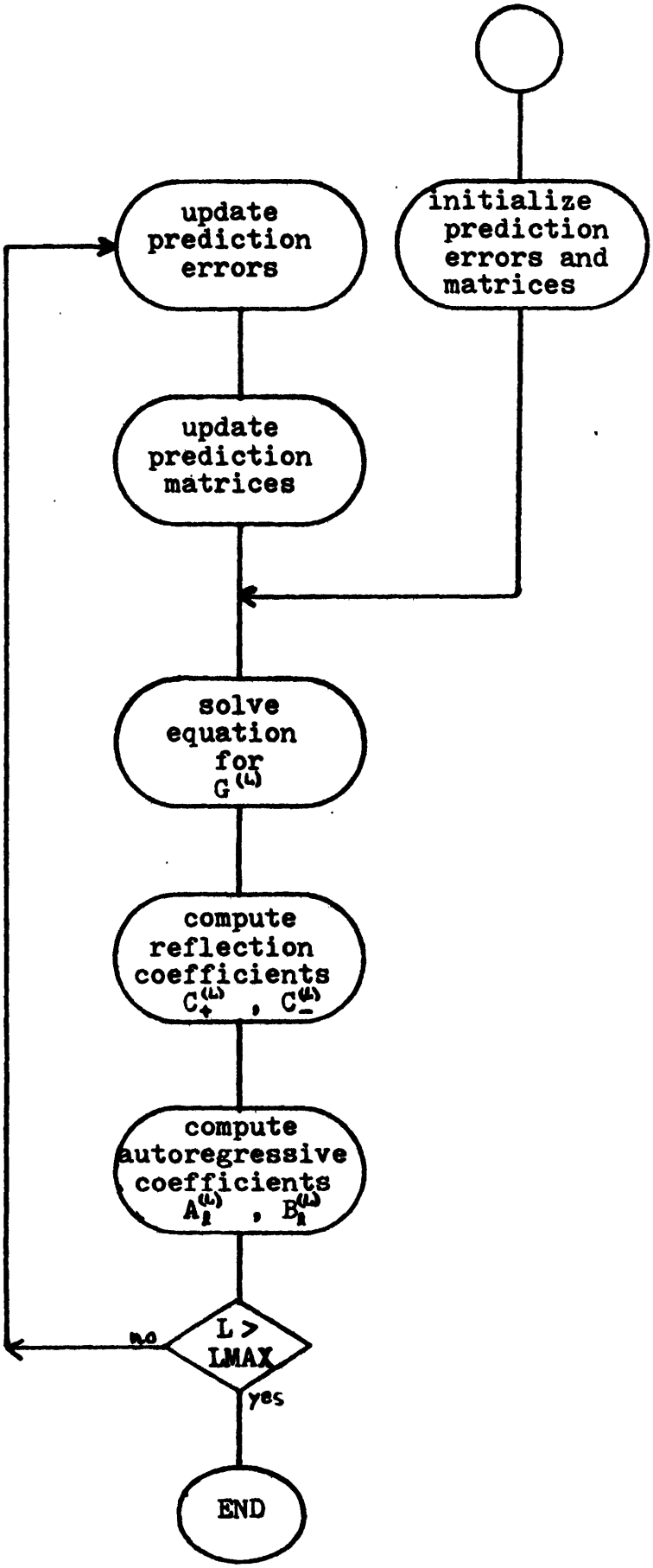
is useful in understanding the mechanism of the Astronomical Theory, it has no direct effect on the validity of the theory.

Owing to the complexity of the field of spectral analysis, and the empirical nature of investigation as reported by the present thesis, these conclusions are far from universal and unchangeable. More experimental studies, both on synthesized data and real data, are definitely needed. Most importantly, a theoretical approach in studying the bias, error, consistency, and efficiency of this estimator needs to be developed. Finally, a spectral estimator of a more general class, like the autoregressive moving average estimator to which the autoregressive estimator belongs as a subclass, requires further development and investigation.

APPENDIX



Flow Diagram of Yule-Walker Algorithm



Flow Diagram of Burg Algorithm

REFERENCES

Akaike, H., "Autoregressive Model Fitting for Control," Ann. Inst. Statist. Math., 23, pp. 163-180, 1971.

Akaike, H., "A New Look at the Statistical Model Identification," IEEE Trans. Auto. Control, AC-19, pp. 716-723, 1974.

Blackman, R. B., and J. W. Tukey, "The Measurement of Power Spectra from the Point of View of Communications Engineering," Dover Press, New York, 1956.

Burg, J. P., "Maximum Entropy Spectral Analysis," paper presented at 37th Annual International SEG Meeting, Oklahoma City, Oklahoma, October 31, 1967.

Burg, J. P., "Maximum Entropy Spectral Estimator," Ph.D. Dissertation, Stanford Univ., Stanford, Calif., 1975.

Capon, J., "High-resolution Frequency-wave Number Spectrum Analysis," Proc. IEEE, 57, pp.1408-1418, 1969.

Hays, J. P., J. Imbrie, and N. J. Shackleton, "Variations in the Earth's Orbit: Pacemaker of the Ice Ages," Science, 194, pp. 1121-1132, 1976.

Jones, R. H., "A Reappraisal of the Periodogram in Spectral Analysis," Technometrics, 7, pp. 531-542, 1965.

Jones, R. H., "Identification and Autoregressive Spectrum Estimation," IEEE Trans. Auto. Control, AC-19, pp. 894-897, 1974.

Jones, R. H., "Multivariate Maximum Entropy Spectral Analysis," Invited paper presented at Applied Time Series Analysis Symposium, Tulsa, Oklahoma, May 14, 1976.

Kavah, M., and G.R. Cooper, "An Empirical Investigation of the Properties of the Autoregressive Spectral Estimator," IEEE Trans. Inform. Theory, IT-22, pp. 313-323, 1976.

Kromer, R., "Asymptotic Properties of the Autoregressive Spectral Estimator, "Ph.D. Dissertation, Stanford Univ., Stanford, Calif., 1970.

Lacoss, R.T., "Data Adaptive Spectral Analysis Methods," Geophysics, 36, pp. 661-675, 1971.

Lacoss, R.T., private communication, 1977.

Lacoss, R.T., and A.F. Shakal, "More $A_1E+EA_2=-D$ and $\dot{X}=A_1X+XA_2+D$, $X(0)=C$," IEEE Trans. Auto. Control, AC-21, pp. 405-406, 1976.

Landers, T.E., and R.T. Lacoss, "Some Geophysical Applications of Autoregressive Spectral Estimates," IEEE Trans. Geoscience Electronics, GE-15, pp.26-31, 1977.

Nuttal, A.H., "Fortran Program for Multivariate Linear Predictive Spectral Analysis, Employing Forward and Backward Averaging," Tech. Report No. 5419, Naval Underwater Systems Center, 1976.

Pagano, M., "Estimation of Models of Autoregressive Signals plus White Noise," Ann. Statist., 1, pp. 99-108, 1974.

Parzen, E., "Statistical Spectral Analysis (Single Channel Case) in 1968," Tech. Report No. 11, Stanford Univ., Stanford, Calif, 1968.

Parzen, E., "Multiple Time Series: Determining the Order of Approximating Autoregressive Schemes," Tech. Report No. 23, State Univ. of New York, Buffalo, New York, 1975.

Parzen, E., "An Approach to Time Series Modeling and Forcasting Illustrated by Hourly Electricity Demands," Tech. Report No. 37, State Univ. of New York, Buffalo, New York, 1976.

Ulrych, T.J., "Maximum Entropy Power Spectrum of Truncated Sinusoids," J.G.R., 77, 1396-1400, 1972.

Ulrych, T.J., and T.N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," Review of Geophysics and Space Physics, 13, pp.183-200, 1975.

Wiggins, R.A., and E.A. Robinson, "Recursive Solution to the Multi-channel Filtering Problem," J.G.R.,70, pp.1885-1891, 1965.

LIST OF FIGURES

13 Phase delay of cross spectrum of multiple sinusoids.

14 Cross spectral amplitudes of single sinusoid. B4 is Burg Estimator of order 4; T4 is Taper and Transform Estimator of lag 4: T50 is Taper and Transform Esitmator of lag 50.

15 Phase spectra of single sinusoid. B4 is Burg Estimator of order 4; T4 is Taper and Transform Estimator of lag 4; T50 is Taper and Transform Estimator of lag 50.

16 Final Prediction Error analysis of broad band signals.

17 Akaike's Information Criterion analysis of broad band signals.

18 Criterion of Autoregressive Transfer Function analysis of broad band signals.

19 Coherence spectra generated by autoregressive estimators of orders 2, 5, 10, 15, 20, 30, 40, 50.

20 Phase spectra generated by autoregressive estimators of orders 2, 5, 10, 15, 20, 30, 40, 50.

21 Coherence and phase spectra for varying signal to noise ratio.

22 Deep ocean core gelolgical data. (a) is surface temperature measurements. (b) is oxygen isotope ratio measurements. (c) is salinity concentration measurements.

23 AIC and CAT analysis of geological data.

24 (a) Power spectral density of surface temperature data.

24 (b) Power spectral density of oxygen isotope data.

24 (c) Power spectral density of salinity data.

25 (a) Coherence spectrum of surface temperature and oxygen isotope data.

25 (b) Coherence spectrum of surface temperature and salinity data.

25 (c) Coherence spectrum of oxygen isotope and salinity data.

26 (a) Phase spectrum of surface temperature and oxygen isotope data.

26 (b) Phase spectrum of surface temperature and salinity data.

26 (c) Phase spectrum of oxygen isotope and salinity data.

27 (a) FPE analysis of low-passed geological data.

27 (b) AIC analysis of low-passed geological data.

27 (c) CAT analysis of low-passed geological data.

28 (a) Power spectral density of low-passed surface temperature data.

28 (b) Power spectral density of low-passed oxygen isotope data.

28 (c) Power spectral density of low-passed salinity data.

29 (a) Coherence spectrum of low-passed surface temperature and oxygen isotope data.

29 (b) Coherence spectrum of low-passed surface temperature and salinity data.

29 (c) Coherence spectrum of low-passed oxygen isotope and salinity data.

30 (a) Phase spectrum of low-passed surface temperature and oxygen isotope data.

30 (b) Phase spectrum of low-passed surface temperature and salinity data.

30 (c) Phase spectrum of low-passed oxygen isotope and salinity data.

31 (a) Phase delay spectrum of low-passed surface temperature and oxygen isotope data.

31 (b) Phase delay spectrum of low-passed surface temperature and salinity data.

31 (c) Phase delay spectrum of low-passed oxygen isotope and salinity data.

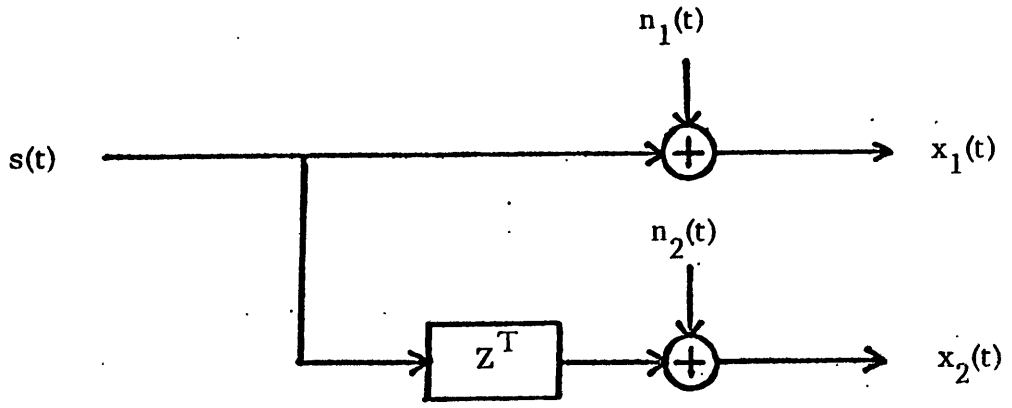


Figure 1(a)

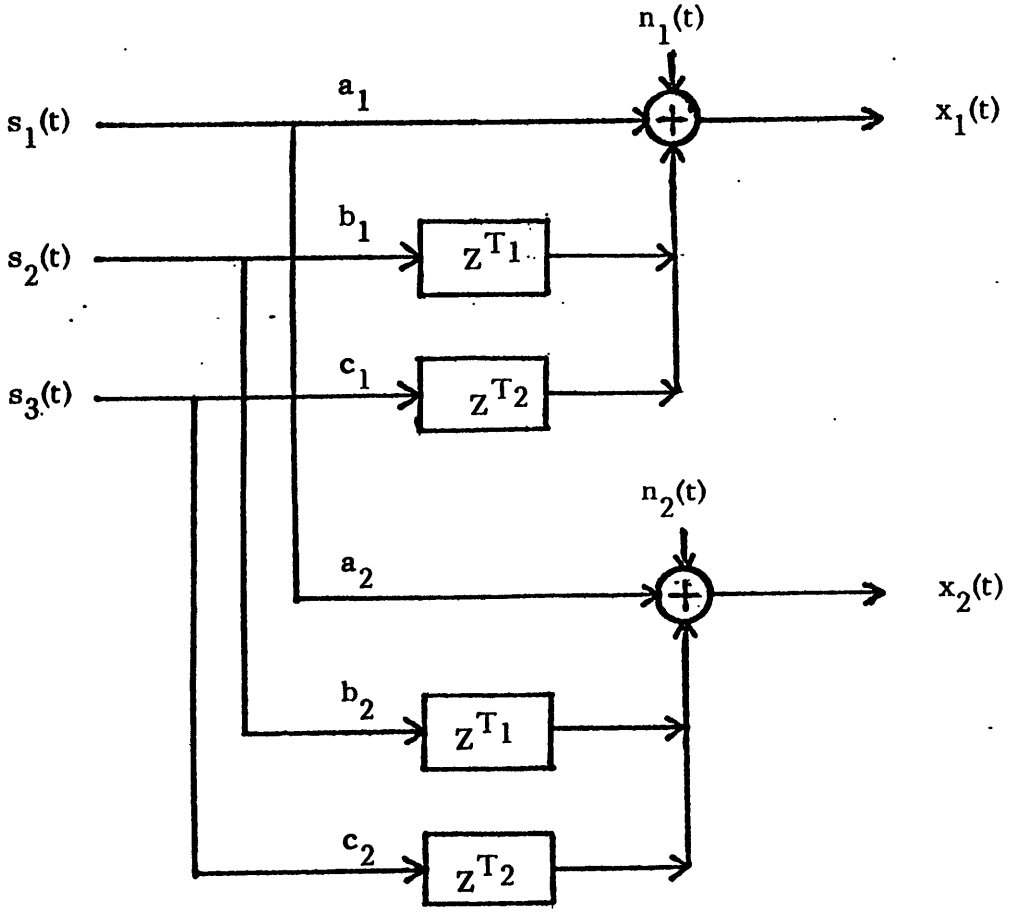


Figure 1(b)

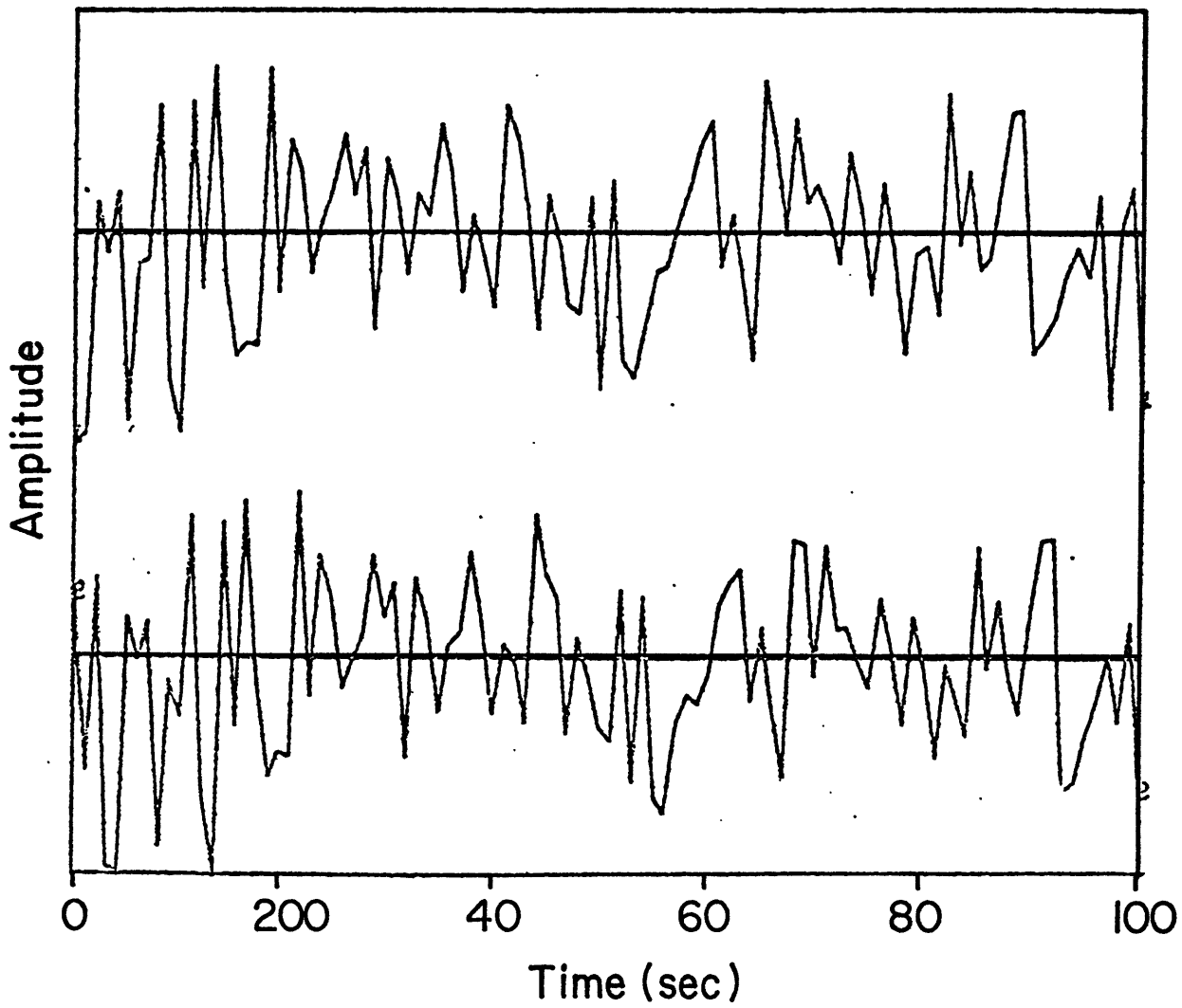


Figure 2

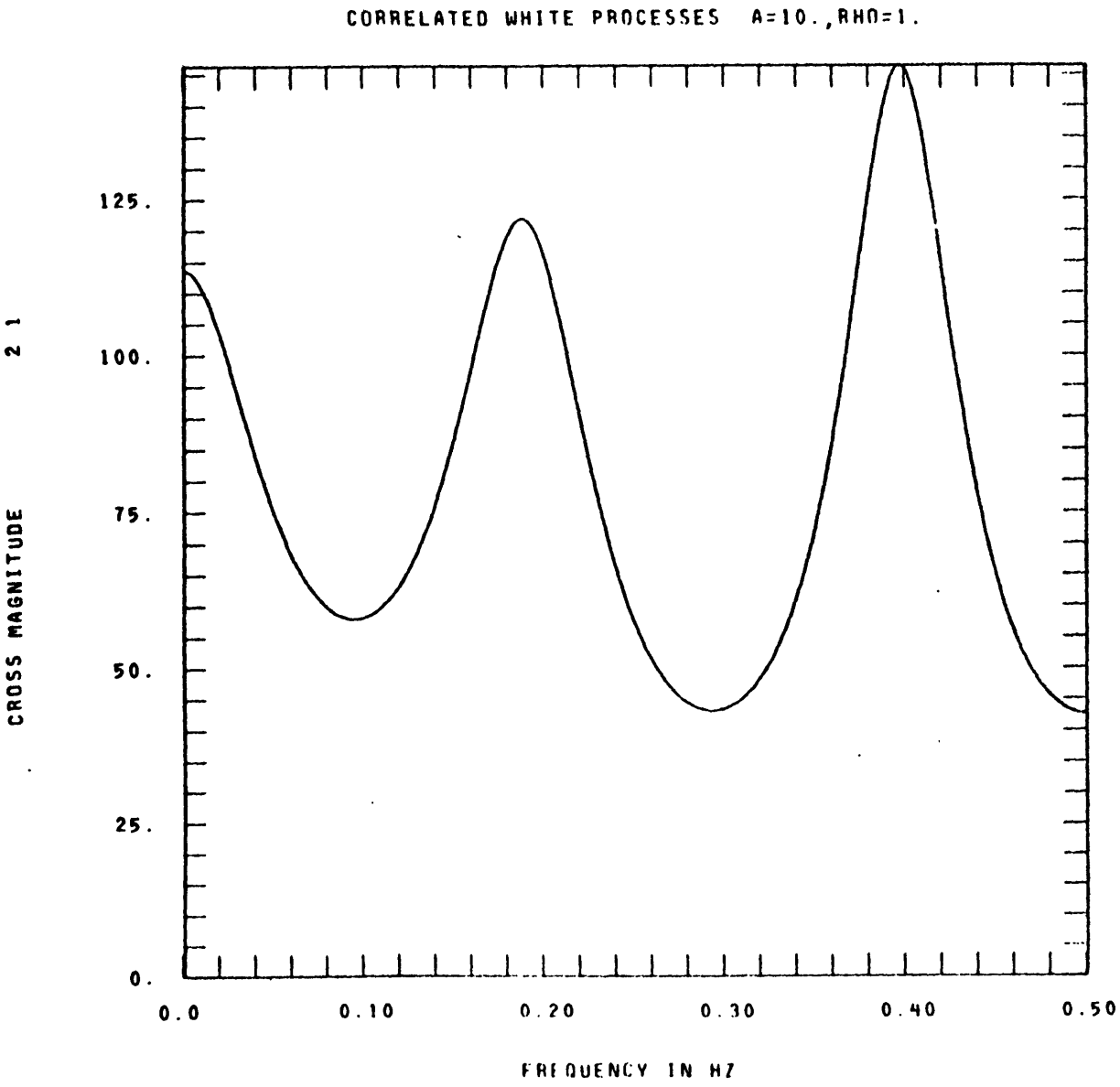

CORRELATED WHITE PROCESSES A=10.,RHO=1.
CROSS MAGNITUDE 2 1
125.
100.
75.
50.
25.
0.
0.0
0.10
0.20
0.30
0.40
0.50
FREQUENCY IN HZ

Figure 3

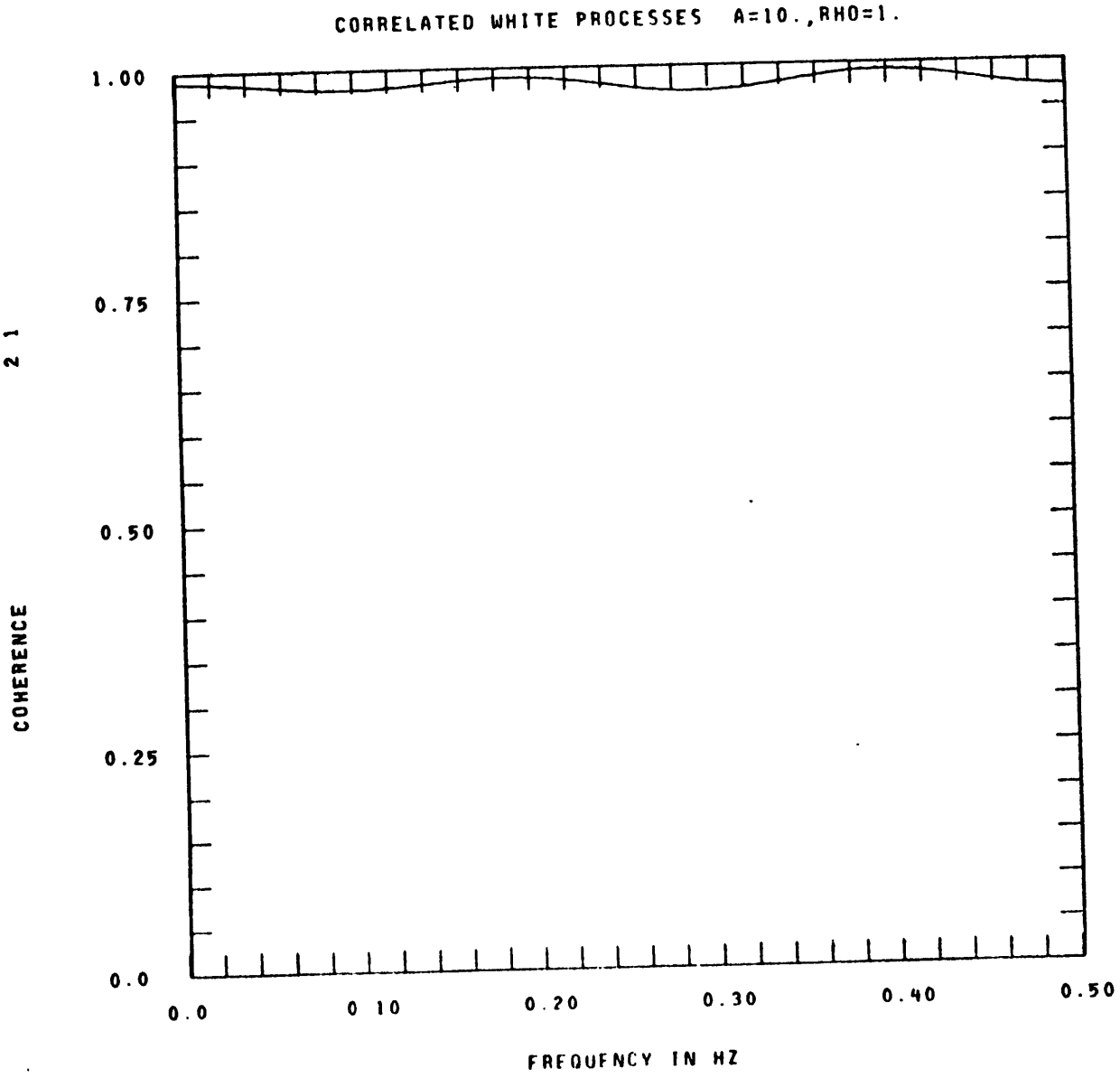

CORRELATED WHITE PROCESSES A=10.,RHO=1.
1.00
0.75
0.50
0.25
0.0
2 1
COHERENCE
0.0
0 10
0.20
0.30
0.40
0.50
FREQUENCY IN HZ

Figure 4

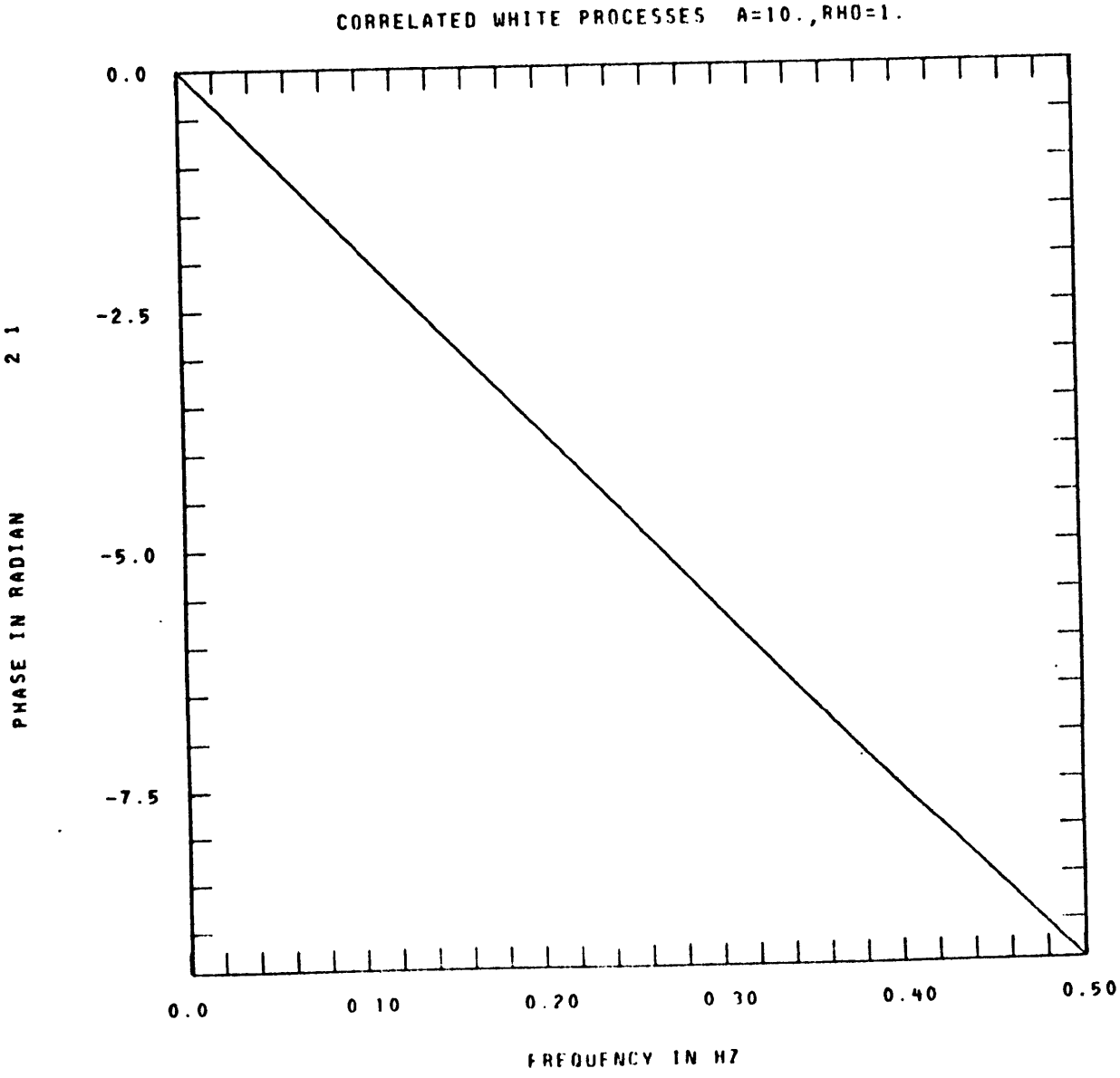

CORRELATED WHITE PROCESSES A=10.,RHO=1.
0.0
-2.5
-5.0
-7.5
PHASE IN RADIAN
2 1
0.0
0 10
0.20
0 30
0.40
0.50
FREQUENCY IN HZ

Figure 5

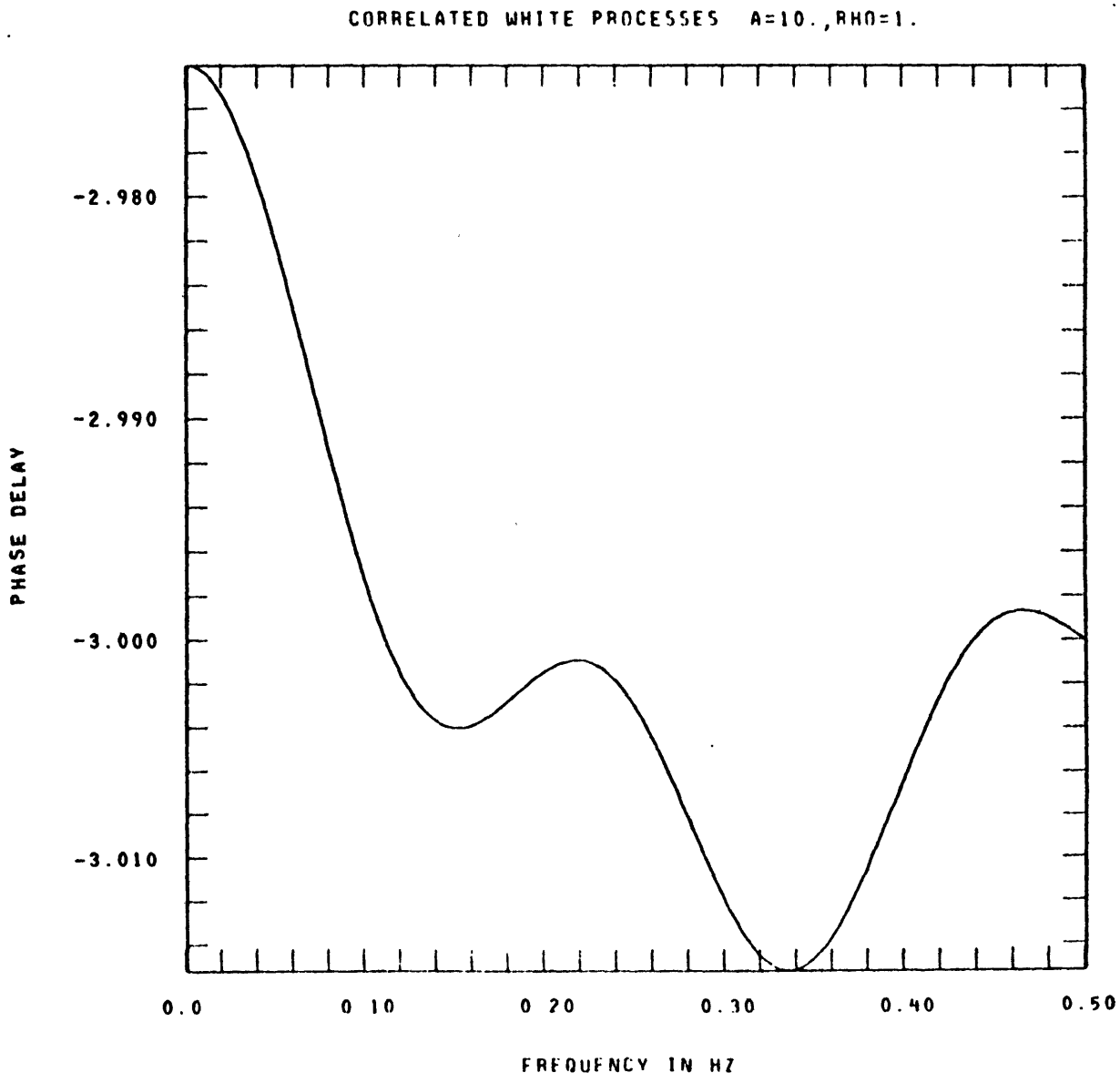

CORRELATED WHITE PROCESSES A=10.,RHO=1.
-2.980
-2.990
-3.000
-3.010
PHASE DELAY
0.0
0 10
0 20
0.30
0.40
0.50
FREQUENCY IN HZ

Figure 6

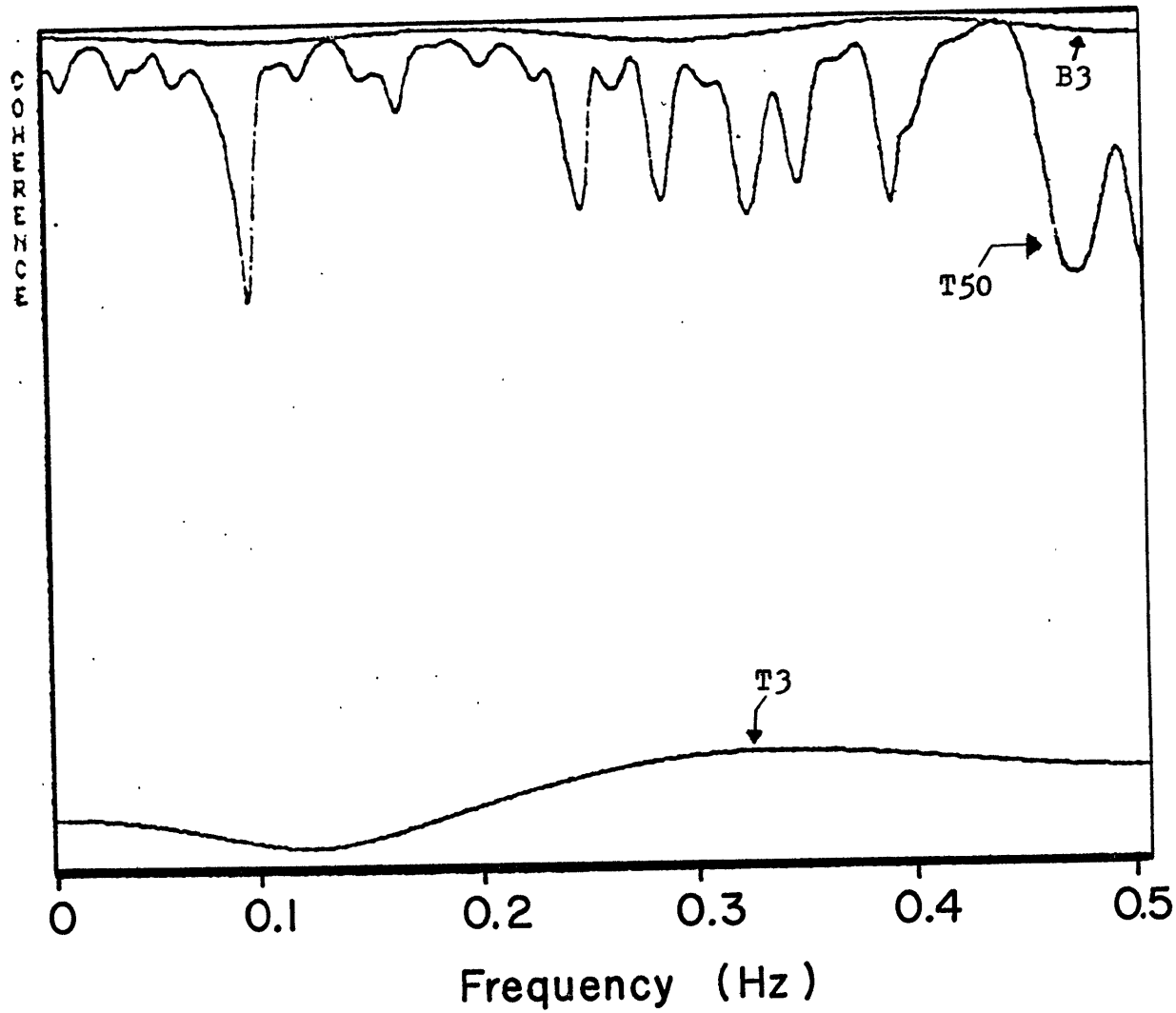


Figure 7

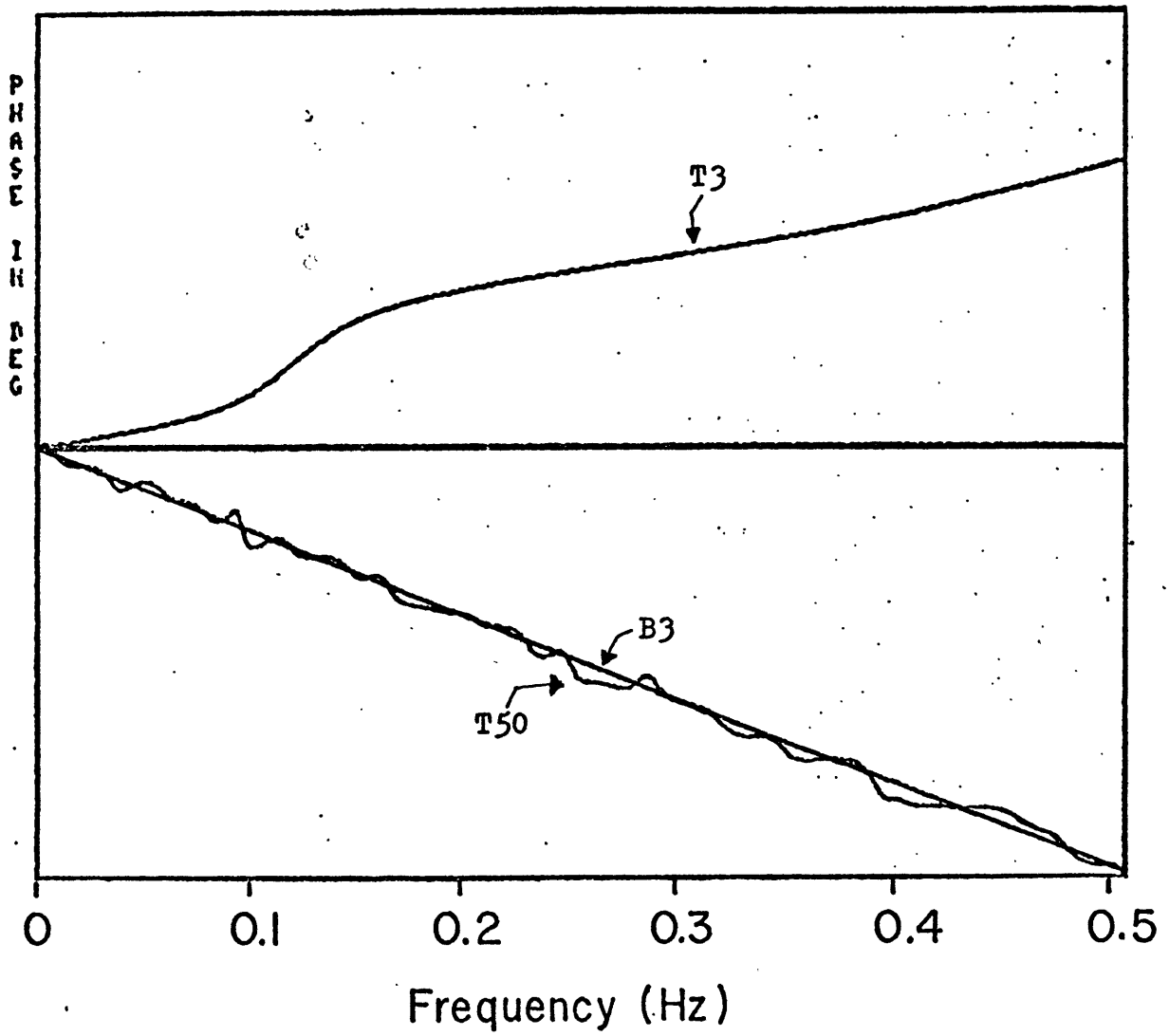


Figure 8

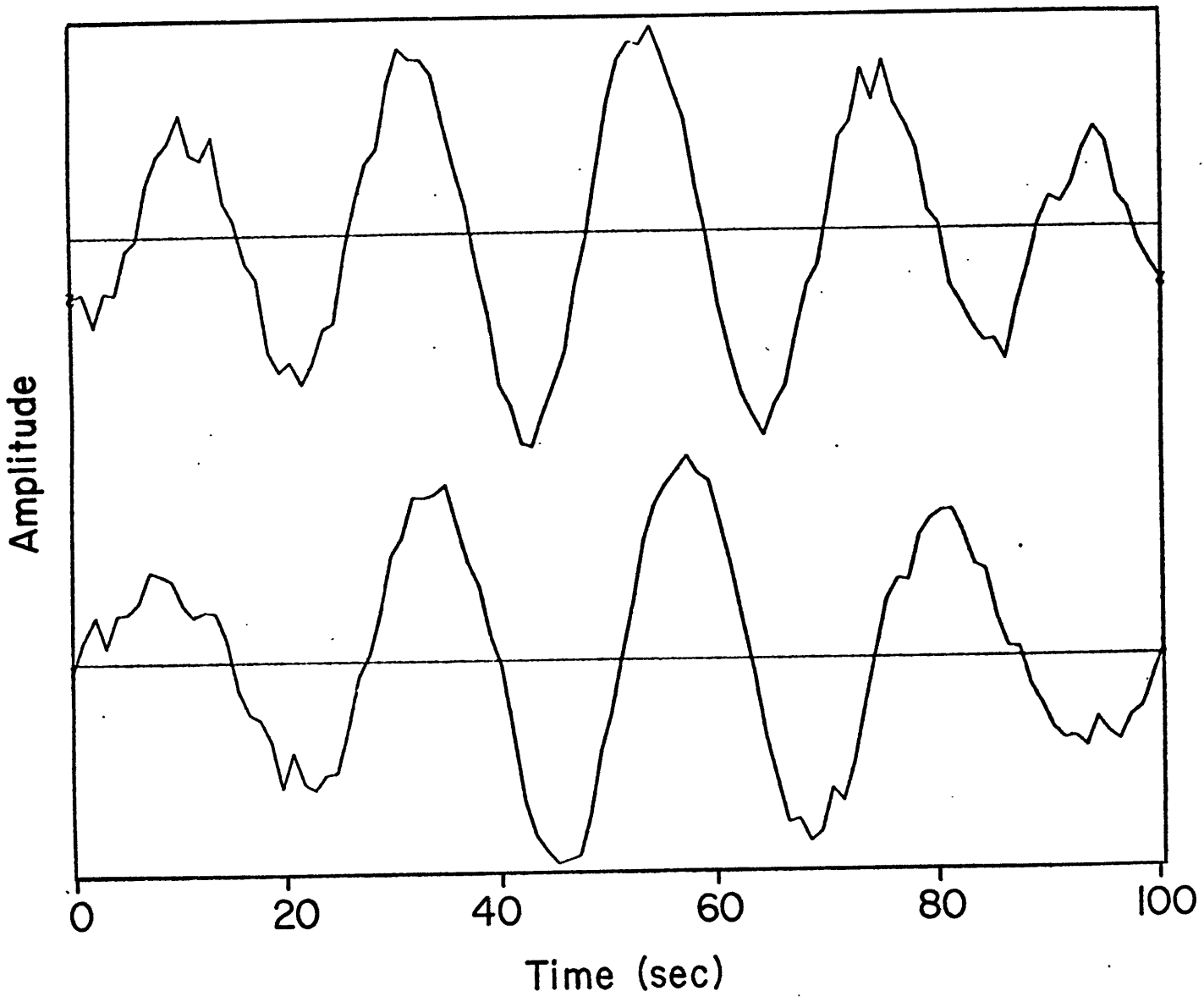


Figure 9

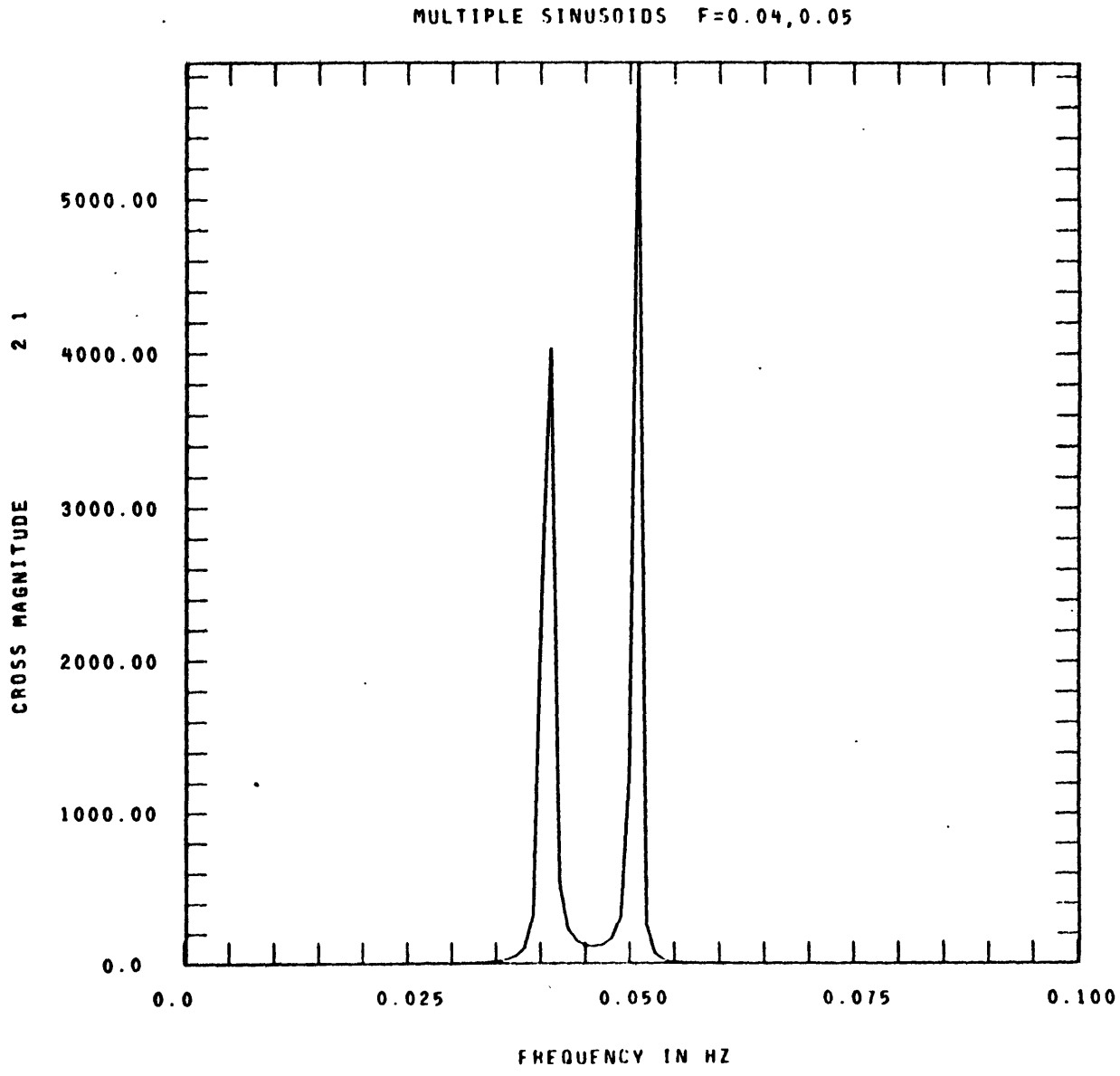

MULTIPLE SINUSOIDS F=0.04,0.05
CROSS MAGNITUDE 2 1
5000.00
4000.00
3000.00
2000.00
1000.00
0.0
0.0
0.025
0.050
0.075
0.100
FREQUENCY IN HZ

Figure 10

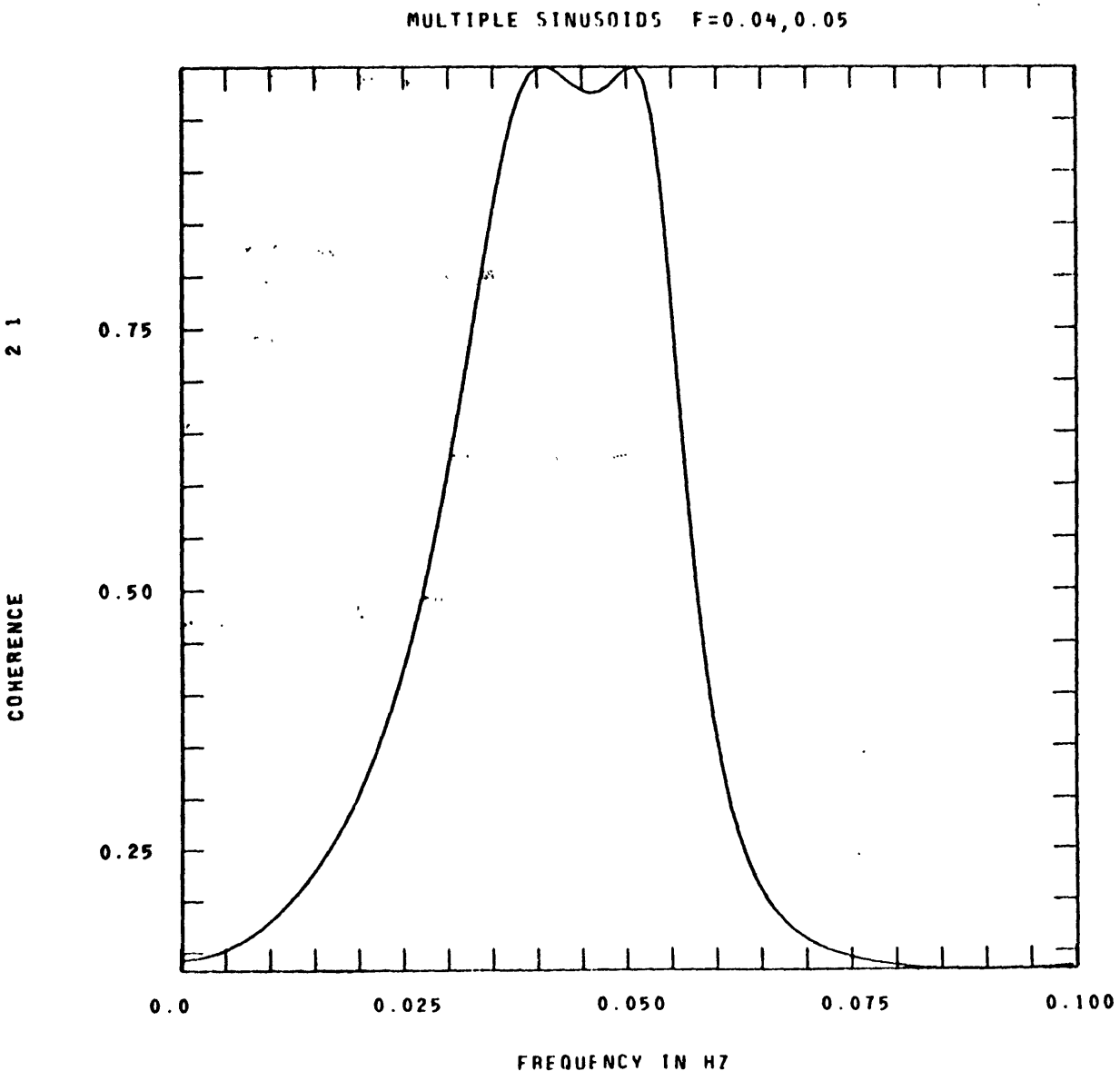

MULTIPLE SINUSOIDS F=0.04,0.05
2 1
COHERENCE
0.75
0.50
0.25
0.0
0.025
0.050
0.075
0.100
FREQUENCY IN HZ

Figure 11

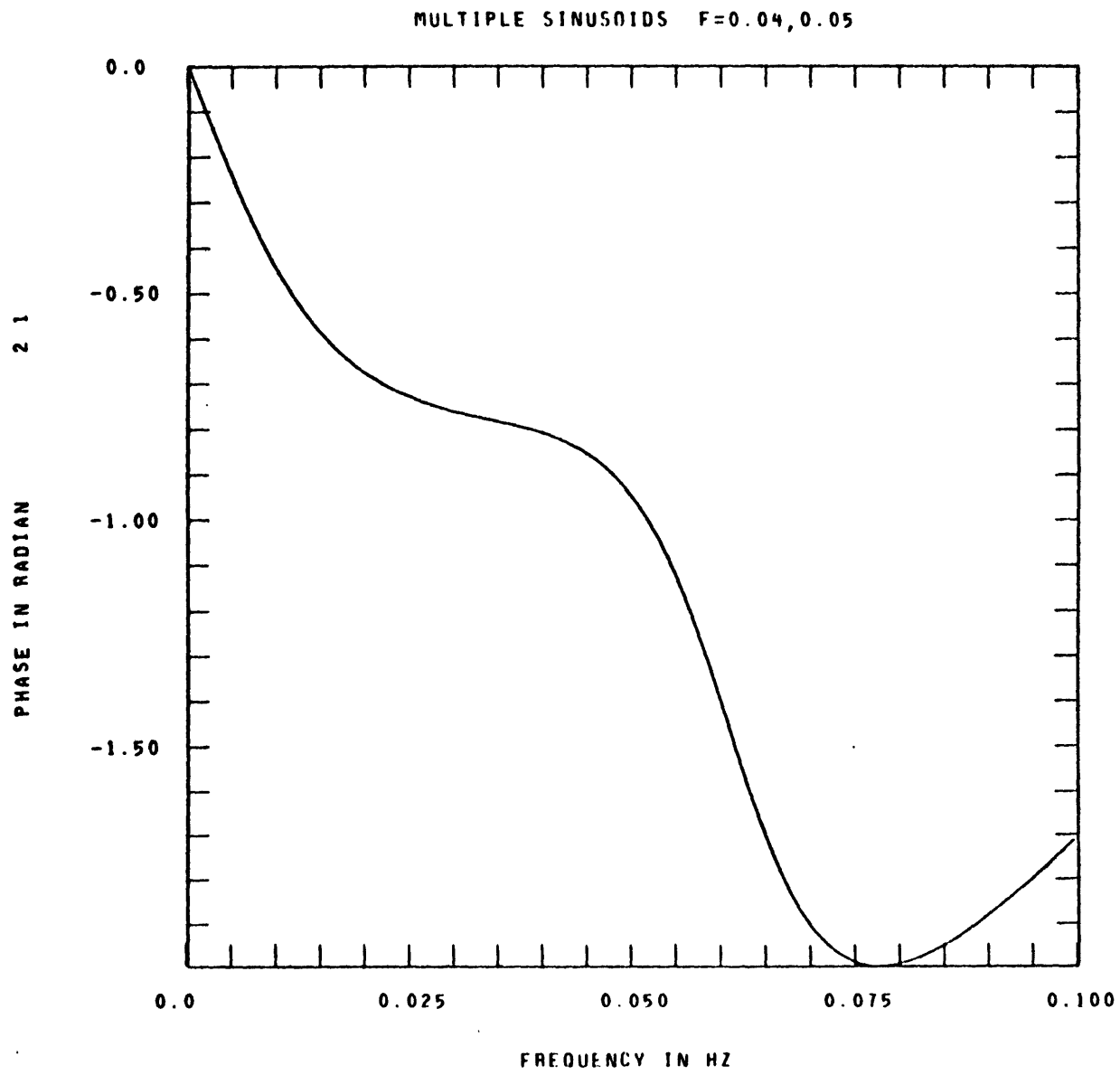

MULTIPLE SINUSOIDS F=0.04,0.05
PHASE IN RADIAN
2 1
0.0
-0.50
-1.00
-1.50
0.0
0.025
0.050
0.075
0.100
FREQUENCY IN HZ

Figure 12

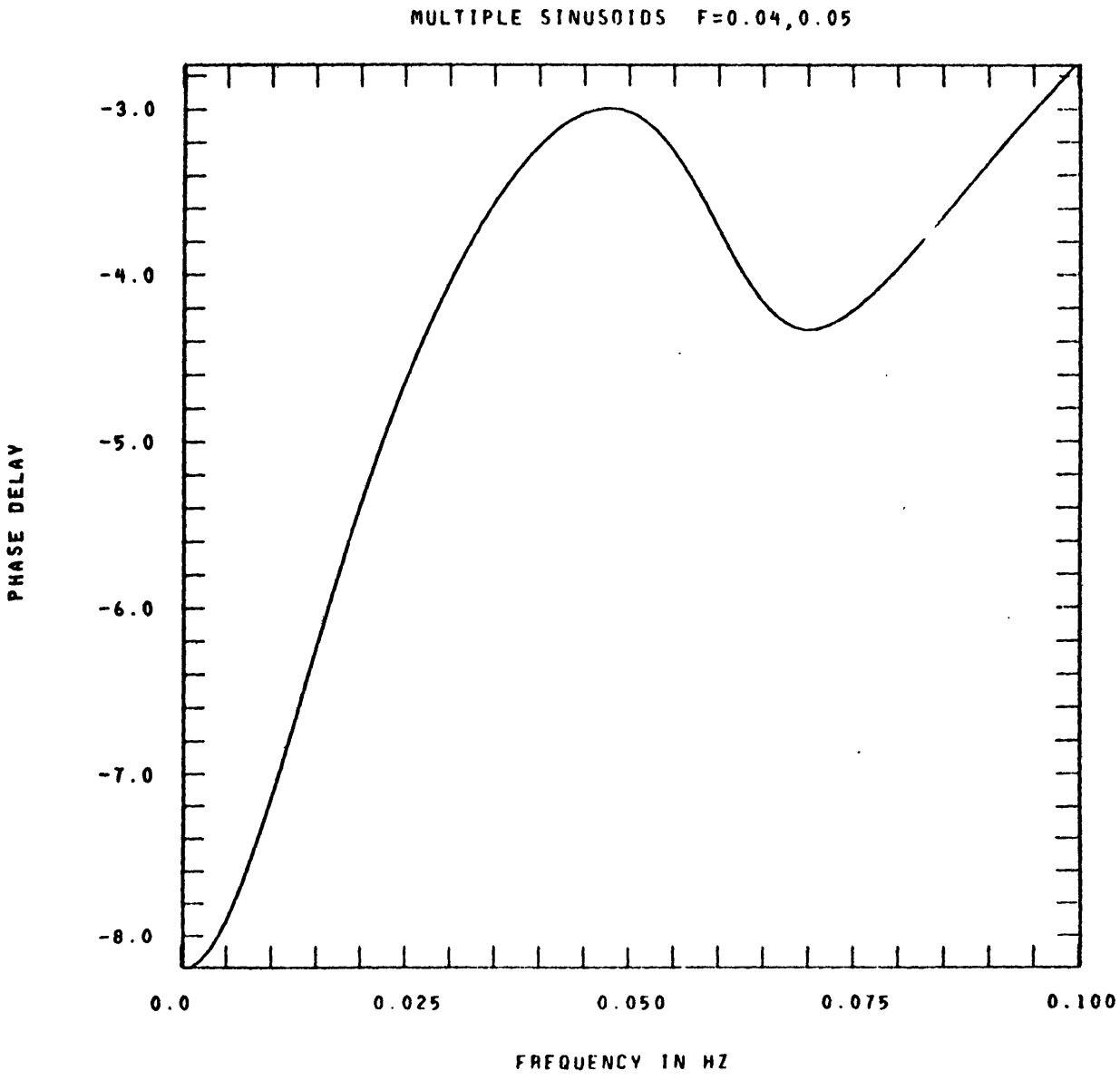

MULTIPLE SINUSOIDS F=0.04,0.05
-3.0
-4.0
-5.0
-6.0
-7.0
-8.0
PHASE DELAY
0.0
0.025
0.050
0.075
0.100
FREQUENCY IN HZ

Figure 13

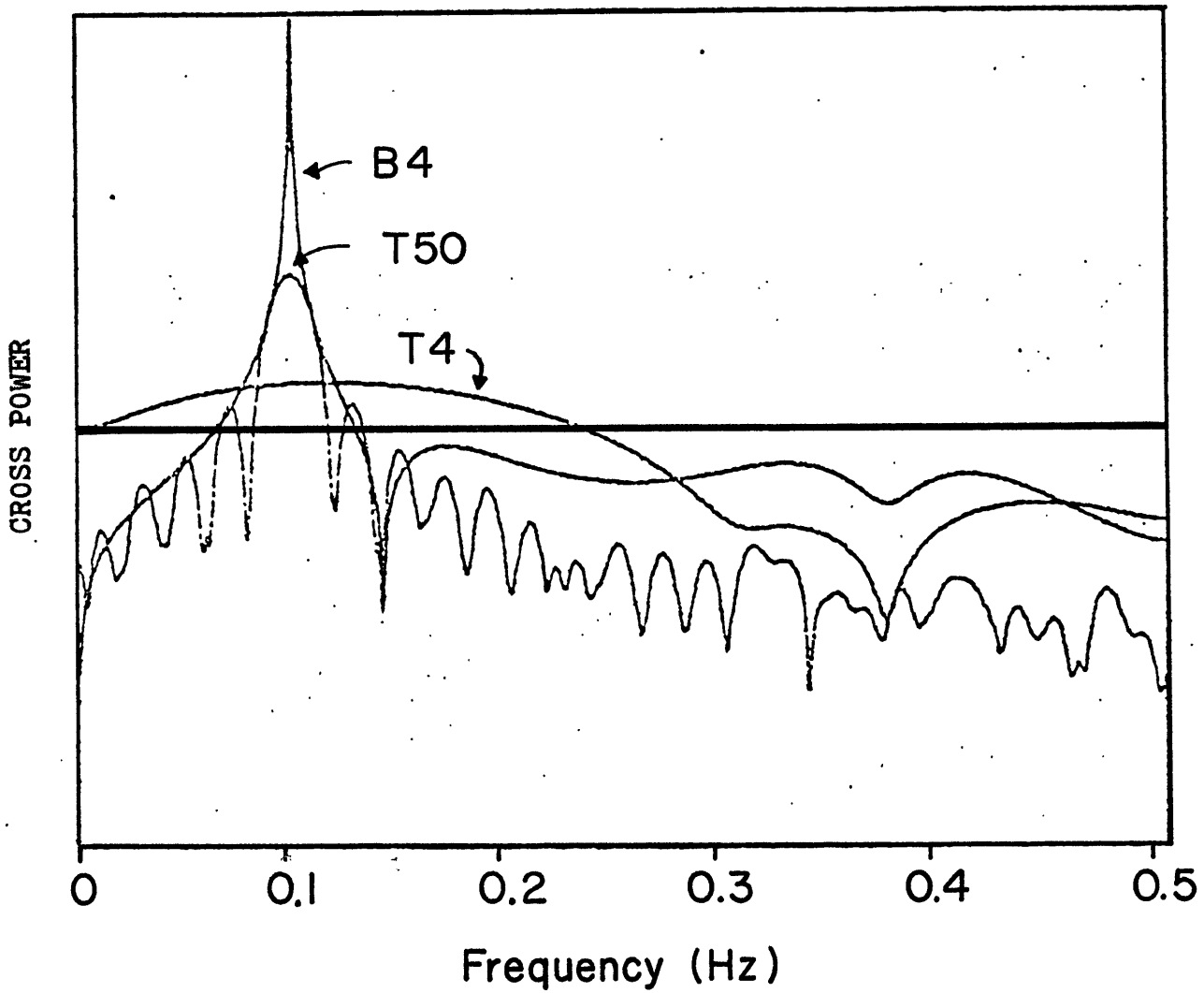


Figure 14

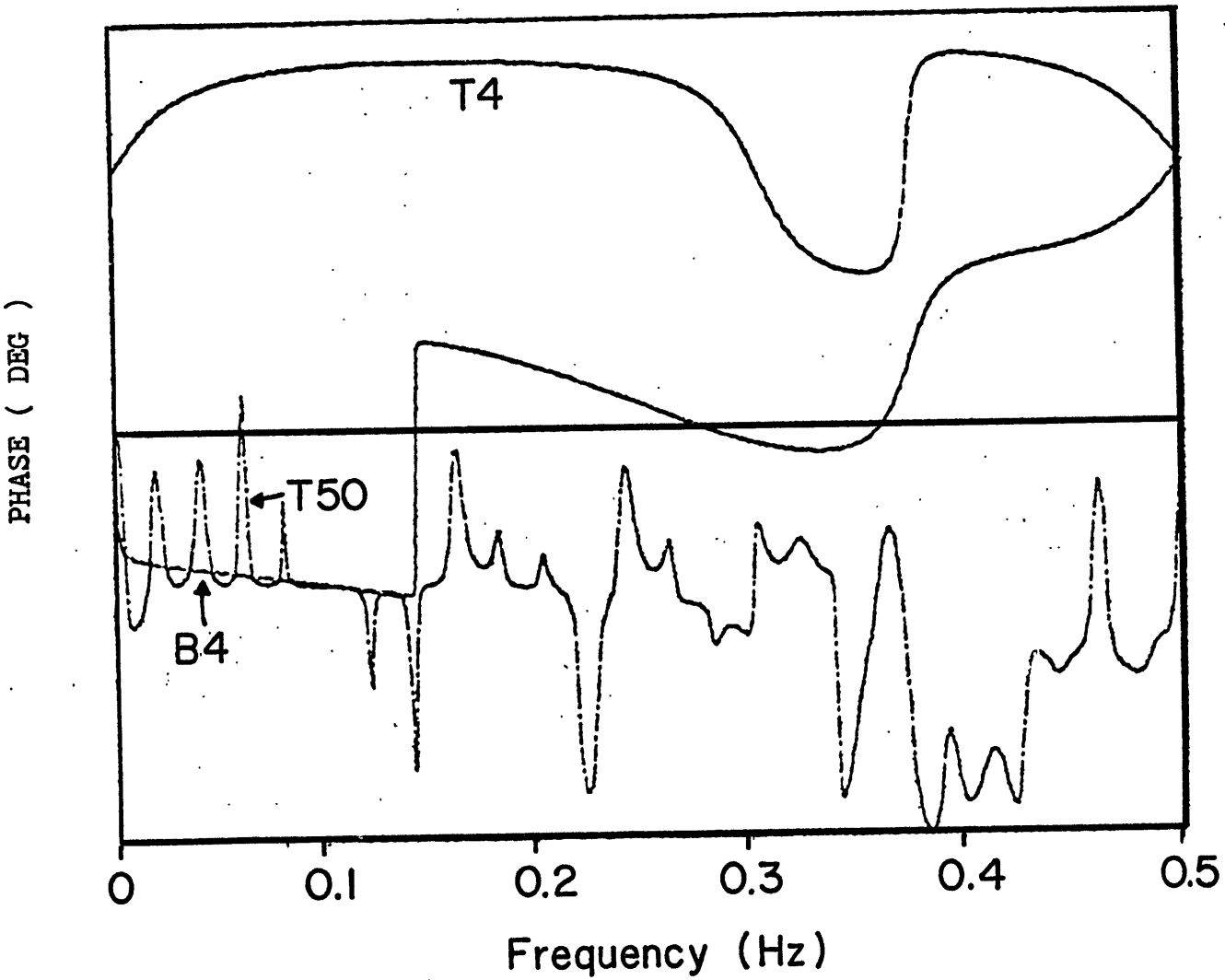


Figure 15

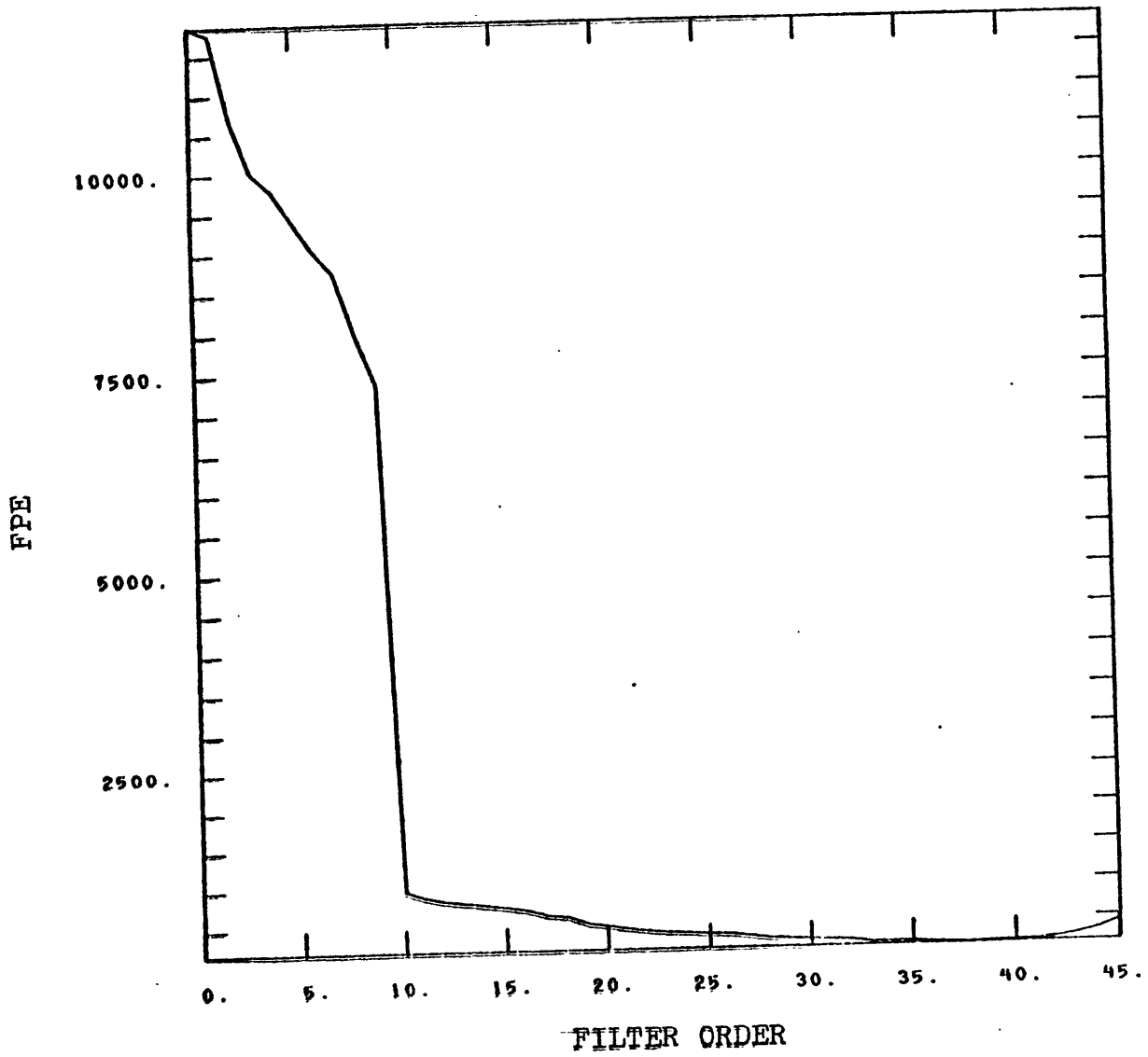


Figure 16

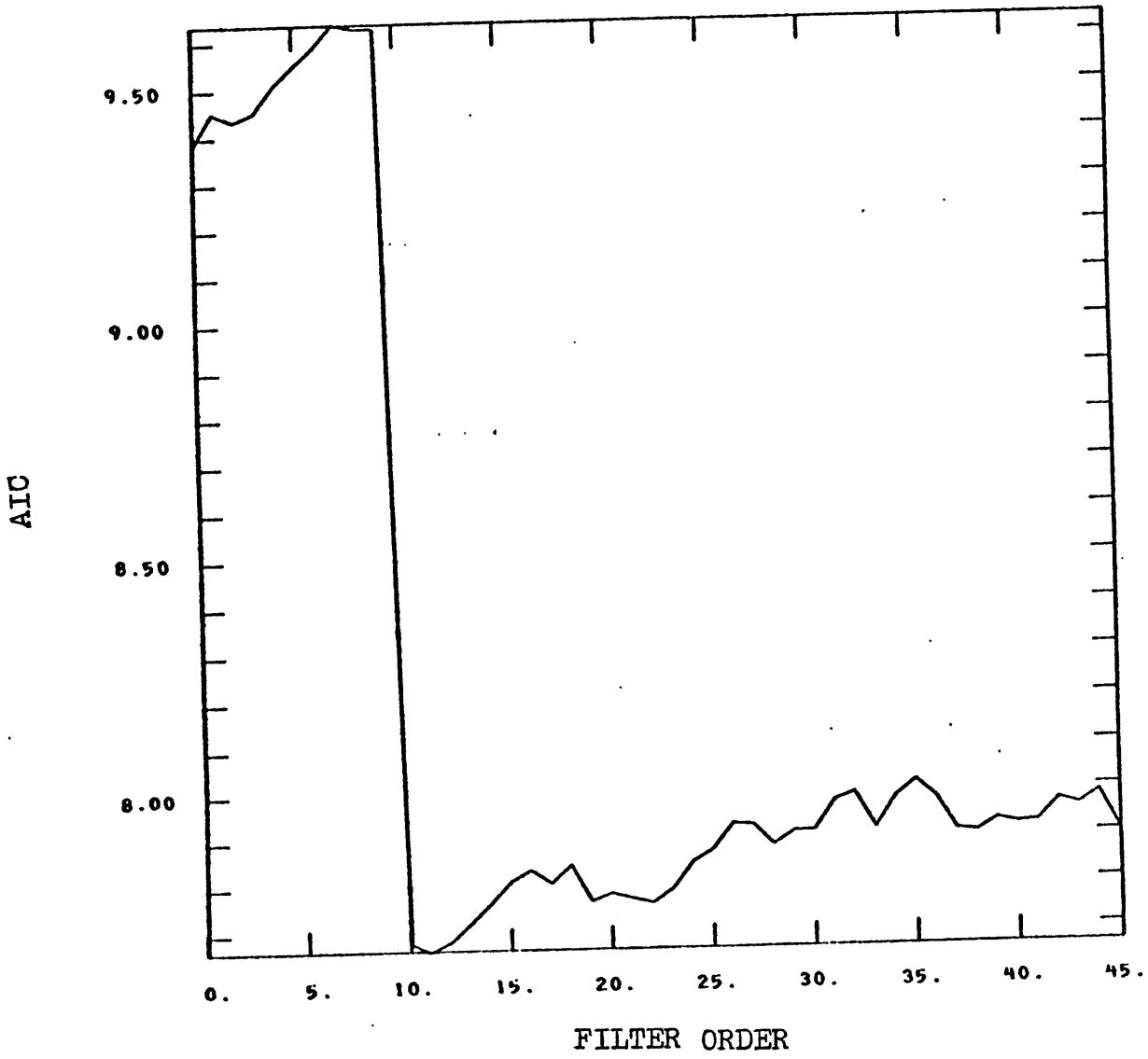


Figure 17

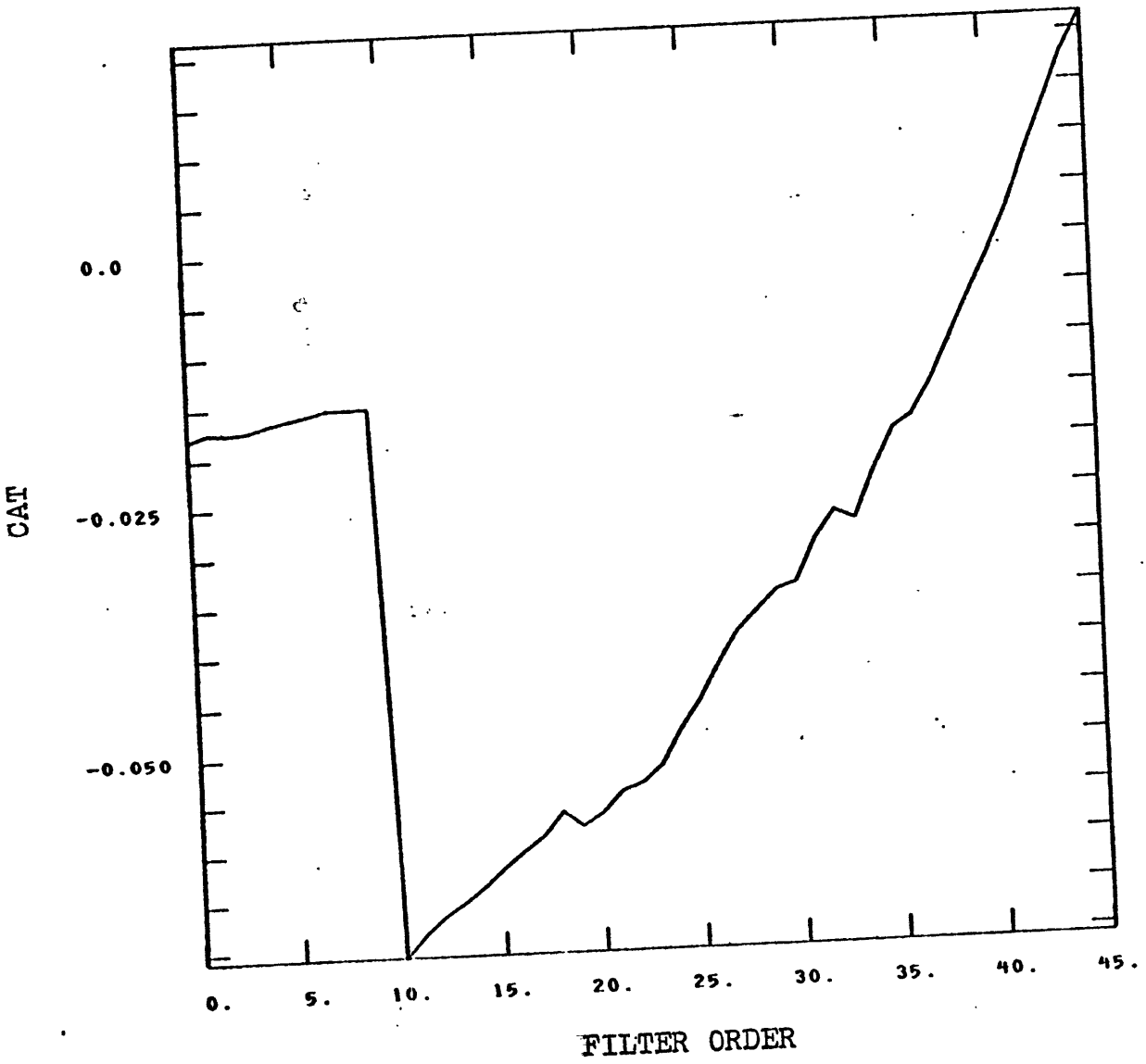
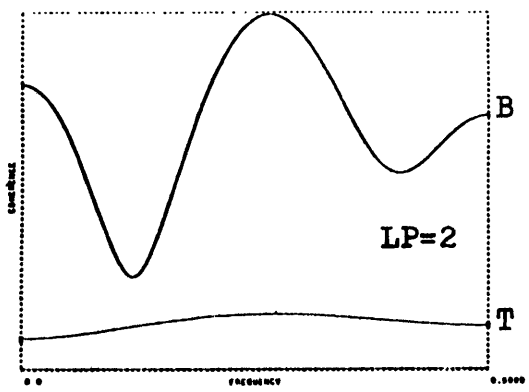
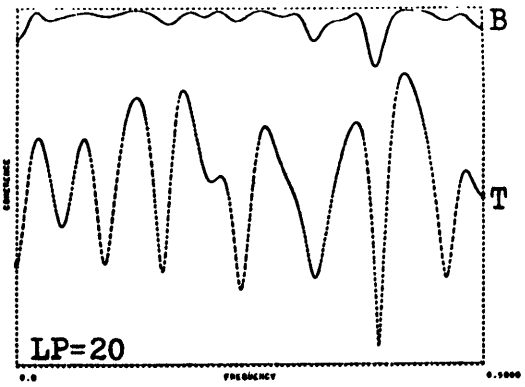
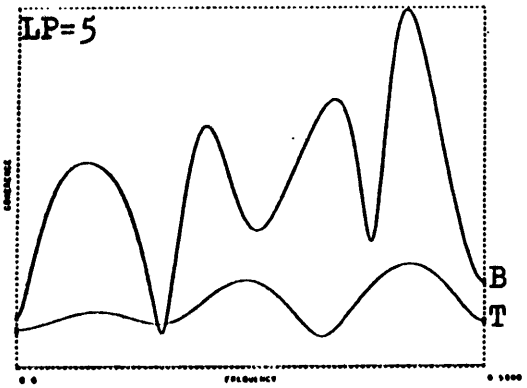
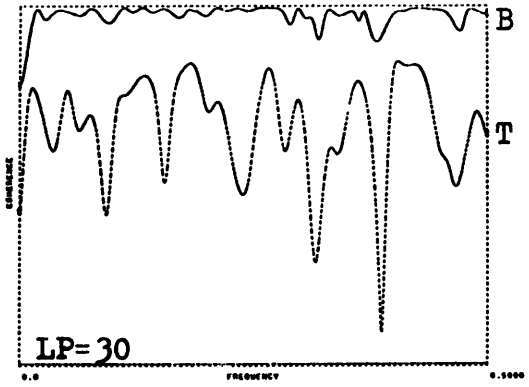


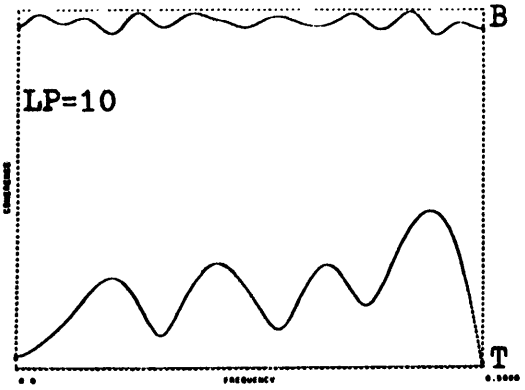
Figure 18

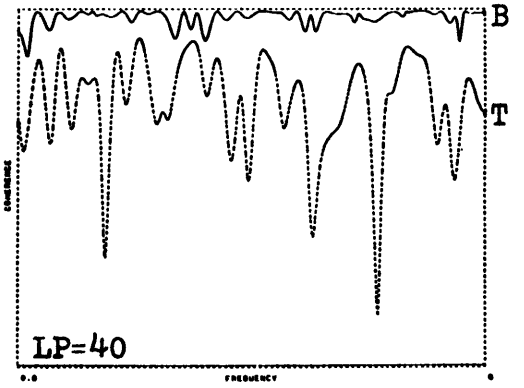


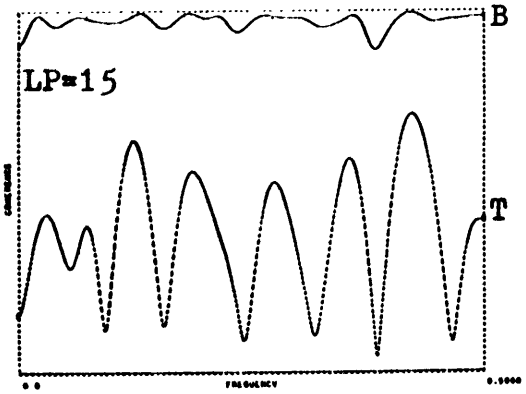












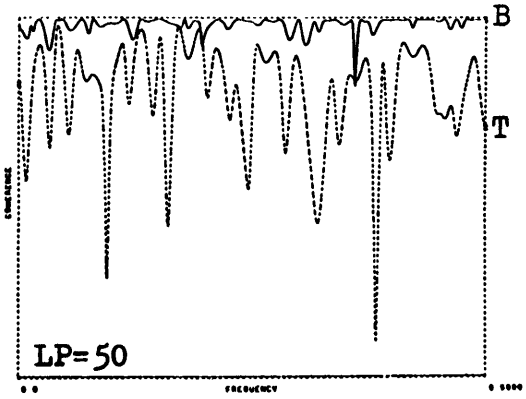


Figure 19

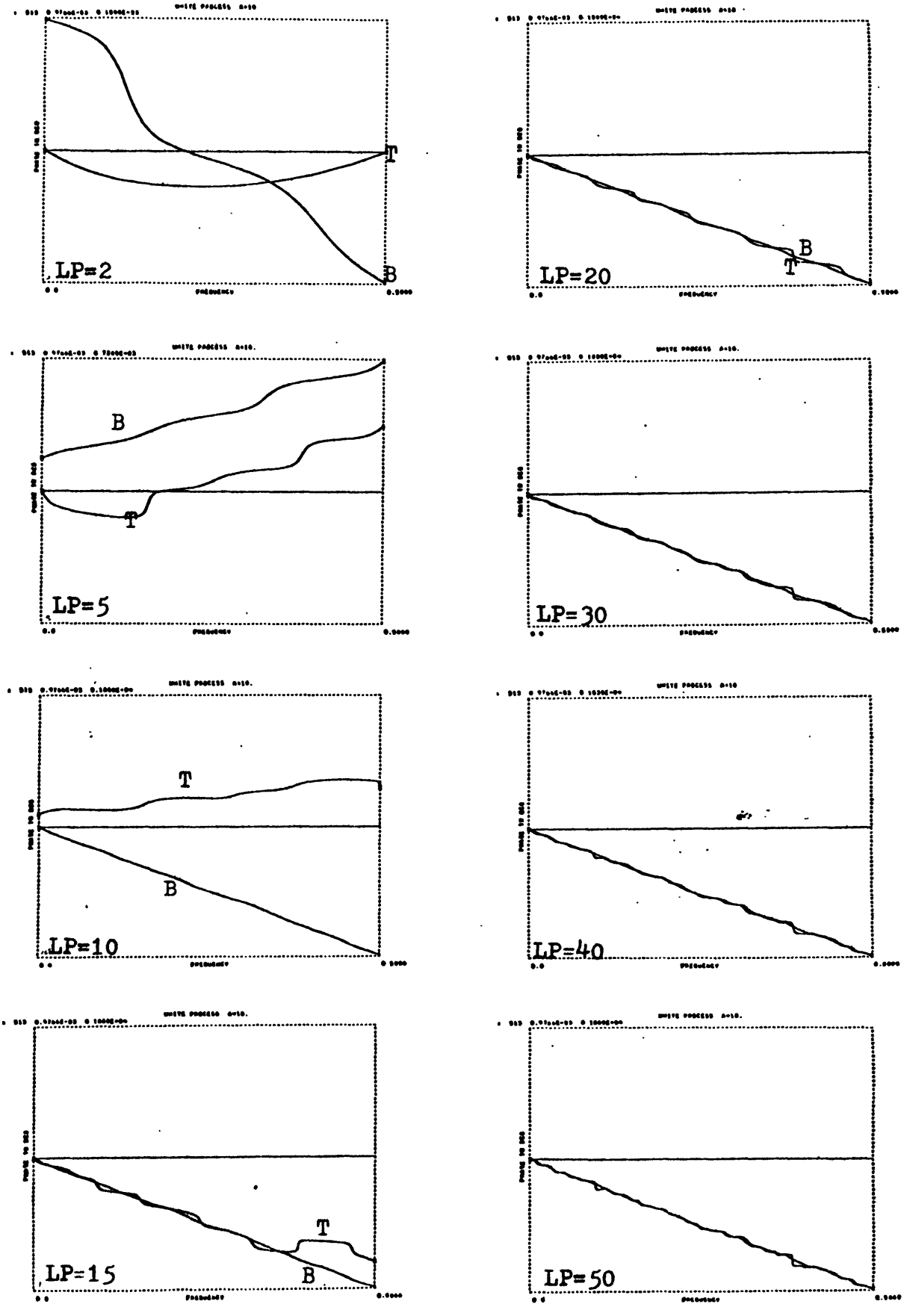


Figure 20

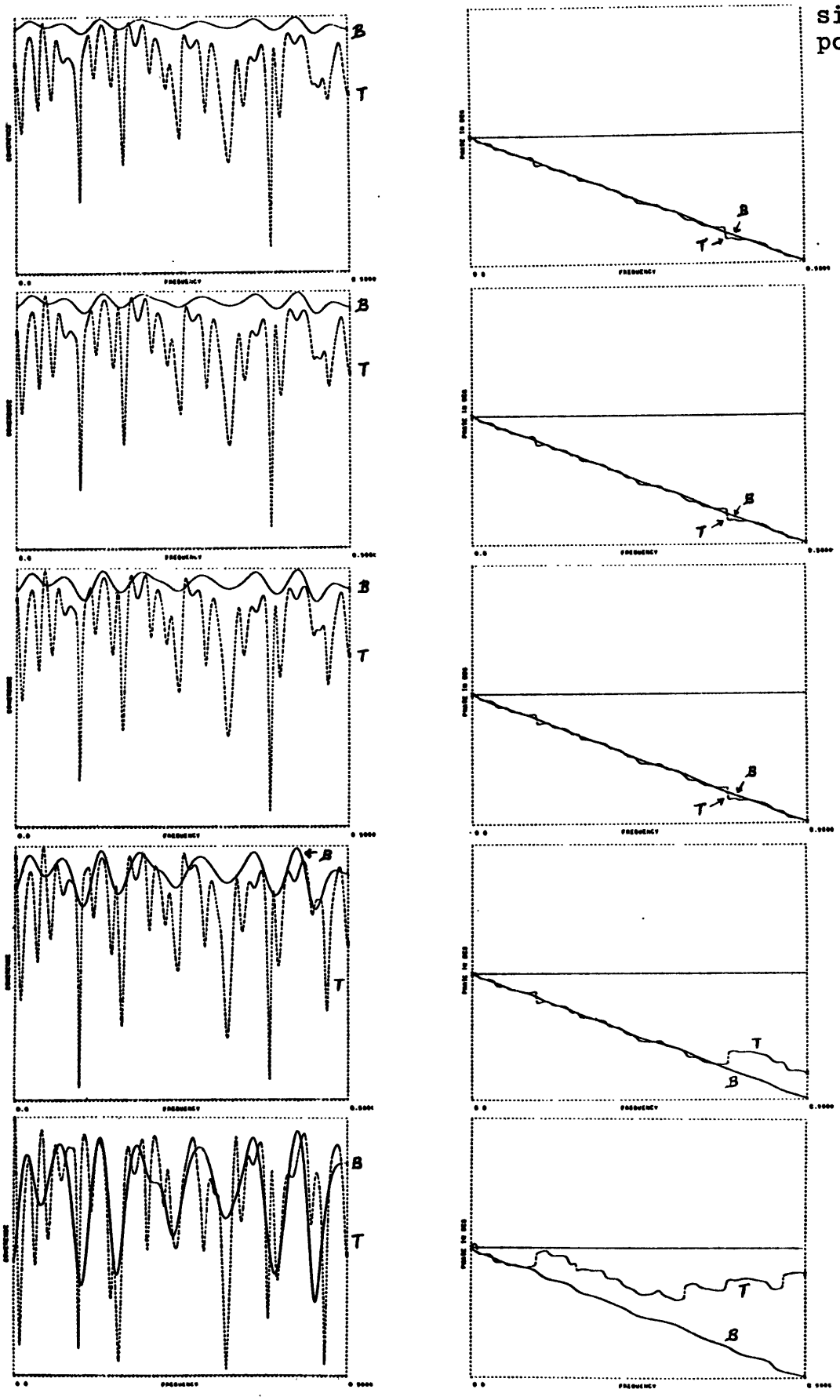


Figure 21

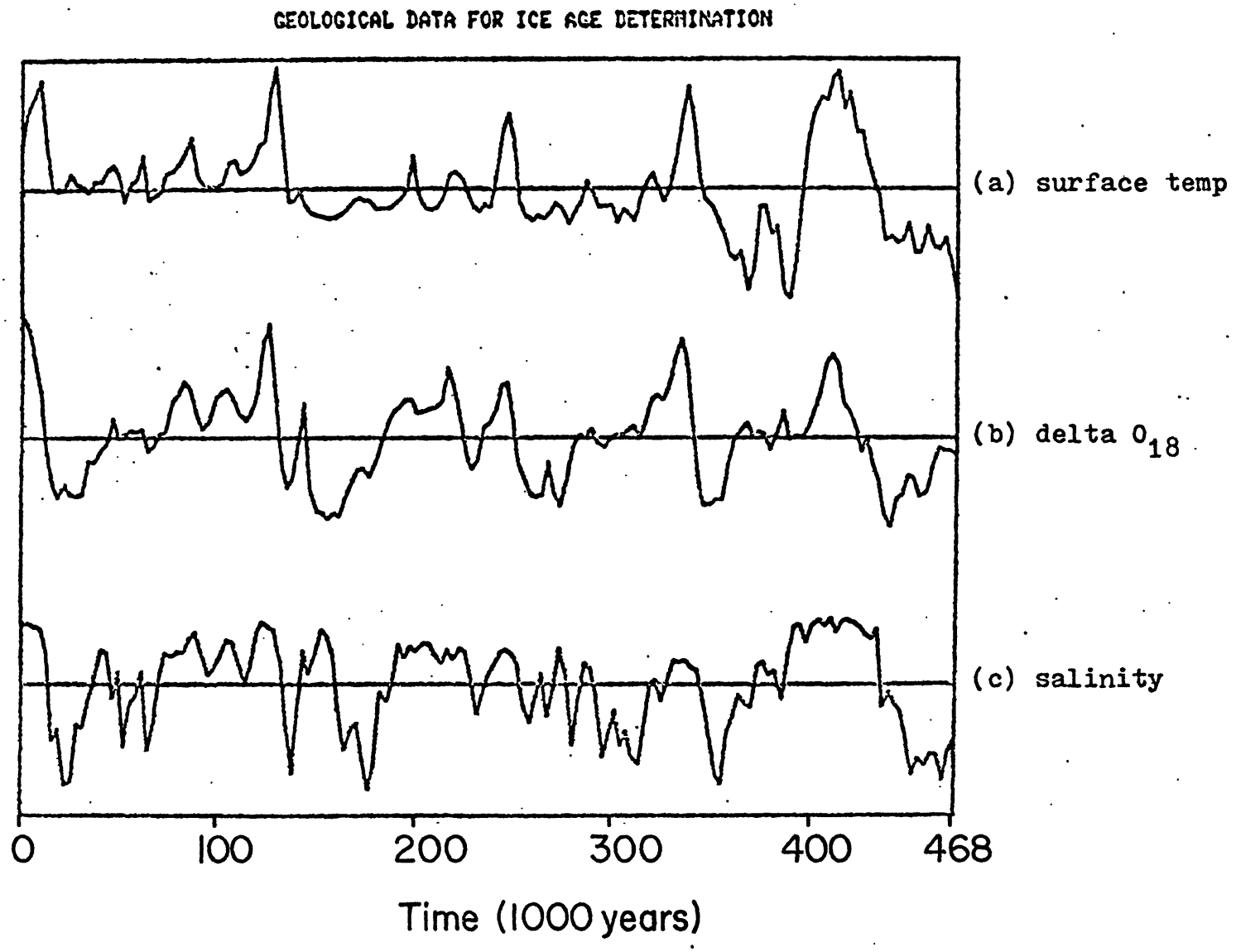

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
(a) surface temp
(b) delta O_{18}
(c) salinity
0
100
200
300
400
468
Time (1000 years)

Figure 22

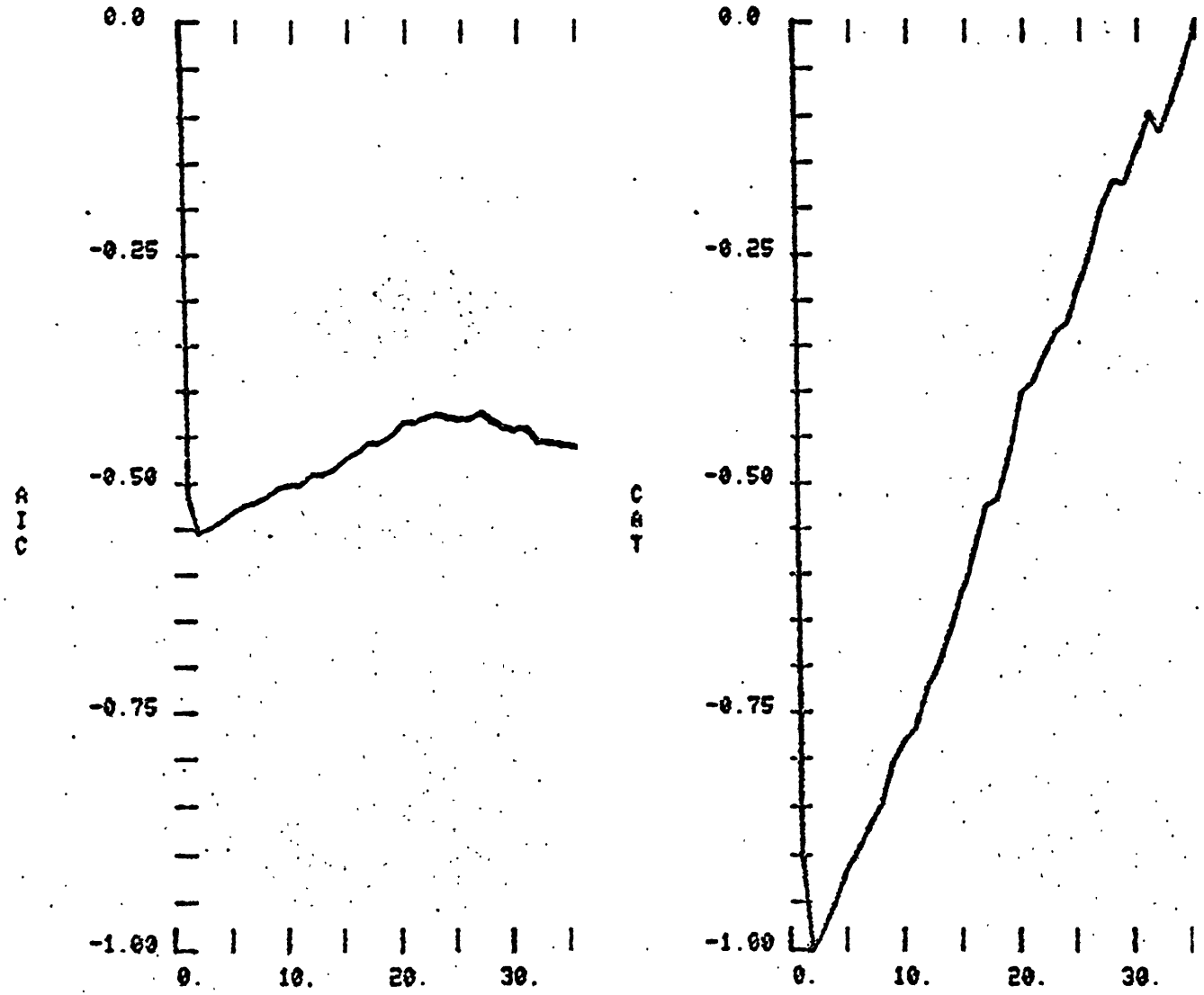

0.0
-0.25
-0.50
-0.75
-1.00
AIC
0.
10.
20.
30.
0.0
-0.25
-0.50
-0.75
-1.00
CAT
0.
10.
20.
30.
FILTER ORDER
AUTOREGRESSIVE FILTER ORDER CRITERION

Figure 23

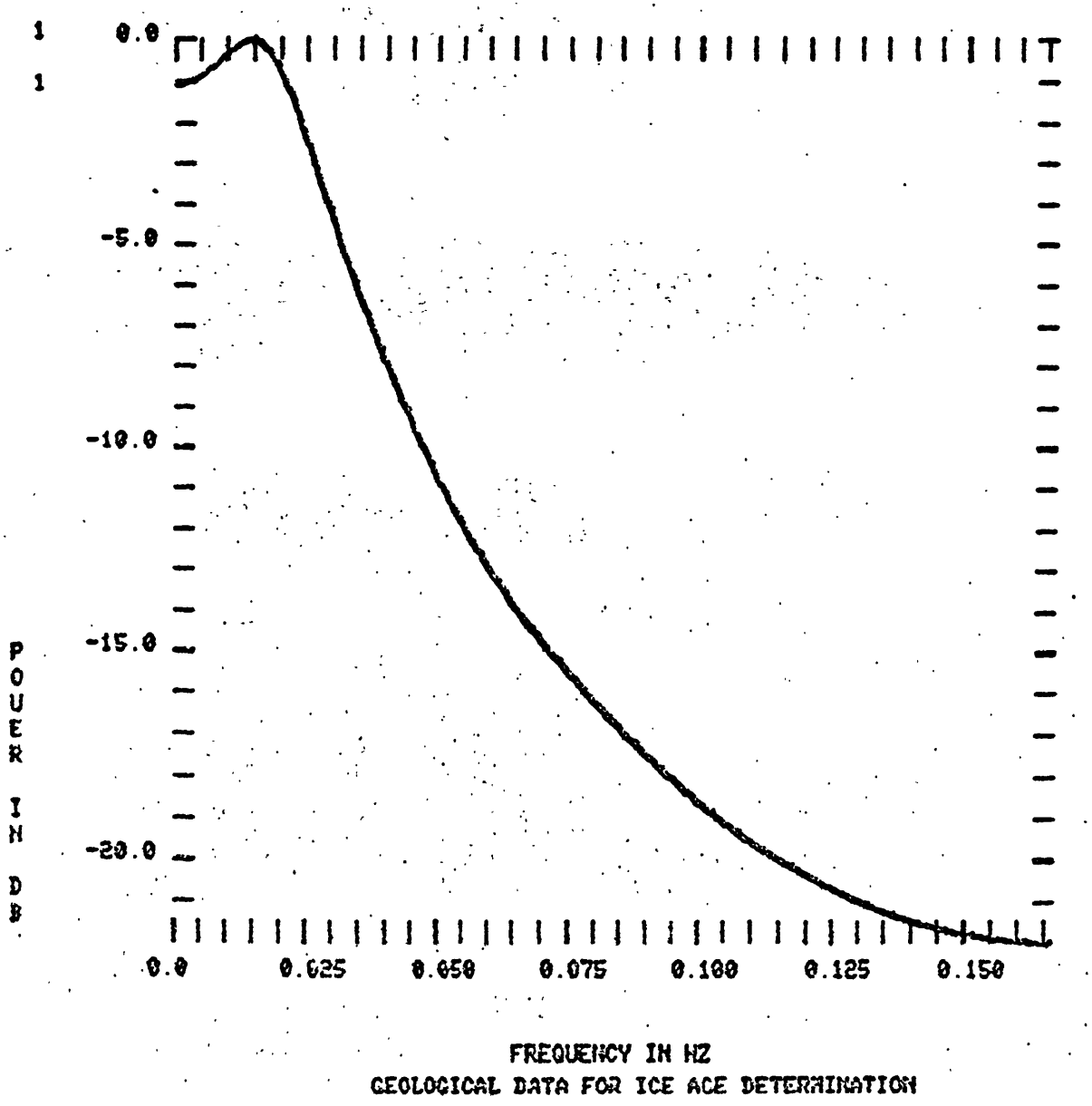

1
1
0.0
-5.0
-10.0
-15.0
-20.0
POWER IN DB
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 24(a)

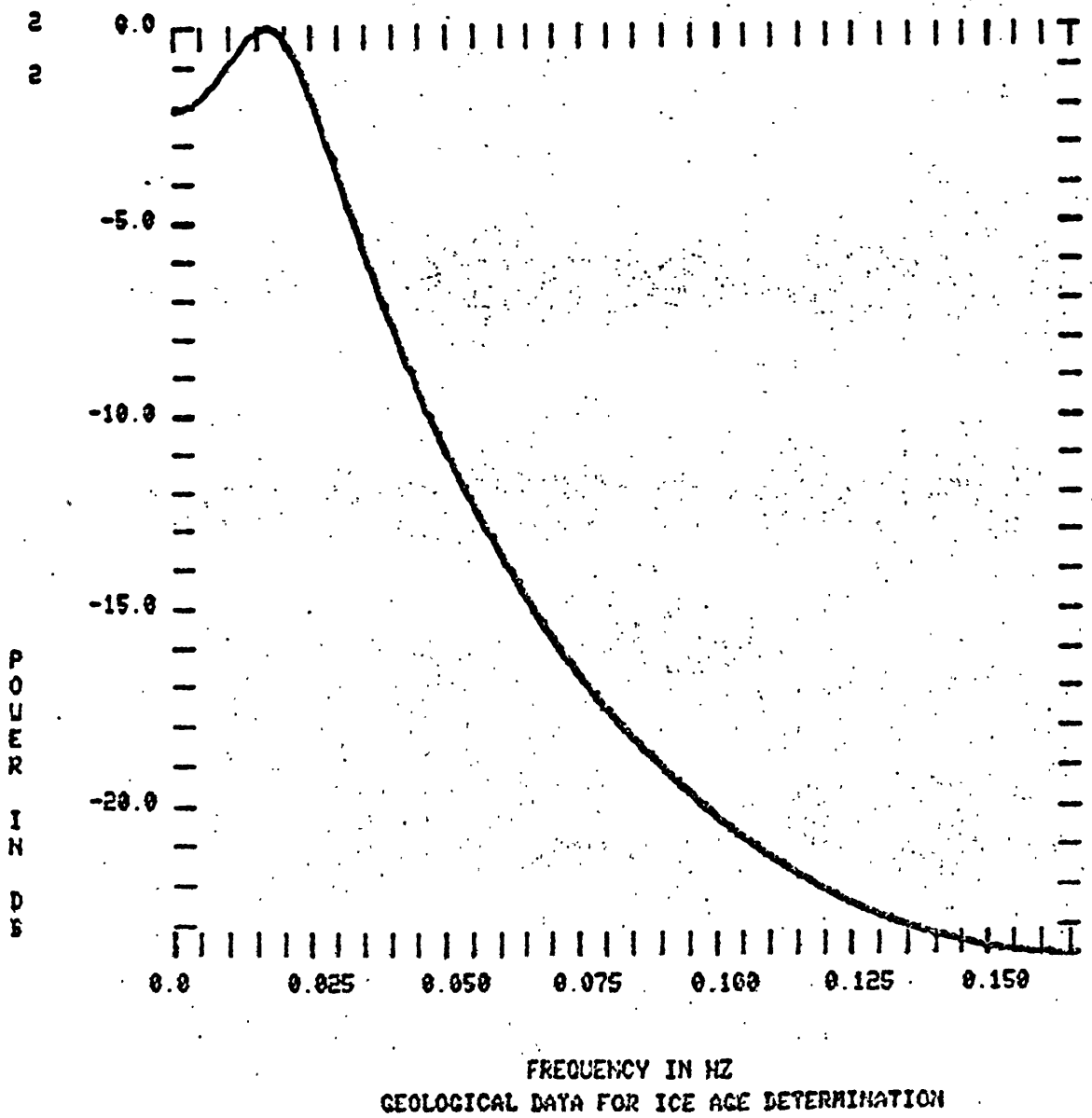

2
2
0.0
-5.0
-10.0
-15.0
-20.0
POWER IN DB
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 24(b)

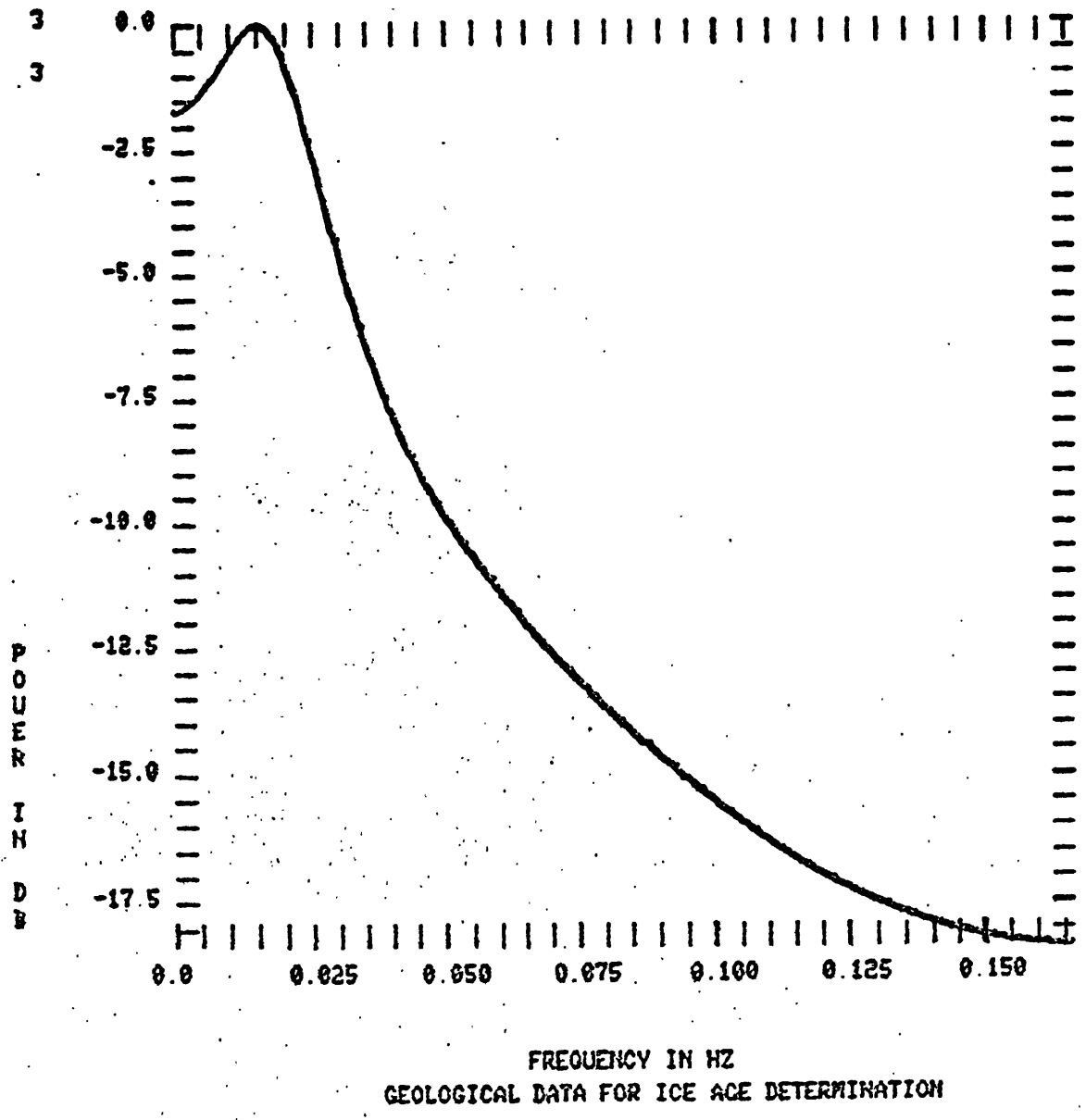

3
3
0.0
-2.5
-5.0
-7.5
-10.0
-12.5
-15.0
-17.5
POWER IN DB
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 24(c)

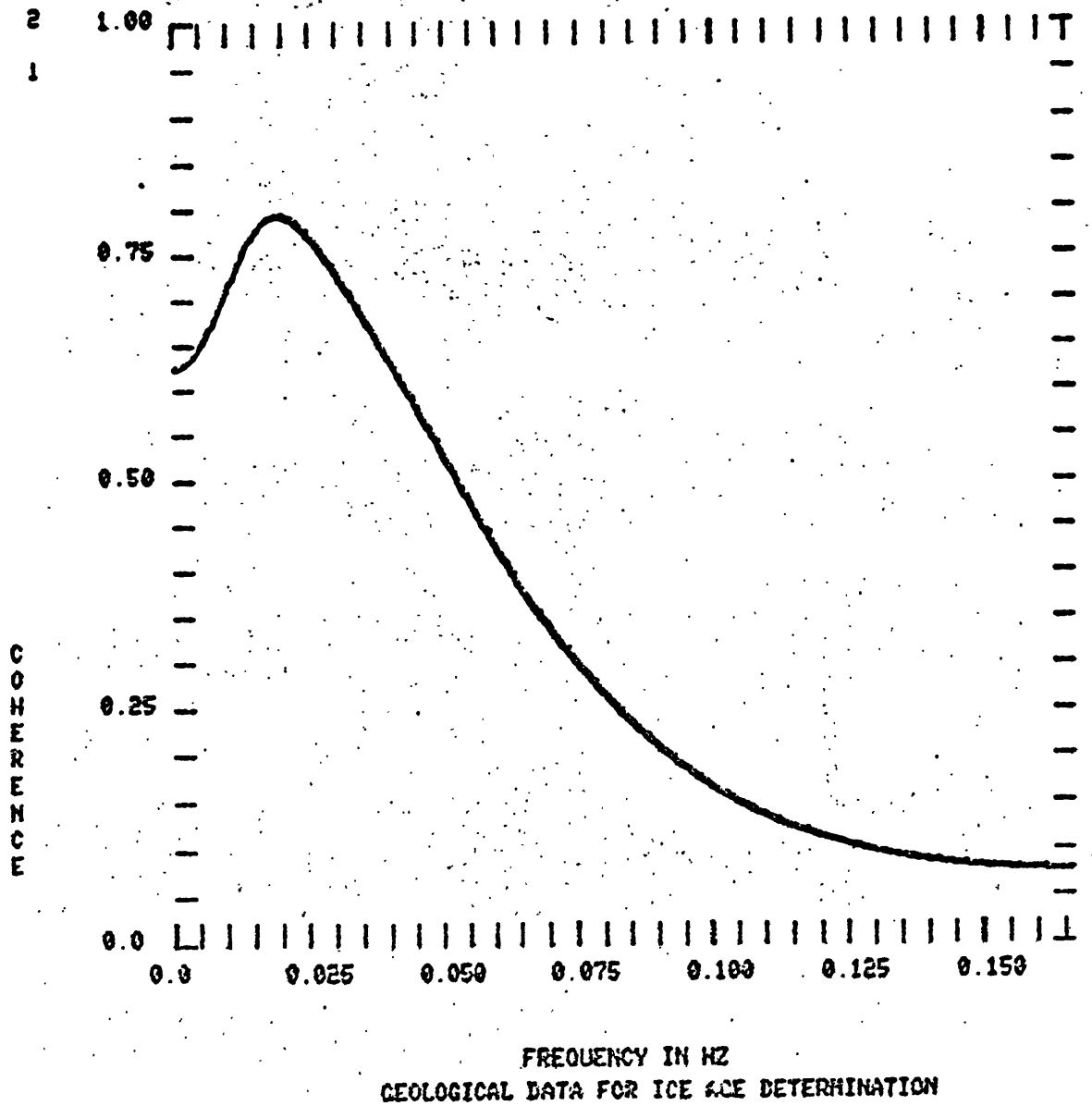

2
1
1.00
0.75
0.50
0.25
0.0
COHERENCE
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 25(a)

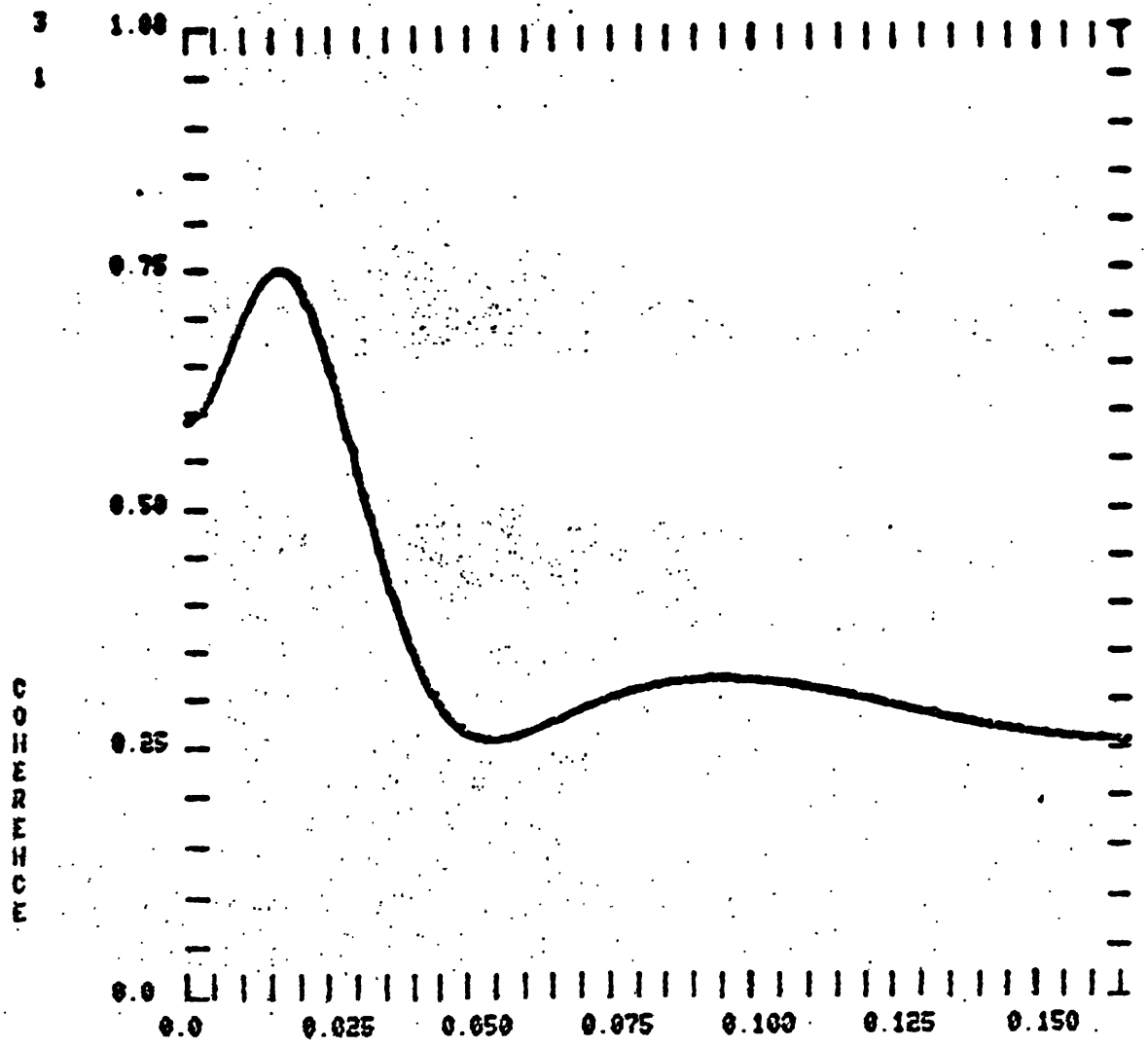

3
1
1.00
0.75
0.50
0.25
0.0
COHERENCE
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 25(b)

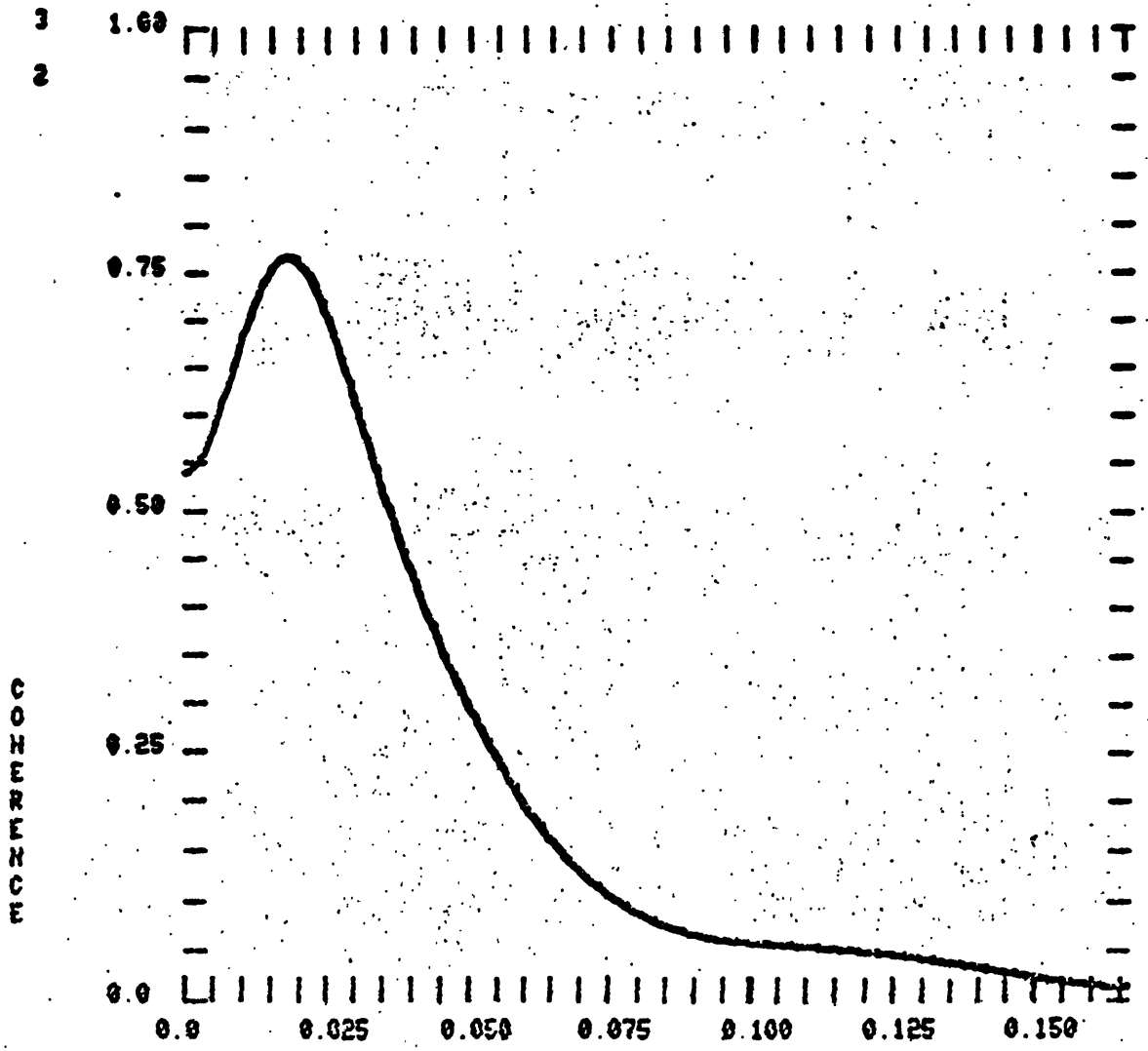

3
2
1.00
0.75
0.50
0.25
0.0
COHERENCE
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
LP= 2
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 25(c)

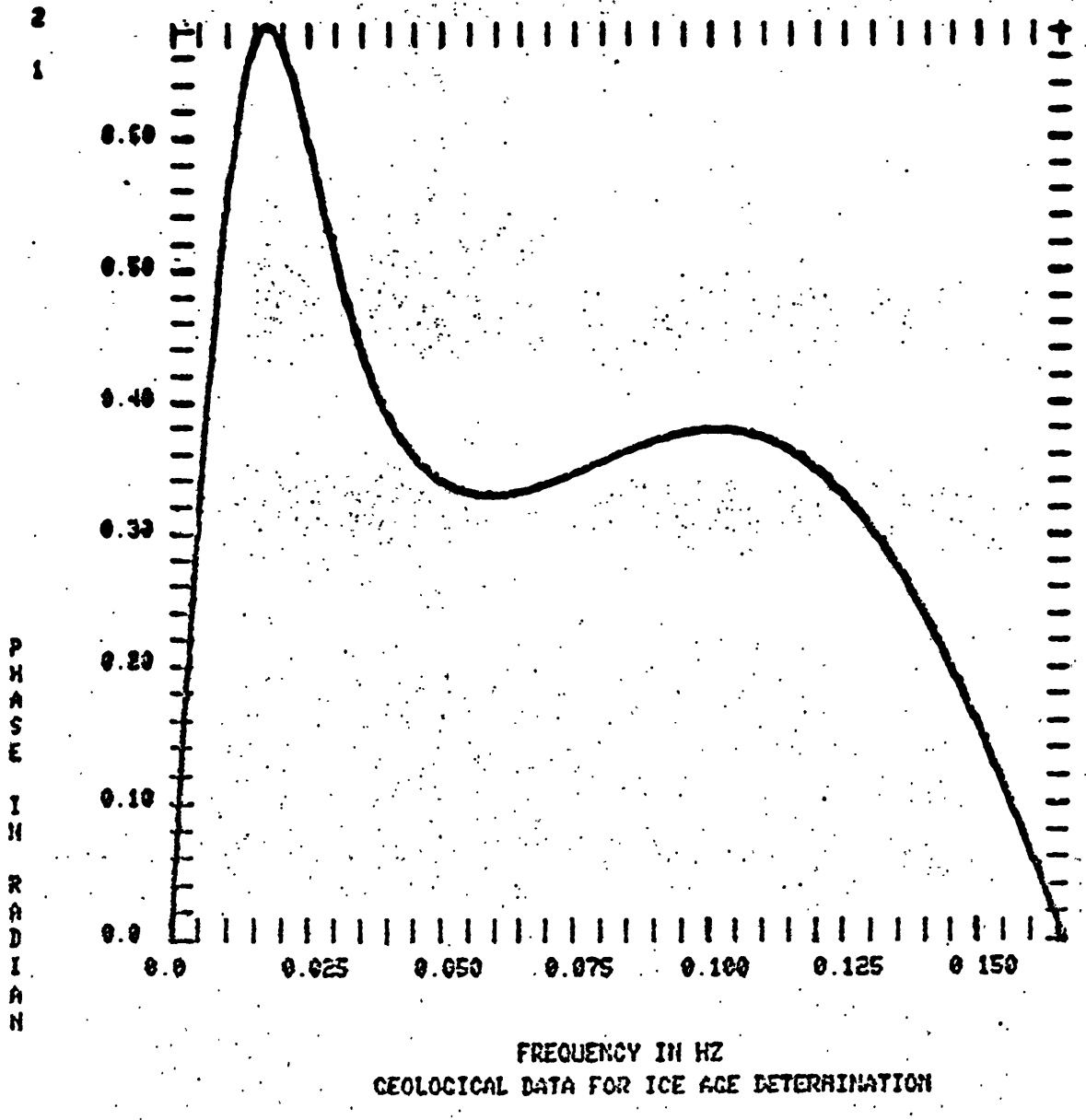

2
1
0.60
0.50
0.40
0.30
0.20
0.10
0.0
PHASE IN RADIAN
0.0
0.025
0.050
0.075
0.100
0.125
0 150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 26(a)

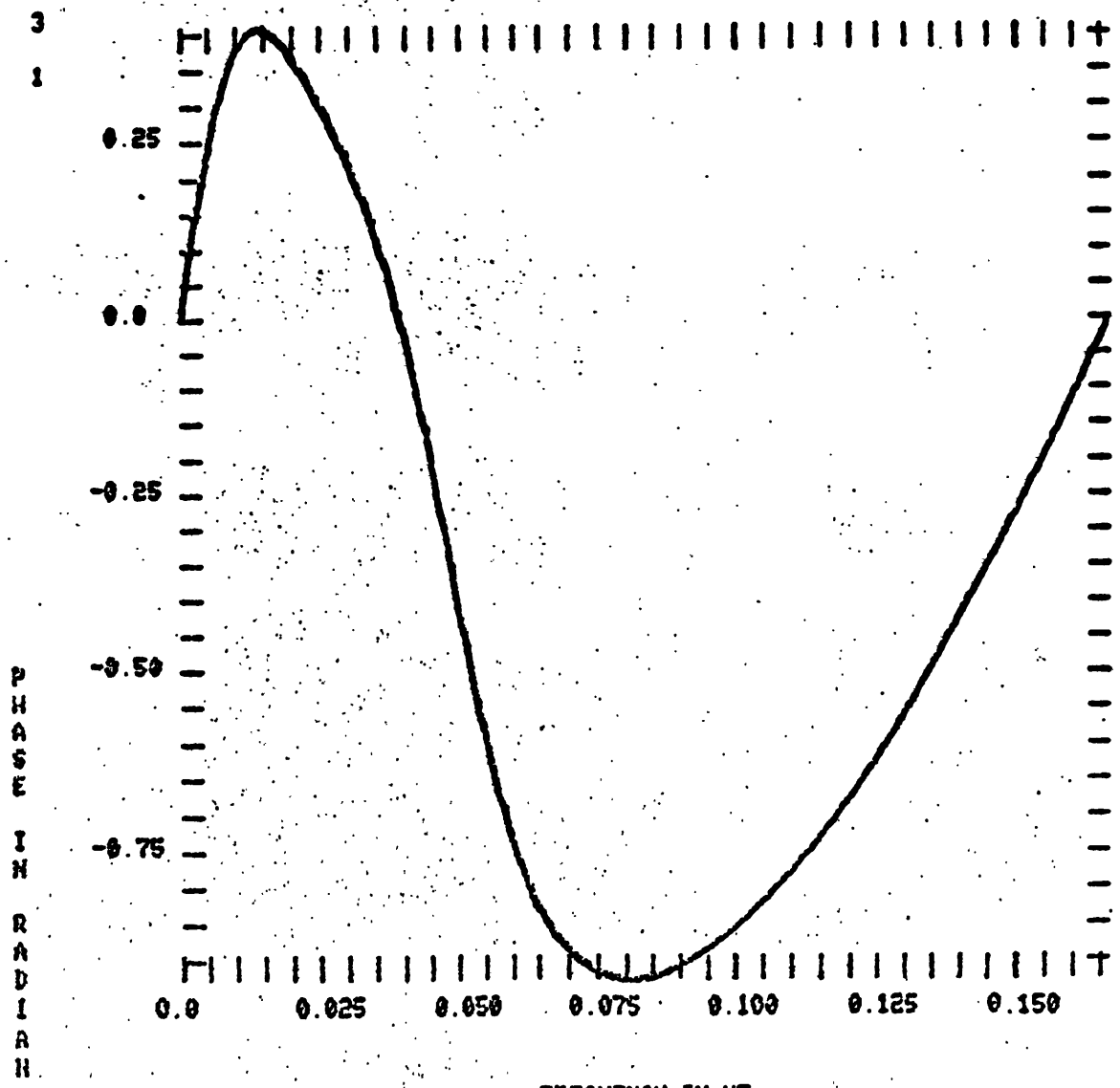

3
1
0.25
0.0
-0.25
-0.50
-0.75
PHASE IN RADIAN
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 26(b)

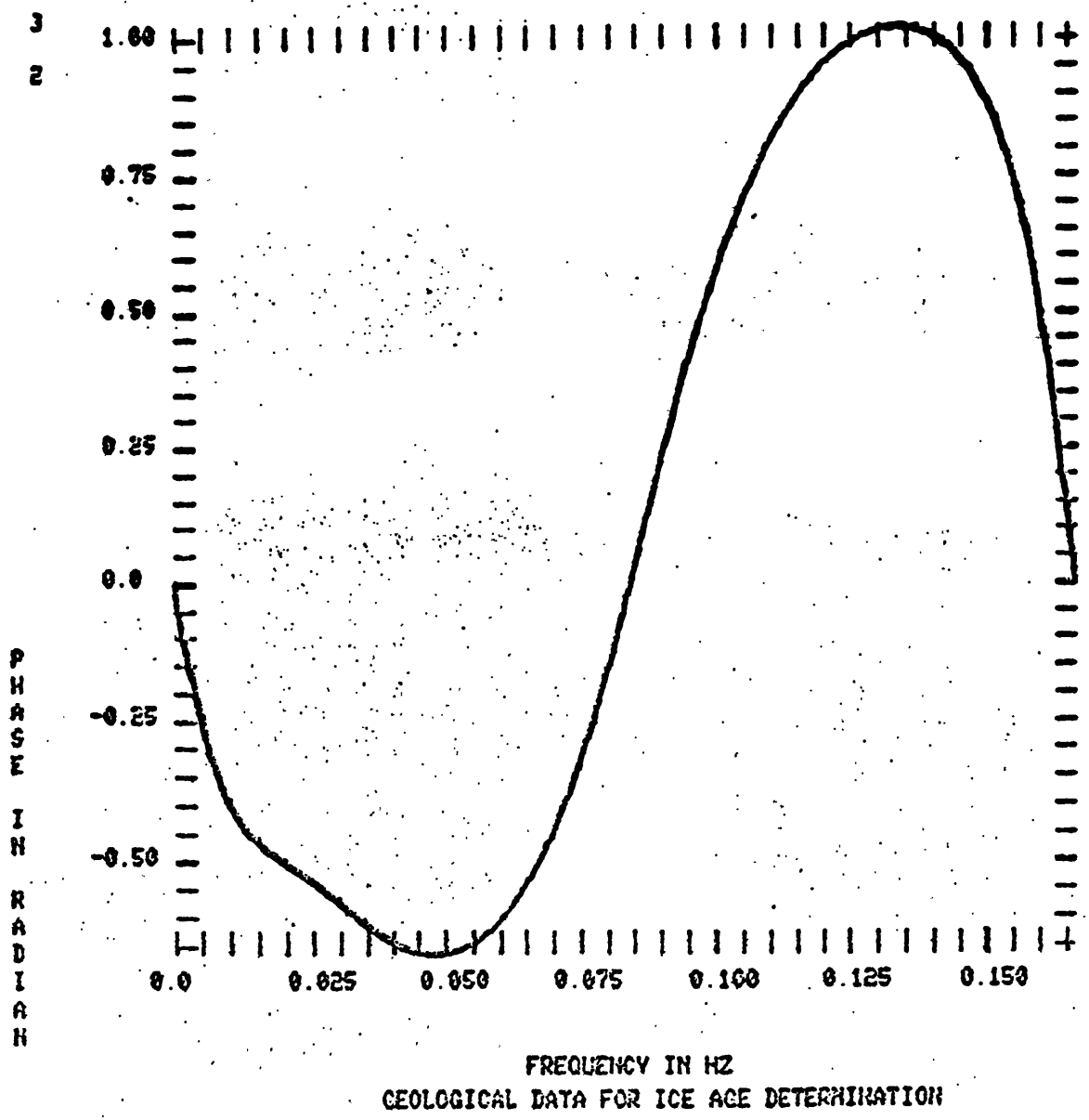

3
2
1.00
0.75
0.50
0.25
0.0
-0.25
-0.50
PHASE IN RADIAN
0.0
0.025
0.050
0.075
0.100
0.125
0.150
FREQUENCY IN HZ
GEOLOGICAL DATA FOR ICE AGE DETERMINATION

Figure 26(c)

MINIMUM IS 0.112204E+01 AT ORDER 1

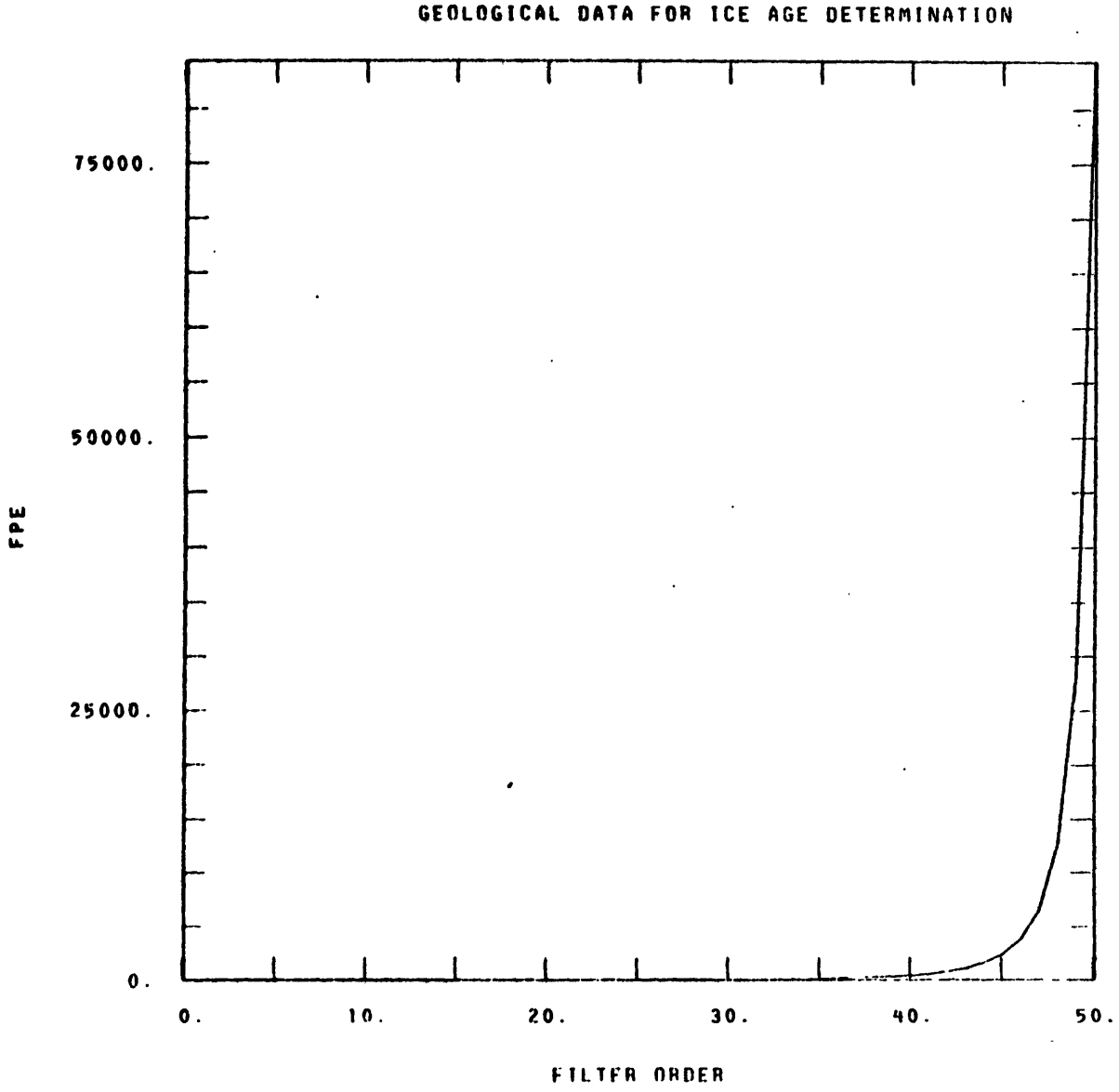


Figure 27(a)

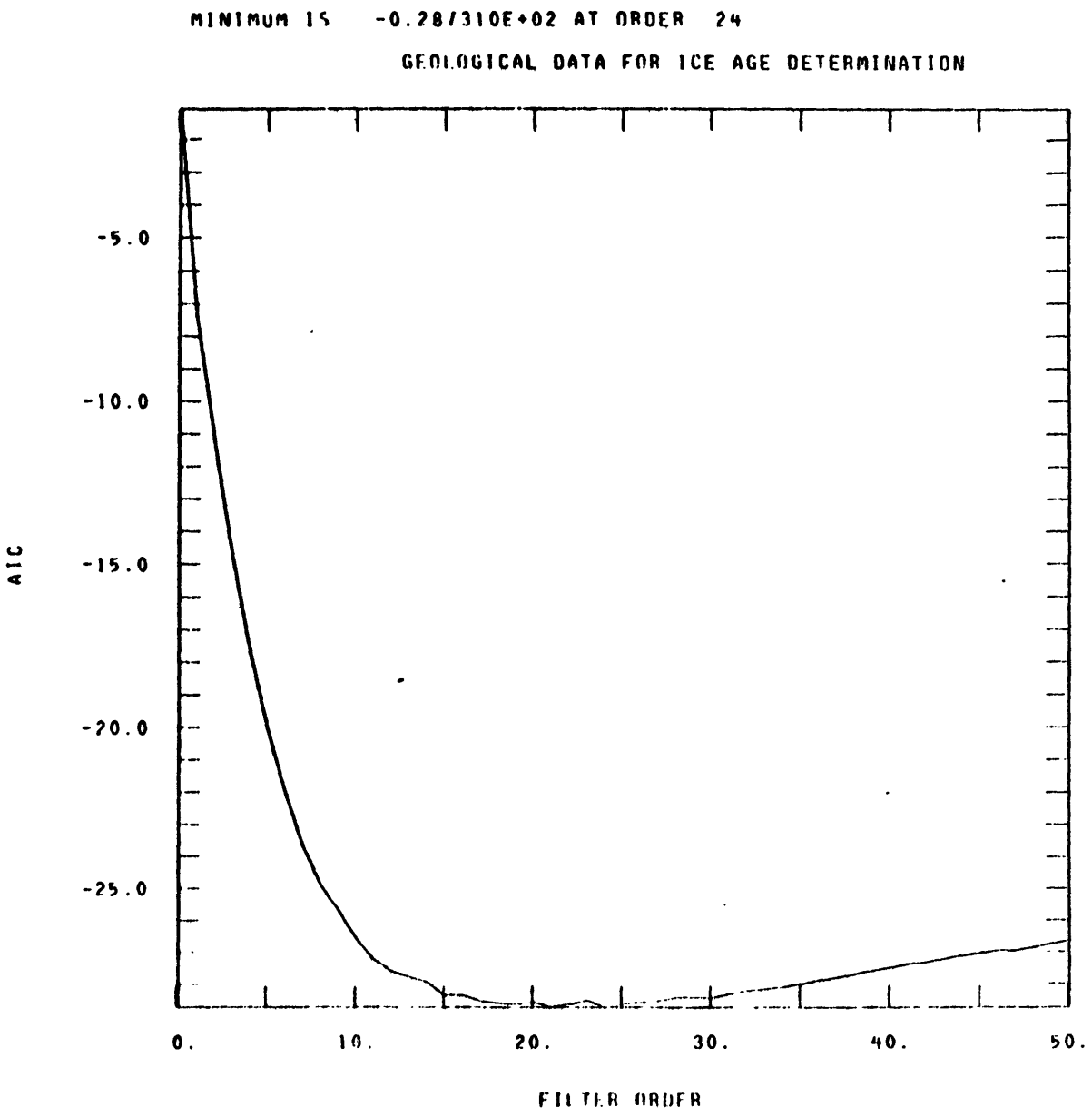

MINIMUM IS -0.28731OE+02 AT ORDER 24
GEOLOGICAL DATA FOR ICE AGE DETERMINATION
-5.0
-10.0
-15.0
-20.0
-25.0
AIC
0.
10.
20.
30.
40.
50.
FILTER ORDER

Figure 27(b)

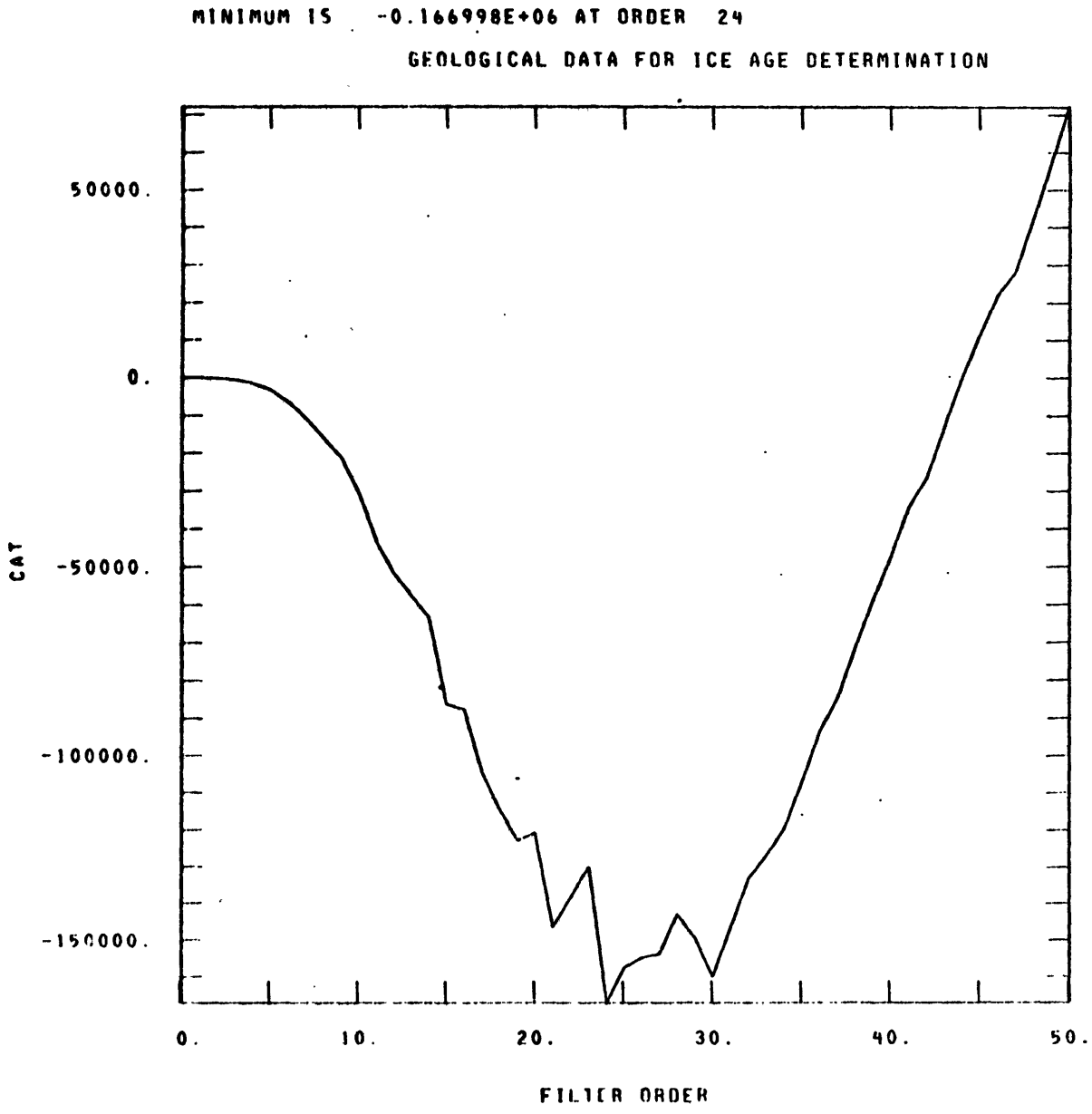

MINIMUM IS -0.166998E+06 AT ORDER 24
GEOLOGICAL DATA FOR ICE AGE DETERMINATION
50000.
0.
-50000.
-100000.
-150000.
CAT
0.
10.
20.
30.
40.
50.
FILTER ORDER

Figure 27(c)

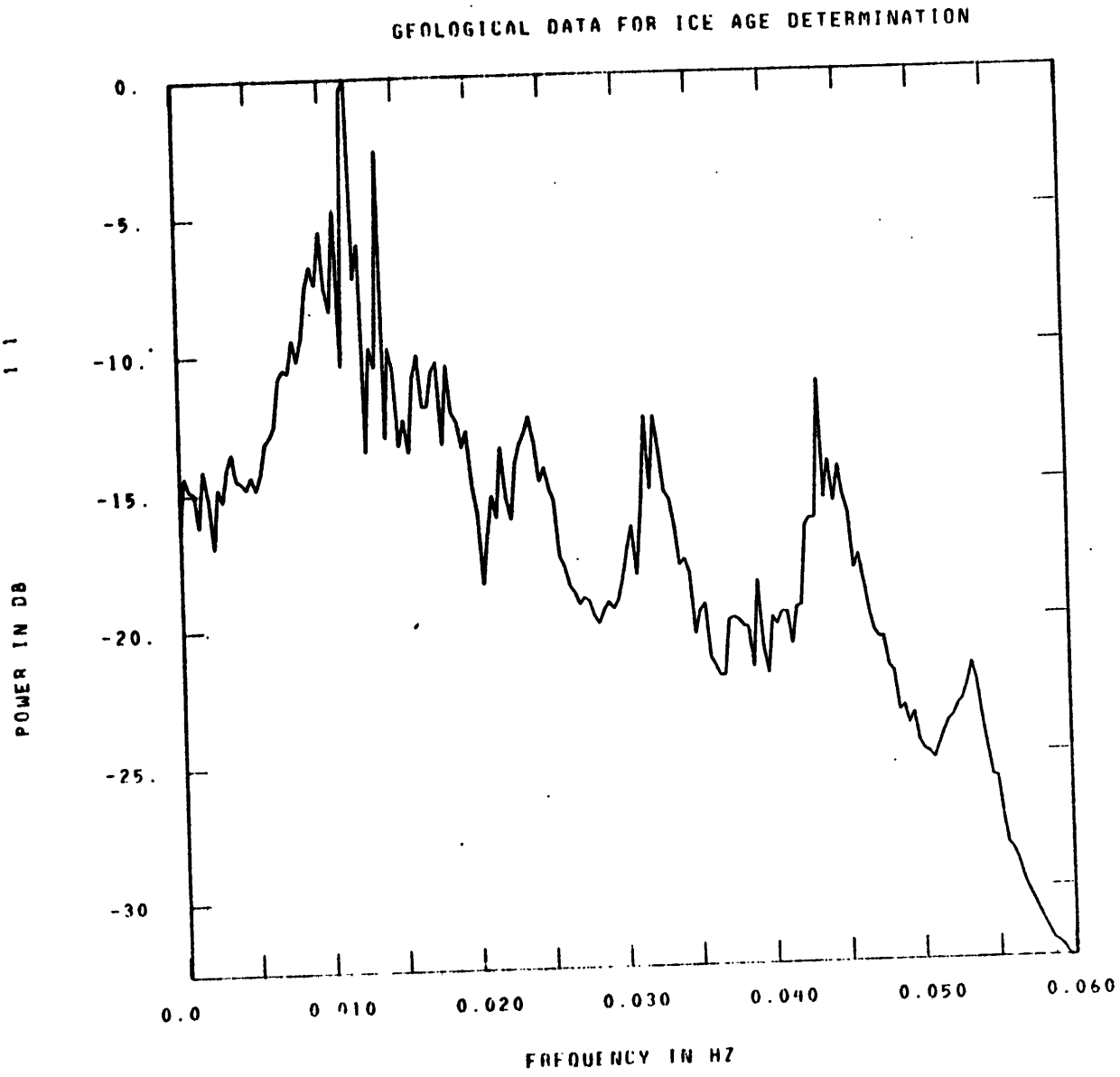

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
0.
-5.
-10.
-15.
-20.
-25.
-30
POWER IN DB
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 28(a)

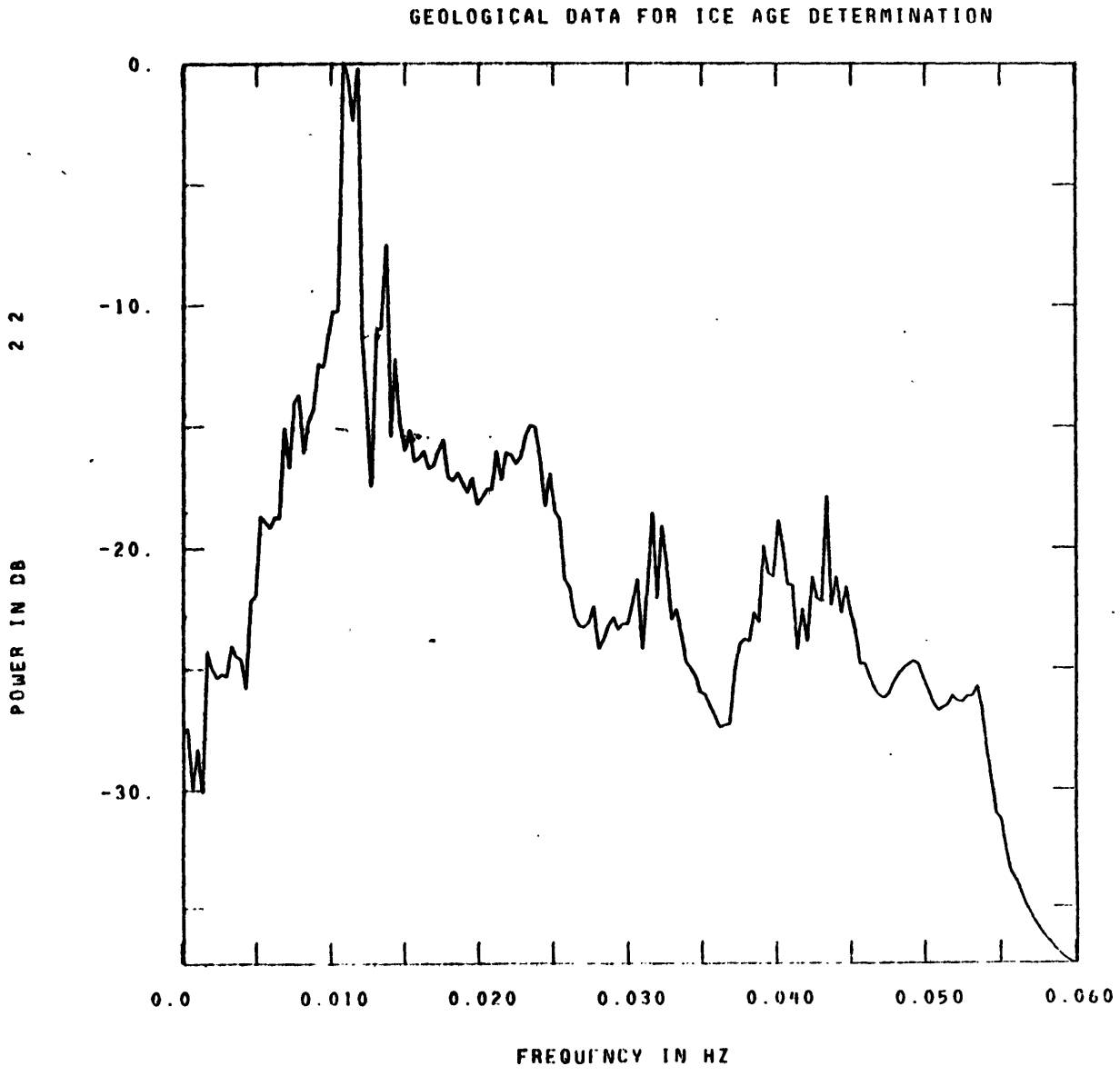

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
0.
-10.
-20.
-30.
POWER IN DB
2 2
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 28(b)

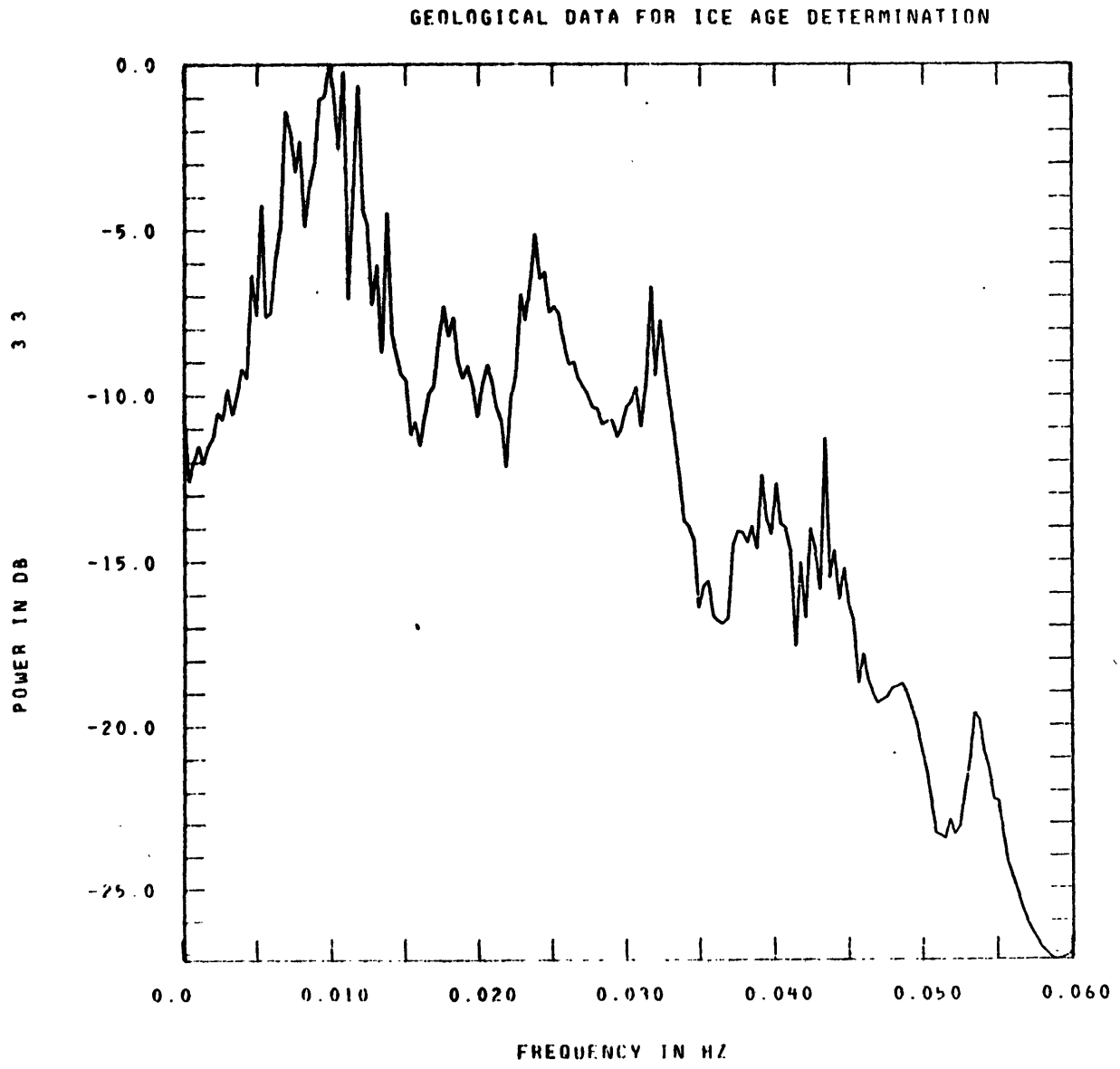

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
POWER IN DB
0.0
-5.0
-10.0
-15.0
-20.0
-25.0
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 28(c)

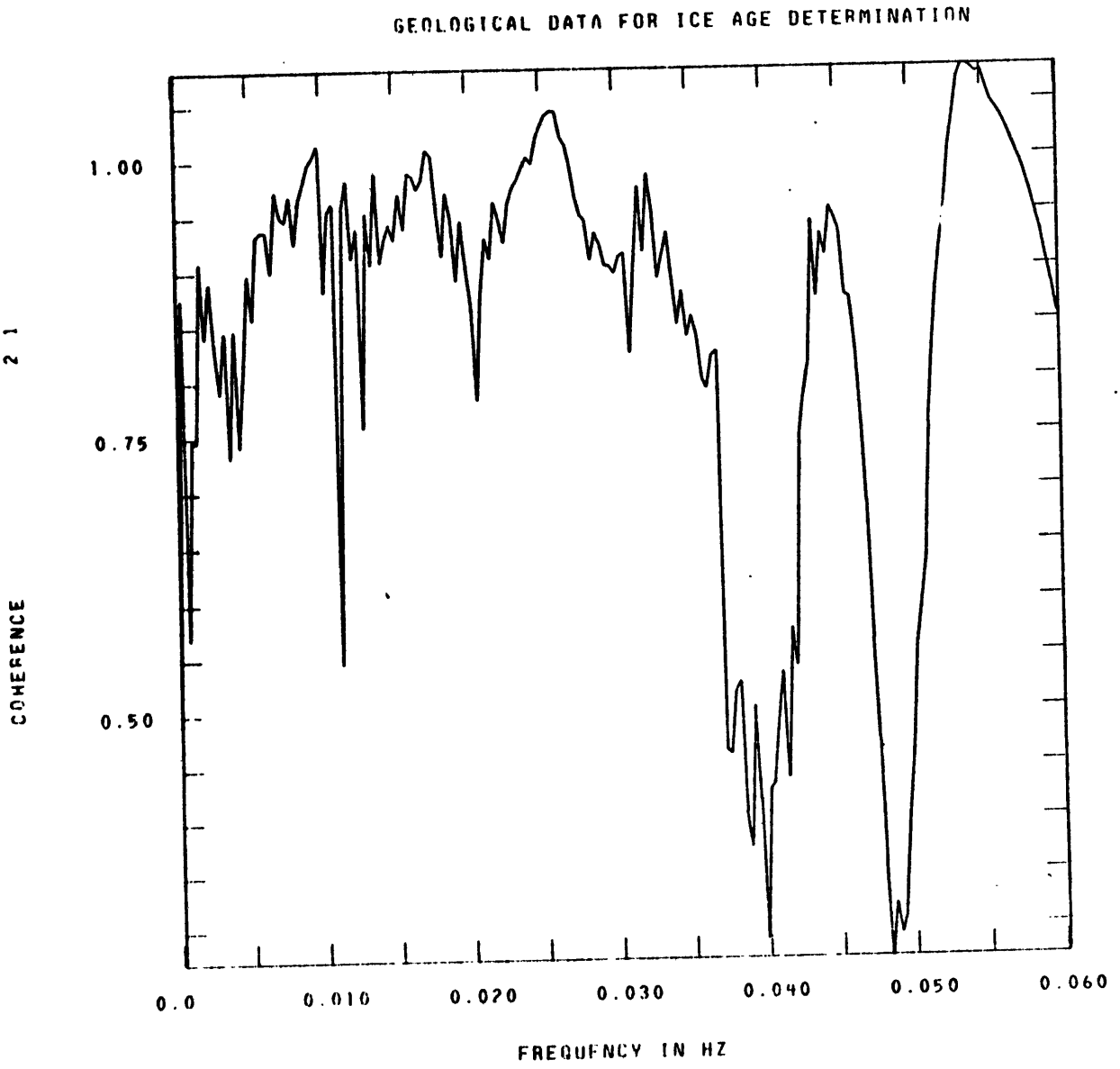

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
1.00
0.75
0.50
COHERENCE
2 1
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 29(a)

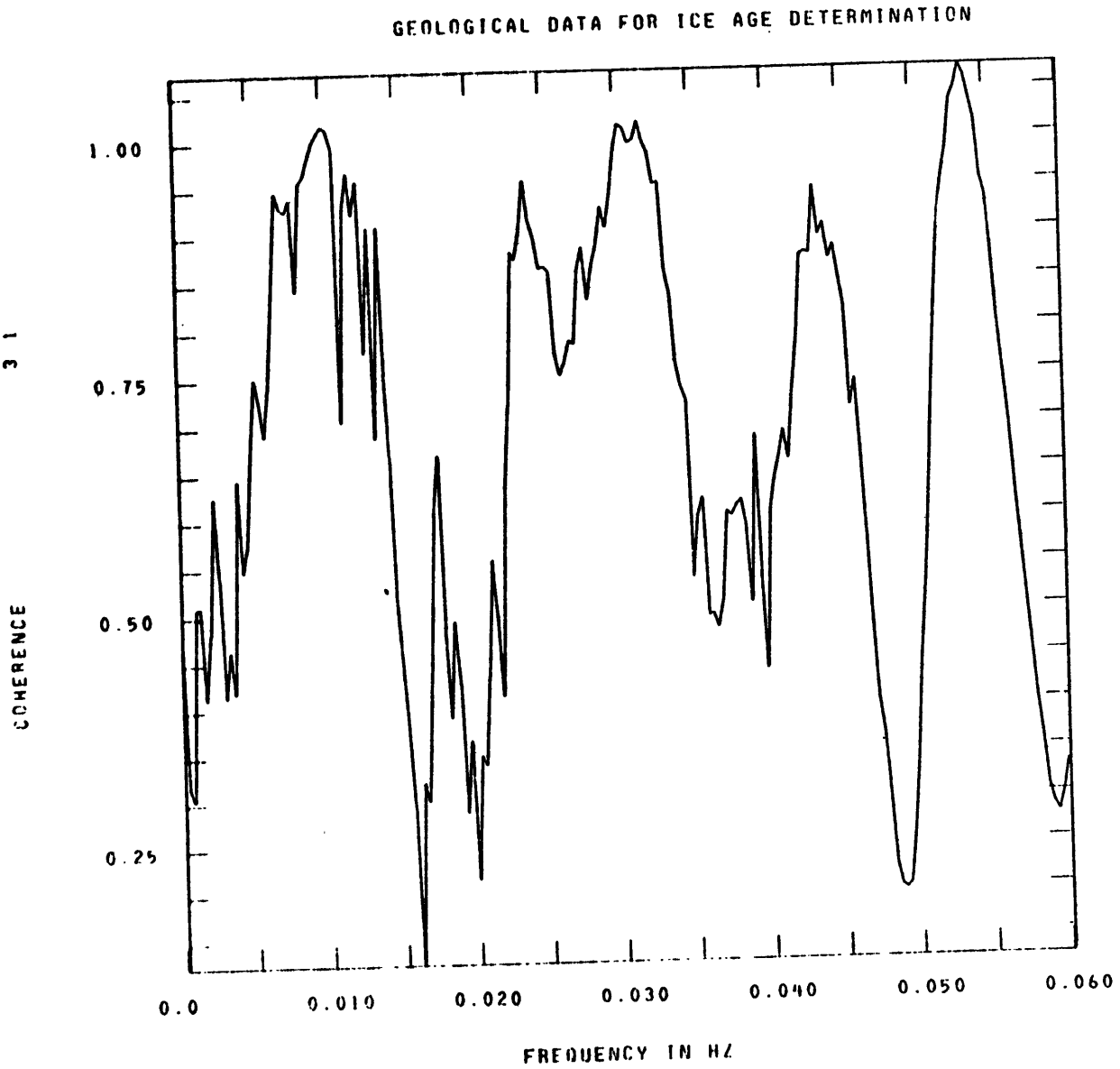

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
1.00
0.75
0.50
0.25
COHERENCE
3 1
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 29(b)

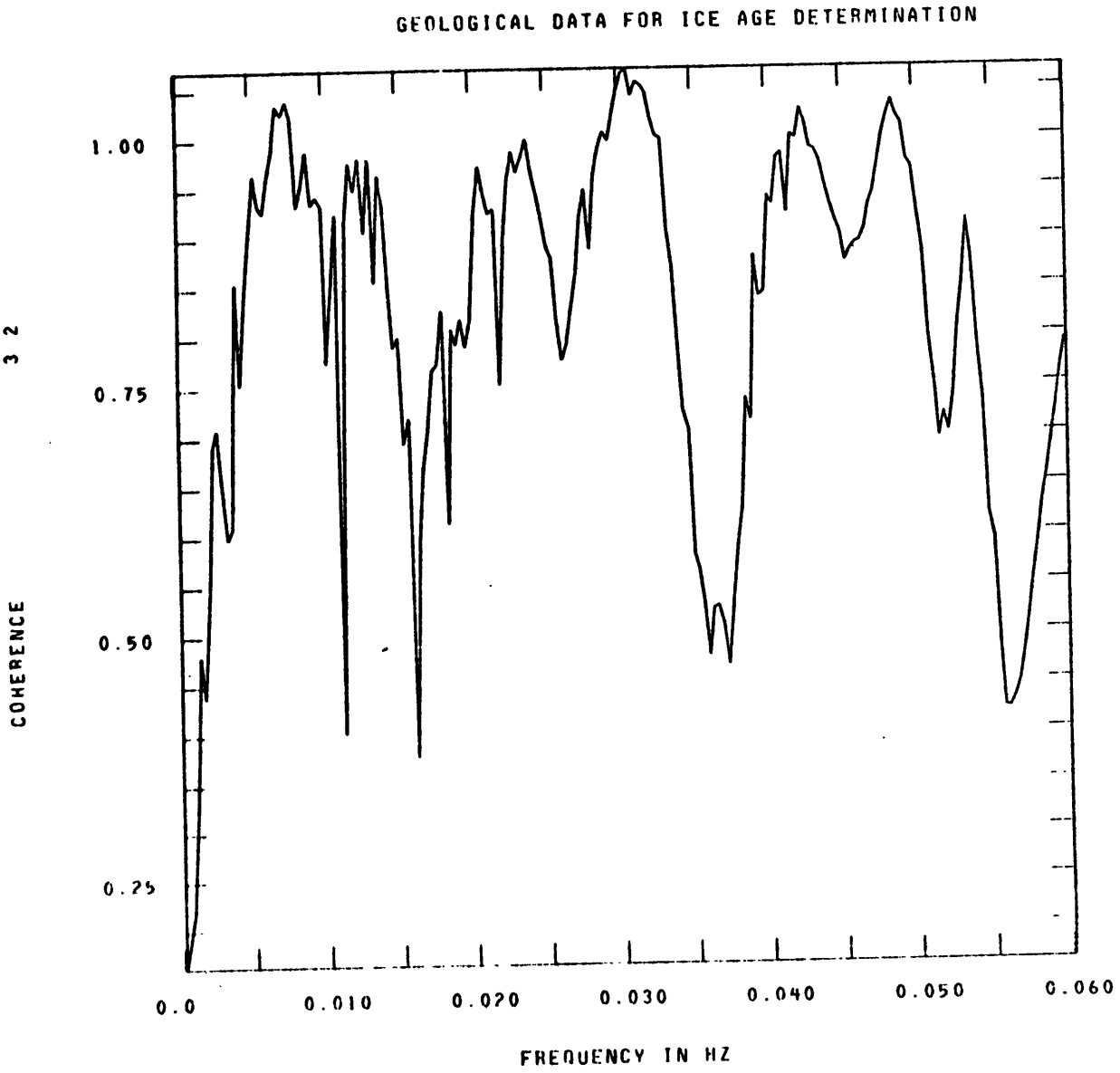

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
3 2
COHERENCE
1.00
0.75
0.50
0.25
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 29(c)

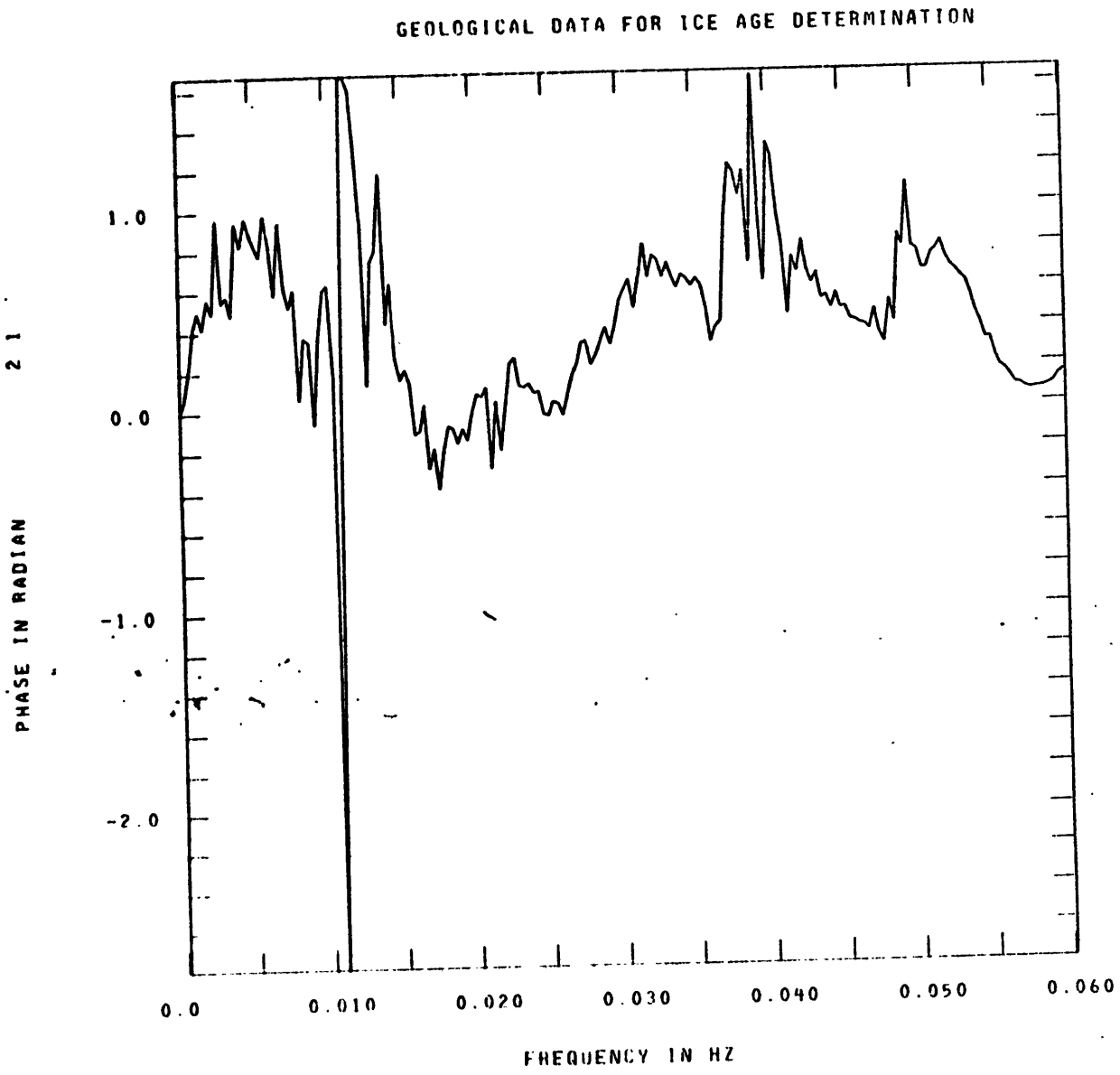

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
PHASE IN RADIAN
2 1
1.0
0.0
-1.0
-2.0
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 30(a)

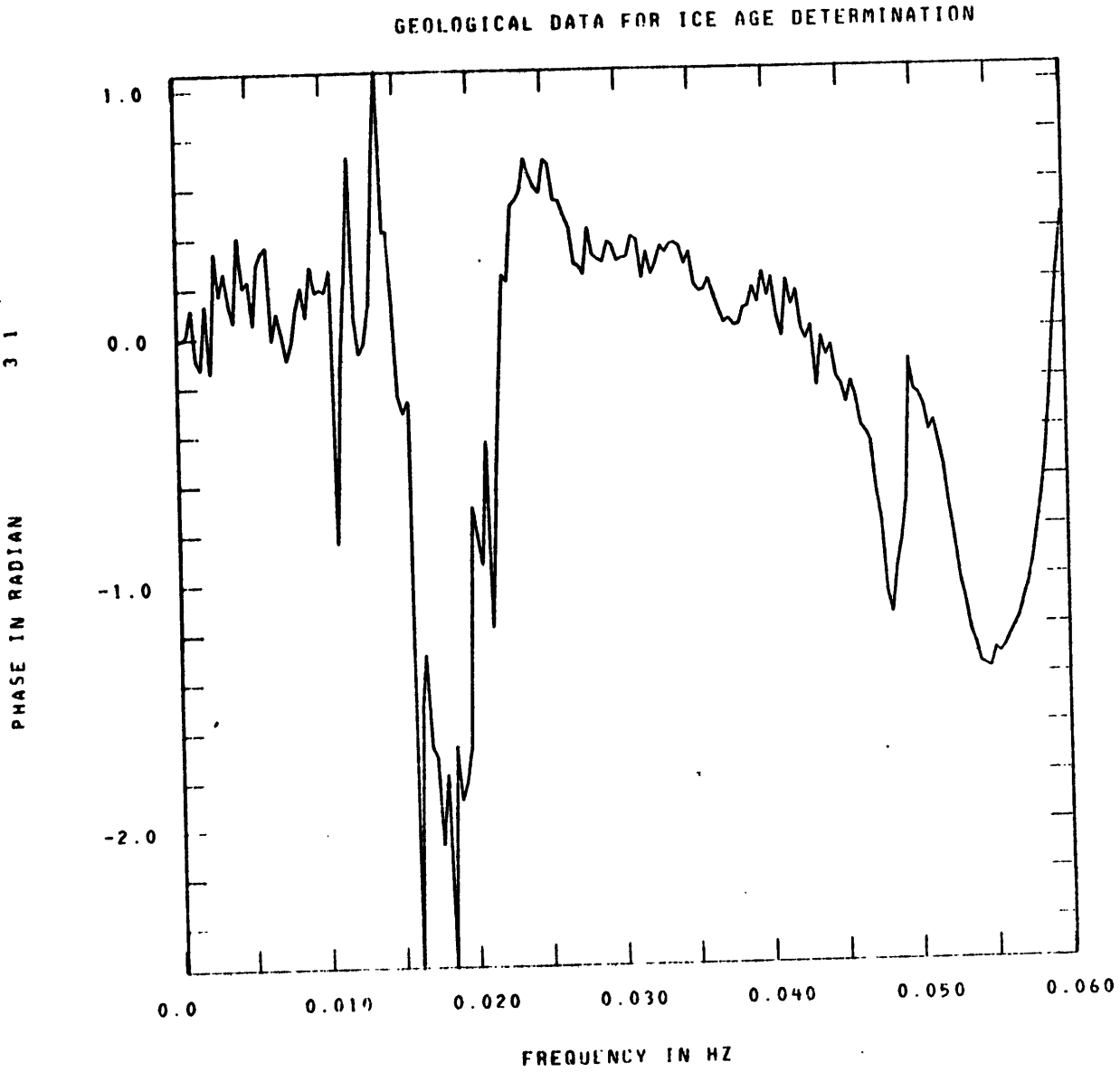

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
PHASE IN RADIAN
3 1
1.0
0.0
-1.0
-2.0
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 30(b)

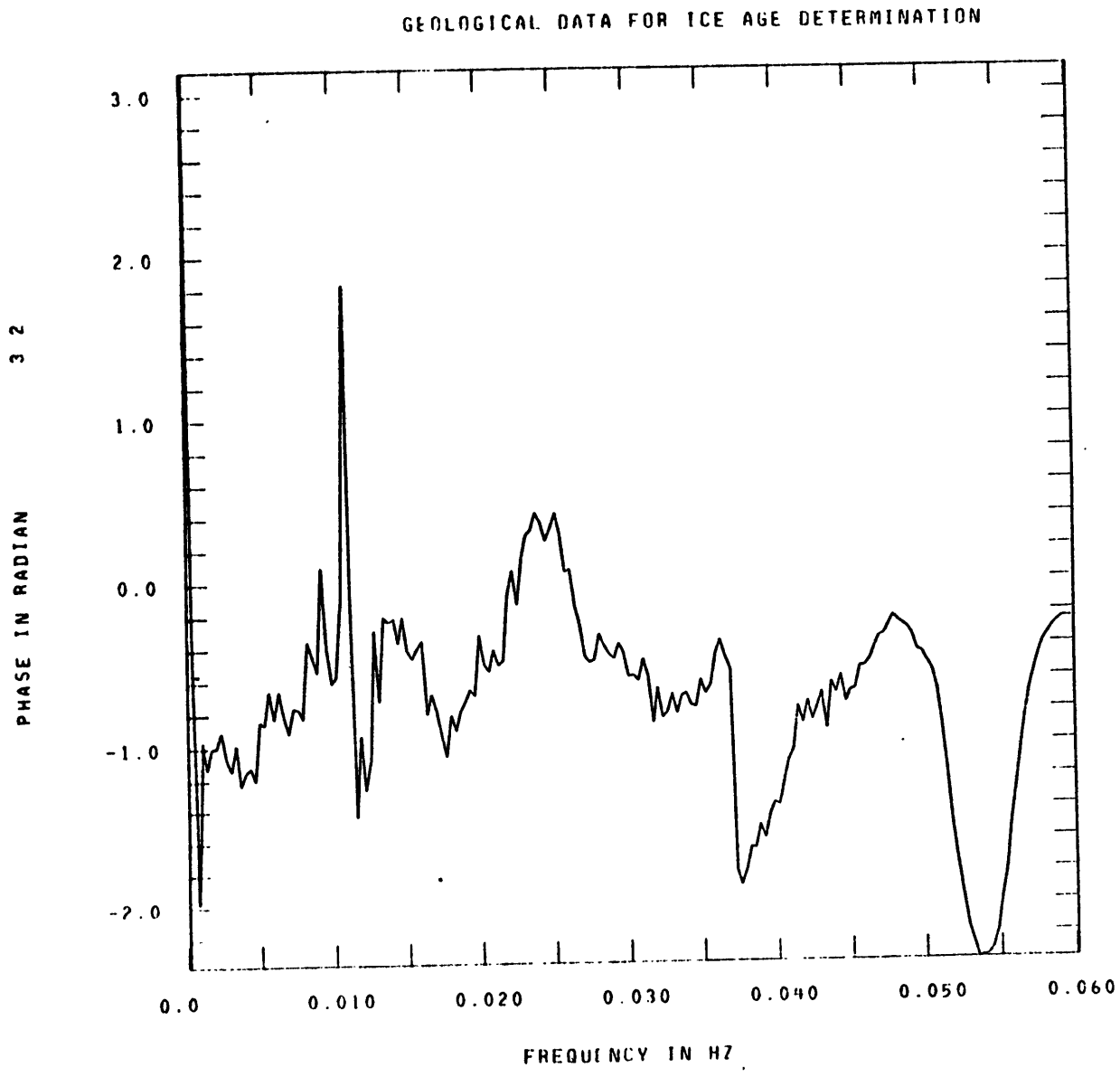

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
3.0
2.0
1.0
0.0
-1.0
-2.0
PHASE IN RADIAN
3 2
0.0
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 30(c)

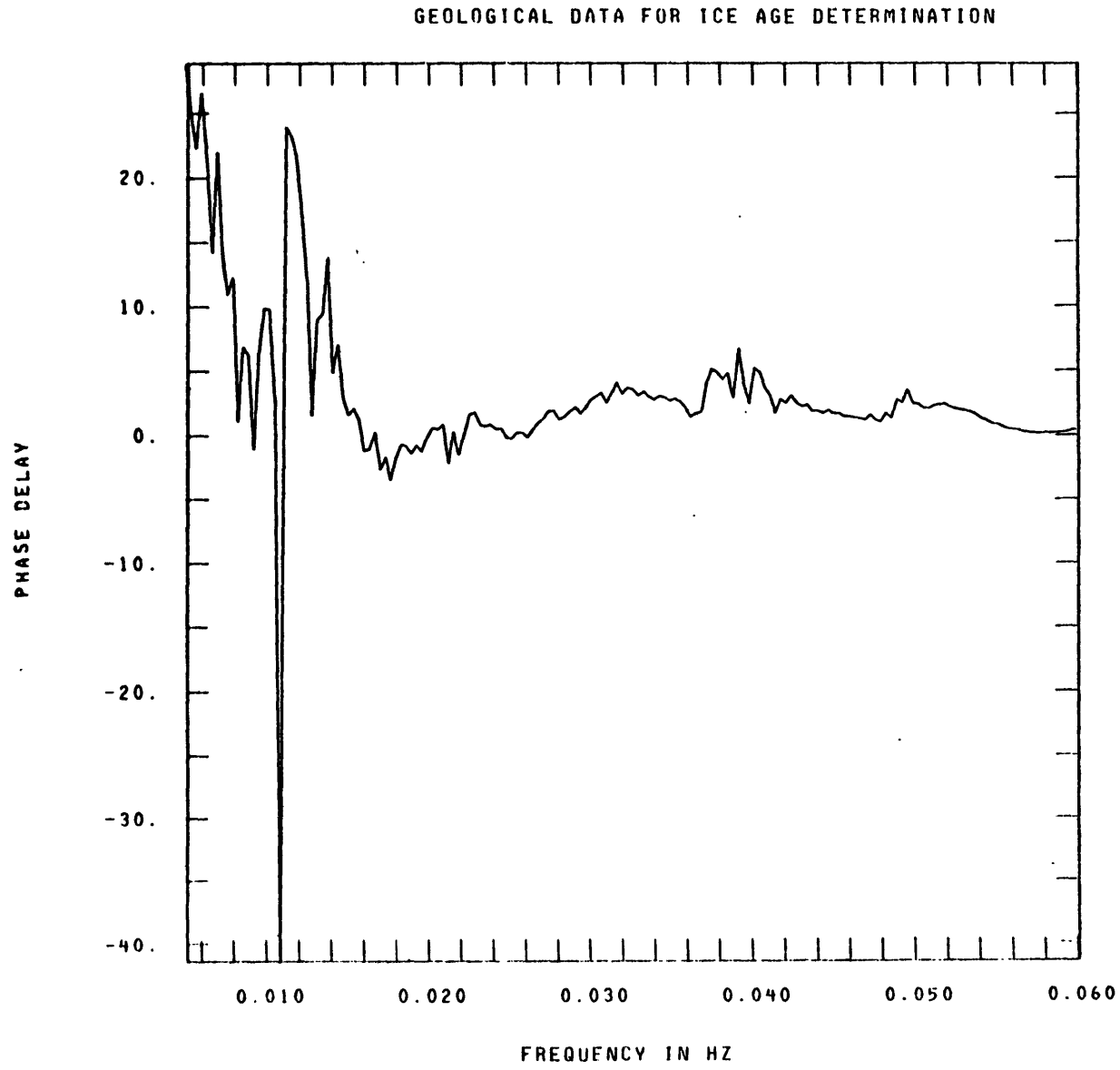

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
PHASE DELAY
20.
10.
0.
-10.
-20.
-30.
-40.
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 31(a)

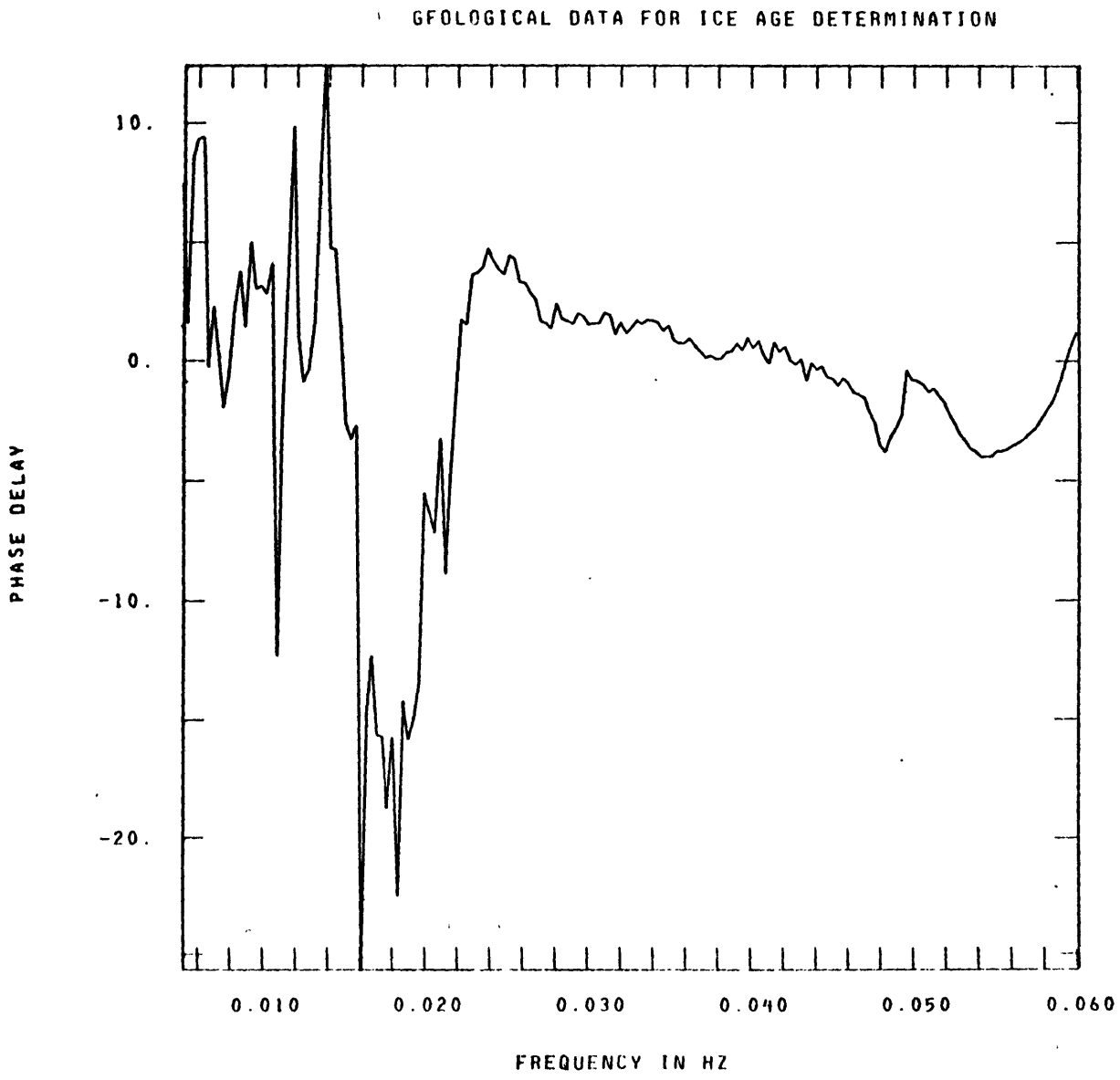

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
10.
0.
-10.
-20.
PHASE DELAY
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 31(b)

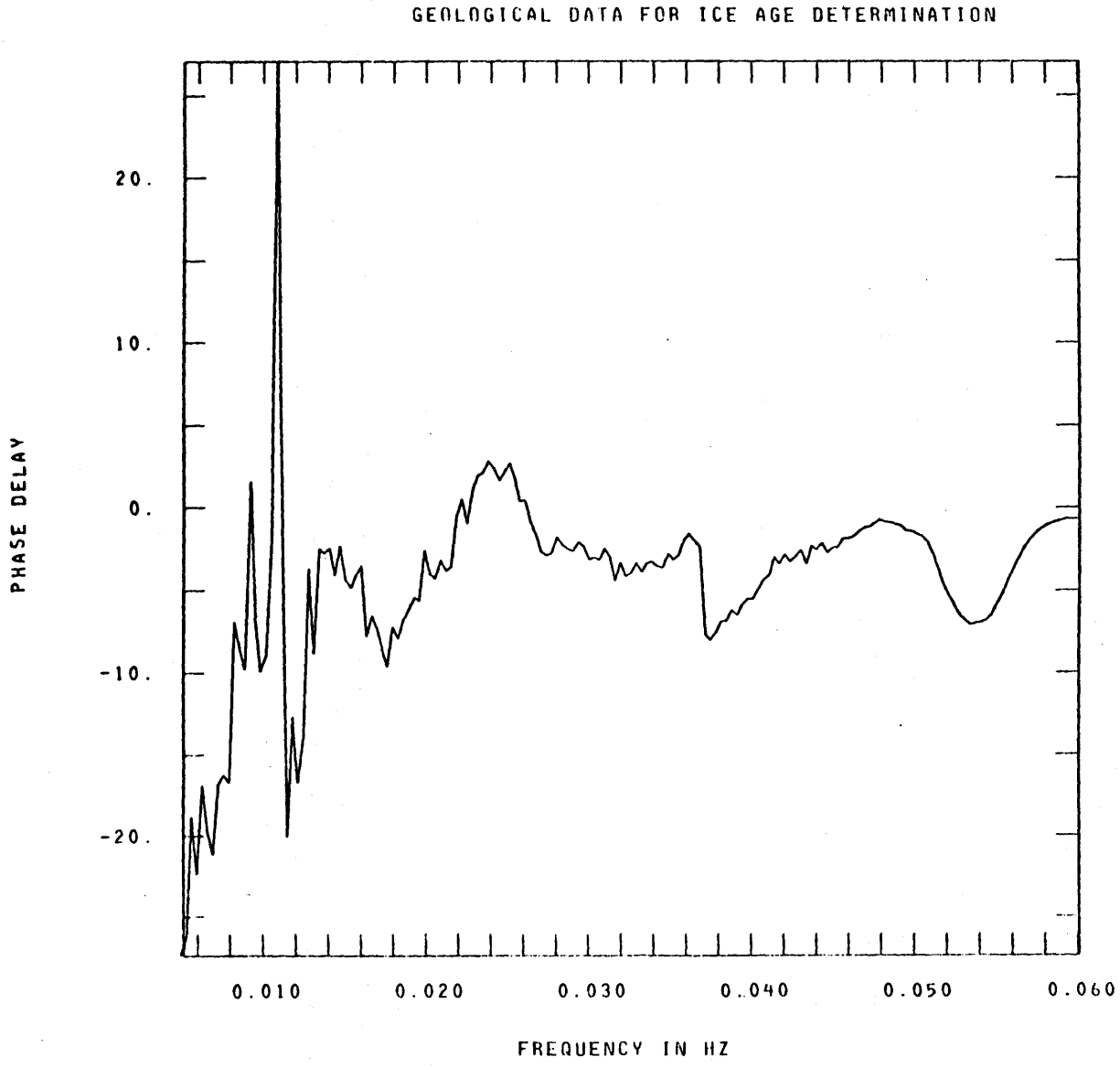

GEOLOGICAL DATA FOR ICE AGE DETERMINATION
PHASE DELAY
20.
10.
0.
-10.
-20.
0.010
0.020
0.030
0.040
0.050
0.060
FREQUENCY IN HZ

Figure 31(c)

ACKNOWLEDGEMENTS

The author is most grateful to Professor M. Nafi Toksoz, his advisor, who provided guidance, support and encouragement in his studies at M.I.T.

Equal gratitude is offered to Dr. Richard Lacoss who suggested this project and to Dr. Tom Landers, both of the Applied Seismology Group, Lincoln Laboratory, who guided the research presented in this thesis.

The autor also wishes to express his gratitude to Mrs. Karen A. North for her excellent typing of the manuscript.

This research was sponsored by the Advanced Research Projects Agency of the Department of Defense.