2-1-1966

Application of Signal Extraction Techniques in the Study of Economic Time Series

David M. Grether

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station

COWLES FOUNDATION DISCUSSION PAPER NO. 202

APPLICATION OF SIGNAL EXTRACTION TECHNIQUES IN THE STUDY OF ECONOMIC TIME SERIES

David Grether

February 25, 1966

Application of Signal Extraction Techniques in the Study of Economic Time Series*

by

David Grether

I. INTRODUCTION

A recent paper [1] showed how some of the results of P. Whittle [4] could be applied to the problem of adaptive forecasting. The purpose of this paper is to extend the results of [1]. In [1] attention was restricted to finding least squares predictors for a certain class of processes, assuming the entire past history of the process to be predicted was known. Here we consider predicting $\{y_t\}$ based upon observed values of a process $\{x_t\}$. We do not restrict ourselves to the case where $\{y_t\}$ is the future of the process $\{x_t\}$. In particular we consider the case where $\{y_t\}$ is an unobservable component of the process $\{x_t\}$. We also consider the case where only a finite number of values of $\{x_t\}$ have been observed. We conclude with an example using the unemployment series and a model introduced in [1].

In order to make the theory to be presented clear we illustrate its application by a series of simple examples. These examples are drawn from the

---

* The research underlying this paper was performed under National Science Foundation Grant NSF GS-142 to Stanford University and its continuation, NSF GS-818 to Yale University. The computations were carried out for the most part at the Stanford University Computation Center (supported by National Science Foundation Grant NSF GP-948).

exercises in [4]. The manner of exposition is first to present the theoretical result, then to give a simple illustration, and finally to show its application to the model presented in [1]. A complete development of the fundamental background for this paper is given in [1] and [4] especially Chapter 2. To prevent repetition the introductory discussion here is highly condensed.

Any stationary process $\{x_t\}$ can be represented as

(1.1) $$x_t = \xi_t + \eta_t$$

where $\{\xi_t\}$ and $\{\eta_t\}$ are mutually uncorrelated. $\{\xi_t\}$ is a so called linearly deterministic process; i.e. given its past, it can be predicted with zero mean square error. $\{\eta_t\}$ is a so called purely nondeterministic process and may be represented as a one-sided moving average.

(1.2) $$\eta_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j}$$

where

$$E(\epsilon_t \epsilon_s) = 0 , \qquad t \neq s$$
$$= \sigma^2 , \qquad t = s$$

Throughout this paper unless otherwise stated all processes are considered to be of the form (1.2).

To facilitate the exposition we introduce the backward shift operator U and the generating function or z-transform.

(1.3) $$U^j x_t = x_{t-j}$$

(1.4) $$B(z) = \sum_j b_j z^j$$

Using (1.3) and (1.4) the process $\{\eta_t\}$ given in (1.2) may be represented as:

(1.5) $$\eta_t = B(U)\epsilon_t$$

where

$$B(z) = \sum_{j=o}^{\infty} b_j z^j$$

We define the covariance generating function $g_{xx}(z)$ by:

(1.6) $$g_{xx}(z) = \sum_{j-\infty}^{\infty} \text{cov}(x_t, x_{t-j}) z^j = \sum_{-\infty}^{\infty} \Gamma_j^{xx} z^j$$

The spectral density function $f_x(\omega)$ is given by:

(1.7) $$f_x(\omega) = g_{xx}(e^{-i\omega}) = \sum_{-\infty}^{\infty} \Gamma_j^{xx} e^{-i\omega j} .$$

For the process (1.2) we have:

(1.8) $$g_{\eta\eta}(z) = \sigma^2 B(z) B(z^{-1})$$

where $B(z)$ is the cannonical factorization of $g_{\eta\eta}(z)$ (see [1] especially pages 4-7). If $B(z)$ has no zeroes on $|z| = 1$ then (1.2) also has an auto-regressive representation:

(1.9) $$A(U)\eta_t = \epsilon_t$$

where

$$A(z) = \sum_{j=o}^{\infty} a_j z^j = \frac{1}{B(z)}$$

In what follows all processes will be assumed to possess both moving average and auto-regressive representations. For the bivariate stationary process

$\{y_t, x_t\}$ we define the cross-covariance generating function $g_{yx}(z)$ by:

(1.10) $$g_{yx}(z) = \sum_{j=-\infty}^{\infty} \operatorname{cov}(y_t, x_{t-j}) z^j = \sum_{j=-\infty}^{\infty} \Gamma_j^{yx} z^j$$

In what follows it will often be useful to consider only some portion of a generating function $h(z)$, so we introduce the following conventions:

(1.11) $$[h(z)]_p^q = \sum_{j=p}^{q} h_j z^j$$
$$[h(z)]^q = \sum_{-\infty}^{q} h_j z^j$$
$$[h(z)]_p = \sum_{j=p}^{\infty} h_j z^j$$

We denote $[h(z)]_0 = [h(z)]_+$ and $[h(z)]^{-1} = [h(z)]_-$.

II. MINIMUM MEAN-SQUARE ERROR PREDICTORS

We now consider determining the minimum mean square error predictor of $\{y_t\}$ given the entire history of some process $\{x_t\}$. Let:

(2.1) $$\hat{y}_t = \sum_{j=0}^{\infty} \gamma_j x_{t-j}$$

taking

(2.2) $$x_t = B(U)\epsilon_t$$

where $\{\epsilon_t\}$ is a noise sequence with variance σ^2 we have

(2.3) $$\hat{y}_t = \gamma(U)\, B(U)\epsilon_t = \varphi(U)\epsilon_t = \sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j}$$

so we see that

(2.4) $$\varphi(z) = \gamma(z)\ B(z)\ .$$

The sequence $\{\varphi_j\}$ may be determined by the minimum mean square error criterion.

(2.5) $$E(\hat{y} - y)^2 = E(\sum_{j=o}^{\infty} \varphi_j\ \epsilon_{t-j} - y)^2$$

$$= \text{var}\ (y) + \text{var}\ (\Sigma\ \varphi_j\ \epsilon_{t-j}) - 2\ \text{cov}\ (y, \Sigma\ \varphi_j\ \epsilon_{t-j})$$

Since $\{\epsilon_t\}$ is an uncorrelated sequence we obtain:

(2.6) $$E(\hat{y} - y)^2 = \text{var}\ (y) + \sigma^2\ \Sigma\ \varphi_j^2 - 2\ \Sigma\ c_j\ \varphi_j$$

where $c_j = \text{cov}\ (y, \epsilon_{t-j})$. Minimize (2.6) by completing the square.

(2.7) $$E(\hat{y} - y)^2 = \text{var}\ (y) + \sigma^2\ [\Sigma\ \varphi_j^2 - \frac{2}{\sigma^2}\ \Sigma\ c_j\ \varphi_j + \frac{1}{\sigma^4}\ \Sigma\ c_j^2 - \frac{1}{\sigma^4}\ \Sigma\ c_j^2]$$

$$= \text{var}\ (y) + \sigma^2\ \Sigma\ (\varphi_j - \frac{c_j}{\sigma^2})^2 - \frac{1}{\sigma^2}\ \Sigma\ c_j^2$$

$$\geq \text{var}\ (y) - \frac{1}{\sigma^2}\ \Sigma\ c_j^2$$

with equality only for

(2.8) $$\varphi_j = \frac{c_j}{\sigma^2}$$

Note that the right hand side of (2.7) is nonnegative as:

(2.9) $$\text{var}\ (y - \frac{1}{\sigma^2}\ \Sigma\ c_j\ \epsilon_{t-j}) = \text{var}\ (y) + \frac{1}{\sigma^2}\ \Sigma\ c_j^2 - \frac{2}{\sigma^2}\ \Sigma\ c_j^2$$

$$= \text{var}\ (y) - \frac{1}{\sigma^2}\ \Sigma = E(y - \hat{y})^2$$

Equation (2.8) implies that

$$(2.10) \qquad \varphi(z) = \sum_{j=0}^{\infty} \varphi_j z^j = \frac{1}{\sigma^2} \sum_{j=0}^{\infty} c_j z^j = \frac{1}{\sigma^2} \left[g_{y\epsilon}(z) \right]_+$$

By the definition given in (1.6) we have:

$$(2.11) \qquad g_{yx}(z) = \sum_{-\infty}^{\infty} \operatorname{cov}(y, x_{t-j}) z^j = \sum_{j=-\infty}^{\infty} \operatorname{cov}(y, \sum_{k=0}^{\infty} b_j \epsilon_{t-j-k}) z^j$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} b_k c_{j+k} z^j = \sum_{j=-\infty}^{\infty} c_{j+k} z^{j+k} \sum_{k=0}^{\infty} b_k z^{-k}$$

$$= g_{y\epsilon}(z) B(z^{-1})$$

So we obtain

$$(2.12) \qquad \varphi(z) = \frac{1}{\sigma^2} \left[\frac{g_{yx}(z)}{B(z^{-1})} \right]_+$$

and from (2.4)

$$(2.13) \qquad \gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{yx}(z)}{B(z^{-1})} \right]_+$$

Consider the special case where $y_t = x_{t+\nu}$. Then for this case we have:

$$(2.14) \qquad g_{xy}(z) = \sum_{-\infty}^{\infty} \operatorname{cov}(y, x_{t-j}) z^j = \sum_{-\infty}^{\infty} \operatorname{cov}(x_{t+\nu}, x_{t-j}) z^j = z^{-\nu} \sum_{-\infty}^{\infty} \Gamma_{j+\nu}^{xx} z^{\nu+j}$$

$$= z^{-\nu} g_{xx}(z) = z^{-\nu} \sigma^2 B(z) B(z^{-1})$$

Thus,

$$(2.15) \qquad \gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{\sigma^2 B(z) B(z^{-1}) z^{-\nu}}{B(z^{-1})} \right]_+ = \frac{1}{B(z)} \left[\frac{B(z)}{z^{\nu}} \right]_+$$

[4] Equation 3.2.7

III. ILLUSTRATION FOR A SIMPLE PROCESS WITH RATIONAL SPECTRAL DENSITY

To illustrate the use of the result shown in (2.15) we now turn to an example ([4] Exercise 3.3.6). We consider the following process:

(3.1) $$x_t = \alpha x_{t-1} + \epsilon_t - \beta \epsilon_{t-1}$$

Which may be written using the shift operator U as:

(3.2) $$(1 - \alpha U) x_t = (1 - \beta U)\epsilon_t$$

(3.3) $$x_t = \frac{1 - \beta U}{1 - \alpha U} \epsilon_t$$

so

$$\beta(z) = \frac{1 - \beta z}{1 - \alpha z}$$

The generating function for the predictor of $x_{t+\nu}$ is given by

(3.4) $$\gamma(z) = \frac{1 - \alpha z}{1 - \beta z} \left[\frac{1 - \beta z}{z^{\nu}(1-\alpha z)} \right]_+$$

To evaluate (3.4) we expand the function inside the $[\cdot]_+$ operator.

(3.5) $$\left[\frac{1 - \beta z}{z^{\nu}(1-\alpha z)} \right]_+ = \left[z^{-\nu} \sum_{j=0}^{\infty} \alpha^j z^j - \beta z^{-\nu+1} \sum_{j=0}^{\infty} \alpha^j z^j \right]_+$$

$$= \left[z^{-\nu} \sum_{j=0}^{\infty} \alpha^j z^j \right]_+ - \beta \left[z^{-\nu+1} \sum_{j=0}^{\infty} \alpha^j z^j \right]_+$$

$$= \alpha^{\nu} \sum_{j=\nu}^{\infty} \alpha^{j-\nu} z^{j-\nu} - \beta \alpha^{\nu-1} \sum_{j=\nu-1}^{\infty} \alpha^{j-\nu+1} z^{j-\nu+1}$$

$$= \frac{\alpha^{\nu} - \beta \alpha^{\nu-1}}{1 - \alpha z}$$

So combining terms (3.4) becomes

$$(3.6) \qquad \gamma(z) = \frac{1 - \alpha z}{1 - \beta z} \cdot \frac{\alpha^{\nu} - \beta \alpha^{\nu-1}}{1 - \alpha z} = \frac{\alpha^{\nu} - \beta \alpha^{\nu-1}}{1 - \beta z}$$

$$(3.7) \qquad \hat{x}_{t+\nu} = (\alpha - \beta)\alpha^{\nu-1} \sum_{j=0}^{\infty} \beta^j x_{t-j}$$

IV. A SIMPLIFIED DERIVATION OF THE OPTIMAL FORECASTS FOR A MORE COMPLICATED MODEL

In (1) the problem of determining $\hat{x}_{t+\nu}$ for the following model was considered.

$$(4.1) \qquad x_t = T_t + S_t + I_t$$

where

$$T_t = \frac{u_t}{(1-\rho U)^2} , \quad S_t = \frac{v_t}{(1-U^L)^2} , \quad I_t = w_t$$

i.e. $T_t - 2\rho\, T_{t-1} + \rho^2 T_{t-2} = u_t$, $S_t - 2S_{t-L} + S_{t-2L} = v_t$, and $\{u_t\}$, $\{v_t\}$, $\{w_t\}$ are mutually uncorrelated noise sequences.

The process $\{x_t\}$ is thus the sum of three independent processes each of the form (1.2). In order for (2.15) to be applied $\{v_t\}$ must be put in the form

$$(4.2) \qquad x_t = B(U)\epsilon_t$$

So the first step in the solution of the prediction problem for this model is the determination of $B(z)$. The procedure is to find $g_{xx}(z)$ and to factor it as described in [1] pages 4-7. To illustrate what is involved we consider another simple example ([4] Ex. No. 3.3.7).

Let

(4.3) $$x_t = u_t + \epsilon_t' \qquad \text{where} \quad u_t = \frac{\epsilon_t}{1 - \alpha U}$$

$\{\epsilon_t\}$, $\{\epsilon_t'\}$ mutually uncorrelated sequences with $E(\epsilon_t^2) = \sigma_1^2$, $E(\epsilon_t'^2) = \lambda\sigma_1^2$, $\lambda > 0$,

(4.4) $$x_t = \frac{\epsilon_t}{1 - \alpha U} + \epsilon_t' = \frac{\epsilon_t + (1 - \alpha U)\epsilon_t'}{1 - \alpha U}$$

(4.5) $$g_{xx}(z) = \frac{\sigma_1^2 + \lambda\sigma_1^2(1 + \alpha^2) - \lambda\,\alpha\,\sigma_1^2 z - \lambda\,\alpha\,\sigma_1^2 z^{-1}}{(1 - \alpha z)(1 - \alpha z^{-1})}$$

Note that the numerator of (4.5) is found by computing $g_{yy}(z)$ for $y_t = (1 - \alpha U)x_t$.

Equation (4.5) is of the form $\sigma^2 \dfrac{(1 - \beta z)(1 - \beta z^{-1})}{(1 - \alpha z)(1 - \alpha z^{-1})}$ where

(4.6) $$\sigma^2\beta = \lambda\,\alpha\,\sigma_1^2$$

$$\sigma^2(1 + \beta^2) = \sigma_1^2 + \lambda\sigma_1^2(1 + \alpha^2)$$

Equation (4.6) is derived by equating coefficients of like powers of z in (4.5). So

(4.7) $$\sigma^2 = \frac{\lambda\,\alpha\,\sigma_1^2}{\beta}$$

and

(4.8) $$\frac{\lambda\,\alpha\,\sigma_1^2}{\beta}(1 + \beta^2) = \sigma_1^2 + \lambda\,\sigma_1^2\,(1 + \alpha^2)$$

Giving

(4.9) $$\lambda\,\alpha\,\sigma_1^2 - (\sigma_1^2 + \lambda\sigma_1^2\,(1 + \alpha^2))\beta + \lambda\,\alpha\,\sigma_1^2\,\beta^2 = 0$$

a quadratic equation in β . Solving (4.9) and taking $|\beta| = 1$ we get:

$$(4.10) \qquad \beta = \frac{\sigma_1^2(1 + \lambda(1 + \alpha^2) - \sqrt{(\sigma_1^2(1 + \lambda(1 + \alpha^2))^2 - 4(\lambda\,\alpha\,\sigma_1^2)^2}}{2\,\lambda\,\alpha\,\sigma_1^2}$$

$$= \frac{1 + \lambda(1 + \alpha^2) - \sqrt{1 + 2\,\lambda(1 + \alpha^2) + \lambda^2(1 - \alpha^2)^2}}{2\,\lambda\,\alpha}$$

$$= \frac{1 + \lambda(1 + \alpha^2) - \Delta}{2\,\lambda\,\alpha}, \quad \text{where} \quad \Delta = (1 + 2\,\lambda(1 + \alpha^2) + \lambda^2(1 - \alpha^2)^2)^{\frac{1}{2}}$$

Using β as given in (4.10) we now have a process of the same form as the example of Section III, and we may apply (3.7) directly.

The problem of predicting $x_t = T_t + S_t + I_t$ is basically the same as the example just given. For the more complicated model we must factor a 52 degree polynomial to obtain $B(z)$ Once this is done we may apply (2.15) or some form of it, and the problem is solved. For a discussion of several convenient forms of (2.15) the reader is referred to [1].

V. PREDICTION OF A COMPONENT OF A TIME SERIES

In some cases we may not wish to predict the observed sequence $\{x_t\}$, but wish to predict some unobserved component of the series e.g. the extraction of a signal in the presence of noise.

Suppose

$$(5.1) \qquad x_t = u_t + \eta \; ; \quad \{u_t\} \; \{\eta_t\} \quad \text{uncorrelated.}$$

In (5.1) $\{u_t\}$ is the signal and $\{\eta_t\}$ is the noise. It is important to note that $\{\eta_t\}$ need not be white noise. Thus for the model in [ ] we could identify $\{u_t\}$ with $\{T_t\}$ and $\{\eta_t\}$ with the seasonal and irregular components $\{S_t + I_t\}$. For the predictor of $u_{t+\nu}$ in (5.1) we have:

(5.2) $$\gamma_u(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{ux}(z)}{z^{\nu} B(z^{-1})} \right]_+ = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{uu}(z) + g_{u\eta}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

$$= \frac{1}{\sigma^2 B(z)} \left[\frac{g_{uu}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

from the assumed independence of $\{u_t\}$ and $\{\eta_t\}$. Similarly for $\hat{\eta}_{t+\nu}$:

(5.3) $$\gamma_\eta(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{\eta\eta}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

Summing the function in (5.2) and (5.3) we see that:

(5.4) $$\gamma_u(z) + \gamma_\eta(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{uu}(z)}{z^{\nu} B(z^{-1})} \right]_+ + \frac{1}{\sigma^2 B(z)} \left[\frac{g_{\eta\eta}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

$$= \frac{1}{\sigma^2 B(z)} \left[\frac{g_{uu}(z) + g_{\eta\eta}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

$$= \frac{1}{\sigma^2 B(z)} \left[\frac{g_{xx}(z)}{z^{\nu} B(z^{-1})} \right]_+$$

$$= \gamma_x(z)$$

So

(5.5) $$\hat{x}_{t+\nu} = \hat{u}_{t+\nu} + \hat{\eta}_{t+\nu}$$

This argument is easily generalized and we see that if the process $\{x_t\}$ is the sum of n independent processes $\{y_t^i, i = 1, 2, \dots n\}$ then

(5.6) $$\hat{x}_{t+\nu} = \sum_{i=1}^{n} \hat{y}^i_{t+\nu}$$

If in the process (5.1) the sequence $\{\eta_t\}$ is in fact white noise, then $g_{\eta\eta}(z) = \sigma_\eta^2$ and we see that in (5.3) $\gamma_\eta(z) = 0$ for $\nu > 0$. This is quite plausible as $\{\eta_t\}$ an uncorrelated sequence means that its past contains no information about its future so the best predictor of the process is simply its mean. For this case $\hat{\lambda}_{t+\nu} = \hat{u}_{t+\nu}$, $\nu > 0$. For $\nu \leq 0$ the result does not hold and in general $\hat{\eta}_{t+\nu}$ is different from zero for $\nu \leq 0$. In the former case (5.2) gives as alternative means of determining $\hat{x}_{t+\nu}$. The reader is warned that the use of (5.2) is likely to be more difficult than simply applying (2.15). To use either formula one must determine $B(z)$. Once $B(z)$ is known it is likely that (2.15) will be the simpler formula to apply.

To illustrate the application of (5.2) we return to the example of Section IV (cf. [4] pages 69-70).

(5.7) $$x_t = u_t + \eta_t$$

$$u_t = \frac{\epsilon_t}{1 - \alpha U} \quad |\alpha| < 1, \ \eta_t = \epsilon'_t$$

$$g_{uu}(z) = \frac{\sigma_1^2}{(1 - \alpha z)(1 - \alpha z^{-1})}, \ g_{\eta\eta}(z) = \lambda\sigma_1^2$$

$$g_{xx}(z) = \frac{\sigma^2(1-\beta z)(1-\beta z^{-1})}{(1-\alpha z)(1-\alpha z^{-1})}$$

where β and σ^2 are the values given in (4.7) and (4.10).

For the prediction of $u_{t+\nu}$ we have

$$(5.8) \qquad \gamma(z) = \frac{(1 - \alpha z)}{\sigma^2(1 - \beta z)} \left[\frac{\sigma_1^2(1 - \alpha z^{-1})}{(1 - \alpha z)(1 - \alpha z^{-1})(1 - \beta z^{-1})z^{\nu}} \right]_+$$

$$= \frac{1 - \alpha z}{\sigma^2(1 - \beta z)} \left[\frac{\sigma_1^2}{z^{\nu}(1 - \alpha z)(1 - \beta z^{-1})} \right]_+$$

To evaluate (5.8) we expand the function inside the operator by partial fractions.

$$(5.9) \qquad \left[\cdot \right]_+ = \frac{\sigma_1^2}{1 - \alpha\beta} \left[\frac{1}{(1 - \alpha z)z^{\nu}} \right]_+ + \frac{\sigma_1^2}{1 - \alpha\beta} \left[\frac{\beta z^{-1}}{(1 - \beta z^{-1})z^{\nu}} \right]_+$$

For $\nu \geq 0$ the second term is zero so we have:

$$(5.10) \qquad \left[\cdot \right]_+ = \frac{\sigma_1^2}{1 - \alpha\beta} \left[\frac{1}{(1 - \alpha z)z^{\nu}} \right]_+ = \frac{\sigma_1^2 \alpha^{\nu}}{(1 - \alpha\beta)(1 - \alpha z)} \qquad \nu \geq 0$$

Substituting (5.10) in (5.8) and cancelling we get:

$$(5.11) \qquad \gamma(z) = \frac{\sigma_1^2 \alpha^{\nu}}{\sigma^2(1 - \alpha\beta)(1 - \beta z} \qquad \nu \geq 0$$

From the argument given above (5.11) should agree with the results obtained for the predictor of $x_{t+\nu}$ in Section III and Section IV. Recalling that

$$(5.12) \qquad \gamma_x(z) = \frac{(\alpha - \beta)\alpha^{\nu-1}}{(1 - \beta z)}$$

we see that the two results will agree if

(5.13) $$\frac{\sigma_1^2}{\sigma^2} = (1 - \alpha\beta)(1 - \alpha^{-1}\beta)$$

Referring to the definition of β given in (4.10) we have:

(5.14) $$(1 - \alpha\beta)(1 - \alpha^{-1}\beta) = \left[\frac{2\lambda - \lambda(1 + \alpha^2) - 1 + \Delta}{2\lambda}\right]\left[\frac{2\lambda\alpha^2 - 1 - \lambda(1 + \alpha^2) + \Delta}{2\lambda\alpha^2}\right]$$

$$= \left[\frac{\Delta - 1 + \lambda(1 - \alpha^2)}{2\lambda}\right]\left[\frac{\Delta - 1 - \lambda(1 - \alpha^2)}{2\lambda\alpha^2}\right]$$

$$= \frac{(\Delta - 1)^2 - \lambda^2(1 - \alpha^2)^2}{4\lambda^2\alpha^2}$$

Expanding (5.14) and remembering the definition of Δ we see that (5.14) becomes:

$$(1 - \alpha\beta)(1 - \alpha^{-1}\beta) = \frac{1 + 2\lambda(1 + \alpha^2) + \lambda^2(1 - \alpha^2) + 1 - 2\Delta - \lambda^2 + 2\lambda^2\alpha^2 - \lambda^2\alpha^4}{4\lambda^2\alpha^2}$$

(5.15) $$= \frac{2\{1 + \lambda(1 + \alpha^2) - \Delta\}}{4\lambda^2\alpha^2}$$

$$= \frac{\beta}{\lambda\alpha}$$

From (4.7) we see that

(5.16) $$\frac{\beta}{\lambda\alpha} = \frac{\sigma_1^2}{\sigma^2}$$ which proves (5.13).

To obtain $\hat{u}_{t+\nu}$ for $\nu \leq 0$ we return to (5.9), writing μ for $-\nu$.

(5.17) $$\left[\frac{\sigma_1^2 z^\mu}{(1 - \alpha z)(1 - \beta z^{-1})}\right]_+ = \frac{\sigma_1^2}{1 - \alpha\beta}\left[\frac{z^\mu}{1 - \beta z} + \frac{\beta z^{\mu-1}}{(1 - \beta z^{-1})}\right]_+$$

$$= \frac{\sigma_1^2 z^\mu}{(1 - \alpha\beta)(1 - \alpha z)} + \frac{\beta\sigma_1^2}{1 - \alpha\beta}\left[\frac{z^{\mu-1}}{1 - \beta z^{-1}}\right]_+$$

as the first term on the right contains only nonnegative powers of z.
Expanding the second term and taking out the nonnegative powers of z we get:

(5.18) $$\left[\frac{\sigma_1^2 z^\mu}{(1 - \alpha^2)(1 - \beta z^{-1})}\right]_+ \quad \frac{\sigma_1^2 z}{(1 - \alpha\beta)(1 - \alpha z)} + \frac{\sigma_1^2\beta}{1 - \alpha\beta}\left(z^{\mu-1}\sum_{j=0}^{\mu-1}\beta^j z^{-j}\right)$$

$$= \frac{\sigma_1^2}{1 - \alpha\beta}\left\{\frac{z^\mu}{1 - \alpha z} + \beta z^{\mu-1}\left(\frac{1 - \beta^\mu z^{-\mu}}{1 - \beta z^{-1}}\right)\right\}$$

$$= \frac{\sigma_1^2}{1 - \alpha\beta}\left\{\frac{z^\mu}{1 - \alpha z} + \beta\left(\frac{z^\mu - \beta^\mu}{z - \beta}\right)\right\}$$

Substituting back into (5.8) and using (5.13) we finally obtain

(5.19) $$\gamma(z) = \frac{\sigma_1^2(1 - \alpha z)}{\sigma^2(1 - \beta z)(1 - \alpha\beta)}\left\{\frac{z^{-\nu}}{1 - \alpha z} + \beta\left(\frac{z^{-\nu} - \beta^{-\nu}}{z - \beta}\right)\right\}$$

$$= \frac{(1 - \alpha^{-1}\beta)}{1 - \beta z}\left\{z^{-\nu} + \beta\frac{(z^{-\nu} - \beta^{-\nu})(1 - \alpha z)}{z - \beta}\right\}$$

We have covered this example in detail to let the reader get some feel for "what is going on." In the next section we will apply the same method to the more complicated model introduced in [1]. For this latter model the manipulations are a good deal more complicated, and it is easy to miss what is happening.

VI. ILLUSTRATION OF COMPONENT PREDICTION FOR A "NEARLY" NONSTATIONARY ECONOMIC TIME SERIES

We now consider the problem of signal extraction for the model introduced in [1].

(6.1) $$X_t = T_t + S_t + I_t$$

$$= \frac{u_t}{(1-\rho U)^2} + \frac{v_t}{(1-\theta U^L)^2} + w_t \qquad |\rho|, |\theta| < 1$$

$\{u_t\}$, $\{v_t\}$, $\{w_t\}$ mutually uncorrelated noise sequences with variances σ_u^2 , σ_v^2 , σ_w^2 respectively. We have introduced a new parameter θ in this form of the model. The purpose is to insure that all the series in the manipulations to follow are convergent. When we have the forms we need, we simply let θ tend to one from below to obtain the solution for the model as originally formulated.

We will consider only prediction $(\nu > 0)$ of the separate components of $\{X_t\}$, so since $\hat{I}_{t+\nu} = 0$, $\nu > 0$ we need examine only predictors for the sequences $\{S_t\}$ and $\{T_t\}$.

(6.2) $$g_{XX}(z) = g_{SS}(z) + g_{TT}(z) + g_{II}(z)$$

$$= \frac{\sigma_u^2}{|1-\rho z|^4} + \frac{\sigma_\nu^2}{|1-\theta z^L|^4} + \sigma_w^2 = \sigma^2 \frac{P(z)\,P(z^{-1})}{|1-\rho z|^4 |1-\theta z^L|^4}$$

$$B(z) = \frac{P(z)}{(1-\rho z)^2(1-\theta z^L)^2}$$

To determine $\hat{T}_{t+\nu}$

$$(6.3)\quad \gamma(z) = \frac{(1-\rho z)^2(1-\theta z^L)^2}{\sigma^2 P(z)} \left[\frac{\sigma_u^2(1-\theta z^{-L})^2(1-\rho z^{-1})^2}{(1-\rho z^{-1})^2(1-\rho z)^2 P(z^{-1})z^\nu}\right]_+$$

$$= \frac{(1-\rho z)^2(1-\theta z^L)^2}{\sigma^2 P(z)} \left[\frac{\sigma_u^2(1-\theta z^{-L})^2}{P(z^{-1})(1-\rho z)^2 z^\nu}\right]_+$$

$$= \frac{(1-\rho z)^2(1-\theta z^{-L})^2}{\sigma^2 P(z)} \left[\frac{Q(z)}{(1-\rho z)^2}\right]_+$$

where

$$(6.4)\qquad Q(z) = \frac{\sigma_u^2(1-\theta z^{-L})^2}{z^\nu P(z^{-1})}$$

The problem now is to evaluate (6.3). To do this we shall use Theorem 1, page 93 of [4] which is approximately as follows:

Let $Q(z)$ be a function of z analytic in $\delta < |z| < \delta^{-1}$ and let φ be a number such that $|\varphi| < 1$, then

$$R(z) = (1-\varphi z)^p \left[\frac{Q(z)}{(1-\varphi z)^p}\right]_+ = \Pi_p(z) + \left[Q(z)\right]_+$$

where

$$\Pi_p(z) = \sum_{i=0}^{p-1} \frac{1}{i!} \left.\frac{d^i\left[Q(z)\right]}{dz^i}\right|_{z=1/\varphi} (z-\varphi^{-1})^i$$

From (6.4) we see that for $\nu > 0$

(6.5) $$\left[Q(z)\right]_+ = 0$$

So for this case the theorem gives

(6.6) $$\gamma(z) = \frac{(1-\theta z^L)^2 (1-\rho^2)}{\sigma^2 P(z)} \left[\frac{Q(z)}{(1-\rho z)^2}\right]_+$$

$$= \frac{(1-\theta z^L)^2}{\sigma^2 P(z)} \left[\pi_o + \pi_1 (z-\varphi^{-1})\right]$$

(6.7) $$\pi_o = Q(\varphi^{-1}),\ \pi_1 = Q'(\varphi^{-1})$$

To obtain the equivalent expression for the predictor of $S_{t+\nu}$ is a good deal more difficult. For this case

(6.8) $$\gamma(z) = \frac{(1-\theta z^L)^2 (1-\rho z)^2}{\sigma^2 P(z)} \left[\frac{\sigma_\nu^2(1-\rho z^{-1})^2}{z^\nu P(z^{-1})(1-\theta z^L)^2}\right]_+$$

We could take

(6.9) $$(1-\theta z^L)^2 = \prod_{i=1}^{L} (1-\lambda_i z)^2$$

where the λ_i are the L^{th} roots of θ, and apply the theorem quoted above using:

(6.10) $$Q(z) = \frac{\sigma_\nu^2(1-\rho z^{-1})^2}{z^\nu P(z^{-1}) \prod_{i \neq j} (1-\lambda_i z)^2}$$

The difficulty is that for this case $\left[Q(z)\right]_+$ is not zero, so we would have to evaluate $\left[Q(z)\right]_+$ and $\left[Q(z)\right]_-$. We can find $\gamma(z)$ for this case by repeated application of the theorem (see Appendix 1).

The straightforward approach would be to calculate the coefficients of the function inside the $\left[\quad\right]_+$ operator in (6.8). The trouble with this is that each coefficient of z in $\frac{(1-\rho z^{-1})^2}{z^{\nu}P(z^{-1})(1-\theta z^{L})^2}$ is the sum of an infinite series. Direct numerical computations will thus be only approximate and cannot take advantage of simplifications that can be found analytically. Further, as is shown in Appendix 1 and Appendix 2:

$$(6.11)\qquad \left[\frac{(1-\rho z^{-1})^2}{z^{\nu}\, P(z^{-1})(1-\theta z^{L})^2}\right]_+ = \frac{G(z)}{(1-\theta z^{L})^2} \quad \text{where} \quad G(z) = \sum_{i=0}^{2L-1} g_i\, z^i$$

Thus

$$(6.12)\qquad \gamma(z) = \frac{\sigma_\nu^2}{\sigma^2}\ \frac{(1-\rho z)^2\, G(z)}{P(z)}$$

a considerable simplification and especially important when θ tends to 1 from below.

The actual calculations are long and tedious. Numerous substitutions are required which make the argument somewhat difficult to follow. The essential elements of the problem are all contained in the example given in Section V. The reader interested in a detailed description of the calculations is referred to Appendix 2.

VII. PREDICTION WITH A FINITE PAST

So far we have assumed that the entire past of the process used in predicting was known. We now consider the case where only a finite number of observations are available. We shall consider two cases. First, we assume $\{x_t\}$ can be represented as a finite autoregression. Second, we consider the prediction of a finite moving average process. A complete discussion of these problems is given in [4] Chapter 7.

Suppose

$$(7.1) \qquad A(U)x_t = \epsilon_t \qquad \text{where} \quad A(z) = \sum_{i=0}^{p} a_i z^i$$

consider the generating function for the following sum:

$$(7.2) \qquad \hat{x}_{t+\nu} + a_1 \hat{x}_{t+\nu-1} + \dots + a_p x_{t+\nu-p} = P(U)x_t$$

From (2.15) we have

$$(7.3) \qquad P(z) = A(z) \left\{ \left[\frac{1}{A(z)z^{\nu}}\right]_+ + a_1 \left[\frac{1}{A(z)z^{\nu-1}}\right]_+ + \dots + a_p \left[\frac{1}{A(z)z^{\nu-p}}\right]_+ \right\}$$

$$= A(z) \left\{ \left[\frac{1}{A(z)z^{\nu}}\right]_+ + \left[\frac{a_1}{A(z)z^{\nu-1}}\right]_+ + \dots + \left[\frac{a_p}{A(z)z^{\nu-p}}\right]_+ \right\}$$

$$= A(z) \left\{ \left[\frac{1}{A(z)z^{\nu}} + \frac{a_1 z}{A(z)z^{\nu}} + \dots + \frac{a_p z^p}{A(z)z^{\nu}}\right]_+ \right\}$$

$$= A(z) \left[\frac{A(z)}{A(z)z^{\nu}}\right]_+$$

$$= A(z) \left[\frac{1}{z^{\nu}}\right]_+$$

$$= 0 \quad \text{for} \quad \nu > 0$$

So (7.2) and (7.3) imply that:

(7.4) $$\hat{x}_{t+\nu} = -\left[a_1 \hat{x}_{t+\nu-1} + \dots + a_p \hat{x}_{t+\nu-p}\right] \quad \text{where } x_{t+\nu-k} = x_{t+\nu-R} \quad k \geq \nu$$

as

(7.5) $$\gamma(z) = A(z) \left[\frac{1}{A(z) z^{\nu-k}}\right]_+$$

But for $k \geq \nu$ the expression inside the $[\ ]_+$ operator already contains only nonnegative powers of z so

(7.6) $$\gamma(z) = z^{k-\nu} \quad \text{and} \quad \hat{x}_{t+\nu-k} = \gamma(U) x_t = x_{t+\nu-k}$$

For this process the predictor of $x_{t+\nu}$ depends upon only a finite number of observations, and so long as the number of observations is at least equal to the order of the autoregression no new restrictions are added, $\hat{x}_{t+\nu}$ being obtained by solving (7.2) recursively for $\nu = 1, 2, \dots r$.

Suppose now that

(7.7) $$x_t = B(U)\varepsilon_t \qquad \text{where } B(z) = \sum_{j=0}^{q} b_j z^j$$

We wish to find the least squares predictor for $x_{t+\nu}$ given by

(7.8) $$\hat{x}_{t+\nu} = \sum_{j=0}^{r-1} \gamma_j x_{t-j}$$

We proceed by minimizing

(7.9) $$E(x_{t+\nu} - \hat{x}_{t+\nu})^2 = E(x_{t+\nu} - \sum_{j=0}^{r-1} \gamma_j x_{t-j})^2$$

Differentiating with respect to the γ_j, setting first derivatives equal to zero.

$$(7.10) \qquad \sum_{k=o}^{r-1} \Gamma_{j-k}^{xx} \gamma_k = \mu_j \quad j = 0, 1, \ldots r-1 , \quad \text{where} \quad \mu_j = E(x_{t+\nu}\, x_{t-j})$$

Multiplying both sides of the j^{th} equation in (7.10) by z^j we get:

$$(7.11) \qquad \sum_{k=o}^{r-1} \Gamma_{j-k}^{xx} \gamma_k z^j = \sum_{k=o}^{r-1} \Gamma_{j-k}^{xx} z^{j-k} \gamma_k z^k = \mu_j z^j$$

Summing over $j = 0, \pm 1, \pm 2, \ldots \pm q$ we have

$$(7.12) \qquad g_{xx}(z)\, \gamma(z) = \mu_1(z) + \mu(z) + \mu_2(z) \quad \text{where}$$

$$\mu_1(z) = \sum_{j=-q}^{-1} \mu_j z^j \quad \mu(z) = \sum_{j=o}^{n-1} \mu_j z^j \quad \mu_2(z) = \sum_{j=n}^{q=n-1} \mu_j z^j$$

Since (7.11) holds only for $j = 0, 1, \ldots n-1$ the functions $\mu_1(z)$ and $\mu_2(z)$ must be treated as unknown. Since $g_{xx}(z) = \sigma^2 B(z) B(z^{-1})$, (7.12) becomes:

$$(7.13) \qquad \sigma^2 B(z)\, B(z^{-1}) = \mu_1(z) + \mu(z) + \mu_2(z)$$

Solving for $\gamma(z)$ we have:

$$(7.14) \qquad \gamma(z) = \frac{\mu_1(z) + \mu(z) + \mu_2(z)}{\sigma^2 B(z)\, B(z^{-1})}$$

But $\gamma(z)$ is a polynomial and can have no poles. So every zero of the denominator of (7.14) is a zero of the numerator i.e.

$$(7.15) \qquad \mu_1(\lambda_i) + \mu(\lambda_i) + \mu_2(\lambda_i) = 0 \quad i = 1, 2, \ldots 2q$$

where

$$B(\lambda_i)\, B(\lambda_i^{-1}) = 0 .$$

Equations (7.15) give 2q equations in the 2q unknowns. Thus we can solve (7.14) for $\gamma(z)$.

To illustrate what is involved in using this method we consider another simple example([4] page 75, Example No. 7.2.1). Let

(7.16) $\quad x_t = \epsilon_t - b\epsilon_{t-1}$. Thus $q = 1$, $g_{xx}(z) = 1 + b^2 - bz - bz^{-1}$, $E(\epsilon_t^2) = 1$,

and

(7.17) $$\mu_j = E(x_{t+\nu} \cdot x_{t-j}) = -b, \quad j = 1$$
$$= 0 ,$$

So $$\mu(z) = -b, \text{ for } \nu = 1$$
$$= 0 \text{ otherwise.}$$

For $\nu > 1$ equation (7.14) is

(7.18) $$\gamma(z) = \frac{1}{\sigma^2}\left[\frac{\mu_{-1}z^{-1} + \mu_n z^n}{1 + b^2 - bz - bz^{-1}}\right]_+$$

and (7.15) becomes

(7.19) $$\mu_{-1}\, b^{-1} + \mu_n b^n = 0$$
$$\mu_{-1}\, b \quad + \mu_n b^{-n} = 0$$

so

$\mu_{-1} = \mu_n = 0$ and $\hat{x}_{t+\nu} = 0$ which agrees with (2.15) for this case.

For $\nu = 1$ equations (7.15) become

(7.20) $$\mu_{-1}b^{-1} + \mu_n b^n - b = 0$$
$$\mu_{-1}b \quad + \mu_n b^{-1} - b = 0$$

where $$\mu_{-1} = \frac{b^{n-1} - b^{-n+1}}{b^{n+1} - b^{-n-1}}, \quad \mu_n = \frac{b^{n-1}}{b^{n+1} - b^{-n-1}}$$

For simple processes such as this one the method is quick and straightforward. For more complicated models the computations become very involved. For the model discussed in the last section each prediction would require the solution of 52 equations in 52 unknowns. For such a model one may prefer to proceed as if the entire past of the process were available and truncate the prediction formula at γ_{n-1} (or earlier if the γ_j become sufficiently small). For cases like the one shown in this section that truncation is likely to do the least harm in the sense of increasing the mean square error.

For a single prediction one may prefer the methods described in this section. If one wants to compute a large number of predictors, using several periods of prediction and based on many different points in time, the simpler procedure has much to recommend it. In [1] for instance, the authors actually obtained about 3600 predictions.

VIII. AN APPLICATION OF SIGNAL EXTRACTION TO AN ECONOMIC TIME SERIES

The theory outlined above allows us to estimate the separate components of models such as the one introduced in [1]. Before describing the results in detail we consider two questions. First, in what kinds of cases can estimation of unobserved components of economic time series be useful? Second, what can we expect to see when we estimate the trend-cycle component of the unemployment series? In particular, can the estimates be considered to be a seasonally adjusted series?

In many cases it may not be helpful to assume that economic time series are composed of several independent components. There are some cases, however, where this approach may be useful. The most obvious area is where we believe that the observed series contains irrelevant material. Some examples are: measurement errors, transitory components, seasonality, and other forms of noise. In these situations it is assumed that the relations of economic theory hold among certain components of the observed series but not among others. Essentially, we are saying that the series contain both interesting and uninteresting components.

Obtaining seasonally adjusted series is similar to assuming that the observed series contains a signal that is corrupted by seasonal and irregular noise. If it is correct to view the process of seasonal adjustment this way, then there are advantages to treating explicitly the problem as signal extraction. First, in order to proceed with the estimation, the components models must be formulated. This requires the investigator to define fully what the separate components are and what meaning they have. An additional advantage of viewing the problem this way is that once the problem is formulated we know what to do. That is, we can now estimate components in a systematic way using a criterion such as least squares. Thus, a considerable body of theory can be applied to the problem of seasonal adjustment.

If certain economic time series do contain independent trend-cycle components, it is probably correct that governmental fiscal policies are primarily concerned with influencing the trend-cycle components. If this is so then it is quite reasonable that policy makers should wish to know the values of the series they are concerned with.

Not all economic time series can be described as partially interesting and partially uninteresting. There is no reason why the usefulness of the approach taken here is limited to these cases. Suppose one believes that a certain adjustment mechanism reacts differently to slow cumulative changes then it does to rapid short term changes. In this case separation of the relevant series into components may aid in studying the different adjustment processes.

Before trying to interpret the components of the model we must consider the model in detail.

(8.1) $$X_t = T_t + S_t + I_t$$

(8.2) $$T_t = \frac{u_t}{(1-U)^2}, \quad S_t = \frac{v_t}{(1-U^L)^2}, \quad I_t = w_t$$

where $\{u_t\}$, $\{w_t\}$, $\{v_t\}$ mutually uncorrelated white noise sequences.

Whether the estimation of T_t can be interpreted as seasonal adjustment depends upon how we characterize seasonality. The seasonal component in this model consists of twelve independent processes of the form:

$$(8.3) \qquad S_t^i = \frac{u_t^i}{(1-\theta_i U^{12})} \qquad i = 1, 2, \dots 12$$

where

$$(8.4) \qquad \theta_i = \theta \qquad i = 1, 2, \dots 12$$

$$E(u_t^i) = \sigma_i^2 = \sigma^2 \qquad i = 1, 2, \dots 12$$

It is important to note that if S_t is of the form (8.3) then (8.4) must hold. If the variances σ_i^2 or the parameters θ_i were not equal for every process, then X_t would not be covariance stationary. Its second moments would depend upon the month of the year. While it is true that if the seasonal processes are independent they must be identical, it is not true that they must be independent. For instance, one could have:

$$(8.5) \qquad S_t = \alpha S_{t-L} + \sum_{i=0}^{k} \beta_i \eta_{t-i}$$

where $\{\eta_t\}$ is some covariance stationary process. A model such as (8.5) could show the usual characteristics of a seasonal process i.e. peaks in its spectrum as the so-called seasonal frequencies, e.g. for monthly data the twelve 12th roots of unity. Consider (8.5) with $k=1$. A model such as this would be appropriate if one felt that the change in last month's seasonal provides information about this month's seasonal. For example, if one believes that "unusually" large Christmas shopping in November implies something about the December rush, then a model like (8.5) would be correct.

If in (8.3) we let θ tend to one from below we obtain the seasonal component used in fitting the unemployment series. The twelve seasonal processes are nonstationary. Their means are zero, but they have infinite variances. Thus we "expect" them to wander around a great deal. In fact there are poles at each of the seasonal frequencies, which includes the origin. If $|\theta|$ less than one then the spectrum of the component will have finite peaks of equal height at each seasonal frequency. If a model similar to (8.5) is used with $|\alpha|$ less than one, the result is to lower the power at the low frequencies relative to the power at the higher frequencies. Unless one assumes that the seasonal component is perfectly periodic, one can expect to find some power at all frequencies.

In specifying a model we must specify the forms of the components and their parameters which include the ratios of the variances of the noise inputs. In the model we are considering here the variance of $\{v\}$ is larger than σ_u^2 or σ_w^2. In this model then the seasonal component is by far the most important. That is it accounts for almost all of the variance of X_t.

In estimating T_t we are also removing the irregular component I_t. Including some irregular component is quite reasonable. There are many examples of events whose effects cannot be sensibly be considered as a change in trend or in seasonal pattern. Obvious examples include strikes, international incidents like the Suez Crisis, and sudden losses of electric power. In some cases assuming that I_t is white noise may not be appropriate. The effects of some large shocks may be observed over several time periods. To account for these cases we might prefer:

(8.6) $$I_t = \sum_{i=0} b_i \epsilon_{t-i}$$

If the b_j in (8.6) are small, then the process would look like white noise except when a large value of ϵ_t occurs.

We turn now to the trend-cycle component T_t. The spectrum of this process has maximum power at zero. The power declines continously to its minimum at π. Any such process may be called relatively smooth. That is, most of its variance is accounted for by the low frequency components. More precisely:

(8.7) $$f(\omega_0) > f(\omega_1) \longleftrightarrow \omega_0 \quad \omega_1, \omega_0, \omega_1 \in [0, \pi]$$

As $\rho \to 1, f(0) \to \infty$ and $f(\pi) \to \frac{\sigma^2}{8\pi}$

Thus large values of ρ tend to accentuate the dominance of the lowest frequencies.

We expect that most of the variance of T_t will be the result of slow cumulative movements. However, this process is stationary and we do not "expect" it to wander far. For many series with strong trends one may prefer to assume that the trend-cycle is nonstationary. For instance:

(8.8) $$T_t = \frac{u_t}{(1-U)(1-\rho U)}$$

Since data are available only for finite intervals and often only for short intervals, it may be that using models that are stationary would do just as well. For instance we could have assumed that:

$$S_t = \frac{v_t}{(1-\theta U^{12})^2} \tag{8.9}$$

where $|\theta|$ is close to one. This introduces some helpful simplifications. For example, if we felt that the seasonal component should have accounted for most of the variation in the model say 90%, the parameters θ and σ_v^2 could have been chosen accordingly.

Having described the model we now turn to the problem of interpreting the separate components. In the preceding discussion we have purposely avoided suggesting interpretations of the individual components. The reason is that it is not possible to interpret any one of them without reference to the others. Though we have referred to the components as "trend-cycle," "seasonal," and "irregular," the reader is warned that he should not allow these terms to make him assume that the components represent the effects of long run, short run, or, heaven help us, middle run forces.

The first thing to note about this or any other model of this type is that each component possesses some power at every frequency. Thus no single frequency band can be associated with any particular component. In this model we cannot safely assume that the trend-cycle will account for the long term movements of the series. This is true for any model in which some other component has high power at the low frequencies.

In the early 1950's drive-in theaters were first introduced in large numbers. The effect of this innovation was a considerable increase in motion picture attendance. This was not simply an increase in attendance, however, but also a change in the seasonal pattern. Previously, the summer months had been months of low attendance. The introduction of drive-ins increased attendance in the summer months until the summer was the period when attendance was highest. How would this change be described by this model? Would this be as upturn in the trend-cycle and some change in the seasonal pattern? Or is it likely that the trend-cycle would be unaffected and the seasonal pattern would change? In this model the entire seasonal process might wander off to a higher level as well as change its pattern. Examples of this kind show the difficulty one can have interpreting components by themselves. Also, examples like this show that even with this univariate approach, which involves only one time series, these models are more than purely descriptive. If one believes that certain kinds of changes should be reflected in the behavior of some given component, then the structure of the model must be chosen appropriately. Models like the one we are dealing with must pass stringent tests. It is not enough that they fit the observed data. Determining the adequacy of this kind of a model is made more difficult by dealing with a single series instead of dealing with a set of related series, since we have little in the way of a priori knowledge we can call on. With a set of related series we would have some a priori knowledge of the relations that should hold between the signals.

In Section VI we showed how to find the sequence γ, such that $\hat{T}_{t+\nu} = \sum_{j=0}^{\infty} \gamma_j^{\nu} x_{t-j}$ is the least squares predictor of $T_{t+\nu}$. To apply the theory discussed above to this model sequences of γ_j^{ν}'s were calculated for ν=0, 1, 3, 6, 8, 12. The coefficients were calculated under the assumption that an infinite number of observations were to be used. The prediction formulas were simply truncated. A total of 240 observations were used. So there is one predictor based on 240 observations, one based on 239, and so on. The estimates can be expected to become less reliable as the end of the series is approached.

The results for ν=0 are shown in Figure 1. For larger values of ν the series are similar. As ν increases the predictions flatten out and in general are closer to zero. Only the estimates based on the larger number of estimates are shown. It is probably fair to say that the estimates do not look like a seasonally adjusted version of the unemployment series.

When estimating an unobserved sequence it is difficult to see if the estimates are behaving well. We can, however, get some evidence from their spectrum. Figure 2 shows the theoretical spectrum of the estimates derived from the model. Figure 3 shows the estimated spectrum of the trend estimates. The spectral estimates were based on 210 observations. This means that the spectral estimates are from trend estimates based on at least 30 observations. Though the empirical spectrum has considerable unpredicted power at some frequencies, it behaves in general quite well. We cannot rely on the spectral estimates to tell us much here. By using a smaller number of observations and fewer lags in estimating the spectrum we may very well be able to get smoother spectral estimates. Similarly, using all the trend estimates and more lags we can probably get a more jagged estimated spectrum. The main conclusion from the spectra is that the estimates seem to behave about as expected.

Is it possible that the estimation procedure used will not work? To answer this question we generated a random series that fit the specifications of the model. A random number generator was used to generate two series of numbers independently and uniformly distributed on the unit interval. These numbers were then converted to standard normal variables by

(8.10) $$\epsilon_t = (-2 \log (u_1))^{\frac{1}{2}} \cdot \cos (2\pi u_2)$$

[cf. [3] p. 334]

This series was then used to generate three random variables with variances 40, 100, and 50 to conform to the ratios specified by the model. The sample means and variances were computed and they were almost exactly as specified. The separate components of the model were then computed using sufficient initial conditions. For example the trend-cycle component was calculated from

(8.11) $$T_t = 1.5 - T_{t-1} - 0.5625T_{t-2} + u_t \quad \text{where} \quad T_0 = T_{-1} = 0$$

For the seasonal component 24 starting values were arbitrarily chosen. They were arbitrarily chosen. They were all close to zero. The largest in absolute value was 2.9.

The random series was then fed into the same prediction routine as had been used for the unemployment series. The only difference was that the maximum number of observations used was 228. The results are shown in Figure 4. Only the "latest" estimates are shown. The figure shows that we can expect to get some fairly decent estimates. Figure 5 shows the random series and the trend. Considering the amount of noise in this series the estimates are quite good. Figure 5 also illustrates our earlier statement that in this model the seasonal component is by far the most important. Remembering that the series was started out near zero we can see that it is evolving rapidly, and that most of the evolution is due to the seasonal component. If this can be interpreted as typical behavior for this model then the estimates shown in Figure 1 are easier to understand. It would appear that we got back just what we put in. If the model is an adequate description of the unemployment series, then the estimates are quite reasonable.

The evidence that this model is appropriate is of two kinds. First, the model fits well. That is, it does a good job of predicting the series. Second, the spectrum of $X_t - 2X_{t-12} + X_{t-24}$ is extremely close to that predicted by the model.

Suppose we estimate the spectrum of the unemployment series using different filters. Will the results still agree with those predicted by the model? The problem here is one of considerable importance. Suppose that we estimate the spectrum of $y_t = \sum_{j=o}^{k} \alpha_i X_{t-j}$, and assume that the estimates are correct. This implies that we know the spectrum of $z_t = \sum_{j=o}^{L} \beta_j y_{t-j}$

As

$$(8.12) \qquad f_z(\omega) = | B(e^{i\omega}) |^2 f_y(\omega)$$

where

$$B(z) = \sum_{j=o}^{L} \beta_j z^j$$

Suppose that we compute z_t anyway and estimate its spectrum. Can we be sure that the results will agree? If the results are very different from what we "knew" they would be, then what can we say?

We wish to know if the spectrum of the unemployment series is similar to the theoretical spectrum of the model. Is it possible that the agreement between the spectra only hold when we apply the operator $(1 - U^{12})^2$? To answer this question we computed two series.

$$(8.13) \quad Y_t = X_t - 1.5\, X_{t-1} + .5625\, X_{t-2} - 2\, X_{t-12} + 3\, X_{t-13} - 1.135\, X_{t-14} - X_{t-24} - 1.5\, X_{t-25} + .5625\, X_{t-26}$$

$$(8.14) \quad W_t = X_t - X_{t-12}$$

Figure 6 shows the spectrum of Y_t , and Figure 7 shows the corresponding theoretical spectrum. The spectrum of W_t is given in Figure 8.

Now

$$(8.15) \quad W_t = x_t - x_{t-12} = T_t - T_{t-12} + S_t - S_{t-12} + I_t - I_{t-12}$$

So

$$(8.16) \quad f_W(\omega) = |1 - e^{-i12\omega}|^2 f_X(\omega)$$

$$= |1 - e^{-i12\omega}|^2 f_T(\omega) + |1 - e^{-i12\omega}|^2 f_S(\omega) + |1 - e^{-i12\omega}|^2 f_I(\omega)$$

since $\{T_t\}$, $\{S_t\}$, and $\{I_t\}$ are uncorrelated.

So

$$(8.17) \quad f_W(\omega) = \frac{2 - 2\cos 12\,\omega}{(1 + \rho^2 - 2\rho\cos\omega)^2}\,\sigma_u^2 + \frac{(2 - 2\cos 12\,\omega)}{(2 - 2\cos 12\,\omega)^2}\,\sigma_v^2 + (2 - 2\cos 12\omega)\sigma_w^2$$

$$= \frac{2 - 2\cos 12\,\omega}{(1 + \rho^2 - 2\rho\cos\omega)^2}\,\sigma_u^2 + \frac{\sigma_v^2}{2 - 2\cos 12\,\omega} + (2 - 2\cos 12\,\omega)\,\sigma_w^2$$

From (8.17) we see that the theoretical spectrum of W_t has poles at $\omega = \frac{\pi j}{6}$ $j = 0, 1, \ldots 11$. These are the so-called seasonal frequencies. But the estimated spectrum of W_t has dips at every seasonal frequency.

It is clear that the empirical results do not correspond with our theoretical expectations. The reason for this difficulty is not hard to see. When we estimate the spectrum of y_t we apply the operator $(1 - U^{12})$ to W_t. The transfer function for this filter is shown in Figure 9. Inspection of Figures 8 and 9 shows that the peaks in the transfer function correspond roughly to the peaks in the spectrum of W_t. Leakage from these peaks can result in estimating nonzero values of the spectrum at the seasonal frequencies. Similarly, giving some weight to the frequencies where power is low can lead to under estimates of the peaks in the spectrum of H_t. For a complete discussion of this problem the reader is referred to [3].

In discussing the estimates we have not commented on the fact that some of the estimates are negative. While one could argue that these results are improbable we have purposely avoided doing so. To see why suppose that we had assumed:

$$(8.15) \qquad x_t = e^{T_t} \cdot e^{S_t} \cdot e^{I_t}$$

In fact for seasonal adjustment it is often assumed that the seasonal factors enter in a multiplicative way. In this case we would have operated on the logarithms of the series and transformed back. It would then be impossible to have any negative values. We would still be faced with the problem of deciding if our estimates were meaningful. Furthermore, it is not impossible to interpret these negative results. One could argue that at some times we would have had over full employment of young males if it were not for seasonal and irregular forces. The plausibility of this argument could be tested by using our knowledge of the state of the world at the times in question. But this is exactly the kind of argument we have been trying to avoid. The trouble is that once we start this there is no end to it. We can also say the estimates should have risen or fallen at certain times. Eventually we end by constructing series of numbers that always go where we thought they would because we made them do it. There are far simpler ways of getting these kinds of series than signal extraction.

The technique of signal extraction can give useful insights into many significant problems. In this particular case the results do not appear to be helpful. The reason for the difficulty appears to be that the model tested does not really describe the unemployment series. This suggests that testing models for single series can be tricky. It is probably true that almost any model that allows for a strong seasonal factor would do a reasonably good job of fitting this series. Picking any one of these models may be aided by examining the implied behavior of their separate components.

FIGURE 1

X_t and $\hat{T}_{t,t}$

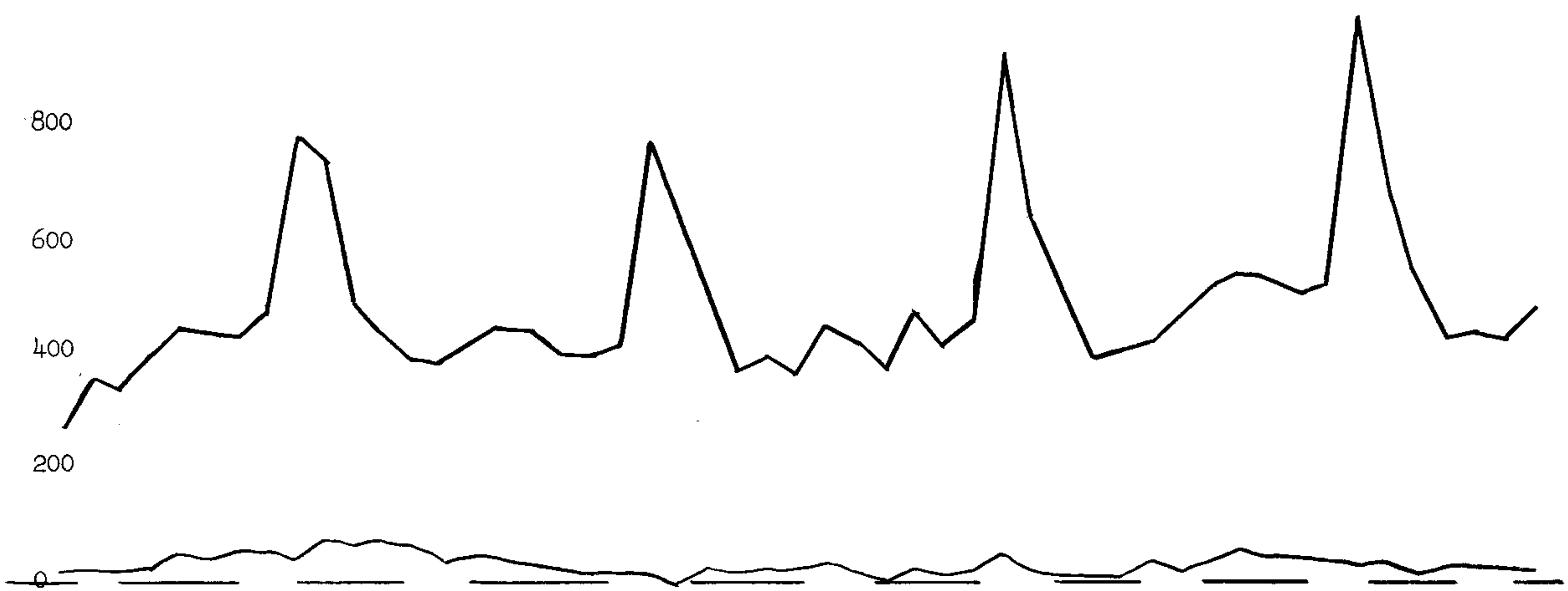


FIGURE 2

Predicted Spectrum of $\hat{T}_{t,t}$

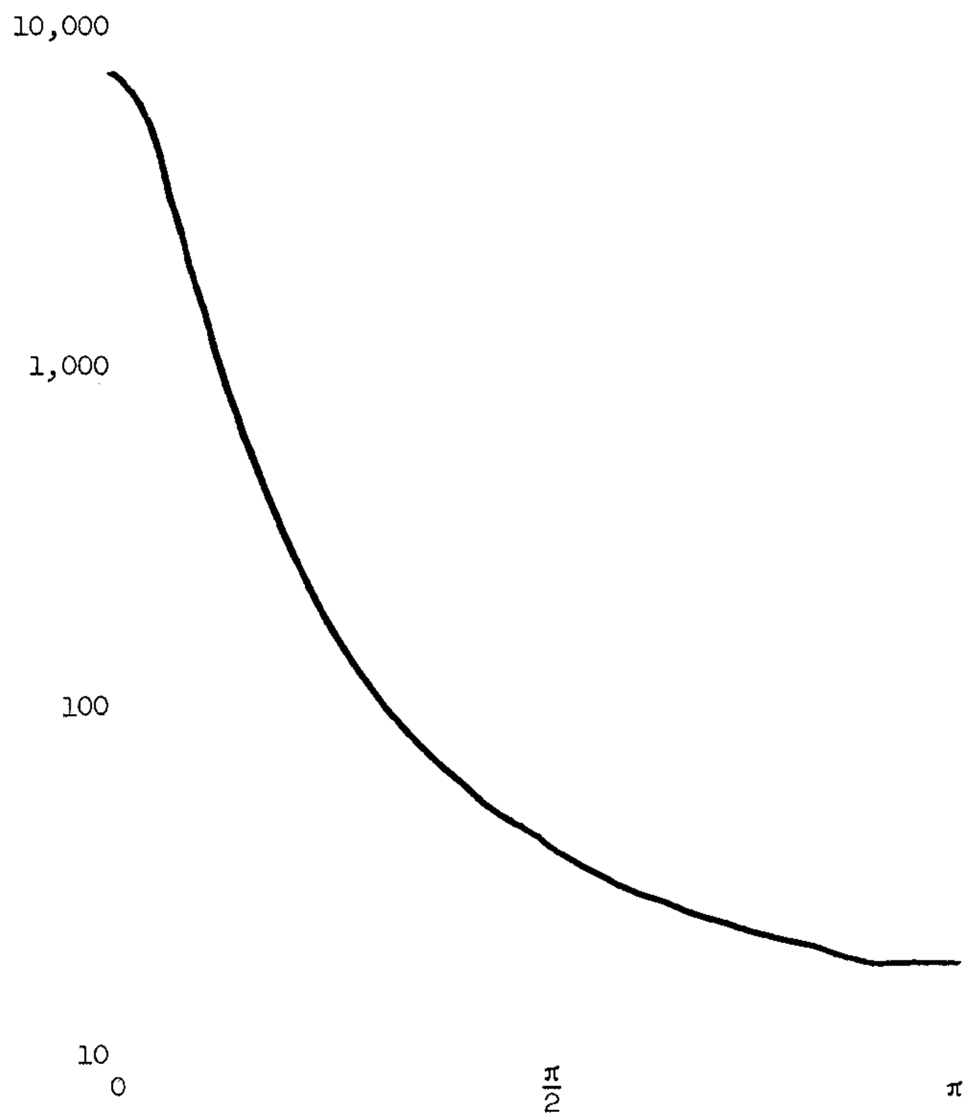


FIGURE 3

Spectrum of $\hat{T}_{t,t}$

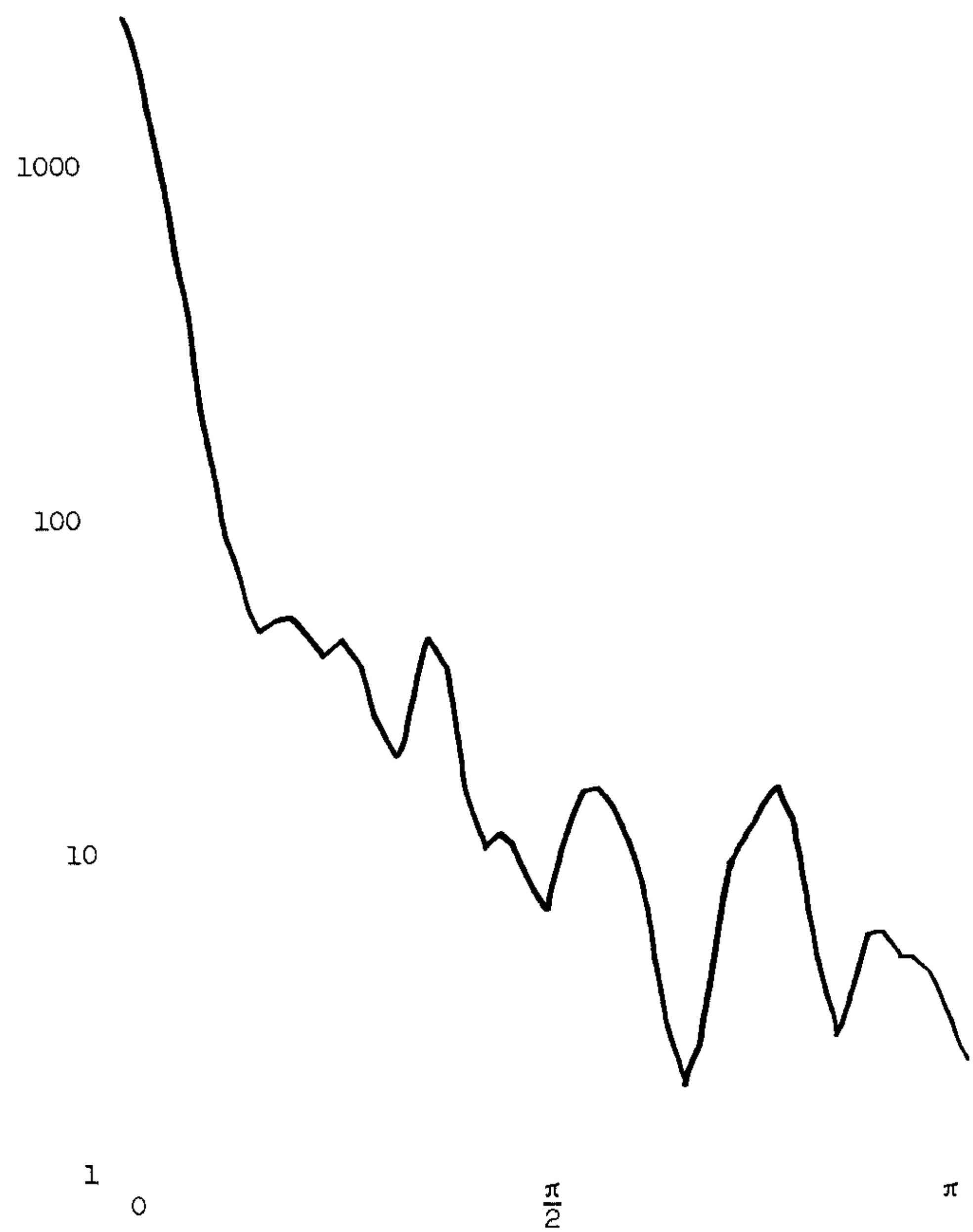


FIGURE 4

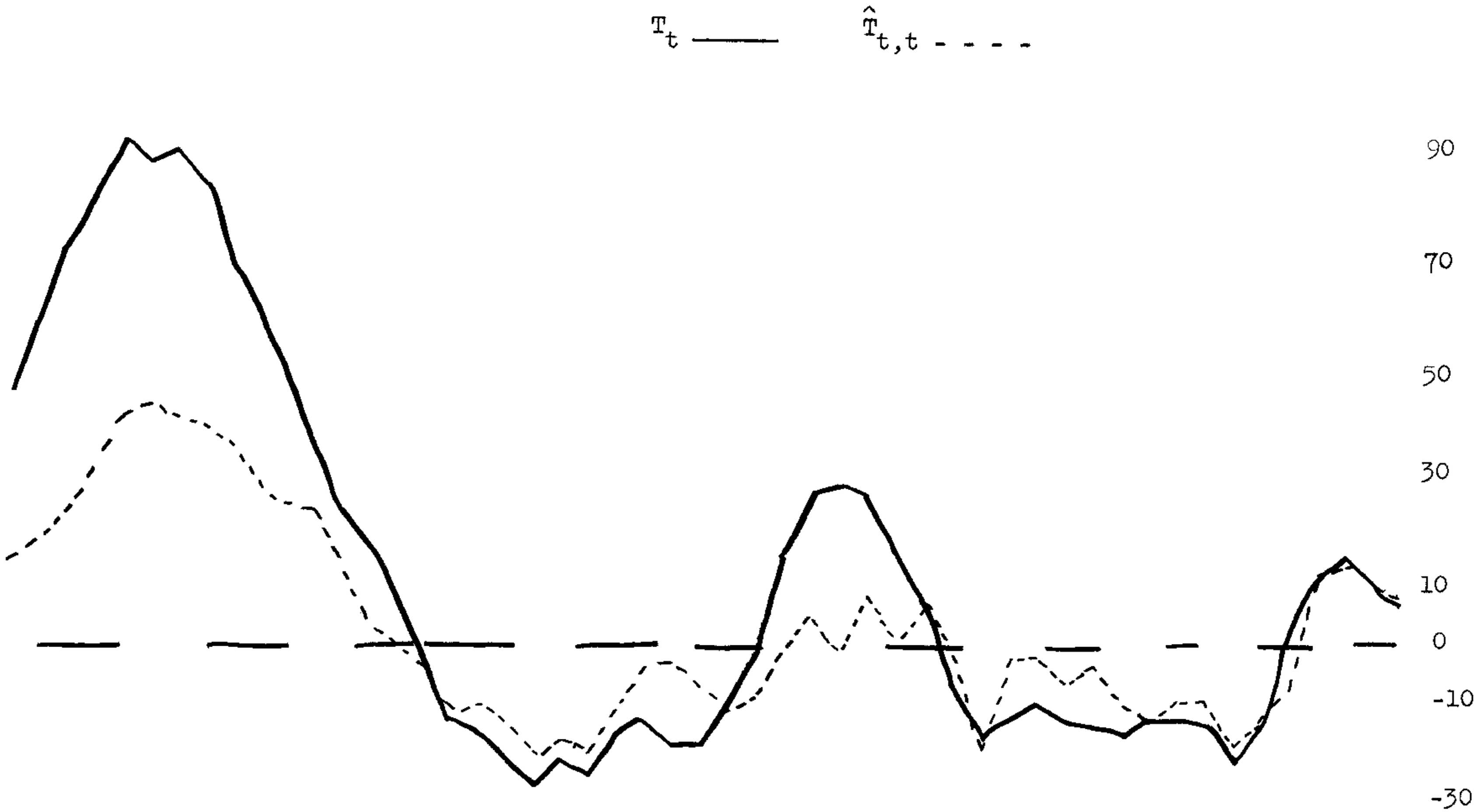


FIGURE 5

Random Series X_t and T_t

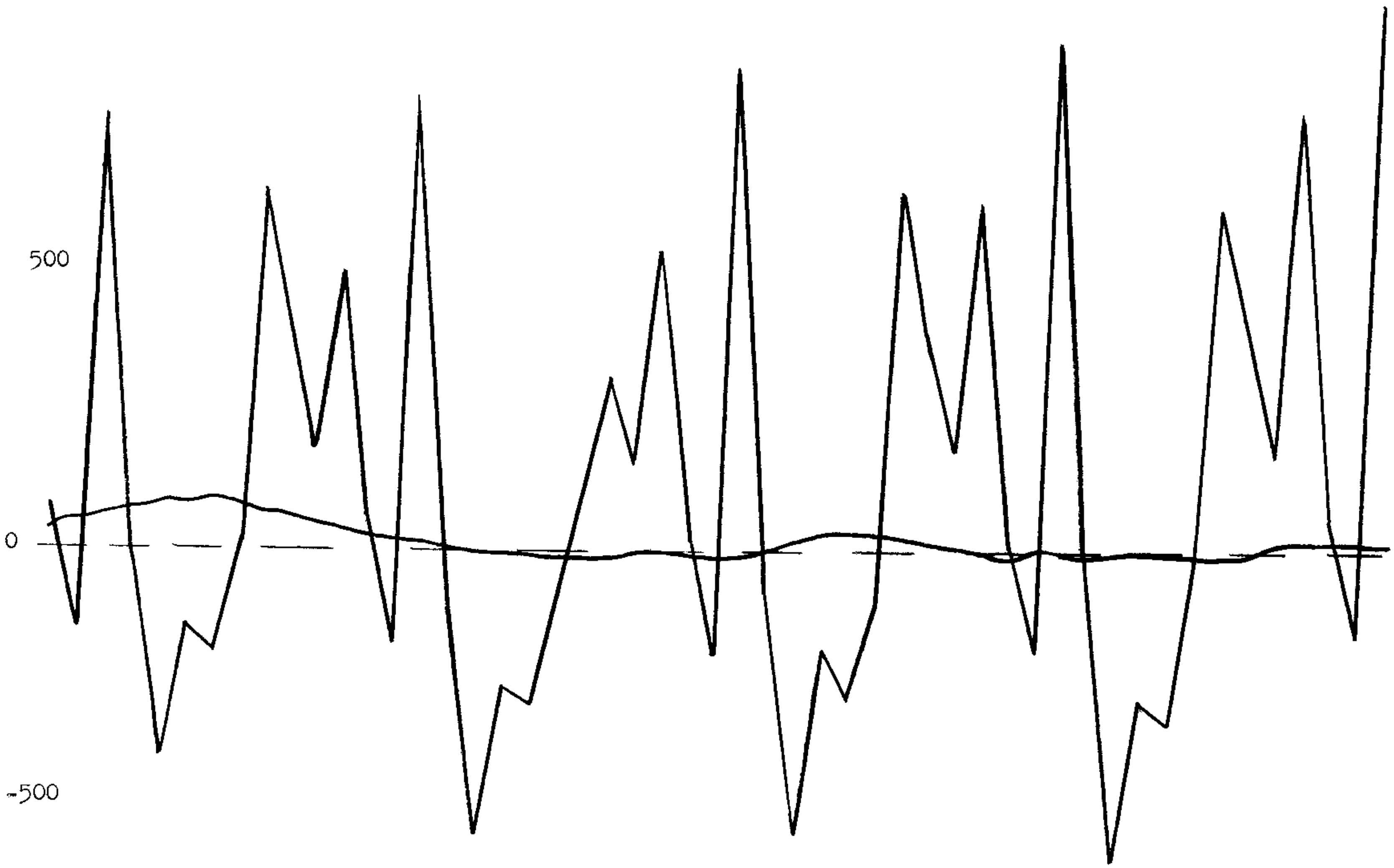


FIGURE 6

Spectrum of $(1-U^{12})^2 (1-\rho U)^2 X_t$

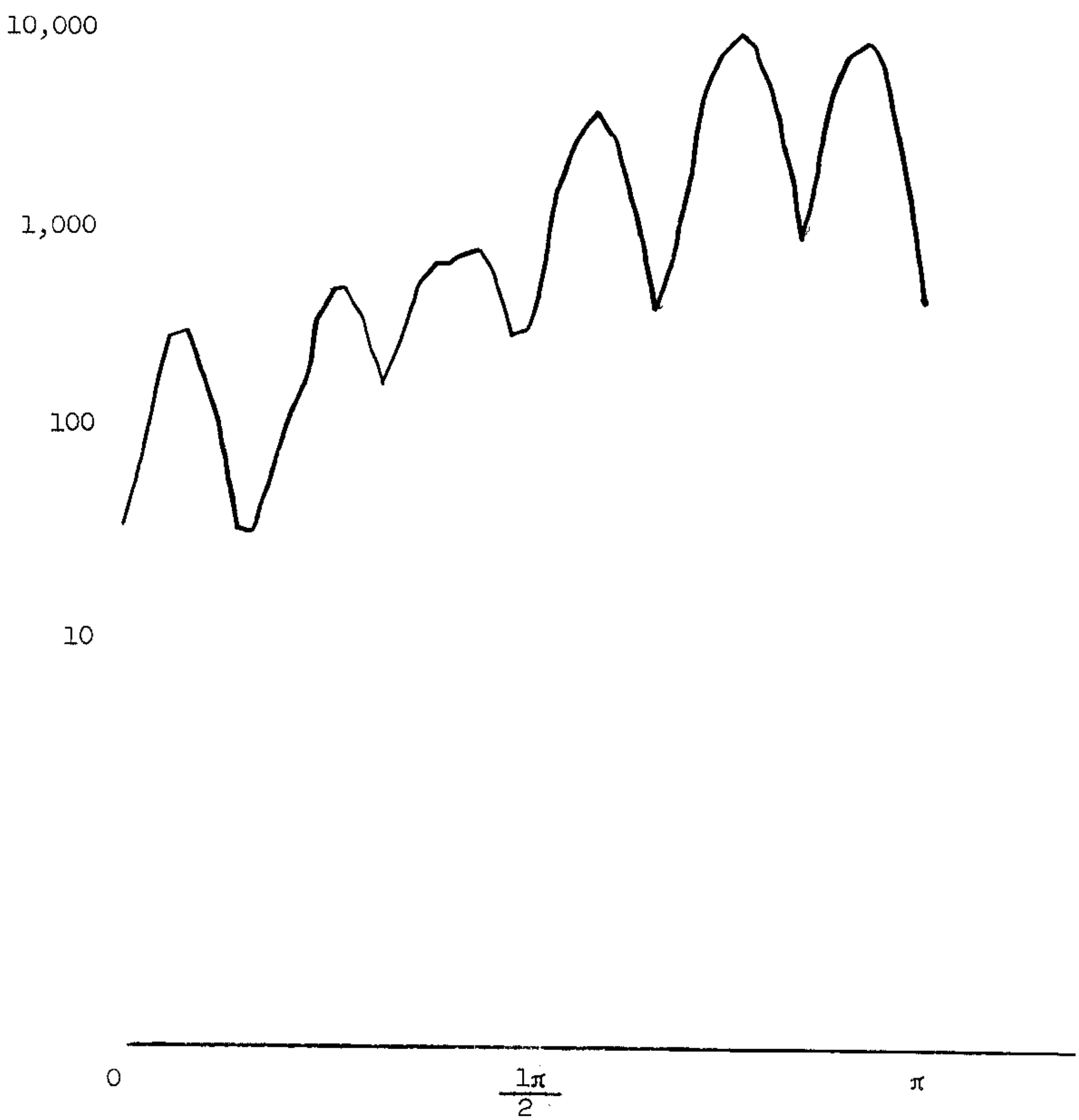


FIGURE 7

Predicted Spectrum of $(1-U^{12})^2(1-\rho U)^2 X_t$

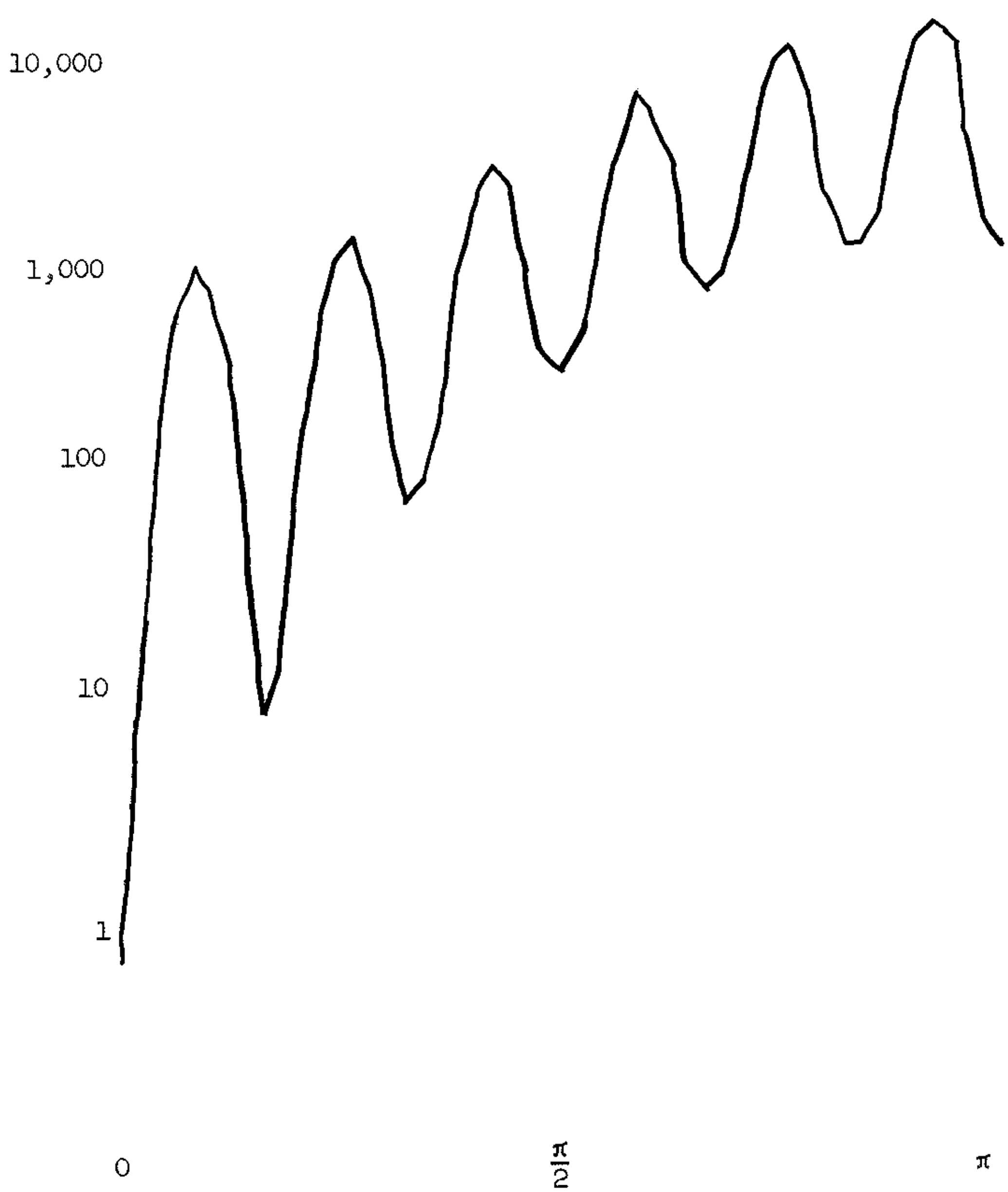


FIGURE 8

Spectrum of $X_t - X_{t-12}$

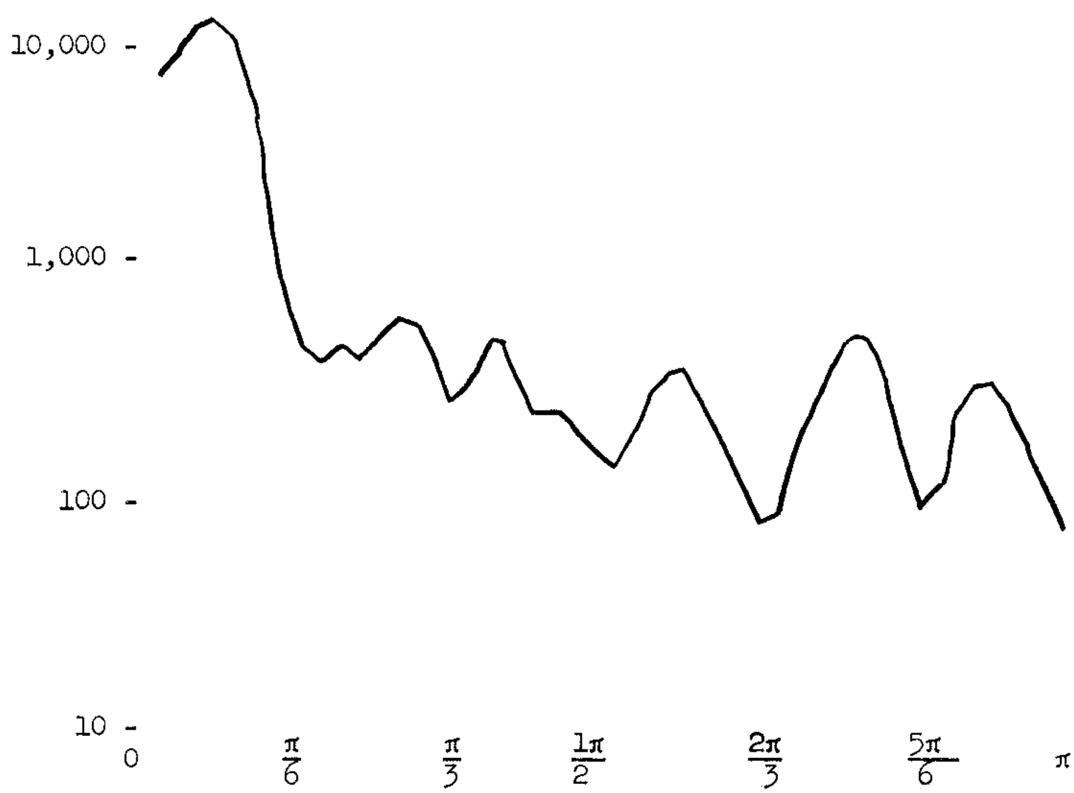
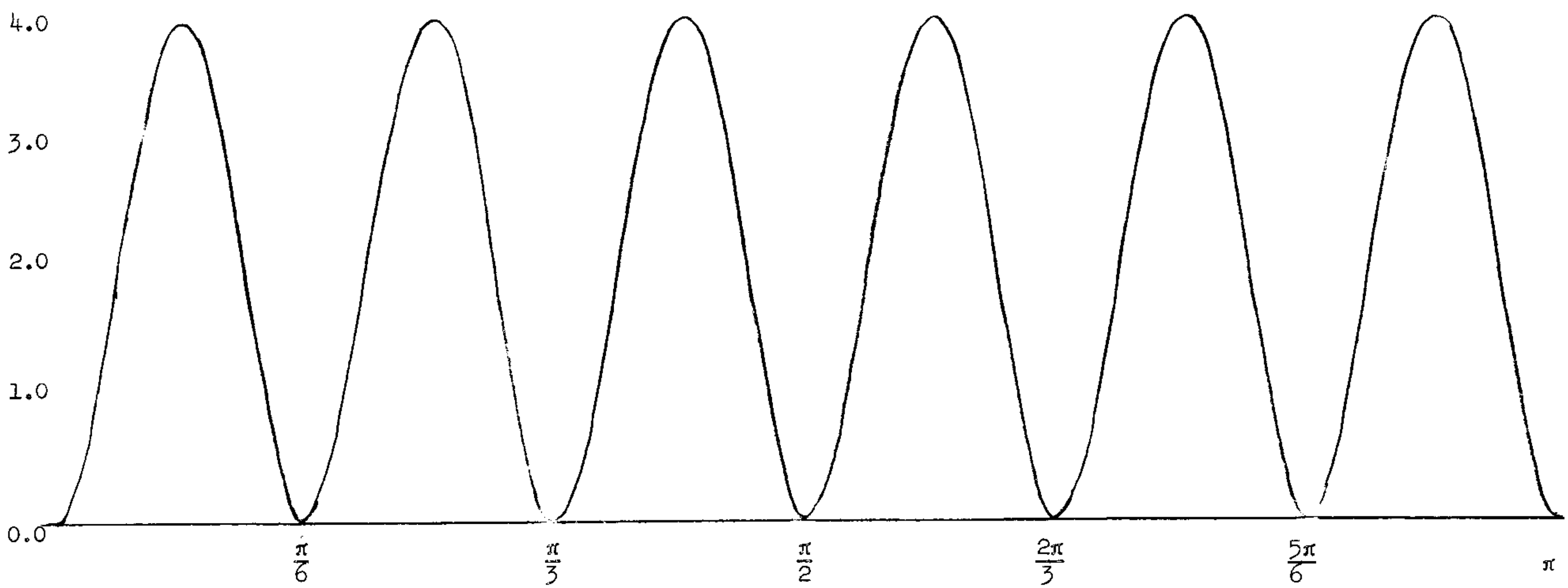


FIGURE 9

Transfer Function for $(1-U^{12})$



APPENDIX 1

We wish to evaluate $\left[\frac{(1-\rho z^{-1})^2}{z^{\nu} P(z^{-1})(1-\theta z^{L})^2}\right]_+$

by means of the following theorem:

Let $Q(z)$ be a function of z analytic in $\delta < |z| < \delta^{-1}$ and let φ be a number such that $|\delta| < 1$, then

$$R(z) = (1 - \varphi z)^P \left[\frac{Q(z)}{(1-\varphi z)^P}\right]_+ = \Pi_p(z) + [Q(z)]_+$$

where

$$\Pi_p(z) = \sum_{i=o}^{p-1} \frac{1}{i!} \left.\frac{d^i[Q(z)]_-}{dz^i}\right|_{z=\varphi^{-1}} (z - \varphi^{-1})^i$$

To apply this theorem to the problem at hand we define

(1) $$Q_o(z) = \frac{(1-pz^{-1})^2}{z^{\nu}P(z^{-1})}$$

Note that for $\nu > 0$ $[Q_o(z)]_+ = 0$

(2) $$(1- \theta z^L)^2 = \theta^{-1} \prod_{i=1}^{L} (1-\lambda_1 z)^2, \; \lambda_i = \theta^{\frac{1}{L}} , \quad i = 1, 2, .. L$$

Now from the theorem stated above:

(3) $$\left[\frac{Q_o(z)}{(1-\lambda_1 z)^2}\right]_+ = \frac{k_o^o + k_o^1 (z-\lambda_1^{-1})}{(1-\lambda_1 z)^2} = \sigma_1(z)$$

where

$$k_o^o = Q_o(\lambda_1^{-1})$$

$$k_o^1 = Q_o^1(\lambda_1^{-1})$$

Now let

$$(4) \qquad Q_1(z) = \frac{Q_o(z)}{(1-\lambda_1 z)^2}$$

then

$$(5) \qquad \left[\frac{Q_1(z)}{(1-\lambda_2 z)^2}\right]_+ = \frac{\sigma_1(z) + k_1^o + k_1^1\,(z-\lambda_2^{-1})}{(1-\lambda_2 z)^2} = \sigma_2(z)$$

where $\sigma_1(z)$ is given by (3) and

$$(6) \qquad k_1^o = [Q_1(\lambda_2^{-1})]_-$$

$$k_1^1 = \left.\frac{d[Q_1(z)]_-}{dz}\right|_{z = \lambda_2^{-1}}$$

To find k_1^o and k_1^1 we note that

$$(7) \qquad Q_1(z) = \frac{Q_o(z)}{(1-\lambda_1 z)^2} = \left[\frac{Q_o(z)}{(1-\lambda_1 z)^2}\right]_+ + \left[\frac{Q_o(z)}{(1-\lambda_1 z)^2}\right]_-$$

$$= \sigma_1(z) + [Q_1(z)]_-$$

so

$$(8) \qquad [Q_1(z)]_- = Q_1(z) - \sigma_1(z)$$

a known function from which we may obtain k_1^o and k_1^1. Clearly we may repeat this process L times until we finally arrive at

$$(9) \qquad \left[\frac{(1-\rho z^{-1})^2}{z^{\nu}P(z^{-1})(1-\theta z^{\nu})^2}\right] = \frac{\sigma(z)}{(1-\theta z^{\nu})^2}$$

If we desire to obtain $\hat{S}_{t+\nu}$ for several $\nu > 0$ we can use

$$(10) \qquad \left[\frac{Q(z)}{z^{\nu}}\right]_{+} = \left[Q(z)\right]_{\nu} \quad z^{-\nu}$$

We could proceed as above and obtain (9) for the case where $\nu = 0$ and use (10). The procedure is exactly the same as outlined above except that in (3) we have

$$(3^*) \qquad \sigma_1(z) = \frac{p_o + k_o^o + k_o^1(z-\lambda_1^{-1})}{(1 - \lambda_1 z)^2}$$

where p_o is the constant term in $P(z)$. From this point on the argument proceeds exactly as before.

APPENDIX 2

In this appendix we give a direct method for expressing

(1) $$\left[\frac{(1-\rho z^{-1})^2}{z^{\nu}\, P(z^{-1})(1-\theta z^{L})^2}\right]_+ = \frac{\sigma(z)}{(1-\theta z^{L})^2}$$

where G(z) is a polynomial in z . We begin by expanding 1/P(z) by partial fractions.

(2) $$\left[\frac{(1-\rho z^{-1})^2}{z^{\nu}\, P(z^{-1})(1-\theta z^{L})^2}\right]_+ = \left[\sum_{i=1}^{2L+2} \frac{T_i(1-\rho z^{-1})^2}{z^{\nu}(1-\theta z^{L})^2(1-\lambda_i z^{-1})}\right]_+$$

$$= \sum_{i=1}^{2L+2} T_i \left[\frac{(1-\rho z^{-1})^2}{z^{\nu}(1-\theta z^{L})^2(1-\lambda_i z^{-1})^2}\right]_+ = \sum_{i=1}^{2L+2} T_i h_i^{\nu}(z)$$

where

$$\sum_{i=1}^{2L+2} \frac{T_i}{(1-\lambda_i z^{-1})} = \frac{1}{P(z^{-1})} \quad \text{and} \quad P(\lambda_i) = 0 , \quad i = 1, 2, \ldots 2L+2$$

$$h_i^{\nu}(z) = \left[\frac{(1-\rho z^{-1})^2}{z^{\nu}(1-\lambda_i z^{-1})(1-\theta z^{L})^2}\right]_+$$

For any function $Q(z) = \sum_{i=-\infty}^{\infty} q_i z^i$ we have

(3) $$\left[\frac{Q(z)}{z^{\nu}}\right]_+ = \left[Q(z)\right]_{\nu} \; z^{-\nu}$$

So to evaluate the expression in (2) it is sufficient to determine the functions $h_i^{0}(z)$ $i = 1, 2, \ldots 2L+2$. We may then use (3) to determine the predictor for any $\nu > 0$.

We now consider evaluating $h_i^o(z)$. In what follows the subscript "i" will be omitted as long as no unnecessary confusion results.

Now

$$(4)\quad h^o(z) = \left[(1-\rho z^{-1})^2 \sum_{i=o}^{\infty} \lambda^i z^{-i} \cdot \sum_{j=o}^{\infty} (j+1)\theta^j z^{L_j} \right]_+$$

$$= \left[\sum_{i=o}^{\infty} \lambda^i z^{-i} \sum_{j=o}^{\infty} (j+1)\, \theta^j z^{L_j} - 2\rho z^{-1} \sum_{i=o}^{\infty} \lambda^i z^{-i} \sum_{j=o}^{\infty} (j+1)\theta^j z^{L_j} \right.$$

$$\left. + \rho^2 z^{-2} \sum_{i=o}^{\infty} \lambda^i z^{-i} \sum_{j=o}^{\infty} (j+1)\theta^i z^{L_j} \right]_+$$

As a first step toward evaluating (4) we obtain an expression for

$$\left[\sum_{o}^{\infty} \lambda^i z^{-i} \sum_{o}^{\infty} (j+1)\theta^j z^{L_j} \right]_+$$

Expanding the second series we have:

$$(5)\quad \left[\sum_{i=o}^{\infty} \lambda^i z^{-i} \sum_{j=o}^{\infty} (j+1)\theta^j z^{L_j} \right] = \left[\sum_{i=o}^{\infty} \lambda^i z^{-i} + 2\theta z^L \sum_{i=o}^{\infty} \lambda^i z^{-i} + \ldots + (n+1)\theta^n z^{nL} \sum_{i=o}^{\infty} \lambda^i z^{-i} + \ldots \right]_+$$

$$= \left[\sum_{i=o}^{\infty} \lambda^i z^{-i} \right]_+ + \ldots + \left[(n+1)\theta^n z^{nL} \sum_{i=o}^{\infty} \lambda^i z^{-i} \right]_+$$

$$(6)\left[\sum_{i=o}^{\infty} \lambda^i z^{-i} \sum_{j=o}^{\infty} (j+1)\theta^j z^{Lj}\right]_+ = 1 + 2\theta z^L \sum_{i=o}^{L} \lambda_i z^{-i} + \ldots + (n+1)\theta^n z^{nL} \sum_{i=o}^{nL} \lambda^i z^{-i} + \ldots$$

$$= 1 + 2\theta z^L \frac{(1-\lambda^{L+1} z^{-L-1})}{1-\lambda z^{-1}} + \ldots + (n+1)\theta^n z^{nL} \frac{(1-\lambda^{nL+1} z^{-nL-1})}{1-\lambda z^{-1}} \ldots$$

$$= \frac{1-\lambda z^{-1} + 2\theta z^L - 2\theta\lambda^{L+1} z^{-1} + \ldots + (n+1)\theta^n z^{nL} - (n+1)\theta^n \lambda^{nL+1} z^{-1} + \ldots}{(1-\lambda z^{-1})}$$

Summing terms of like sign in (6) we obtain:

$$(7) \quad \left[\sum_{i=o}^{\infty} \lambda_i z^{-i} \sum_{i=o}^{\infty} (j+1)\theta^j z^{Lj}\right]_+ = \left(\frac{1}{(1-\theta z^L)^2} - \frac{\lambda z^{-1}}{(1-\theta\lambda^L)^2}\right) \frac{1}{1-\lambda z^{-1}}$$

$$= \frac{(1-\theta z^L)^2 - \lambda z^{-1}(1-\theta z^L)^2}{(1-\lambda z^{-1})(1-\theta z^L)^2(1-\theta\lambda^L)^2}$$

To verify that (7) contains only nonnegative powers of z we multiply out the numerator and collect terms.

$$(8) \quad \left[\sum_{i=o}^{\infty} \lambda_i z^{-i} \sum_{j=o}^{\infty} (j+1)\theta^j z^{Lj}\right] = \frac{(1-\lambda z^{-1}) + 2\theta\lambda z^{L-1} - 2\theta\lambda^L + \theta^2\lambda^{2L} - \theta^2\lambda z^{2L-1}}{(1-\theta\lambda^L)^2 (1-\theta z^L)^2 (1-\lambda z^{-1})}$$

$$= \frac{1-\lambda z^{-1} + 2\theta\lambda z^{L-1}(1-\lambda^{L-1} z^{-L+1}) - \theta^2\lambda z^{2L-1}(1-\lambda^{2L-1} z^{-2L+1})}{(1-\theta z^L)^2 (1-\lambda z^{-1}) (1-\theta\lambda^L)^2}$$

$$= \frac{1 + 2\theta\lambda z^{L-1} \sum_{i=o}^{L-2} \lambda^i z^{-i} - \theta^2 \lambda z^{2L-1} \sum_{i=o}^{2L-2} \lambda^i z^{-i}}{(1-\theta z^L)^2 (1-\theta\lambda^L)^2}$$

$$= \frac{G(z)}{(1-\theta z^L)^2(1-\theta\lambda^L)^2} \qquad \text{where} \qquad G(z) = \sum_{i=o}^{2L-1} g_i z^i$$

$$(9)\qquad \begin{cases} g_o = 1 \\ g_i = 2\theta\lambda^{L-i} - \theta^2\,\lambda^{2L-1}, & i = 1, 2, \ldots L-1 \\ g_i = \qquad\quad - \theta^2\,\lambda^{2L-i}, & i = L, L+1, \ldots 2L-1 \end{cases}$$

From (4) we see that

$$(10)\quad h^o(z) = \frac{G(z)}{(1-\theta\lambda^L)^2(1-\theta z^L)^2} - \frac{2\rho}{(1-\theta\lambda^L)^2}\left[\frac{G(z)}{z(1-\theta z^L)^2}\right]_+ + \frac{\rho^2}{(1-\theta\lambda^L)^2}\left[\frac{G(z)}{(1-\theta\lambda^L)^2 z^2}\right]_+$$

To evaluate the second term on the right hand side of (10) note that

$$(11)\quad \left[\frac{G(z)}{(1-\theta z^L)^2 z}\right] = \left[\frac{g_o}{z(1-\theta z^L)^2} + \frac{g_1}{(1-\theta z^L)^2} + \ldots + \frac{g_{2L-1}\,z^{2L-2}}{(1-\theta z^L)^2}\right]_+$$

$$= z^{-1}\,g_o \sum_{j=1}^{\infty} (j+1)\theta^j z^{Lj} + \left[\frac{[G(z)]z^{-1}}{(1-\theta z^L)^2}\right]_+$$

Now $z^{-1}\sum_{j=1}^{\infty}(j+1)\,\theta^j z^{Lj} = \left(\frac{1}{(1-\theta z^L)^2} - 1\right) z^{-1} = \frac{2\theta z^{L-1} - \theta^2 z^{2L-1}}{(1-\theta z^L)^2}$

so

$$(12)\quad \left[\frac{G(z)}{(1-\theta z^L)^2 z}\right]_+ = \frac{2\theta z^{L-1} - \theta^2\,z^{2L-1} + [G(z)]_1\,z^{-1}}{(1-\theta z^L)^2}$$

Evaluating the last term on the right of (10) in a similar fashion we obtain

$$(13)\quad h^{o}(z) = \frac{1}{(1-\theta\lambda^{L})^2(1-\theta z^{L})^2} \Big\{G(z) - 2\rho\Big[2\theta z^{L-1} - \theta^2 z^{2L-1} + [G(z)]_1\, z^{-1}\Big]$$

$$+ \rho^2 \Big[(2\theta z^{L-1} - \theta^{\theta} z^{2L-1})(z^{-1} + g_1) + [G(z)]_2\, z^{-2}\Big] \Big\}$$

$$= \frac{K(z)}{(1-\lambda^{L}\theta)^2(1-\theta z^{L})^2}$$

$$(14)\qquad k_1 = g_i - 2\rho\, g_{i+1} + \rho^2 g_{i+2} \qquad i = 0, 1 .. L-3, L, L+1 .. 2L-3$$

$$k_{L-2} = g_{L-2} - 2\rho\, g_{L-1} + \rho^2 g_L + \rho^2\, 2\theta$$

$$k_{1-1} = g_{L-1} - 2\rho\, g_L + \rho^2 g_{L+1} - 4\rho\theta + 2\rho^2\theta g_1$$

$$k_{2L-2} = g_{2L-2} - 2\rho\, g_{2L-1} - \rho^2\theta^2$$

$$k_{2L-1} = g_{2L-1} + 2\rho\theta^2 - \rho^2 g_1\theta^2$$

Equations (13) and (14) provide a means of computing each element of the sum in (2). Thus, we may write:

$$(15)\qquad \left[\frac{(1-\rho z^{-1})^2}{P(z^{-1})(1-\theta z^{L})^2}\right]_+ = \frac{J(z)}{(1-\theta z^{L})^2} \qquad \text{where} \qquad J(z) = \sum_{i=o}^{2L-1} j_i\, z^i$$

Now:

$$(16)\qquad \gamma(z) = \frac{(1-\theta z^{L})^2(1-\rho z)^2}{P(z)} \cdot \frac{\sigma_\nu^2}{\sigma^2} \left[\frac{J(z)}{(1-\theta z^{L})^2 z^{\nu}}\right]_+$$

By exactly the same method used in obtaining (13) and (14) we can write:

$$\gamma(z) = \frac{(1-\theta z^{L})^{2}(1-\rho z)^{2}}{P(z)} \quad \frac{\sigma_{\nu}^{2}}{\sigma^{2}} \quad \frac{J_{\nu}(z)}{(1-\theta z^{L})^{2}} \tag{17}$$

$$= \frac{(1-\rho z)^{2}}{P(z)} \quad \frac{\sigma_{\nu}^{2}}{\sigma^{2}} \; J_{\nu}(z)$$

The exact form of $J_{\nu}(z)$ depends on the value of ν. For $\nu < L$

$$J_{\nu}(z) = \left[J(z) \right] z^{-\nu} + (2\theta z^{L-1} - \theta^{2} z^{2L-1})(j_{\nu-1} + z^{-1}\, j_{\nu-2} + \dots + z^{-\nu+1} j_{o}) \tag{18}$$

The numerous substitutions in the above procedure make it somewhat difficult to follow. It should work satisfactorily on an electronic computer however.

REFERENCES

[1] Couts, D., D. Grether, M. Nerlove, Forecasting Non-Stationary Economic Time Series, Technical Report No. 10, Institute for Mathematical Studies in the Social Sciences, May 10, 1965.

[2] Hext, George R., A Note on Pre-Whitening and Recolouring; Technical Report No. 5, Institute for Mathematical Studies in the Social Sciences, September 24, 1965.

[3] Parzen, E., Modern Probability Theory and its Applications, John Wiley & Sons, Inc., 1960.

[4] Whittle, P., Prediction and Regulation by Linear Least-Squares Methods, London: English Universities Press, 1963.