Statistics Canada Statistique Canada

Methodology Branch

Time Series Research and Analysis Division

Direction de la méthodologie

Division de la recherche et de l'analyse des chroniques

Canada

WORKING PAPER TSRA 85-028E

CAHIER DE TRAVAIL RASC 85-028 E

TIME SERIES RESEARCH AND ANALYSIS DIVISION
METHODOLOGY BRANCH

DIVISION DE LA RECHERCHE
ET DE L'ANALYSE DES CHRONIQUES

MODELS FOR STATIONARY STOCHASTIC PROCESSES

by

Estela Bee Dagum

This is a preliminary version. Do not quote without author's permission.
Comments are welcome.

II. MODELS FOR STATIONARY STOCHASTIC PROCESSES

2.0 Introduction

Two important categories of stochastic processes, the Normal Linear Stationary and the Normal Homogeneous Linear Non-Stationary processes have proved to be the easiest to deal with from a mathematical point of view. Furthermore, they seem to describe quite accurately the generating mechanism of many physical problems. The properties that make these types of processes very useful are that, by the assumption of normality they are fully characterized by their moments of the first and second order and, by being assumed stationary or stationary in the differences (homogeneous non-stationary) the mean and variance are constants and, thus, the autocovariance functions depend only on the time lags. Linear stochastic processes have often been applied to describe phenomena that belong to the natural and social sciences.

The importance of introducing the assumptions of linearity is that all linear stationary processes fulfill the principle of ergodicity, i.e. we can obtain consistent estimators of the mean and covariance functions with only <u>one single realization</u>. In other words, averages computed from one sample (one set of data) called <u>time averages</u> can ultimately be identified with corresponding <u>ensemble averages</u>, that is, averages over the whole span of outcomes or possible realizations of the process. The ergodic property is very important when dealing with time series, the data of which are the result of a single experiment as for example in economics. In the next section we will introduce parametric and non-parametric models for linear stationary stochastic processes. Both forms are statistically equivalent but depending on the problem one form can be more convenient than the other

to apply. For example, the description of the stochastic nature of a time series by means of a parametric model (in the sense of a model with few parameters) is useful for forecast purposes. On the other hand, for frequency response studies, it is more useful to have non-parametric models (in the sense that actually an infinite number of parameters is required to specify the process) such as the autocovariance function or its Fourier transform the spectrum.

Among the classical books that discuss parametric models for stationary and homogeneous stationary stochastic processes, the reader may refer to Anderson (1971), Box and Jenkins (1970), Cramer and Leadbetter (1967), Fuller (1976) and Yaglom (1962). Non-parametric models are treated in detail by Brillinger (1975), Jenkins and Watts (1969), Granger and Hatanaka (1964), Priestley (1981), Koopmans (1974) and Nerlove, Grether and Carvalho (1979).

2.1 Parametric Models of Linear Stationary Stochastic Processes

A stochastic process X_t is said to be linear if there exists a purely random process U_t and a sequence of parameters $\alpha_0, \alpha_1, \alpha_2, \ldots$ such that,

$$(2.1) \quad X_t - m = U_t + \alpha_1 U_{t-1} + \alpha_2 U_{t-2} + \ldots = \sum_{k=0}^{\infty} \alpha_k U_{t-k}; \; \alpha_o = 1$$

In continuous time, equation (2.1) becomes,

$$(2.2) \quad X(t) - m = \int_o^{\infty} \alpha(v) U(t-v) dv.$$

The $E(U_t) = E[U(t)] = 0$ and $E(X_t) = E[X(t)] = m$

If the series or the integral is convergent, then equation (2.1) or (2.2) define a stationary linear stochastic process, where m is the expected

value of the process.

If the series or the integral is not convergent, then equations (2.1) or (2.2) define a non-stationary linear process and m has no precise meaning, except as a point of reference for the level of the process. The equation (2.1) is sometimes called an <u>infinite moving average</u> although the infinite sum of the coeficients is not constrained to be equal to 1.

For a finite number of terms, equation (2.1) defines a <u>linear parametric stochastic process with a discrete time parameter.</u>

In this section we shall deal only with discrete time parameter processes. (The analysis for continuous time follows a similar procedure .)

Using the backshift operator B, where $B^0 = 1$, $BU_t = U_{t-1}$ and $B^nU_t = U_{t-n}$, the equation (2.1) for m=0 can be written in compact form as follows:

$$(2.3) \qquad X_t = (1+\alpha_1 B+\alpha_2 B^2+\ldots)U_t = \sum_{k=0}^{\infty}\alpha_k B^k U_t = \alpha(B)U_t$$

where $\sum_{k=0}^{\infty}\alpha_k B^k = \alpha(B)$ is the linear filter or operator that transforms the input U_t into the output X_t. It is also called the <u>transfer function</u> of the linear system or generating function of the weights. In the latter, B can be considered as a dummy variable whose k-th power is the coefficient of α_k.

For equation (2.3) to be stationary in the wide sense, it is a necessary and sufficient condition that the transfer function $\alpha(B)$ be convergent for $|B| \leq 1$ which implies that $\sum_{k=0}^{\infty}|\alpha|<M$, M finite.

2.1.1 <u>Autoregressive Processes (AR)</u>

Autoregressive processes were first introduced by U. Yule (1921) and

are a subclass of linear processes. A process X_t is said to be an autoregressive process of order p, AR(p) if it can be expressed as a linear combination of p previous values of the process plus a purely random component U_t. It may be written in the following form

$$(2.4)\quad X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_p X_{t-p} = U_t$$

Using the backshift operator B, equation (2.4) results in,

$$(2.5)\quad U_t = (1 + \beta_1 B + \beta_2 B^2 + \ldots + \beta_p B^p) X_t = \sum_{r=0}^{p} \beta_r B^r X_t, = U_t;\ \beta_0 = 1$$

If the X_t process is stationary, then equation (2.5) may be written as an infinite moving average

$$(2.6)\quad X_t = \frac{1}{\sum_{r=o}^{p} \beta_r B^r} U_t = \beta^{-1}(B)\, U_t = \sum_{r=o}^{\infty} \psi_r B^r U_t$$

In effect, each X_{t-m}, m=1,..., p can be replaced by an autoregressive process and finally we end up with an infinite series of U's.

There are several equivalent conditions for the stationarity of the process (2.6). Written as an infinite moving average, we already said in the previous section that the infinite series of weights must be convergent for $|B| \leq 1$, which implies $\sum_{r=o}^{\infty} |\psi_r|$ finite.

These conditions for convergence are equivalent to $\sum_{r=o}^{p} \beta_r B^r \neq 0$

since if $\sum_{r=o}^{p} \beta_r B^r = 0$, then the transfer function becomes infinite.

We shall see which are the conditions for $\sum_{r=o}^{p} \beta_r B^r \neq 0$

$$(2.7)\quad \sum_{r=o}^{p} \beta_r B^r = (1 + \beta_1 B + \beta_2 B^2 + \ldots + \beta_p B^p) = \beta(B)$$

Observe that $\beta(B)$ is a polynomial in B of order p. Factoring out B^p and

making $B = \frac{1}{G}$ we have

$$(2.8) \quad \beta(\frac{1}{G}) = \frac{1}{G^p} \; (G^p + \beta_1 G^{p-1} + \beta_2 G^{p-2} + \ldots + \beta_p) = \frac{1}{G^p} \; (G-G_1)\,(G-G_2)\,\ldots\,(G-G_p)$$

Where $G_1, G_2, \ldots, G_p$ are the p roots of the characteristic equation of $\beta(\frac{1}{G})$.

It is obvious that the equation (2.8) may be written

$$(2.9) \quad \beta(B) = (1-G_1B)(1-G_2B)\;.\;.\;.\;(1-G_pB)$$

Therefore, if there is a $G_i = \frac{1}{B}$, $\beta(B) = 0$, the process (2.6) is non-stationary. for it to be stationary, the roots of $\beta(B)$, namely $B_i = G_i^{-1}$, $i = 1.2, \ldots, p$ must all be in absolute value greater than 1 or equivalently $|G_i| < 1$. The conditions $|G_i| < 1$ coincide with the requirements for the stability of the deterministic component of (2.4). Observe that if $U_t = 0$, then (2.4) reduces to a homogenous finite difference equations of order p, whose general solution is

$$(2.10) \quad X_t = A_1\, G_1^{\,t} + A_2\, G_2^{\,t} + \ldots + A_p G_p^{\,t}$$

Where G_i, $i = 1,2,\ldots,p$ are the roots of its characteristic equation and for the process (2.10) to be stationary, $|G_i| < 1$, for $i = 1,2,..,p$.

For a continuous time parameter, the autoregressive process (2.4) is defined by a differential equation of order p and to be stationary, the roots of its characteristic equation must all have negative real parts.

Autoregressive processes of order 1 and 2 are often used to represent business and economic time series.

The first order autoregressive process also known as Markov process is usually written as,

(2.11) $X_t = \phi_1 X_{t-1} + U_t$

where $-1<\phi_1 < 1$ for the process to be stationary.

Figure 6 shows an AR(1) process generated from the model $X_t=0.60X_{t-1}+U_t$ where $\sigma_u^2 = 1.0$.

The second-order autoregressive process may be written

(2.12) $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + U_t$

For the process (2.12) to be stationary, the roots of

(2.13) $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$

must be greater than one, which implies that the parameters ϕ_1 and ϕ_2 must satisfy the following conditions

$$(2.14) \quad \begin{aligned} &\phi_2 + \phi_1 < 1 \\ &\phi_2 - \phi_1 < 1 \\ &-1 < \phi_2 < 1 \end{aligned}$$

Figure 7 shows an autoregressive process AR(2) generated from the model $X_t = .75X_{t-1} - .50X_{t-2}+U_t$ where $\sigma_U^2 = 1.0$

(place Figures 6 and 7 about here)

2.1.2 Finite Moving Average Processes (MA)

The autoregressive model discussed above, expressed X_t and a finite weighted sum of p previous values of the process X_t, plus a purely random component U_t. Equivalently, if the process is stationary, it can be

expressed as an infinite weighted sum of the U_t's, where $E(U_t) = 0$, $E(U_t)^2 = \sigma_U^2$ and $E(U_t\, U_{t+\tau}) = 0$ for all $\tau \neq 0$.

A process X_t is said to be a <u>finite moving average of order q</u> MA(q) if it is a linear combination of q+1 purely random variables U_t, $U_{t-1}, \ldots, U_{t-q}$. That is,

$$(2.15)\quad X_t = U_t + \alpha_1 U_{t-1} + \alpha_2 U_{t-2} + \ldots + \alpha_q U_{t-2} = \sum_{r=o}^{q} \alpha_r U_{t-r};\ \alpha_o = 1.$$

Using the backshift operator B equation (2.15) becomes,

$$(2.16)\quad X_t = \sum_{r=o}^{q} \alpha_r B^r U_t = \alpha(B) U_t$$

Where $\alpha(B) = 1 + \alpha_1 B + \alpha_2 B^2 + \ldots + \alpha_q B^q$, is the generating function of the weights.

The process (2.15) is always stationary, that is $\alpha(B)$ is convergent for $|B| \leq 1$ but for the model (2.15) to be invertible i.e. to be expressed as an <u>infinite autoregressive process,</u>

$$(2.17)\quad U_t = \left(\sum_{r=o}^{q} \alpha_r B^r\right)^{-1} X_t = \alpha^{-1}(B) X_t = \sum_{r=o}^{\infty} \pi_r B^r X_t$$

the roots of $\alpha(B) = 0$ must lie outside the unit circle.

The first-order moving average process MA(1) is usually written as

$$(2.18)\quad X_t = U_t - \theta_1 U_{t-1} = (1 - \theta_1 B) U_t$$

where $-1 < \theta_1 < 1$ for the process to be invertible. However, the process (2.18) is stationary for all values of θ.

Figure 8 shows a realization of a MA(1) process from $X_t = U_t - .80\, U_{t-1}$ where $\sigma_u^2 = 1.0$.

The second-order moving average process MA(2) may be written

$$(2.19)\quad X_t = U_t - \theta_1 U_{t-1} - \theta_2 U_{t-2} = (1-\theta_1 B-\theta_2 B^2)U_t$$

and is stationary for all values of θ_1 and θ_2. For the process (2.19) to be invertible the roots of the characteristic equation

$$(2.20)\quad 1 - \theta_1 B - \theta_2 B^2 = 0$$

must be greater than one, which implies that the parameter values must satisfy the following conditions

$$(2.21)\quad \begin{aligned} &\theta_2 + \theta_1 < 1 \\ &\theta_2 - \theta_1 < 1 \\ &-1 < \theta_2 < 1 \end{aligned}$$

Figure 9 shows a moving average process MA(2) generated from a model

$$X_t = U_t + .40\ U_{t-1} + .40\ U_{t-2} \text{ where } \sigma_u^2 = 1.0$$

(Place Figures 8 and 9 about here)

2.1.3 Autoregressive-Moving Average Processes (ARMA)

For empirical applications, a combination of an autoregressive process, say of order p with a moving average, say of order q, has the advantage of involving very few parameters.

Thus, an ARMA (p,q) process may be written

$$(2.22)\quad X_t = U_t + \alpha_1 U_{t-1} + \ldots + \alpha_q U_{t-q} - \beta_1 X_{t-1} - \ldots - \beta_p X_{t-p}$$

or equivalently

$$(2.23)\quad (\sum_{r=o}^{p} \beta_r B^r) X_t = (\sum_{r=o}^{q} \alpha_r B^r) U_t; \quad \alpha_o = \beta_o = 1$$

and therefore,

$$(2.24)\quad X_t = \frac{\sum_{r=o}^{q} \alpha_r B^r}{\sum_{r=o}^{p} \beta_r B^r} U_t$$

The ARMA (p,q) process is considered as the output X_t obtained from an input U_t purely random or white noise, where the transfer function is the quotient of two polynomials. The number of parameters of the model (2.22) is p+q+2, including the mean of X_t and the variance of U_t.

For the ARMA (p,q) process (2.24) to be stationary the same conditions discussed in the previous sections are required namely, the roots of the characteristic equation $\sum_{r=o}^{p} \beta_r B^r = \beta(B) = 0$ must all be in absolute value greater than 1. The process (2.24) is invertible if $\sum_{r=o}^{q} {}_r B^r = \alpha(B) = 0$ has all its roots outside the unit circle. Then,

$$(2.25)\quad U_t \quad \frac{\sum_{r=o}^{p} \beta_r B^r}{\sum_{r=o}^{q} \alpha_r B^r} X_t$$

Figures 10 and 11 show an ARMA (1,1) realization from $X_t = 0.60X_{t-1}+U_t - 0.80U_{t-1}$ and an ARMA (2.2) generated from $X_t = .75X_{t-1} - .50X_{t-2} + U_t + .40U_{t-1} + .40U_{t-2}$ where $\sigma_u^2 = 1.0$ in both models.

(place Figures 10 and 11 about here)

2.2 Non-Parametric Models of Stationary Stochastic Processes: The Autocovariance and autocorrelation Functions and their Fourrier Transforms

The models previously discussed are all parametric, in the sense that they have a finite number of parameters. Another way to describe the generating process of a stationary time series is by means of non-parametric models -

models with an infinite number of parameters. Among the non-parametric approaches, the analysis of the autocovariance and autocorrelation functions and their Fourier transforms, the power spectrum also known as the non-normalized spectral density function or simply spectrum and the normalized power spectrum or normalized spectral density function are the most relevant.

From a mathematical point of view, these functions are Fourier pairs and consequently, they are equivalent. Both provide the same type of probability information, in the sense that both characterize all the second order moments of a stationary stochastic process. The use of either the autocovariance function analysis or the spectral analysis depends on the particular properties of the data that one needs to stress. The latter stresses the frequency domain, whereas the former conveys the same information in the time domain.

In the spectral representation, a stationary process is seen as a linear combination of random oscillatory components where the total variance is distributed over frequency. If the process is defined for discrete time parameter, it is possible to determine the proportion of variance attributable to each component with a particular frequency λ, but for time continuous processes, we refer to the contribution of a <u>band of frequencies</u> around a particular λ.

An important use of the autocorrelation function and the normalized spectral density function (normalized spectrum) is to permit the identification of linear filters that minimize the mean square error when the systematic component of a process is corrupted by a purely random

component.

Both functions are also very useful as initial guides in constructing a probability model for the mechanism that has generated the time series. Thus, for example, an autocorrelation function that is positive for successive values of τ (time lag) and tends to zero as τ increases, will reflect both a smooth behaviour of the time series and the fact that the process is more of a finite autoregressive type than purely random. The order of the autoregressive process can also be obtained from the partial autocorrelation function. The same information is given in the frequency domain by a normalized spectral density function with predominancy of low frequencies. See for example Figures 12 and 13 which show the autocorrelation function and the normalized spectrum corresponding to the AR(1) process shown in Figure 6. On the other hand, when adjacent values of $\rho(\tau)$ are negatively correlated, the process generating the time series will show a great fluctuation for short periods of time and the corresponding normalized spectral density function will have predominancy of high frequencies.

(Place Figures 12 and 13 about here)

However, although important for model building (especially in engineering and physics) spectral analysis has shown to be more relevant in frequency response studies and in the area of design of experiments to optimize the performance of industrial processes.

In the analysis of economic time series the first non-parametric approaches were based on the autocovariance and autocorrelation functions. At the end of the fifties and during the decade of the sixties, the latter was almost abandoned and spectral analysis became fashionable. In the decade of the

seventies, however, the analyses in the time domain regained acceptance mainly because of new powerful algorithms and new methods to modelized stochastic processes in a parametric form as those developed by Box and Jenkins (1970).

The autocovariance function of a stationary process X_t is by definition,

$$(2.26) \quad \sigma_{XX}(\tau) = E[\{X_t - \mu\}\{X_{t+\tau} - \mu\}]; \quad \tau = \ldots -1,0,1,\ldots$$

where τ is the time lag, assumed here to be an integer and μ is the mean. If the time parameter of the process is continuous, then τ can assume any value between $\pm\infty$.

There are situations in which it is necessary to compare time series which have different scales of measurement and in this case, it is useful to normalize the autocovariance function dividing by the variance of the process. This function is called the autocorrelation function. That is,

$$(2.27) \quad \rho_{XX}(\tau) = \frac{\sigma_{XX(\tau)}}{\sigma_{XX(0)}} = \frac{\sigma_{XX(\tau)}}{\sigma_X^2} \qquad \tau = \ldots, -1,0,1,\ldots$$

Observe that,

$$(2.28) \quad \sigma_{XX}(0)\, \rho_{XX}(\tau) = \sigma_{XX}(\tau)$$

and therefore, if we know the autocorrelation function and the variance of the process X_t we have all the information provided by the autocovariance.

The graph of equation (2.27) is also known as the <u>correlogram</u>. The basic properties of the autocorrelation function for a real process are: (We shall suppress the subindex X to abbreviate the notation):

(1) $\rho(0)=1$

(2) $\rho(-\tau) = \rho(\tau)$ The function is symmetric with respect to the origin because of the stationarity assumption and therefore it needs only to be calculated for positive lags.

(3) $|\rho(\tau)| \leq 1$. This is a consequence of the fact that the variance of a random variable or of linear combination of random variables is positive. In effect, assume $Y_t=\lambda_1 X_t+\lambda_2 X_{t-\tau}$ then the variance of Y_t is

(a) var. $Y_t = \lambda_1^2$ var. $X_t + \lambda_2^2$ var. $X_{t-\tau} + 2\lambda_1\lambda_2$cov. $(X_t X_{t-\tau})$. The right member is non-negative for all λ_1, λ_2 real and the second member is a quadratic form in λ_1, λ_2. For it to be positive, its roots must be imaginary, which implies

(b) var. X_t var. $X_{t-\tau} \geq [\text{cov}(X_t X_{t-\tau})]^2$

or equivalently

$$\text{(c)} \quad \rho^2(X_t, X_{t-\tau}) = \frac{[\text{cov}\ (X_t X_{t-\tau})]^2}{\text{var}\ X_t\ \text{var}\ X_{t-\tau}} \leq 1$$

For a stationary process, (3) reduces to,

$$|\rho(\tau)| = \left|\frac{\sigma(\tau)}{\sigma(o)}\right| \leq 1$$

(4) The autocorrelation matrix is positive semi-definite. That is, the determinant of the autocorrelation matrix and all its principal minors are positive or zero. Property 4 is a generalization of property 3 and shows that the autocorrelation function is always positive semi-definite. The converse is also true; that is, every positive semi-definite function of a real (or integral) argument is the autocorrelation function of a continuous (discrete) stochastic process. (This was proven by Khintchine and Kolmogorov, see Yaglom, 1962).

(5) if the process is continuous, then $\rho(\tau)$ is defined for τ taking values between $+\infty$ and $-\infty$ and it is a necessary and sufficient condition that the function be continuous at $\tau = 0$ since this implies continuity everywhere (Yaglom, 1962). If the process is assumed to be purely random, this continuity property poses problems.

For a discrete purely random process U_t, the autocorrelation function is $\rho(0) = 1$ and $\rho(\tau) = 0$, for all $\tau \neq 0$. If U_t is a time continuous process we would have a discontinuity at $\tau = 0$. To avoid this, the autocovariance function is redefined as

$$(2.29)\quad q_{UU}(\tau) = \sigma(0)\,\delta(\tau) = \sigma_U^2\,\delta(\tau)$$

Where $\delta(\tau)$ is a Dirac delta or impulse function, interpreted as zero for $\tau \neq 0$ and infinite for $\tau = 0$, then the covariance between neighbouring points is zero but at the expense of making the variance of the process infinite (Jenkins and Watts, 1969). A delta function is defined as a sequence of function $\delta_n(t)$ such that $\int_{\infty}^{\infty}\delta_n(t)\,dt = 1$, for every n and in the limit as n tends to ∞

$$\delta(t) = \begin{matrix} o & t \neq 0 \\ \infty & t = 0 \end{matrix}$$

Since any analytical function, periodic or not, can be approximated to any degree using any class of periodic functions, the <u>spectral representation</u> of a stochastic process can be done using Fourier series or Fourier integrals, depending on the time parameter being discrete or continuous. In Fourier analysis, the periodic functions are sines and cosines. They have the important properties that an approximation of a given number of terms gives the minimum mean square error between the function and its approximation, and also that they are orthogonal, so the coefficients may

be determined independently of one another.

The use of Fourier series to describe phenomena evolving through time was suggested in several studies by Lagrange (1772-78), Buys-Ballot (1847) and Stokes (1879), but the best known work was the periodogram method used by Schuster (1898) in the search of hidden periodicities in sunspot data. In economics, the periodogram was used by Moore (1914) and by Beveridge (1922). The use of the periodogram to describe time series failed because of the assumptions of fixed amplitudes, frequencies and phases in the Fourier components. The modern spectral analysis uses the Fourier series (or Fourier integral) assuming that the amplitudes and phases are random variables.

It is shown (see, Yaglom, 1962) that every stationary stochastic process $X(\omega,t)$ can be approximated by a linear combination or harmonic oscillations of form

$$(2.30)\quad X(\omega,t) = \sum_k X_k(\omega)f(t) = \sum_k X_k(\omega)Re^{i(\lambda_k t+\theta)}$$

where $X_k(\omega)$ is a time independent random variable and f (t) is a numerical function of t. The numerical factor $Re^{i\theta}$ can be included in the random variable $X_k(\omega)$ and the product $X_k(\omega)Re^{i\theta}$ will be simply denoted here by X_k;then equation (2.30) becomes,

$$(2.31)\quad X_k(t) = \sum_k X_k\, e^{i\lambda_k t} = \sum_k X_k(\cos\lambda_k t + i\sin\lambda_k t)$$

where X_k is a complex random variable with mean value zero, and λ is a constant.

Then, each component of the form (2.31) describes a periodic oscillation of angular frequency λ_k with random amplitude R and random phase θ. The

angular frequency $\lambda = 2\Pi f = 2\Pi\frac{1}{T}$, is the number of cycles around the unit circle per unit of time; f is simply the frequency and it is the reciprocal of the period T or length of time required for one complete oscillation.

If the process is defined for a discrete time parameter, we can represent it by

$$(2.32) \quad X_t = \sum_{k=1}^{\infty} X_k e^{i\lambda_k t}$$

and for continuous time parameter by,

$$(2.33) \quad X(t) = \int_{\infty}^{\infty} e^{i\lambda t} dZ(\lambda)$$

where $Z(\lambda)$ is a stochastic process indexed on λ.

Equation (2.32) is the <u>spectral representation</u> of a stationary process with a <u>discrete spectrum</u> and the set of numbers $\{\lambda_1, \lambda_2 \ \ldots\}$ is called the <u>spectrum</u> of the process.

Equation (2.33) is the spectral representation of a stationary process, where $Z(\lambda)$ is a continuous spectrum with mean value zero and <u>uncorrelated increments</u>. The possibility of such representations for arbitrary stationary processes was first shown by Kolgomorov (see Yaglom, 1962).

The spectral representation of a stationary process is then a "decomposition" of the process into separate pairwise uncorrelated periodic oscillations. It is possible to separate spectral components corresponding

to different parts of the spectrum by using suitable chosen linear operators or filters. In engineering, a filter is a device which passes harmonic oscillations in certain frequency range (the pass band) while suppressing oscillations with different frequencies. In practice, the filters used are of three types, the low-pass filters, passing all oscillations with frequencies less than a certain critical frequency λ_0, the high-pass filters, passing all oscillations greater than λ_0 and the band-pass filters passing only oscillation with frequencies that belong to a given interval (pass-band) $[\lambda_0, \lambda_1]$.

We shall see now, that the information contained in the autocovariance function is equivalent to the one given by its Fourier-Stieljes transform, the power spectrum or non-normalized spectral distribution function.

Since any arbitrary stationary process can have a spectral representation, its corresponding autocovariance function can also be expressed in the spectral form.

Thus, for a process such as (2.32), which is assumed stationary and therefore $E(X_k \bar{X}_\ell)=0$, $k \neq \ell$ (by $\bar{X}$ we denote the conjugate of X), the autocovariance function is,

$$(2.34) \quad \sigma(\tau) = \sum_{k=1}^{\infty} E|X_k|^2 e^{i\lambda_k \tau} = \sum_{k=1}^{\infty} b_k e^{i\lambda\tau}; \quad b_k > 0$$

The autocovariance function (2.34) exists if the series is convergent, that is, if

$$(2.35) \quad \sum_{k=1}^{\infty} E|X_k|^2 = \sum_{k=1}^{\infty} b_k < \infty \quad ; \qquad b_k > 0$$

It was shown by Slutsky (1938) that the converse is true, every stationary stochastic process with autocovariance function of the form (2.34) can be

represented in form of (2.32) with $E(X_k\bar{X}_\ell) = 0$ for $k=\ell$.

Setting $\tau=0$, equation (2.34) becomes,

$$(2.36)\quad \sigma(0) = \sum_{k=1}^{\infty} E|X_k|^2 = \sum_{k=1}^{\infty} b_k \quad ; \quad b_k>0$$

which shows that in the superposition of uncorrelated periodic oscillations, the total variance of the process is equal to the sum of the variances of the separate periodic components.

Observe that variances b_k of the separate periodic components are the mean values of the squares of the amplitude X_k of the harmonic components $X_k e^{i\lambda k\tau}$ of the process X_t.

The representation (2.34) was generalized by Khintchine (1934) who proved that the autocovariance function of any stationary stochastic process can be represented in the form of an integral,

$$(2.37)\quad \sigma(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dG(\lambda)$$

where $G(\lambda)$ is the integrated spectrum or also called spectral distribution function because it is equivalent to a probability distribution function and it is the Fourier-Stieljes transform of the autocovariance function $\sigma(\tau)$. The spectral distribution function is a monotonically non-decreasing function, symmetric with respect to the origin and bounded $G(-\infty)=o$ and $G(\infty)=\sigma(0)$. When $G(\lambda)$ is normalized, that is, divided by the variance, then $F(\lambda)=\frac{G(\lambda)}{\sigma(0)}$ is called the normalized spectral distribution function which is the Fourier-Stieljes transform of the autocorrelation function $\rho(\tau)$. That is,

$$(2.38)\quad \rho(\tau) = \int_{-\infty}^{\infty} e^{i\lambda\tau} dF(\lambda)$$

The normalized spectral distribution function $F(\lambda)$ is also non-decreasing, symmetric with respect to the origin and bounded $F(-\infty)=0$ and $F(\infty)=1$. It can be decomposed as,

$$(2.39)\quad F(\lambda)=F_1(\lambda)+F_2(\lambda)+F_3(\lambda)$$

where $F_1(\lambda)$, $F_2(\lambda)$ and $F_3(\lambda)$ are each non-decreasing, $F_1(\lambda)$ is a pure step function, $F_2(\lambda)$ is absolutely continuous, that is, $F_2(\lambda)=\int_{-\infty}^{\infty}F_2'(u)\,du$ and $F_3(\lambda)$ is a singular function, continuous and with $F_3'(\lambda)=0$ almost everywhere.

Thus $F(\lambda)$ can be seen as a distribution function and $\rho(\tau)$ as its characteristic function. Since $G(\lambda)$ and $F(\lambda)$ are odd functions, for every real process the (2.37) and (2.38) are real integrals and can be written as

$$(2.40)\quad \sigma(\tau)=\int_{-\infty}^{\infty}\cos\lambda\tau\, dG(\lambda)=\int_{0}^{\infty}\cos\lambda\tau\, dG_1(\lambda)$$

where $G_1(\lambda)=2G(\lambda)$ and

$$(2.41)\quad \rho(\tau)=\int_{-\infty}^{\infty}\cos\lambda\tau\, dF(\lambda)=\int_{0}^{\infty}\cos\lambda\tau\, dF_1(\lambda)$$

where $F_1(\lambda)=2F(\lambda)$.

When $F(\lambda)$ and $G(\lambda)$ have derivatives (which are the interesting cases)

$$(2.42.a)\quad dF(\lambda)=f(\lambda)d\lambda$$

$$(2.42.b)\quad dG(\lambda)=g(\lambda)d\lambda$$

then $f(\lambda)$ is the normalized spectral density function or normalized spectrum, and $g(\lambda)$ the non-normalized spectral density function or power spectrum. $\sigma(\tau)$ and $\rho(\tau)$ are the inverse Fourier transforms of $g(\lambda)$ and

$f(\lambda)$ respectively.

For τ discrete, the normalized spectral density function $f(\lambda)$ is the Fourier transform of a sequence of autocorrelations and we have

$$(2.43) \quad f(\lambda) = \frac{1}{2\Pi} \sum_{\tau=-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau}; \qquad -\Pi \leq \lambda \leq \Pi$$

and for τ a continuous time parameter,

$$(2.44) \quad f(\lambda) = \frac{1}{2\Pi} \int_{-\infty}^{\infty} \rho(\tau) e^{-i\lambda\tau} d\tau: \qquad -\infty \leq \lambda \leq \infty$$

Similarly the spectrum $g(\lambda)$ for a discrete process is

$$(2.45) \quad g(\lambda) = \frac{1}{2\Pi} \sum_{\tau=-\infty}^{\infty} \sigma(\tau) e^{-i\lambda\tau}; \quad -\Pi \leq \lambda \leq \Pi$$

and for a continuous process,

$$(2.46) \quad g(\lambda) = \frac{1}{2\Pi} \int_{-\infty}^{\infty} \sigma(\tau) e^{-i\lambda\tau} d\tau \qquad -\infty \leq \lambda \leq \infty$$

Since $F(\lambda) = \int_{-\infty}^{\lambda} f(u)du$, integrating (2.43) and (2.44) we obtain the normalized spectral distribution function $F(\lambda)$.

For a real process X_t, the (2.43) reduces to

$$(2.47) \quad f(\lambda) = \frac{\rho(o)}{2\Pi} + \frac{1}{\Pi} \sum_{\tau=1}^{\infty} \rho(\tau) \cos\lambda\tau = \frac{1}{2\Pi} \sum_{\tau=-\infty}^{\infty} \rho(\tau) \cos\lambda\tau; \quad -\Pi \leq \lambda \leq \Pi$$

and the (2.44) takes the form,

$$(2.48) \quad f(\lambda) = \frac{1}{2\Pi} \int_{-\infty}^{\infty} \rho(\tau) \cos\lambda\tau d\tau; \qquad -\infty \leq \lambda \leq \infty$$

Then, the corresponding normalized spectral distribution functions are:

$$(2.49) \quad F(\lambda) = \frac{\rho(o)}{2\Pi} + \frac{1}{\Pi} \sum_{\tau=-\infty}^{\infty} \frac{\rho(\tau) \sin\lambda\tau}{\tau}; \qquad -\Pi \leq \lambda \leq \Pi$$

and

$$(2.50)\quad F(\lambda) = \frac{1}{2\Pi}\int_{-\infty}^{\lambda}\int_{-\infty}^{\infty}\rho(\tau)\cos\lambda\tau d\tau d\lambda; \qquad -\infty\leq\lambda\leq\infty$$

A similar procedure is followed to obtain $G(\lambda)$.

In the next section we deal with the autocovariance and autocorrelation functions and the spectrum and normalized spectral density corresponding to the linear processes in (2), (2.1), (2.2) and (2.3), and illustrate with some theoretical examples.

2.3 AUTOCOVARIANCE AND AUTOCORRELATION FUNCTIONS OF LINEAR STATIONARY PROCESSES AND THEIR FOURIER TRANSFORMS

We saw in Section 2.1 that a linear stationary process X_t can be interpreted as the output obtained from an input U_t (a purely random process) that has been passed through a transfer function $\alpha(B)$ which is a convergent infinite sum of weights α for $|B|\leq 1$. That is,

$$(2.51)\quad X_t = \sum_{k=o}^{\infty}\alpha_k U_{t-k} = \sum_{k=o}^{\infty}\alpha_k B^k U_t = \alpha(B)U_t;\ \sum_{k=o}^{\infty}|\alpha_k| < M; \quad M \text{ finite}$$

Then, according to the definition of the autocovariance function (2.26) and remembering that the process is assumed to be stationary with zero mean, we obtain,

$$(2.52)\quad \sigma_{XX}(\tau) = E(X_t X_{t+\tau}) = \sigma_U^2 \sum_{k=o}^{\infty}\alpha_k\alpha_{k+\tau}$$

The variance of X_t is then,

$$(2.53)\quad \sigma_{XX}(o) = \sigma_U^2 \sum_{k=o}^{\infty}\alpha_k^2$$

and the autocorrelation function is

$$(2.54)\quad \rho_{XX}(\tau) = \frac{\sigma_{XX}(\tau)}{\sigma_{XX}(o)} \qquad \frac{\sum_{k=o}^{\infty}\alpha_k\alpha_{k+\tau}}{\sum_{k=o}^{\infty}\alpha_k^2}$$

The autocovariance function can be obtained in an easier way using the autocovariance generating function, which also can be used to obtain the spectrum of the process.

The autocovariance generating function is

$$(2.55) \quad \sigma_{XX}(B) = \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) B^{\tau}, \qquad \tau = o, \pm 1, \pm 2, \ldots$$

Since for a stationary process $\sigma_{XX}(\tau)$ is an even function, then $\sigma_{XX}(k) = \sigma_{XX}(-k)$, is the coefficient of B^k and B^{-k}.

For the infinite moving average linear process (2.51) the autocovariance generating function is shown to be

$$(2.56) \sigma_{XX}(B) = \sigma_U^2 \alpha(B) \alpha(B^{-1})$$

For $B = e^{-i\lambda}$, the equation (2.55) becomes

$$(2.57) \sigma_{XX}(B) = \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) e^{-i\lambda\tau} ; \qquad \tau = 0, \pm 1, \pm 2, \ldots$$

Comparing (2.57) with the spectrum $g_{XX}(\lambda) = \frac{1}{2\Pi} \sum_{\tau=-\infty}^{\infty} \sigma_{XX}(\tau) e^{-i\lambda\tau}$, we see that $2\pi g_{XX}(\lambda) = \sigma_{XX}(B)$. If we limit λ to be non-negative, then, $\pi g_{XX}(\lambda) = \sigma_{XX}(B)$. Therefore, multiplying the autocovariance generating function by $\frac{1}{\pi}$ gives us the power spectrum of the process. Then, using equation (2.56), the spectrum of the linear process (2.51) can also be written in the form,

$$(2.58) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\pi} |\alpha(B)|^2 ; \qquad -\pi \leq \lambda \leq \pi$$

where $|\alpha(B)|^2 = |\alpha(e^{-i\lambda})|^2$ is called the filter gain and it is the square of the filter transfer function. Equation (2.58) shows that the spectrum of

the output X_t of the linear process (2.51) can be obtained from the constant uniform spectrum $\frac{\sigma_U^2}{\pi}$ of a white noise process U_t multiplied by a factor (the filter gain) that depends only on the characteristics of the filter. The frequencies for which $|\alpha(e^{-i\lambda})|^2$ is large are magnified and those for which the filter gain is small are reduced. The variance of X_t is,

$$(2.59) \quad \sigma_U^2 = \int_o^\pi g_{XX}(\lambda)d\lambda = \frac{\sigma_U^2}{\pi}\int_o^\pi |\alpha(e^{-i\lambda})|^2 \, d\lambda$$

Dividing equation (2.58) by the variance (2.59) we obtain the normalized spectral density function $f_{XX}(\lambda)$ which is the Fourier transform of the autocorrelation function $\rho_{XX}(\tau)$. Then,

$$(2.60) \quad f(\lambda) = \frac{|\alpha(e^{-i\lambda})|^2}{\int_o^\pi |\alpha(e^{-i\lambda})|^2 d\lambda} \quad ; \qquad o \leq \lambda \leq \pi$$

It is easy to show that the autocovariance and autocorrelation functions of the autoregressive process AR(p) (2.4) satisfy the same form of difference equation corresponding to the deterministic part of it. In effect, multiplying (2.4) by $X_{t-\tau}$ and using (2.6) we obtain,

$$(2.61) \quad \sum_{r=o}^{p} \beta_r X_{t-r} X_{t-\tau} = \sum_{r=o}^{\infty} \Psi_r U_t U_{t-\tau-r}$$

Since $E(X_{t-r}X_{t-\tau}) = \sigma_{XX}(\tau-r)$; $E(U_t^2) = \sigma_U^2$; $E(U_tU_s) = 0$ for all $t \neq s$; the expected values of the two sides of (2.61) satisfy for $\tau = o$ and for $\tau > o$, respectively,

$$(2.62) \quad \sum_{r=o}^{p} \beta_r \sigma_{XX}(-r) = \sigma_U^2$$

and

$$(2.63) \quad \sum_{r=o}^{p} \beta_r \sigma_{XX}(\tau-r) = o \qquad \tau = 1, 2, \ldots$$

These are often called the Yule-Walker equations. Thus, the sequence

$\sigma_{XX}(\tau-1)$, $\sigma_{XX}(\tau-2)$, ..., $\sigma_{XX}(\tau-p)$, $\tau = 1, 2, ...$

satisfies the homogeneous difference equation (2.63). Dividing equation (2.63) throughout by σ_X^2 , we obtain the autocorrelation function which also satisfies a homogeneous difference equation analogous to the one of the process X_t itself. We can write equation (2.63) using the backshift operator B in the form of,

$$(2.64) \qquad \sum_{r=o}^{p} \beta_r B^r \sigma_{XX}(\tau) = \beta(B)\sigma_{XX}(\tau) = o; \qquad \tau = 1, 2, ...$$

and equivalently for the autocorrelation function,

$$(2.65) \qquad \sum_{r=o}^{p} \beta_r B^r \rho_{XX}(\tau) = \beta(B)\rho_{XX}(\tau) = o; \qquad \begin{array}{l} \tau = 1, 2, ... \\ \beta o = 1 \end{array}$$

where B operates on τ.

The same conditions for stationarity that were required for the finite autoregressive processes of the form (2.4) apply here. If all the roots $|G_i| < 1$ are distinct, we have two situations:

(a) A root G_i is real, in which case $A_i G_i^t$ (see equation 2.10) decreases geometrically to zero as t increases. If the root is positive, we will have a decreasing exponential function; and, if it is negative, we will have an exponential function alternating in sign and decreasing in absolute value.

(b) A pair of complex conjugate roots G_i, G_j, in which case they generate a term that is an oscillating trigonometric function, decreasing in absolute value, and whose period of oscillation depends on the argument of the complex roots.

The variance $\sigma_{XX}(o)$ of an autoregressive process X_t can be obtained from equation (2.62) and also can be expressed in terms of the autocorrelation function by dividing (2.62) by $\sigma_{XX}(0)$ and making $\sigma_{XX}(-r)=\sigma_{XX}(r)$.

Then,

$$(2.66) \quad \sigma_X^2 = \frac{\sigma_U^2}{1+\beta_1\rho(1)+\beta_2\rho(2)+\ldots+\beta_p\rho(p)}$$

The spectrum $g_{XX}(\lambda)$ for the AR(p) process can be obtained using equation (2.58) where the filter gain is $|\beta(e^{-i\lambda})|^{-2}$.

Then,

$$(2.67) \quad g_{XX}(\lambda) = \frac{\sigma_U^2}{\Pi|\beta(e^{-i\lambda})|^2} = \frac{\sigma_U^2}{\Pi|1+\beta_1 e^{-i\lambda}+\ldots+\beta_p e^{-ip\lambda}|^2} ; \qquad \sigma\leq\lambda\leq\Pi$$

The normalized spectral density function $f(\lambda)$ is obtained dividing $g_{XX}(\lambda)$ by σ_X^2.

Figures 14 and 15 show the autocorrelation function and the normalized spectral density function, respectively, of the autoregressive process AR(2) of Figure 7.

(Place figures 14 and 15 about here)

Though the autocorrelation function of an AR(p) process is infinite in extent, by its own nature it can be described in terms of p non-zero functions of the autocorrelations. This information is provided by the partial autocorrelation function which helps to determine the order of an autoregressive process to fit to an observed time series. For an autoregressive process of order p, the partial autocorrelation function has a cutoff after the p lag.

Denoting by β_{kj} the jth coefficient in an AR(k) process so that β_{kk} is the last coefficient, then from equation (2.65) β_{kk} satisfies the set of

equations.

$$(2.68)\quad \rho(j)=\beta_{k1}\rho(j-1)+\beta_{k2}\rho(j-2)+\ldots+\beta_{kk}\rho(j-k); \qquad j=1,2,\ldots k$$

leading to the Yule-Walker equations, which may be written as;

$$(2.69)\quad \begin{matrix} 1 & \rho & \ldots & \rho(k-1) \\ \rho & 1 & \ldots & \rho(k-2) \\ \cdot & \cdot & & \cdot \\ \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & & 1 \end{matrix} \begin{matrix} \beta_{k1} \\ \beta_{k2} \\ \cdot \\ \vdots \\ \beta_{kk} \end{matrix} = \begin{matrix} \rho(1) \\ \rho(2) \\ \cdot \\ \vdots \\ \rho(k) \end{matrix}$$

Solving the system (2.69) for k=1, 2, 3, ... we obtain the partial autocorrelations β_{11}, β_{22}, β_{33},.... In general, β_{kk} is the partial autocorrelation function of the lag k.

For the finite moving average process of order q (MA-q) (2.15) the autocovariance function is

$$(2.70)\quad \sigma_{XX}(\tau)=E(X_tX_{t+\tau})=\sigma_U^2\sum_{k=o}^{q-\tau}\alpha_k\alpha_{k+\tau}\ ; \qquad \tau\leq q$$

and

$$\sigma_{XX}(\tau)=o; \qquad \tau>q$$

Then the variance of the MA(q) process is

$$(2.71)\qquad \sigma_X^2=\sigma_U^2\sum_{k=o}^{q}\alpha_k^2$$

and the autocorrelation function is

$$(2.72)\quad \rho_{XX}(\tau)=\frac{\sum_{k=o}^{q-\tau}\alpha_k\alpha_{k+\tau}}{\sum_{k=o}^{q}\alpha_k^2}\ ; \qquad \tau\leq q$$

and

$$\sigma_{XX}(\tau)=o; \qquad \tau>q$$

Consequently, the correlogram of a MA(q) process is zero for $\tau=q+1$ onwards.

According to equation (2.61), the spectrum of a MA(q) process is then

$$(2.73) \quad g_{XX}(\lambda)=\frac{\sigma_U^2}{\Pi}|\alpha(e^{-i\lambda})|^2=\frac{\sigma_U^2}{\Pi}\left|\sum_{k=o}^{q}\alpha_k e^{-i\lambda k}\right|^2 ; \quad \alpha_{o=1}; \quad o\leq\lambda\leq\Pi$$

and the normalized spectral density function $f_{XX}(\lambda)$ is $g_{XX}(\lambda)/\sigma_X^2$.

Figures 16 and 17, 18 and 19 show the autocorrelation function and the normalized spectral density of the moving average process MA(1) of Figure 8. and MA(2) of Figure 9, respectively.

(Place Figures 16, 17, 18 and 19 about here)

For the ARMA(pq) process X_t defined in (2.22), the $\sigma_{XX}(\tau)$, $\rho_{XX}(\tau)$, $g_{XX}(\lambda)$ and $f_{XX}(\lambda)$ can be obtained in a similar way. Thus,

$$(2.74) \quad \sigma_{XX}(\tau)=\sigma_{XU}(\tau)+\alpha_1\sigma_{XU}(\tau-1)+\ldots+\alpha_q\sigma_{XU}(\tau-q)-\beta_1\sigma_{XX}(\tau-1)-\ldots-\beta_p\sigma_{XX}(\tau-p)$$

where $\sigma_{XU}(\tau)$ is the cross covariance function between X_t and U_t and is defined by $\sigma_{XU}(\tau)=E(X_{t-\tau}U_t)$.

Since $X_{t-\tau}$ depends only on random components which have occurred up to time $t-\tau$ uncorrelated with U_t it follows that $\sigma_{XU}(\tau)=o$ for $\tau>o$ and $\sigma_{XU}(\tau)\neq o$ for $\tau\leq o$.

The (2.74) reduces to,

$$(2.75) \quad \sigma_{XX}(\tau)=-\beta_1\sigma_{XX}(\tau-1)-\ldots-\beta_p\sigma_{XX}(\tau-p); \qquad \tau\geq q+1$$

Hence, the autocorrelation function is:

$$(2.76) \quad \rho_{XX}(\tau)=-\beta_1\rho_{XX}(\tau-1)-\ldots-\beta_p\rho_{XX}(\tau-p), \qquad \tau\geq q+1$$

or

$$\beta(B)\rho_{XX}(\tau)=o; \qquad \tau\geq q+1$$

where B operates on τ.

Then for an ARMA (p,q) process there will be q autocorrelations whose values depend on the choice of the q moving average parameters α as well as on the p autoregressive parameters β. Now for $\tau>q+1$, the p autocorrelations already obtained provide the initial values for the homogeneous difference equation $\beta(B)\rho_{XX}(\tau)=o$ which then entirely determines the autocorrelations of higher lags.

If $q<p$, the whole autocorrelation function will consist of a mixture of damped exponentials and/or damped sine functions. If $q\geq p$, the q-p+1 autocorrelations used as initial values will not follow this general pattern.

For $\tau=0$, equation (2.74) gives the variance of the process

$$(2.77) \quad \sigma_X^2=\sigma_U^2+\alpha_1\sigma_{XU}(-1)+\ldots+\alpha_q\sigma_{XU}(-q)-\beta_1\sigma_{XX}(1)-\ldots-\beta_p\sigma_{XX}(p)$$

which has to be solved along with the p equations (2.74) for $\tau=1, 2, \ldots p$ to obtain $\sigma_X^2, \sigma_{XX}(1), \ldots, \sigma_{XX}(p)$.

The spectrum of the process is

$$(2.78) \quad g_{XX}(\lambda)=\frac{\sigma_U^2}{\Pi}\cdot\frac{\left|\sum_{r=o}^{q}\alpha_r e^{-i\lambda r}\right|^2}{\left|\sum_{k=o}^{p}\beta_r e^{-i\lambda r}\right|^2}; \qquad o\leq\lambda\leq\Pi$$

and the normalized spectral density function $f_{XX}(\lambda)$ is the quotient between $g_{XX}(\lambda)$ and σ_X^2.

Figures 20 and 21 show the autocorrelation function and the normalized spectral density function of the ARMA(1,1) process of Figure 10. Similarly, Figures 22 and 23 show the autocorrelation function and normalized spectrum of the ARMA(2,2) process of Figure 11.

(place Figures 20, 21, 22 and 23 about here)

2.3.1 Relation Between the Autocorrelation Function and the Normalized Spectral Density Function

We shall now discuss for some particular cases the relationship between the autocorrelation function and the normalized spectrum.

Consider an autocorrelation function of the form

$$\rho(\tau)=e^{-\beta|\tau|} \tag{2.79}$$

where $o<\beta<1$ serves as a measure of the rapidity of decrease of the $\rho(\tau)$ with the increase of the time lag τ. We see in Figure 24, that the greater the β, the more damped the autocorrelation function, thus implying a less smooth stochastic process

(Place Figure 24 about here)

The corresponding normalized spectral density function $f(\lambda)$ is,

$$f(\lambda)=\frac{1}{2\Pi}\int_{-\infty}^{\infty}\rho(\tau)e^{-i\lambda\tau}d\tau = \frac{1}{2\Pi}\int_{-\infty}^{\infty}e^{-i\lambda\tau-\beta|\tau|}d\tau = \frac{1}{\Pi}\frac{\beta}{\lambda^2+\beta^2} \tag{2.80}$$

and it gives the same information contained in (2.79) but in the frequency domain. The normalized spectral density function is shown in Figure 25.

(Place figure 25 about here)

We can see that for small β the normalized spectrum has predominancy of low frequencies implying a smooth process, whereas as β increases, the curve is compressed toward the λ - axis, at the same time becoming flatter. This kind of behaviour of the function $f(\lambda)$ enables us to illustrate a purely random process or white noise, the normalized spectral density of which is a constant equal to $\frac{1}{2\Pi}$ for $-\Pi \leq \lambda \leq \Pi$. Observe that the ordinate of $f(\lambda)$ at the origin is $\frac{1}{\Pi\beta}$ and as β increases, the interception decreases. In fact, it is assumed that β can assume very high values and the $\rho(\tau)$ is transformed into a spike-shaped function, different from zero only in a very small neighbourhood around $\tau=0$.

The total variance of the process is the area under $f(\lambda)$. For small β, a low frequency band accounts for most of the variance whereas for large β the variance is distributed almost uniformly in the frequency band capable of exerting an effect on the process under consideration. It is impossible for absolutely white noise to exist, since for the spectral density to be constant in the whole range of variations of λ, the autocovariance function for $\tau=0$, would have an infinite variance which cannot take place in any real process. In effect.

$$(2.81) \qquad \sigma_{XX}(o)=\int_{-\infty}^{\infty} g(\lambda)\,d\lambda = c\int_{-\infty}^{\infty} d\lambda = \infty$$

Let us now consider an autocorrelation function $\rho(\tau)$ of the form

$$(2.82) \qquad \rho(\tau)=e^{-\beta|\tau|}\cos\omega\tau$$

which differs from (2.79) by the presence of the factor $\cos\omega\tau$ that gives to $\rho(\tau)$ the form of a damped harmonic oscillation as shown in Figure 26. If we were to observe the generating process, some periodicity would be apparent.

(Place Figure 26 about here)

The corresponding normalized spectral density function can be obtained replacing $\cos \omega\tau$ by $\frac{1}{2}(e^{i\omega\tau}+e^{-i\omega\tau})$ and by replacing λ by $(\lambda-\omega)$ and $(\lambda+\omega)$ in the integrals of $f(\lambda)$. Then,

$$(2.83) \quad f(\lambda) = \frac{1}{2\Pi}\left|\frac{\alpha}{(\lambda-\omega)^2+\alpha^2} + \frac{\alpha}{(\lambda+\omega)^2+\alpha^2}\right| = \frac{\alpha}{\Pi}\left|\frac{\lambda^2+\alpha^2+\omega}{(\lambda^2-\omega^2-\alpha^2)^2+4\alpha^2\lambda^2}\right|$$

The representation of $f(\lambda)$ shown in Figure 27 presents peaks in the neighbourhood of the angular frequency ω. For $\omega=\frac{\Pi}{3}$, the fundamental seasonal frequency in the time domain would correspond to a period of 6 months; for $\omega=\frac{\Pi}{6}$, to a period of 12 months. Then, the corresponding model for the generating process would have an oscillatory seasonal component whose amplitude and phase are changing slowly compared with the fundamental seasonal period. The larger the β (that is, the more rapid is the rate of change) the more obscure is the seasonal component and therefore, the less sharp is the peak in the normalized spectral density function.

(Place Figure 27 about here)

REFERENCES

(1) Anderson, T.W. (1971) The Statistical Analysis of Time Series, Wiley, New York

(2) Beveridge, W.H. (1922) "Wheat Prices and Rainfall in Western Europe" J. Roy. Stat. Soc. 85.

(3) Box, G. and G. Jenkins (1970) Time Series Analysis Forecasting and Control, Holden Day, California

(4) Brillinger, D.R. (1975): Time Series: Data Analysis and Theory, Holt, Rinehart and Winston, New York.

(5) Buys-Ballot, G. (1847) "Les changements périodique de temperature" Kemink et Fils; Utrecht.

(6) Cramer, H. and M.R. Leadbetter (1967) Stationary and Related Stochastic Processes, Wiley, New York.

(7) Fuller, Wayne A. (1976): Introduction to Statistical Time Series, Wiley, New York.

(8) Granger, C.W. and M. Hatanaka, (1964) Spectral Analysis of Economic Time Series, Princeton University Press, New Jersey.

(9) Jenkins, G. and D. Watts (1969) Spectral Analysis and its Applications Holden Day, California

(10) Khintchine, A.Y. (1934) "Korrelations Theorie der Stationarem Stochastischen Prozesse" Math. Ann. 109, 604-615.

(11) Koopmans, L.H. (1974): The Spectral Analysis of Time Series, Academic Press, New York.

(12) Lagrange, J.L.(1972-78) Oevres, Vol. 6.

(13) Moore, Henry L. (1914): Economic Cycles: Their Law and Cause, Macmillan, New York.

(14) Nerlove, M., Grether, D.M. and Carvalho, J.L. (1979): Analysis of Economic Time Series, Academic Press, New York.

(15) Priestley, M.B. (1981): Spectral Analysis and Time Series, Academic Press, New York, Vol. I and II.

(16) Schuster, A. (1898) "On the Investigation of Hidden Periodicities with Application to a Supposed Twenty-six Day Period of Meteorological Phenomena", Terrestrial Magnetism, and Atmospheric Electricity (now Journal of Geographical Research) vol. 3.

(17) Slutsky, E.E. (1938) "Sur les fonctions aléatoires presque périodiques et sur la décomposition de fonctions aléatoires stationnaires en composantes" Actualités Scient. et Ind.738, Paris.

(18) Stokes, G.C. (1879) "Note on Searching for Periodicities" Proc. Roy. Soc. V. 29.

(19) Yaglom, A.M. (1962) Theory of Stationary Random Functions, Prentice Hall, New Jersey.

(20) Yule, V. (1921) "On the Time Correlation Problem with Special Reference to the Variate-Difference Correlation Method" J. Roy. Stat. Soc. Vol. 84, 497-526.

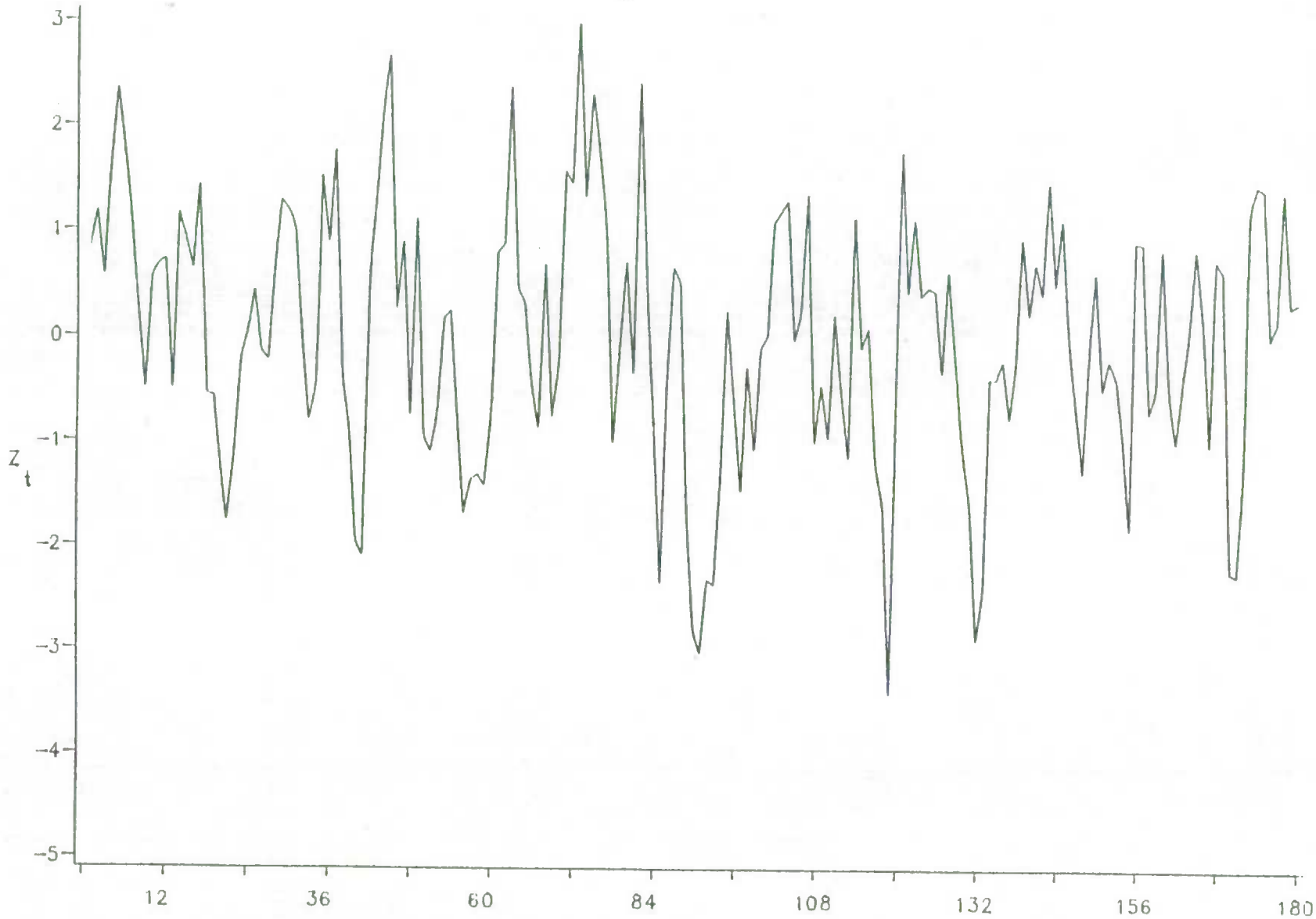


Figure 6 - A Realization of an AR(1) Process

$X_t = .60\ X_{t-1} + U_t$ and $\sigma_u^2 = 1.0$

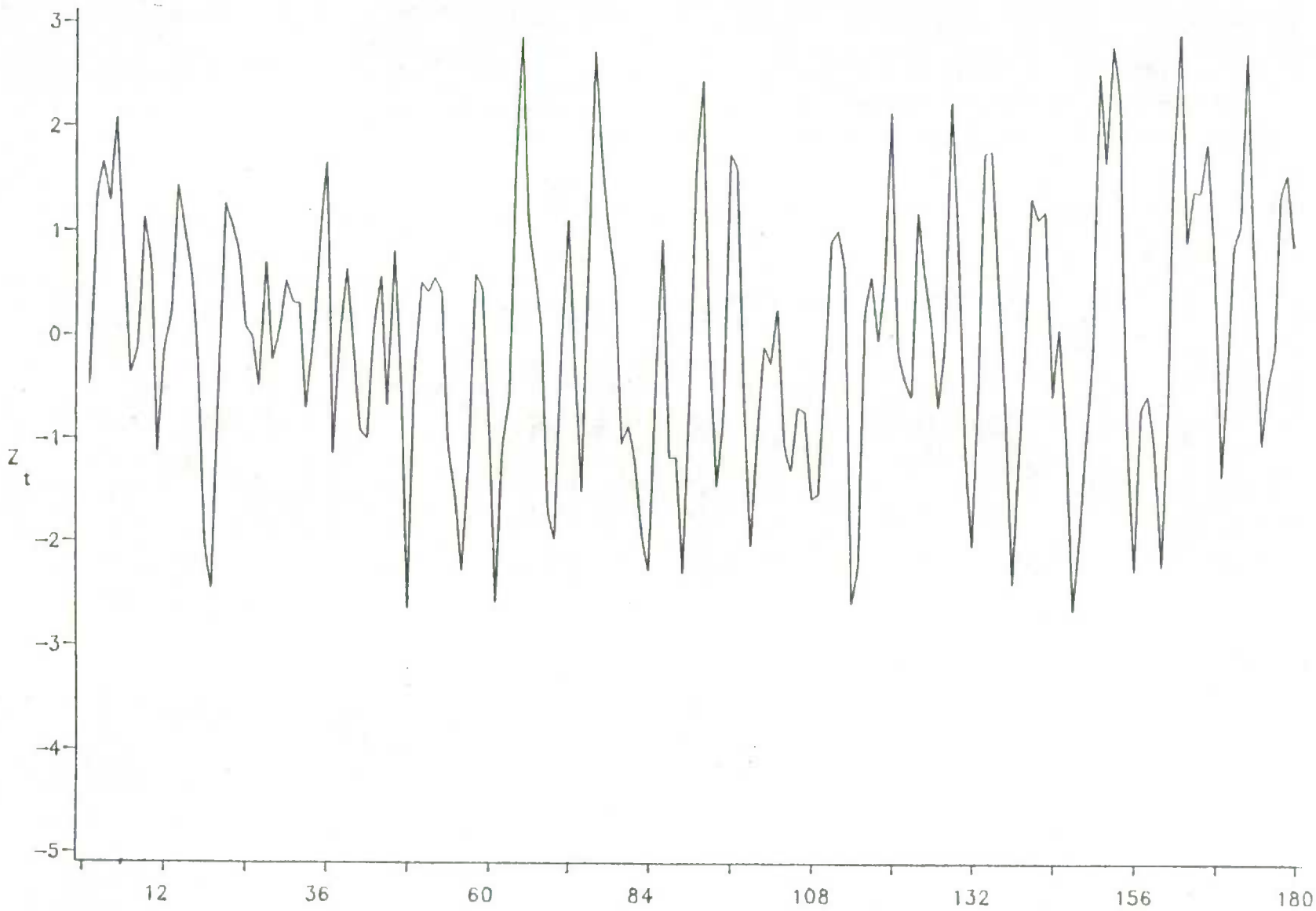


Figure 7 - A Realization of an AR(2) Process

$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t$ and $\sigma_u^2 = 1.0$

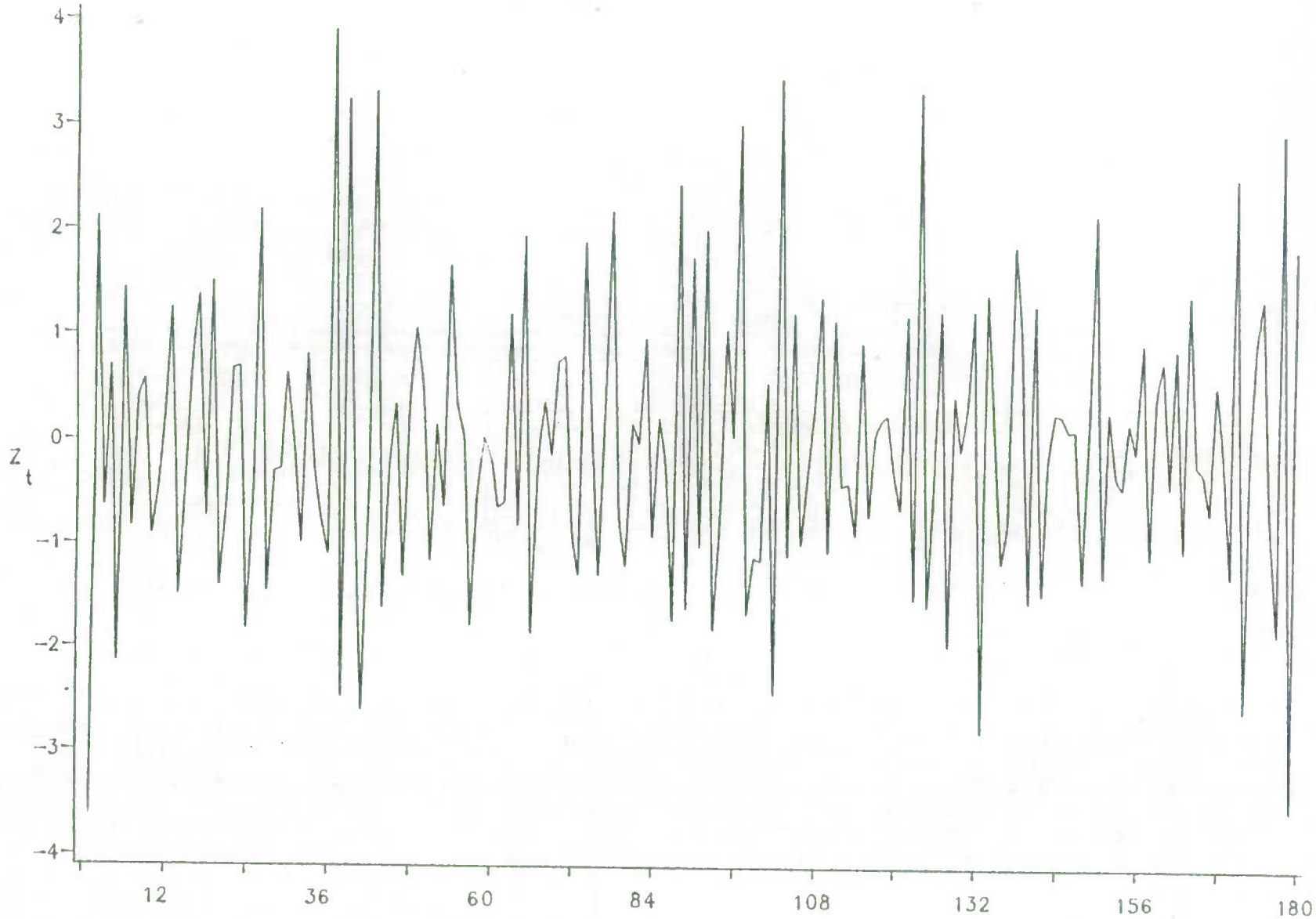


Figure 8 - A Realization of an MA(1) Process

$X_t = U_t - .80\ U_{t-1}$ and $\sigma_u^2 = 1.0$

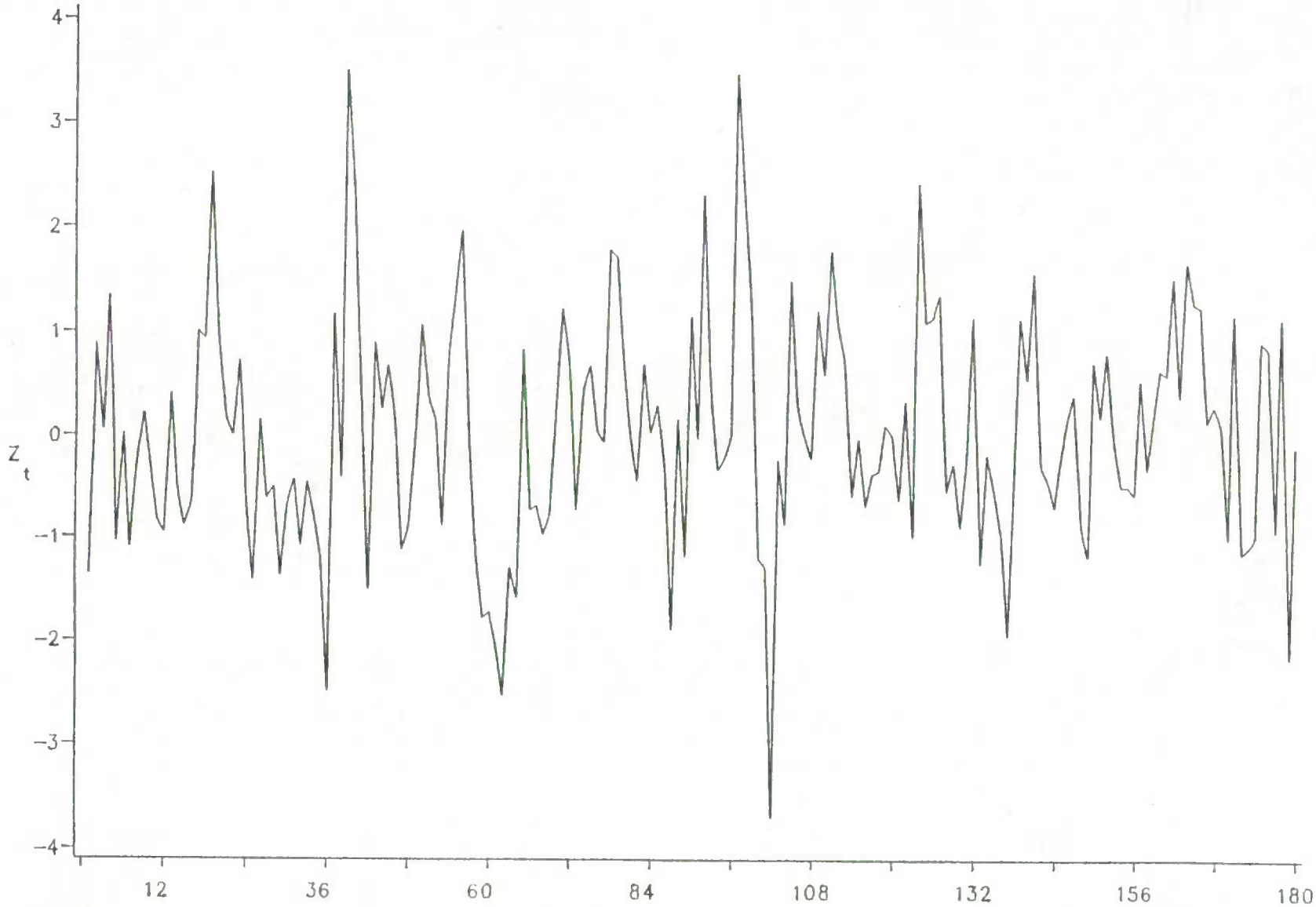


Figure 9 - A Realization of an MA(2) Process

$X_t = U_t + .40\ U_{t-1} + .40\ U_{t-2}$ and $\sigma_u^2 = 1.0$

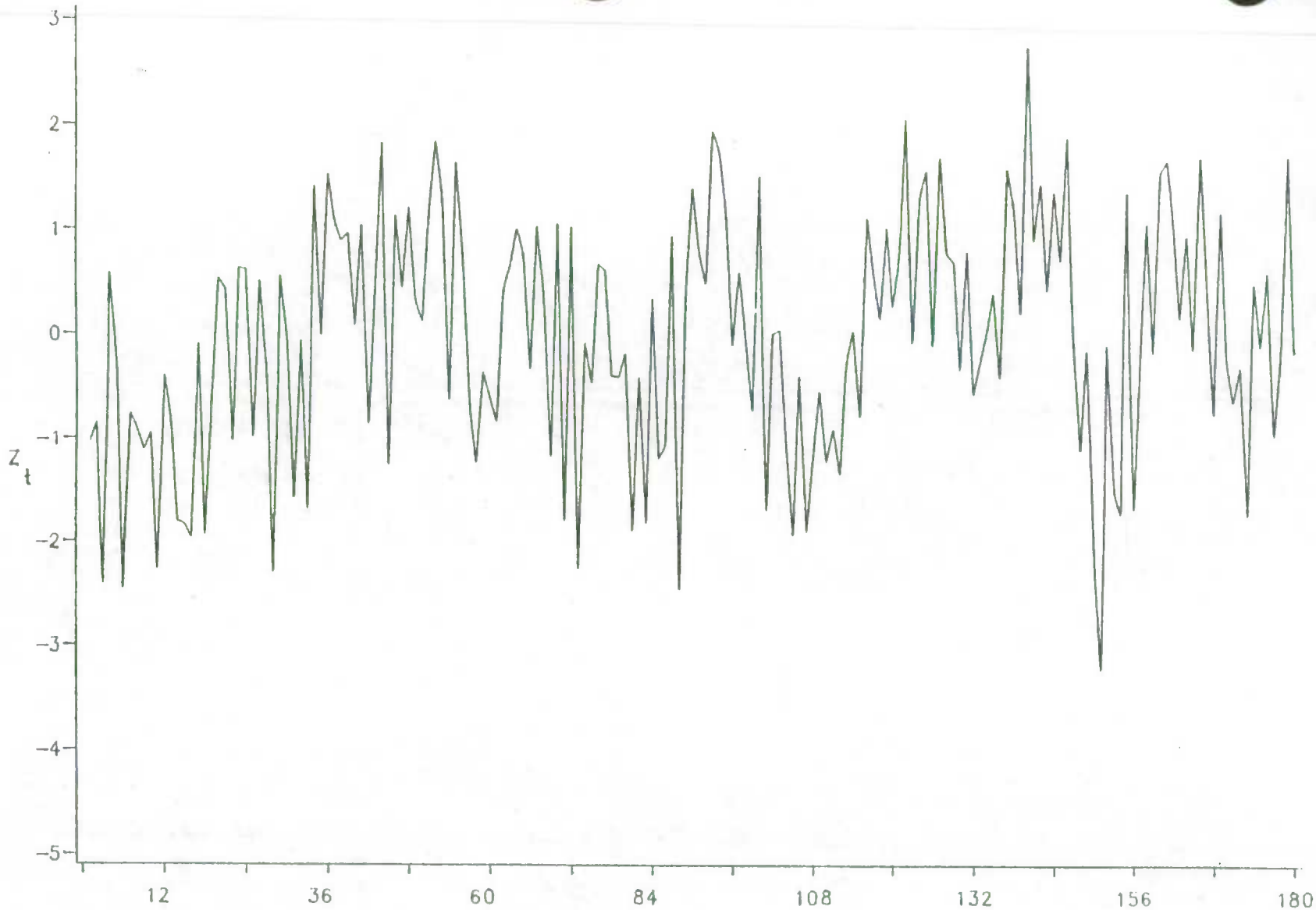


Figure 10 - A Realization of an ARMA(1,1) Process

$X_t = .60\ X_{t-1} + U_t - .80\ U_{t-1}$ and $\sigma_u^2 = 1.0$

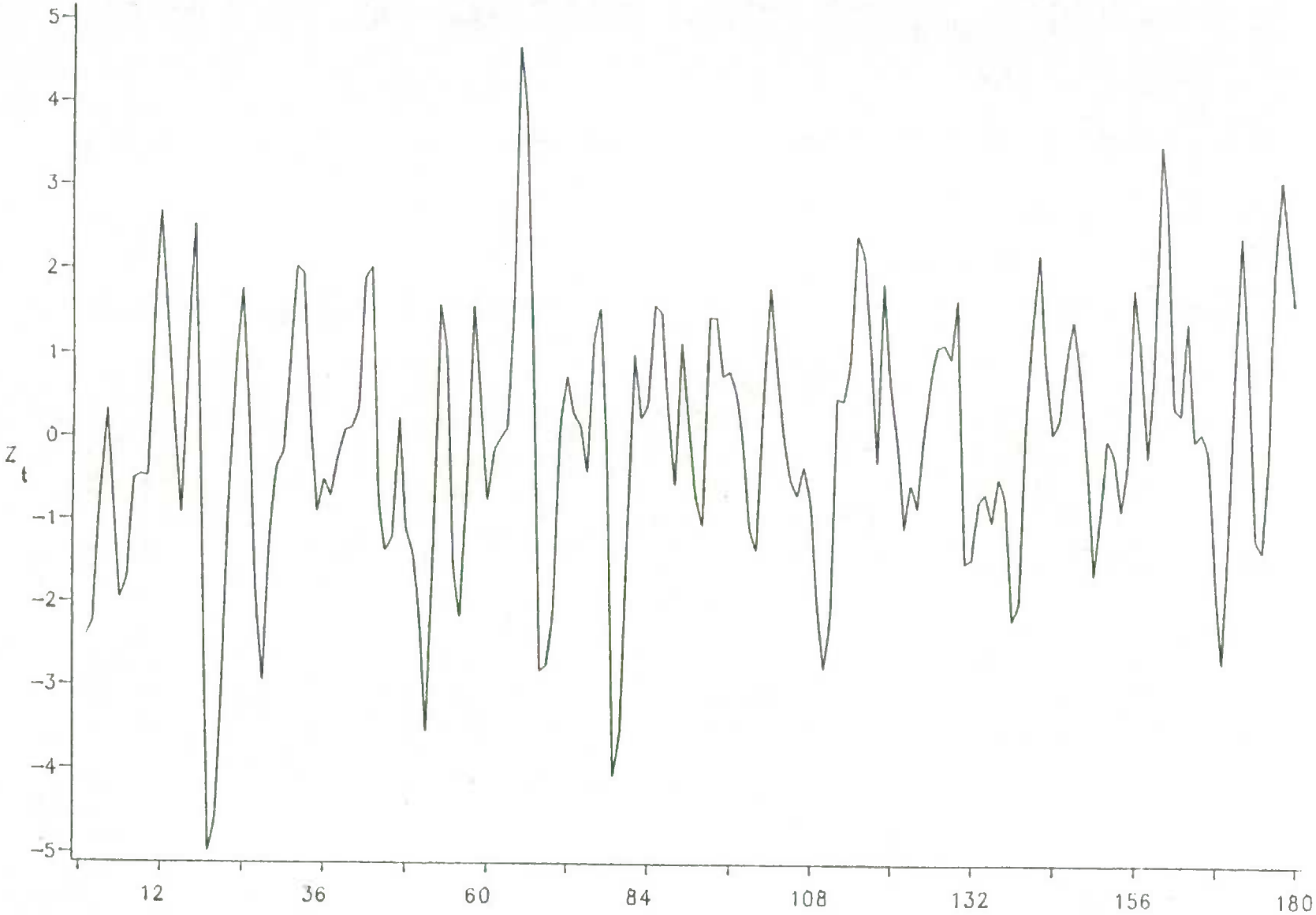


Figure 11 - A Realization of an ARMA(2,2) Process

$$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t + .40\ U_{t-1} + .40\ U_{t-2} \text{ and } \sigma_u^2 = 1.0$$

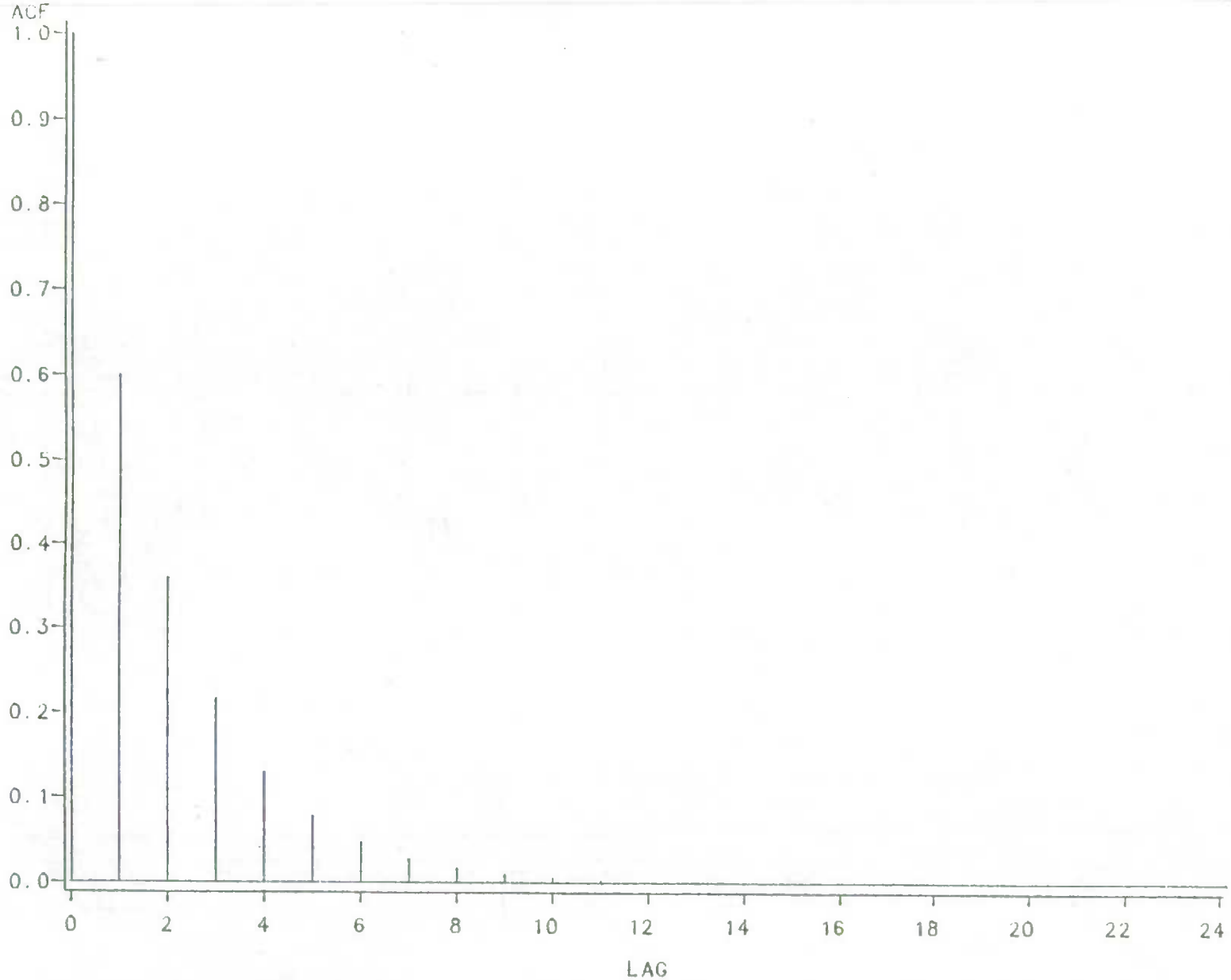


Figure 12 - Autocorrelation Function of AR(1) Process

$X_t = .60\ X_{t-1} + U_t$ and $\sigma_u^2 = 1.0$

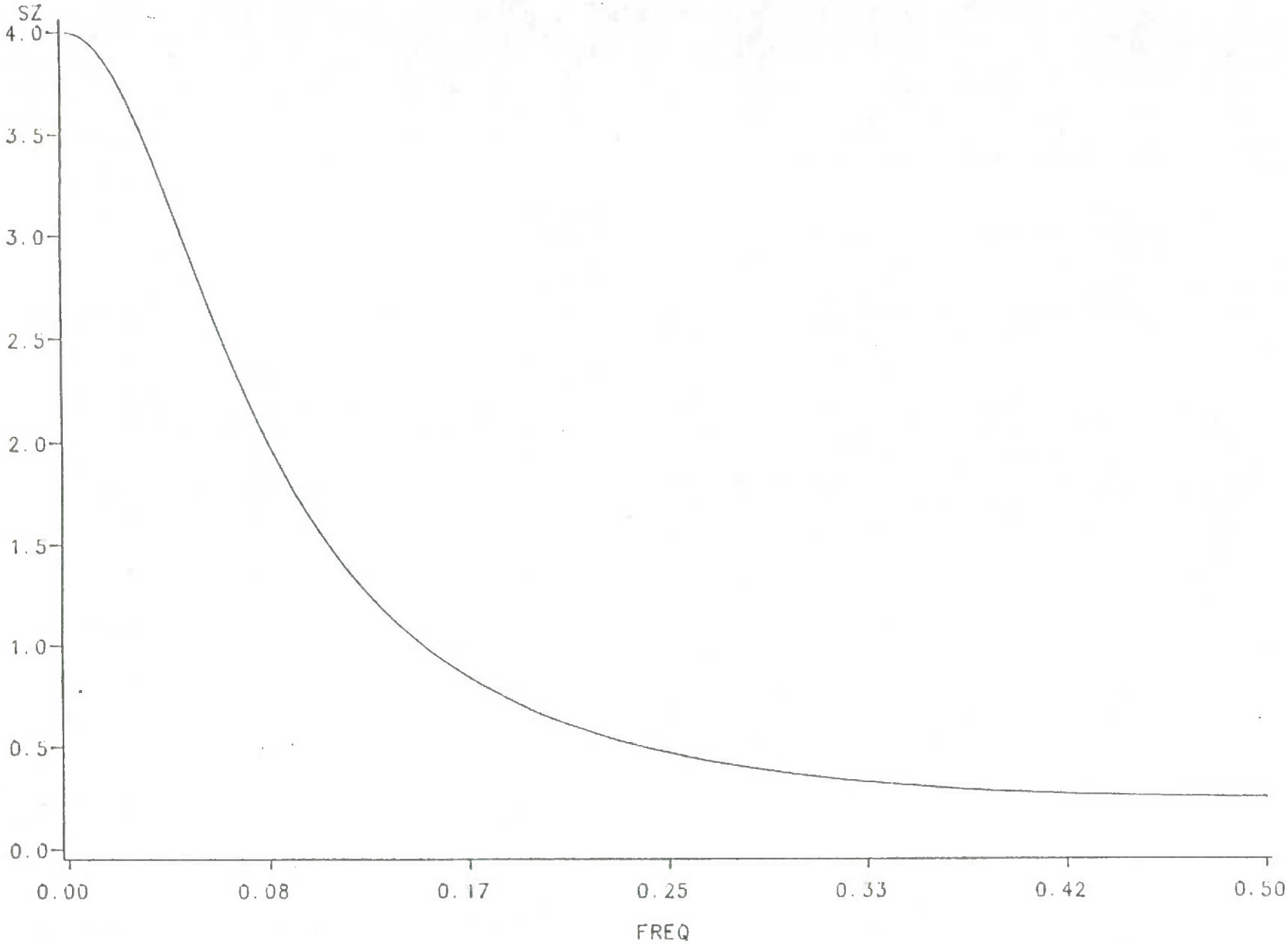


Figure 13 - Normalized Spectrum of an AR(1) Process

$X_t = .60\ X_{t-1} + U_t$ and $\sigma_u^2 = 1.0$

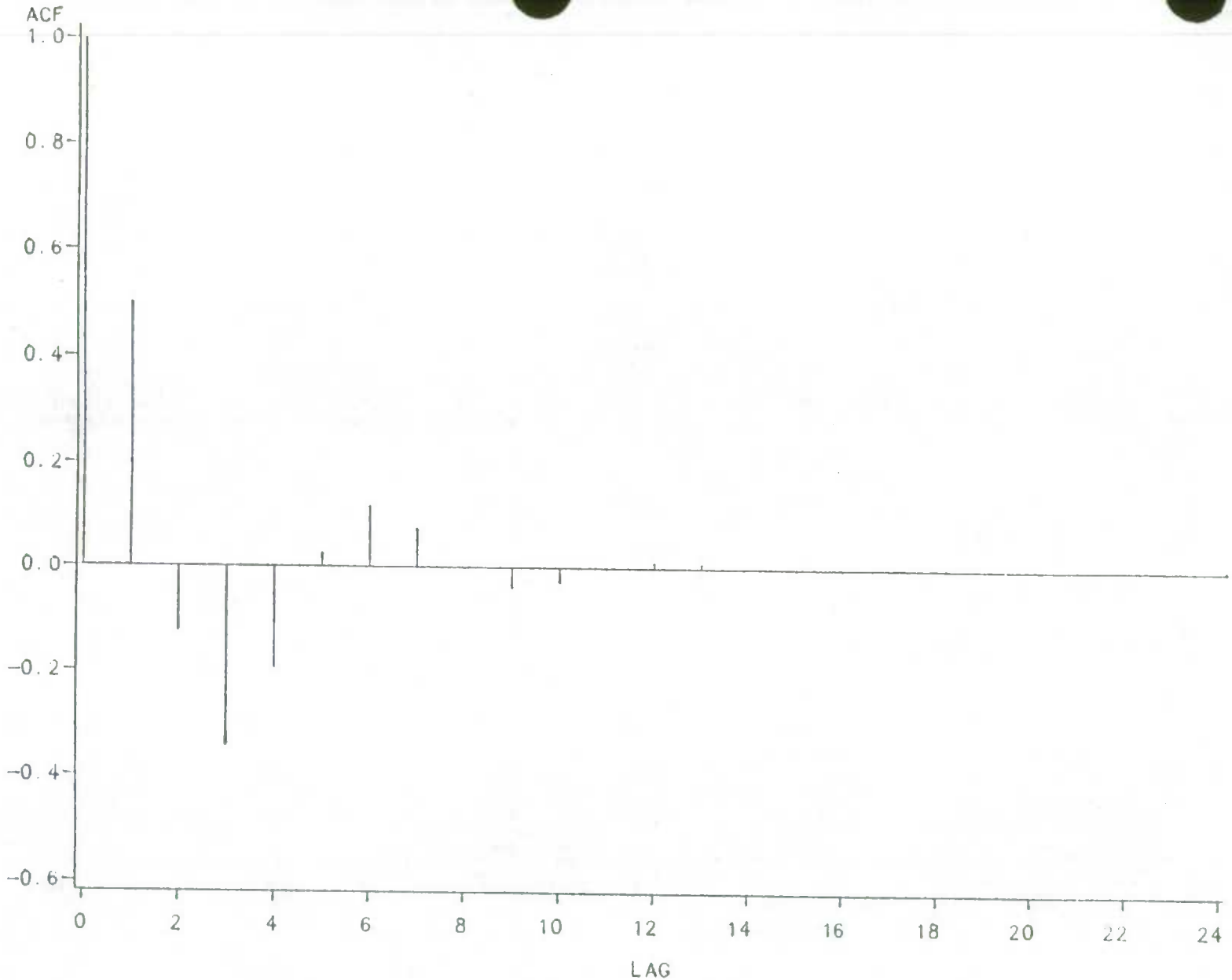


Figure 14 - Autocorrelcation Function of an AR(2) Process

$$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t \quad \text{and} \quad \sigma_u^2 = 1.0$$

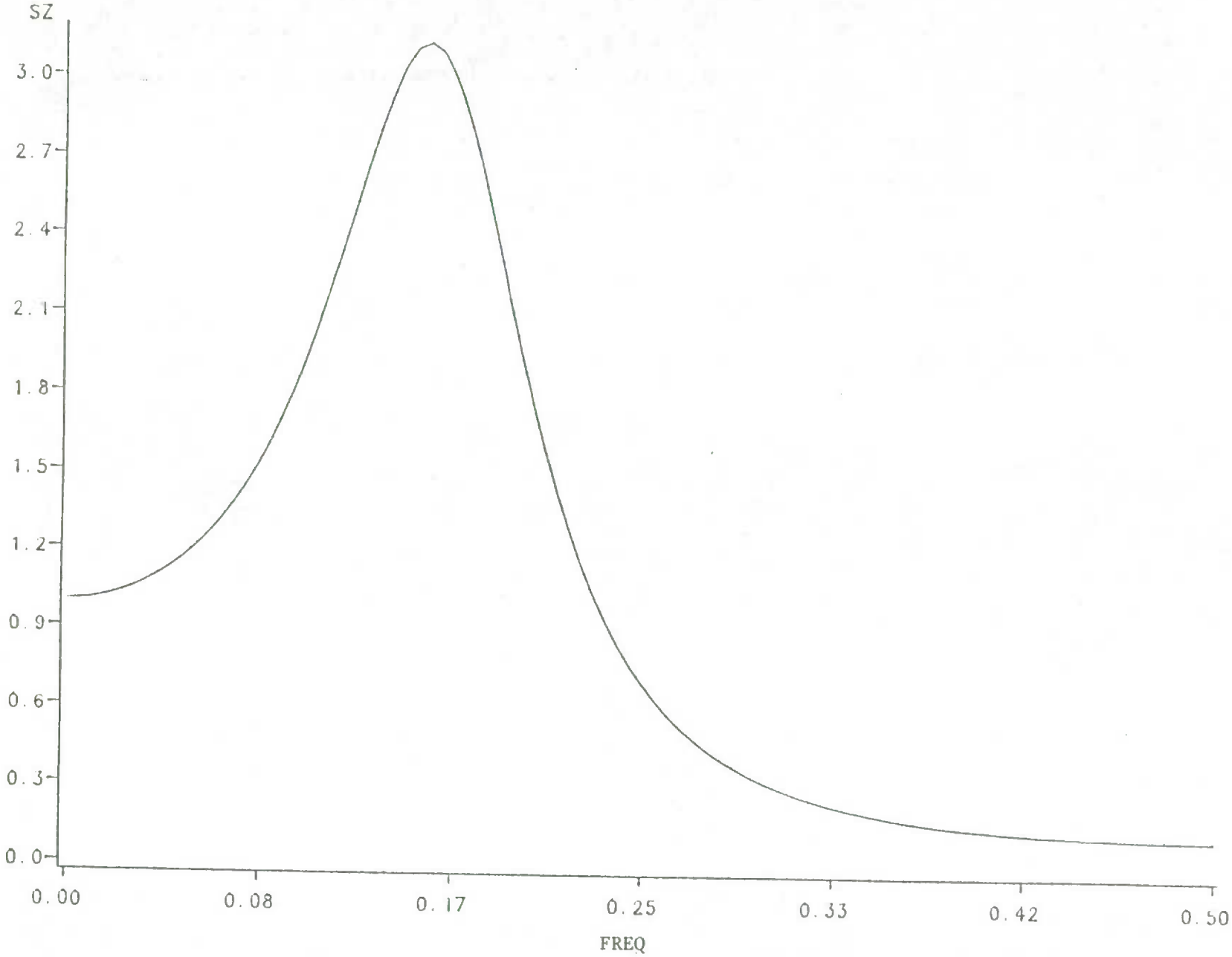


Figure 15 - Normalized Spectrum of an AR(2) Process

$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t$ and $\sigma_u^2 = 1.0$

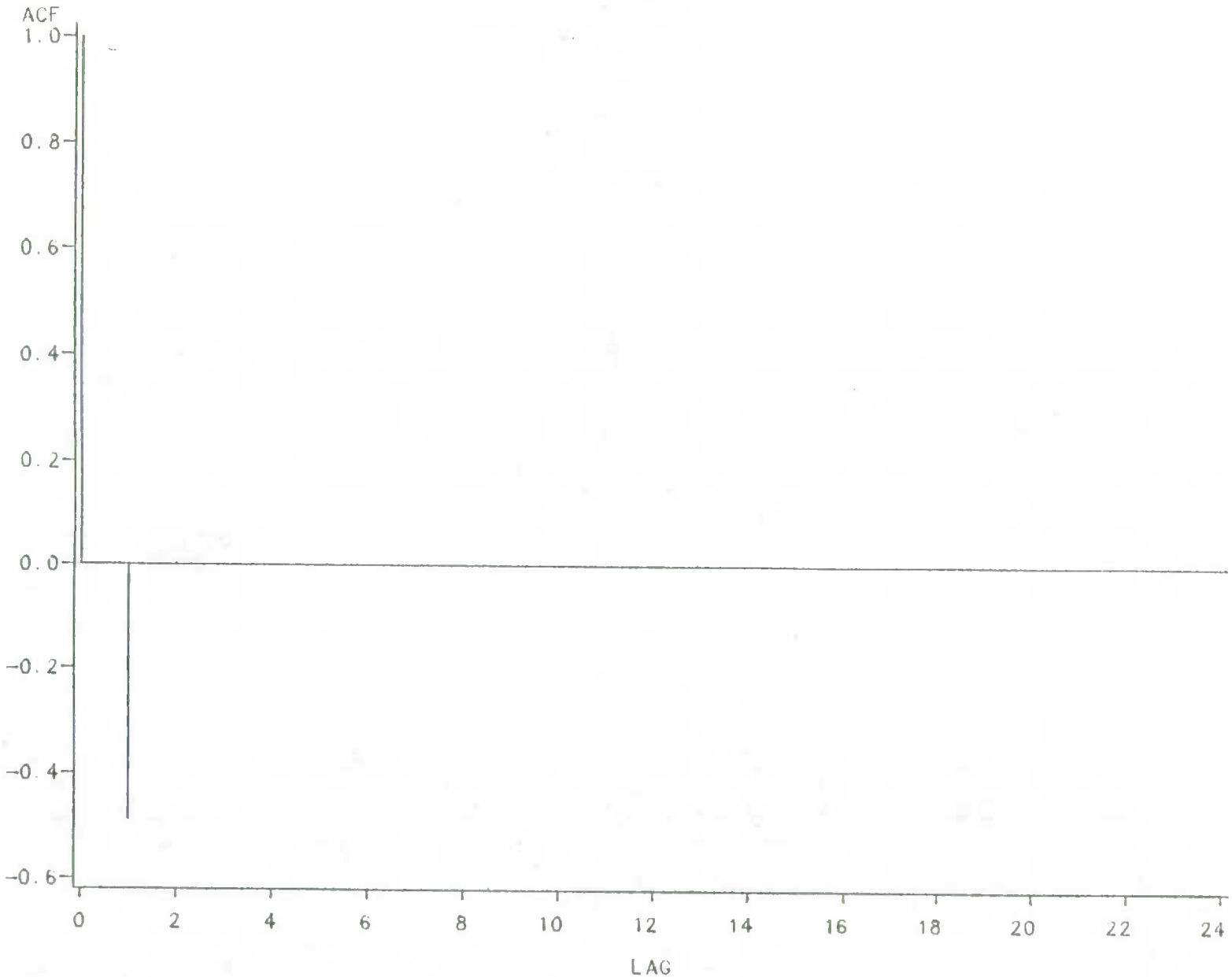


Figure 16 - Autocorrelation Function of an MA(1) Process

$X_t = U_t - .80\ U_{t-1}$ and $\sigma_u^2 = 1.0$

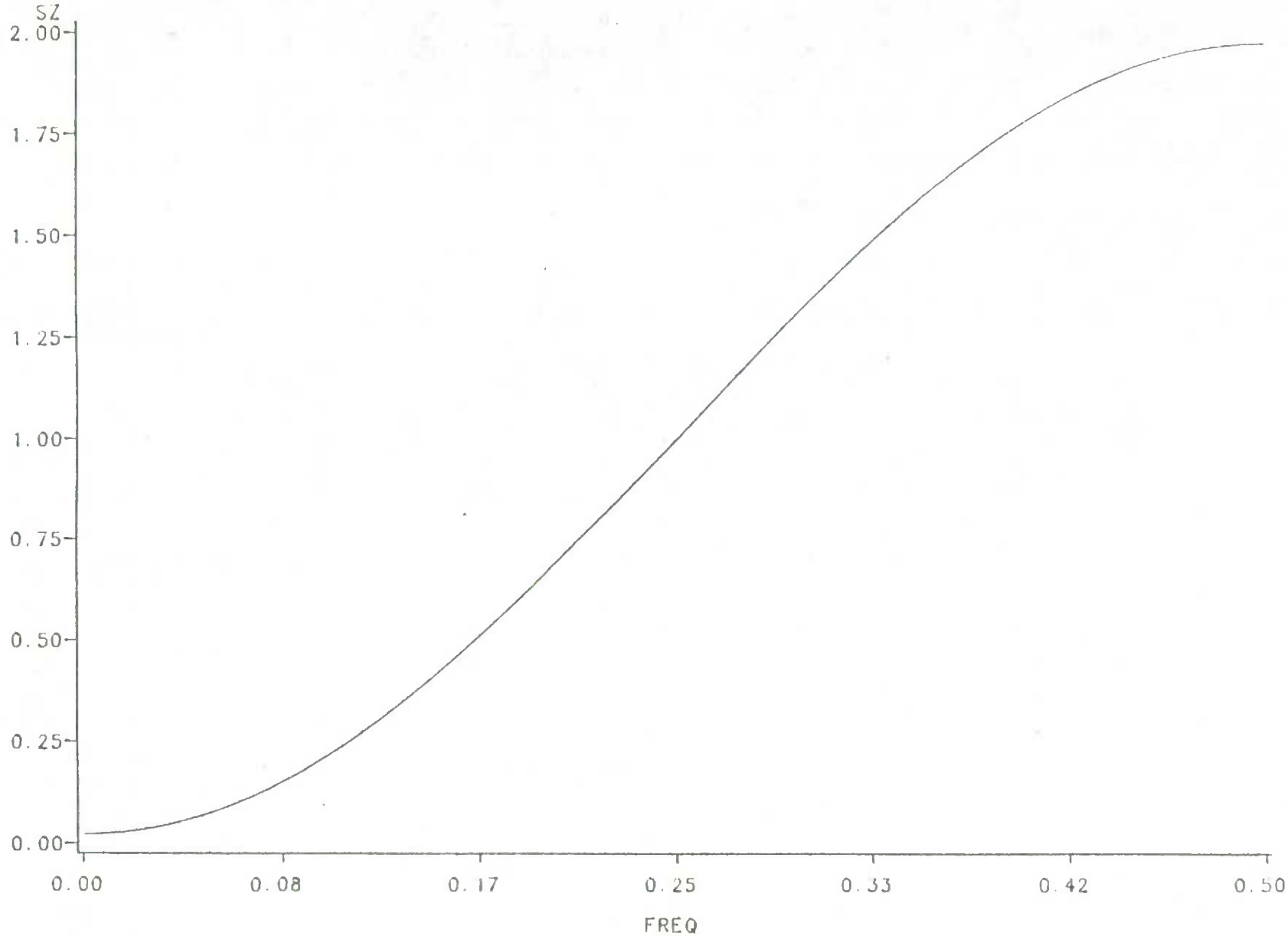


Figure 17 - Normalized Spectrum of an MA(1) Process

$X_t = U_t - .80\ U_{t-1}$ and $\sigma_u^2 = 1.0$

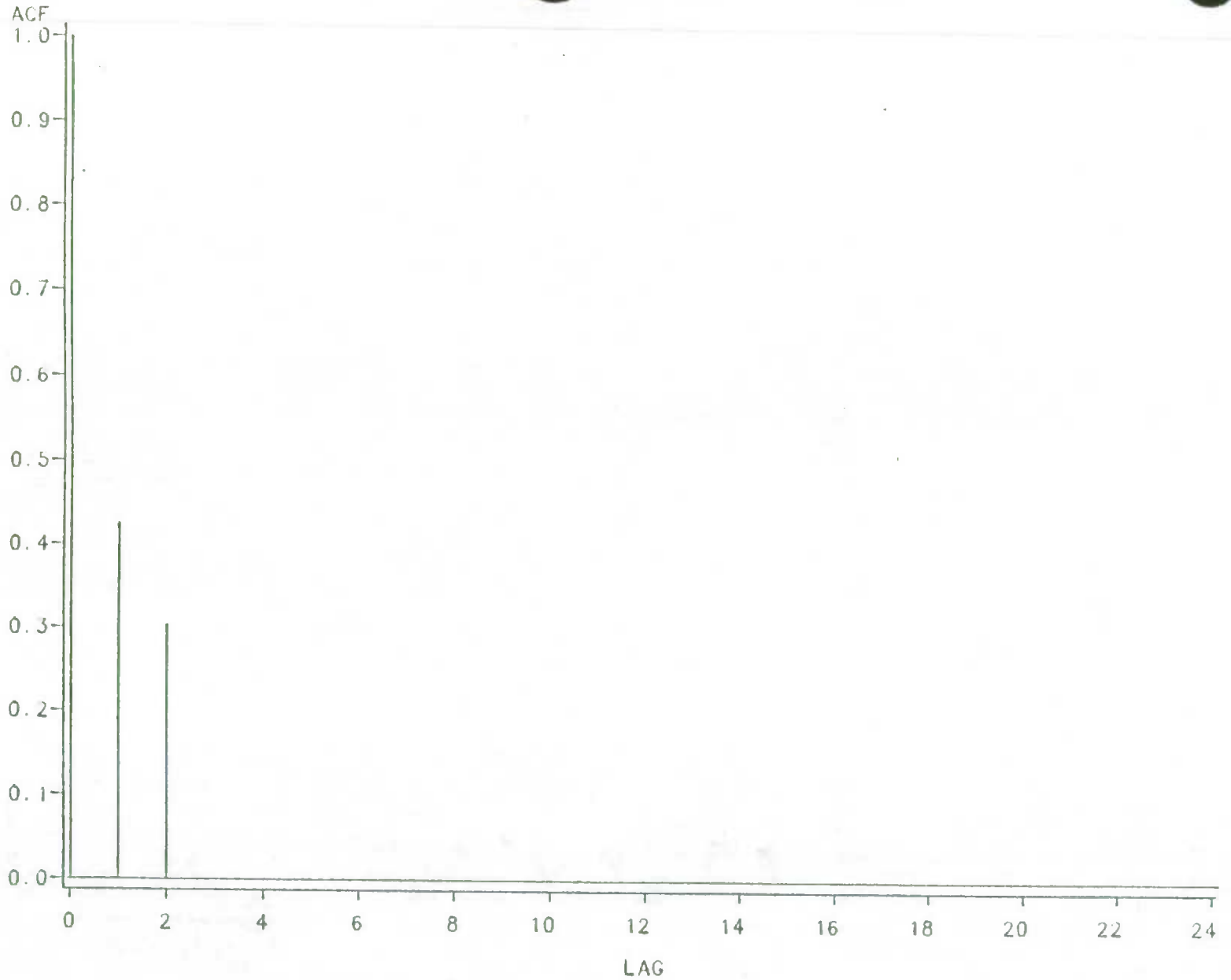


Figure 18 - Autocorrelation Function of an MA(2) Process

$$X_t = U_t + .40\ U_{t-1} + .40\ U_{t-2} \quad \text{and} \quad \sigma_u^2 = 1.0$$

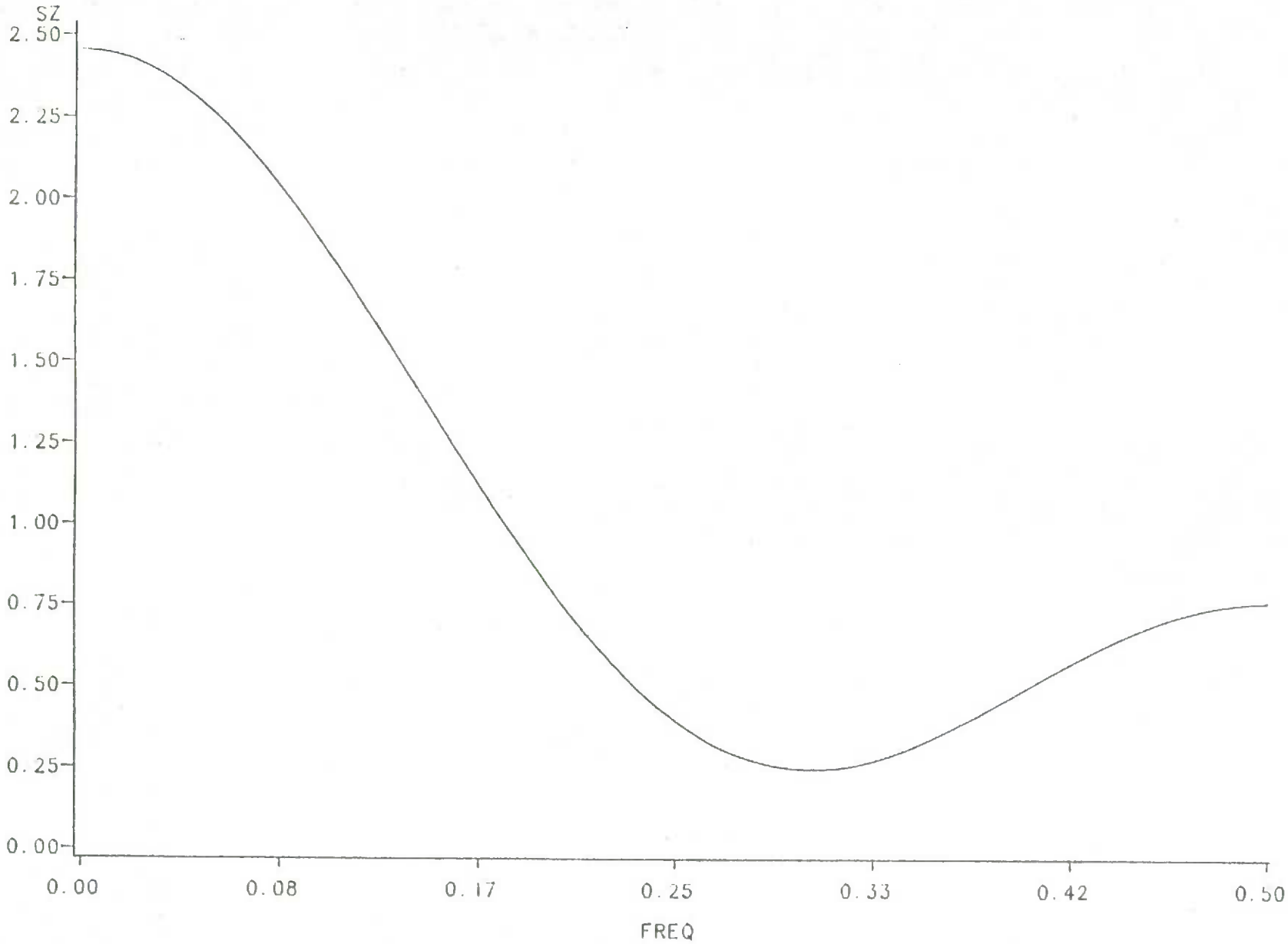


Figure 19 - Normalized Spectrum of an MA(2) Process

$X_t = U_t + .40\ U_{t-1} + .40\ U_{t-2}$ and $\sigma_u^2 = 1.0$

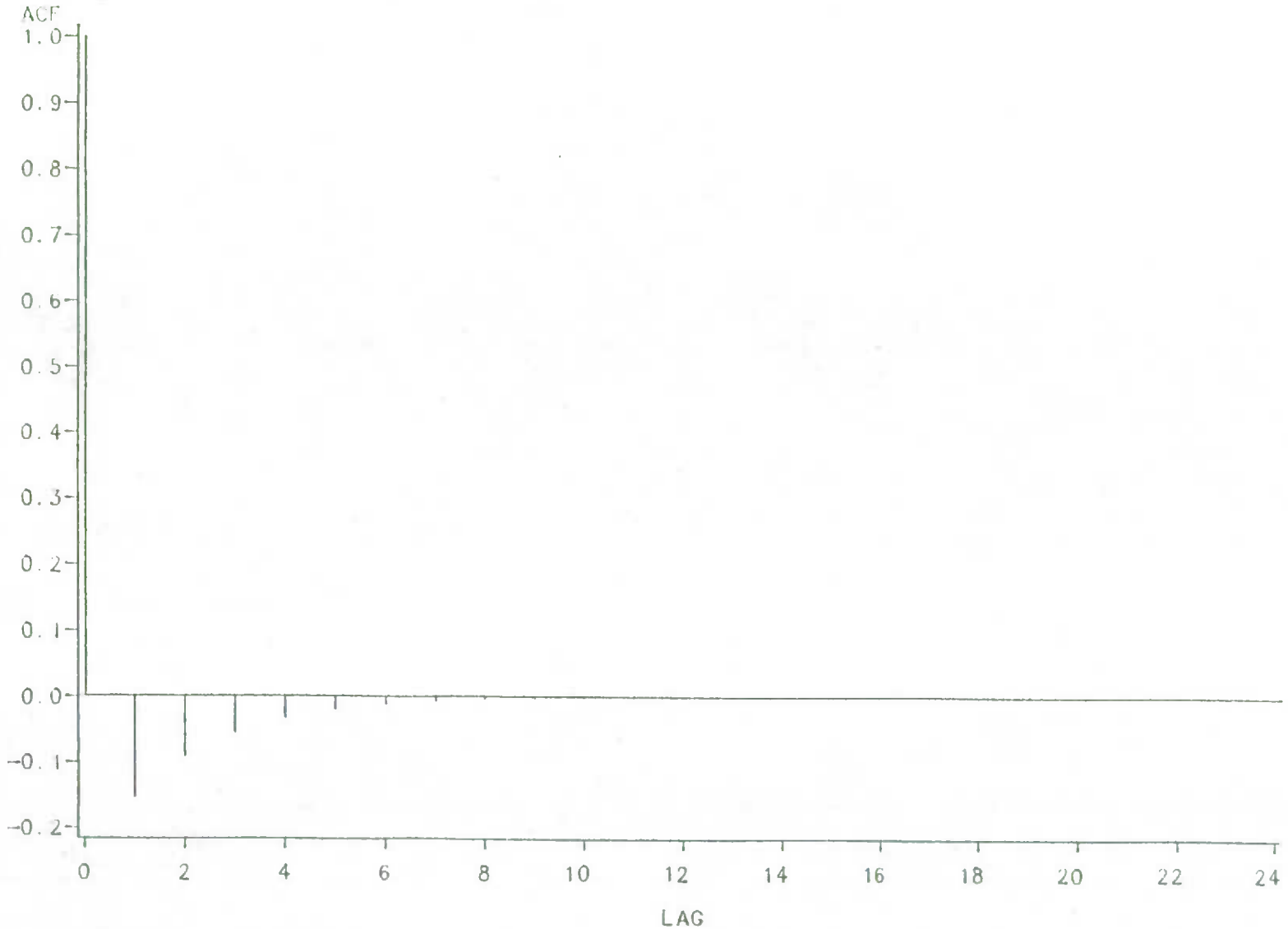


Figure 20 - Autocorrelation Function of an ARMA(1,1) Process

$$X_t = .60\ X_{t-1} + U_t - .80\ U_{t-1} \quad \text{and} \quad \sigma_u^2 = 1.0$$

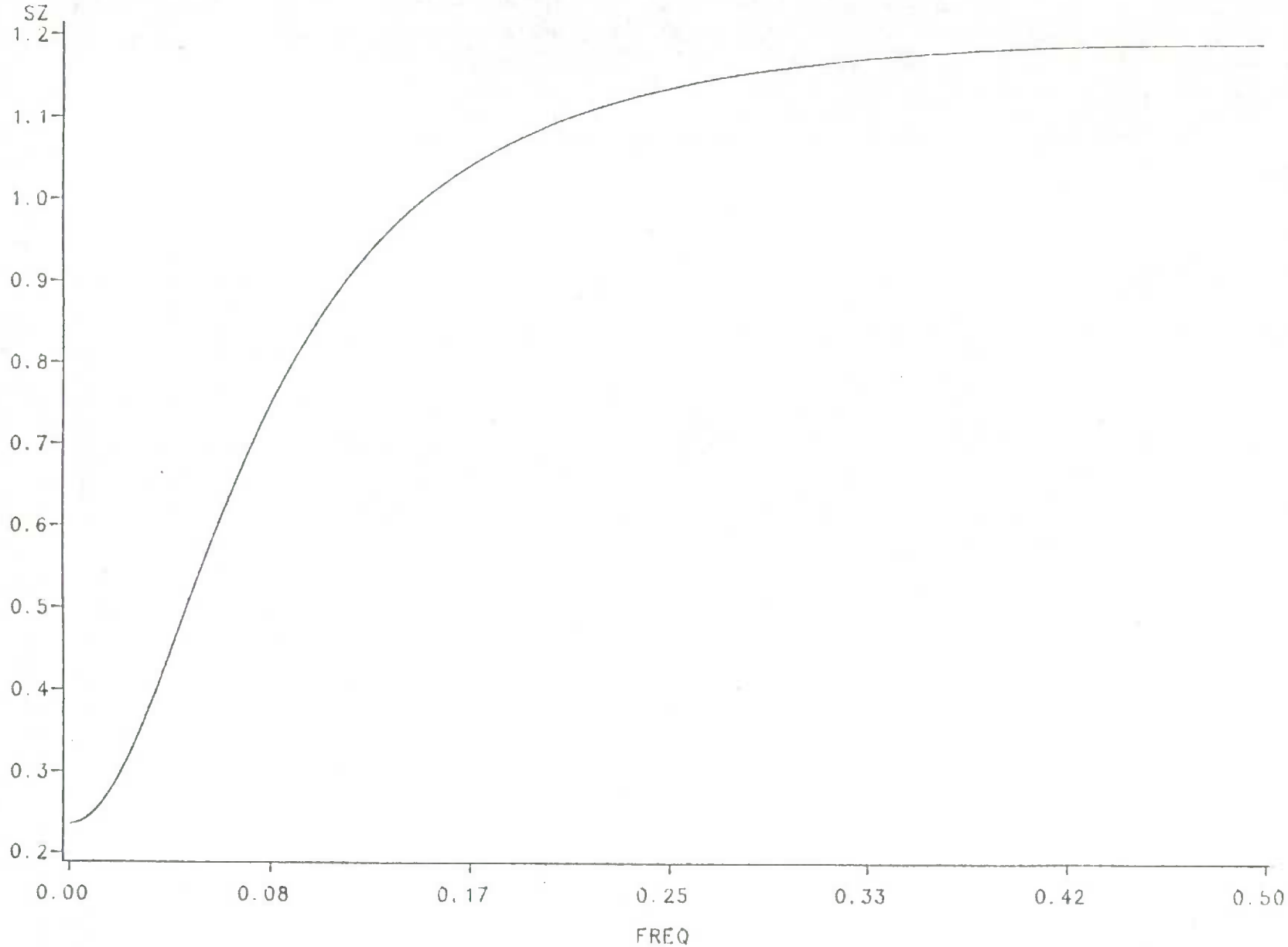


Figure 21 - Normalized Spectrum of an ARMA(1,1) Process

$$X_t = .60\ X_{t-1} + U_t - .80\ U_{t-1} \quad \text{and} \quad \sigma_u^2 = 1.0$$

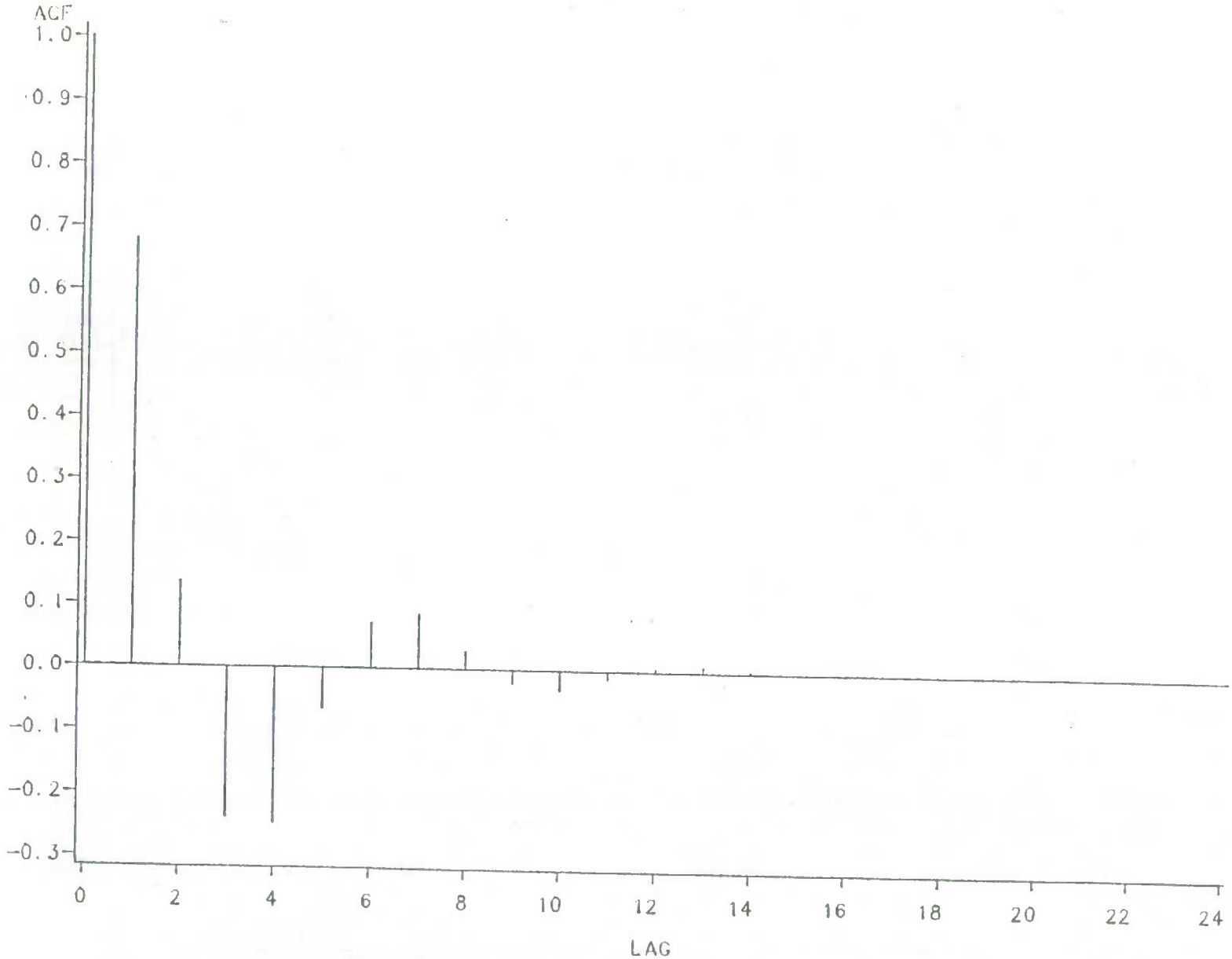


Figure 22 - Autocorrelation Function of an ARMA(2,2) Process

$$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t + .40\ U_{t-1} + .40\ U_{t-2} \quad \text{and} \quad \sigma_u^2 = 1.0$$

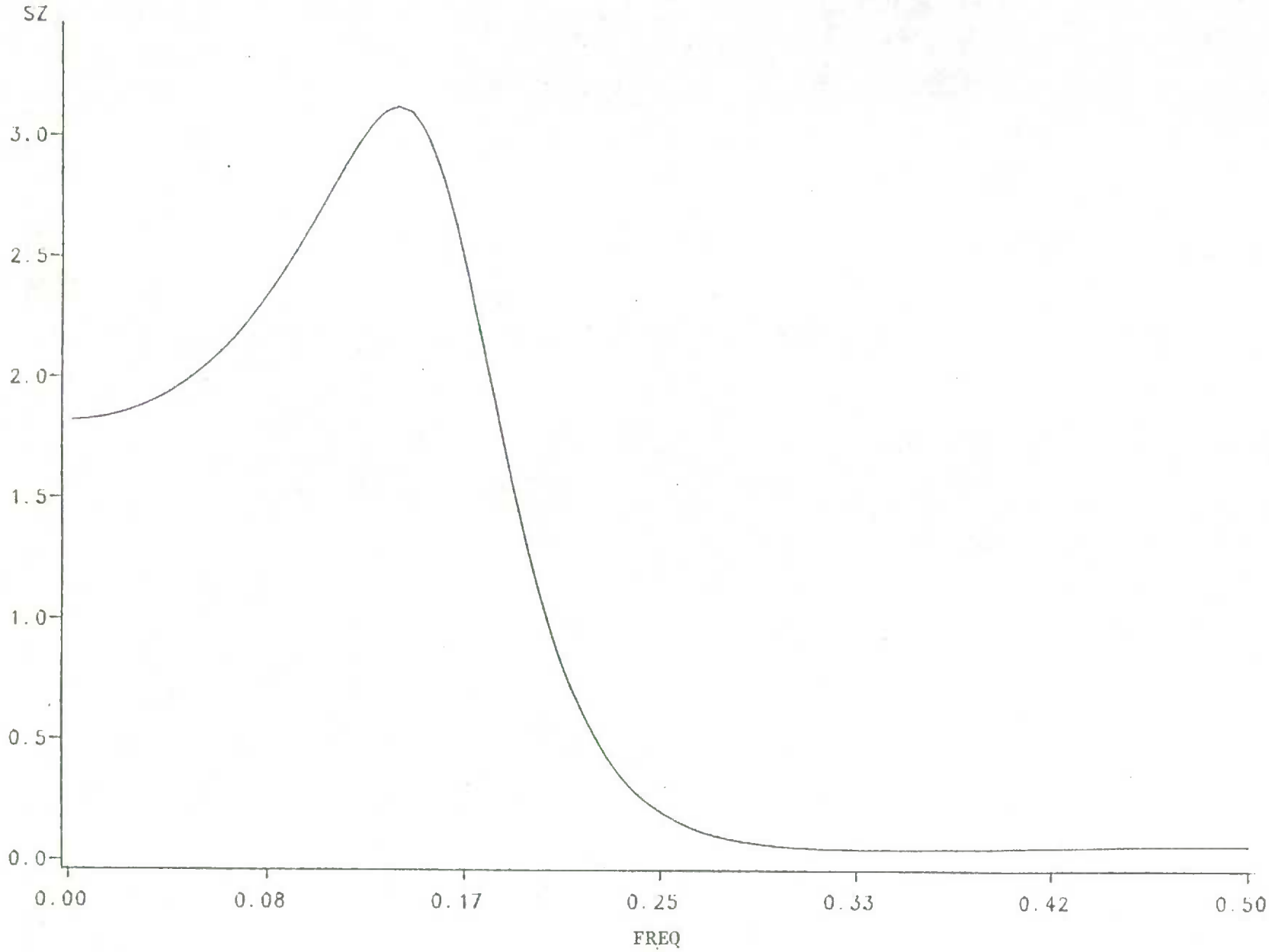


Figure 23 - Normalized Spectrum of ARMA(2,2) Process

$$X_t = .75\ X_{t-1} - .50\ X_{t-2} + U_t + .40\ U_{t-1} + .40\ U_{t-2} \quad \text{and} \quad \sigma_u^2 = 1.0$$

Figure 24

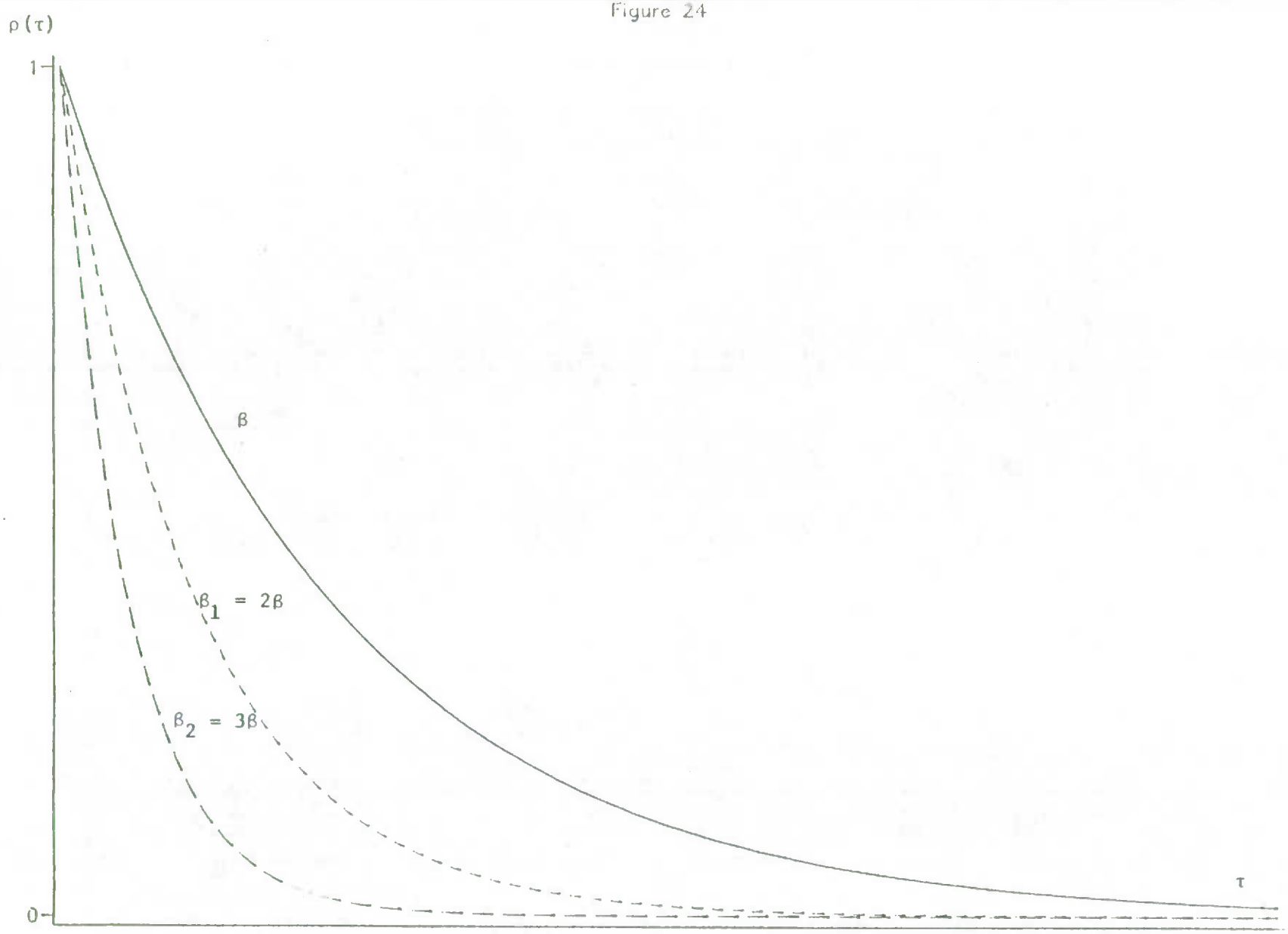


Figure 25

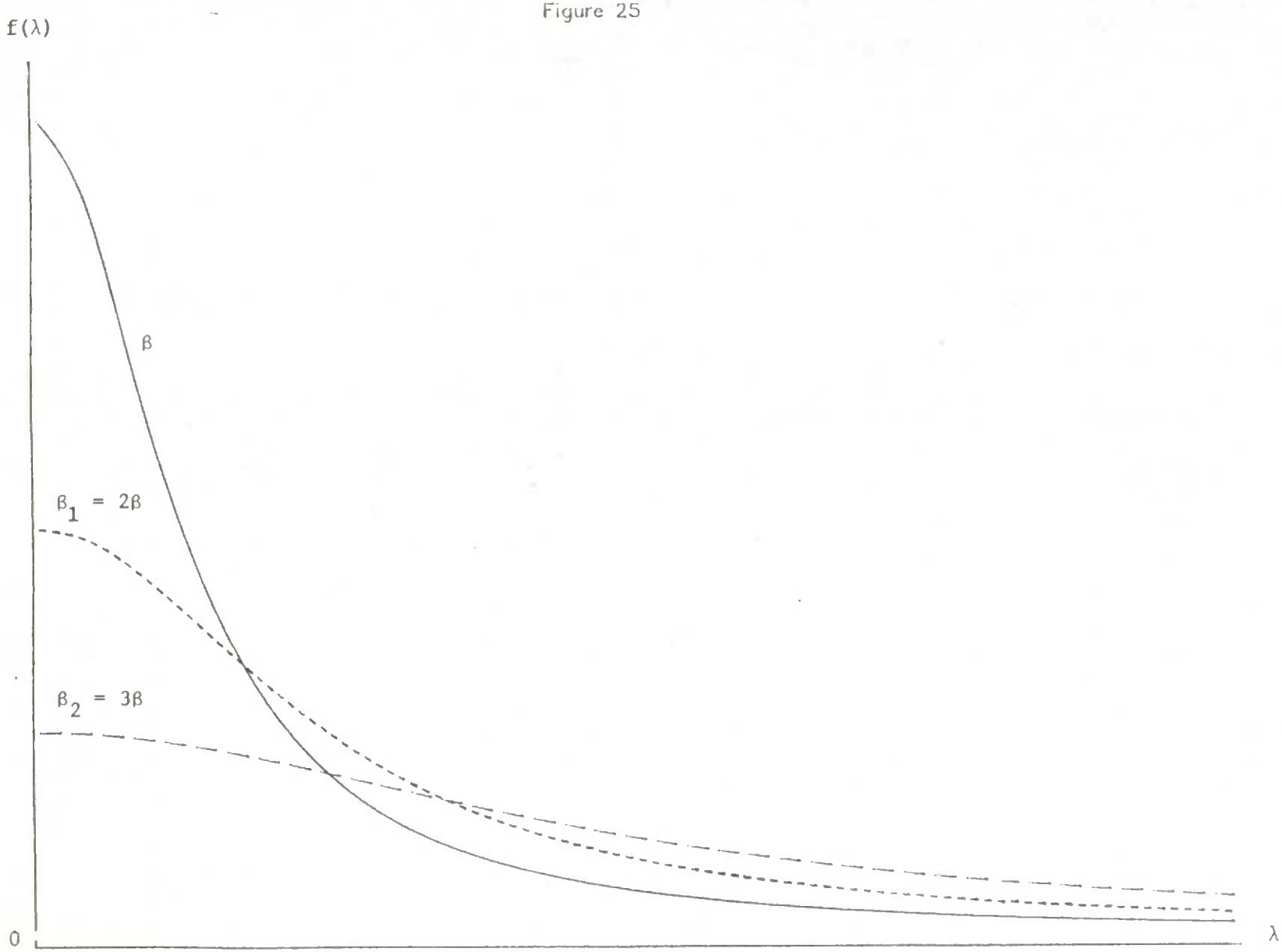


Figure 26

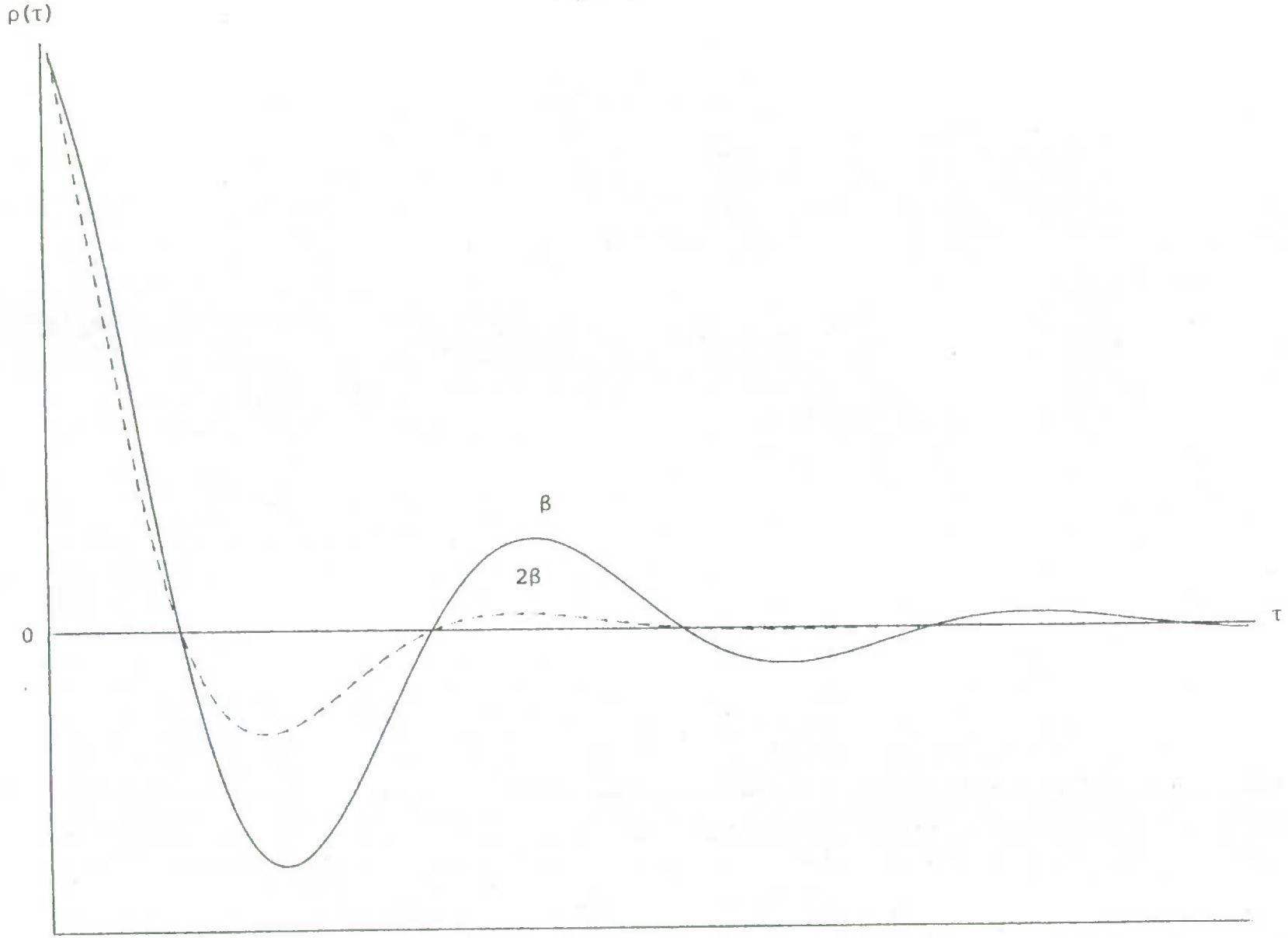
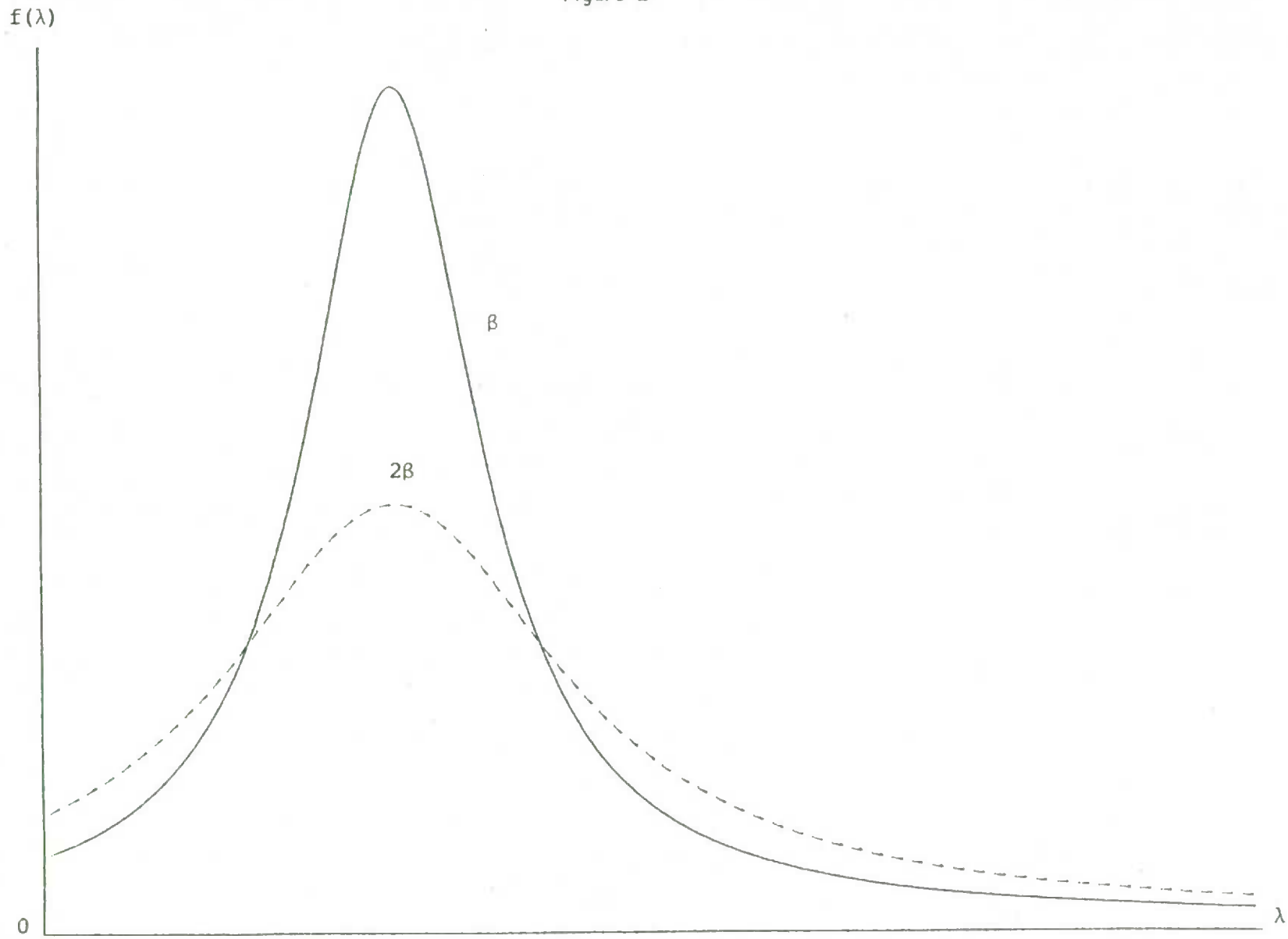


Figure 27



c.3