

UMVU Estimation of Phase and Group Delay With Small Samples

by

Philip J. Ramsey

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Statistics

APPROVED:

Robert V. Foutz, Chairman

Jesse C. Arnold

Klaus H. Hinkelmann

Raymond H. Myers

Marion R. Reynolds

June 16, 1989

Blacksburg, Virginia

UMVU Estimation Of Phase And Group Delay With Small Samples

by

Philip J. Ramsey

Robert V. Foutz, Chairman

Statistics

(ABSTRACT)

Group delay between two univariate time series is a measure, in units of time, of how one series leads or lags the other at specific frequencies. The only published method of estimating group delay is Hannan and Thomson (1973); however, their method is highly asymptotic and does not allow inference to be performed on the group delay parameter in finite samples. In fact, spectral analysis in general does not allow for small sample inference which is a difficulty with the frequency domain approach to time series analysis. The reason that no statistical inference may be performed in small samples is the fact that distribution theory for spectral estimates is highly asymptotic and one can never be certain in a particular application what finite sample size is required to justify the asymptotic result.

In the dissertation the asymptotic distribution theory is circumvented by use of the Box-Cox power transformation on the observed sample phase function. Once transformed, it is assumed that the sample phase is approximately normally distributed and the relationship between phase and frequency is modelled by a simple linear regression model. In order to estimate group delay it is necessary to inversely transform the predicted values to the original scale of measurement and this is done by expanding the inverse Box-Cox transformation function in a Taylor Series expansion. The group delay estimates are generated by using the derivative of the Taylor Series expansion for phase. The UMVUE property comes from the fact that the Taylor Series expansions are functions

of complete, sufficient statistics from the transformed domain and the Lehmann-Scheffe' result (1950) is invoked to justify the UMVUE property.

Acknowledgements

I would sincerely like to thank my thesis advisor Dr. Robert V. Foutz for his patience, support, and guidance in my academic career at V.P.I.&S.U. There is no doubt that none of this would have been possible without his assistance. I would also like to thank Dr. Raymond H. Myers for his confidence in me and support when it was needed and I certainly extend the same statement of gratitude to Dr. Marion R. Reynolds. I also want to sincerely thank Dr. Klaus H. Hinkelmann and Dr. Jesse C. Arnold for their willingness to serve on my committee and their time and assistance in completing my program and the final preparation of this dissertation. In fact, I wish to thank all of the members of the faculty for providing an education which is of great value in today's highly competitive job market. The time, effort, and occasional anxiety of completing a Ph.D. program is certainly a small price to pay for the potential rewards of a degree from our Statistics Department. Lastly, but far from least, I wish to dedicate this dissertation to my wife and daughters , , and . Although much of the time expended in completing my de-

gree has been agreeably spent, a good deal of that time was arguably theirs and hopefully the future rewards will adequately compensate their past sacrifices.

Table of Contents

Introduction to Group Delay Estimation.	1
I.1. Introduction	1
I.2. The Concept Of Group Delay.	2
I.3. Review Of The Literature On Group Delay Estimation.	4
Concepts In Spectral Analysis.	8
II.1. Introduction.	8
II.2. The Autospectrum.	10
II.3. Bivariate Processes.	13
II.4. The Linear Filter.	19
The Box-Cox Power Transformation.	22
III.1. Introduction.	22

III.2. Selection Of The Power Parameter.	25
III.2.1. Approximate Confidence Intervals For The Parameter.	27
III.3. Variance Inflation In The Box-Cox Transformation.	28
UMVU Estimation of Phase and Group Delay.	31
IV.1. Introduction.	31
IV.2. The Problem Of Inverse Transformation Bias.	34
IV.3. Correcting For Bias In the Inverse Transformation.	35
IV.4. Modeling Phase and Group Delay By Transformations.	41
IV.4.1. The UMVUE of The Phase Function.	42
IV.4.2. An UMVUE For The Group Delay Function.	50
IV.4.3 Estimating Group Delay By Finite Difference.	59
IV.5. Estimating The Variance Of The UMVU Estimators.	61
IV.5.1. Approximate Confidence Intervals For The UMVUE Values.	63
IV.6. Alternative Inverse Transformation Strategies.	64
Appendix A.IV.	67
Simulation Study of The UMVUE's For Phase and Group Delay.	71
V.1. Introduction.	71
V.1.1. The Continuous Time "Scratch" Filter.	74
V.2. Generating A Random Sample Phase Function.	77
V.3. Box-Cox Transformation Of The Sample Phase Observations.	80
V.4. UMVU Estimation Of Phase And Group Delay.	87
V.4.1. The Results Of The Study For The Cube Root Case.	89

V.4.2. Results Of The Study For Square Root Case.	98
Applications Of Group Delay Estimation.	113
VI.1. Introduction.	113
VI.2. Lobster Harvest And Sea Water Temperature.	114
VI.3. Sunspots vs. Consumption Of Alcohol By Londoner's.	133
Summary And Topics For Future Research.	148
VII.1. Summary And Conclusions.	148
VII.2. Proposed Topics For Future Research.	149
BIBLIOGRAPHY	152
Vita	157

List Of Illustrations

Figure V.1 Plot Of Random And Actual Phase Function, $t = 1.0$.	79
Figure V.2 L_{max} For 2 Parameter Box-Cox Transformation.	83
Figure V.3 L_{max} For 2 Parameter Box-Cox Transformation.	84
Figure V.4 Plot Of Transformed Phase $\text{Eta}_1 = .3333$ And $T = 1.0$.	86
Figure V.5 RSTUDENT Residuals For Transformed PHase $T = 1$,	88
Figure V.6 UMVUE And Naive Estimate Of Phase $T = 1$, $\text{Eta}_1 = 1/3$,	92
Figure V.7 Plot Of Actual And UMVUE Phase For $T = 1$, $\text{Eta}_1 = 0.3333$,	93
Figure V.8 Actual, UMVUE, And Finite GD For $T = 1$ $\text{Eta}_1 = 0.3333$,	95
Figure V.9 Comparison Of Variance For Phase And GD $T = 1$ $\text{Eta}_1 = 0.3333$.	96
Figure V.10 Comparison Of MSE For Phase And GD $T = 1$ $\text{Eta}_1 = 0.3333$.	99
Figure V.11 Comparison Of Variance For UMVUE And Delta GD $T = 1$	100
Figure V.12 Comparison Of MSE For UMVUE And Delta GD $T = 1$	101
Figure V.13 Plot Of Estimated Phase And .95 CI, $\text{Eta}_1 = 0.3333$ $T = 1.0$.	102

Figure V.14 UMVUE And Actual Phase For $N_{rep} = 1,000$ $E_{tal} = 0.5000$	105
Figure V.15 UMVUE, Finite And Actual GD For $N_{rep} = 1,000$ $E_{tal} = 0.500$	106
Figure V.16 Comparison Of Bias For UMVUE And Delta GD $T = 1$ $E_{tal} = 0.50$	107
Figure V.17 Comparison Of Variance For UMVUE And Delta GD $T = 1$	109
Figure V.18 Comparison Of MSE For UMVUE And Delta GD $T = 1$ $E_{tal} = 0.50$	110
Figure V.19 Comparison Of Variance For Phase And GD $T = 1$ $E_{tal} = 0.50$	111
Figure V.20 Comparison Of MSE For Phase And GD $T = 1$ $E_{tal} = 0.50$	112
Figure VI.1 Plot Of Annual Lobster Catch 1897-1984.	116
Figure VI.2 Plot Of Detrended Annual Lobster Catch 1897-1984.	118
Figure VI.3 Plot Of Annual Sea Surface Temperature 1897-1984.	120
Figure VI.4 Plot Of Spectral Density Of Harvest.	122
Figure VI.5 Plot Of Spectral Density Of Temperature.	123
Figure VI.6 Plot Of Squared Coherence Function.	124
Figure VI.7 Plot Of Raw Phase Function.	125
Figure VI.8 Plot Of Unwrapped Phase Function.	127
Figure VI.9 RSTUDENT For Transformed Phase Values Of Catch And Temp.	128
Figure VI.10 UMVUE And .95 CI Estimates For Catch And Temp.	130
Figure VI.11 UMVUE's For GD Of Lobster Harvest And Water Temperature.	131
Figure VI.12 Plot Of Coefficients For GD Model Of Catch And Temp.	132
Figure VI.13 London Per Capita Spirit Consumption 1870-1938.	134
Figure VI.14 Wolfer Sunspot Numbers For 1870-1938.	135
Figure VI.15 Spectral Density Of Sunspots For 1870-1938.	136
Figure VI.16 Spectral Density Of Consumption For 1870-1938.	137
Figure VI.17 Coherence Between Consumption And Sunspots 1870-1938.	138

Figure VI.18 Sample Phase For Consumption And Sunspots From 1870-1938	141
Figure VI.19 Unwrapped Phase For Consumption And Sunspots	142
Figure VI.20 RSTUDENT For Fit Of Trans. Phase For Sunspots And	143
Figure VI.21 UMVUE And .95 CI For Phase Of Sunspots And Consumption.	144
Figure VI.22 UMVUE's For GD Of Sunspots And Consumption.	146
Figure VI.23 Beta Hat For GD Model Of Consumption And Sunspots.	147

Chapter I

Introduction to Group Delay Estimation.

I.1. Introduction

The dissertation is divided into 6 chapters. The reader who is unfamiliar with time series concepts, particularly spectral analysis, may wish to read Chapter II before reading the remaining sections of Chapter I and the other 4 chapters. Chapter II is designed to provide intuitive and concise definitions of the concepts required for the discussion in much of the dissertation. Chapter III may be skipped if one is familiar with the particulars of the Box-Cox power transformation. However, section III.3 contains a review of controversial topics surrounding the use and interpretation of the Box-Cox power transformation; and the reader may find it of interest even though it is not required for any of the subsequent discussion. The principle result of the dissertation is

contained in Chapter IV which is a theoretical development of an UMVUE technique for group delay and phase. Chapter V is a simulation study to demonstrate the application of the theoretical results in Chapter IV. Finally Chapter VI is an application of the methods developed in Chapter IV to a couple of actual problems. The first problem discusses the estimation of group delay between the annual harvest of the Maine lobster and the annual mean sea water temperature. The second problem addresses estimation of the group delay relationship between the occurrence of sunspots and the consumption of alcoholic beverages by residents of London (a previously unknown scientific result).

The remainder of Chapter I is an intuitive motivation of the concepts of time delay and group delay. This discussion is followed by a general literature review specifically on group delay and a brief review of the literature on simple time delay estimation. A number of other topics raised in the present research (e.g., Box-Cox transformation) are also given specific literature reviews and these are addressed where the topic is introduced.

1.2. The Concept Of Group Delay.

In studying the relationships between two univariate time series, one important property is the time--lag relationship between the series. The study of the time--lag relationship is often referred to as time delay analysis. Although time delay analysis is important to time series analysis, it has been neglected by statisticians. We quote Hannan and Thomson (1988)

There are, however, many other situations in which time delay problems arise, for example, in economics, or in examination of tree rings (Foutz, 1980). In spite of its importance, the problem has not been widely considered by statistical time series analysts.

The term time delay implies a situation in which one time series simply leads or lags another series by some fixed time lag parameter τ ; i.e., the parameter τ is constant for all frequency components of the two time series. Time delay analysis is an important subject in the field of signal processing and can be nicely motivated in that context.

Suppose we have a sonar transducer sending an acoustic signal which is received by two different transducers set at two different distances from the transmitter. Since the acoustic wave travels different distances to each receiver, the signal at the farther transducer is a time delayed version of the signal at the nearer transducer.

We may describe the above example mathematically in the following way. Let $s(t)$ be the state of a signal at time t (assume the signal is stationary and time continuous) and the signal $s(t)$ is received at two different recorders which are at two different distances from the transmitter. Let $x(t)$ be the state of the signal at the nearer receiver at time t and $y(t)$ be the state of the signal at the farther receiver. The time series $\{X_t\}$ and $\{Y_t\}$ may be written as

$$Y(t) = X(t + \tau) + \varepsilon(t) \quad [1]$$

where $\varepsilon(t)$ is an independent and uncorrelated white noise process. In this case, $y(t)$ is a simply--delayed version of $x(t)$ and would be identical except for the two noise processes.

Group delay is a more general time delay parameter than simple time delay. The concept of group delay can be motivated by the same sonar example. Suppose that some of the frequency components of the signal propagate at different velocities in water. Then, one can no longer consider a simple time delay parameter τ between the two received signals. The actual time delay is now a function of the various frequency components in the transmitted signal; i.e.,

$$Y_{\Lambda}(t) = X_{\Lambda}(t + \tau(\Lambda)) + \varepsilon(t) \quad [2]$$

where Λ is a specific Fourier frequency where [2] is valid.

The goal of the present research is to find a method of estimating group delay between two univariate time series. We are particularly interested in an estimator for short data records, say several hundred observations or fewer. Since researchers rarely know the exact relationship between group delay and frequency (linear or nonlinear), we are interested in a general method which allows for curvature in that relationship. In the present research a method is developed which allows the estimation of the group delay function from a general phase function and furthermore provides UMVUE properties in small samples.

1.3. Review Of The Literature On Group Delay Estimation.

The estimation of group delay is discussed by Akaike and Yamanouchi (1963), Hannan and Thomson (1971, 1973), and Cleveland and Parzen (1974). Also, Foutz (1980) adapts the results

of Hannan and Thomson (1973) to estimation of common group delay between two multiple time series. There does not appear to be any original published literature on group delay estimation after Foutz (1980).

The estimation of time delay is discussed by Hamon and Hannan (1974), Hannan and Thomson (1981, 1988) and Chan, et. al. (1978, 1980, 1984). Most of the time delay literature is addressed to highly specialized applications, such as speech analysis or sonar detection and is not of direct interest in a general statistical treatment of time delay estimation. Since the subject is not of direct interest to the present topic of group delay estimation, no attempt is made to extensively review the literature on this topic; however, the papers referenced above appear to be the more important ones for general time delay estimation and the reader is referred to those articles if more information is desired on the subject.

Hannan and Thomson (1973) is the only published general estimation technique for group delay. They assume, in their 1973 paper, that one has two weakly stationary, univariate time series, which have a twice differentiable phase function and the phase function is possibly nonlinear. Under these assumptions, one selects a center Fourier frequency, say ω_0 , where it is desired to estimate the group delay between the two time series. Having selected ω_0 , one then selects a band B of m contiguous Fourier frequencies centered at ω_0 . It is further assumed by Hannan and Thomson (1973) that the relationship between phase and frequency is simple linear over the band B; i.e.,

$$\phi_{xy}(\omega) = \tau\omega, \quad \forall \omega \in B. \quad [3]$$

An intercept is not included in [3] because it is assumed that $\phi_{xy}(\omega) = 0$ for $\omega = 0$. The group delay parameter at the center frequency is then $\phi'_{xy}(\omega) = \tau$ and may have different values over alternative bands B.

The actual Hannan and Thomson (1973) estimate $\hat{\tau}$ of the group delay parameter τ is determined by optimizing the objective function

$$P(\hat{\tau}) = \frac{1}{m} \sum_{\omega \in B} I_{xy}(\omega_j) e^{i\tau\omega_j} \quad [4]$$

where $I_{xy}(\omega_j)$ is the j^{th} cross-periodogram ordinate in the band B. Since [4] is complex valued in general, they define a second objective function

$$Q(\hat{\tau}) = |P(\hat{\tau})|^2 \quad [5]$$

which is the squared norm of [4] and is real valued. The Hannan and Thomson (1973) estimate of τ is that value $\hat{\tau}$ which maximizes [5]. The estimate $\hat{\tau}$ is found by a direct search over a predetermined interval of possible group delay values. Hannan and Thomson (1973) show that their estimator is asymptotically normal, but no distributional properties exist for small samples. Furthermore, no useable guidelines are available as to the finite sample size necessary for the asymptotic theory to be applicable.

Deaton and Foutz (1980) developed the theoretical justification for the interpretation of group delay as a time delay parameter between two time series. Their work justifies the frequency based model given in [2]; i.e., they extend the concept of time delay to the case where phase and frequency

have a nonlinear relationship. Based upon the work of Deaton and Foutz (1980) we have a theoretical basis for the estimation of group delay as a measure of time lag of one series behind another at a specific frequency and this interpretation is used throughout the remainder of the dissertation.

Chapter II

Concepts In Spectral Analysis.

II.1. Introduction.

The majority of statistical theory and practice is concerned with the analysis of random samples of independent observations. Time series analysis is concerned with the analysis of samples where we assume the observations are taken with respect to a time axis (or a spatial dimension such as length). Furthermore, we assume dependencies exist between the observations. The difference between time series and non-time series analysis might be demonstrated by the fact that in regression analysis one is interested in estimating a trend between a dependent and independent variables, and it is assumed that the errors about the trend line are independent and random. In time series one would usually remove the trend by some means and actually attempt to model the

structure of the error about the trend, which is now considered to have some sort of dependent structure between observations.

Dichotomous approaches exist to the analysis of time series data. The first works directly with the observed sample in the time domain; e.g., the Box and Jenkins (1976) ARIMA models. The second approach requires that we transform our sample function from the time domain to frequency domain by use of a Fourier transform. Once in the frequency domain one attempts to analyze the behavior of the time series based upon the principles of spectral analysis.

Kendall (1976, p. 18) points out that the term spectral analysis is actually misleading in that most naturally occurring time series do not contain actual spectra; i.e., they do not exhibit true cyclic behavior that one would generate with a classical oscillator (e.g., signal generator). In general, time series contain "wave-like" patterns of regular variability over time (Priestly, 1981, p. 2). In this more typical case we have aperiodic patterns which we approximate by frequency. Of course if the patterns of variability were sufficiently irregular then we could not perform a meaningful analysis using spectral theory. In general aperiodic functions which can be approximated as periodic functions are called almost periodic functions (Koopmans, 1974, p. 21). In the present work we are concerned with time series which have almost periodic behavior, are 0-mean, and are weakly stationary (the mean and variance do not change with time).

II.2. The Autospectrum.

The autospectrum refers to the spectral analysis of a single univariate time series, say $X(t)$. The true spectral density of $X(t)$ at Fourier frequency (defined subsequently in this section) ω_o is denoted by $f_x(\omega_o)$ and the spectral density is a measure of the variability in the series $X(t)$ attributable to frequency ω_o . We observe $X(t)$ at discrete, equally spaced points in time, i.e., t takes on only integer values and we assume that $X(t)$ has a continuous spectrum (aperiodic functions are modelled as having continuous spectra although they may also contain true periodic components) irrespective of whether $X(t)$ is actually a discrete time or continuous time process. In naturally occurring time series we are only able to sample at finite intervals so that continuous time processes are only observed at finite periods of time. This has an important implication for the analysis of the time series in the frequency domain, since we are not able to observe the continuous spectrum. Intuitively, if one can only sample at some finite interval then one can not possibly detect effects which have a frequency higher than the sampling rate. In other words any effects with a frequency above $(2\pi/\Delta t)$ are aliased in with lower frequencies. Here, Δt is the sampling interval and the highest observable frequency is sometimes called the Nyquist or folding frequency. We assume in the present research that our time series have negligible aliasing from frequencies beyond the Nyquist frequency.

The transformation to the frequency domain is accomplished by the use of the discrete Fourier transform and we denote the transform at frequency ω_o by

$$Z_x(\omega_o) = \frac{1}{(2\pi N)^{\frac{1}{2}}} \sum_{t=1}^N x(t) e^{-i\omega_o t} \quad [1]$$

where N is the sample size. The observable frequencies are referred to as Fourier frequencies and they are calculated as

$$\omega_j = \frac{2\pi j}{N}, j = 0, \pm 1, \pm 2, \dots \quad [2]$$

For a given sample size there are $(N/2) + 1$ observable Fourier frequencies.

Once the data are transformed to the frequency domain, we calculate estimates of the spectral density at each Fourier frequency by means of the method of periodograms. The periodogram for the Fourier frequency ω_o is denoted by

$$I_x(\omega_o) = Z_x(\omega_o) \bar{Z}_x(\omega_o) \quad [3]$$

where $\bar{Z}_x(\omega_o)$ is the complex conjugate of $Z_x(\omega_o)$. It can be shown that the periodogram is actually related to the Fourier transform of the sample autocovariance function of the time series in the time domain; i.e.,

$$I_x(\omega_o) = \sum_{k=-(N-1)}^{(N-1)} c_k e^{-i\omega_o k} \quad [4]$$

where c_k is the sample autocovariance between observations which are k lags apart on the time axis. In other words the periodogram is a measure of the variability in the time series attributable to Fourier frequency ω_0 .

Although the periodogram may appear to be an ideal estimate of the true spectral density $f_x(\omega_0)$, it can be shown to be MSE inconsistent (see, for example, Brockwell and Davis, 1987, p. 333). The periodogram is asymptotically unbiased, but the variance asymptotically goes to a non-zero constant. Theorem 10.3.2 of Brockwell and Davis (1987, p. 337) shows also that the periodogram ordinates are asymptotically independent and exponentially distributed. In fact the periodogram ordinates are approximately independent even in small samples and some authors (see Bloomfield, 1976, p. 224 or Hannan and Thomson, 1988) simply consider them independent random variables. As a result of their independence and inconsistency, contiguous periodogram ordinates tend to fluctuate significantly in value even in very large samples. Therefore, the periodogram is a poor estimator of the spectral density.

A consistent estimator of the spectral density at a particular Fourier frequency can be achieved by applying smoothing algorithms to the sample periodograms. Many types of smoothers exist, however, one of the easiest is a technique due to Daniell (1946). The Daniell smoother or window is actually a moving average algorithm in which adjacent periodogram ordinates are arithmetically averaged to generate a spectral density estimate at a specific Fourier frequency. The band of m Fourier frequencies is selected such that it is symmetric about a center Fourier frequency where a spectral density estimate is desired. The Daniell spectral estimate is denoted by

$$\hat{f}(\omega_o) = \frac{1}{m} \sum_{j=1}^m I_x(\omega_j) \quad [5]$$

and asymptotically as $N \rightarrow \infty$, $m \rightarrow \infty$, $m/N \rightarrow 0$

$$E(\hat{f}(\omega_o)) = f(\omega_o). \quad [6]$$

Since the periodogram ordinates are asymptotically independent, we have that

$$\text{Var}(\hat{f}(\omega_o)) = \frac{1}{m^2} \sum_{j=1}^m \text{Var}(I_x(\omega_j)) \quad [7]$$

and as $m \rightarrow \infty$ the RHS of [7] goes to 0. As a result the Daniell spectral estimate is asymptotically MSE consistent. In the present research the Daniell window is used to generate spectral estimates whenever they are necessary to the work under consideration.

II.3. Bivariate Processes.

In this section we now consider the frequency domain analysis of the relationship between 2 univariate time series, say $X(t)$ and $Y(t)$. Fortunately, techniques used in the previous section, for a single univariate time series, generalize very easily to the bivariate case.

The first step to a frequency domain analysis of the 2 time series is to perform a separate Fourier transform on each of the 2 time series using [1]. The cross-periodogram ordinate is then generated by

$$I_{xy}(\omega_o) = Z_x(\omega_o)\bar{Z}_y(\omega_o). \quad [8]$$

The cross-periodogram is generally complex valued and is theoretically the Fourier transform of the cross-covariance function for the 2 time series. Therefore, $I_{xy}(\omega_o)$ is a measure of the covariation between the 2 time series at frequency ω_o .

The distributional properties of the cross-periodogram are given by Brockwell and Davis (1987, p. 431). Asymptotically the cross-periodogram ordinates are independently distributed with a complex Wishart distribution (Priestly, 1981, p. 696). As in the univariate case, the independence is assumed even in small samples (Bloomfield, 1976, p. 224 or Hannan and Thomson, 1988). The real part of $I_{xy}(\omega_o)$ is sometimes referred to as the cospectrum and the imaginary part as the quadrature spectrum; the terms carry over to the smoothed spectral estimates. The inconsistency of the univariate periodogram carries over to the cross-periodogram, i.e., both the real and imaginary parts of the cross-periodogram are unbiased estimates of their underlying spectral density counterparts. However, the variances of the real and the imaginary parts both go to a nonzero constant as the sample size becomes infinite. Once again the inconsistency and independence result in highly variable estimates of the true cross-spectral density (see, Brockwell and Davis, 1987, p. 429 or Bloomfield, 1976, p. 221).

Consistent estimates of the cross-spectral density can be developed by applying a smoothing algorithm to the real and the imaginary parts of the cross-periodogram (Chatfield, 1984, p. 183); i.e.,

$$\operatorname{Re}(\hat{f}_{xy}(\omega_o)) = \frac{1}{m} \sum_{j=1}^m \operatorname{Re}(I_{xy}(\omega_j)) \quad [9]$$

and

$$\operatorname{Im}(\hat{f}_{xy}(\omega_o)) = \frac{1}{m} \sum_{j=1}^m \operatorname{Im}(I_{xy}(\omega_j)). \quad [10]$$

It is shown by Brockwell and Davis (1987, p. 432) that the smoothed spectral density estimates become independent asymptotically. However, the independence is not generally assumed in finite samples. In fact, there does not appear to be a published study which identifies the necessary sample size to assume independence for the smoothed spectral estimates.

An important cross-spectral parameter is the coherence squared function

$$\rho_{xy}^2(\omega_o) = \frac{|\hat{f}_{xy}(\omega_o)|^2}{\hat{f}_x(\omega_o)\hat{f}_y(\omega_o)} \quad [11]$$

and it can be shown by the Cauchy-Schwarz inequality that $0 < \rho_{xy}^2(\omega_o) < 1$ (Bloomfield, 1976, p. 214). In other words the coherence between two time series is a measure of the correlation between them at a specific Fourier frequency.

The phase between the 2 time series at a specific Fourier frequency is denoted by

$$\phi_{xy}(\omega_o) = -\tan^{-1} \left\{ \frac{\text{Im}[f_{xy}(\omega_o)]}{\text{Re}[f_{xy}(\omega_o)]} \right\}. \quad [12]$$

The phase at Fourier frequency ω_o is a radian measure of how one time series is leading or lagging the other time series at that Fourier frequency. By looking at [12], it is noted that the phase is not a unique function of frequency because the arc tangent function only generates unique values up to added multiples of 2π . By a convention in spectral analysis the phase is plotted on a circle from $-\pi$ to π (Priestly, 1981, p. 660). As a result the calculated phase spectrum is discontinuous. Chatfield (1984, p. 177) notes that the actual phase between two stationary time series is unique and continuous with respect to frequency. Unfortunately, in naturally occurring time series we are not able to directly observe phase and it must be calculated by [12].

Sample estimates of coherence and phase are achieved by substituting the appropriate sample estimates into [11] and [12]. In the case of [12] one is able to use the cross-periodogram ordinates or the smoothed spectral estimates to generate sample phase values. If periodogram ordinates are inserted into [11], then one has a sample coherency of 1 at every frequency and therefore the smoothed spectral estimates must be used in [11]. When the periodogram ordinates are used to generate sample phase, this is sometimes referred to as the raw sample phase and if the smoothed spectral estimates are used this is sometimes referred to as the smoothed sample phase (Priestly, 1981, p. 695). The advantage of the raw phase is the fact that the values are generated by the approximately independent periodogram ordinates with the result that the raw sample phase values may also be considered approximately independent.

Brockwell and Davis (1987, p. 434) show that the asymptotic distribution of the smooth sample phase is

$$\hat{\phi}_{xy}(\omega_o) \sim \text{AN}\left\{\phi_{xy}(\omega_o), \frac{1}{a_n^2 \alpha_{xy}^2(\omega_o)} \left[\frac{(1 - \rho_{xy}^2(\omega_o))}{\rho_{xy}^2(\omega_o)} \right]\right\} \quad [13]$$

where a_n^2 is the sum of the weights squared that are used to generate the smooth spectral estimates, in the case of the Daniell window the sum would be $\Sigma \frac{1}{m^2}$ and asymptotically as the sample size goes to infinity the sum of the weights squared goes to 0. The quantity $\alpha_{xy}^2(\omega_o)$ is called the cross-amplitude and is the squared norm of the cross-spectral density. The result is important because it demonstrates that the sample phase is an asymptotically unbiased estimator of the true phase. The raw sample phase is also asymptotically unbiased, since the unbiasedness is based upon averaging unbiased spectral estimates together in the smoothing algorithm. Further, the smoothed sample phase is also asymptotically consistent. However, the raw sample phase would not in general be consistent because the variance of the estimator would not go to 0 as the sample size increases. Another important point is brought out in [13] and that is the fact that the variance of the sample phase is asymptotically proportional to the coherence between the time series at a specific Fourier frequency; i.e., as the coherence becomes small the variance of the sample phase becomes increasingly large and as the coherence goes to 1 the variance approaches 0. Although we cannot state that the sample phase variance is exactly proportional to the coherence in finite samples it clearly has some impact on the sample phase variance. The exact relationship of coherence and sample phase variance in finite samples is unknown.

Since one is only able to compute non-unique sample phase values using [13], one must perform further processing on the sample phase values if unique values are desired. The idea is to realign

the sample phase values calculated by [12] such that they represent a sample from an unknown family of continuous phase distributions. The process of realigning the sample phase values is sometimes referred to as unwrapping the phase. Very little literature exists on techniques for performing the unwrapping and it appears to be a neglected topic by statisticians. Complicated algorithms for engineering applications can be found in Tribolet (1977) or Bonzanigo (1978), but these algorithms assume the underlying phase function has a linear relationship with frequency, which is not an assumption of the present work. Therefore, their algorithms are not directly applicable to the present research.

At present only one general unwrapping technique is available as an accepted practice. The method is discussed and applied by Fuller (1976, p. 320). The method is based upon the assumption that there exists an underlying phase function which is continuous with respect to frequency. Application of the method consists of adding or subtracting multiples of 2π to each sample phase value calculated by [12] such that the absolute distance between the present sample phase value and the previously chosen phase value is minimized; i.e.,

$$\tilde{\phi}_{xy}(\omega_o) = \hat{\phi}_{xy}(\omega_o) \pm j\pi; j = 0 \pm 2, \pm 4, \pm \dots \quad [14]$$

Although the method is somewhat ad hoc, no other method is currently available and it is a topic for future research in spectral analysis.

One other important cross-spectral parameter is the group delay function. The group delay function is a measure of how one of the time series leads or lags the other time series in units of time. Group delay requires the assumption that the underlying phase function is differentiable with

respect to frequency, which is not a restrictive assumption for stationary time series with continuous spectra (Hannan and Thomson, 1973) and we denote group delay by

$$GD_{xy}(\omega_o) = \frac{d\phi_{xy}(\omega_o)}{d\omega_o}. \quad [15]$$

The present work is an attempt to develop an estimation technique for group delay.

II.4. The Linear Filter.

Linear filter theory is a major topic in the theory of time series analysis. The present discussion is a brief overview of the subject since a complete discussion would require a considerable amount of time and space. In fact the best known linear filters are the Box and Jenkins (1976) ARIMA models. In the present discussion we confine ourselves to time invariant linear filters, which simply means that the parameters of the filter do not change with respect to the time axis. A linear filter is a type of model which allows us to model one time series as a function of another time series. By convention we call the time series to be modelled as a function of another, the output series, say $Y(t)$. The other time series is then referred to as the input series, say $X(t)$ and the linear filter is often denoted by

$$Y(t) = L(X(t)) \quad [16]$$

where L is the linear filter.

An important property of the linear filter, in fact why it is referred to as a linear filter, is the fact that the frequency content of the input series is linearly passed to the output series (the frequency content is conserved by the transformation of the filter), while the phase and amplitude of the frequency components are modified by the filter as they are passed from the input to the output series. This fact can be verified by reviewing Theorem 9.1 of Chatfield (1984, p. 194). The theoretical properties of linear filters are discussed by Koopmans (1974, Chapter 4).

We may write the general form of the filters for either a process continuous in time or discrete in time by

$$Y(t) = \int_{-\infty}^{\infty} h(u)X(t-u)du \quad [17]$$

and the discrete time process can be modelled by the convolution filter

$$Y_t = \sum_{k=-\infty}^{\infty} h_k X_{t-k}. \quad [18]$$

The linear filter describes the output at time t as a weighted sum of the inputs over past time (theoretically it may be extended to future time but such filters are not physically realizable and are not discussed here). The weighting function $h(u)$ or h_k is often referred to as the impulse response function of the filter, since it measures the effect of the input on the output over time. By taking the Fourier transform of the impulse response function one generates what is called the frequency response function (Koopmans, 1974, and some other authors refer to it as the transfer function of the filter). Just as the impulse response function measures the effect of an input in the time domain,

the frequency response function measures the effect of an input in the frequency domain. Again, the frequency response function can only change the phase or the magnitude of an input it cannot change the frequency content in a linear filter. The frequency response function is denoted by

$$H(\omega_o) = \int_{-\infty}^{\infty} h(u)e^{-i\omega_o u} du \quad [19]$$

for continuous time and the discrete time is analogous. In general $H(\omega_o)$ is complex valued and if we take the norm of the frequency response function we generate what is known as the gain function of the filter; i.e.,

$$G(\omega_o) = |H(\omega_o)|. \quad [20]$$

$G(\omega_o)$ is actually a regression coefficient measuring the amount of effect in the input at frequency ω_o which is passed to the output at that specific frequency.

When a linear filter has high values for $G(\omega_o)$ at low frequencies and small values at high frequencies, it is referred to as a low pass filter. In such a low pass filter the effects of the low frequency components are amplified while the effects of the high frequency components are attenuated. For the case of a high pass filter, the low frequency components are attenuated and the high frequency components are amplified. These concepts are applied in the simulation study of Chapter V.

Chapter III

The Box-Cox Power Transformation.

III.1. Introduction.

The purpose of this chapter is to give a brief overview of the Box-Cox power transformation and some of the issues associated with its use as an analytic tool for statisticians. The discussion is not intended to be a complete review of all aspects and literature on the subject, since they are rather prolific and broad in scope. Rather, the present discussion is simply intended as an aid to those readers not familiar with the Box-Cox transformation.

The method of transforming variables has long been a useful tool that allows statisticians to develop parsimonious representations and interpretations of observed data. In many cases the

transformation schemes may be applied to the independent variable(s) in a linear model (e.g., Box and Tidwell, 1962) or perhaps to both the independent variable(s) and dependent variable(s) simultaneously (Snee, 1986). But, in the present context we are only concerned with a single transformation which is applied only to the dependent variable in a linear model.

Although a wide range of transformation methods and objectives exist for transforming the dependent variable (e.g., stabilizing variance), we are restricting the present discussion to the Box-Cox (1964) power transformation. The three objectives of the Box and Cox (1964) approach to variable transformation is to achieve

1. A simplified structure for Ey for a given set of independent variables,
2. A constant variance for the error term, i.e., $\varepsilon \sim (0; I\sigma^2)$,
3. Normality of the distributions for the linear model.

The Box-Cox power transformation is a nonlinear transformation of the response vector \underline{Y} to a new response vector \underline{g} such that the three objectives stated above are achieved simultaneously. The Box-Cox transformation is well known for its ability to achieve normality in the transformed variables and the stabilizing of the variance (Box and Cox, 1964, 1982; Andrews, 1971; Carroll and Ruppert, 1981; Hinkley and Runger, 1984; Snee, 1986; and so on). The actual form of the transformation is

$$\underline{g} = \begin{cases} (\underline{y}^\eta - 1)/\eta & \eta \neq 0 \\ \ln(\underline{y}) & \eta = 0. \end{cases} \quad [1]$$

and we note from [1] that all elements of \underline{y} must necessarily be positive. After a suitable selection of the parameter η we assume that

$$\underline{g} = X\underline{\beta} + \varepsilon \quad [2]$$

where we have that

$$\underline{g} \sim N(X\underline{\beta}; I\sigma^2), \quad \varepsilon \sim N(0; I\sigma^2), \quad [3]$$

and

$$\hat{E}\underline{g} = \hat{\underline{g}} = X\hat{\underline{\beta}}. \quad [4]$$

We further assume in the present discussion that η is selected such that the model in [2] is a simple linear one, which is a common objective of the Box-Cox transformation method.

It is quite reasonable to assume that one would want to use the Box-Cox transformation when the observed data contains negative values. Box and Cox (1964) or Draper and Smith (1981, p. 236) modify [1] to deal with the situation of negative values. The approach is simply to add a shift constant to all of the observed response values such that all of them are made greater than 0. This form of the Box-Cox power transformation is commonly referred to as the 2 parameter transformation and is denoted by (for the i^{th} observation)

$$g_i = \begin{cases} ((y_i + \eta_2)^{\eta_1} - 1)/\eta_1 & \eta_1 \neq 0 \\ \ln(y_i + \eta_2) & \eta_1 = 0. \end{cases} \quad [5]$$

After a suitable selection of the two parameters η_1 and η_2 , we then make all of the assumptions of [2], [3], and [4] for g . The two parameter transformation is discussed further in section V.3. The remainder of the discussion in Chapter III is for the single parameter case, since the results generalize naturally to the two parameter case.

III.2. Selection Of The Power Parameter.

It is unlikely that one would know a priori what value η should have for the transformation in [1], such that the three objectives of the Box-Cox transformation are achieved. There are two basic approaches to selecting a value for the power parameter η . The method used most often in practice to obtain an estimate of η is a maximum likelihood approach with the assumptions of normality of error terms invoked; i.e., the likelihood equation is a normal density and the method is discussed in detail by Draper and Smith (1981, p. 235) or Box and Cox (1964). The second approach is Bayesian equivalent of the MLE method and is discussed in detail by Box and Cox (1964). An approximate MLE method is given by Hinz and Eagles (1976) which is computationally efficient for large samples.

The maximum likelihood method of estimating η is the method preferred by Draper and Smith (1981, p. 235) and the present discussion is limited to the MLE method. The approach is to select a range of values for η , say $(-3,3)$, and then perform a direct search at some predetermined interval across the range of possible values. For each selected value of η , in the range of possible values, we evaluate an MLE objective function

$$L_{\max}(\eta) = -\frac{n}{2} \ln \hat{\sigma}^2 + (\eta - 1) \sum_{i=1}^n \ln y_i \quad [6]$$

where

$$\hat{\sigma}^2 = \underline{\mathbf{g}}'(I - X(X'X)^{-1}X')\underline{\mathbf{g}}/n \quad [7]$$

which is just the MSE after fitting a linear regression model to the transformed data $\underline{\mathbf{g}}$. The second term, on the RHS of [6], is a Jacobian to account for the change in scale with each unique value of η . We select as our estimate of η that value which maximizes [6], which is equivalent to minimizing the MSE for the specified model in [2] (see Draper and Smith, 1981, p. 225, or Box and Cox, 1964).

After evaluating [6] for a number of values of η in the specified interval of possible values, Draper and Smith (1981, p. 226) recommend a graphical method to select the optimum choice of η . The idea is to plot the values of [6] against the values of η and then draw a smooth curve through the plotted points. The value of η associated with the high point on the curve is selected as the optimum value for η . In actual practice, one usually selects a convenient value for η in the neighborhood of the stationary point, rather than the precise value of η at that location. As an example, it would be computationally simpler to work with a value of $\eta = 0.5$ rather than, say, $\eta = 0.599$; however, Draper and Smith (1981, p. 226) point out that this is mostly matter of personal preference.

III.2.1. Approximate Confidence Intervals For The Parameter.

It is possible to calculate approximate $100(1 - \alpha)\%$ confidence intervals for η . Draper and Smith (1981, p. 227) or Box and Cox (1964) base the confidence interval method on the Chi-Squared distribution; i.e.,

$$L_{\max}(\hat{\eta}) - L_{\max}(\eta) < \frac{1}{2} \chi_{1,(1-\alpha)}^2. \quad [8]$$

Although the method is not motivated by Draper and Smith (1981, p. 227) or Box and Cox (1964), it can be seen that the LHS of [8] is the log of a likelihood ratio and the inequality is achieved by the large sample distributional assumption for the log likelihood ratio statistic and the 1 degree of freedom comes from the fact that η is the only parameter assigned a value under H_0 (see, for example, Theorem 10.2 of Mendenhall, et. al., 1986, p. 431). In practice the confidence interval is usually approximated by computing the value $L_{\max}(\hat{\eta}) - \frac{1}{2} \chi_{1,(1-\alpha)}^2$. A horizontal line is then drawn on the vertical axis of the plot of $L_{\max}(\eta)$ for the computed value; the values of η where the horizontal line intersects the curve are the endpoints of the confidence interval. Atkinson (1973) performed a simulation study of the confidence interval procedure under the assumption of normal error and found that the coverage probability of the calculated confidence intervals were in close agreement with their theoretical values.

III.3. Variance Inflation In The Box-Cox Transformation.

There appears to be no agreement in the literature as to whether one should consider η a fixed parameter, once its value has been fixed, or whether it should be treated as a random variable which has been estimated. The question is not a trivial one, since the treatment of $\hat{\eta}$ as a random variable has deleterious implications for the use of the Box-Cox transformation (Bickel and Doksum, 1981, or Spitzer, 1984). Box and Cox (1964, 1982) consider $\hat{\eta}$ a fixed parameter once its value is determined and any post-transformation analysis of the data is performed subject to the fixed value of $\hat{\eta}$ which is equivalent to fixing a scale of measurement. Hinkley and Runger (1984) strongly support the concept of post-transformation analysis subject to the fixed value of $\hat{\eta}$, while Bickel and Doksum (1981) present an equally strong position that $\hat{\eta}$ must be treated as a random variable.

Bickel and Doksum (1981) present a mathematically impressive analysis of the Box-Cox transformation under the assumption that post-transformation inference must treat $\hat{\eta}$ as stochastic. Their argument is that the parameter estimates in [4] must necessarily be dependent upon the choice of $\hat{\eta}$ since the model coefficients are scale dependent. As a result of this dependency the $\text{Var}(\hat{\beta}) = (X'X)^{-1}\hat{\sigma}^2$ must be biased downward, since we need to add an additional variance term for the estimate of $\hat{\eta}$. In their paper, Bickel and Doksum (1981) work out the asymptotic distribution theory for the parameter estimates under the assumption that $\hat{\eta}$ is stochastic and they conclude that the variance inflation that occurs can be of such a magnitude that prediction with the transformation model is worthless (they also consider bias of the estimated coefficients under the assumption the incorrect scale has been selected, however this result is not accepted as meaningful in a scientific

setting, Hinkley and Runger, 1984). Spitzer (1982, 1984) makes a similar, but less theoretical, conclusion to Bickel and Doksum (1981).

The arguments of Bickel and Doksum (1981) are strongly rebutted by Box and Cox (1982) and Hinkley and Runger (1984). Box and Cox (1982) view the work of Bickel and Doksum (1981) as, and we quote, "It seems to us that this general conclusion (Bickel and Doksum, 1981) is qualitatively obvious and scientifically irrelevant." Hinkley and Runger (1984) make a nearly identical statement in their assessment of the work of Bickel and Doksum (1981). The key point of Box and Cox (1982) or Hinkley and Runger (1984) is that the parameters of a linear model only have meaning with respect to a specific scale of measurement and the selection of a value for η is simply selecting a scale of measurement; and we should pay no penalty for having used the data to select that scale. Furthermore, Hinkley and Runger (1984) point out that the work of Bickel and Doksum (1981) assumes that there are some unknown but correct values for the parameters in [2] and one must select the correct transformation to achieve those values; however, Hinkley and Runger (1984) note that the comparison of coefficients measured on two different scales makes no scientific sense.

A way to describe the controversy is to consider the situation where two scientists independently measure the lengths of some objects and one uses a scale in millimeters and the other inches. Having taken a sufficient number of measurements each calculates a mean length; however, a contrast of the two means has no direct interpretation, since they are measured on separate scales. This is one of the criticisms Hinkley and Runger (1984) raise with the work of Bickel and Doksum (1981). Box and Cox (1982) further make the point that one could solve the problem of Bickel and Doksum (1981) by resorting to parameter estimates which are scale free; however, Box and Cox

(1982) raise the question of scientific relevance in that one could talk about the standard deviation of heights for human adults as being 1000, but without a scale of reference what scientific interpretation does the number have?

A compromise position to the controversy surrounding the interpretation of $\hat{\eta}$ is taken by Carroll and Ruppert (1981), where they performed a simulation study of the effects of considering $\hat{\eta}$ fixed for the purposes of post-transformation analysis. In their simulation study they did observe some inflation in prediction variance due to the estimation of η . However, they found the amount of inflation to be small and not of consequence in application. Since the present research only uses the Box-Cox transformation as a tool and in light of the work of Carroll and Ruppert (1981), the position of Box and Cox (1964, 1982) and Hinkley and Runger (1984) is adopted and the selected value of $\hat{\eta}$ is considered fixed once it is determined. The controversy surrounding the interpretation of $\hat{\eta}$ is provided as information for the reader.

Chapter IV

UMVU Estimation of Phase and Group Delay.

IV.1. Introduction.

A classic problem encountered in spectral analysis is that distributional properties of estimators are limited to asymptotic results. Kay (1988, p. 8) points out that exact distributional properties are not known explicitly for finite data records. Furthermore, the finite sample size required for the asymptotic result to apply is not known for any particular problem. As a result Kay (1988) warns against the use of asymptotic distribution theory to make inferences about spectral parameters, based upon a short record of observation.

Generally, the properties of estimators in finite sample sizes are investigated by simulation studies. Unfortunately, such studies provide useful guidelines for the behavior of small sample spectral estimators, but no global properties may be stated. It would be particularly unique if spectral estimators existed for which it is possible to state such global optimality criteria as BLUE or UMVUE. The work of this chapter is an attempt to construct estimators of the phase function and group delay function which have UMVUE properties. The development of estimators with global properties based upon finite samples is a significant advancement over the existing methods, which are highly asymptotic. A case in point is the Hannan and Thomson (1973) group delay estimator. The estimator for group delay in this chapter is not a direct competitor with the Hannan and Thomson estimator since their estimator is restricted to specific points on the group delay function, while the method of this chapter attempts to construct an estimate of the entire function.

The fundamental approach of this chapter is to perform a suitable transformation on the observed spectral function; the phase function in our particular case. The choice of transformation is such that the transformed values may be assumed to have a normal distribution with constant variance. Once transformed to the new scale of measurement, the new set of observations are modeled using standard linear models methodology under the normality assumption. In our particular case we will model the relationship between the phase function and frequency by performing regression analysis in the new metric. Of course with the assumption of normality, the OLS estimates are UMVUE. A nice feature of this approach is that no explicit distributional assumptions are made about the observations in the original scale of measurement.

A theoretical difficulty with this approach is the fact that UMVUE properties in the transformed metric are not directly useful in spectral analysis. This is specifically true for estimation of the group

delay function since this is the derivative of the phase function. Therefore, estimates of phase and the derivative of phase in the transformed metric are not directly applicable to the functions in the original scale of measurement. The principle focus of this chapter is to find a way to perform an inverse transformation of our UMVUE's without destroying the UMVUE properties.

Although it is not necessary to make explicit distributional assumptions about the phase function in the original scale of measurement, it is necessary to be concerned about two properties of the phase function. First, the phase function exhibits skewed distributional behavior in finite samples and, secondly, it tends to exhibit heterogeneity in variance. There appears to be no published literature thoroughly examining the phase function in finite samples. In fact, Chatworth (1984, p. 184) points out that the properties of cross-spectral estimates, including the phase, have not been adequately studied and much work needs to be done by statisticians. In light of the absence of published results for the phase function, a small scale simulation was performed to verify the behavior of the phase function in finite samples. The simulation study suggested that the phase function exhibits marked skewness in distribution, even in sample sizes as large as 7,000. It was also noted that the amount of skewness varies from frequency to frequency across the function. The variance also showed a tendency to vary across the frequency interval $(0, \pi]$. This simply confirmed the usual assumption that the phase function is skewed, in distribution, and has heterogeneous variance. The simulated behavior is particularly relevant because the raw sample phase values generated from the periodograms usually exhibit some level of skewness. In fact the exact distribution of the sample phase (raw or smoothed) in large samples is unwieldy (Jenkins and Watts, 1968, p. 380) and Hannan (1970, p. 258) states that it is not of great interest in general. The approximate smoothed sample phase distribution is given in equation [13] of Chapter II.

It is clear that if the current method is to succeed we must use a transformation method which is effective in stabilizing variance and changing skewed distributions to approximately normal ones. Since the Box-Cox power transformation was shown to be effective in these situations, in chapter III, we assume the use of that transformation method in the current work. However, the method to be discussed is directly applicable to a wide range of transformation schemes.

IV.2. The Problem Of Inverse Transformation Bias.

Although the transformed data are analyzed by a straightforward application of linear models methodology, a problem arises when it is desirable to perform an inverse transformation of the estimates to the original scale of measurement. In the present case we are interested in generating predicted values of phase and group delay in the original scale, but based upon the UMVUE values in the transformed scale. It is assumed that the transformation scheme is monotonic and more will be assumed in section IV.3.

Under monotonic transformations, Hald (1952) showed that the median of the distribution is invariant under transformation; i.e., the 50th percentile of the transformed values inversely transforms to the 50th percentile on the original scale. Intuitively, the mean of the transformed values does not, in general, inversely transform to the mean on the original scale. Since we assume the transformed data to constitute a sample from a normal distribution, it follows that the population median and mean have the same value. Therefore, an UMVU estimate of the mean also provides an UMVU estimate of the median (the reverse is not necessarily true); i.e., under normality, the

predicted values from a regression equation are estimates of the mean and the median at that particular setting of the regressors.

When one performs an inverse transformation on the UMVUE's (e.g., predicted values from a regression model) generated in the new scale of measurement, the UMVUE property for the mean estimate does not carry over to the original scale, in general. Once the, say predicted values, are transformed back to the original scale of measurement, the mean and the median of the underlying distribution are no longer considered coincident. As a result, the inversely transformed predicted values still constitute unbiased estimates of the median, but in general are biased estimates of the mean. The result, of course, is the loss of an UMVU estimate of the mean. Miller (1984) shows that the resulting inverse transformation bias for estimates of the mean may be quite severe and, in the case of regression, the inverse values constitute a median regression analysis. The remainder of this chapter is an attempt to develop an inverse transformation scheme which preserves the UMVUE properties of estimates of expected function values in the transformed metric.

IV.3. Correcting For Bias In the Inverse Transformation.

At this point it is convenient to set the notation which is used in the remainder of the chapter. It should also be noted that appendix A to this chapter contains various theoretical results from advanced calculus and measure theory which, are required at various points in the discussion contained in the remaining sections of this chapter.

Let the $n \times 1$ vector of observations in the original scale of measurement be denoted by \underline{Y} . and further the observations may be a function of an $n \times p$ matrix of observations for p non-stochastic variables and we denote this matrix by Ω . We define the the following:

$$E(y_i) = \theta_i, \quad \text{Var}(y_i) = \phi_i^2$$

or

$$\theta(\omega_i) = E(y(\omega_i)), \quad \text{Var}(y(\omega_i)) = \phi^2(\omega_i)$$

Next, let g_i be the i^{th} observed value after a suitable transformation by a monotonic function f^{-1} ; i.e.

$$g_i = f^{-1}(y_i)$$

or

$$g(\omega_i) = f^{-1}(y(\omega_i)).$$

The transformation of the independent variables is not considered in the present work, but the results naturally generalize to this case. After the transformation, it is assumed that

$$g(\omega_i) \sim N(\mu(\omega_i); \sigma^2) \quad [1]$$

and

$$\hat{E}g(\omega_i) \sim N(\mu(\omega_i); \lambda^2 \sigma^2) \quad [2]$$

where λ^2 is a scalar constant dependent upon specific values of the independent variables; a typical example is the 'hat' diagonal h_{ii} in regression analysis.

In the present research we are particularly concerned with modelling the relationship between the transformed random variables and the independent variable(s) by the use of OLS regression analysis. Therefore we further use the notation

$$\underline{\mathbf{g}} = \underline{\Omega}\underline{\beta} + \underline{\varepsilon}, \quad \underline{\varepsilon} \sim N(0; I\sigma^2) \quad [3]$$

and

$$\hat{\underline{\mathbf{g}}} = \hat{\underline{\mathbf{E}}}\underline{\mathbf{g}} = \underline{\Omega}\hat{\underline{\beta}}, \quad \underline{\mathbf{g}} \sim N(\underline{\Omega}\underline{\beta}; \underline{\Omega}(\underline{\Omega}'\underline{\Omega})^{-1}\underline{\Omega}'\sigma^2). \quad [4]$$

also

$$\hat{\underline{\beta}} \sim N(\underline{\beta}; (\underline{\Omega}'\underline{\Omega})^{-1}\sigma^2) \quad [5]$$

and

$$\hat{\sigma}^2 = \frac{\text{SSE}}{\nu}, \quad [6]$$

where ν are the degrees of freedom and SSE is the error sum of squares. It is further noted that the estimates generated by OLS, under normality, constitute complete, sufficient, statistics and are UMVUE by the result of Lehmann and Scheffe' (1950).

The key publication on correcting inverse transformation bias appears to be Neyman and Scott (1960) which developes a general method for calculating UMVUE values in the original scale of measurement. The UMVUE's are based upon complete, sufficient statistics from the transformed domain and rely upon the Lehmann and Scheffe' (1950) result to provide UMVUE properties. The Neyman and Scott (1960) estimator is referred to as the N-S estimator in the remainder of the paper in order to simplify notation. The N-S estimator generates an UMVUE for $\theta(\omega_i)$ by an inverse transformation of an UMVUE estimate $\hat{\mathbf{g}}(\omega_i)$ in the new metric. Hoyle (1968) extended the

N-S procedure to the estimation of $\phi^2(\omega_i)$ from the UMVUE statistics in the transformed domain. Schmetterer (1960) developed a general solution to the inverse transformation problem in terms of a solution to an integral equation. His approach appears intractable in actual application because the integral equation is difficult to solve in general. Granger and Newbold (1975) develop a solution based upon complex generating functions which also does not appear tractable in most problems. Miller (1984) has proposed a number of bias adjustment factors which are easy to apply. However, the adjusted estimates are still biased and therefore not UMVUE.

The present work is an extension of the N-S estimator to the generation of UMVUE values for expected sample phase and expected sample group delay at each of the n Fourier frequencies. We let ω_i be the i^{th} Fourier frequency and $\theta(\omega_i)$ be the i^{th} expected sample phase value. The i^{th} expected sample group delay value is denoted by $\theta'(\omega_i)$. The basic approach is to observe n sample phase values $y(\omega_i)$, $i = 1, \dots, n$. After unwrapping (see section II.3) the observed sample phase values, the observations are transformed by the Box-Cox power transformation such that the transformed observations $g(\omega_i)$ are assumed normally distributed and of constant variance. After transformation the relationship between the phase and frequency is modelled by OLS regression analysis. After estimating the coefficients of the model and generating predicted values, these predicted values are then inversely transformed to UMVUE values for the expected sample phase in the original scale of measurement. The N-S estimator is used to preserve the UMVUE property and is further extended to develop UMVUE values for the derivative of the expected sample phase or the expected sample group delay.

The N-S method assumes that the inverse transformation function f is at least second order entire; i.e.,

$$y(\omega_i) = f(g(\omega_i)) \quad [7]$$

where f is the second order entire inverse transformation function. The method further assumes that $\theta(\omega_i) = E y(\omega_i) < \infty \forall \omega \in (0, \pi]$. In the present case it is also assumed that $\theta'(\omega_i) < \infty \forall \omega \in [0, \pi]$. The second order requirement simply refers to the rate of growth of a function as its argument increases. It should be noted that every indefinitely differentiable function, at a particular point, whose derivatives are all bounded at that particular point is a second order entire function (Boaz, 1954). Polynomials and exponentials are entire functions, although they need not be second order. An example of an entire function which is not second order is $e^{-\omega^2}$. Boaz (1954) is an often cited reference for a good discussion of entire functions. Fortunately, many of the functions f that are of interest satisfy the second order entire requirement.

It is formally shown in section IV.4 that a function say $f(\omega)$ is second order entire if and only if the radii of absolute convergence of the two series

$$\sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(2m)} \omega^m \quad \text{and} \quad \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(2m+1)} \omega^m \quad [8]$$

are infinite and $f_o^{(m)}$ represents the m^{th} derivative of f evaluated at a . Neyman and Scott (1960) point out that this a much stronger requirement than the usual requirement for a Taylor series expansion of f to be absolutely convergent. Neyman and Scott (1960) then go on to show that if f is second order entire then

$$\theta(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} E(g(\omega_i))^m \quad [9]$$

and the proposed N-S estimator for expected sample phase is

$$\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} T_m(\omega_i) \quad [10]$$

where $T_m(\omega_i)$ is selected such that

$$ET_m(\omega_i) = Eg^m(\omega_i). \quad [11]$$

The key to the development is to notice that $Eg^m(\omega_i)$ is the m^{th} noncentered moment of a normal random variable. UMVUE's are given by Neyman and Scott (1960) or Hoyle (1968) as

$$T_{2m}(\omega_i) = \sum_{k=0}^m \frac{(2m)!}{(2k)!(m-k)!} [\hat{g}(\omega_i)]^{2k} \left[\frac{1}{4} \text{SSE}(1 - h_{ii}) \right]^{m-k} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)}, \quad [12]$$

$$T_{2m+1}(\omega_i) = \sum_{k=0}^m \frac{(2m+1)!}{(2k+1)!(m-k)!} [\hat{g}(\omega_i)]^{2k+1} \left[\frac{1}{4} \text{SSE}(1 - h_{ii}) \right]^{m-k} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)}. \quad [13]$$

The fact that $\hat{\theta}(\omega_i)$ is UMVUE follows from the sufficiency of $\hat{g}(\omega_i)$ and $\hat{\sigma}^2$ (calculated from the SSE by dividing by the degrees of freedom) for $Eg(\omega_i)$ and σ^2 respectively. Then, using the result of Lehmann and Scheffe' (1950) that any function of complete, sufficient statistics is the UMVUE of its expectation, we have the UMVUE property of $\hat{\theta}(\omega_i)$. Furthermore, since $\hat{g}(\omega_i)$ and $\hat{\sigma}^2$ are complete, sufficient statistics under normality, we have that $\hat{\theta}(\omega_i)$ is unique.

IV.4. Modeling Phase and Group Delay By Transformations.

In this section we specialize the N-S estimator for the expected sample phase function and the expected sample group delay function. We assume an initial set of unwrapped sample phase values and we further restrict our attention to raw sample phase values calculated from the periodograms for the observed data. As a result we consider our set of phase values to be approximately independently distributed for each of n Fourier frequencies. As in section IV.3. we assume that the Box-Cox transformation has been successful in generating normally distributed random variables with constant variance.

We begin by standardizing notation specifically required in this section. This is in addition to the notation given in section IV.3. Let

$\Omega_{n \times p}$ be our model matrix in original units and $\Omega'_i = (1, \omega_i, \omega_i^2, \dots, \omega_i^p)$ be the i^{th} row of Ω

$$g(\omega_i) = \frac{(y(\omega_i) + \eta_2)^{\eta_1} - 1}{\eta_1}$$

and

$$\underline{g}(\omega) = \Omega \underline{\beta} + \underline{\varepsilon}, \quad \underline{\varepsilon} \sim N(0; I\sigma^2)$$

We make the following distributional assumptions

$$\underline{g}(\omega) \sim N(\Omega \underline{\beta}; I\sigma^2) \quad [13]$$

$$\hat{\underline{\beta}} \sim N(\underline{\beta}; (\Omega' \Omega)^{-1} \sigma^2) \quad [14]$$

$$\hat{\underline{g}}(\omega) = \hat{\underline{g}}(\omega) \sim N(\Omega \hat{\underline{\beta}}; \Omega (\Omega' \Omega)^{-1} \Omega' \sigma^2) \quad [15]$$

We also need notation for the derivative of $\hat{\underline{g}}(\omega_i)$ with respect to ω . Let

$$\frac{d}{d\omega} \hat{\underline{g}}(\omega_i) = \hat{\underline{g}}'(\omega_i) = \hat{\beta}_1 + 2\hat{\beta}_2 \omega_i + \dots + p\hat{\beta}_p \omega_i^{p-1} \quad [16]$$

We denote the $n \times (p-1)$ model matrix for the differentiated regressors by Ω_d and we assume

$$\hat{\underline{g}}'(\omega) \sim N(\Omega_d \hat{\underline{\beta}}_d; \Omega_d (\Omega_d' \Omega_d)^{-1} \Omega_d' \sigma^2) \quad [17]$$

The i^{th} diagonal of Ω , h_{ii} is the i^{th} hat diagonal or leverage measure for the that observation.

IV.4.1. The UMVUE of The Phase Function.

We now develop an UMVUE of expectation of the sample phase function based upon the N-S estimator. The theoretical justification for the method is also provided in this section. The

only explicit assumptions made about the phase function is that it is differentiable over the interval $(0, \pi]$. Of course, if the underlying phase function were not differentiable, then it would not even make sense to discuss the estimation of the group delay function. Also we reiterate, from the previous section, the requirements of the N-S method

1. $\vartheta(\omega) < \infty \forall \omega \in (0, \pi]$,
2. $f(g(\omega_i)) - \eta_2 = y(\omega_i)$ is entire,
3. The Taylor series expansion of f is absolutely convergent with infinite radius $\forall \omega \in (0, \pi]$

Clearly, assumption 1 is necessary if one is even to discuss the estimation of the expectation of the sample phase at each Fourier frequency.

Since the inverse transformation function f is assumed entire at each Fourier frequency, it has a MacLaurin series representation; i.e.,

$$y(\omega_i) = f(g(\omega_i)) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} g^m(\omega_i) - \eta_2. \quad [18]$$

Further, because of assumption 3 we may take term by term expectation of this infinite series. This result is subsequently justified by formal proof. Therefore,

$$E y(\omega_i) = E f(g(\omega_i)) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} E g^m(\omega_i) - \eta_2. \quad [19]$$

Under the normality assumption for the $g(\omega_i)$, $E g(\omega_i)$ is the m^{th} noncentered moment of a normal random variable. In order to formally develop our phase UMVUE we use the UMVUE estimators for the noncentered moments; i.e.,

$$T_m(\omega_i) = \sum_{k=0}^m C_{m,k} [\hat{g}(\omega_i)]^k [\text{SSE}]^{m-k} \quad [20]$$

where $C_{m,k}$ is a nonrandom coefficient, $\hat{g}(\omega_i)$ is the i^{th} predicted value from [15]. Under the normality assumption $\hat{g}(\omega_i)$ and $\hat{\sigma}^2$ are complete, sufficient statistics (see, for example, Graybill, 1976, Theorem 10.2.3). They are also UMVUE by the Lehmann and Scheffe' (1950) result. Since,

$$E[T_m(\omega_i)] = E g^m(\omega_i) \quad [21]$$

our moment estimator is an unbiased estimator and a function of complete, sufficient statistics, it is UMVUE and unique for each of the m noncentered moments.

If we substitute [20] into $E g^m(\omega_i)$, then we have our expected sample phase estimator

$$\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} T_m(\omega_i) - \eta_2 \quad [22]$$

where η_2 is the Box-Cox shift parameter required to make all the observed sample phase values positive. It follows that if we allow the interchange of summation and expectation in the infinite series expansion (which we justify subsequently), then

$$E\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} E T_m(\omega_i) - \eta_2 \quad [23]$$

which is by the unbiasedness of $T_m(\omega_i)$

$$E\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} E g^m(\omega_i) - \eta_2 = \theta(\omega_i). \quad [24]$$

Therefore, if the interchange is permissible, then $\hat{\theta}(\omega_i)$ is an unbiased estimator of $\theta(\omega_i)$ and is UMVUE by Lehmann and Scheffe' (1950). With completeness it is also unique.

In order to justify the interchange of expectation and summation, Neyman and Scott (1960) give the formulas for the noncentered normal moments as follows;

$$E[SSE]^m = (2\sigma^2)^m \frac{\Gamma(\frac{1}{2} \nu + m)}{\Gamma(\frac{1}{2} \nu)}. \quad [25]$$

For any even normal noncentered moment we have

$$Eg^{2m}(\omega_i) = \sum_{k=0}^m \frac{(2m)!}{(2k)!(m-k)!} [(Eg(\omega_i))^2]^k \left(\frac{\sigma^2}{2}\right)^{m-k} \quad [26]$$

The formula for any odd noncentered normal moment is then

$$\text{Eg}^{2m+1}(\omega_i) = \text{Eg}(\omega_i) \sum_{k=0}^m \frac{(2m+1)!}{(2k+1)!(m-k)!} [(\text{Eg}(\omega_i))^2]^k \left[\frac{\sigma^2}{2}\right]^{m-k} \quad [27]$$

Absolute bounds on these moments are given by Neyman and Scott (1960) and a more theoretical development of them is given by Hoyle (1968). The absolute bounds are

$$\frac{(2m)!}{m!} \left[\frac{\sigma^2}{2}\right]^m \leq \text{Eg}^{2m}(\omega_i) \leq \frac{(2m)!}{m!} [(\text{Eg}(\omega_i))^2 + \frac{\sigma^2}{2}]^m \quad [28]$$

and

$$|\text{Eg}(\omega_i)| \frac{(2m+1)!}{m!} \left(\frac{\sigma^2}{2}\right) < \text{E} | \text{g}^{2m+1}(\omega_i) | < \frac{(2m+1)!}{m!} [((\text{Eg}(\omega_i))^2 + \sigma^2)]^m \cdot (|\text{Eg}(\omega_i)| + \sigma) \quad [29]$$

Please note that the upper bound given in Neyman and Scott (1960) disagrees with the RHS of [29], which is the correct upper bound. These bounds are now used to prove a key theorem which allows the interchange of summation and expectation of the Taylor series expansion. The following theorem is adapted from Neyman and Scott (1960, p. 647).

Theorem IV.1 In order that the Taylor series expansion of $\theta(\omega_i)$ be absolutely convergent for all values of $\text{Eg}(\omega_i)$ and σ^2 it is necessary and sufficient that the radii of convergence of the two series

$$\sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(2m)} g^{2m}(\omega_i) \text{ and } \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(2m+1)} g^{2m+1}(\omega_i) \quad [30]$$

both be infinite, so that

$$\lim_{m \rightarrow \infty} \frac{1}{2m} (|f_o^{(2m)}|)^{\frac{1}{m}} = \lim_{m \rightarrow \infty} \frac{1}{2m+1} (|f_o^{(2m+1)}|)^{\frac{1}{m}} = 0. \quad [31]$$

The theorem basically gives the root test criteria for the absolute convergence of the two power series with infinite radius. This assures the the two series are convergent irrespective of specific values of $Eg(\omega_i)$ or σ^2 . A proof is given in Neyman and Scott (1960) and makes use of the bounds on the noncentered moments given in [28] and [29]. Since the Taylor series expansion is absolutely convergent if the inverse transformation function is second order entire, we now have the necessary conditions to allow the interchange of summation and expectation; i.e., the interchange is permissible for all possible values of the random variables in the expansion. This is now formally stated as a theorem after Neyman and Scott (1960, p. 648)

Theorem IV.2. Under the conditions of theorem IV.1, that is equation [31],

$$\theta(\omega_i) = Ef(g(\omega_i)) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} Eg^m(\omega_i) - \eta_2, \quad \forall Eg(\omega_i) \text{ and } \sigma^2. \quad [32]$$

A short justification of the theorem is given in Neyman and Scott (1960, p. 649). However, the result is actually implied by the fact that absolute bounds exist for all the noncentered moments

which implies the interchange of infinite summation and expectation. See the note to theorem A.IV.7. corollary 1.

Adopting the terminology of Neyman and Scott we call any function satisfying the criteria of [31] a second order entire function. As one would expect, Neyman and Scott (1960) mention that two second order entire functions may be added to produce another second order entire function. This fact will be useful in examining finite difference equations involving the Taylor series expansions of second order entire functions.

Having established the form of the expansion for $\theta(\omega_i)$ in theorem IV.2, we now develop its UMVUE. Our goal is to establish the exact form of the UMVUE's $T_m(\omega_i)$. We invoke the standard linear models assumptions (under normality) that $\hat{g}(\omega_i)$ and s^2 are stochastically independent. With the SSE properly normalized by σ^2 , it has a chi-squared distribution with ν degrees of freedom. For the even moments we have from equation [12]

$$T_{2m}(\omega_i) = \sum_{k=0}^m \frac{(2m)!}{(2k)!(m-k)!} [\hat{g}(\omega_i)]^{2k} \left[\frac{1}{4} \text{SSE}(1 - h_{ii}) \right]^{m-k} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)}, \quad [33]$$

and for the odd moments from equation [13]

$$T_{2m+1}(\omega_i) = \sum_{k=0}^m \frac{(2m+1)!}{(2k+1)!(m-k)!} [\hat{g}(\omega_i)]^{2k+1} \left[\frac{1}{4} \text{SSE}(1 - h_{ii}) \right]^{m-k} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)}, \quad [34]$$

Next we state a theorem which gives the exact form of the UMVUE for the i^{th} expected sample phase value $\theta(\omega_i)$. The theorem is adapted from Neyman and Scott (1960, p. 650).

Theorem IV.3. If f is second order entire function, then

$$\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} T_m(\omega_i) - \eta_2 \quad [35]$$

is absolutely convergent $\forall \hat{g}(\omega_i)$ and s^2 and is an unbiased estimator of $\theta(\omega_i)$.

The proof is given by Neyman and Scott (1960) and essentially follows that for theorem IV.1 and is based upon the following absolute bounds for the T_m

$$\frac{\Gamma(\frac{1}{2}v)}{\Gamma(\frac{1}{2} + m - k)} \leq (\frac{2}{v})^{n-k}, \quad [36]$$

$$|T_{2m}(\omega_i)| < \frac{(2m)!}{m!} \left[\frac{v(\hat{g}(\omega_i))^2 + \text{SSE}(1 - h_{ii})}{2v} \right]^m, \quad [37]$$

and

$$|T_{2m+1}(\omega_i)| < |\hat{g}(\omega_i)| \frac{(2m+1)!}{m!} \left[\frac{v(\hat{g}(\omega_i))^2 + \text{SSE}(1 - h_{ii})}{2v} \right]^m. \quad [38]$$

These bounds are simply developed from those given for the noncentered normal moments given in equations [25], [28] and [29]. Since the Taylor series expansion is absolutely convergent by theorem IV.1, we may take expectation inside the summation and demonstrate the unbiasedness of the estimator; i.e.,

$$E\hat{\theta}(\omega_i) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} E T_m(\omega_i) - \eta_2 \quad [39]$$

Because of the unbiasedness of the $T_m(\omega_i)$ we have that $E\hat{\theta}(\omega_i) = \theta(\omega_i)$ as is desired. Furthermore, since our estimator is a function of complete, sufficient statistics, our estimator is unique and an UMVUE of the phase.

IV.4.2. An UMVUE For The Group Delay Function.

In this section we develop an UMVUE for $\theta'(\omega_i)$ by extending the method of Neyman and Scott (1960). Our goal is to find a way of using the UMVUE values in the transformed metric to generate UMVUE values of the expected sample group delay function in the original metric. Essentially we are interested in finding the UMVUE for the derivative of the i^{th} expected sample phase value $\theta(\omega_i)$. It is straightforward to see that this derivative may be evaluated by the usual rule for taking derivative of a power series (see theorems A.IV.2 and A.IV.3); i.e.,

$$\theta'(\omega_i) = \frac{d}{d\omega} E f(g(\omega_i)) = \sum_{m=0}^{\infty} \frac{1}{m!} f_o^{(m)} \frac{d}{d\omega} E g^m(\omega_i) \quad [40]$$

and we wish to find the UMVUE for the derivative of the m^{th} noncentered moment.

The justification for [40] relies on the fact that $f(g(\omega_i))$ is second order entire, which we get from theorem IV.1. Notice, the coefficients of the power series $f_0^{(m)}$ are not affected by the differentiation, so we see that the derived series itself is second order entire and we know from advanced calculus that the radius of convergence of the derived series is the same as the original series. In order to develop the UMVUE expressions it is necessary to be able to exchange the expectation operator and the derivate operator. That is, we wish to show that

$$\frac{d}{d\omega} \text{Eg}^m(\omega_i) = \text{E} \frac{d}{d\omega} g^m(\omega_i). \quad [41]$$

In order to make the interchange it is necessary to find absolute bounds for the LHS of [41] and the bounding function must be integrable with respect to a probability measure in our case. This satisfies the requirements of theorem A.IV.5 corollary 1 for the interchanging derivatives and integrals; we are essentially invoking the Lebesgue dominated convergence criterion for uniform integrability.

In order to find bounding functions for the even noncentered moments notice that by taking derivative of the RHS of [26] we have

$$\frac{d}{d\omega} \text{Eg}^{2m}(\omega_i) = \frac{d}{d\omega} \text{Eg}(\omega_i) \sum_{k=1}^m \frac{(2m)!}{(2k-1)!(m-k)!} [\text{Eg}(\omega_i)]^{2k-1} \left(\frac{\sigma^2}{2}\right)^{m-k} \quad [42]$$

and it can be shown by direct computation that this is absolutely bounded by

$$\left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| < \left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \left| \text{Eg}(\omega_i) \right| \left\{ \frac{(2m)!}{(m-1)!} \left[(\text{Eg}(\omega_i))^2 + \frac{\sigma^2}{2} \right]^{m-1} \right\}. \quad [43]$$

To show that this absolutely bounds the derivative simply compute the terms of the expansion in the RHS of [42] and take term by term absolute value, then use the result that the absolute value of the sum is less than or equal to sum of absolute values. We proceed as follows, using theorem A.IV.8,

$$\left| \frac{d}{d\omega} \text{Eg}^{2m}(\omega_i) \right| \leq \left| \frac{d}{d\omega} \text{Eg}^{2m}(\omega_i) \right| = \left| \frac{d}{d\omega} \text{Eg}(\omega_i) \sum_{k=1}^m \frac{(2m)!}{(2k-1)!(m-k)!} (\text{Eg}(\omega_i))^{2k-1} \cdot \left(\frac{\sigma^2}{2} \right)^{m-k} \right|. \quad [44]$$

Using the triangle inequality and Cauchy-Schwarz's inequality we have

$$\left| \frac{d}{d\omega} \text{Eg}^{2m}(\omega_i) \right| \leq \left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \sum_{k=1}^m \frac{(2m)!}{(2k-1)!(m-k)!} |(\text{Eg}(\omega_i))^{2k-1}| \left(\frac{\sigma^2}{2} \right)^{m-k} \quad [45]$$

Expanding the RHS of [45] we have

$$\left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \left\{ \frac{(2m)!}{(m-1)!} |\text{Eg}(\omega_i)| \left[\frac{\sigma^2}{2} \right]^{m-1} + \frac{(2m)!}{3!(m-2)!} |(\text{Eg}(\omega_i))^3| \left[\frac{\sigma^2}{2} \right]^{m-2} + \frac{(2m)!}{5!(m-3)!} |(\text{Eg}(\omega_i))^5| \left[\frac{\sigma^2}{2} \right]^{m-3} + \dots + \frac{(2m)!}{(2m-1)!} |(\text{Eg}(\omega_i))^{2m-1}| \right\}. \quad [46]$$

Now, expanding the RHS of [43] and noting that $(\text{Eg}(\omega_i))^{2m} |\text{Eg}(\omega_i)| = |(\text{Eg}(\omega_i))^{2m+1}|$ we have

$$\left| \frac{d}{d\omega} \text{Eg}^{2m}(\omega_i) \right| \leq \left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| |\text{Eg}(\omega_i)| \left\{ \frac{(2m)!}{(m-1)!} \left[\frac{\sigma^2}{2} \right]^{m-1} + \frac{(2m)!}{(m-2)!} [\text{Eg}(\omega_i)]^2 \left[\frac{\sigma^2}{2} \right]^{m-2} + \frac{(2m)!}{2!(m-3)!} [\text{Eg}(\omega_i)]^4 \left[\frac{\sigma^2}{2} \right]^{m-3} + \dots + \frac{(2m)!}{(m-1)!} [\text{Eg}(\omega_i)]^{2m-2} \right\}. \quad [47]$$

If a term by term comparison is made between the RHS of [46] and the RHS of [47], it can be seen that the coefficients in [47] are uniformly larger than those in [46] and hence the absolute bound is established. The integrability of the bounding function can be established by noting that it is comprised of terms containing powers of normal and chi-square random variables and hence has finite expectation with respect to a Lebesgue probability measure (the standard definition of expectation). Therefore, for even moments we may interchange expectation and derivative in accordance with theorem A.IV.5 corollary 1.

The interchange of derivative and expectation can be justified for the odd noncentered moments by the same procedure. First we give the form of the derivative of the odd moments by taking derivative of the RHS of [27]

$$\frac{d}{d\omega} \text{Eg}^{2m+1}(\omega_i) = \frac{d}{d\omega} \text{Eg}(\omega_i) \sum_{k=0}^m \frac{(2m+1)!}{(2k)!(m-k)!} [\text{Eg}(\omega_i)]^{2k} \left[\frac{\sigma^2}{2}\right]^{m-k} \quad [48]$$

Again, by direct computation, it can be shown that [48] is absolutely bounded by

$$\left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \frac{(2m+1)!}{m!} \left[(\text{Eg}(\omega_i))^2 + \frac{\sigma^2}{2} \right]^m \quad [49]$$

If we take term by term absolute value of the RHS of [48] and expand the terms we have

$$\begin{aligned} \left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \left\{ \frac{(2m+1)!}{m!} \left[\frac{\sigma^2}{2}\right]^m + \frac{(2m+1)!}{2(m-1)!} [\text{Eg}(\omega_i)]^2 \left[\frac{\sigma^2}{2}\right]^{m-1} + \right. \\ \left. \frac{(2m+1)!}{4!(m-2)!} [\text{Eg}(\omega_i)]^4 \left[\frac{\sigma^2}{2}\right]^{m-2} + \dots + \frac{(2m+1)!}{(2m)!} [\text{Eg}(\omega_i)]^{2m} \right\} \end{aligned} \quad [50]$$

Next we expand the bounding function in [49] and compare the results with the expansion of the LHS of [48] in [50],

$$\left| \frac{d}{d\omega} \text{Eg}(\omega_i) \right| \left\{ \frac{(2m+1)!}{m!} \left[\frac{\sigma^2}{2} \right]^m + \frac{(2m+1)!}{(m-1)!} [\text{Eg}(\omega_i)]^2 \left[\frac{\sigma^2}{2} \right]^{m-1} + \frac{(2m+1)!}{2!(m-2)!} [\text{Eg}(\omega_i)]^4 \left[\frac{\sigma^2}{2} \right]^{m-2} + \dots + \frac{(2m+1)!}{m!} [\text{Eg}(\omega_i)]^{2m} \right\}. \quad [51]$$

By comparing the expansion in [50] to that in [51], we see that the absolute value of the LHS of [48] is absolutely bounded by expression [49]. Once again we appeal to the fact that the terms in the bounding function are powers of normal and chi-square random variables to justify integrability. Therefore we conclude that the interchange of expectation and derivative is justified for all non-centered normal moments of order $m = 1, 2, 3, \dots$

With the assumption that we can interchange the expectation and derivative, our expression for the expected value of the group delay function becomes

$$\theta'(\omega_i) = \sum_{m=1}^{\infty} \frac{1}{m!} f_o^{(m)} \text{E} \frac{d}{d\omega} g^m(\omega_i) \quad [52]$$

and our task is to find UMVUE's for $\text{E} \frac{d}{d\omega} g^m(\omega_i)$. Since the $g(\omega_i)$ are normal random variables we assume their differentiability in general, although in our case this is specifically justified because we are taking derivatives of polynomial expressions in ω . What is to be shown is that

$$\text{E} \frac{d}{d\omega} T_m(\omega_i) = \text{E} \frac{d}{d\omega} g^m(\omega_i) = \frac{d}{d\omega} \text{E} g^m(\omega_i). \quad [53]$$

Let,

$$T'_m(\omega_i) = \frac{d}{d\omega} T_m(\omega_i) \quad [54]$$

where $T_m(\omega_i)$ is our UMVUE for the noncentered moments given in [33] and [34]. What remains to be established is the interchange of the expectation and derivative in LHS of [53]; i.e., we need to show that

$$ET'(\omega_i) = \frac{d}{d\omega} ET_m(\omega_i) = \frac{d}{d\omega} Eg^m(\omega_i). \quad [55]$$

In order to justify the interchange we must show, as we did above, that the derivative of the estimator is absolutely bounded by an integrable function; then we may appeal to theorem A.IV.5 corollary 1. First we proceed for the even moments

$$T'_{2m}(\omega_i) = \sum_{k=0}^m \frac{(2m)!}{(2k)!(m-k)!} \{2k\hat{g}'(\omega_i)[\hat{g}(\omega_i)]^{2k-1} [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k} - (m-k)[\hat{g}(\omega_i)]^{2k} (\frac{\text{SSE}}{4} \frac{d}{d\omega} h_{ii}) [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k-1}\} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2}+m-k)}, \quad [56]$$

$$\leq \sum_{k=1}^m \frac{(2m)!}{(2k)!(m-k)!} \{2k|\hat{g}'(\omega_i)[\hat{g}(\omega_i)]^{2k-1} [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k} - (m-k)[\hat{g}(\omega_i)]^{2k} (\frac{\text{SSE}}{4} \frac{d}{d\omega} h_{ii}) [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k-1}\} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2}+m-k)}. \quad [57]$$

Unfortunately, there does not appear to be a compact re-expression for the RHS of [57], however we will justify the interchange of expectation and derivative using the results of [56] and [57].

Theorem IV.4. Let $T_{2m}(\omega_i)$ be the UMVUE of $Eg^{2m}(\omega_i)$ and by [56] $T_{2m}(\omega_i)$ is differentiable with respect to ω . Furthermore, [56] is an unbiased estimator of $\frac{d}{d\omega} Eg^{2m}(\omega_i)$.

The proof of the theorem depends upon being able to make the interchange of the derivative and the expectation of [56] and we proceed as follows. Using the differentiability shown in [56], define the finite difference function

$$T_{2m}(\Delta\omega) = \frac{T_{2m}(\omega + \Delta\omega) - T_{2m}(\omega)}{\Delta\omega} \quad [A]$$

and given any $\varepsilon > 0 \exists$ a positive $\delta(\varepsilon) \ni |\Delta\omega| < \delta \Rightarrow |T_{2m}(\Delta\omega) - T'_{2m}(\omega)| < \varepsilon$, now using the triangle inequality we have

$$|T_{2m}(\Delta\omega)| \leq |T'_{2m}(\omega)| + \varepsilon \quad \forall \Delta\omega \in B_\delta(0)$$

where $B_\delta(0)$ is an open ball about 0. We appeal to the fact that the RHS of [57] is comprised of terms with powers of normal and chi-square random variables to imply that it has finite expectation and is integrable with respect to a Lebesgue probability measure. Having established a bound for [A], we now appeal to the LDCT or theorem A.IV.5 to justify the interchange of derivative and expectation. The remainder of the proof demonstrates the unbiasedness of [56]. The proof goes as follows, first define

$$\lim_{\Delta\omega \rightarrow 0} \int \frac{T_{2m}(\omega + \Delta\omega) - T_{2m}(\omega)}{\Delta\omega} d_{\mu}(\omega) = \int \lim_{\Delta\omega \rightarrow 0} \frac{T_{2m}(\omega + \Delta\omega) - T_{2m}(\omega)}{\Delta\omega} d_{\mu}(\omega) \quad [\text{B}]$$

where $d_{\mu}(\omega)$ is a measure and by the LDCT we have the interchange of limits and integration.

We now have the desire conclusion

$$\therefore \lim_{\Delta\omega \rightarrow 0} \frac{E[T_{2m}(\omega + \Delta\omega) - T_{2m}(\omega)]}{\Delta\omega} = \frac{d}{d\omega} ET_{2m}(\omega) \quad [\text{C}]$$

and interchanging limit and expectation on the LHS of [C], we have

$$E \lim_{\Delta\omega \rightarrow 0} \frac{T_{2m}(\omega + \Delta\omega) - T_{2m}(\omega)}{\Delta\omega} = E \frac{d}{d\omega} T_{2m}(\omega) = \frac{d}{d\omega} Eg^{2m}(\omega). \quad [\text{D}]$$

■

The result for the odd moments follows in a straightforward manner.

$$\begin{aligned} T'_{2m+1}(\omega_i) = & \sum_{k=0}^m \frac{(2m+1)!}{(2k)!(m-k)!} \{2k\hat{g}'(\omega_i)[\hat{g}(\omega_i)]^{2k} [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k} - \\ & (m-k)[\hat{g}(\omega_i)]^{2k+1} (\frac{\text{SSE}}{4} \frac{d}{d\omega} h_{ii}) [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k-1}\} \frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)} \end{aligned} \quad [\text{58}]$$

and the absolute bound is

$$\begin{aligned}
|T'_{2m+1}(\omega_i)| &= \sum_{k=0}^m \frac{(2m+1)!}{(2k)!(m-k)!} \{ |2k\hat{g}'(\omega_i)[\hat{g}(\omega_i)]^{2k} [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k} - \\
(m-k)[\hat{g}(\omega_i)]^{2k+1} (\frac{\text{SSE}}{4} \frac{d}{d\omega} h_{ii}) [\frac{1}{4} \text{SSE}(1-h_{ii})]^{m-k-1} | \} &\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2} + m - k)}. \tag{59}
\end{aligned}$$

Next we state a theorem to justify [58] as the UMVUE for odd noncentered normal moments.

Theorem IV.5. Let $T_{2m+1}(\omega_i)$ be the UMVUE of $Eg^{2m+1}(\omega_i)$ and by [58] $T_{2m+1}(\omega_i)$ is differentiable with respect to ω . Furthermore, [58] is an unbiased estimator of $\frac{d}{d\omega} Eg^{2m+1}(\omega_i)$.

The proof follows exactly as that for theorem IV.4 and is omitted.

With the establishment of UMVUE's for the odd and even noncentered moments, we now are able to justify our UMVUE for $\theta'(\omega_i)$ by a theorem.

Theorem IV.6. Under the conditions of [31] and theorems IV.4 and IV.5,

$$\hat{\theta}'(\omega_i) = \sum_{m=1}^{\infty} \frac{1}{m!} f_o^{(m)} T'_m(\omega_i) \tag{60}$$

is absolutely convergent \forall values of $\hat{g}(\omega_i)$ and s^2 and is an unbiased estimator of $\theta'(\omega_i)$.

The proof is very straightforward and follows from the fact that theorem IV.1 allows the interchange of summation and expectation and the unbiasedness of [56] and [58]; i.e.,

$$\begin{aligned}
E\hat{\theta}'(\omega_i) &= \sum_{m=1}^{\infty} \frac{1}{m!} f_o^{(m)} E\Gamma'_m(\omega_i) = \\
&= \sum_{m=1}^{\infty} \frac{1}{m!} f_o^{(m)} E \frac{d}{d\omega} g^m(\omega_i) = \theta'(\omega_i)
\end{aligned}$$

■

Since [56] and [58] are both functions of complete, sufficient, statistics, we appeal to Lehmann and Scheffe' (1950) to justify the UMVUE property of [60].

IV.4.3 Estimating Group Delay By Finite Difference.

Although theoretically we established an UMVUE for $\theta'(\omega_i)$ in section IV.2, a computationally simpler approach is to use finite difference equations to estimate expected sample group delay at each frequency. The theory of finite difference equations can be reviewed in Boole (1970) or Richardson (1954). However, the application is straightforward. We define the finite difference of the sample phase function at the Fourier frequency ω_o by

$$\frac{\Delta y(\omega_o)}{\Delta\omega} = \frac{y(\omega_o + \Delta\omega) - y(\omega_o)}{\Delta\omega} \tag{61}$$

We introduce the notation

$$GD(\Delta\omega_o) = \frac{\Delta y(\omega_o)}{\Delta\omega} \tag{62}$$

for the finite difference at Fourier frequency ω_o .

Since $GD(\Delta\omega_o)$ is a simple linear difference, it is easy to see that

$$\begin{aligned} EGD(\Delta\omega_o) &= \frac{Ey(\omega_o + \Delta\omega) - Ey(\omega_o)}{\Delta\omega} \\ &= \frac{\theta(\omega_o + \Delta\omega) - \theta(\omega_o)}{\Delta\omega}. \end{aligned} \quad [63]$$

Assuming the differentiability of the expected sample phase function

$$\lim_{n \rightarrow \infty} EGD(\Delta\omega_o) = \lim_{\Delta\omega \rightarrow 0} \frac{Ey(\omega_o + \Delta\omega) - Ey(\omega_o)}{\Delta\omega} = EGD(\omega_o). \quad [64]$$

Therefore, as the sample size increases the finite differences approach the derivative of the phase function, i.e., group delay. The approach is particularly appealing because Fourier frequencies are evenly spaced across the band $(0, \pi]$. But, we have UMVUE's for all the $\theta(\omega_i) \in (0, \pi]$, so we have that

$$\begin{aligned} \hat{EGD}(\Delta\omega_o) &= \frac{\hat{\theta}(\omega_o + \Delta\omega) - \hat{\theta}(\omega_o)}{\Delta\omega} \\ &= \frac{1}{\Delta\omega} \sum_{m=1}^{\infty} \frac{1}{m!} t_o^{(m)} [\Gamma_m(\omega_o + \Delta\omega) - \Gamma_m(\omega_o)]. \end{aligned} \quad [65]$$

The RHS of [65] is justified by the usual result that two convergent power series may be subtracted term by term within their radius of convergence, which is infinite in our case (see, theorem A.IV.1).

We can also demonstrate the absolute convergence of the RHS of [65] by using the absolute bounds given in [37] and [38]; i.e.,

$$\begin{aligned}
|T_m(\omega_o + \Delta\omega) - T_m(\omega_o)| &\leq |T_m(\omega_o + \Delta\omega)| + |T_m(\omega_o)| \\
&\leq \frac{(2m)!}{(2\nu)^m m!} \{[\hat{g}(\omega_o + \Delta\omega)]^{2\nu} + \text{SSE}(1 - h_{\Delta\Delta})\}^m + \{[\hat{g}(\omega_o)]^{2\nu} + \text{SSE}(1 - h_{oo})\}^m. \quad [66]
\end{aligned}$$

Therefore, the finite difference of the terms in the expansion of [65] are absolutely bounded for all values of $\hat{g}(\omega_i)$ and s^2 .

The finite difference approach would appear to be computationally simpler than the direct computation of the series expansion for $\hat{\theta}'(\omega_i)$ in cases where it is necessary to have a large number of terms in the series. Assuming that some series will not converge rapidly, it would seem a useful alternative to estimating group delay. The approach is examined in the simulation study of chapter V.

IV.5. Estimating The Variance Of The UMVU Estimators.

In section IV.4 we developed UMVUE's for the expected sample phase function and expected sample group delay function. Although this is certainly a desirable accomplishment, one would also desire some estimate of the error involved in any particular estimate generated by these methods. The purpose of this section is to discuss some possible strategies for estimating the standard error (s.e.) of our estimates and then approximate a confidence interval (CI) based upon this s.e.. The question of estimating the variance of $\hat{\theta}(\omega_i)$ was addressed by Hoyle (1968) as an addition to his method for finding the UMVUE for $\phi^2(\omega_i)$. Hoyle's approach was to solve the equation

$$\text{Var}(\hat{\theta}(\omega_i)) = E(\hat{\theta}(\omega_i))^2 - (\theta(\omega_i))^2. \quad [67]$$

In order to find an estimate of $\text{Var}(\hat{\theta}(\omega_i))$ Hoyle used the approximation

$$\hat{\text{Var}}(\hat{\theta}(\omega_i)) = (\hat{\theta}(\omega_i))^2 - E^{-1}[(\theta(\omega_i))^2]. \quad [68]$$

where $E^{-1}[(\theta(\omega_i))^2] \equiv \text{UMVUE}$ of $(\theta(\omega_i))^2$. In order to find $E^{-1}[(\theta(\omega_i))^2]$, Hoyle further solves the equation

$$(\theta(\omega_i))^2 = E(y(\omega_i))^2 - \phi^2(\omega_i). \quad [69]$$

Hoyle (1968) develops an UMVUE for $\phi^2(\omega_i)$ and an UMVUE of $E(y(\omega_i))^2$ can be obtained from the N-S estimator. In section 8 of his paper, Hoyle gives specific variance estimates for $\hat{\theta}(\omega_i)$ for different transformation schemes. None of the variances apply to the Box-Cox transformation.

The scope of the present work does not include developing an UMVUE for $\text{Var}(y(\omega_i))$ or $\text{Var}(y'(\omega_i))$ for the Box-Cox transformation, so the work of Hoyle does not directly apply to the current problem of finding an estimated variance for our UMVUE's of phase and group delay; although this is a topic for further research. An alternative to the approach of Hoyle (1968), is to estimate the variance directly from the inverse transformation Taylor series. In most cases the transformation either truncates after a small number of terms or the coefficients go to 0 sufficiently fast to truncate the expansion after a relatively small number of terms. Since the random terms in the expansion are normal and chi-square random variables, it is possible to solve for the variance of the individual terms and their covariance terms. However, the covariance terms do

involve higher normal and Chi-squared moments. As a result some of the covariance terms may be difficult to estimate in practice.

IV.5.1. Approximate Confidence Intervals For The UMVUE Values.

Land (1974) performed a simulation study on confidence interval estimation for mean estimates derived by the inverse transformation method. He states,

A familiar inference problem, for which there appears to be no standard solution, occurs when data are transformed to meet the distributional requirements of a spherical normal linear model; and a confidence interval is sought for the mean of the variate in the original, untransformed scale.

Hoyle (1973) and Land (1974) separate the various approaches to the problem into two broad categories; (1) the direct methods, and (2) transformation methods. The direct methods estimate the s.e. $\hat{\theta}(\omega_i)$ and assume some distribution for $\hat{\theta}(\omega_i)$. Using these approximations, an estimated CI is calculated for $\hat{\theta}(\omega_i)$. The transformation methods develop direct $1 - \alpha$ level CI's in the transformed domain; i.e., we would find $\hat{g}_{\frac{\alpha}{2}}(\omega_i) \leq Eg(\omega_i) \leq \hat{g}_{1-\frac{\alpha}{2}}(\omega_i)$. In order to approximate the CI for $\hat{\theta}(\omega_i)$, the endpoints of the CI are inversely transformed by the N-S estimator (in our particular case). Interestingly, the most common transformation approach is to find the CI in the transformed domain and naively perform a simple inverse transformation on the endpoints of the interval. The method does generate a correct $1 - \alpha$ level confidence interval for the median, but in a simulation study Land (1974) found that the method does not approximate the $1 - \alpha$ level CI for the mean very well. In fact, no general approach appears satisfactory and much work appears to remain on this topic.

In his simulation study, Land (1974) did find that the transformation method works reasonably well when UMVUE methods are employed to inversely transform the endpoints. However, his study was based upon one-sided CI procedures and was somewhat narrow in scope. A number of more involved approaches are discussed by Land (1974) which involve developing confidence regions based upon various values for our sufficient statistics $\hat{g}(\omega_i)$ and $\hat{\sigma}^2$. The methods are more difficult to implement than those discussed in the previous paragraph and it is not clear that they perform any better in application. The reader is referred to Land's paper for more details on the various approaches to the problem.

In the present work, the method of inversely transforming the endpoints of a CI on a mean estimate in the new metric is adopted. The method is relatively easy to apply and the limited work by Land (1974) gave some indication that it works relatively well; although the actual performance depends upon the transformation used and the variability in the data (Land, 1974). The topic is discussed further in the simulation study of chapter V, and is a topic for future research.

IV.6. Alternative Inverse Transformation Strategies.

The emphasis of the present research is the extension of the N-S estimator to the problem of estimating spectral parameters. However, a number of alternative transformation strategies exist which generate UMVUE's or approximate UMVUE's and may be of use in cases where the Taylor series expansion in the N-S method does not exist or is unwieldy.

Neyman and Scott (1960) propose an alternative set of inverse transformation schemes based upon the theory of recursive differential equations; i.e., a function f is said to be of the recursive type if it satisfies the second order differential equation

$$f''(\omega) = A + Bf(\omega) \quad [70]$$

where A and B are arbitrary constants (at least one not zero). It turns out that every recursive function is second order entire, but the converse is not true. As an example, the cube root inverse transformation function is second order entire, but it is not recursive. Therefore, this approach does not apply, in general to the Box-Cox transformation and has not been pursued in the present work. However, the UMVUE for the inverse log transform and the inverse arcsin transform appear to be useful for general application; they are given in equations [54] and [58], respectively, of Neyman and Scott (1960). They do involve Bessel functions of complex argument, but appear solvable in application.

Interestingly, a group of commonly used transformations, the reciprocal transformations, do not have an exact solution for the UMVUE of the mean of the inverse transformation. In other words

$$g(\omega_i) = \frac{1}{y(\omega_i)} \quad [71]$$

and

$$\theta(\omega_i) = E\left[\frac{1}{y(\omega_i)} \right] \quad [72]$$

does not exist, in general, so it can not be directly estimated. Box (1971) considered generating an UMVUE for the inverse mean from a reciprocal transformation by defining a pseudo-expectation function and estimating the UMVUE by that function. An alternative is provided by Miller (1984) where he gives a low bias estimator for reciprocal transformations; and his estimator appears quite easy to implement in practice. Since reciprocal transformations do not seem useful for phase functions, the topic is not considered further in the present research.

Appendix A.IV.

In this section are a number of definitions, and theoretical results from advanced calculus and measure theory, which are intended to supplement the theoretical results given in chapter IV. Although formal proofs are not given, appropriate references are given where such proofs are available.

Definition A.IV.1. A function $f(x)$ is analytic at $x = a$ if and only if it has a Taylor series representation at $x = a$ that represents the function in some neighborhood of $x = a$. (Olmsted, 1961).

Definition A.IV.2. Every polynomial is everywhere represented by all its Taylor series (over its entire domain). If $f(x)$ is a polynomial, $f^{(n)}(x)$ exists for all n and x , and $f^{(n)}(x) = 0$ for all n greater than the degree of the polynomial (Olmsted, 1961).

Theorem A.IV.1. (Olmsted, 1961) Let $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ be two power series representing functions $f_1(x)$ and $f_2(x)$, within their intervals of convergence, and let γ be an arbitrary constant. then

1. The power series $\sum \gamma a_n x^n$ represents $\gamma f_1(x)$ throughout the interval of convergence,
2. The power series $\sum (a_n \pm b_n) x^n$ represents $f_1(x) \pm f_2(x)$ throughout their radii of convergence,

3. if $C_n \equiv \sum_{k=0}^n a_k b_{n-k}$, $n = 0, 1, 2, \dots$, then the power series $\sum_{n=0}^{\infty} C_n x^n$ represents the function $f_1(x)f_2(x)$ for all points interior to both intervals of convergence.

The above theorem simply reiterates the well known fact that power series may be added, subtracted, or multiplied term by term within their radius of convergence.

Theorem A.IV.2. (Olmsted, 1961): A power series and its derived series have the same radius of convergence.

Theorem A.IV.3. (Olmsted, 1961): A function $f(x)$ represented by a power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ in the interior of its interval of convergence is differentiable there, and its derivative is represented there by the derived series $f'(x) = \sum_{n=1}^{\infty} n a_n (x-a)^{n-1}$.

Theorem A.IV.4. (Rainville, 1967): the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ either

1. Converges only at $x = a$,
2. Converges absolutely and uniformly in every finite closed interval, or
3. \exists a positive number R \ni the series is absolutely convergent $\forall |x-a| < R$, divergent $\forall |x-a| > R$, and is uniformly convergent in the closed interval $|x-a| \leq r \leq R$

The following are standard results from measure theory.

Theorem A.IV.5. (Billingsley, 1986) The Lebesgue Dominated Convergence Theorem: If $|f_n| \leq g$ a.e., where g is integrable, and if $f_n \rightarrow f$ a.e., then f and the f_n are integrable and $\int f_n d\mu \rightarrow \int f d\mu$ or $\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu$.

Corollary 1. (Rao, 1973): Let $dg(\omega, x)/dx$ exist in an interval (a,b) of x , and $|dg(\omega, x)/dx| < G(\omega)$ integrable, then in (a,b)

$$\frac{d}{dx} \int g(\omega, x) d\mu = \int \frac{d}{dx} g(\omega, x) d\mu.$$

The following 2 theorems are specifically for infinite series.

Theorem A.IV.6. (Billingsley, 1986) If $f_n \geq 0$, then $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$. (monotone convergence theorem).

Theorem A.IV.7. (Billingsley, 1986) If $\sum_n f_n$ converges a.e. and $|\sum_{k=1}^n f_k| \leq g$ a.e., where g is integrable, then $\sum_n f_n$ and f sub n are integrable and $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$.

Corollary 1(Billingsley, 1986): If $\sum_n |f_n| d\mu < \infty$, then $\sum_n f_n$ converges absolutely a.e. and is integrable and $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$.

Note: the corollary justifies the interchange of expectation and summation of infinite sums, if the sums of expected values converge absolutely. (see Rao, 1973, p. 111 or Gnedenko, 1967, p. 216) for more details.

Theorem A.IV.8. (Milton and Tsokos, 1976): $E x$ exists if and only if $E |x|$ exists and furthermore $|E x| \leq E |x|$.

Chapter V

Simulation Study of The UMVUE's For Phase and Group Delay.

V.1. Introduction.

In sections IV.4.1 and IV.4.2, the UMVUE's for the phase and group delay functions, respectively, were developed. The development in Chapter IV was entirely theoretical. Therefore, it is of interest to perform a simulation study to demonstrate the practicality of the theoretical results of Chapter IV.

The simulation study of this chapter is an attempt to show how the UMVUE procedures work in actual practice; but, under simulated conditions. The simulation study given here is not an attempt to thoroughly investigate every aspect of the proposed methods of estimating phase and group delay given in Chapter IV. Since there do not appear to be any published, large scale simulation studies of phase and group delay estimation, there is no past history to supplement the current simulation effort. In fact, no published, comprehensive simulation study of inverse transformation bias could be located by this author. With little supplemental information from past simulation studies, it was not feasible to attempt to address every open issue on the subject of phase and group delay estimation and inverse transformation bias. Indeed, such an ambitious study would entail a considerable expenditure of resources and time, to a point beyond the scope of the present research.

The basic approach of the present simulation study is to use linear filter theory, as discussed in section II.4, to generate simulated sample phase and group delay functions. The idea is to select a linear filter (we always assume a time-invariant filter in the present discussion) which has a known phase and group delay function relating a univariate input time series to a univariate output time series. With a known phase and group delay relationship between the two series, one can generate random time series and apply the UMVUE methods of Chapter IV to estimate the underlying phase and group delay functions.

There are two approaches one could take in conducting the simulation, once a desirable filter has been selected. The first approach is to simulate two random time series, in the time domain, and then transform the observed series to the frequency domain via the Fourier transform method. Once in the frequency domain, one can generate random phase observations by the arc tangent

method discussed in section II.3. The estimates of phase and group delay are then generated by working with the sample phase observations. The second approach is more direct and works entirely in the frequency domain. In this approach, one directly generates a random phase function and applies UMVUE techniques to the random phase observations to generate estimates of the underlying phase and group delay functions.

The first approach has the advantage of simulating the entire procedure one would follow in estimating phase and group delay; but, it has the disadvantage of requiring far more computational cost, since one must repeatedly transform between the time and frequency domain. An additional disadvantage of the first approach is that one must attempt to simulate the phase function by a digital filter and it is not possible to simulate all continuous time phase functions with digital filters. The second approach has the advantage of working directly with the phase function and thereby eliminating the need for repetitive Fourier transformations. Furthermore, one can directly generate any desired phase function without attempting to simulate the phase function with a digital filter. This is particularly important because the current research has a lot of applications to time series that are related by continuous time filters that are not easily approximated by digital filters. For the present simulation study, the second approach is taken since it is less costly and one can easily generate a random phase function with any desired characteristics.

V.1.1. The Continuous Time "Scratch" Filter.

An important class of continuous time linear filters are described by the equation

$$y(t) = \int_0^{\infty} g(s)x(t-s)ds \quad [1]$$

where; $g(s)$ is a time-invariant parameter of the filter, $x(t)$ is the input to the filter at time t , and $y(t)$ is a continuous, weighted combination of past inputs to the system. Equation [1] is the general form for linear, continuous time filters.

An important subset of these filters described by [1] can be described by the relationship

$$y(t) = \frac{1}{T} \int_0^{\infty} e^{-s/T} x(t-s)ds, \quad [2]$$

where

$$g(s) = \frac{1}{T} e^{-s/T} \quad [3]$$

is referred to as time constant of the filter. This class of filter is discussed by Priestly (1981, p. 21 and p. 272) or Chatfield (1984, p. 201). The type of filter in [2] is sometimes referred to as a "scratch" filter, because it is the model for a simple noise filter in an electronic amplifier. In other words the filter is a low pass type, which removes high frequency components of the input series.

However, this class of filter has wide applicability to continuous time processes in many fields of study where the effects of inputs to the process die-off exponentially with passing time.

The time constant of the filter is also the impulse response function of the filter, as discussed in Chapter II; i.e.,

$$g(s) = \begin{cases} \frac{1}{T} e^{-s/T} & s \geq 0 \\ 0 & s < 0. \end{cases} \quad [4]$$

Taking the Fourier transform of $g(s)$ we generate the frequency response function

$$H(\omega) = \frac{(1 - i\omega T)}{(1 + \omega^2 T^2)} \quad [5]$$

By taking the norm of [5], one generates the gain function of filter

$$G(\omega) = \frac{1}{(1 + \omega^2 T^2)^{\frac{1}{2}}}. \quad [6]$$

Using straightforward calculations one can calculate the phase, group delay, and coherence squared functions of the filter; i.e.,

$$\phi_{xy}(\omega) = -\text{Tan}^{-1}(T\omega), \quad [7]$$

$$\text{GD}_{xy}(\omega) = -\frac{T}{1 + T^2\omega^2}, \quad [8]$$

and

$$\rho_{xy}^2(\omega) = \frac{\sigma_x^2}{\sigma_x^2 + (1 + T^2 \omega^2)(1 - 2\alpha \cos(\omega) + \alpha^2)\sigma_y^2} \quad [9]$$

From [7] it can be seen that the phase function is a monotonically decreasing function of increasing frequency. In fact, as $\omega \rightarrow \infty$, $\phi_{xy}(\omega) \rightarrow \frac{-\pi}{2}$ which implies that the high frequency components are approximately 90° out of phase. The coherence between the input and output series is controlled by the parameters σ_x^2 , σ_y^2 , T and α . Here, σ_x^2 and σ_y^2 are the variances of the input and the output series, respectively. The parameter α is a characteristic of input series $X(t)$. Examining [9], it can be noticed that the coherence is also a decreasing function of increasing frequency; this confirms the fact that the filter is a low pass type. Therefore, the input and output series are less coherent at the higher frequency components. The coherence function is particularly important to the present work, because the variance of the sample phase function is inversely proportional to the coherence squared at each frequency; i.e., the variance of the phase function becomes infinite as the coherence goes to 0. As a result, in the low pass filter the variance of the sample phase increases as the frequency increases.

In order to perform the simulation, it is necessary to assign values to all the parameters in equations [5] through [9]. Each of the parameters actually has an infinite set of possible values, although there are usually reasonable ranges of possible values. In the present simulation, values for each of the parameters were empirically selected in an attempt to generate sample phase functions which were realistic in character. However, a large scale study simulation of the phase and group delay function would involve examining many different types of filters and varying values of

the parameters of those filters. Such a study is simply beyond the scope of the present work, so the study was performed with the single set of empirically determined values. The following parameter settings were used to conduct the present simulation study:

1. $T = 1.00$
2. $\alpha = 0.45$
3. $\sigma_x^2 = 3.00$
4. $\sigma_y^2 = 0.20$

The theoretical sample size was set at 200 observations, which is small for spectral analysis.

V.2. Generating A Random Sample Phase Function.

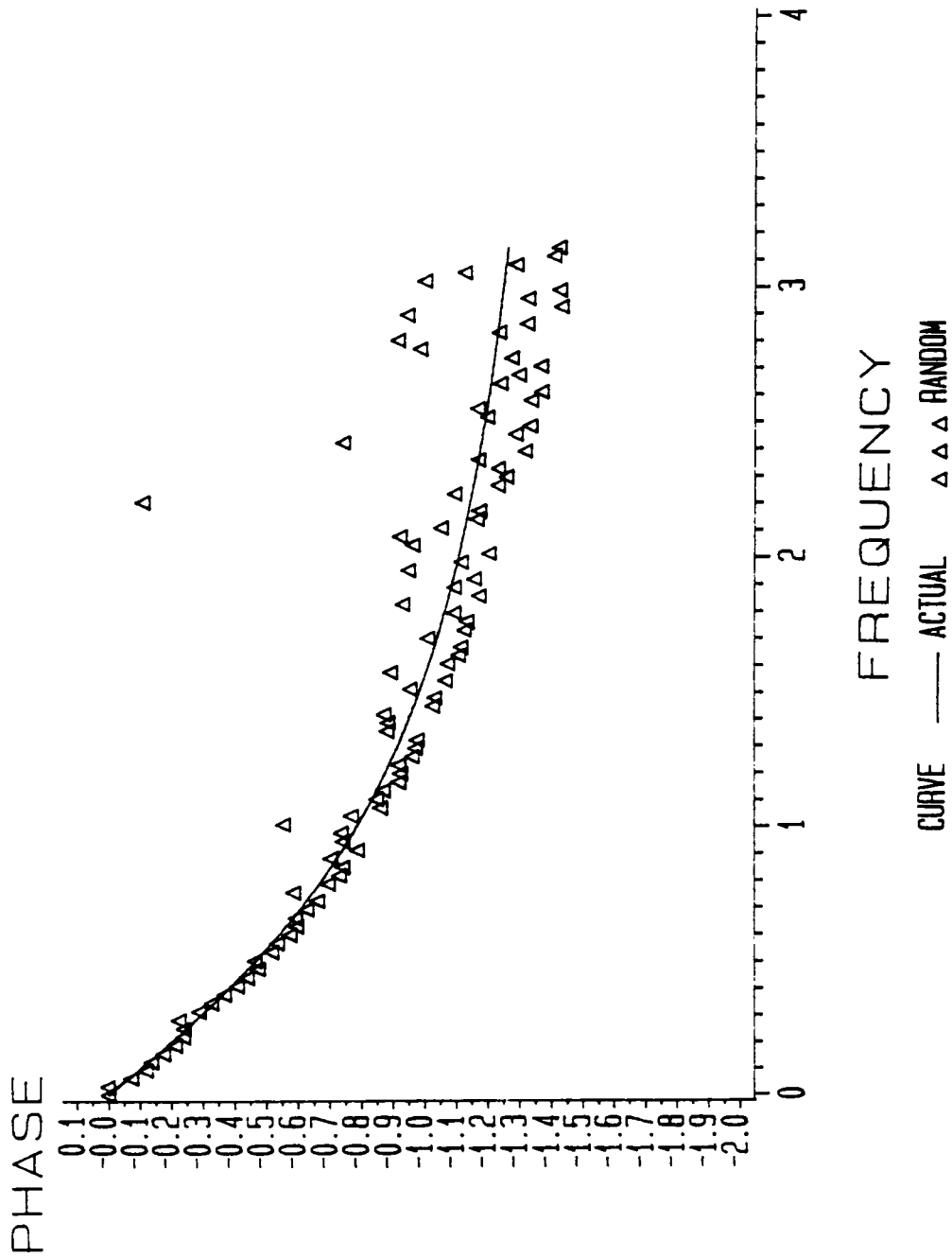
By use of [7] and [8], it is possible to generate nonrandom phase and group delay functions of virtually any desired level of nonlinearity. Recall, that group delay implies a nonlinear relationship between phase and frequency. However, our goal is to study the UMVU estimation of phase and group delay using random sample phase observations. Therefore, one must convert the nonrandom phase values to a simulated random sample of observations for each of the Fourier frequencies. In the process of generating the random phase observations it is also desirable to simulate the skewness and heterogeneity of variance which was discussed in section IV.1.

In order to simulate the skewness, a random vector of error values was generated such that it had a skewed distribution. The skewness of the distribution was accomplished by generating a set of random observations from a log normal distribution using the IMSL subroutine DRNLNL. The vector of log normal observations were then centered to obtain a simulated set of 0-mean, random error terms. The variance of the random error term could be adjusted to any desired value by use of the IMSL subroutine DSCAL. In the present study, the variance of the error term was set to 0.10 in order to keep the magnitude of the error values within the range of the sample phase observations.

In order to simulate the heterogeneity in variance of the sample phase observations, a weighting function was applied to the random error term at each Fourier frequency. The error at each frequency was assigned the weight

$$\left[\frac{1 - \rho_{xy}^2(\omega)}{\rho^{2xy}(\omega)} \right]^{\frac{1}{2}}. \quad [10]$$

Recall, from Chapter II equation [13] that the asymptotic variance of the sample phase is directly proportional to [10]. Since our filter is a low pass type, [10] implies that the sample phase variance is an increasing function of increasing frequency. Figure V.1 is a typical random, sample phase function for the simulation study.



PLOT OF RANDOM VS ACTUAL PHASE FUNCTION, $T=1.0$

Figure V.1

V.3. Box-Cox Transformation Of The Sample Phase

Observations.

In the present simulation study, and in naturally occurring phase functions, negative values of phase are often encountered. As a result, the 2 parameter Box-Cox transformation scheme has to be employed. Recall from Chapter III equation [5] that the two parameter form is

$$g(\omega) = \begin{cases} \frac{[y(\omega) + \eta_2]^{\eta_1} - 1}{\eta_1} & \eta_1 \neq 0 \\ \ln[y(\omega) + \eta_2] & \eta_1 = 0 \end{cases} \quad [11]$$

where $\eta_2 > -\min(y(\omega_i))$. There is not a great deal of discussion, in the literature, for the 2 parameter transformation case, compared to the single parameter case. However, it is a straightforward extension of the well studied single parameter case. The application of the two parameter case is discussed by Draper and Smith (1981, p. 236).

The objective function for the two parameter case is

$$L_{\max}(\eta) = -\frac{n}{2} \ln \left(\frac{\text{SSE}(\eta)}{n} \right) + (\eta_1 - 1) \sum_{i=1}^n \ln(y(\omega_i) + \eta_2). \quad [12]$$

Here, $\text{SSE}(\eta)$ is the residual sum of squares after fitting a simple linear model to the transformed response values. The goal is to find those values of η_1 and η_2 which maximize [12] subject to the

constraint on η_2 . Although the response surface is degenerate at $\eta_2 = -\min_i(y(\omega_i))$, Draper and Smith (1981, p. 236) point out that it is irrelevant; one is only concerned with the local maximum subject to the constraint on η_2 .

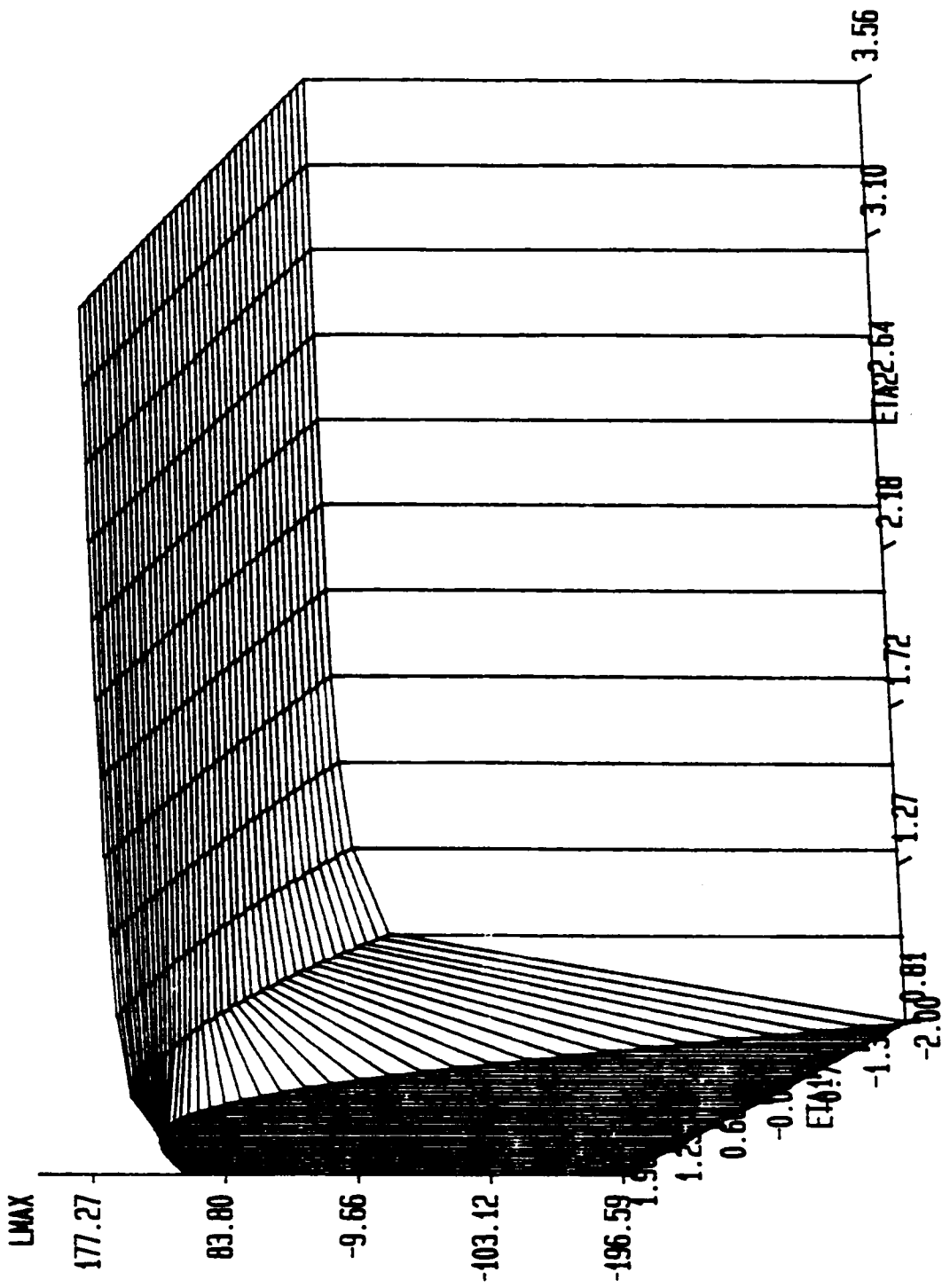
In the initial stages of the simulation study, 10 independent sets of random, sample phase functions were generated to examine the Box-Cox transformation required for the study. The optimization of [12] consistently gave the same range of possible values for the power parameter η_1 which was $1/2 \leq \eta_1 \leq 2/3$. Draper and Smith (1981, p. 226) note that one usually selects a convenient value for η_1 in the range of possible values; rather than using the exact value indicated by the optimization of the objective function [12]. The range of possible values is selected by the confidence interval method discussed in section III.2.1. Unfortunately, in the two parameter case we have to work with a confidence region, which has only approximate properties and makes interpretation difficult (see Box and Cox, 1964, p. 225).

Over the 10 independent cases, the maximum for [12] always occurred at the minimum possible distance η_2 could be set from the singularity; i.e., $\eta_2 = \min_i(y(\omega_i)) + \delta$, where δ is a small positive constant dependent upon the numerical precision of the computing system. In the present case it was necessary to set $\delta = 1.0 \times 10^{-15}$ in order to avoid run time errors during the maximization of [12]. The local maximum value of [12] always occurred with δ set to the minimum possible value. Since, η_2 always took the value closest to the singularity, the range of values for η_1 was determined by treating it as a single parameter case and applying a single confidence interval method to η_1 with a confidence coefficient of 0.95. In the range of values for η_1 , the values of 1/2 and 1/3 are both convenient values, since they lead to Taylor series expansions with a finite number of nonzero terms. Figures V.2 and V.3 are 2 different perspectives of a typical response surface for [12].

The general shape of the response surface shown in figures V.2 and V.3 was typical of that observed for all 10 of the independent cases examined before running the actual simulation. As η_2 increased in value, the response surface consistently transitioned into a broad plateau with subtle contours. Large radial searches of the surface, away from the singularity, showed the surface to remain a broad plateau over increasingly very large values of η_2 .

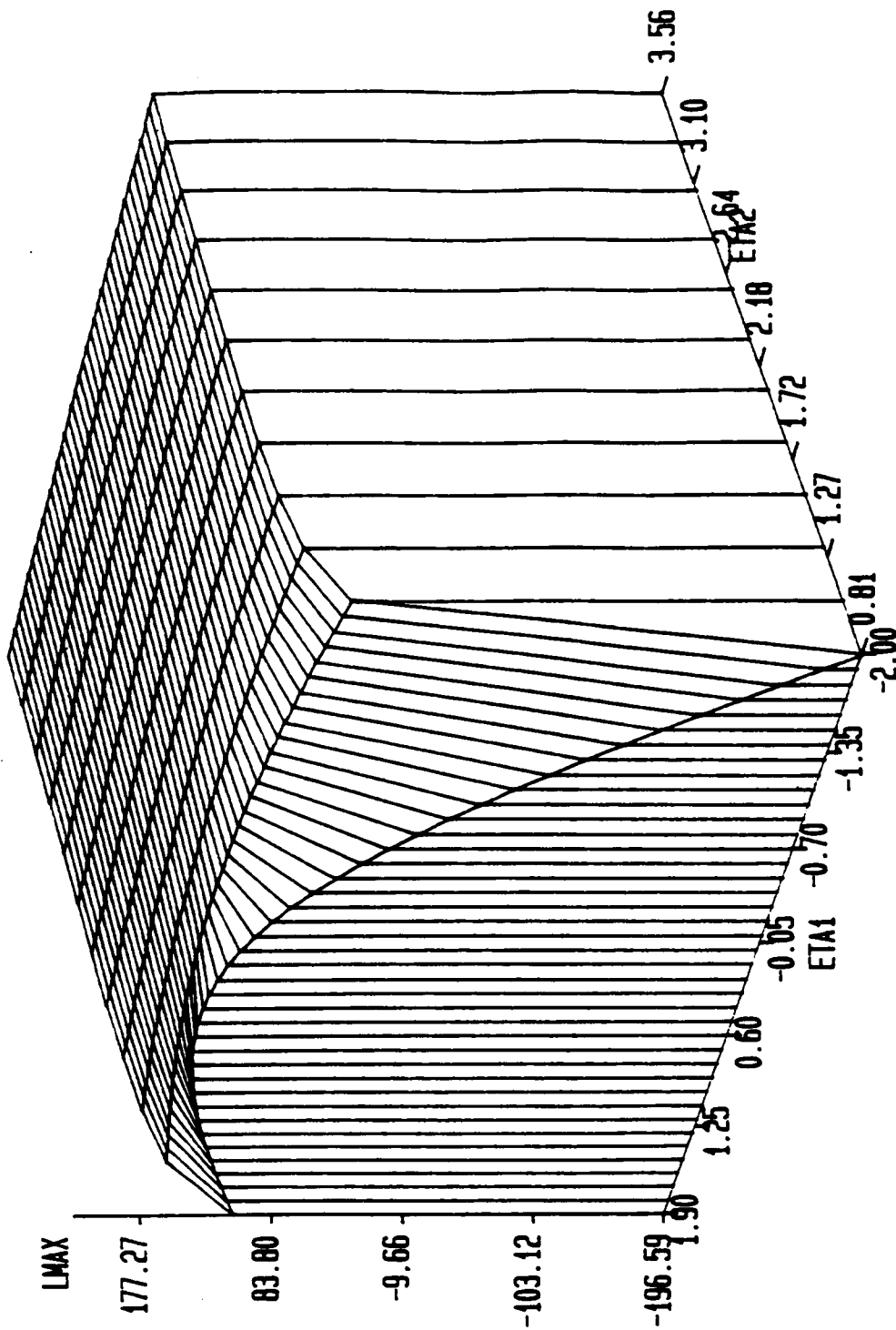
Based upon the analysis of the 10 cases it was decided that η_2 be set as close to the singularity as possible and that η_1 could be set to a predetermined value for all iterations of the simulation. The alternative would have been to re-optimize on each repetition of the simulation which would have been computationally infeasible and would result in the comparison of estimates from different transformations of the sample data.

Figure V.4 is a plot of a typical transformed sample of random phase observations used in the simulation. The power parameter was $\eta_1 = 1/3$ and $\eta_2 = \min_i(y(\omega_i)) + \delta$, for the transformation. It appears that the transformed values can be adequately modelled by a simple linear model, which is one of the goals of the Box-Cox transformation. It also appears that the transformation has not been completely successful in stabilizing variance. However, the sample phase values have a very large variance at the higher frequencies and it would be difficult to completely stabilize the variance over the entire range of frequencies with a Box-Cox transformation. In any case, it was felt that the stabilization of the variance from the Box-Cox transformation was sufficient for the present purposes; and the variance was assumed constant for the transformed phase observations. A more complicated transformation scheme to completely stabilize the variance would also result in more complicated inverse transformation procedure to generate the UMVUE's in the original



LMAX FOR 2 PARAMETER BOX-COX TRANSFORMATION

Figure V.2



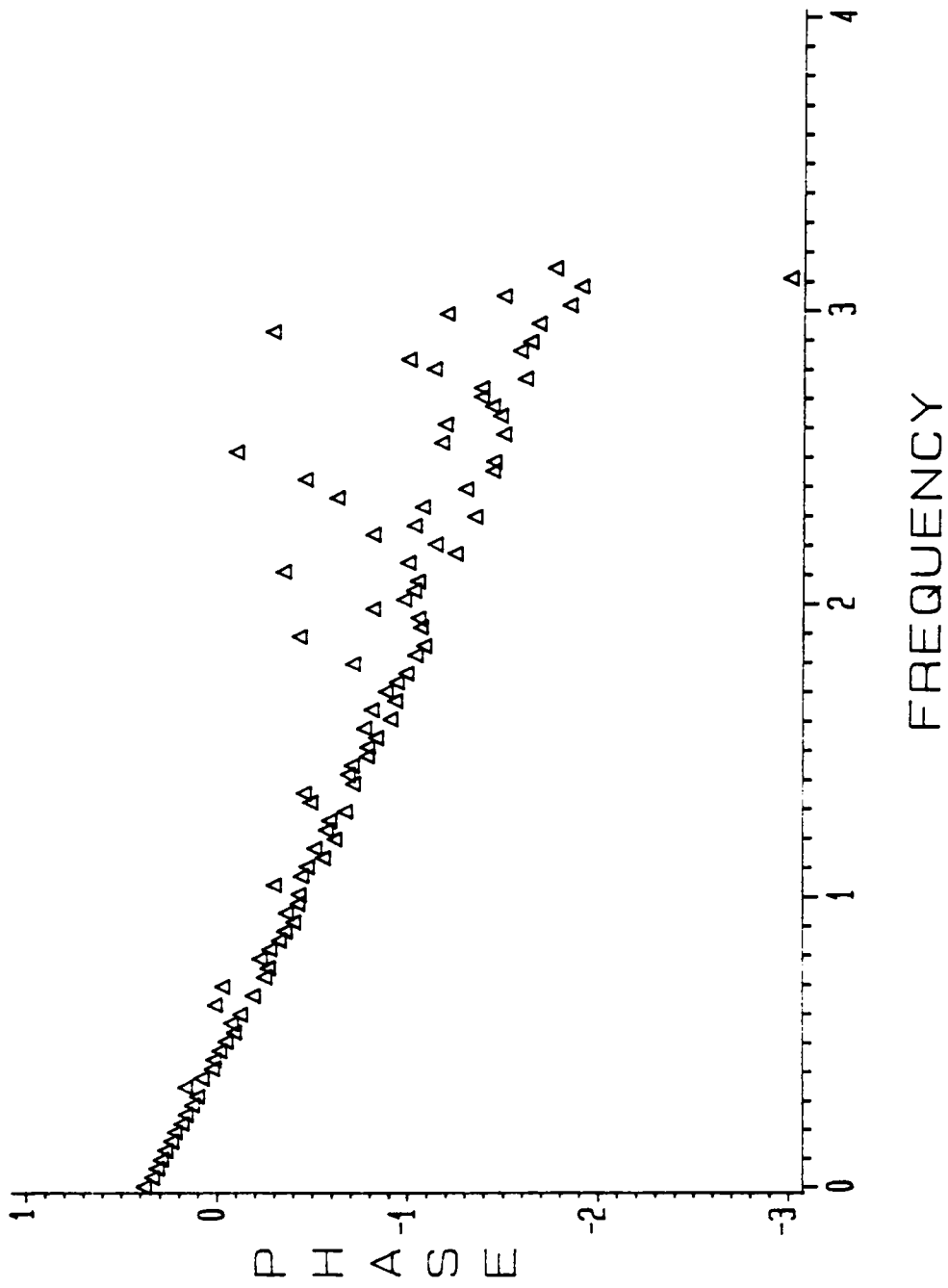
LMAX FOR 2 PARAMETER BOX-COX TRANSFORMATION

Figure V.3

metric and it was decided that any benefit to further stabilization of the variance would be offset by the increased complexity in the inverse transformation.

An interesting point of the initial study of the 10 cases, was the fact that the natural log transform ($\eta_1 = 0$) was never indicated for any of the 10 cases. In fact, attempts to use the natural log transformation consistently resulted in UMVUE's of phase and group delay which were significantly more biased than those obtained using the cube root or square root power parameter in the Box-Cox transformation. Although one would intuitively expect the natural log transformation to work rather well, it simply did not appear to be the optimal choice of transformation in the Box-Cox method. Perhaps this might be explained by the well known result that chi-squared random variables (properly normalized by their degrees of freedom) can be transformed to approximate normality by the square root and the cube root transformation (see Hoyle, 1973). Since the error function in the simulation has a highly skewed distribution it is possible that the cube root or the square root power parameter were the best choice to generate approximate normality for the transformed observations. Recall from chapter III that one of the primary goals of the Box-Cox transformation is to generate normally distributed transformed observations.

Although the plot of transformed values in figure V.4 indicates that a simple linear model may be adequate to model the relationship between transformed phase and frequency, it is of interest to verify that a higher order model (in terms of frequency) would not provide a better fit to the data. Figure V.5 is a plot of the RSTUDENT residuals after fitting a simple linear model to the transformed phase and frequency as plotted in figure V.4. There does not appear to be any compelling evidence in the residual plot that a higher order term is needed in the model; although the heterogeneity in variance does appear at the plot of the residuals for the higher frequencies. To



PLOT OF TRANSFORMED PHASE $\eta_{A1} = .3333$ AND $T = 1.0$

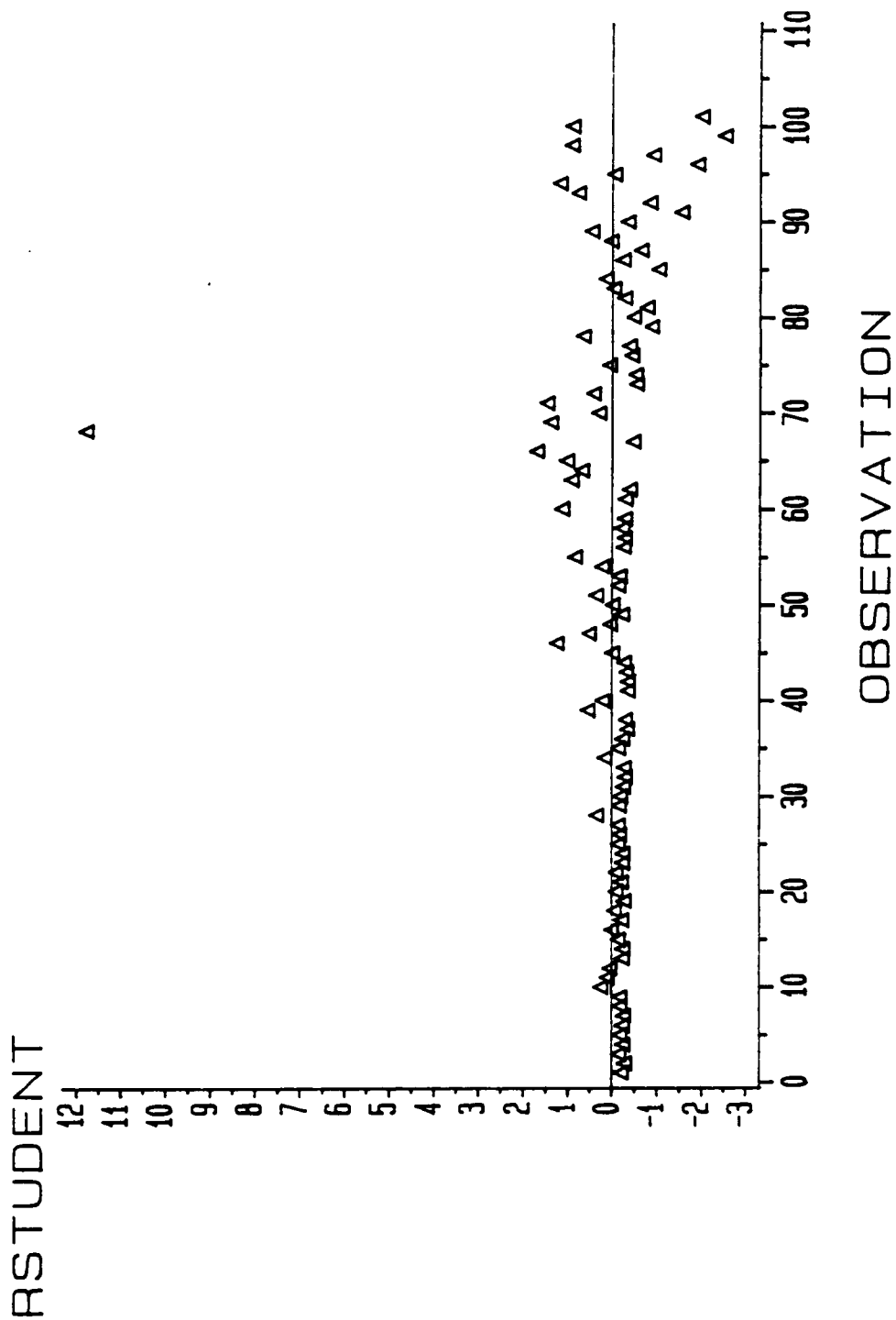
Figure V.4

further verify that higher terms were not necessary in the model of transformed phase and frequency, quadratic terms were included in the model relating the two sets of variables, for each of the 10 test cases, in every case the quadratic term was nonsignificant and only served to increase the prediction variance. Therefore, it was determined that a simple linear model was adequate for modelling the relationship between transformed phase and frequency; and this was the assumption made during the conduct of the simulation study.

V.4. UMVU Estimation Of Phase And Group Delay.

In this section we specifically study the UMVU estimation of expected sample phase and expected sample group delay, using the methods of section IV.4. Specifically, equations [35] and [60] of sections IV.4.1 and IV.4.2, respectively, were used to generate UMVU estimates. The finite difference method of calculating group delay, discussed in section IV.4.3, is also considered during this aspect of the simulation study. It was found that 1,000 repetitions of the simulation, for each set of test conditions, was sufficient to establish the properties of the estimators. It was found that the difference in observed properties changed slightly in going from 100 to 1,000 repetitions. However, in going from 1,000 to 2,000 repetitions the observed properties were virtually identical; and it was decided that 1,000 repetitions was sufficient to investigate the properties of the proposed estimators of section IV.4.

The transformation method of generating confidence intervals, discussed in section IV.5.1, was also evaluated during the conduct of the simulation. The present investigation was simple in scope



RSTUDENT RESIDUALS FOR TRANSFORMED PHASE T=1, ETA1=.3333

Figure V.5

and only intended to produce a guideline as to the actual coverage probabilities for the theoretical $(1-\alpha)100\%$ confidence intervals. The approach was to calculate a 95% confidence interval for the expected sample phase on the first iteration of the simulation. On the subsequent 999 iterations, a running count was kept of how many expected sample phase estimates fell into the calculated interval; the count was kept for each of the 101 Fourier frequencies. The final result is an observed coverage probability for the interval at each frequency.

V.4.1. The Results Of The Study For The Cube Root Case.

We first consider the UMVU estimation of expected sample phase and expected sample group delay using a Box-Cox transformation with a cube root power parameter. The transformation is

$$g(\omega_i) = \frac{(y(\omega_i) + \eta_2)^{\frac{1}{3}} - 1}{\frac{1}{3}} \quad [12]$$

and the inverse transformation becomes

$$y(\omega_i) = \left[\frac{1}{3} g(\omega_i) + 1 \right]^3 - \eta_2 = f(g(\omega_i)) - \eta_2. \quad [13]$$

It is $f(g(\omega_i))$ that is expanded in a Taylor series to achieve our UMVUE's of expected sample phase and expected sample group delay. The exact form of the expansions, given in equations [35] and [60] of Chapter IV, for the cube root Box-Cox transformation are (using the notation of section IV.4)

$$\hat{\theta}(\omega_i) = 1.0 + \hat{g}(\omega_i) + \frac{1}{3} [\hat{g}(\omega_i)]^2 + \frac{1}{3} s^2(1 - h_{ii}) + \frac{1}{9} \hat{g}(\omega_i)s^2(1 - h_{ii}) + \frac{1}{27} [\hat{g}(\omega_i)]^3 - \eta_2 \quad [14]$$

and for [60]

$$\begin{aligned} \hat{\theta}'(\omega_i) = \hat{\beta}_1 + \frac{2}{3} [\hat{\beta}_1 \hat{g}(\omega_i)] - \frac{1}{3} s^2 \left[\frac{d}{d\omega} h_{ii} \right] + \frac{1}{9} \hat{\beta}_1 s^2 (1 - h_{ii}) \\ - \frac{1}{9} s^2 \left[\frac{d}{d\omega} h_{ii} \right] \hat{g}(\omega_i) + \frac{1}{9} \hat{\beta}_1 [\hat{g}(\omega_i)]^2. \end{aligned} \quad [15]$$

The UMVUE for phase is generated by [14] and that for group delay by [15].

Before actually conducting the simulation study, an investigation was made of the naive or simple inverse transformation method of generating estimates on the original scale of measurement. The predicted values $\hat{g}(\omega_i)$ were substituted into [13] for the $g(\omega_i)$ to generate the estimates of $Ey(\omega_i)$ in the original metric. Recall from the discussion in section IV.2, that this method does generate unbiased estimates of the median in the original scale of measurement, but not necessarily the mean of the distribution. The purpose of this exercise was to see if the naive inverse transformation was actually adequate for estimation of the mean and whether the additional complication of the bias correction methods of [14] and [15] was actually required. Figure V.6 represents the average UMVUE and naive estimates, for 1,000 repetitions of the simulation, of the underlying phase function generated by [7]. The solid line in the plot of figure V.6 is the actual underlying phase function and the dashed line lying about the actual phase curve is the average UMVUE function. The average naive transformation estimates lie on the dashed line lying well below the actual phase function. The increasing distance between the naive estimation curve and the actual phase curve, with increasing frequency, can be explained by the fact that the sample phase function

is highly skewed and of increasing dispersion as frequency increases; i.e., the mean and the median of the distribution are moving farther apart as the frequency increases. It was concluded from this exercise that the transformation bias can indeed be significant and the additional computation required to compensate for the bias is justified.

For the 1,000 repetitions of the simulation, the average expected sample phase estimate was calculated at each of the 101 Fourier frequencies. Since we have not made distributional assumptions about our estimators in [14] and [15], we must approximate the bias and variance of the UMVU estimators by sample estimates; i.e., let $\hat{\theta}_m(\omega_i) = (1/n) \sum_{j=1}^n \hat{\theta}_j(\omega_i)$, then

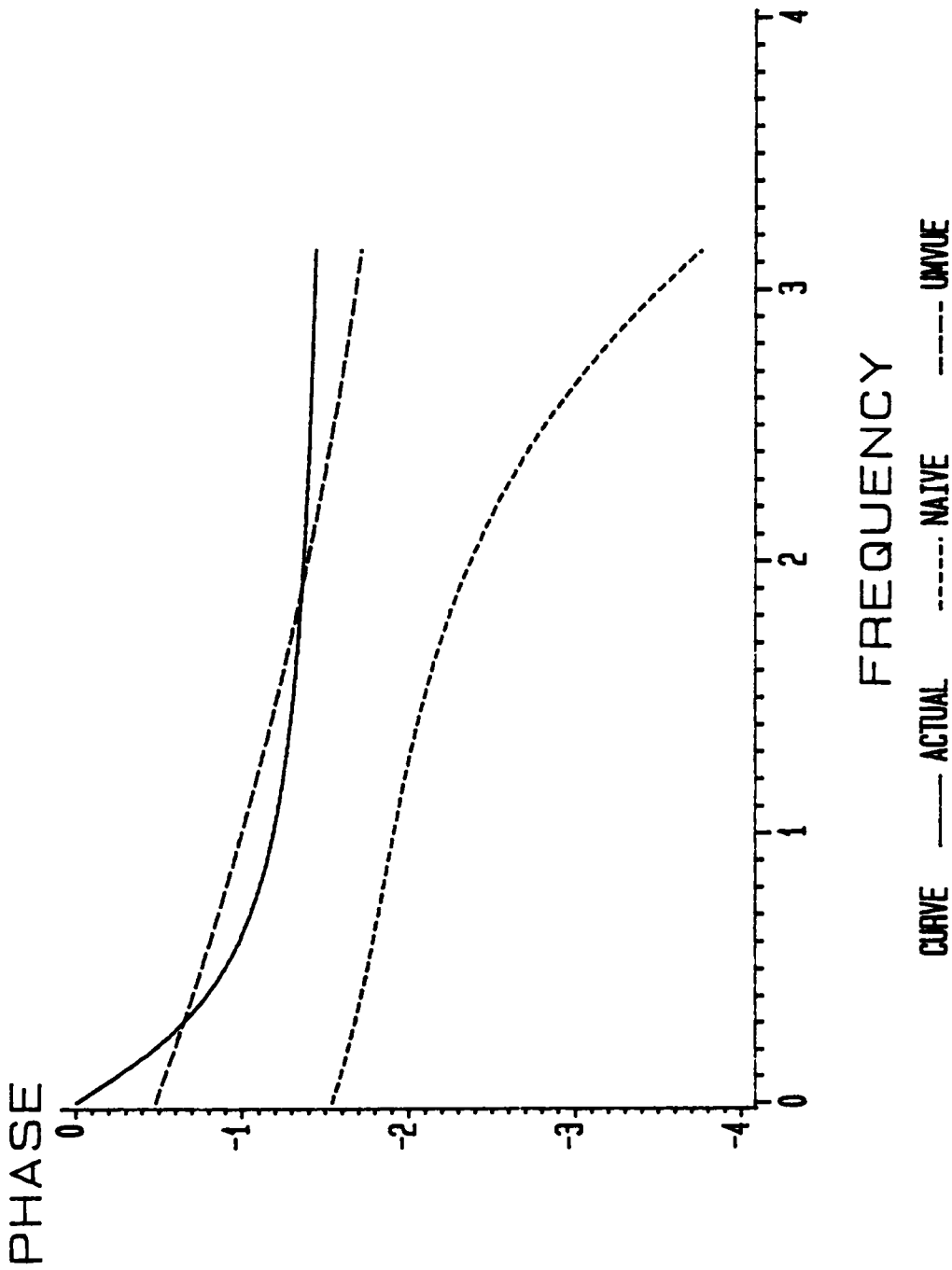
$$\hat{\text{Var}}(\hat{\theta}(\omega_i)) = \frac{1}{n-1} \sum_{j=1}^{101} (\hat{\theta}_j(\omega_i) - \hat{\theta}_m(\omega_i))^2 \quad [16]$$

and the bias estimate

$$\hat{\text{Bias}}(\hat{\theta}(\omega_i)) = \hat{\theta}_m(\omega_i) - \phi_{xy}(\omega_i) \quad [17]$$

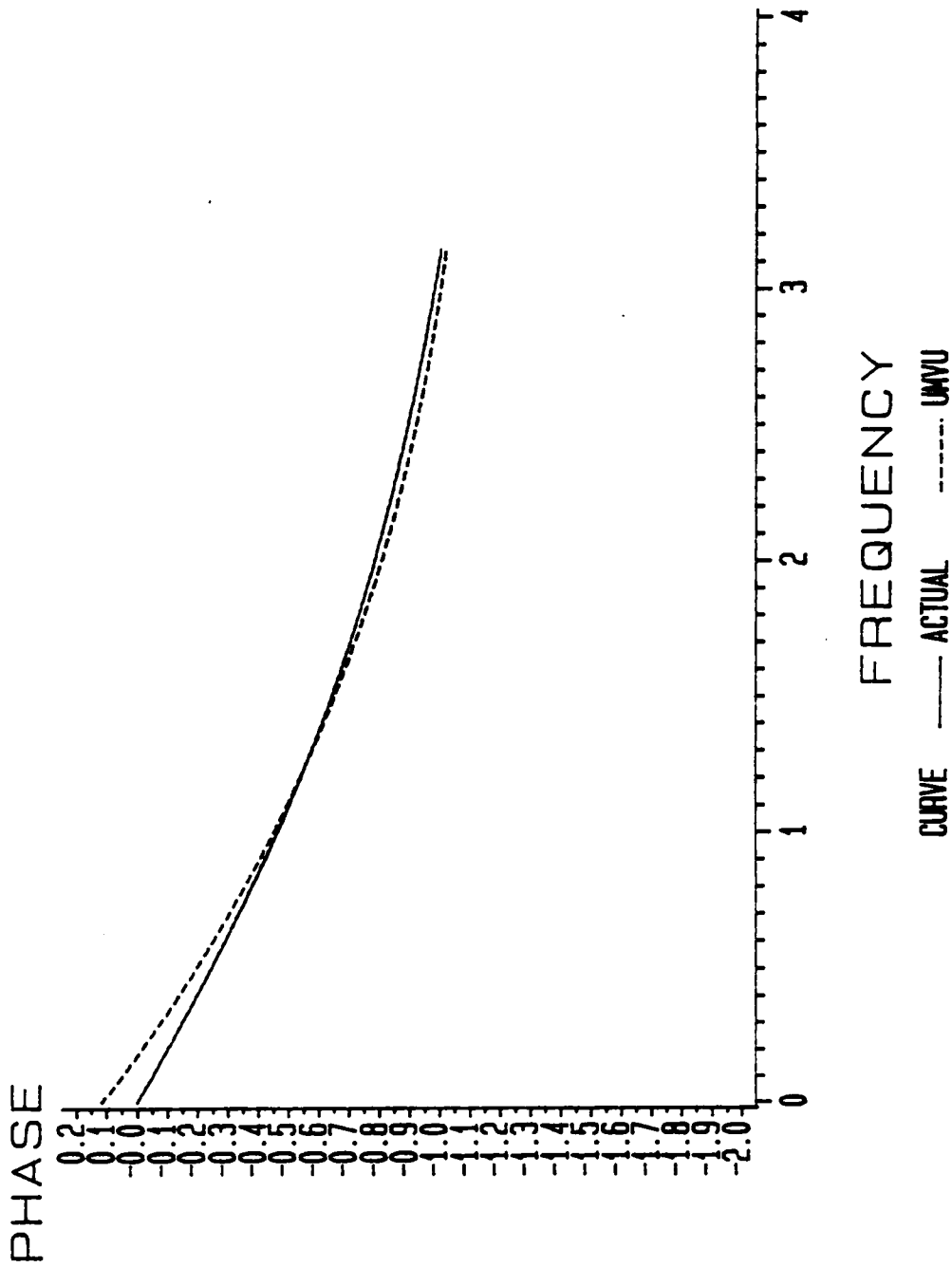
where $\phi_{xy}(\omega_i)$ is the known phase value generated from [7]. Figure V.7 is a plot of the average estimates for each of the Fourier frequencies, for the 1,000 repetitions. In the plot of figure V.7 some sample bias is present for the phase estimates, particularly at the low frequencies. However the overall approximation to the actual phase function appears reasonable.

In a similar manner, average expected sample group delay estimates were calculated at each of the Fourier frequencies, for the 1,000 repetitions. The same calculations were also performed for the finite difference method of computing expected sample group delay estimates. Furthermore,



UMVUE AND NAIVE ESTIMATE OF PHASE $T=1$ $\text{ETA1}=1/3$ $\text{NREP}=1000$

Figure V.6



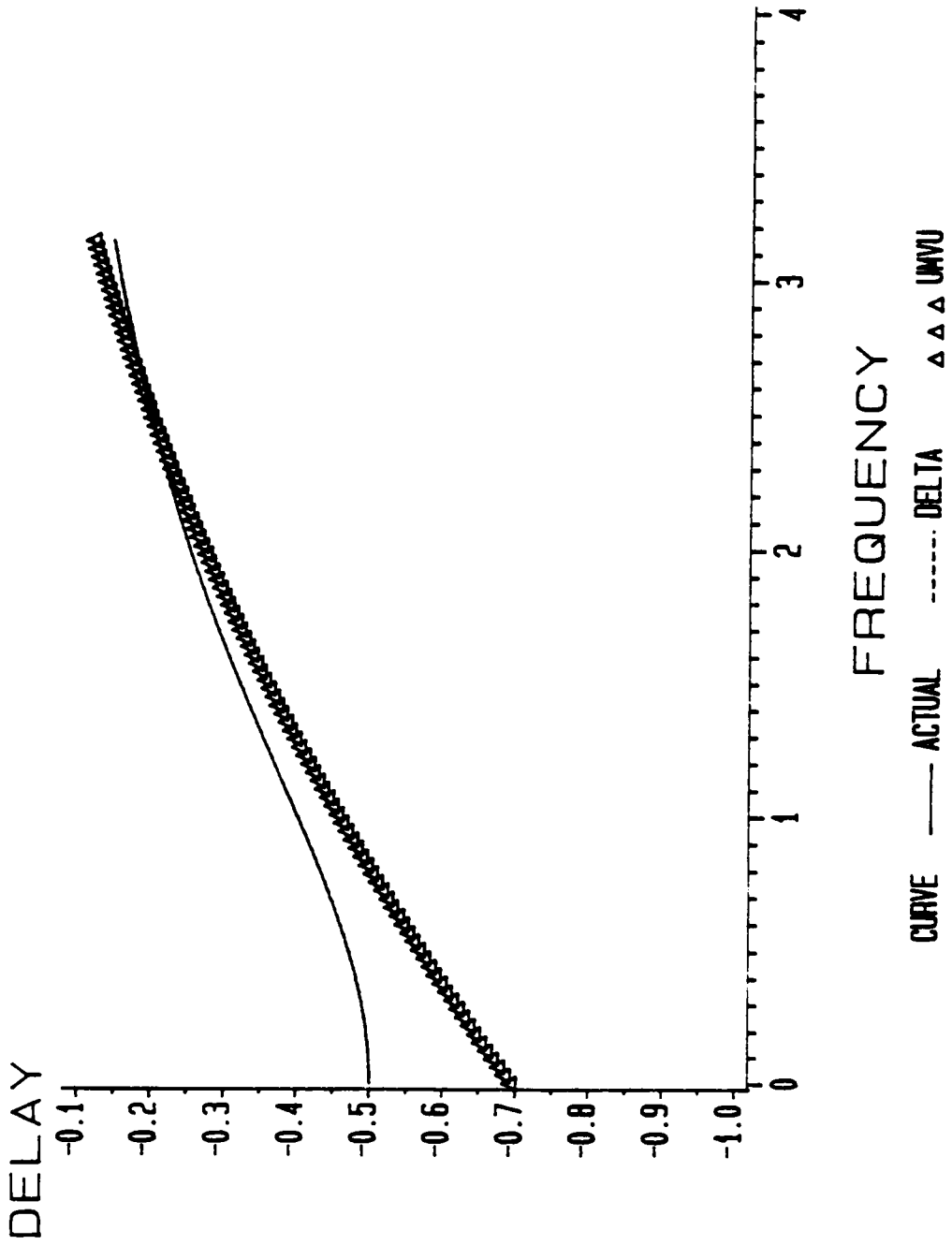
PLOT OF ACTUAL AND UMVUE PHASE FOR $T=1$ $\text{ETA1}=0.3333$, $\text{NREP}=1000$

Figure V.7

sample estimates of the estimator variance and bias were calculated using the analogous forms of equations [16] and [17] for both the UMVUE method of [15] and the finite difference method. Figure V.8 contains plots of the average estimate at each frequency for the UMVUE method and the finite difference method. The plots of the finite difference and UMVUE estimates were virtually coincident and the finite difference plot is obscured in figure V.8.

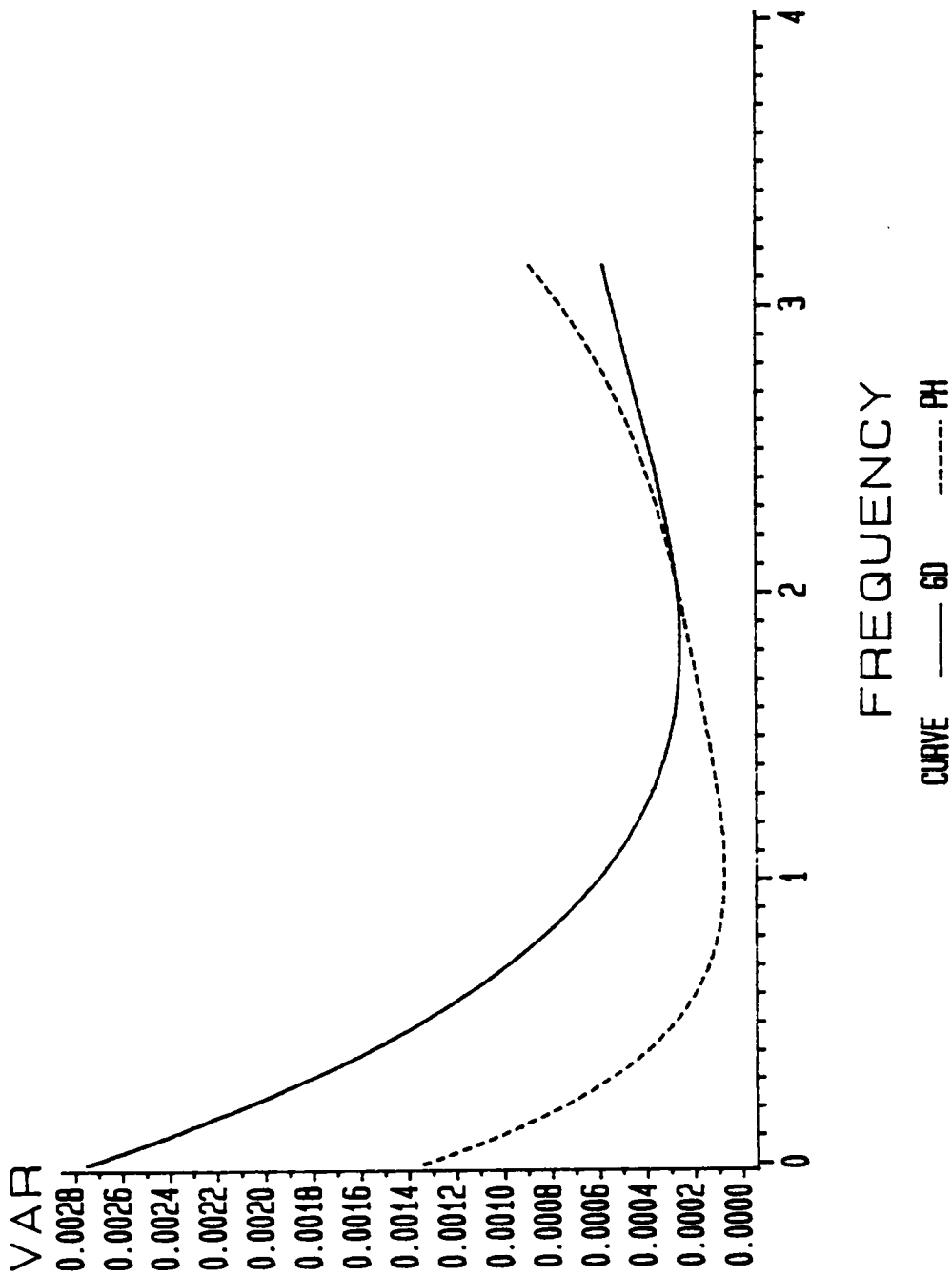
From the plots in figure V.8 it is again evident that both methods of calculating estimates of the actual group delay function have noticeable sample bias. However, in the case of both estimation procedures the magnitude of the sample bias is acceptable for the size of the group delay values at those frequencies where the sample bias is greatest. In fact we must qualify the UMVUE property of [14] and [15] as only approximate, for the actual underlying phase and group delay, since we have obvious sample bias for our average UMVU estimates. However, the actual level of the bias observed in figures V.7 and V.8 is not at a significant enough level to preclude using the UMVUE methods to estimate the expected sample phase and group delay. In the simulation, the actual values and the expected sample values are coincident, but in finite samples we can not explicitly state this equivalence relationship for expected sample values and actual parameter values. Finite sample sizes required to assume the equivalence of the expected sample values and actual parameter values are unknown and this subject requires further investigation. Although, from section II.3 equation [13], we know that the sample estimate of the phase is asymptotically an unbiased estimate of the actual phase parameter.

The sample variances, computed by [16], for the UMVUE of expected sample phase and the expected sample group delay, are plotted in figure V.9. If one compares the the observed sample variances for these estimators with the actual phase and group delay values (which are also the ex-



ACTUAL, UMVUE, AND FINITE GD FOR $T=1$ $\text{ETA1}=0.3333$, $\text{NREP}=1000$

Figure V.8



COMPARISON OF VARIANCE FOR PHASE AND GD T=1 ETA1=0.3333

Figure V.9

pected sample values) shown in figures V.7 and V.8, respectively, then it can be seen that the observed sample variances are uniformly smaller than the actual values. We conclude from this that, under the conditions of the simulation, the observed variances of the two estimators are acceptable for estimation of expected sample phase and group delay. The observed standard errors of the estimates can be determined from the sample MSE plot (for estimated phase and group delay) in figure V.10. If one compare these observed standard error values with the actual phase and group delay values shown on figures V.7 and V.8, then again we see that the magnitude of the standard errors is uniformly smaller than the actual values at each Fourier frequency. Although this is a qualitative judgement, it does appear that, under the conditions of the simulation, the observed error in the UMVUE estimators of [14] and [15] is acceptable for spectral analysis.

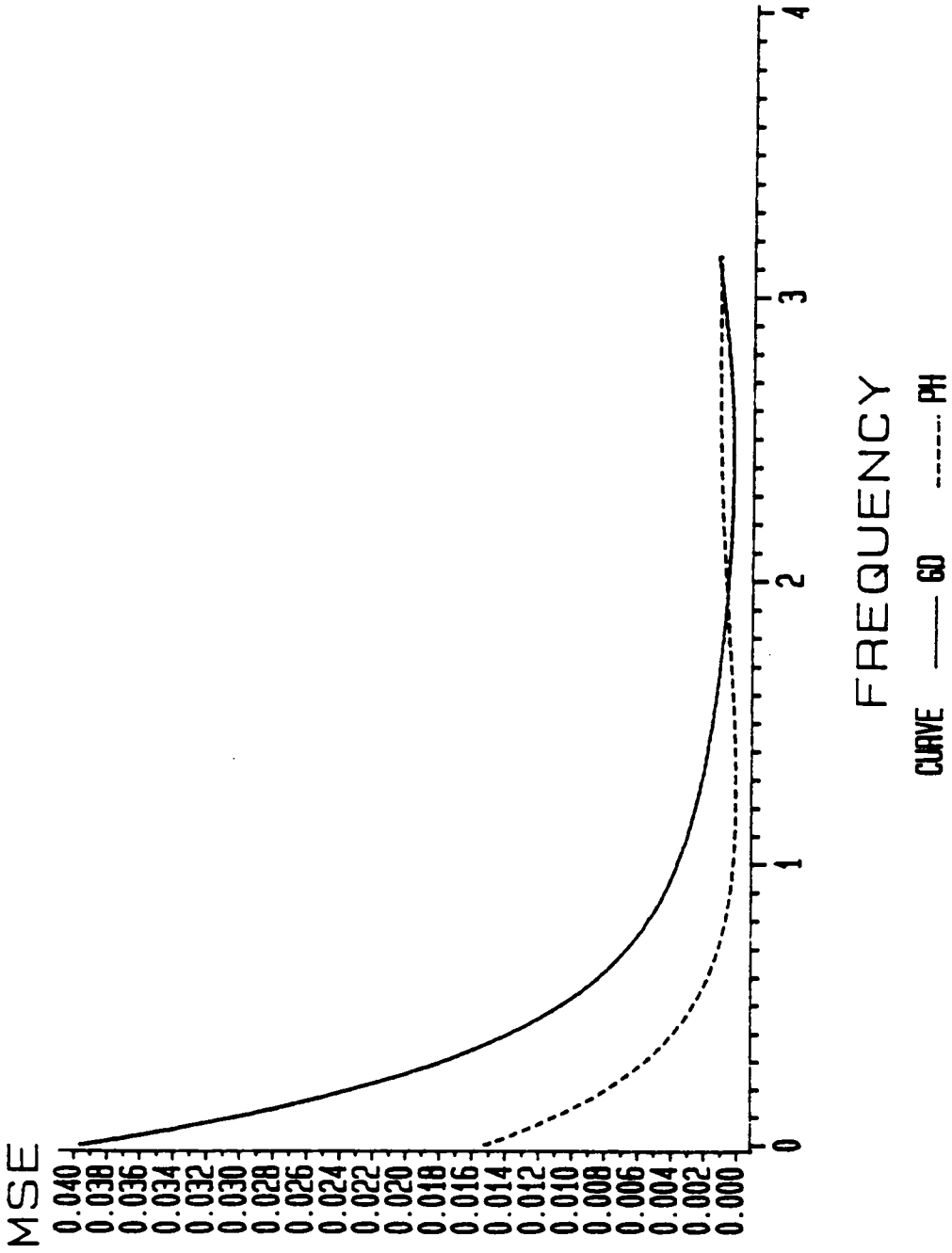
A comparison of the observed error for the direct UMVUE method of calculating expected sample group delay and finite difference method of calculating expected sample group delay estimates was conducted. In figures V.11 and V.12 we have a comparison of the sample variance and sample MSE, respectively, for 1,000 repetitions of the simulation, for each of the two calculation methods. Examining the sets of plots in figures V.11 and V.12, it can be seen that the performance of the two methods of estimating group delay were virtually identical. Since we have already concluded that the UMVUE method of calculating expected sample group delay has exhibited acceptable levels of error, under the conditions of the simulation, we must then conclude that the finite difference calculation method has also performed acceptably in calculating estimates of expected sample group delay; and it is concluded that the finite difference method of calculation is a viable alternative to the UMVUE method of calculation [15].

Figure V.13 is plot of a typical estimate of the expected sample phase function for one iteration of the simulation. The solid curve represents the UMVUE of the expected sample phase function while the dashed curves represent the upper and lower 95% confidence limits, calculated by the transformation method of section IV.5.1. Since very little is known about the performance of this method, the subject is clearly one for future research.

The observed coverage probabilities for the theoretical 95% confidence interval for the expected sample phase function were 1.0 for all the 101 Fourier frequencies. It would appear that the inverse transformation method generates conservative confidence intervals in the present application. However, as previously mentioned, we do not have enough knowledge of the performance of this method in actual application to extrapolate this simulated result to real problems. In other words it appears to generate conservative interval estimate, but we do not know how it performs under general conditions. Also, with very limited historical work on confidence interval methods for inverse transformations, the topic requires a lot of future attention.

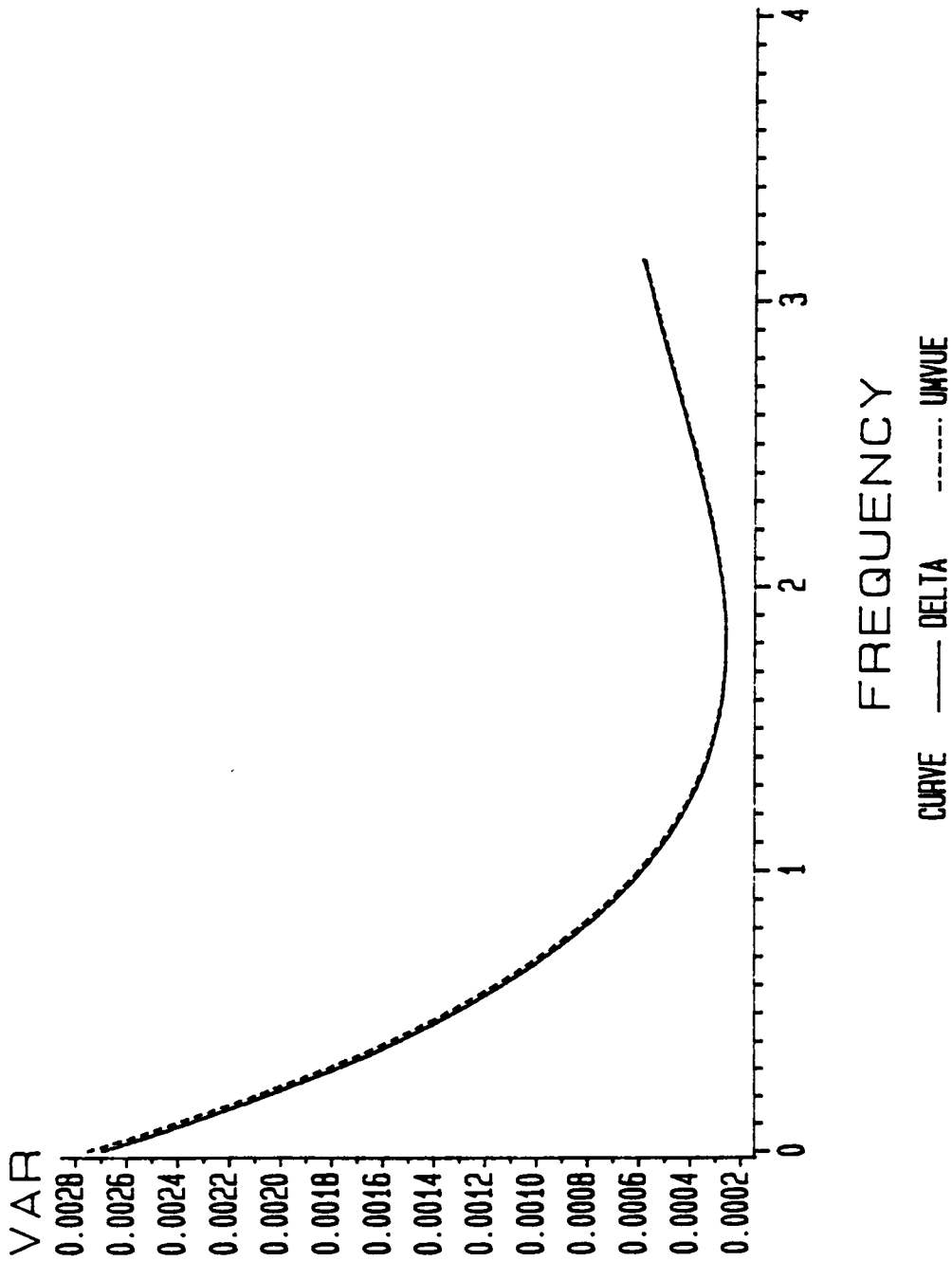
V.4.2. Results Of The Study For Square Root Case.

In this section we consider a simulation study of the UMVUE methods with a Box-Cox transformation using a square root power parameter. It was of interest to see if a different, but valid, change in the power parameter would produce significant changes in the UMVU estimate (in an MSE sense). Once again the finite difference method of calculating expected sample group delay was also examined.



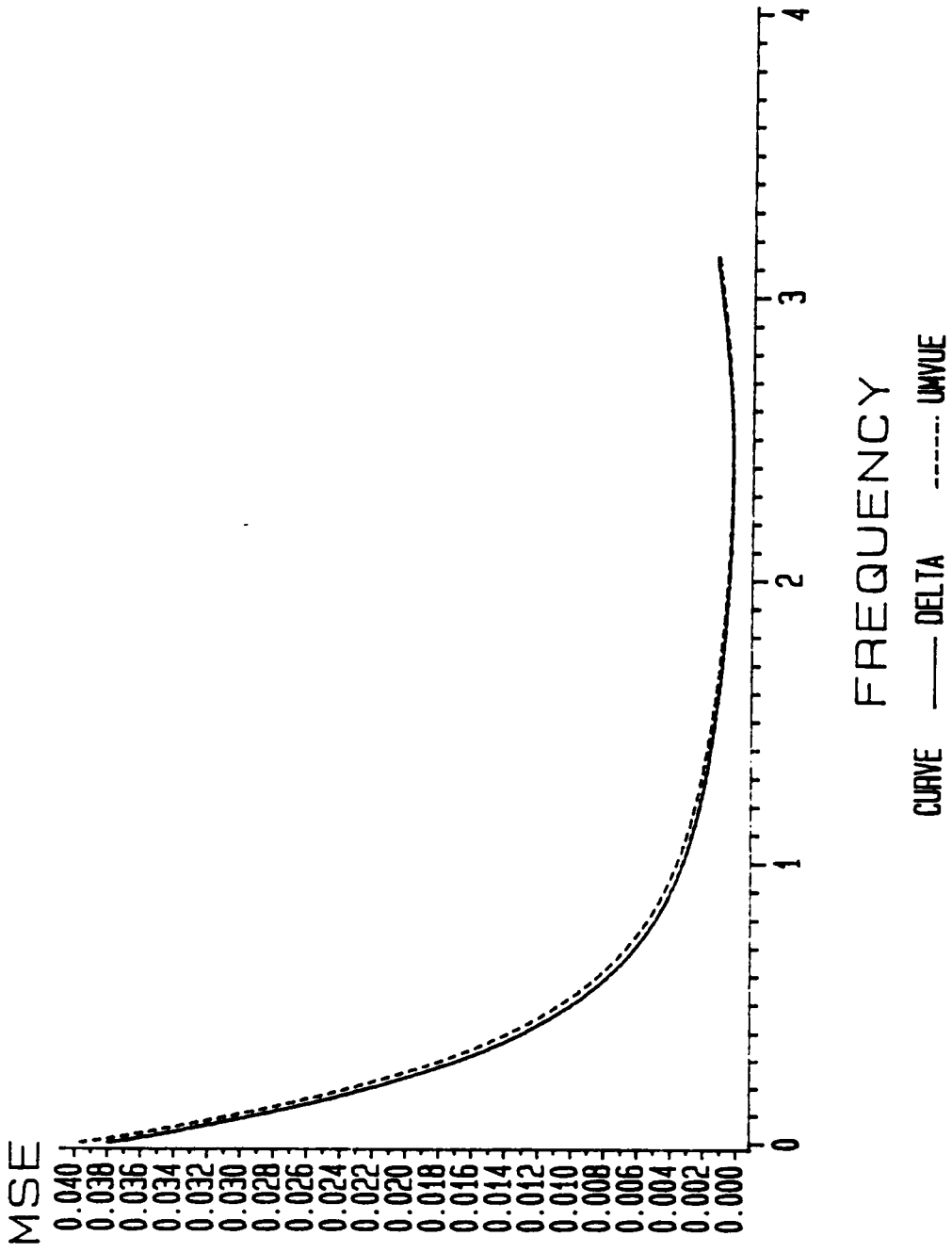
COMPARISON OF MSE FOR PHASE AND GD T=1 ETA1=0.3333

Figure V.10



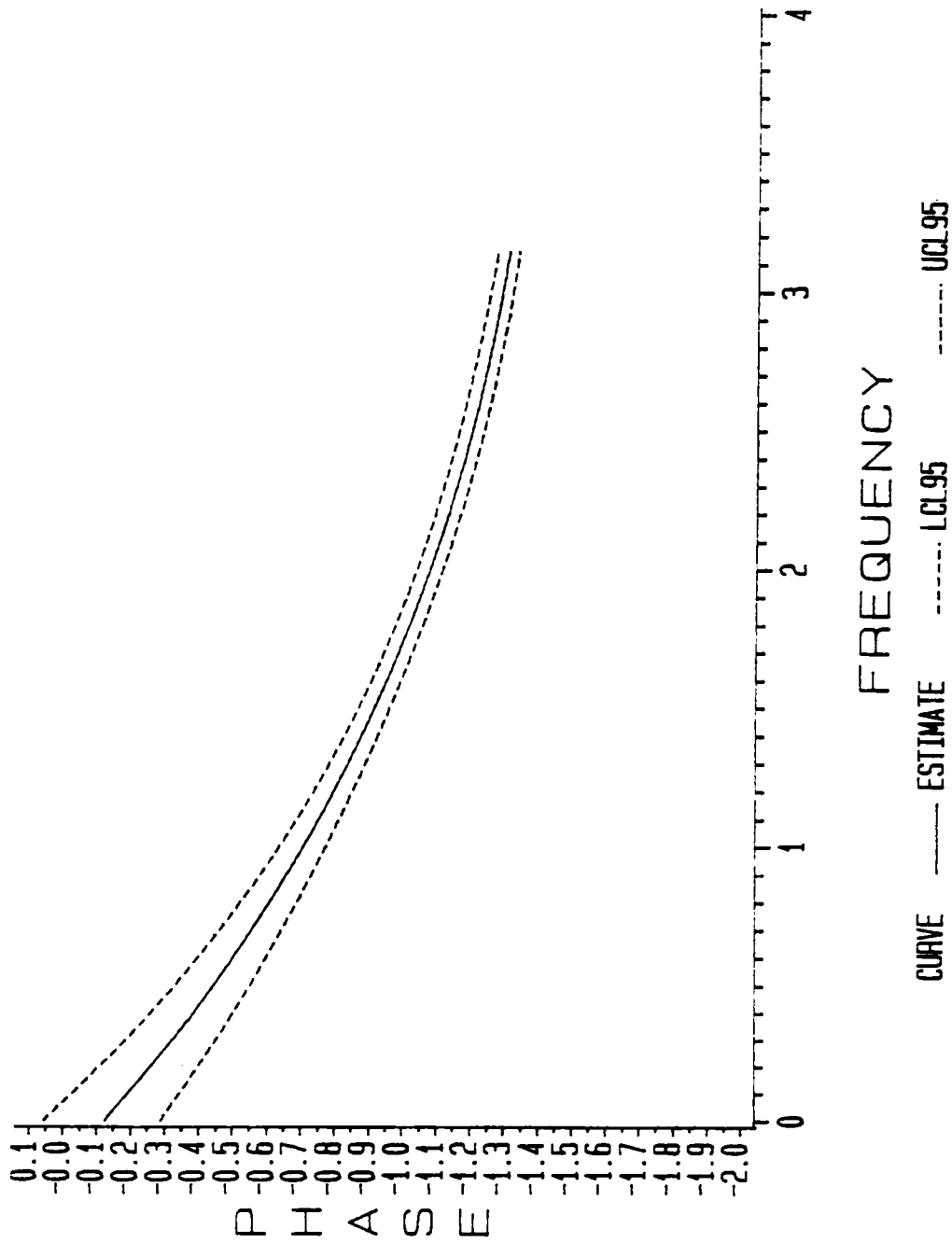
COMPARISON OF VARIANCE FOR UMVUE AND DELTA GO T=1 ETA1=0.3333

Figure V.11



COMPARISON OF MSE FOR UMVUE AND DELTA 6D T=1 ETA1=0.3333

Figure V.12



PLOT OF ESTIMATED PHASE AND 95% C.I., $\eta_1 = .3333$ AND $T = 1.0$

Figure V.13

The Box-Cox transformation for this case is

$$g(\omega_i) = \frac{(y(\omega_i) + \eta_2)^{\frac{1}{2}} - 1}{\frac{1}{2}} \quad [18]$$

and the inverse transformation is

$$y(\omega_i) = \left[\frac{1}{2} g(\omega_i) + 1 \right]^2 - \eta_2 = f(g(\omega_i)) - \eta_2. \quad [19]$$

As in the previous section it is $f(g(\omega_i))$ that we wish to expand in a Taylor series to generate UMVUE's of expected sample phase and group delay. The exact form of the expansions are

$$\hat{\theta}(\omega_i) = 1.0 + \hat{g}(\omega_i) + \frac{1}{4} [\hat{g}(\omega_i)]^2 + \frac{1}{4} s^2(1 - h_{ii}) - \eta_2 \quad [20]$$

and for group delay

$$\hat{\theta}'(\omega_i) = \hat{\beta}_1 + \frac{1}{2} \hat{\beta}_1 \hat{g}(\omega_i) - \frac{1}{4} s^2 \left(\frac{d}{d\omega} h_{ii} \right). \quad [21]$$

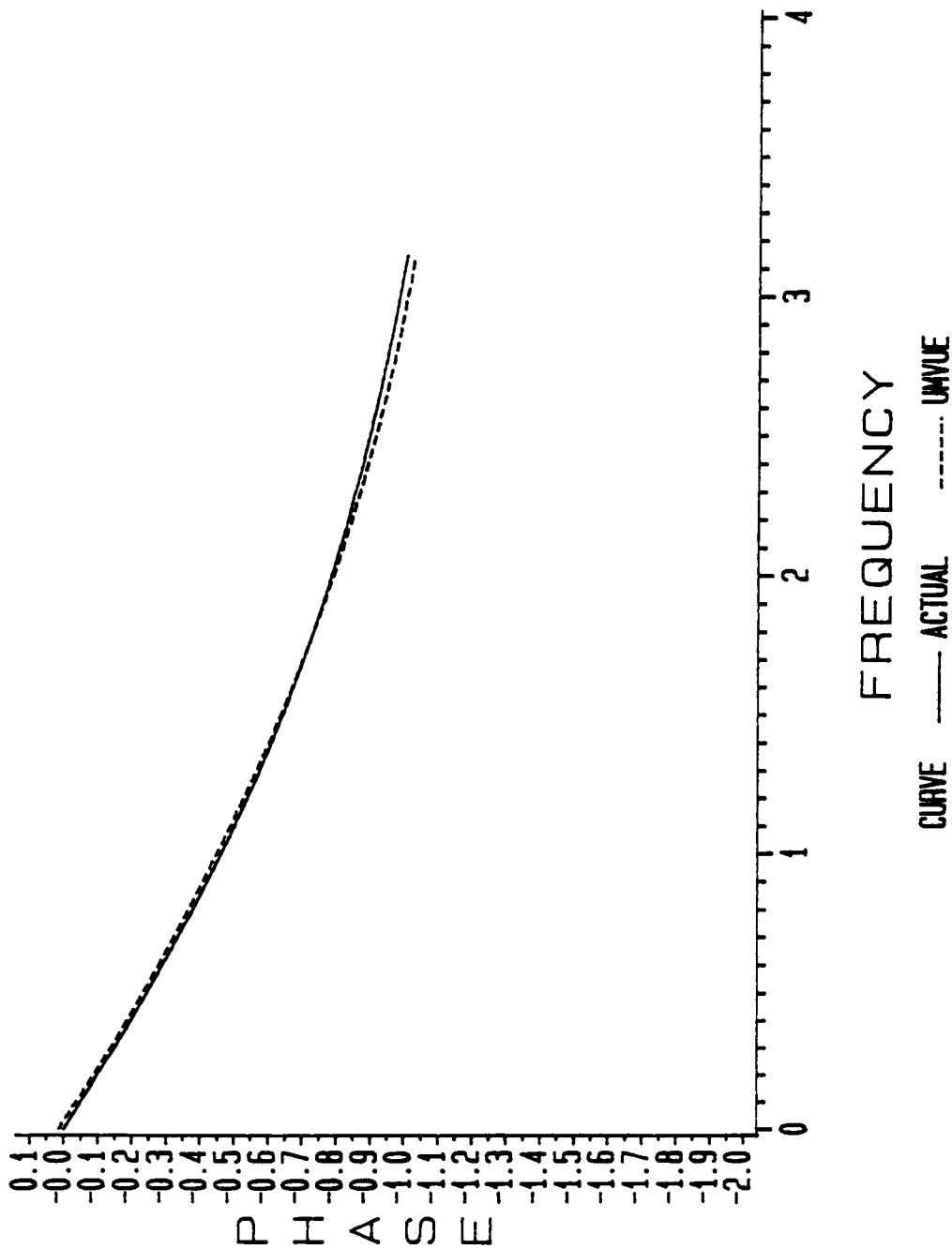
Comparing [20] and [21] with [14] and [15], it can be seen that the square root parameter case has the advantage of requiring fewer parameter estimates in the Taylor series expansions for $\hat{\theta}(\omega_i)$ and $\hat{\theta}'(\omega_i)$.

As in section V.5.1, 1,000 repetitions of the simulation were performed and sample estimates of variance and bias were computed for the UMVUE method for phase, the UMVUE method for group delay, and the finite difference method for group delay. The average expected sample phase

estimate at each Fourier frequency is plotted in figure V.14. The overall approximation to the actual phase function (which is equivalent to the expected sample phase function) appears to be quite good with some sample bias apparent at the higher frequencies. Therefore, we once again conclude that under the conditions of the simulation, the UMVUE property is an approximate one. Similarly, the average expected sample group delay estimates, both by the UMVUE and finite difference methods of calculation, are plotted in figure V.15. Once again the UMVUE plot and the finite difference plot are coincident and therefore the finite difference plot is obscured. We can see from figure V.14 that the nonlinear group delay function is being approximated by a straight line. However, we could have determined this from [21] since it contains no second order terms.

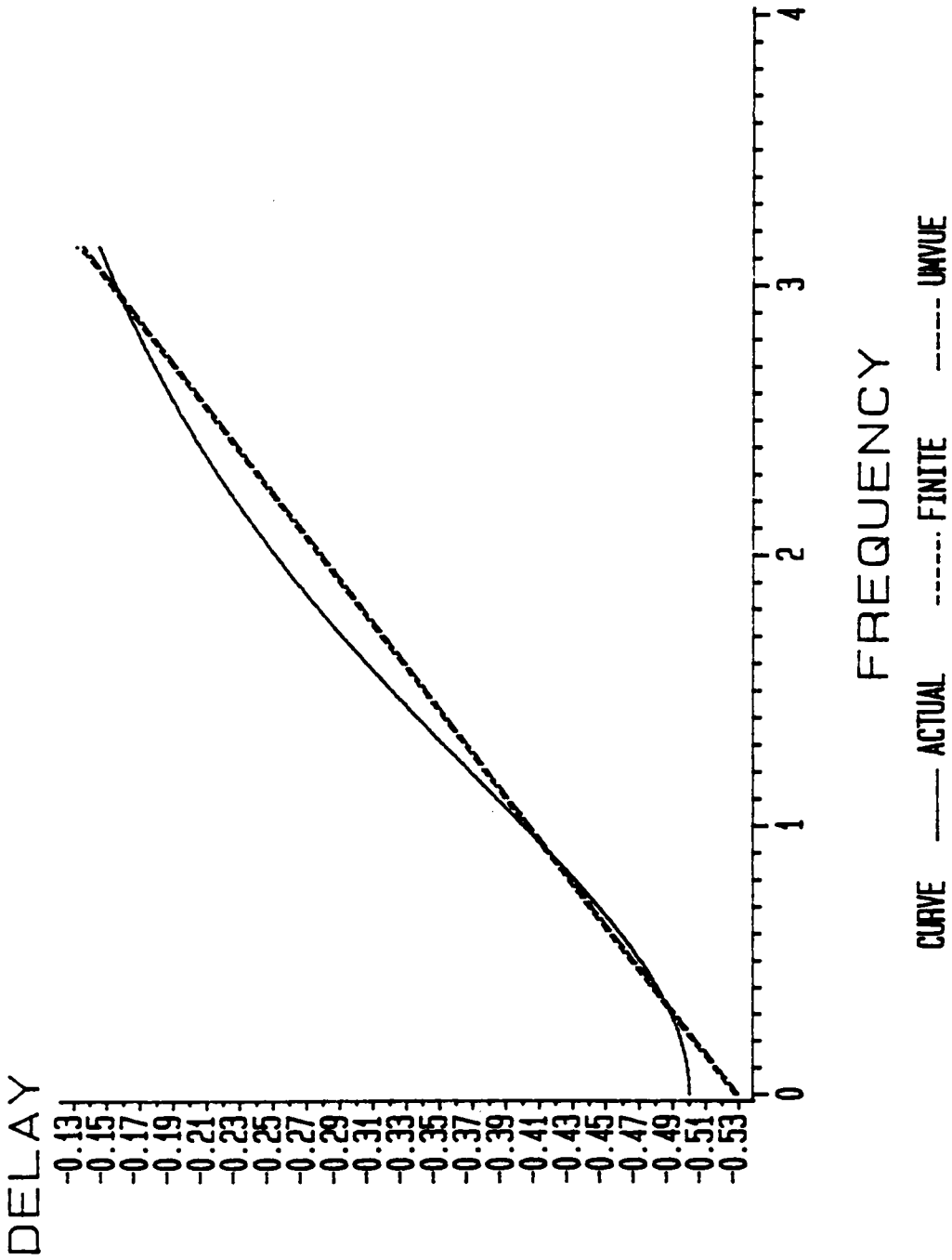
The effect of approximating the nonlinear group delay function by a linear function is graphically illustrated in figure V.16. Here, we have a plot of the sample bias for the UMVUE and the finite difference expected sample group delay estimates. The sinusoidal shape of the sample bias curves is explained by the fact that both estimates cross two inflection points on the actual group delay curve. The sample variances for both group delay estimates are plotted in figure V.17 and we see that the sample variances are virtually identical for the two expected sample group delay estimates. If we compare the magnitude of the sample variances in figure V.17 with those in figure V.11 ($\eta_1 = 1/3$), then we observe that the group delay estimates for the $\eta_1 = 1/2$ case exhibit uniformly smaller variances than for the cube root parameter case. Of course this was expected, since [20] and [21] have fewer parameters to estimate than [14] and [15].

We can further compare the performance of the 2 group delay estimates in the square root case against their performance in the cube root case, by examining the MSE plots in figures V.12 and V.18. In figure V.18 the observed sample bias over the midrange frequencies, as noted in figure



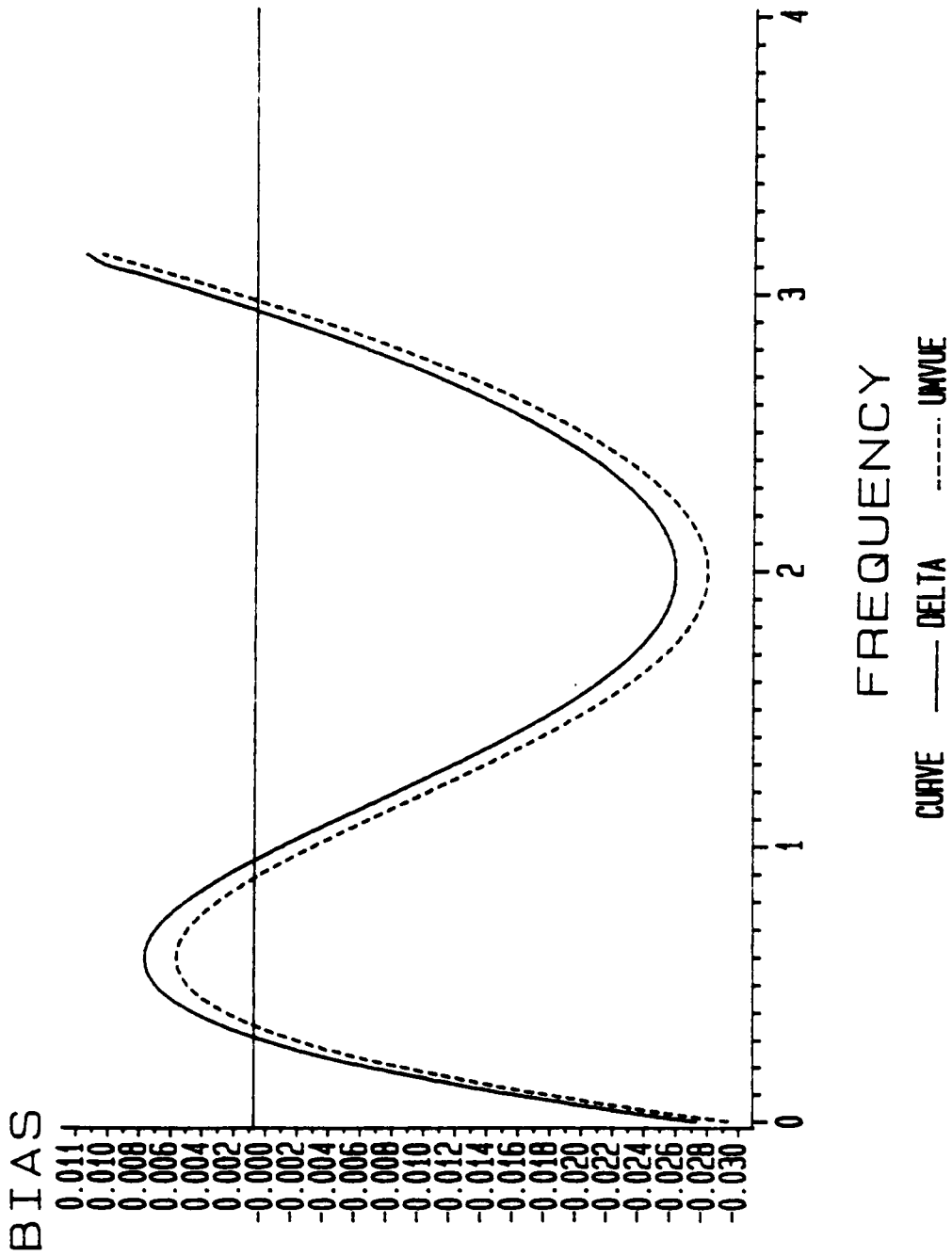
UMVUE AND ACTUAL PHASE FOR NREP=1000 T=1 ETA1=0.5000

Figure V.14



UMVUE, FINITE AND ACTUAL GD FOR NREP=1000 T=1 ETA1=0.5000

Figure V.15

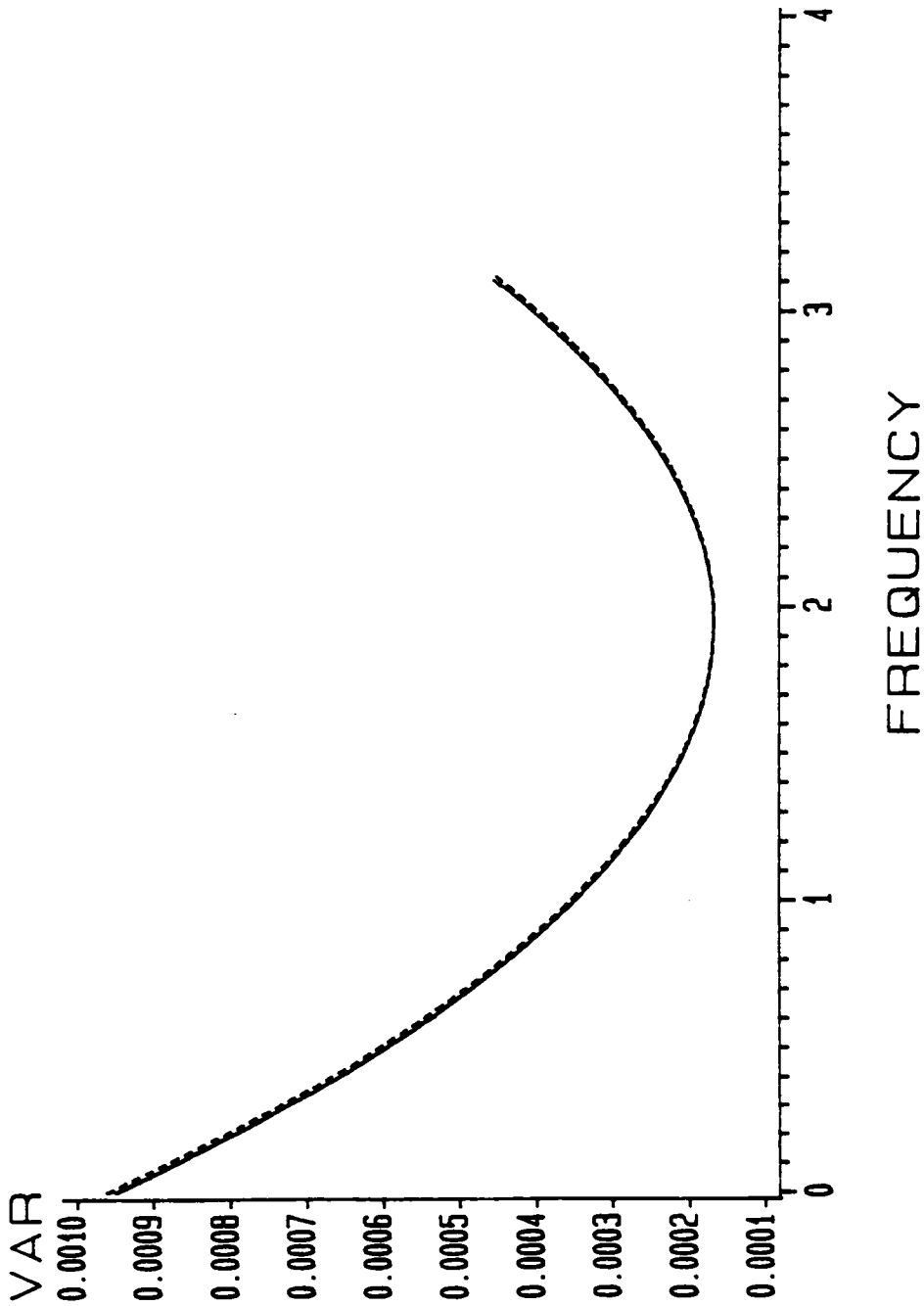


COMPARISON OF BIAS FOR UMVUE AND DELTA $\tau=1$ $\eta_1=0.50$

Figure V.16

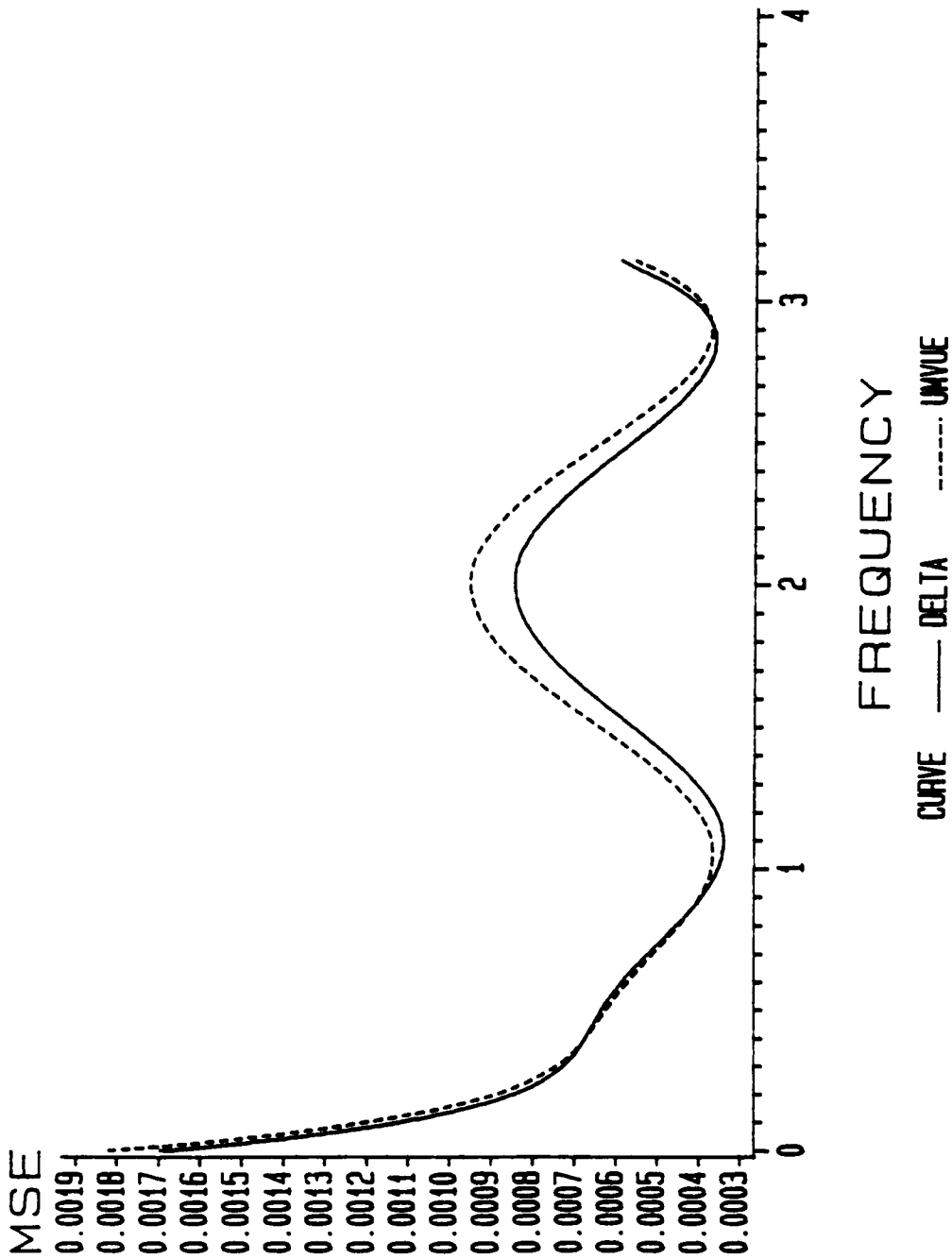
V.15, is quite noticeable. However, the overall MSE values for the square root case are lower than those for the cube root case. In other words, any increase in bias that was incurred by the use of a square root transformation was more than compensated for by the reduction in variance for the simpler inverse transformation. It would appear that the selection of the power parameter in the Box-Cox transformation must take into account the classic tradeoffs between bias and variance to generate the smaller mean square error for our estimates. A basic strategy for the parameter selection, however, can not be stated on just the basis of the present simulation. This is clearly another topic for future research.

A similar result for the UMVUE of the expected sample phase function, i.e., equation [20], can be observed by comparing the performance (in an MSE sense) for the cube root transformation case and the square root transformation case. As noted in the previous paragraph, [20] has fewer parameters to estimate than [14] which results in lower sample variances for the phase estimates generated under the square root transformation case. We can verify this by comparing the variance plots in figures V.19 with V.9 (cube root case). The sample variances are uniformly lower for the square root phase estimates than for the cube root phase estimates. We can also gain some insight into the two choices of transformation parameter by comparing the sample MSE plots in figures V.20 and V.10. The MSE values for the UMVUE phase and group delay estimates are uniformly lower under the square root transformation case (despite the increased bias in the group delay estimates) than their counterparts under the cube root transformation case. The conclusion might be made that one should always select the simplest, in terms of the inverse transformation, value for η_1 . However, such a conclusion can not be made without much further study on the topic.

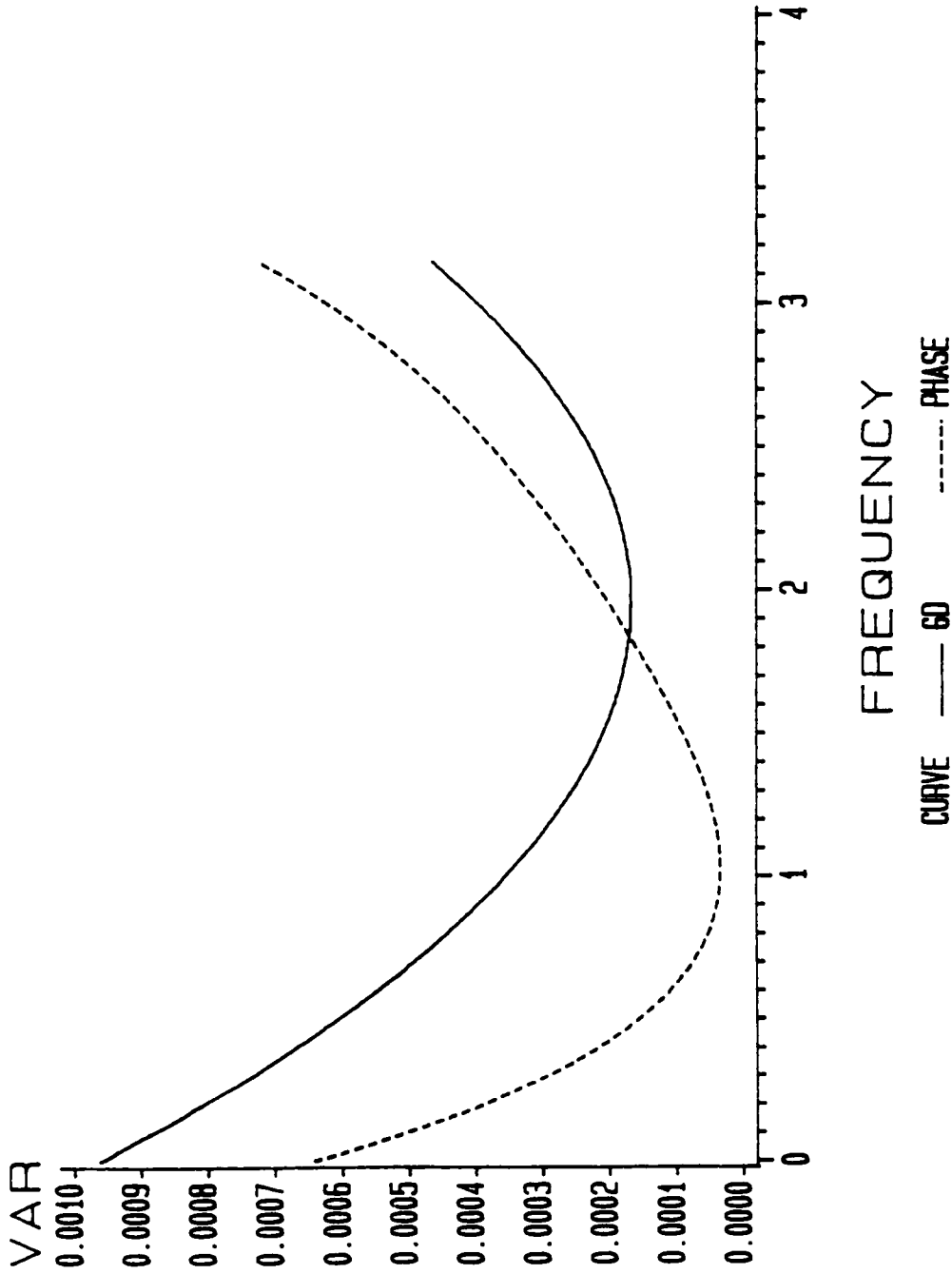


COMPARISON OF VARIANCE FOR UMVUE AND DELTA 60 T=1 ETA1=0.50

Figure V.17

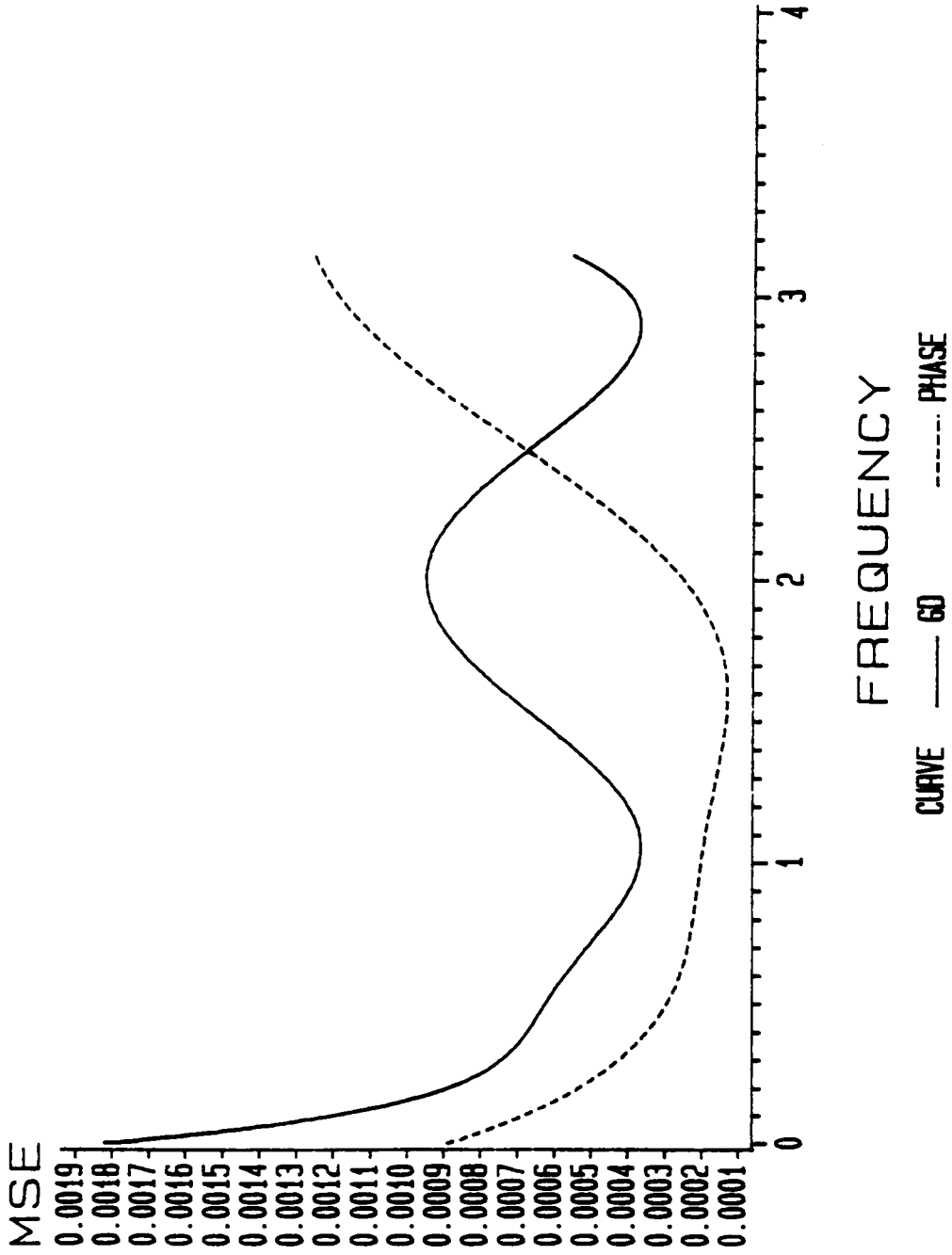


COMPARISON OF MSE FOR UMVUE AND DELTA $\tau=1$ $\eta_1=0.50$
 Figure V.18



COMPARISON OF VARIANCE FOR PHASE AND GD T=1 ETA1=0.50

Figure V.19



COMPARISON OF MSE FOR PHASE AND 60 T=1 ETA1=0.50

Figure V.20

Chapter VI

Applications Of Group Delay Estimation.

VI.1. Introduction.

Chapter IV is concerned with developing theoretical procedures to generate UMVUE's of expected sample phase and group delay. Chapter V is a demonstration of the methods of Chapter IV by means of a simulation study. The present chapter is an application of the methods of Chapter IV to the estimation of expected sample phase and group delay between naturally occurring time series. The primary focus of the chapter is the estimation of expected sample group delay, since phase is measured in dimensionless radian units and its interpretation in naturally occurring time series is difficult (Hannan and Thomson, 1973). On the other hand, since group delay is measured in units of time its physical interpretation is obvious.

Two different pairs of univariate time series are examined in this chapter. The first set of time series consist of the annual harvest of Maine clawed lobsters (*H. Americanus*) and the annual sea water temperature as measured at Boothbay Harbor, Maine; the two data sets cover the years 1897 through 1984. The second set of time series consist of the annual per capita consumption of alcoholic beverages by residents of London and the annual Wolfer sunspot numbers for the years 1870 through 1938.

VI.2. Lobster Harvest And Sea Water Temperature.

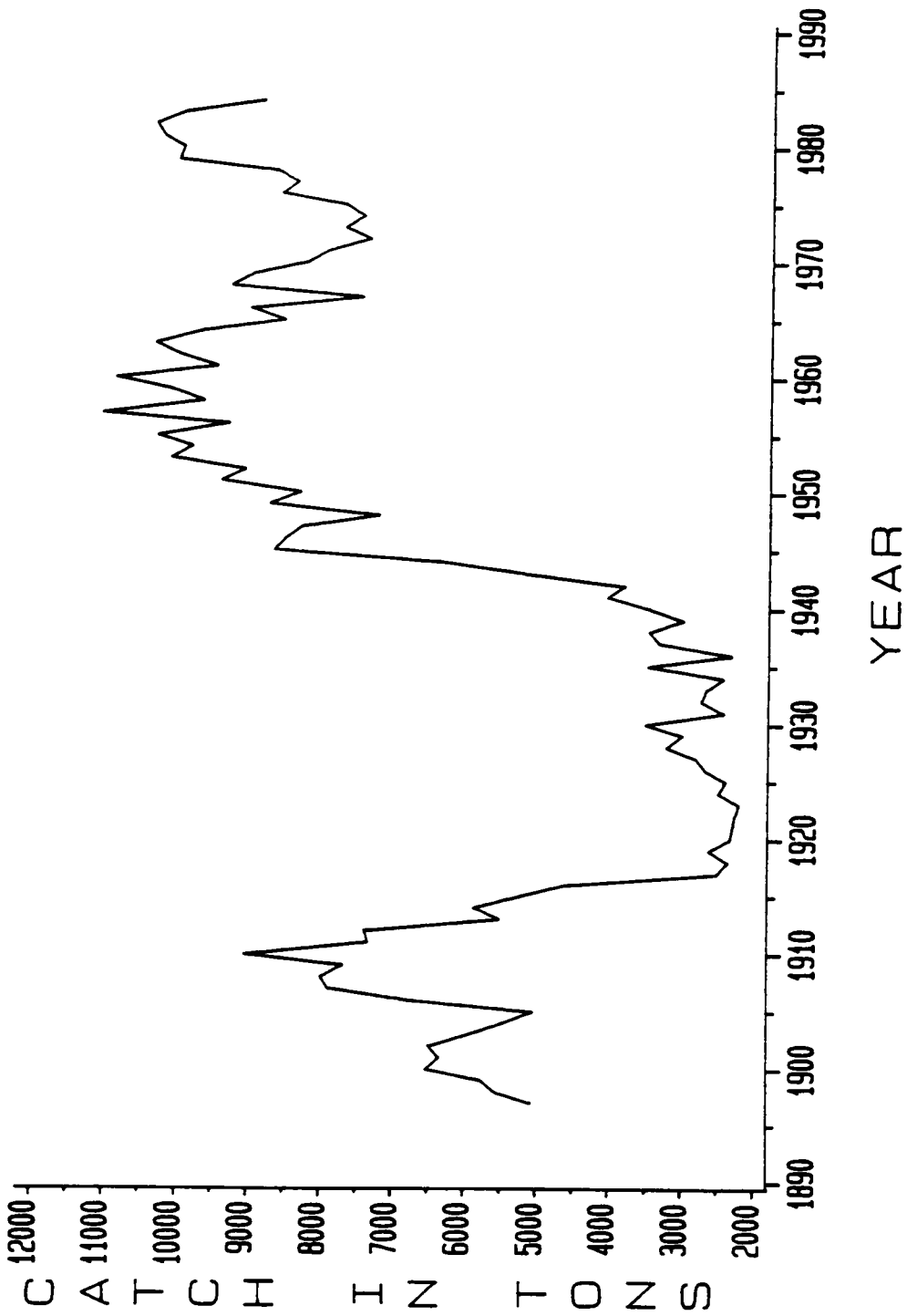
The clawed lobster is a commercially important marine resource which is harvested about the perimeter of the mid-North Atlantic Basin. Three species are of particular commercial importance: (1) *Nephrops norvegicus*, which accounts for 50% of the worldwide harvest, (2) *Homarus americanus*, which accounts for 46% of the world harvest, and (3) *Homarus gammarus*, which accounts for about 4% of the total world harvest (Cobb and Phillips, 1980, p. 266). Although no accurate values for the annual harvest are available, it is believed to be in excess of \$250 million in North America and at least \$50 million in the State of Maine alone (Lewis, 1988). Since worldwide lobster markets are poorly organized, exact monetary values for lobsters are not available (Cobb and Phillips, 1980, p. 286).

In North America, the primary source of clawed lobsters is the species *H. americanus* which is fished from Newfoundland in the Canadian Meritimes to the southern terminus of the clawed lobster in North Carolina (Dow, 1978). The single most important source of lobsters in the world

is the coastal waters of Maine which produces 30% of the total North American catch per year (Dow, 1978). Figure VI.1 is a plot of the annual harvest of Maine lobsters, in metric tons, for the years 1897 through 1984. Since the demand for lobster has also been slowly increasing over this time, a linear trend was observed in the catch, which violates stationarity assumptions. Therefore, a simple linear model was fit to the data to remove the trend. Figure VI.2 is a detrended plot of the annual catch for the years 1897 through 1984.

All three of the commercially important species are known to experience natural fluctuations in abundance which appear to be associated with cyclic fluctuations in sea water temperatures (Cobb and Phillips, 1980, p. 304). In reviewing the literature it was noticed that the abundances of the clawed lobsters are primarily measured by recorded landings of the different species and the terms harvest or catch and abundance are used almost interchangeably, so in the present research the two terms are treated as synonymous. Since *H. americanus* is heavily fished in North America and is well known to vary in abundance with sea water temperature, it is very important to understand this relationship and attempt to adjust the level of fishing to the predicted level of abundance (Cobb and Phillips, 1980, p. 294). In fact Cobb and Phillips (1980, p. 304) go on to state that all three commercial lobster fisheries are in serious economic and biological trouble and without better management they may well collapse. It would seem particularly relevant, to the ability to effectively manage the fisheries, if one could precisely characterize the actual relationship between abundance and sea water temperature, which is attempted in the present work by the use of group delay analysis of the time series.

Although group delay analysis can have implications of causality, one must have some physical mechanism for the causal effect before it has real meaning. The exact mechanism by which sea

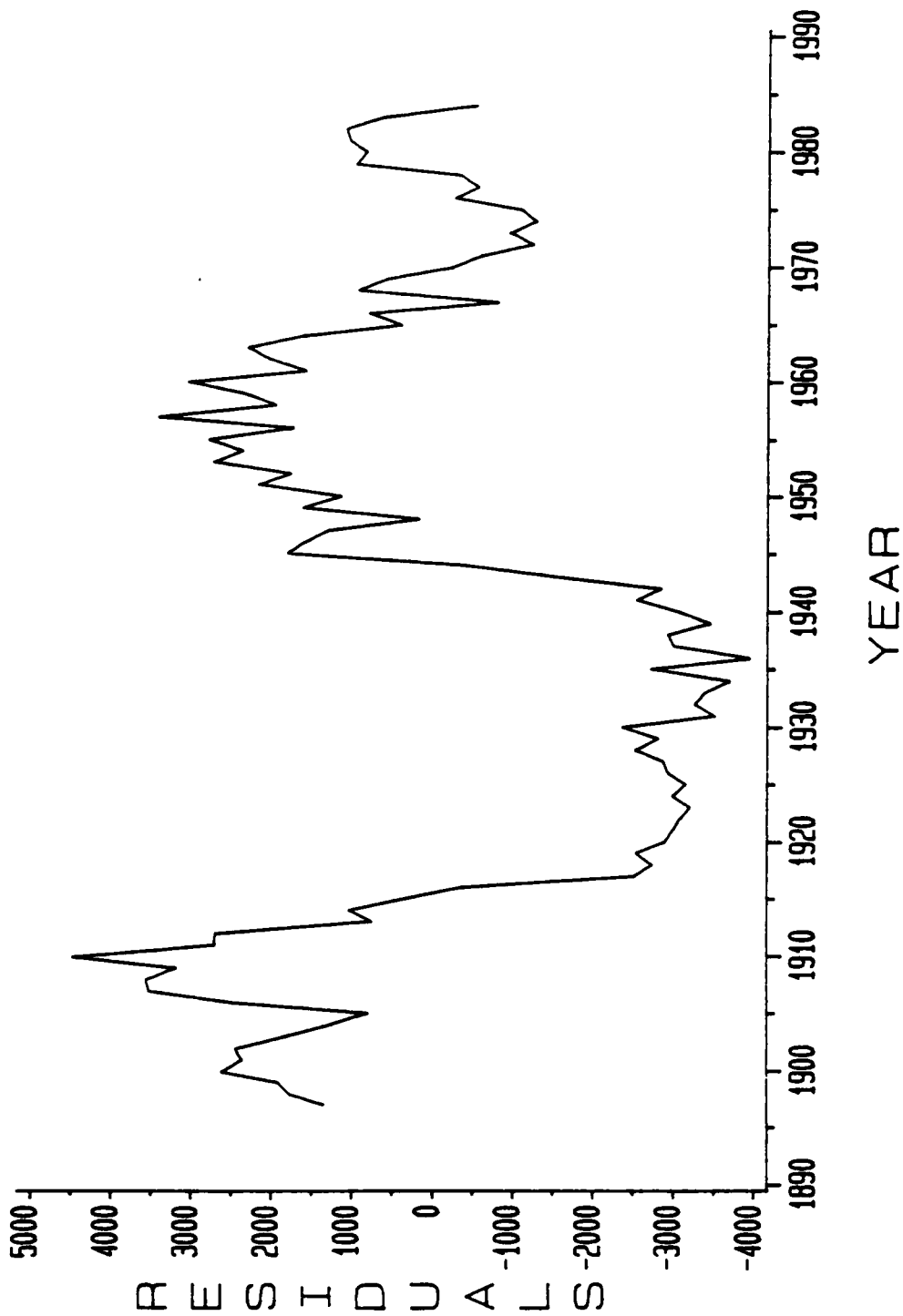


PLOT OF ANNUAL LOBSTER CATCH 1897-1984

Figure VI.1

water temperature influences lobster abundance is not completely understood (Cobb and Phillips, 1980, p. 290). However, the primary sources of the influence are thought to be: (1) the length of time and number of lobsters molting from sublegal to legal size (important in North America where size restrictions are strictly enforced on lobster fishing), and (2) the survival rate for each years class of larval lobsters (Dow, 1977). In fact, the larval stage (first year of life) survival rate is known to be strongly dependent upon sea water temperature (Faustino, 1978). Studies performed on the stocks around Boothbay Harbor, Maine indicate that the optimal sea water temperature for lobster abundance appears to be in the range of 9.2°C to 10.7°C. Looking at the plot of annual sea water temperature for 1897 to 1984, in figure VI.3, it can be seen that the sea water temperature varies below and above this optimal range.

Since World War II, many attempts have been made to model the relationship between lobster abundance and environmental and economic variables; but, most of these models have relied upon standard economic or fisheries models to explain the dynamic behavior of abundance and sea water temperature and the models simply are able to deal with this type of dynamic behavior (Faustino, 1978). A good example of why such (these are non-time series models) models fails is the attempt to characterize the present level of lobster abundance as a function of past sea water temperatures (Faustino, 1978). Faustino (1978) appears to be among the first researchers to formally model the relationship between abundance and sea water temperature by use of a time series model. His model is a distributed lag model or simply a linear filter where the output is abundance and the inputs are annually lagged values of sea water temperature. Based upon the Faustino (1978) model, it was determined that a 1°C change in sea water temperature caused a 12.46% change in abundance within the same year (lag 0), a 7.45% change for the following year (lag 1), and 6.45% change in the third year (lag 2). This analysis is somewhat complicated by the fact that *H. americanus*



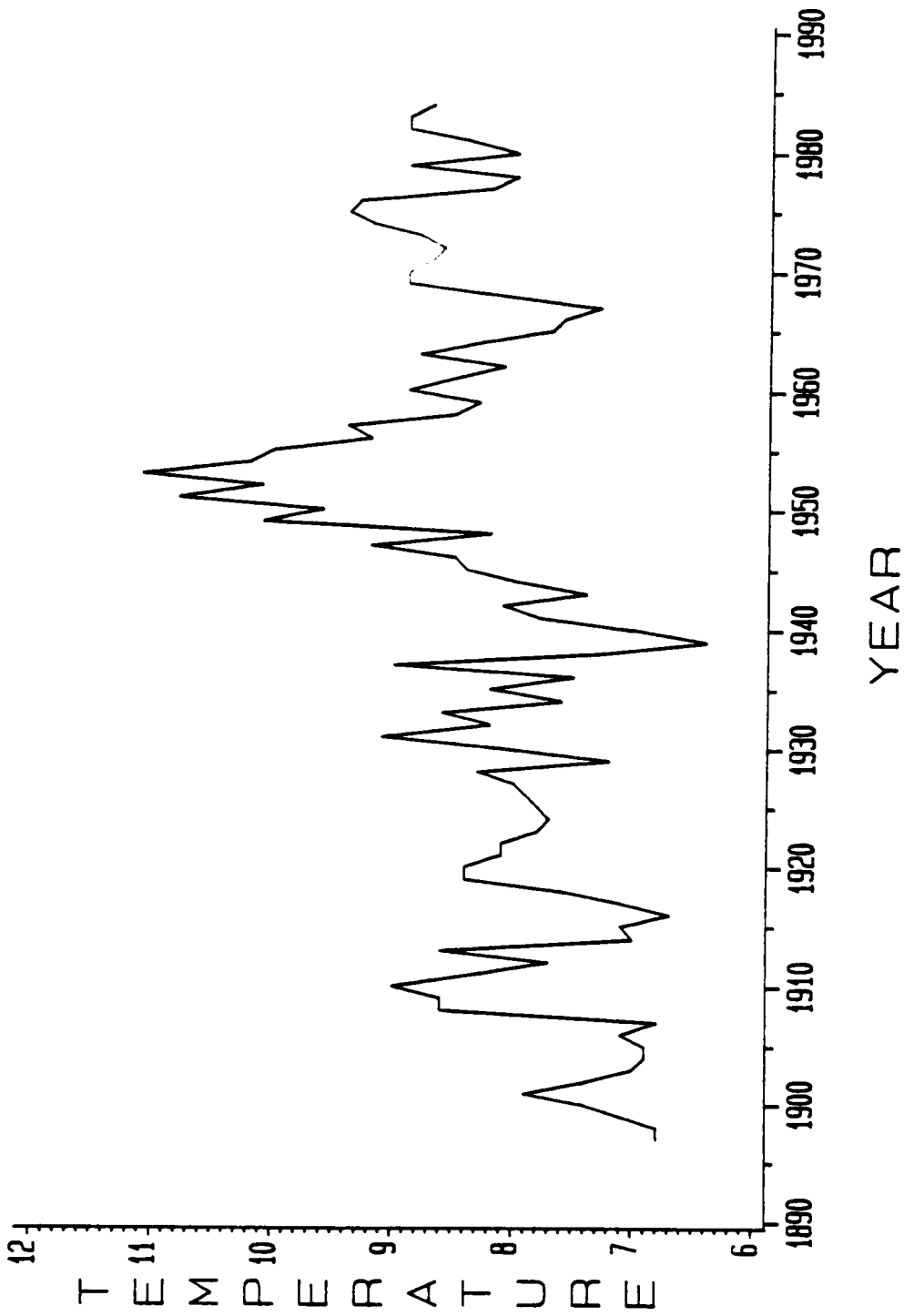
PLOT OF DETRENDED ANNUAL LOBSTER CATCH 1897-1984

Figure VI.2

shows very dynamic growth rates even under identical conditions (Dow, 1978); i.e., the effect of sea water temperature could easily be spread over several years because of the varying rates of growth of same age lobsters in the same stocks.

Faustino (1978) worked entirely in the time domain and there is no published evidence that frequency domain analysis has ever been applied to the problem. The use of frequency domain analysis would be indicated by the fact that ocean temperatures are known to have distinct climatic cycles and subcycles (Dow, 1969). In fact, if true patterns exist, then the observed catch (or abundance) at one observation period would represent an aggregate affect of all the cyclic patterns in the sea water temperatures. As a result, a time domain only approach can result in misleading conclusions because the true time lag relationships may be frequency dependent, i.e., group delay. If one works solely in the time domain, then the estimated time lag relationships actually represent an averaging of the various lags at the different frequencies. The present approach is to work in the frequency domain and attempt to model the relationship between the time series on an individual frequency basis using group delay.

Figures VI.4 and Figures VI.5 represent plots of the spectral densities of annual lobster harvest (in the State of Maine) and annual sea water temperature (measured at Boothbay Harbor, Maine), respectively. These were generated by taking the Fourier transform of the data and calculating the individual periodograms for the two data sets, as as described in section II.2. The spectral densities were calculated by using a Daniell window with a smoothing parameter $m = 3$ (see section II.2). Both spectral density plots show a number of significant peaks at the lower frequencies. The spectral density plot for sea water temperature in particular seems to indicate the existence of many cycles with widely different frequencies which agrees with the statement of Dow (1969).



PLOT OF ANNUAL SEA SURFACE TEMPERATURE 1897-1984

Figure VI.3

Even though the individual time series have apparant cyclical patterns, our real interest is in the covariation between the two series at these various frequencies. In figure VI.6 we have a plot of the coherence squared (see section II.3) between catch and sea water temperature; there appear to be a number of frequencies with the coherence values above 0.5. This would indicate that the two series have coherent cyclic patterns which agrees with the large body of scientific evidence linking the two time series. In any case, the existence of substantial coherence indicates that meaningful phase and group delay relationships exist and an analysis of these relationships is meaningful.

The raw sample phase (see section II.3) is plotted in figure VI.7. The sample phase values were calculated by using the real and imaginary parts of the cross-periodograms and then computing

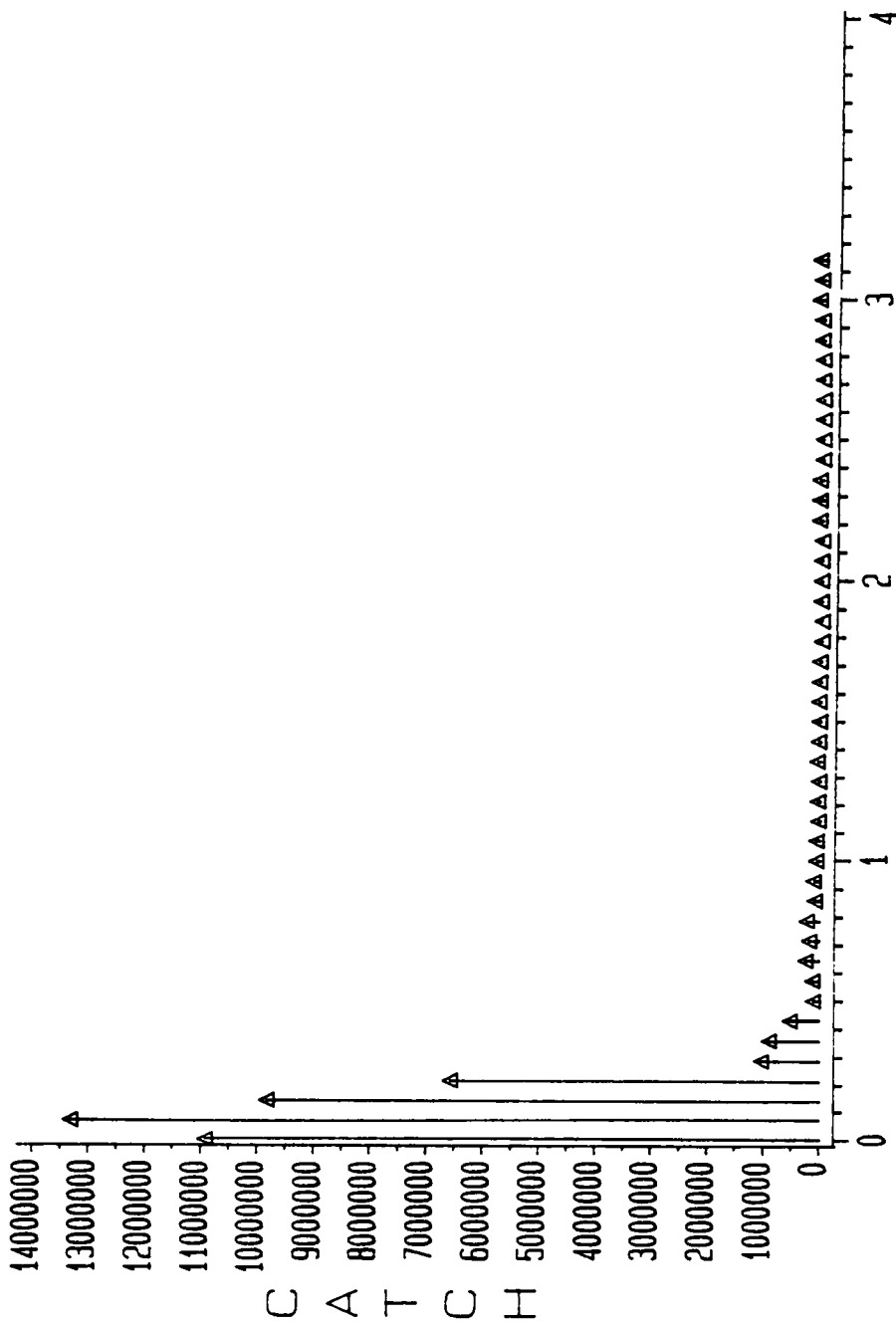
$$\hat{\phi}_{xy}(\omega_i) = \tan^{-1} \left\{ \frac{\text{Im}(I_{xy}(\omega_i))}{\text{Re}(I_{xy}(\omega_i))} \right\} \quad [1]$$

The resulting sample phase function was then unwrapped by

$$\tilde{\phi}_{xy}(\omega_i) = \hat{\phi}_{xy}(\omega_i) \pm j\pi, \quad j = 0, \pm 2, \pm 4, \dots \quad [2]$$

Figure VI.8 is a plot of the $\tilde{\phi}_{xy}(\omega_i)$ from [2]. Notice the general similarity in the shape of the plot in figure VI.8 and the continuous time filter phase function in figure V.1. It would seem quite reasonable that the relationship between the two series might actually be modelled by the 'scratch' filter of Chapter V. However, at present no methods exist for determining a linear filter for two time series based upon the sample phase function.

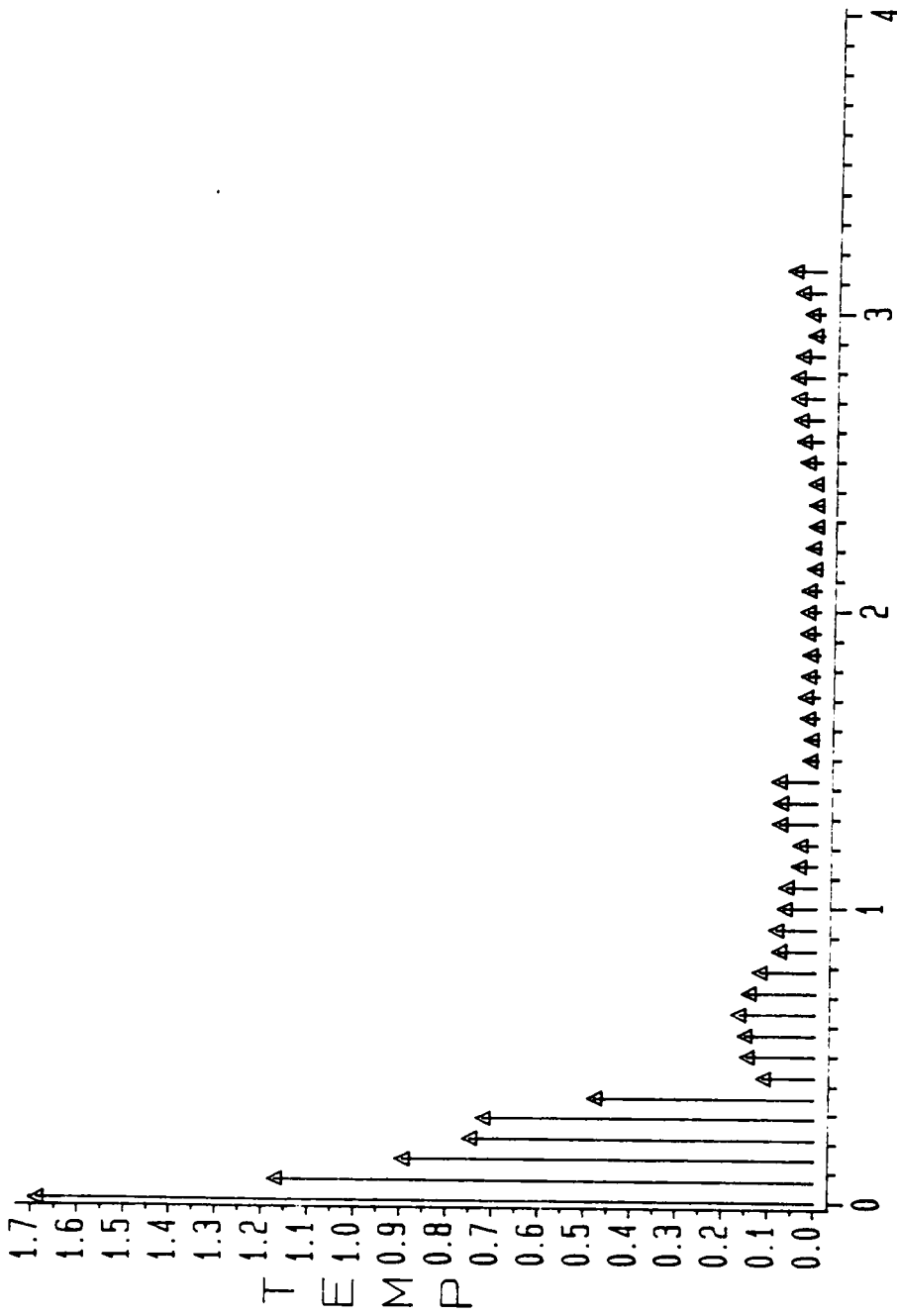
A Box-Cox transformation, using equation [11] of Chapter V, was applied to the unwrapped phase plotted in figure VI.8. Optimizing equation [12] of Chapter V, the power parameter η_1 was



FREQUENCY

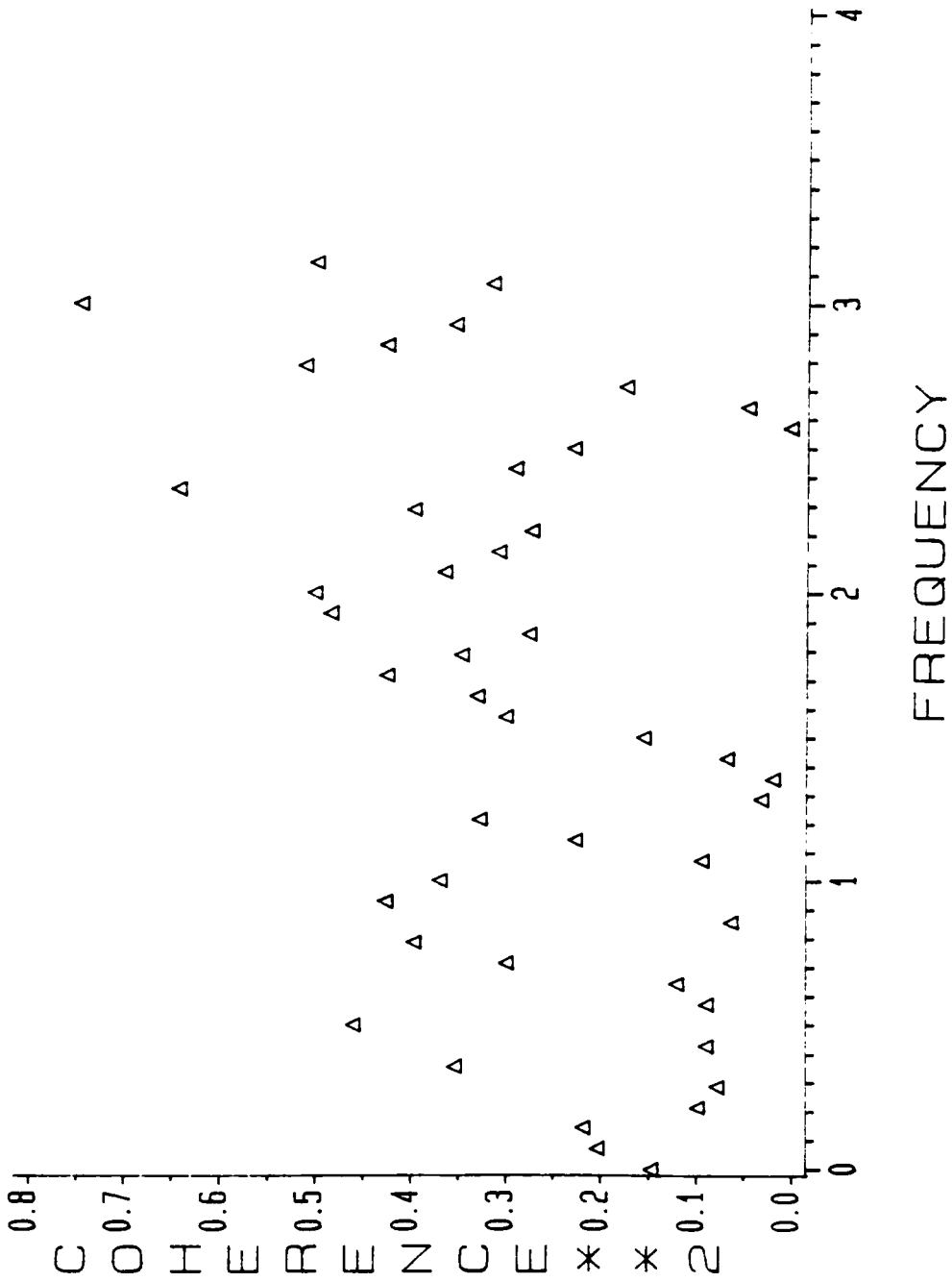
PLOT OF SPECTRAL DENSITY OF HARVEST

Figure VI.4



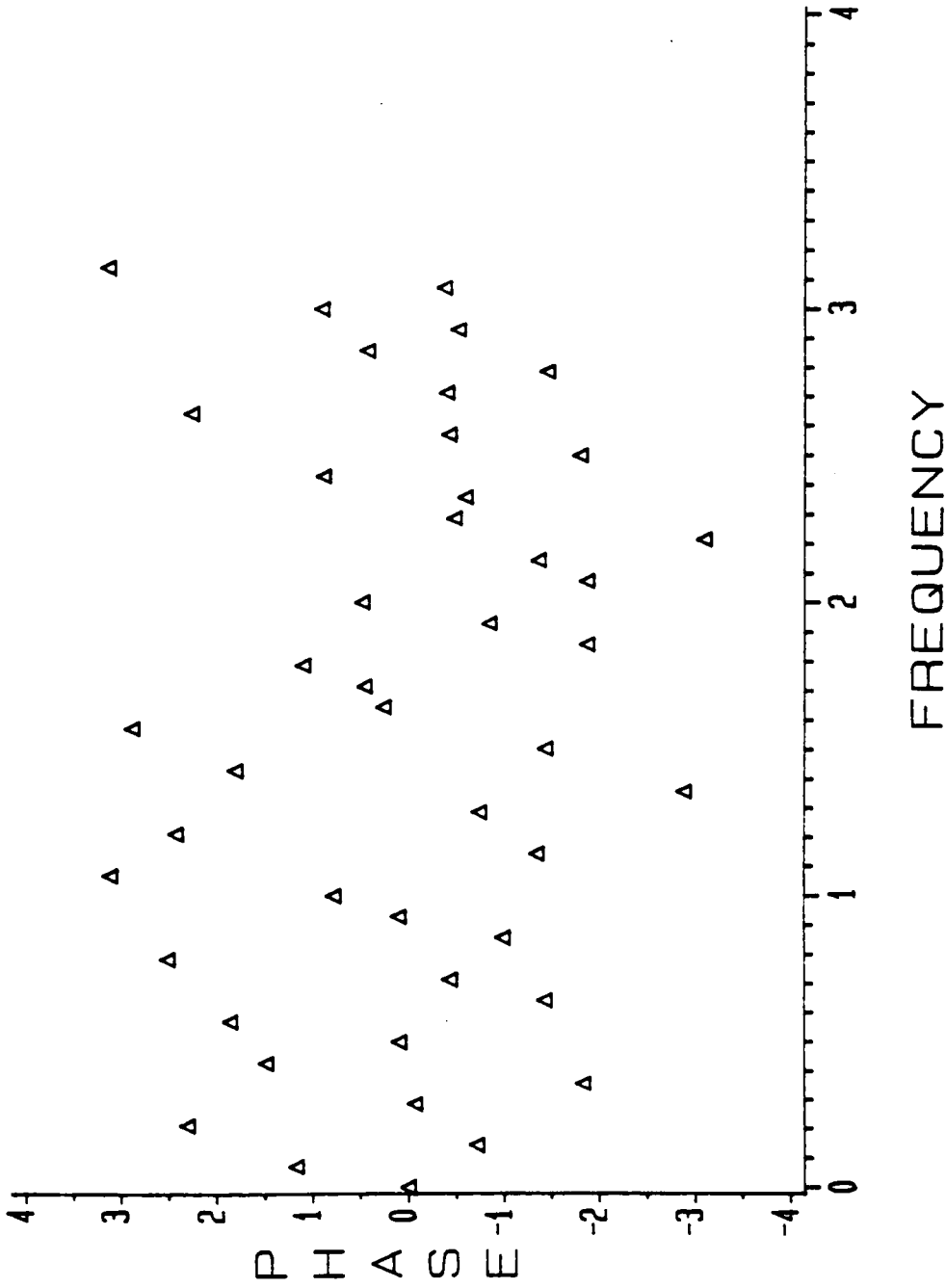
PLOT OF SPECTRAL DENSITY OF TEMPERATURE

Figure VI.5



PLOT OF COHERENCE SQUARED FUNCTION

Figure VI.6



PLOT OF RAW PHASE FUNCTION

Figure VI.7

determined to have a value of 1/3 with η_2 set as close to the singularity as possible (see section V.3). The resulting transformed sample phase values were then modelled with a simple linear model; i.e.,

$$g(\omega_i) = \beta_0 + \beta_1 \omega_i + \varepsilon \quad [3]$$

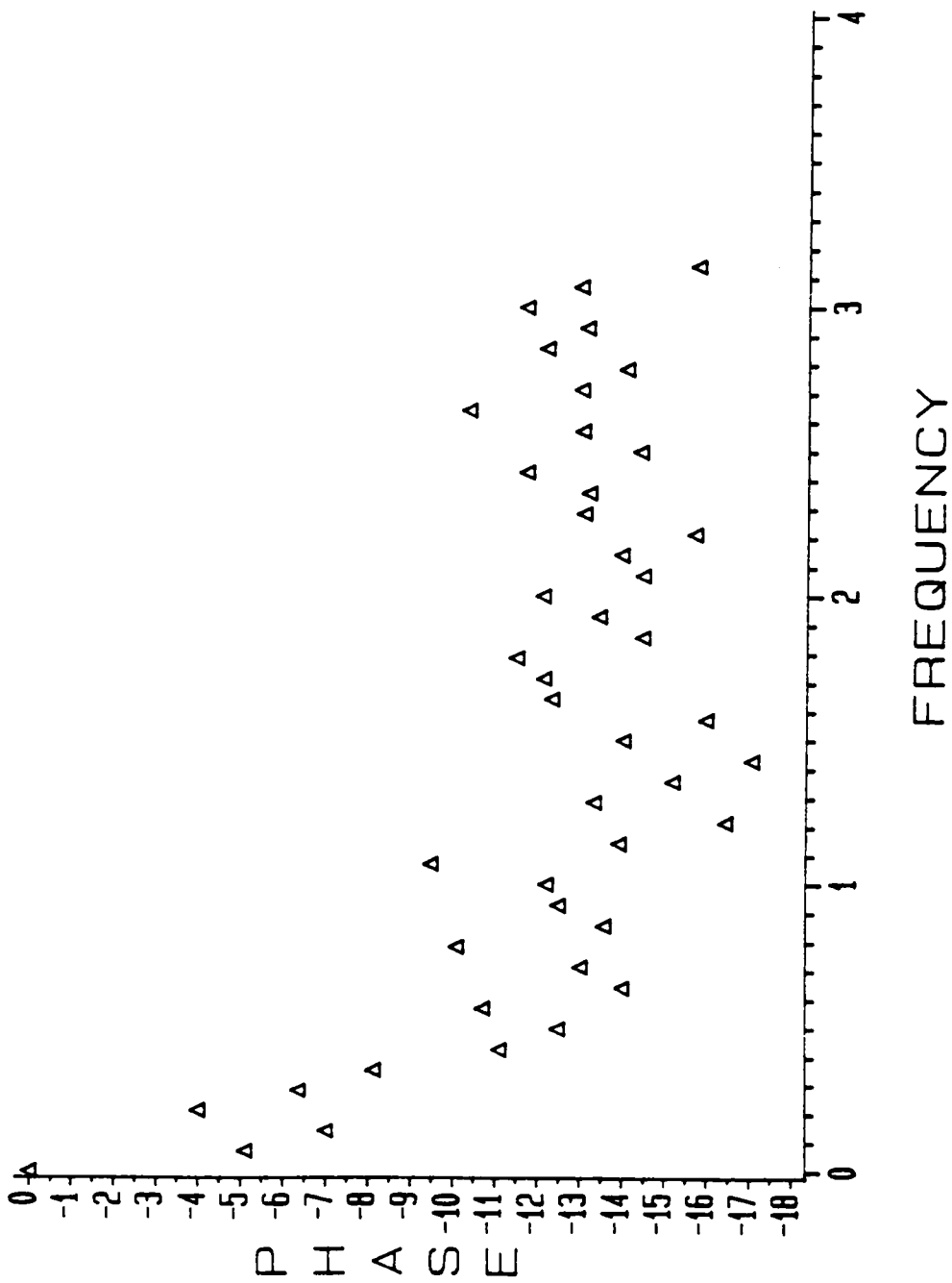
and the estimate of the model is

$$\hat{g}(\omega_i) = \hat{\beta}_0 + \hat{\beta}_1 \omega_i. \quad [4]$$

Figure VI.9 is a plot of the RSTUDENT residuals resulting from [4] and there is some evidence of pattern. However, an attempt to fit a higher order model did not show higher order terms to be significant. In any case the simple linear model was considered adequate and the statistics generated by [4] were then used to generate the UMVUE's of expected sample phase and group delay.

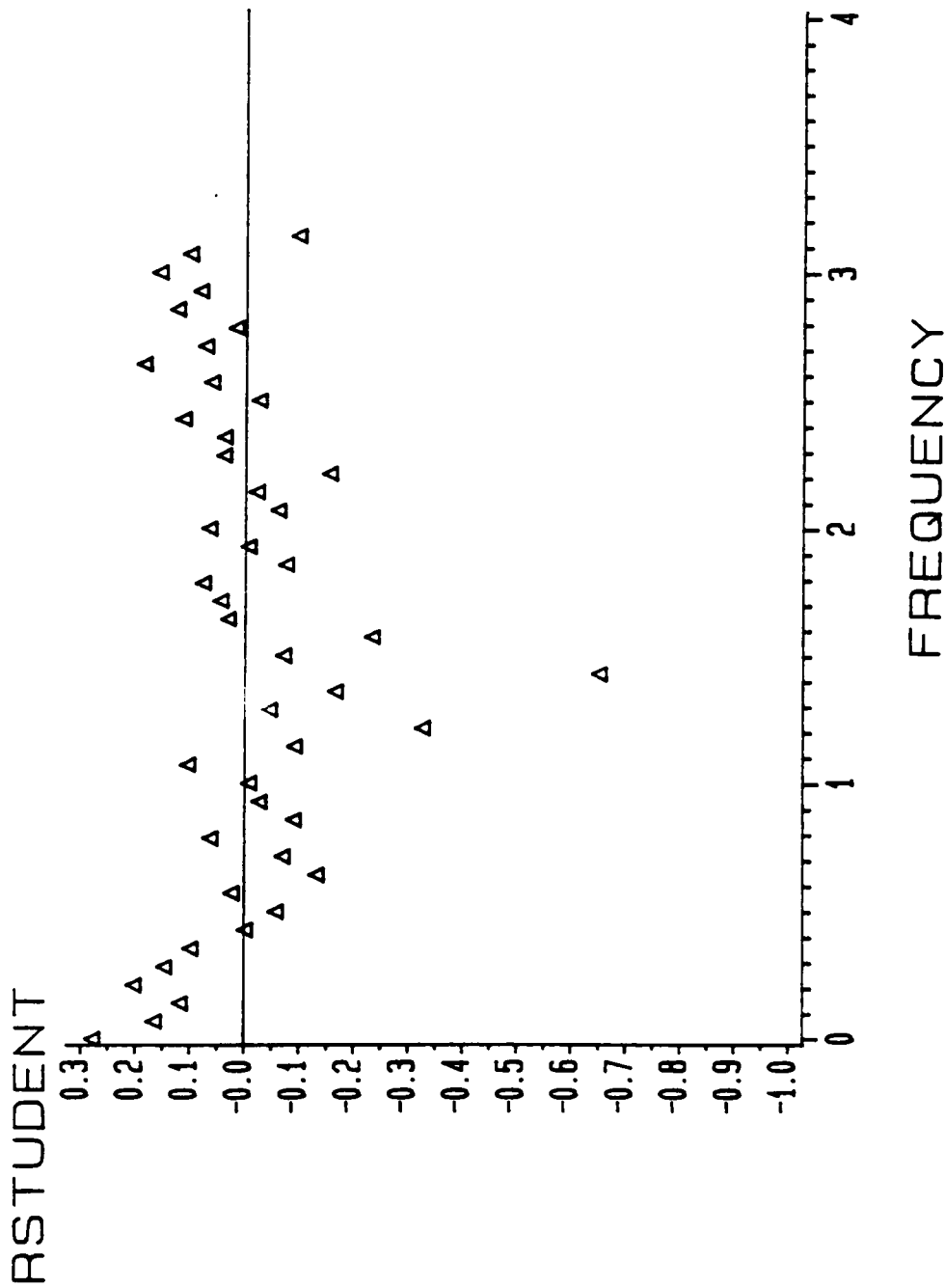
The UMVUE's for expected sample phase and group delay were generated by equations [14] and [15], respectively, of Chapter V. Also, a 95% confidence interval for the expected sample phase function was generated by the transformation method of section IV.5.1. Figure VI.10 is a plot of the UMVUE's of the expected sample phase at each Fourier frequency and the upper and lower confidence limits.

Similar results for the expected sample group delay function are given in figure VI.11. The estimates calculated by the direct UMVUE method and the finite difference method are shown in the plot. Overall the ranges of values shown in figure VI.11 do seem to agree with values one would expect, based upon a review of the scientific literature on the problem. Therefore, one may at least



PLOT OF UNWRAPPED PHASE FUNCTION

Figure VI.8



RSTUDENT FOR TRANSFORMED PHASE VALUES OF CATCH AND TEMP.

Figure VI.9

conclude that the UMVUE procedures for phase and group delay have performed acceptably in the analysis of the lobster and sea water data.

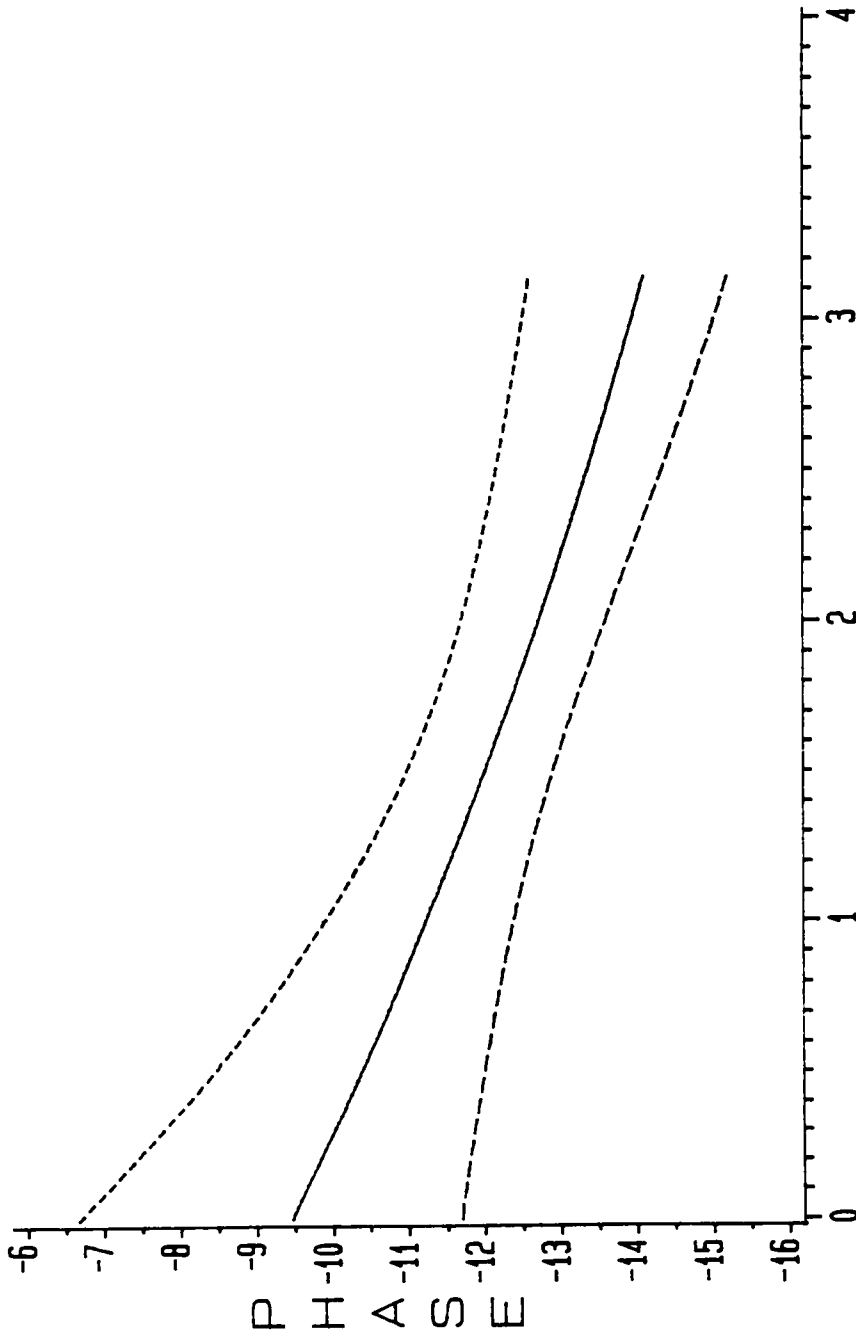
In figure VI.11 we have plotted the time lag estimates, at each Fourier frequency, and the next step is to generate estimates of the coefficients relating the two time series and each of the frequencies. Estimation of the frequency dependent coefficients is straightforward, since they are actually estimates of the gain at each frequency (see Chatfield, 1984, p. 178 or section II.3). Therefore, we may estimate the coefficient at frequency ω_i by

$$\hat{\beta}_{xy}(\omega_i) = \frac{|f_{xy}(\omega_i)|}{f_x(\omega_i)}. \quad [5]$$

The RHS of [5] is directly analogous to the estimate of the slope parameter in simple linear regression models, except the estimate here only applies at a specific frequency. Figure VI.12 is a plot of the coefficients for all the Fourier frequencies, computed by [5]. As an example of our final model relating catch to sea water temperature, we select the third Fourier frequency coefficient plotted in figure VI.12 and our model estimate (at $\omega_i = \Lambda$) using figures VI.11 and VI.12 is

$$\text{Catch}_{\Lambda}(t) = 1550\text{Temp}_{\Lambda}(t - 1.8). \quad [6]$$

Recall, catch is in metric tons, temperature is in celsius, and time is in years. At this Fourier frequency we conclude that increases in lobster abundance are preceded by increases in sea water temperature by about 2 years.

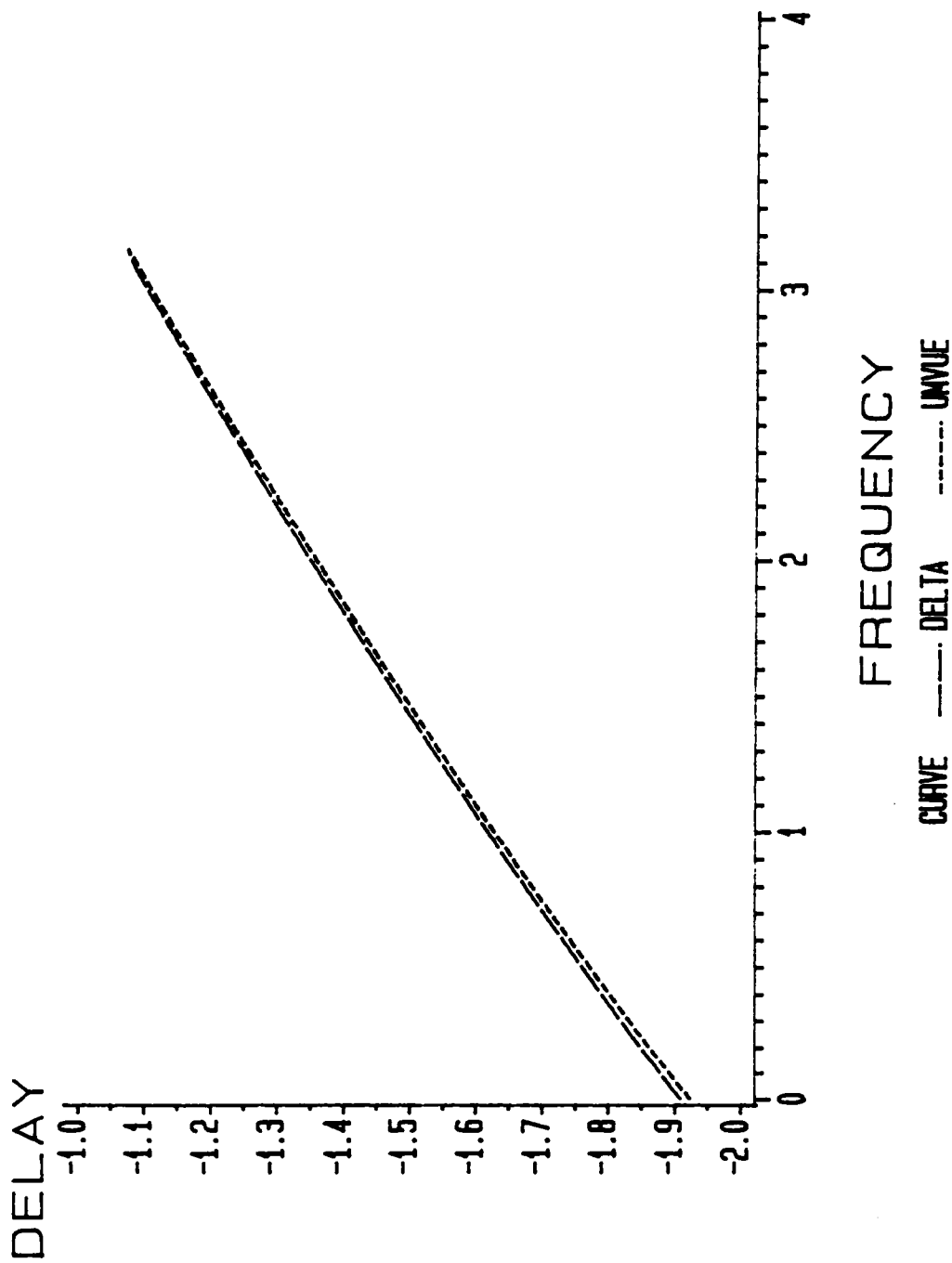


FREQUENCY

CURVE — UMVUE - - - - 95%LC - . - . - 95%UC

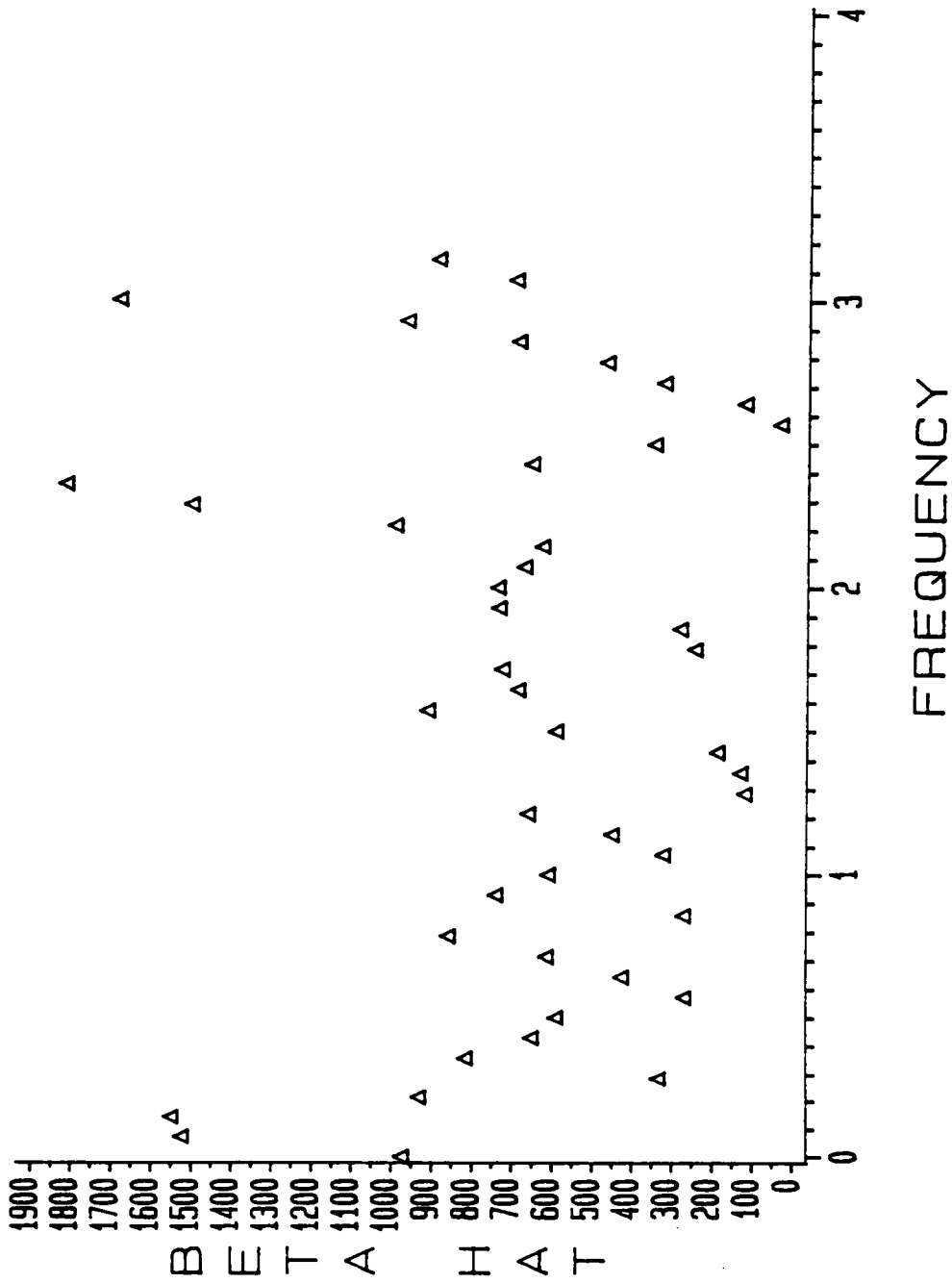
UMVUE AND 95% CI ESTIMATES OF PHASE FOR CATCH AND TEMP.

Figure V1.10



UNVUE'S FOR 60 OF LOBSTER HARVEST AND WATER TEMPERATURE

Figure VI.11



PLOT OF COEFFICIENTS FOR GD MODEL OF CATCH AND TEMP.

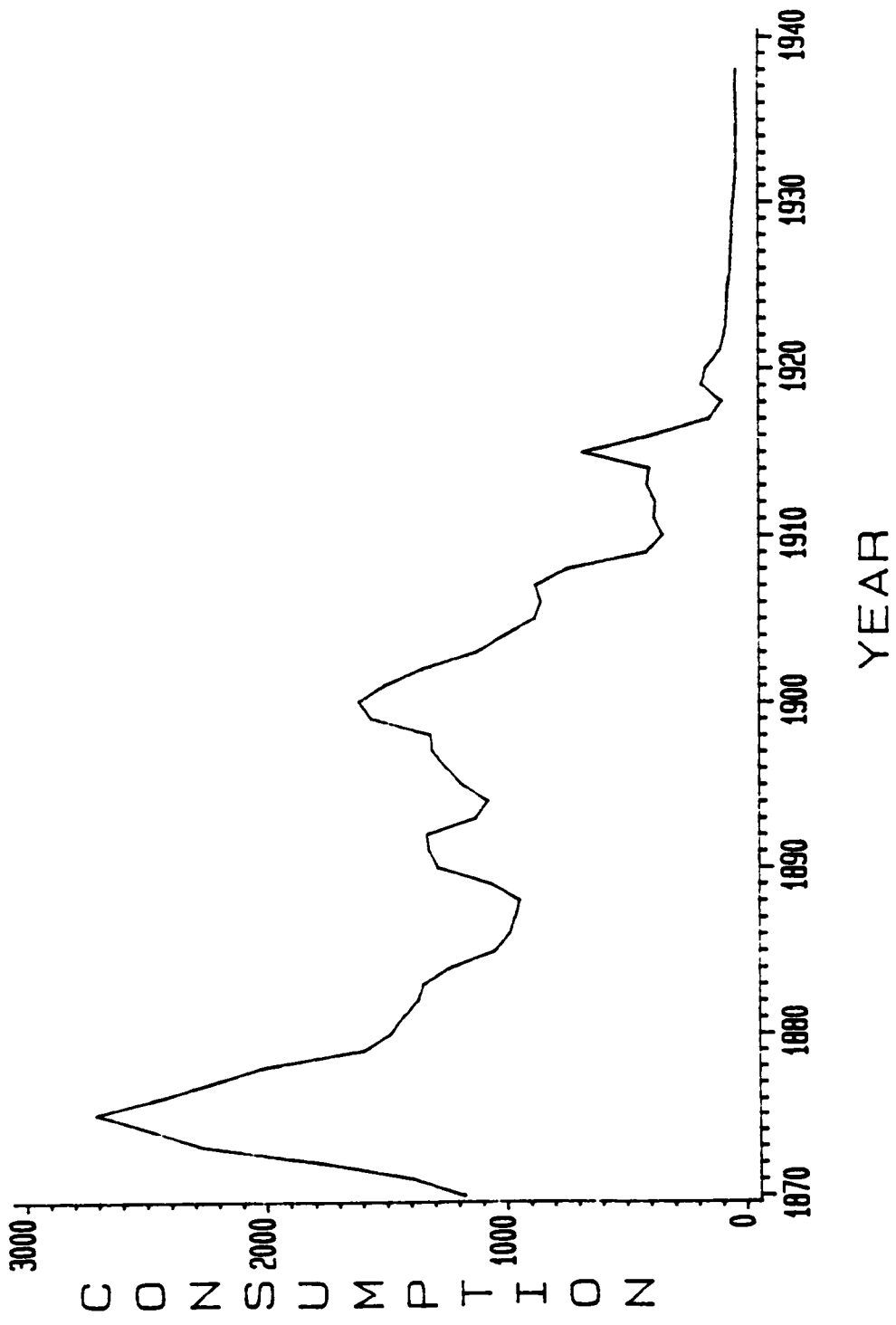
Figure VI.12

VI.3. Sunspots vs. Consumption Of Alcohol By Londoner's.

Whereas the first example was based upon a significant body of scientific research, which could be used in performing group delay analysis, this second example is more a case of data exploration. In the example we explore the possible relationship between alcohol consumption by residents of London from 1870 through 1938 and the Wolfer sunspot numbers for the same time period. In this example we have no prior knowledge of a relationship and therefore we are simply performing an exploration to see if a possible relationship exists, in the form of group delay.

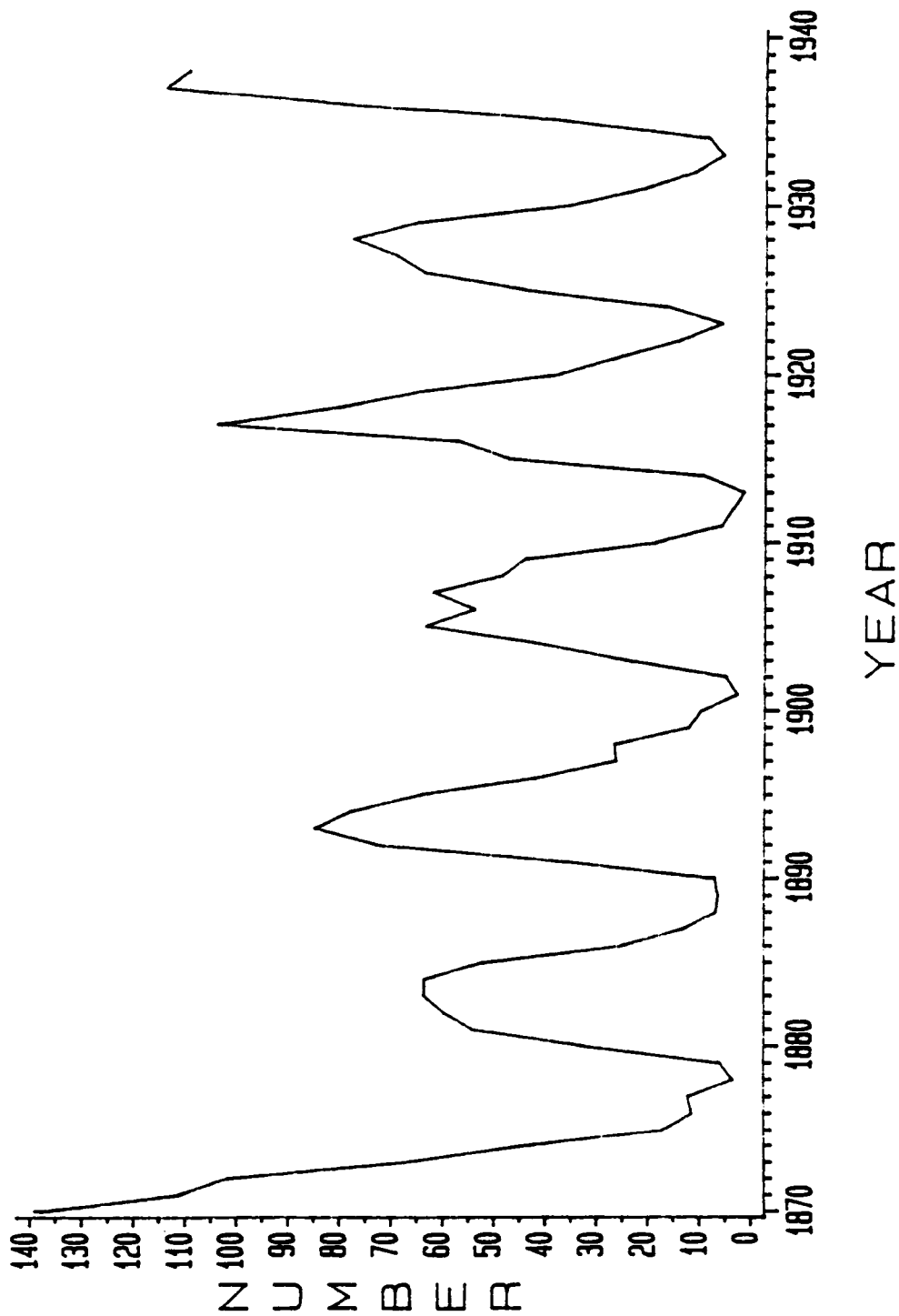
Figure VI.13 is a plot of the per capita consumption of alcoholic beverages by residents of London from 1870 to 1938. The data are discussed by Fuller (1976, p. 427), Prest (1949), and Durbin and Watson (1951). Although a direct relationship between sunspots and beverage consumption might, at first, seem rather unscientific, there is a growing body of scientific evidence that sunspot activity directly affects the Earth's climate. In other words, if one accepts that the increase levels of sunspot activity directly effect global temperatures, then one would actually have a scientific basis for a cause and effect relationship between sunspots and alcohol consumption. The idea of sunspot activity effecting economic systems has been considered by economists at least as far back as Jevons (1884). A recent discussion on the possible link between sunspots and economic systems can be found in the New Palgrave dictionary (Eatwell, 1987).

Although there is much speculation on the impact of sunspots on weather, no concrete links have been found; i.e., a causal mechanism. However, a recent article in Sky And Telescope (1988) indicates that sunspot activity is now known to directly effect the strength and temperature of the



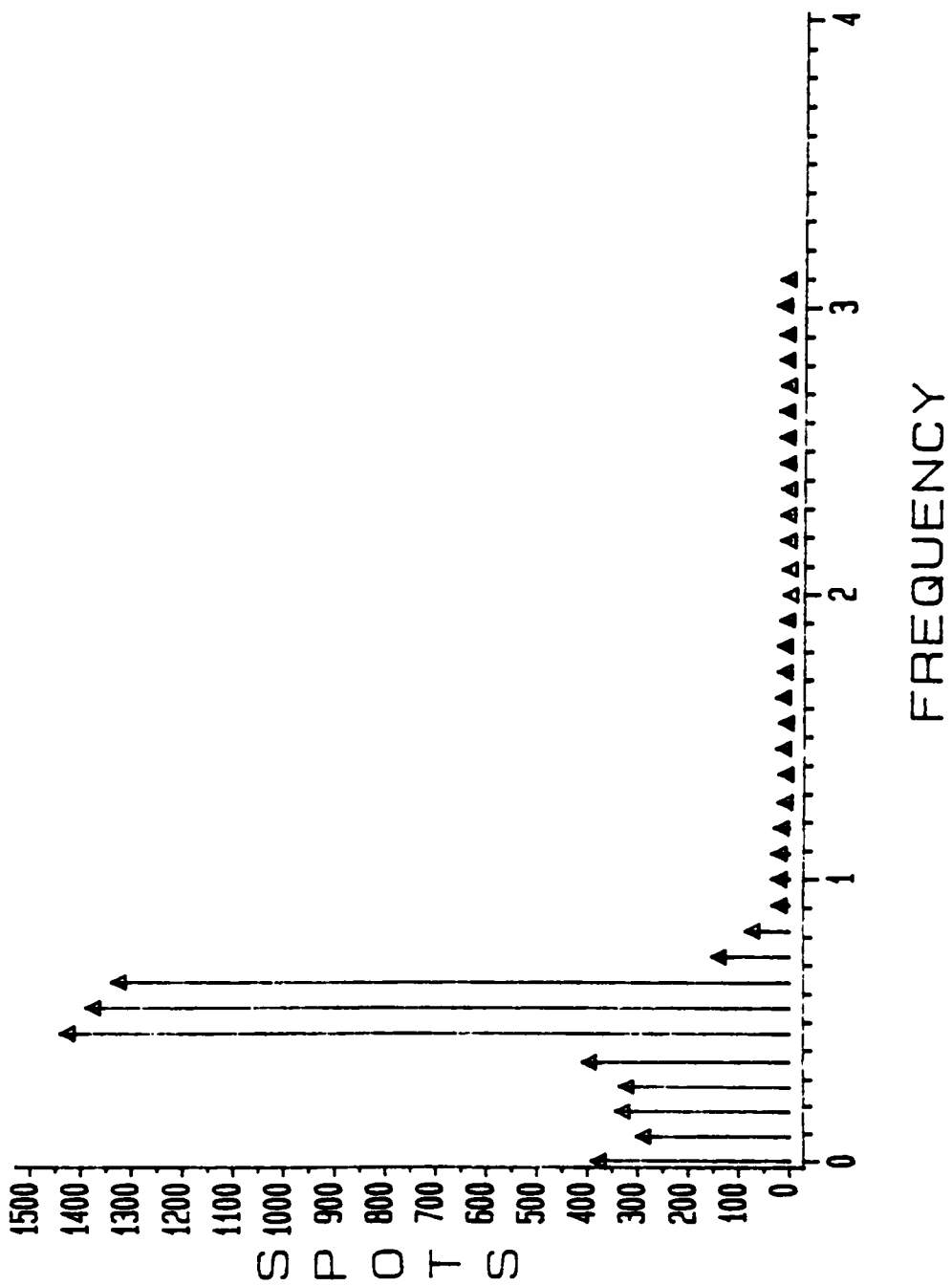
LONDON PER CAPITA SPIRIT CONSUMPTION 1870-1938

Figure VI.13



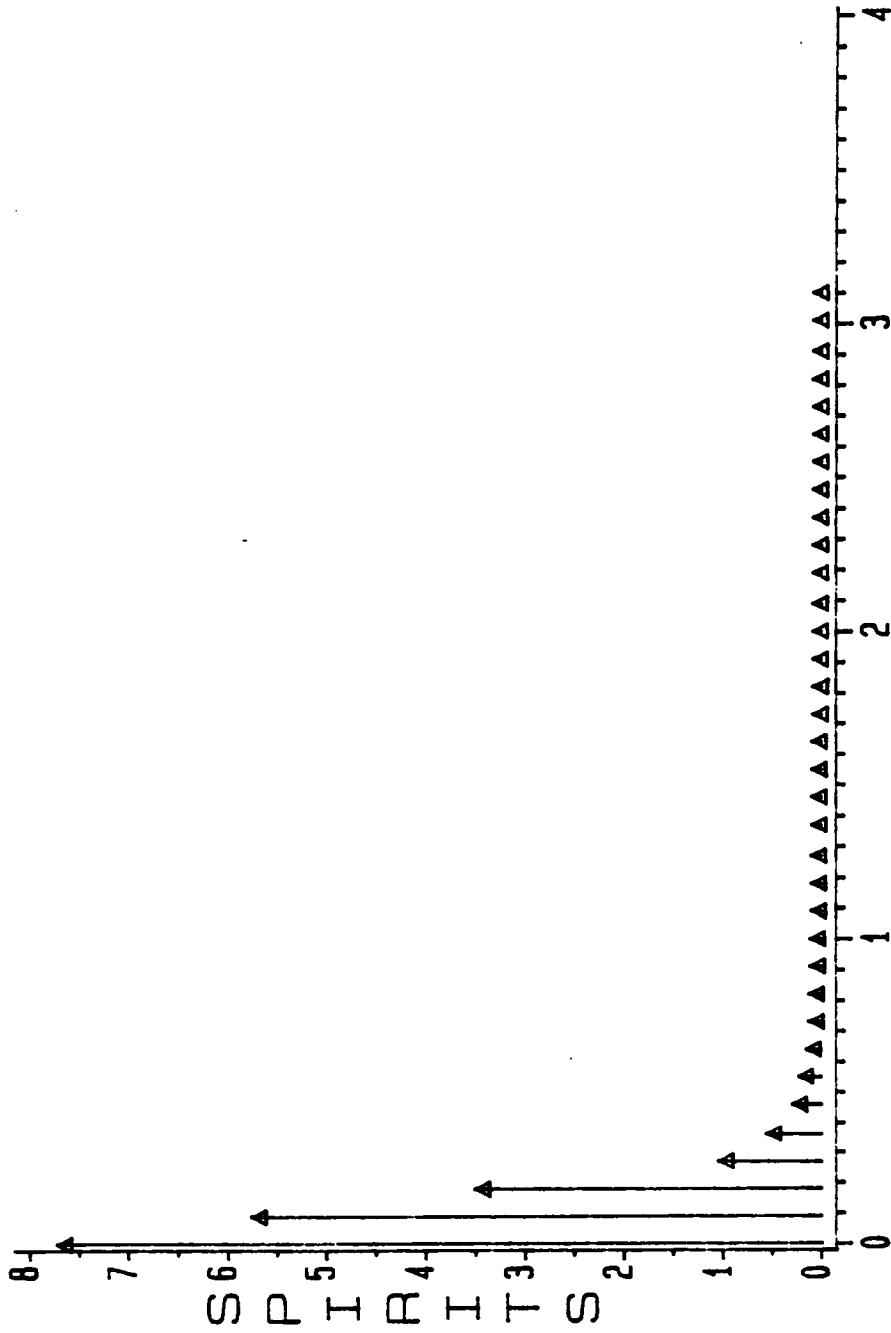
WOLFER SUNSPOT NUMBERS FOR 1870-1938

Figure VI.14



SPECTRAL DENSITY OF SUNPOTS FOR 1870-1930

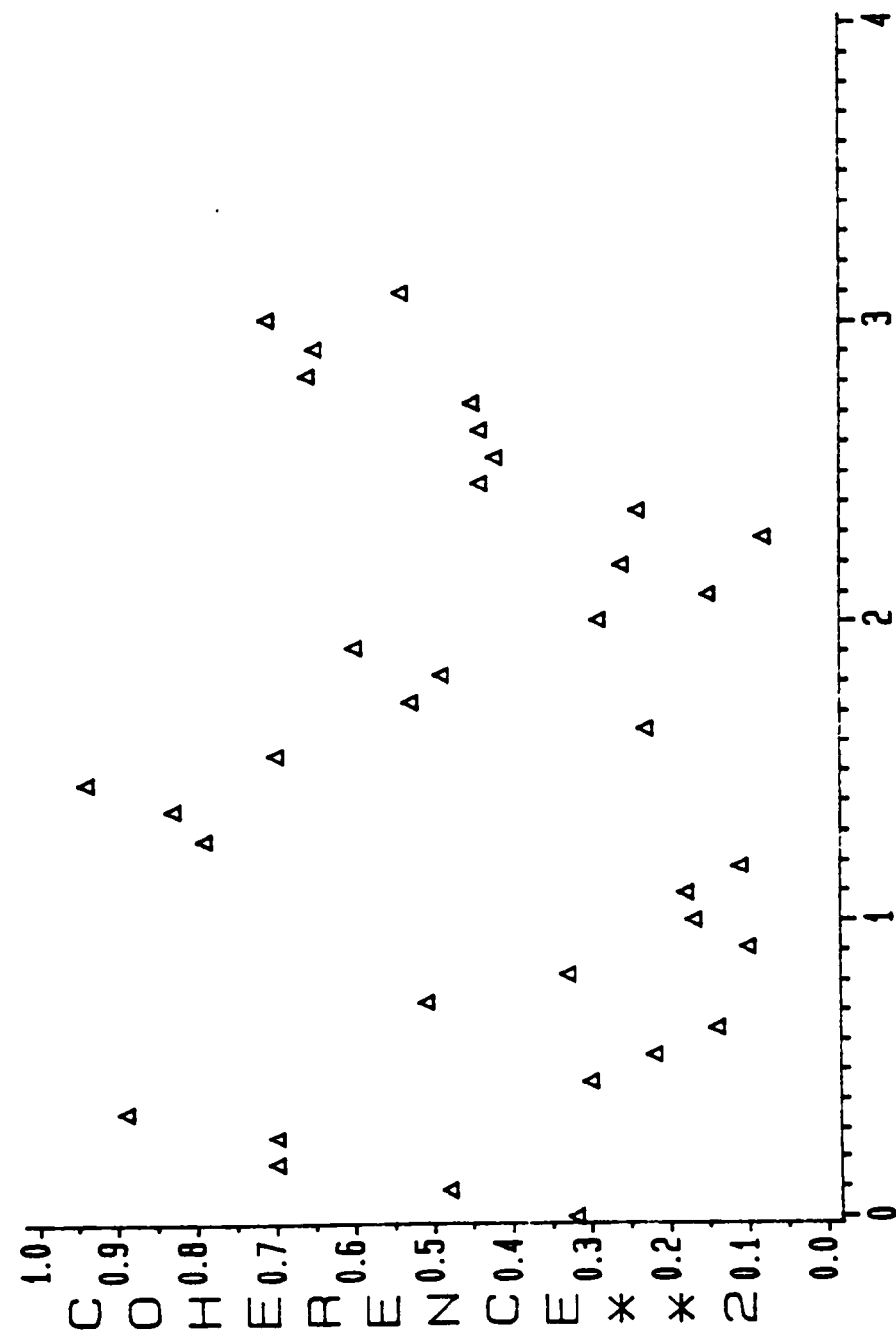
Figure VI.15



FREQUENCY

SPECTRAL DENSITY OF CONSUMPTION FOR 1870-1938

Figure VI.16



COHERENCE BETWEEN CONSUMPTION AND SUNSPOTS FOR 1870-1938

Figure VI.17

Northern Polar winds. The Earth's climatic temperature is then directly influenced and becomes warmer with increased sunspot activity. The link of sunspots and global temperatures at least provides an initial model for a cause and effect relationship between consumption and sunspot activity.

Figure VI.14 is a plot of the Wolfer sunspot numbers for 1870 through 1938 and the source for these numbers is Menzel (1959, p. 107). Initially the usual Pearson product moment correlation coefficient was calculated for the two data sets in figures VI.13 and VI.14 and the analysis gave a value $r = -.101$ which is hardly an interesting correlation. As a followup the possibility of cross correlations between the two time series were explored with SAS's Proc Arima and the analysis showed broad bands of time lags where the two series may be related, but none of the lags were in excess of the confidence band for 0 correlation. However, under group delay analysis we do not have to assume that the two time series are highly correlated in the time domain, we only assume that the two series are significantly coherent at specific patterns (frequencies) within the two series.

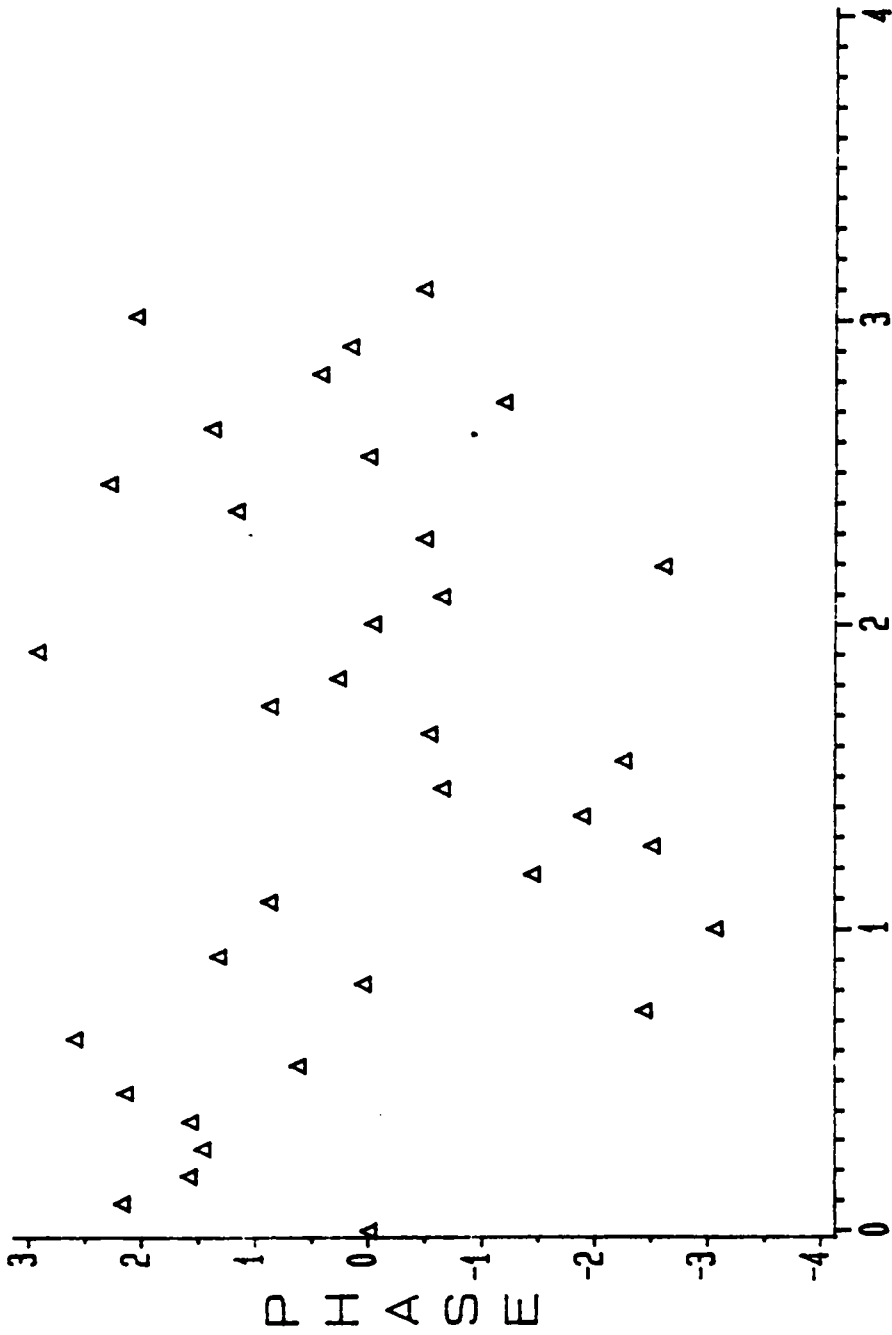
In order to examine the relationship between the time series in the frequency domain, the two data sets were transformed by Fourier transforms and periodograms generated from the transformed data sets. The spectral density estimates were generated, as in the example of section VI.2, by a Daniell window with $m = 3$. Figures VI.15 and VI.16 are plots of the spectral densities of sunspot activity and alcohol consumption, respectively; in figure VI.16 the ordinate for the famous 11 year sunspot cycle is the ordinate of the sixth Fourier frequency. Figure VI.17 is a plot of the coherence squared estimates between the two time series. It can be noticed that there are a number of Fourier frequencies at which very large values of coherence squared exist. It would appear that there may well be some strong relationships between consumption and sunspots when the two series are ex-

amined in the frequency domain. Recalling that a time-domain-only analysis averages together affects of all the cycles, it is entirely possible that one would not find significant relationships in a time domain approach.

The sample phase values were generated from the cross-periodograms by use of [1]. The observed sample phase values are plotted in figure VI.18. The sample phase values were then unwrapped by using [2] and the resulting observed phase values are plotted in figure VI.19. The unwrapped phase values were then transformed by the Box-Cox transformation with a power parameter $\eta_1 = 1/2$. The value of η_1 was determined by optimizing [12] over Chapter V. Once transformed, the sample phase values were modelled by a simple linear model, as in [3] and [4]. Figure VI.20 is a plot of the RSTUDENT residuals after fitting the simple linear model between transformed phase and frequency. There is not an indication in the plot that the model or transformation has been inadequate in generating approximate normality and a simple linear relationship.

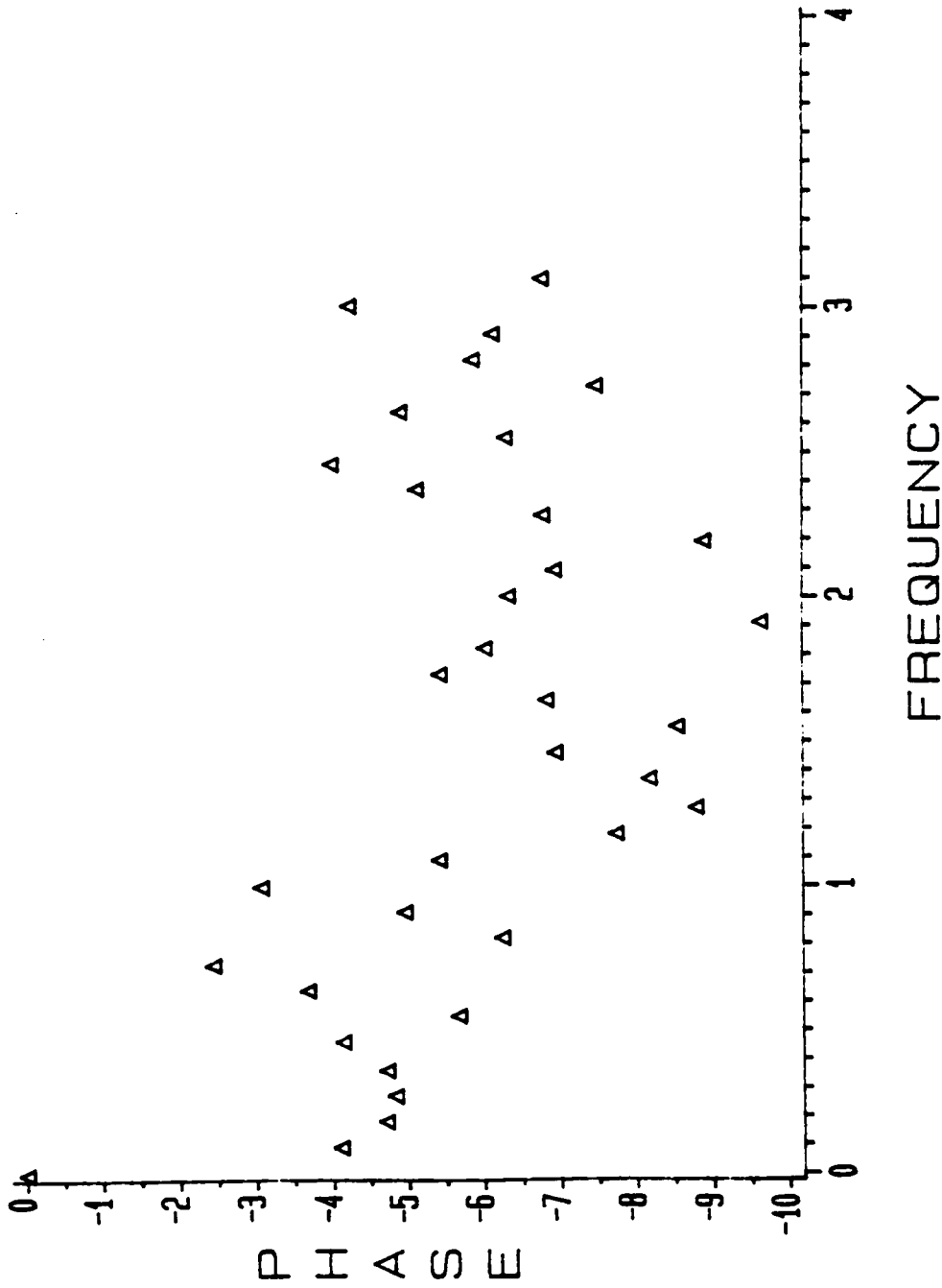
Using the statistics generated by the fit of a simple linear model, UMVUE's for the expected sample phase and group delay were generated by equations [20] and [21] of Chapter V, respectively. Figure VI.21 is a plot of the UMVUE's of expected sample phase along with the 95% confidence interval generated by the transformation method. Likewise, figure VI.22 is a plot of the UMVUE's of the expected sample group delay between consumption and sunspots.

Looking at the plot of group delay in figure VI.22 it appears that the events in the sunspot cycle seem to precede events in alcohol consumption by Londoner's by about 1 year to 6 months. Since the effect of sunspot activity on Earth's weather is believed rather direct (Kerr, 1988), one would



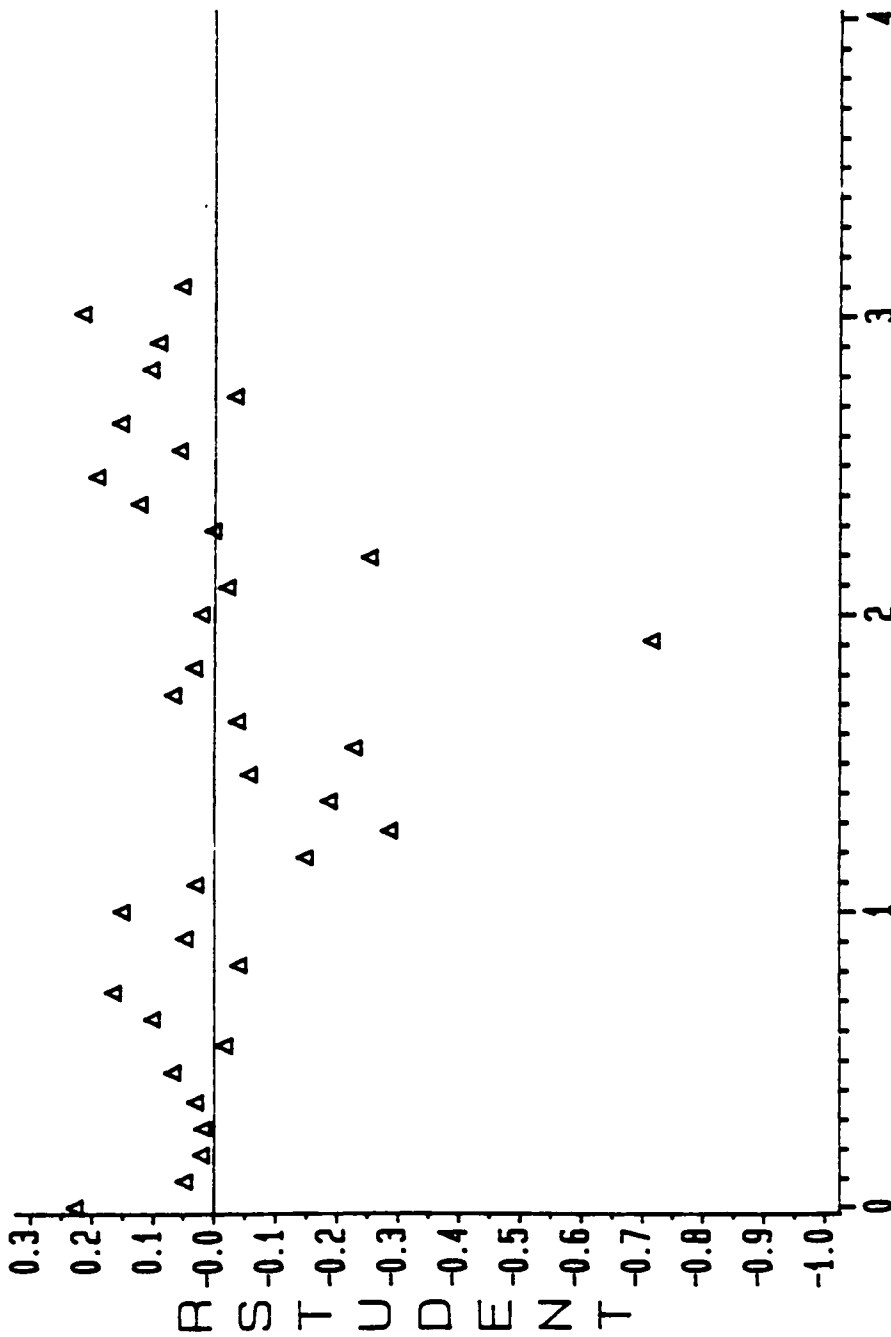
SAMPLE PHASE FOR CONSUMPTION AND SUNSPOTS FROM 1870-1938

Figure VI.18



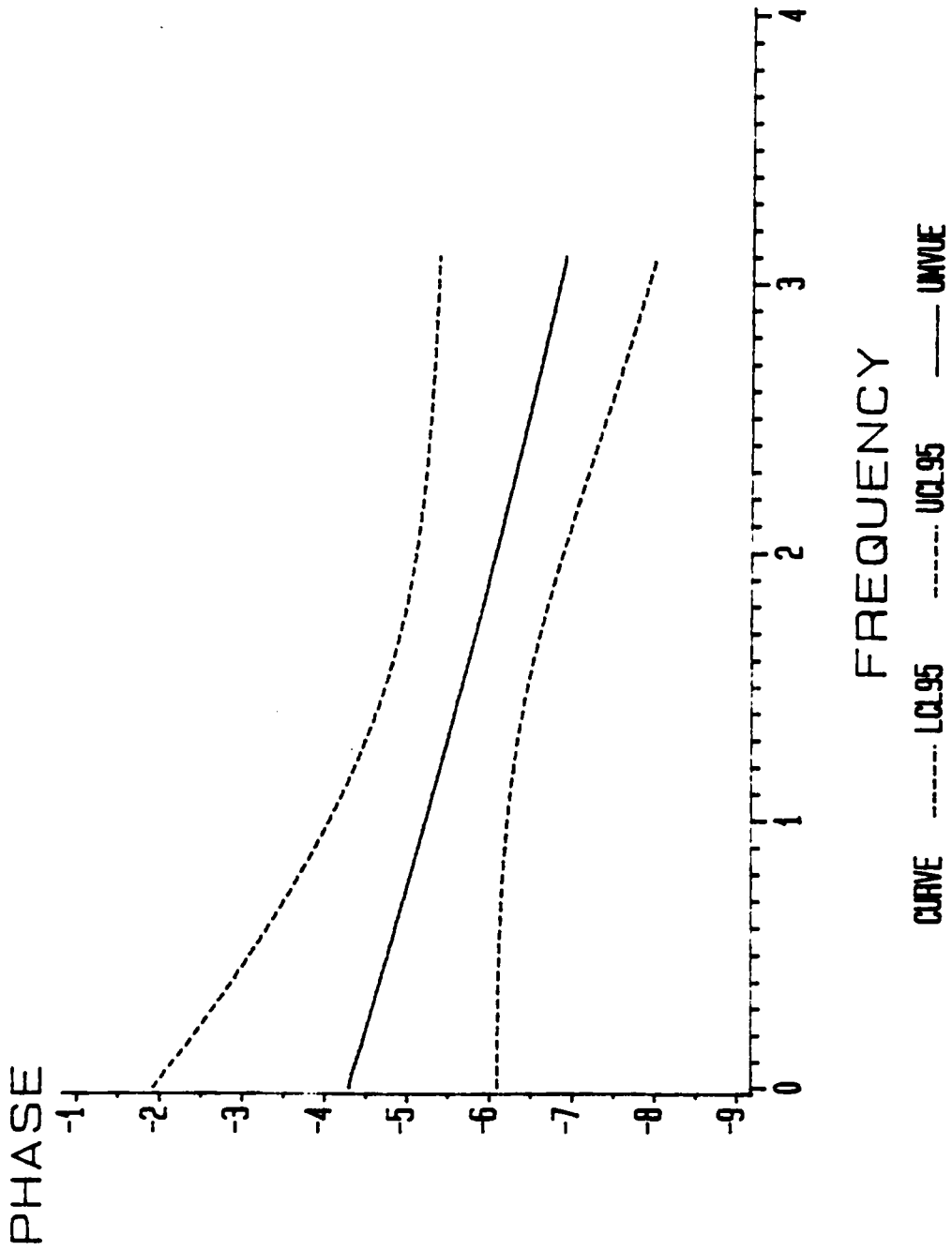
UNWRAPPED PHASE FOR CONSUMPTION AND SUNSPOTS FROM 1870-1938

Figure VI.19



RESIDUAL FOR FIT OF TRANS. PHASE FOR SUNSPOTS AND CONSUMPTION

Figure VI.20



CURVE LCL.95 - · - · - · UCL.95 ——— UMVUE
 UMVUE AND 95% CI FOR PHASE OF SUNSPOTS AND CONSUMPTION

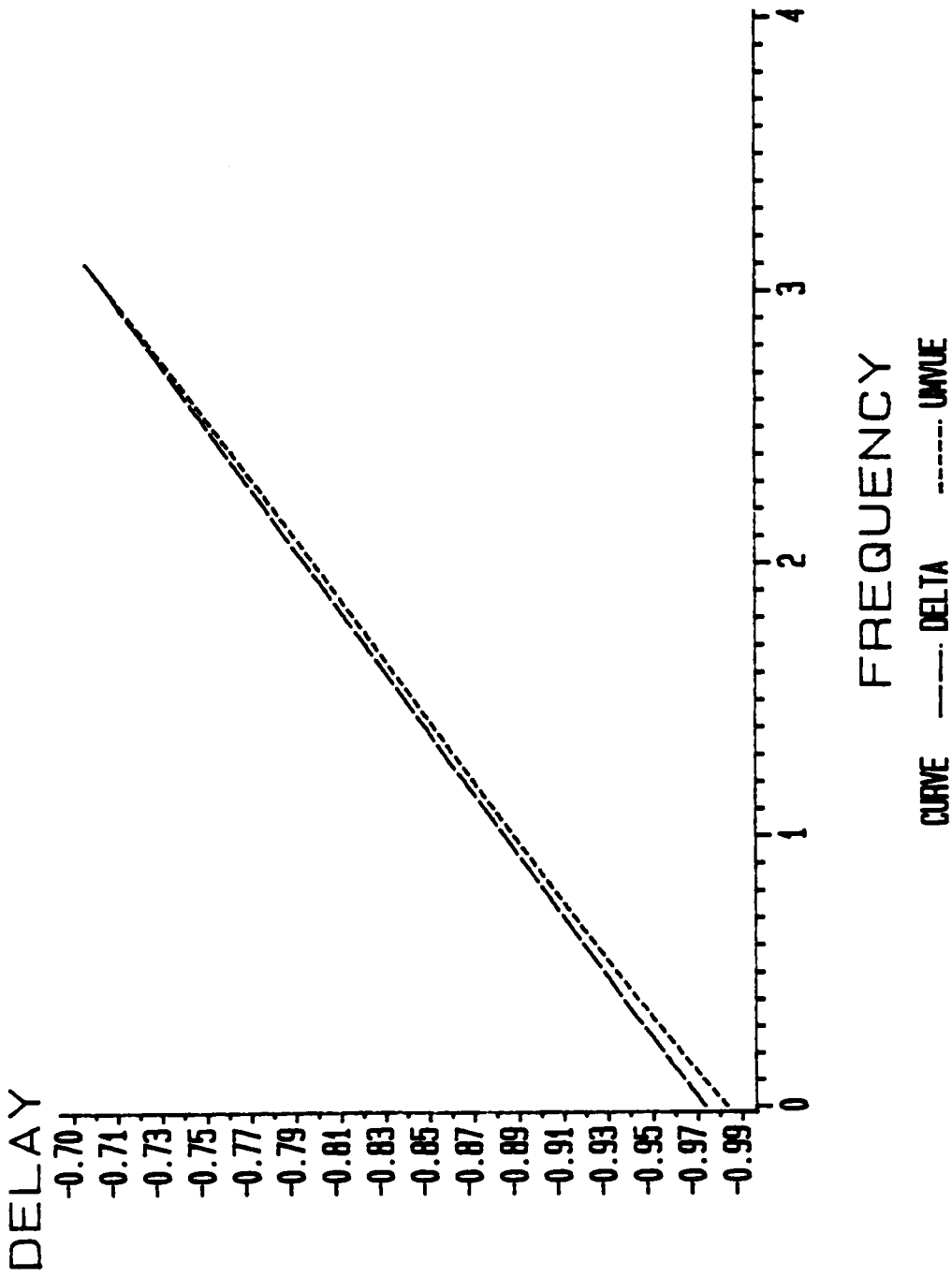
Figure VI.21

not expect long delays between the events in sunspot activity and their affects upon various phenomena on Earth; i.e., the delay values calculated do not appear to be scientifically unrealistic. Figure VI.23 is a plot of the coefficients relating the 2 time series at the various Fourier frequencies and the coefficients were computed using [5]. Therefore, using the values plotted in figures VI.22 and VI.23 we are able to develop simple linear models between the time series on a frequency basis. As an example, for the 11 year sunspot cycle we have the following model (Λ = the frequency for the 11 year cycle)

$$\text{consumption}_{\Lambda}(t) = 0.03\text{Sunspots}_{\Lambda}(t - 1). \quad [8]$$

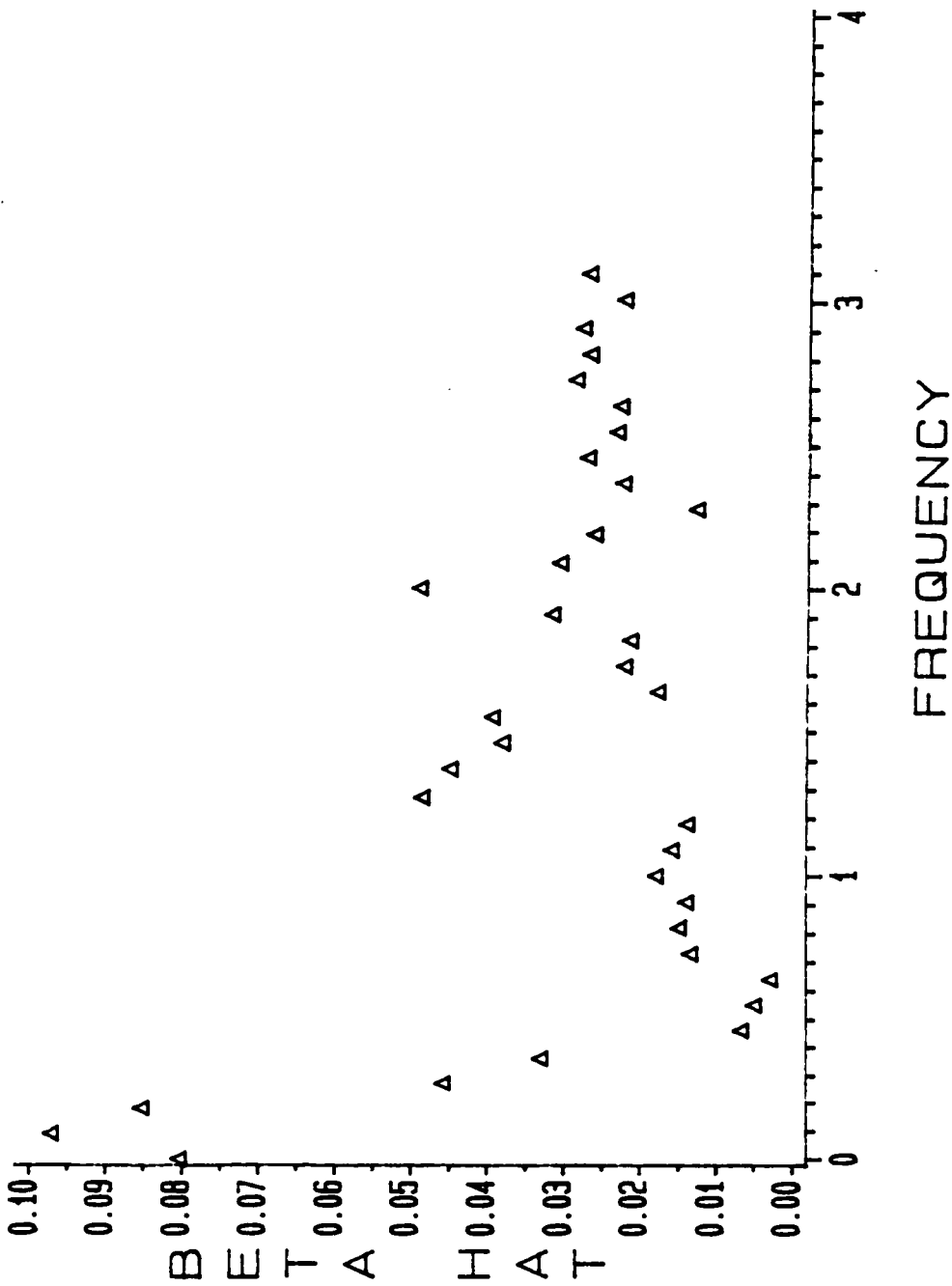
In other words at the 11 year cycle we expect increases in sunspot activity to precede increases in beer drinking by Londoners in the next year following the increase. With the sun apparently entering the most intense period of sunspot activity in 2 centuries (Hilts, 1988), perhaps the London distillers and brewers should begin a corresponding increase in production. Although we must really view this analysis as exploratory and really not an attempt to forecast consumption by sunspots.

The analysis of the this data does raise an interesting issue and that is the topic of partial group delay; i.e., the actual group delay between two time series can actually be a function of a third, completely analogous to partial correlation. In other words what would provide a more meaningful analysis would be to inspect the group delay between London's temperatures and alcohol consumption, with the delay between the two adjusted for sunspots. The topic is a serious one since the possibility that the relationship between the two time series is often affected by other time series related to both of them, is quite significant. Although parital group delay is of considerable interest, it is beyond the scope of the present and is a topic for future research.



UMVUE'S FOR GD OF SUNSPOTS AND CONSUMPTION

Figure VI.22



BETA HAT FOR GQ MODEL OF CONSUMPTION AND SUNSPOTS

Figure VI.23

Chapter VII

Summary And Topics For Future Research.

VII.1. Summary And Conclusions.

The following list is a brief summary of the results of the research and a number of concluding remarks are included in the list.

1. A method has been developed which allows the estimation of expected sample phase and expected sample group delay with small samples.
2. The estimation techniques provide small sample estimates which have theoretical UMVUE properties even with finite sample sizes.

3. A simulation study demonstrated that the theoretical UMVUE techniques are feasible for small sample estimation and that the UMVUE property is an approximate one; because of the presence of sample bias in the simulated estimates of expected sample phase and group delay.
4. When the theoretical UMVUE techniques were applied to two actual problems they generated acceptable estimates of expected sample phase and group delay. Most importantly, the expected sample group delay estimates were always of the correct sign and magnitude for scientific feasibility.
5. It is concluded that the methods developed for estimating expected sample phase and group delay are practical and can easily be used by any time series analyst.

We next discuss a number of proposed topics for future research.

VII.2. Proposed Topics For Future Research.

Although numerous topics for future research have been noted throughout the body of the dissertation, the following list is a summary of those topics deemed most important.

1. A better unwrapping technique is needed for the sample phase values generated by the arc tan method. This would seem a priority research item if the UMVUE technique for expected sample group delay is to be improved significantly.

2. Since we have developed UMVUE's for expected values of sample phase and group delay, it would be beneficial to develop UMVUE's for the variance of the sample phase and group delay.
3. In order to perform exact inference with our UMVUE estimators, it is important to develop an expression for the variance of the UMVUE for the expected sample phase and the expected sample group delay.
4. Another topic that is an important aspect of the inferences to be made with the UMVUE estimators is to develop exact confidence interval expressions for both the expected sample phase and expected sample group delay.
5. A great deal of work needs to be performed to explore the relationship between expected sample values for spectral parameters and the relationship to the true spectral parameters, when only small samples of observations are available for estimation purposes. At present there appears to be no published work on this topic. In the present research the relationship between expected sample phase and the true phase parameter is particularly important and at present it is unknown for finite samples.
6. In the present approach it is important that the Box-Cox transformation exactly produce normality in the distribution of transformed values, since the UMVUE property relies upon the normality assumption to generate UMVUE's. If GLIM procedures could be used, instead of the standard OLS techniques, to estimate the model between transformed phase and frequency, then it would be possible to develop approximate UMVUE's even if the distributions were significantly non-normal. Since the assumption

of transformation to normality is a rather idealistic one, the GLIM approach to estimation may be extremely useful in the context of the present research.

7. The present research concentrated upon estimation of phase and group delay, but it would appear feasible to extend the same approach to the estimation of other spectral parameters such as spectral density. This would be an attractive area of research since present estimation methods rely so heavily upon asymptotic distribution theory.

Many other topics of research will certainly come to mind as one reads through the present research. However, the topics listed here appear to be the ones requiring the most urgent attention, if the proposed approach to small sample spectral estimation is to become an accepted practice in spectral analysis.

BIBLIOGRAPHY

- Andrews, D.F. (1971). A Note On The Selection Of Data Transformations. *Biometrika* 58:2, 249-254.
- Atkinson, A.C. (1973). Testing Transformations To Normality. *J. R. Statist. Soc. B* 35:3, 473-479.
- Akaike, H. and Yamanouchi, Y. (1963). On The Statistical Estimation Of The Frequency Response Function. *Ann. Inst. Statist. Math. (Tokyo)* 14, 23-56.
- Billingsley, P. (1986). *Probability And Measure*. New York: John Wiley & Sons.
- Bickel, P.J. and Doksum, K.A. (1981). An Analysis Of Transformations Revisited. *J. Am. Statist. Assoc.* 76:374, 296-311.
- Bloomfield, P. (1976). *Fourier Analysis Of Time Series: An Introduction*. New York: John Wiley & Sons.
- Boaz, R.P. (1954). *Entire Functions*. New York: Academic Press.
- Bonzanigo, F. (1978). An Improvement Of Tribolet's Phase Unwrapping Algorithm. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-26:1, 104-105.
- Boole, G. (1970). *Finite Difference* (5th ed.). New York: Chelsea Publishing Co.
- Box, G.E.P. (1971). Bias In Nonlinear Estimation. *J. R. Statist. Soc. B* 33, 171-201.
- Box, G.E.P. and Cox, D.R. (1964). An Analysis Of Transformations (With discussion). *J. R. Statist. Soc. B* 26:2, 211-243.

- _____ (1982). An Analysis Of Transformations Revisited, Rebutted. *J. Am. Statist. Assoc.* 77:377, 209-211.
- Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. Oakland, CA: Holden-Day.
- Box, G.E.P. and Tidwell, P.W. (1962). Transformations Of The Independent Variable. *Technometrics* 4, 531-550.
- Brockwell, P.J. and Davis, R.A. (1987). *Time Series: Theory And Application*. New York: Springer-Verlag.
- Carroll, R.J. and Ruppert, D. (1981). On Prediction And The Power Transformation Family. *Biometrika* 68:3, 609-615.
- Chan, Y.T., Hattin, R.V. and Plant, J.B. (1978). The Least Squares Estimation Of Time Delay And Its Use In Signal Detection. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-26:3, 217-222.
- Chan, Y.T., Riley, J.M. and Plant, J.B. (1980). A Parameter Estimation Approach To Time Delay Estimation And Signal Detection. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-28:1, 8-15.
- Chan, Y.T. and Miskowicz, R.K. (1984). Estimation Of Coherence And Time Delay With ARMA Models. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-32:2, 295-303.
- Chatfield, C. (1984). *The Analysis Of Time Series: An Introduction*. New York: Chapman and Hall.
- Cleveland, W.S. and Parzen, E. (1975). The Estimation Of Coherence, Frequency Response And Envelope Delay. *Technometrics* 17:2, 167-172.
- Cobb, J.S. and Phillips, B.F. (1980). *The Biology And Management Of Lobsters, Volume II*. New York: Academic Press.
- Daniell, P.J. (1946). On The Theoretical Specification And Sampling Properties Of Autocorrelated Time Series, by M.S. Bartlett (see discussion following). *J. R. Statist. Soc. Suppl.* 8, 27-41.
- Deaton, M.L. and Foutz, R.V. (1980). Group Delay And The Time-Lag Relationship Between Stochastic Processes. *J. Time Series Anal.* 1, 111-118.
- Dow, R.L. (1969). Cyclic And Geographic Trends In Sea Water Temperature And Abundance Of American Lobster. *Science* 164, 1060-1063.
- _____ (1977). Relationship Of Sea Surface Temperature To American And European Lobster Landings. *J. Cons. Int. Explor. Mer.* 37:2, 186-191.

- _____ (1978). Effects Of Sea Surface Temperature Cycles On Landings Of American, European and Norway Lobsters. *J. Cons. Int. Explor. Mer.* 38:2, 271-272.
- Draper, N.R. and Smith, H. (1981). *Applied Regression Analysis*. New York: John Wiley & Sons.
- Durbin, J. and Watson, G.S. (1951). Testing For Serial Correlation In Least Squares Regression, I. *Biometrika* 38, 159-178.
- Eatwell, J.E., et. al. editors (1987). *The New Palgrave: A Dictionary Of Economics*. New York: Stockton Press.
- Faustino, L., Orach-Meza and Salar, S.B. (1978). Application Of A Polynomial Distributed Lag Model To Tahe Maine Fishery. *Trans. Am. Fish. Soc.* 107:3, 402-411.
- Foutz, R.V. (1980). Estimation Of A Common Group Delay Between Two Multiple Time Series. *J. Am. Statist. Assoc.* 75:372, 779-788.
- Fuller, W.A. (1976). *Introduction To Statistical Time Series*. New York: John Wiley & Sons.
- Gnedenko, B.V. (1967). *The Theory Of Probability*. New York: Chelsea Publishing Co.
- Granger, C.W.J. and Newbold, P. (1976). Forecasting Transformed Series. *J. R. Statist. Soc. B* 38, 189-203.
- Graybill, F.A. (1976). *Theory And Application Of The Linear Model*. Boston: Duxbury Press.
- Hald, A. (1952). *Statistical Theory With Engineering Applications*. New York: John Wiley & Sons.
- Hamon, B.V. and Hannan, E.J. (1974). Spectral Estimation For Dispersive And Non-dispersive Systems. *Appl. Stat.* 23:2, 134-142.
- Hannan, E.J. and Thomson, P.J. (1971). The Estimation Of Coherence And Group Delay. *Biometrika* 58:3, 469-481.
- _____ (1973). Estimating Group Delay. *Biometrika* 60:2, 241-253.
- _____ (1981). Delay Estimation And The Estimation Of Coherence And Phase. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-29:3, 485-490.
- _____ (1988). Time Delay Estimation. *J. Time Series Anal.* 9:1, 21-33.
- Hilts, P.J. (1988). Sunspots May Hit 200 Year Peak, Play Havoc With Communications. *The Washington Post*, Aug 27, col 4, p 4.
- Hinkley, D.V. and Runger, G. (1984). The Analysis Of Transformed Data. *J. Am. Stat. Assoc.* 79:386, 302-309.

- Hinz, P.N., and Eagles, H.A. (1976). Estimation Of A Transformation For The Analysis Of Some Agronomic And Genetic Experiments. *Crop Science* 16, 280-283.
- Hoyle, M.H. (1968). The Estimation Of Variances After Using A Gaussianating Transformation. *Ann. Math. Stat.* 39:4, 1125-1143.
- _____ (1973). Transformations-An Introduction And A Bibliography. *Int. Stat. Rev.* 41:2, 203-223.
- IMSL. (1987). *User's Manual, Stat/Math Library: Fortran Subroutines for Statistical Analysis*. Houston: IMSL.
- Jenkins, G.M. and Watts, D.G. (1968). *Spectral Analysis And Its Applications*. San Francisco: Holden-Day.
- Jevons, W.S. (1884). *Investigations In Currency And Finance* (see the article, Periodicity Of Commercial Crises And Its Physical Explanation). London: MacMillan.
- Kay, S.M. (1988). *Modern Spectral Estimation: Theory And Application*. Englewood Cliffs, NJ: Prentice Hall.
- Kendall, M.G. (1976). *Time Series* (2nd ed.). London: Charles Griffin and Co., Ltd.
- Kerr, R.A. (1988). Sunspot Weather Link Holding Up. *Science* 242 Nov 25, 1122-1124.
- Koopmans, L.H. (1974). *The Spectral Analysis of Time Series*. New York: Academic Press.
- Land, C.E. (1974). Confidence Interval Estimation For Means After Data Transformations To Normality. *J. Am. Stat. Assoc.* 69:3, 795-802.
- Lehmann, E.L. and Scheffe', H. (1950). Completeness, Similar Regions, And Unbiased Estimation, Part I. *Sankhya* 10, 305-340.
- Lewis, R.D. (1988). Private Communication. Department Of Marine Resources, Augusta, ME, 04333.
- Mendenhall, W., Scheaffer, R.L. and Wackerly, D.D. (1986). *Mathematical Statistics With Applications*. Boston: Duxbury Press.
- Menzel, D.H. (1959). *Our Sun*. Cambridge, MA: Harvard Univ. Press.
- Miller, D.M. (1984). Reducing Transformation Bias In Curve Fitting. *The American Statistician* 38:2, 124-126.
- Milton, J.S. and Tsokos, C.P. (1976). *Probability Theory With The Essential Analysis*. Reading, MA: Addison-Wesley Publishing Co.

- Neyman, J. and Scott, E.L. (1960). Correction For Bias Introduced By A Transformation Of Variables. *Ann. Math. Stat.* 31, 643-655.
- Olmsted, J.M.H. (1961). *Advanced Calculus*. New York: Appleton-Century-Crofts, Inc.
- Prest, A.R. (1949). Some Experiments In Demand Analysis. *Rev. Econ. And Stat.* 31, 33-49.
- Priestley, M.B. (1981). *Spectral Analysis And Time Series*. San Diego: Academic Press.
- Rainville, E.D. (1967). *Infinite Series*. New York: MacMillan Co.
- Rao, C.R. (1973). *Linear Statistical Inference And Its Applications*. New York: John Wiley & Sons.
- Richardson, C.H. (1954). *An Introduction To The Calculus Of Finite Differences*. New York: Van Nostrand.
- Schmetterer, L. (1960). On A Problem Of J. Neyman And E. Scott. *Ann. Math. Stat.* 31, 656-661.
- Sky & Telescope (1988). Sunspot Weather Link Found. *Sky & Telescope* 75, March, 243.
- Snee, R.D. (1986). An Alternative Approach To Fitting Models When Re-Expression Of The Response Is Useful. *J. Q. T.* 18:4, 211-225.
- Spitzer, J.J. (1982). A Primer On Box-Cox Estimation. *Rev. Econ. And Stat.* 64, 307-313.
- _____ (1984). Variance Estimates In Models With The Box-Cox Transformation: Implications For Estimation And Hypothesis Testing. *Rev. Econ. And Stat.* 66, 645-652.
- Tribolet, J.M. (1977). A New Phase Unwrapping Algorithm. *IEEE Trans. Acoust., Speech, Signal Processing* Vol. ASSP-25:2, 170-177.

**The two page vita has been
removed from the scanned
document. Page 1 of 2**

**The two page vita has been
removed from the scanned
document. Page 2 of 2**