COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 261

SOME PROPERTIES OF "OPTIMAL" SEASONAL ADJUSTMENT

D. M. Grether and M. Nerlove

December 6, 1968

SOME PROPERTIES OF "OPTIMAL" SEASONAL ADJUSTMENT*

by

D.M. Grether and M. Nerlove
Yale University

1. Introduction

In two previous papers, [14] and [15], Nerlove attempted to analyze the effects of various procedures for seasonal adjustment of economic time series on the characteristics of the series to which these procedures were applied. The analysis consisted of a comparison of the estimated spectra of the two series, original and seasonally adjusted, and an examination of the cross-spectrum of the two series, particularly of the coherence and phase shift at various frequencies. In the course of these investigations several informal criteria for judging the adequacy of seasonal adjustment were developed. These criteria were:

First, the coherence of the original and the seasonally adjusted series should be high at all frequencies except, possibly, seasonal ones.

---

*The research on which this paper is based was performed with the support of the National Science Foundation under grants NSF GS-818 and NSF GS-1721 to Yale University, and under a grant from the Ford Foundation.

Second, although phase shifts are generally impossible to avoid altogether in any method of seasonal adjustment which uses past data to adjust current observations, such shifts should be minimized especially at low frequencies at which most of the power in economic time series is typically concentrated.

Finally, seasonal adjustment should remove the peaks in the original series which typically appear at the so-called seasonal frequencies, but should affect the remainder of the spectral density as little as possible; in particular, the process of seasonal adjustment should not remove excessive power at other than seasonal frequencies.

A subsidiary consideration involved the possibility that seasonal adjustment might remove more than enough power at the seasonal frequencies, thus producing "dips" at those frequencies. While this was not regarded as especially serious in and of itself, corresponding to the "dips" there must exist intermediate peaks at frequencies between the seasonal ones. Such peaks, if large enough, might induce spurious fluctuations in the adjusted series -- a disturbing possibility. Nettheim [17], however, showed how the seasonally adjusted series could be corrected for "over-adjustment" at the seasonal frequencies. When such corrections were made in a number of seasonally adjusted series (male unemployment, 20+, total imports, and total civilian labor force), the overall movement and general appearance of the resulting series differed very little from the uncorrected seasonally adjusted series; however, in all three series the locations of turning points were frequently altered, in some cases by as much as two to three months.

It is not possible to assess the economic significance of such effects outside of a particular substantive context, but it is unlikely that most consumers of seasonally adjusted series would fail to be concerned about the possibility that seasonal adjustment methods may affect the location of turning points by as much as two to three months.

In the first of the two studies referred to above, [14], it was found that, for the BLS method of adjusting unemployment, a considerable loss of power occurred at nearly all non-seasonal frequencies higher than those corresponding to a sinusoidal fluctuation of 12-months duration and that this loss was most severe for the age-sex groups unemployed males 14-19 and unemployed females 14-19, for which the seasonal pattern is most pronounced and regular. Little difference in this respect was found between the BLS method then in use and the proposed "residual" method which involved first adjusting employment and labor force for seasonal influences and then deriving the seasonally adjusted unemployment series as the difference between the two. However, in the second study, [15], a method closely related to a proposal made by E.J. Hannan [10] proved to be markedly superior in this respect. The two studies showed that all three methods of seasonal adjustment produced series which had low coherence with its original series at most frequencies above that corresponding to a sinusoidal fluctuation of 12-months duration. Rather violent phase shifts were found at many frequencies including some of the lower frequencies; however, the significance of such shifts in frequency bands where coherence is low is quite limited.[1]

The purpose of this paper is a reassessment of the earlier findings reported above and of recent work along similar lines by Rosenblatt ([21], [22], [23], and [24]). It has never been denied, and, indeed, is repeatedly emphasized by Rosenblatt in his own work, that the effects noted in the frequency domain are significant only to the degree to which these same effects, translated into the behavior of the adjusted series over time, affect the interpretation of the movements of that series. It has been found, however, extraordinarily difficult to obtain general agreement on what constitutes "good" seasonal adjustment in the time domain, and, hence, there has been continued reliance on the examination of the behavior of the adjusted series in the frequency domain.[2]

The chief problem in the interpretation of the effects observed in the frequency domain is the formulation of a definite theory of what constitutes "good" seasonal adjustment. Without a clear and rigorous notion of the purpose of seasonal adjustment, optimal procedures cannot be formulated nor short-comings assessed.

Suppose, for example, that the observed series, X_t is the sum of two components, N_t, a non-seasonal component, and S_t, a seasonal component

$$X_t = N_t + S_t \,. \tag{1.1}$$

We leave aside, for the time being, the precise meaning of these components, except we assume, by definition, that they are uncorrelated

at any lag:

$$\text{cov}\,(N_t,\ S_{t-k}) = 0\ ,\quad k = 0,\ \pm 1,\ \pm 2,\ \ldots \tag{1.2}$$

In this case, the spectral density of the observed series is simply the sum of the spectral densities of each of the two components:

$$f_{XX}(\lambda) = f_{NN}(\lambda) + f_{SS}(\lambda) \tag{1.3}$$

where $-\pi \leq \lambda \leq \pi$. Since $f_{SS}(\lambda) \geq 0$, the spectral density of N_t is less than the spectral density of the observed series, X_t . Unless S_t has a line spectrum, e.g.,

$$S_t = \Sigma c_j e^{i\lambda_j}\ , \tag{1.4}$$

where $\lambda_j = 2\pi j/12$, $j = 1,\ \ldots,\ 12$, for monthly observations, so that it is positive only at the so-called seasonal frequencies,[3] the non-seasonal component will have a spectral density lying below that of the observed series at frequencies other than these. For example, if

$$S_t = \alpha S_{t-12} + \epsilon_t\ , \tag{1.5}$$

where ϵ_t is a serially uncorrelated random variable with variance σ^2 and mean zero independent of t , again assuming monthly observations, $f_{SS}(\lambda)$ will be non-zero at every frequency. Thus, a comparison of the spectral densities of N_t and X_t would show "loss of

power at non-seasonal frequencies" as well as seasonal ones. Conceivably, our objective in seasonal adjustment might be to isolate N_t, in which case we should not object to some loss of power at non-seasonal frequencies, particularly if the seasonal peaks were not expected to be exceptionally sharp.

Continuing the above example, note that the coherence between N_t and X_t is

$$0 \leq R_{XN}(\lambda) = \frac{1}{1 + \frac{f_{SS}(\lambda)}{f_{NN}(\lambda)}} \leq 1 \tag{1.6}$$

with the upper inequality strict when $f_{SS}(\lambda) \neq 0$. If $f_{SS}(\lambda)$ is large at any frequency, the coherence between N_t and X_t will be low at that frequency. Some plausible models for the seasonal component and the non-seasonal component lead to contribution by the seasonal component of much of the power at high frequencies, seasonal or otherwise. Hence, low estimated coherence of seasonally adjusted and seasonally unadjusted series at other than seasonal frequencies requires careful interpretation. Furthermore, any method of seasonal adjustment which consists of linear combinations of past and, possibly, future values of the series in question, with constant weights, will produce a series always having coherence with the unadjusted series equal to one. By itself, then, the coherence can tell us little about the adequacy of a seasonal adjustment procedure.

It is thus clear that spectral criteria of the sort used in [14] are inadequate alone for the proper assessment of methods of seasonal adjustment. Below, we outline several closely related "theories" of seasonal adjustment and, in terms of the desiderata suggested by those theories, we derive "optimal" seasonal adjustment procedures. The methods so derived are applied to a simulated time series having the observed spectral characteristics of many real economic time series, and it is shown that these "optimal" seasonal adjustment procedures produce time series which bear a very similar relation to the original time series in the frequency domain as that which has been found in previous studies. The investigation thus casts considerable doubt on the adequacy of the spectral criteria adopted for the assessment of seasonal adjustment procedures. It should be emphasized, however, that the results of our study do not call into question the utility of spectral techniques for the description and comparison of different methods of adjustment. What is clear is that without an adequate theory of what seasonal adjustment is supposed to do, no proper criteria for the assessment of seasonal adjustment procedures can be formulated either in the time domain or in the frequency domain.

2. Desiderata of Seasonal Adjustment

When most consumers of seasonally adjusted series -- and that includes nearly every economically literate person -- are confronted by the question of why they prefer such a series to the original, the most common and natural reaction is that the answer is obvious. Yet

on further reflection the basis for such a preference becomes less clear, and those who give the matter extensive thought often finish by becoming hopelessly confused. While we must confess that we belong to the hopelessly confused category, the following remarks may serve to provide a framework, albeit perhaps not the proper framework, for analysis.

Fundamental to existing discussions of seasonal adjustment is the idea that an economic time series may be divided into components, individually unobserved but distinct. The notion that a series of observations might be divided into several unobserved components which summed or multiplied together, or otherwise combined, to give the observed value of the series appears to have its origin in the work of the astronomers of the eighteenth and early nineteenth centuries.[4] When meteorological studies became important in the early nineteenth century, the notion of unobserved components was carried over from astronomy and applied to the analysis of temperature variations and variations in barometric pressure by men such as Forbes [5] and Buys Ballot [3]. The basic idea in many of these studies was that the observed records were the superposition of several unobserved components, each of which depended upon a distinct set of causal factors. Buys Ballot, for example, averaged temperature data using different periods (e.g. monthly, weekly, etc.) in order to isolate individual periodic components which he assumed depended on various astronomical factors such as the position of the moon.

Similar techniques were applied in the analysis of economic phenomena beginning in the middle of the nineteenth century. Charles

Babbage, analyzed daily clearing house records and determined the normal weekly and monthly patterns as well and calculated the consequences of disturbances by taking the difference between the observed and normal levels of clearings [1]. J.W. Gilbart, a mid-nineteenth century banker, used the different seasonal patterns in the circulation of bank notes to argue against tying the issue of notes of all banks to that of the Bank of England [2]. He also assessed the impact of the Bank Reform Act of 1844 by the change in the seasonal pattern and in the monthly pattern of note circulation [6], [7]. W.S. Jevons applied similar techniques in his studies of financial variables and made policy recommendations on the basis of the regularities he found [11]. The idea that an economic time series may be divided meaningfully into several unobserved components appears to have been firmly established in economics since the time of Jevons.[5]

It is, of course, quite debatable whether the idea of unobserved components, appropriate as it may be in the analysis of astronomical observations, is usefully applied to economic data or even to meteorological data. Nonetheless, we believe that this idea lies behind both present methods of seasonal adjustment and the desire for seasonally adjusted time series. In itself, the division of a time series into several unobserved components is of little significance, it is, rather, that the components are themselves ascribable to separate and distinct groups of causes or influences. These in turn may or may not be directly observable in principle or in practice and may consist of other unobserved components in other observed time series. Again, note the astronomical overtones that attribution of the unobserved components to separate

groups of causal factors has. The great problem in the strict division into separate causal groups in this way is the difficulty in the interpretation of the components. Seasonal fluctuations as we understand them, for example, do depend upon meteorological phenomena, but they also depend upon customs and institutions. "Unseasonal" weather, if we may be permitted such a phrase, will, under current practice, be allocated to another component. Thus, separation into causual groups of factors affecting a given series is not in itself sufficient for division into unobserved components. Social phenomena are a great deal more complex than are astronomical phenomena; the unobserved components into which it seems natural to divide an economic time series may represent very heterogeneous and complex groups of causal factors.

Consider, for example, the analysis of the behavior of the monthly unemployment series in the United States: one might attribute some of the month-to-month variation to sampling errors, some to the general level of economic activity and to demand conditions for various types of labor in relation to supply, some to weather variations and such more-or-less regular events as holidays, school opening or closing dates, and the like. Some of these causal factors have different import for potential future changes, some may be affected one way by certain economic policies, others another, and some not at all. If this point of view is adopted, it is natural to ask: Why should we not attempt to relate the observed series directly to all of the causal factors which determine its values over time? There are at least two reasons why such an approach may not be feasible.

First, not all of the causal factors may be directly measurable. For example, some of them may themselves be unobserved components, or on others there may be only incomplete, fragmentary, or no data available. Second, the relationship between the causal factors may be so complex and there may be so many different factors involved, that the task of sorting everthing out with a relatively limited amount of data is quite hopeless.

While it is surely not the case that it is "all or nothing" as far as isolating distinct groups of causal factors affecting a given economic time series, it is not unnatural to begin by a careful examination of the series itself in the hope of isolating regularities which might be ascribed to the different causal groups. Such an approach is characteristic of the pioneers in the use of unobserved components in economics. However, one might also examine the effects of weather and other variables more directly. Unfortunately, it does not seem possible, for the United States as a whole at least, to get much beyond the effect of the number of trading days in a month. The reason is clear: "Weather," for a country the size of the U.S., is not an unambiguous concept. The task of summarizing in a few variables this complex and geographically diverse phenomenon, which then could be related to economic variables such as retail sales, is not possible at the present time. Thus, for the most part, economic statisticians have tended to treat each series in isolation and have attempted to "decompose" each observation into unobserved components each of which can be attributed to the effects of different causal factors.

In this paper we accept the approach which concentrates on the analysis of the behavior of a single time series. At the same time, we must emphatically state that such an extreme position can only be a temporary expedient, a prelude to further work designed to suggest meaningful hypotheses concerning the origins of different components and the relations among various components in several time series. Despite the great simplification achieved by restricting the analysis to individual time series, we are not yet in a position to state formal criteria for optimal seasonal adjustment. Even though one may accept the notion that the observed value of an economic time series may be divided into several unobserved components, for example, the traditional trend-cycle, seasonal, and irregular, it is still necessary to specify in more detail the purpose of the analysis before formal optimality criteria may be developed. Is it clear that seasonal adjustment should be designed to obtain the best estimate of the so-called trend cycle component? Or should the goal be an estimate of the seasonal compent itself so that this may be removed leaving everything else? Alternatively, perhaps what is really wanted is a prediction, not just of next month perhaps, but rather of a full year. All of these <u>desiderata</u> are implicit in what has been said or written on the question of what constitutes good seasonal adjustment.[6] Furthermore, they are not entirely distinct, for good forecasts of an annual total early in the year may be nearly equivalent to good estimation of the trend-cycle component. The statement, for example, that unemployment is higher than usual for this time of year os often interpreted to mean that unemployment <u>will</u> be higher for the year.

The purposes of seasonal adjustment are both analytic and predictive. If we believe that an economic time series results from the superposition of several "sub-series" which are not directly observable, we may wish to isolate such sub-series in order to study their movements in relation to observable variables or other such sub-series. This in turn may permit more accurate prediction of future values of the observed series. Alternatively we may be interested in a whole set of predictions, e.g., a manufacturer planning production not merely on the basis of a forecast of next month's sales but using forecasts for each of many future periods, and the seasonally adjusted series may be an attempt to summarize a rather complicated vector result in the form of a single number. In what follows, we do not attempt to choose between these various points of view, but consider seasonal adjustment both as an "extraction", or estimation of one or more unobserved components, and as a form of prediction. In each case, we do, however, adopt a formal criterion of optimality.

Suppose that the observed value of a time series $\{x_t\}$ may be represented as the sum of, say, three unobserved components: $\{y_t\}$, trend-cycle; $\{s_t\}$, seasonal; and $\{u_t\}$, irregular. Thus

$$x_t = y_t + s_t + u_t \,. \tag{2.1}$$

Further, let us suppose that the stochastic structure of each of these unobserved components can be specified in such a way that the stochastic structure of the observed time series $\{x_t\}$ is thereby determined.

Following the above discussion the problem of seasonal adjustment may be specified in the following three ways as the problem of obtaining, at a given moment of time t :

A. An estimate or extraction of y_t based on the observed x_t either up to that time or beyond it as well;

B. An estimate or extraction of s_t, and subtraction of this from x_t, based on observations either up to that time or beyond it as well; or

C. A series of predictions of $x_t, x_{t+1}, \ldots x_{t+v}$, up to v periods ahead (e.g., 12 if we are dealing with monthly data), based on past observations alone, which may or may not be summarized in the form of an average or mean value, weighted or unweighted.

All of these problems are closely related and may be solved by essentially the same method. However, in order to apply the method a number of further simplifications are necessary with regard to the criterion of optimality we shall adopt and to the nature of our specification of the stochastic structure of the unobserved components.

In some fundamental sense, optimal estimation of something should depend upon what it costs to make an error. Such cost or loss functions, can only be specified within the context of a specific decision problem, that is, we have to know precisely how and by whom an estimate or prediction is going to be used, before we can determine an appropriate criterion of optimality and derive an estimate or prediction

which satisfies that criterion. Such a specification is obviously impossible in the present context, for, whatever seasonal adjustment may or may not be, it is surely designed to serve a great variety of users in many different situations. The classic solution to this problem is to minimize the expected value of the squared error between whatever it is we are trying to estimate or predict and the true value. Thus, the criterion of optimality we adopt is that of minimum mean-square error.

As regards the stochastic structure of the unobserved components, $\{y_t\}$, $\{s_t\}$, and $\{u_t\}$, a number of possibilities exist. Because we wish to impose a decomposition into several unobserved components, a relatively assumption-free or nonparametric approach is ruled out. However, rather general parametric schemes can be considered relatively easily. To facilitate our discussion, we introduce the backward shift operator U defined by

$$U^k x_t = x_{t-k} . \tag{2.2}$$

Using this notation we may write the two-sided moving average

$$\sum_{-p}^{q} a_j x_{t-j} = A(U)x_t , \tag{2.3}$$

where the $A(U)$ is a polynomial in the lag operator U,

$$A(U) = \sum_{-p}^{q} a_j U^j . \tag{2.4}$$

It has been found possible to reproduce the spectral shapes of many economic time series by the following sort of model:

The observed series $\{x_t\}$ may be decomposed into three components as in (2.1) above and these three components satisfy the following relations

$$(2.5) \qquad \begin{cases} Q(U)y_t = P(U)v_t \\ \\ S(U)s_t = R(U)w_t \ , \end{cases}$$

where $Q(U)$, $P(U)$, $S(U)$, and $R(U)$ are polynomials in the lag operator U , of relatively low order except for $S(U)$, which, however, is of low order in U^L , where L is the number of times per year the series $\{x_t\}$ is observed.[7] The series $\{u_t\}$, $\{v_t\}$ and $\{w_t\}$ are assumed to be stationary with the following properties:

$$(2.6) \qquad \begin{cases} Eu_t u_{t'} = \begin{cases} \sigma_u^2 \ , & t = t' \ , \\ 0 \ , & t \neq t' \ , \end{cases} \\ \\ Ev_t v_{t'} = \begin{cases} \sigma_v^2 \ , & t = t' \ , \\ 0 \ , & t \neq t' \ , \end{cases} \\ \\ Ew_t w_{t'} = \begin{cases} \sigma_w^2 \ , & t = t' \ , \\ 0 \ , & t \neq t' \end{cases} \\ \\ Eu_t v_{t'} = Eu_t w_{t'} = Ev_t w_{t'} = 0 \ , \quad \text{all} \quad t \quad \text{and} \quad t' \ . \end{cases}$$

How complicated one may wish to make the polynomials Q, P, S, and R depends in part on how much data is available and how closely one wishes to approximate the observed characteristics of the series $\{x_t\}$. For most series of economic relevance, rather simple, i.e., low-order, polynomials are desirable.

The question which now must be considered is to what extent the parameters characterizing the polynomials and the series $\{u_t\}$, $\{v_t\}$, and $\{w_t\}$ should be taken as given in the problem of seasonal adjustment as specified in one of the three ways outlined above. If none of the parameters are taken as given, including the degree of the polynomials Q, P, S, and R, the problem of seasonal adjustment becomes an extraordinarily difficult one. On the other hand, complete numerical specification a priori is clearly a most unreasonable simplification. Possibly, a more realistic approach lying somewhere between these extremes would be to assume that the orders of the polynomials, and perhaps something of their internal structure could be specified in advance, but that numerical values of the remaining parameters would be determined in the course of statistical analysis. Unfortunately, even this simplification leaves the problem of seasonal adjustment too complex for us to handle at the present time. It seems useful, therefore, to separate the problem of estimation of the structure of the time series to be adjusted from the problem of seasonal adjustment proper. At the same time, we must recognize that, in principle, the two problems should be solved simultaneously and hope to

return to this question in subsequent research.

In the simulation studies reported below, we assume that the stochastic structure of the series $\{x_t\}$ is known a priori. This means, in effect, that our "optimal" seasonal adjustment procedures are more optimal than realizable procedures could ever be. That they, in fact, possess many of the same spectral properties as the Census and Bureau of Labor Statistics Methods do, is rather strong evidence, therefore, in support of our conclusion that the spectral criteria developed earlier do not in fact provide an adequate basis for the evaluation of such methods.

3. Minimum Mean-Square-Error Extraction and Prediction: An Outline of the Theory[8]

Our aim in this paper as suggested in the previous section, is to obtain certain predictions or extractions for a time series of given structure which are minimum mean-square error. This section gives a brief outline of the theory of such predictions or extractions for stationary time series whose entire past is known up to a certain point in time. In the next section, this theory is applied to the simple model used in our simulation study.

It is convenient for present purposes to deal with processes which are stationary, at least to second order,[9] and defined at discrete points in time. Let such a process be denoted by $\{y_t\}$ where the index t ranges over the positive and negative integers and 0.

In 1938, Wold [27] showed that every discrete stationary process $\{y_t\}$ could be decomposed into the sum of two mutually independent processes $\{\xi_t\}$ and $\{\eta_t\}$ such that

$$y_t = \xi_t + \eta_t \quad , \tag{3.1}$$

where $\{\xi_t\}$ is the so-called linearly deterministic process which may be predicted with zero mean-square error from all past observations and where η_t is a stationary, possibly infinite, moving average process:

$$\eta_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j} \quad , \qquad b_0 = 1 \quad , \tag{3.2}$$

where

$$\Sigma |b_j|^2 < \infty \quad ,$$

and

$$E\epsilon_t \epsilon_{t'} = \begin{cases} \sigma^2 \quad , & t = t' \\ 0 \quad , & \text{otherwise} \end{cases} \quad .$$

The process $\{\epsilon_t\}$ is often called the "white noise input of the moving average." We have found that most economic time series, possibly after some transformation to render them stationary, can be represented in the form (3.2). Such processes are said to be purely non-deterministic and can be shown to possess absolutely continuous spectral distribution functions. In terms of the backward shift operator U introduced in the previous section, η_t may be written

$$\eta_t = B(U)\epsilon_t \,, \tag{3.3}$$

where

$$B(U) = \sum_{j=0}^{\infty} b_j U^j$$

is an infinite series rather than a polynomial of finite degree in the lag operator U.

In deriving the minimum mean-square-error predictors and extractors as linear functions of past, and possibly some future, values of the observed time series, it is convenient to express our results, not in terms of the coefficients attaching to various lagged values, but rather in terms of the generating transform of these coefficients. The generating transform or z-transform of a sequence $\{ \ldots a_{-1}, a_0, a_1, a_2, \ldots \}$ is defined as

$$A(z) = \sum_{k=-\infty}^{\infty} a_k z^k \tag{3.4}$$

when the summation on the right converges. (When it does in some region it represents the Laurent expansion of the function $A(z)$ there.) Note that z is complex. Clearly,

$$B(z) = \sum_{j=0}^{\infty} b_j z^j$$

converges in a closed region bounded by the unit circle. We further assume, largely to simplify the ensuing discussion, that it converges

inside a circle slightly larger than the unit circle. Then

$$B(z^{-1}) = \sum_{j=0}^{\infty} b_j z^j$$

converges outside a circle which is inside the unit circle, so that the function $B(z)B(z^{-1})$ is defined and analytic in an annulus about the unit circle. As we shall see this function evaluated on the unit circle is proportional to the spectral density of the time series x_t.

The autocovariance function of a stationary time series is defined as

$$c(k) = Ex_t x_{t-k} \tag{3.5}$$

and is a function only of the lag k. For processes which are stationary and contain no purely linearly deterministic component, the autocovariance generating transform exists and is given by

$$g_{xx}(z) = \sum_{k=-\infty}^{\infty} c(k)z^k = \sigma^2 B(z)B(z^{-1}) \tag{3.6}$$

as can be readily deduced from (3.2) and (3.5), replacing η_t by x_t. On the unit circle, i.e., for $z = e^{-i\lambda}$, $-\pi \leq \lambda \leq \pi$, we have

$$g_{yy}(e^{-i\lambda}) = 2\pi f_{xx}(\lambda) = \sum_{k=-\infty}^{\infty} c(k)e^{-ik\lambda} \tag{3.7}$$

$$= c(0) + 2\sum_{k=1}^{\infty} c(k) \cos k\lambda ,$$

so we see that on the unit circle the autocovariance generating transform is proportional to the spectral density function. Furthermore,

$$f_{xx}(\lambda) = \frac{\sigma^2}{2\pi} \left|B(e^{i\lambda})\right|^2 , \tag{3.8}$$

so that, because $B(z)B(z^{-1})$ is analytic in an annulus containing the unit circle, we see that spectral distribution functions for processes of the type considered are absolutely continuous functions of λ. Equation (3.8) shows why the representation of $g_{xx}(z)$ as $\sigma^2 B(z)B(z^{-1})$ is often called the canonical factorization of the spectral density function. This factorization must evidentally exist for all processes of the type considered but it may not be unique unless one sets conditions on the zeros of $g_{xx}(z)$.

Although all stationary time series with no linearly deterministic component have a one-sided moving average representation, not all have an autoregressive representation. A process defined by a sequence $\{x_t\}$ satisfying

$$A(U)x_t = \epsilon_t \tag{3.9}$$

where $A(U)$ is a polynomial in U and ϵ_t is a white noise input is called an autoregressive process and may or may not be stationary. In stationary cases it is not necessary to restrict the degree of $A(U)$ to be finite. When $g_{xx}(z)$ is the autocovanciance generating transform of a stationary process which has no zeros on the unit circle

and is analytic both on the unit circle and in an annulus about the unit circle, then the process $\{x_t\}$ has both a moving average and an autoregressive representation. In this case the generating transform of the weights in the autoregressive representation and those in the moving average representation are related by

$$A(z) = \frac{1}{B(z)} . \tag{3.10}$$

Clearly a necessary condition that $A(z)$ exist for a stationary process, whose spectral density in canonical form is $\sigma^2 B(z)B(z^{-1})$, is that $B(z)$ shall have no zeros on the unit circle. Indeed to make the factorization unique we observe that $g_{yy}(z)$, being symmetric, has a zero outside the unit circle corresponding to every one inside the unit circle so that we can separate these zeros by appropriate choice of the factors $B(z)$ and $B(z^{-1})$. If this is done so that $B(z)$ has zeros only outside the unit circle the factorization will be unique and $A(z)$ will be given as in (3.10).

Although perfectly acceptable stationary processes such as $y_t = \epsilon_t - \epsilon_{t-1}$, do not possess autoregressive representations, we generally suppose throughout the remainder of this paper that the processes with which we deal have both moving average and autoregressive representations.

Of substantial practical importance is the case of a time series with a rational spectral density function. In this case, by

definition, the autocovariance generating function may be written as the ratio of two polynomials:

$$g_{xx}(z) = \frac{P(z)}{Q(z)} . \tag{3.11}$$

If $Q(z)$ has roots on the unit circle, then $\{x_t\}$ cannot be regarded as stationary for it has no moving average representation. On the other hand, if $Q(z)$ has no such roots, we know that because $g_{xx}(z)$ is symmetric in z and z^{-1}, both $P(z)$ and $Q(z)$ must be as well, and, hence, can be factored as

$$g_{xx}(z) = \frac{\sigma^2 \prod_{k=1}^{m} (1 - \beta_k z)(1 - \beta_k z^{-1})}{\prod_{k=1}^{m} (1 - \alpha_k z)(1 - \alpha_k z^{-1})} . \tag{3.12}$$

Note σ^2 has been chosen so that the leading coefficients of $P(z)$ and $Q(z)$ are both one.[11] In line with the convention mentioned earlier to ensure a unique factorization, we suppose $|\beta_k| \leq 1$ and $|\alpha_k| < 1$. If the strict inequality holds in the first instance the process has an autoregressive as well as a moving average prepresentation; the latter has generating transform

$$B(z) = \frac{\prod_{k=1}^{m} (1 - \beta_k z)}{\prod_{k=1}^{n} (1 - \alpha_k z)} \tag{3.13}$$

and the former $A(z) = [B(z)]^{-1}$. Processes with rational spectral density are typically represented as an autoregression equal to a noise input which is not white, i.e.,

$$(3.14) \qquad \prod_{k=1}^{n} (1 - \alpha_k U)x_t = \prod_{k=1}^{m} (1 - \beta_k U)\epsilon_t$$

We are now in a position to show that the process described in the previous section by equations (2.1), (2.5) and (2.6) has a rational spectral density function of the form (3.12). Let

$$(3.15) \qquad \begin{cases} \lambda = \sigma_v^2/\sigma_u^2 \\ \mu = \sigma_w^2/\sigma_u^2 \end{cases}$$

be the ratios of the variances of the noise inputs to the trend-cycle and the seasonal, respectively, to the variance of the irregular component. Because $\{u_t\}$, $\{v_t\}$, and $\{w_t\}$ are mutually uncorrelated sequences, we may write the spectral density or, better, autocovariance generating transform, of the observed series $\{x_t\}$ as the sum of the spectral densities or autocovariance generating transforms of its unobserved components. Thus,

$$(3.16) \qquad g_{xx}(z) = g_{yy}(z) + g_{ss}(z) + g_{uu}(z)$$

$$= \frac{P(z)P(z^{-1})}{Q(z)Q(z^{-1})}\sigma_v^2 + \frac{R(z)R(z^{-1})}{S(z)S(z^{-1})}\sigma_w^2 + \sigma_u^2$$

by (2.5) above. Factor out σ_u^2 and collecting terms, we obtain

(3.17) $g_{xx}(z) =$

$$\sigma_u^2 \frac{Q(z)Q(z^{-1})S(z)S(z^{-1}) + \lambda P(z)P(z^{-1})S(z)S(z^{-1}) + \mu R(z)R(z^{-1})Q(z)Q(z^{-1})}{Q(z)Q(z^{-1})S(z)S(z^{-1})}$$

$$= \sigma_u^2 \frac{\prod_{k=1}^{m}(1 - \beta_k z)(1 - \beta_k z^{-1})}{\prod_{k=1}^{n}(1 - \alpha_k z)(1 - \alpha_k z^{-1})}$$

where β_k and $1/\beta_k$, $k = 1, \ldots, m$ are the $2m$ roots of the polynomial

$$Q(z)Q(z^{-1})S(z)S(z^{-1}) + \lambda P(z)P(z^{-1}) + R(z)R(z^{-1})Q(z)Q(z^{-1}),$$

assumed to be of degree $2m$, and α_k and $1/\alpha_k$, $k = 1, \ldots, n$, are the $2n$ roots of

$$Q(z)Q(z^{-1})S(z)S(z^{-1})$$

assumed to be of degree $2n$. Note, however, that σ^2 in (3.12) is the variance of $\{\epsilon_t\}$, the white noise input to the moving average representation of the time series, whereas, σ_u^2 is not this at all, but only the variance of the irregular component. The method of obtaining σ^2 for processes with covariance generating transforms of the form given in (3.17) is explained in detail in Nerlove [16, Appendix, sec. B] and need not be repeated here.

Consider now two jointly stationary nondeterministic processes $\{y_t\}$ and $\{x_t\}$. The k^{th} lag covariance of y_t and x_t, in that order, is given by

$$c_{yx}(k) = Ey_t x_{t-k} \qquad k = 0, \pm 1, \pm 2, \ldots . \tag{3.18}$$

Note this is different from $Ey_{t-k}x_t = c_{xy}(k) = c_{yx}(-k)$. The generating transform of $c_{yx}(k)$ is

$$g_{yx}(z) = \sum_{k=-\infty}^{\infty} c_{yx}(z)z^k \tag{3.19}$$

and may be termed the cross-covariance generating function inasmuch as its value on the unit circle is proportional to the cross-spectral density function of the series $\{y_t\}$ and $\{x_t\}$.

Consider the problem of estimating y_t for fixed t given the entire past of the series $\{x_t\}$ up to and including that time. Let us consider only predictors which can be expressed as linear combinations of past x's,

$$\hat{y}_t = \sum_{j=0}^{\infty} \gamma_j x_{t-j} = \gamma(U)x_t , \tag{3.20}$$

and consider optimal that choice of γ_j for which $E(\hat{y}_t - y_t)$ is minimized.

Because $\{x_t\}$ is stationary we can write

$$x_t = B(U)\epsilon_t \,, \tag{3.21}$$

where $\{\epsilon_t\}$ is a white noise sequence with variance σ^2 and $b_o = 1$. If $\varphi(z)$ represents the generating transform

$$\varphi(z) = \gamma(z)B(z) \,, \tag{3.22}$$

we can express $\hat{y}_t$ in terms of the past of $\{\epsilon_t\}$ to t :

$$\hat{y}_t = \varphi(U)\epsilon_t = \sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j} \tag{3.23}$$

It is, in fact, more convenient to find $\varphi(z)$ or $\hat{y}_t$ in the form (3.23), then determine $\gamma(z)$ from (3.22), provided $[B(z)]^{-1}$ exists, and so express $\hat{y}_t$ in the form (3.20).

Under the minimum mean-square-error criterion, we seek to minimize

$$\begin{aligned} E(\hat{y}_t - y_t)^2 &= E[\sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j} - y_t]^2 \\ &= \mathrm{var}(y) + \mathrm{var}(\sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j}) - 2\mathrm{cov}(\sum_{j=0}^{\infty} \varphi_j \epsilon_{t+j},\ y_t) \\ &= \mathrm{var}(y) + \sigma^2 \sum_{j=0}^{\infty} \varphi_j^2 - 2\sum_{j=0}^{\infty} c_j \varphi_j \end{aligned} \tag{3.24}$$

where $c_j = E(y_t - Ey_t)\epsilon_{t-j}$. Completing the square, we obtain

$$(3.25)\quad E(\hat{y}_t - y_t)^2 = \mathrm{var}(y) + \sigma^2 \sum_{j=0}^{\infty} [\varphi_j - \left(\frac{c_j}{\sigma^2}\right)]^2 - \frac{1}{\sigma^2} \sum_{j=0}^{\infty} c_j^2$$

$$\geq \mathrm{var}(y) - \frac{1}{\sigma^2} \sum_{j=0}^{\infty} c_j^2 ,$$

with equality only for $\varphi_j = c_j/\sigma^2$. Whence

$$(3.26)\qquad \min E(\hat{y}_t - y_t)^2 = \mathrm{var}(y) - \sigma^2 \sum_{j=0}^{\infty} \varphi_j^2 .$$

The following notation will be used extensively throughout the remainder of this paper: If $\{\ldots h_{-1}, h_0, h_1, \ldots\}$ is a sequence with generating transform

$$H(z) = \sum_{j=-\infty}^{\infty} h_j z^j ,$$

we denote by $[H(z)]_+$ that part of $H(z)$ having only nonnegative powers of z , i.e.,

$$[H(z)]_+ = \sum_{j=0}^{\infty} h_j z^j ,$$

and that part having only negative powers by

$$[H(z)]_- = \sum_{j=-\infty}^{-1} h_j z^j .$$

Using this notation we see that (3.25) implies

$$(3.27)\qquad \varphi(z) = \frac{1}{\sigma^2} \sum_{j=0}^{\infty} c_j z^j = \frac{1}{\sigma^2} [g_{y\epsilon}(z)]_+ .$$

We may assume $\{y_t\}$ has zero mean without loss of generality. Then

$$g_{yx}(z) = \sum_{k=-\infty}^{\infty} (Ey_t x_{t-k}) z^k \tag{3.28}$$

$$= \sum_{k=-\infty}^{\infty} z^k Ey_t \sum_{j=0}^{\infty} b_j \epsilon_{t-j-k}$$

$$= \sum_{k=-\infty}^{\infty} z^k \sum_{j=0}^{\infty} b_j Ey_t \epsilon_{t-j-k}$$

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

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

Hence,

$$\varphi(z) = \frac{1}{\sigma^2} \left[\frac{g_{yx}(z)}{B(z^{-1})} \right]_+ \tag{3.29}$$

or, using (3.22) and assuming the process has an autoregressive representation,

$$\gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{g_{yx}(z)}{B(z^{-1})} \right]_+ . \tag{3.30}$$

Equation (3.30) is the fundamental formula for optimal signal extraction and prediction. To obtain the result for prediction we set $y_t = x_{t+\nu}$, $\nu > 0$. Then

$$(3.31) \qquad g_{yx}(z) = \sum_{k=-\infty}^{\infty} z^k E x_{t+\nu} x_{t-k}$$

$$= z^{-\nu} \sum_{k=-\infty}^{\infty} z^{k+\nu} c(k+\nu)$$

$$= \frac{\sigma^2 B(z) B(z^{-1})}{z^{\nu}} ,$$

whence

$$(3.32) \qquad \gamma(z) = \frac{1}{\sigma^2 B(z)} \left[\frac{\sigma^2 B(z) B(z^{-1})}{B(z^{-1}) z^{\nu}} \right]_+$$

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

where $\gamma(z)$ is now the generating transform for the prediction of $x_{t+\nu}$. There is no need to restrict $\{y_t\}$ to observable series, it is only necessary that we have sufficient information about its stochastic properties to be able to specify $g_{yx}(z)$. Equation (3.30) then expresses the generating transform for the estimate $\hat{y}_t$. If y_t is an unobserved component of the series x_t, we say $\gamma(z)$ is the generating transform of the optimal extraction.

Let $y_{t+\theta}$ be a weighted average of future values of x_t

$$(3.33) \qquad y_{t+\theta} = \sum_{k=1}^{\theta} w_k x_{t+k} .$$

Suppose we wish to forecast not $x_{t+\nu}$ but $y_{t+\theta}$. We may apply (3.30). Since

$$Ey_{t+\theta}x_{t-j} = E\sum_{k=1}^{\theta} w_k x_{t+k}x_{t-j}$$

$$= \sum_{k=1}^{\theta} w_k c(k+j) \ , \quad j = 0, \pm 1, \ldots,$$

$$g_{yx}(z) = \sum_{j=-\infty}^{\infty} \sum_{k=1}^{\theta} w_k c(k+j) z^j \tag{3.34}$$

$$= \sum_{k=1}^{\theta} w_k z^{-k} \sum_{j=-\infty}^{\infty} z^{k+j} c(k+j)$$

$$= \sigma^2 \sum_{k=1}^{\theta} w_k z^{-k} B(z)B(z^{-1})$$

$$= \sigma^2 W(z^{-1})B(z)B(z^{-1})$$

where $W(z) = \sum_{k=1}^{\theta} w_k z^k$. Let the generating transform of the optimal weights in

$$\hat{y}_{t+\theta} = \sum_{j=0}^{\infty} \eta_{\theta j} x_{t-j}$$

by $\eta_\theta(z)$. From (3.30) and (3.34), we obtain

$$\eta_\theta(z) = \frac{1}{B(z)} \left[W(z^{-1})B(z) \right]_+ \tag{3.35}$$

$$= \sum_{k=1}^{\theta} w_k \left\{ \frac{1}{B(z)} \left[\frac{B(z)}{z^k} \right]_+ \right\}$$

$$= \sum_{k=1}^{\theta} w_k \gamma_k(z)$$

where $\gamma_k(z)$ is the generating transform for the optimal autoregressive predictor of x_{t+k}. Inserting the backward shift operator in place of z in (3.35) and applying the resulting operator to x_t gives us

$$\hat{y}_{t+\theta} = \eta_\theta(U)x_t \tag{3.36}$$

$$= \sum_{k=1}^{\theta} w_k \gamma_k(U)x_t$$

$$= \sum_{k=1}^{\theta} w_k \hat{x}_{t+k} ,$$

so that the optimal prediction for a weighted sum is just the weighted sum of separately optimal predictions. This result is useful in deriving optimal seasonal adjustments of type C as described in Section 1 above.

4. Choice of a Simulation Model and Derivation of Optimal Extractions and Predictions

In this section we show that a simple model of the general form (2.1), (2.5) and (2.6) yields a spectral density function with characteristics similar to those of many observed time series. The simple model is used to generate a simulated time series and to obtain theoretically optimal seasonal adjustments according to various criteria.

Figure 1 shows estimated spectral densities for automotive sales and inventories. The estimates exhibit two features commonly

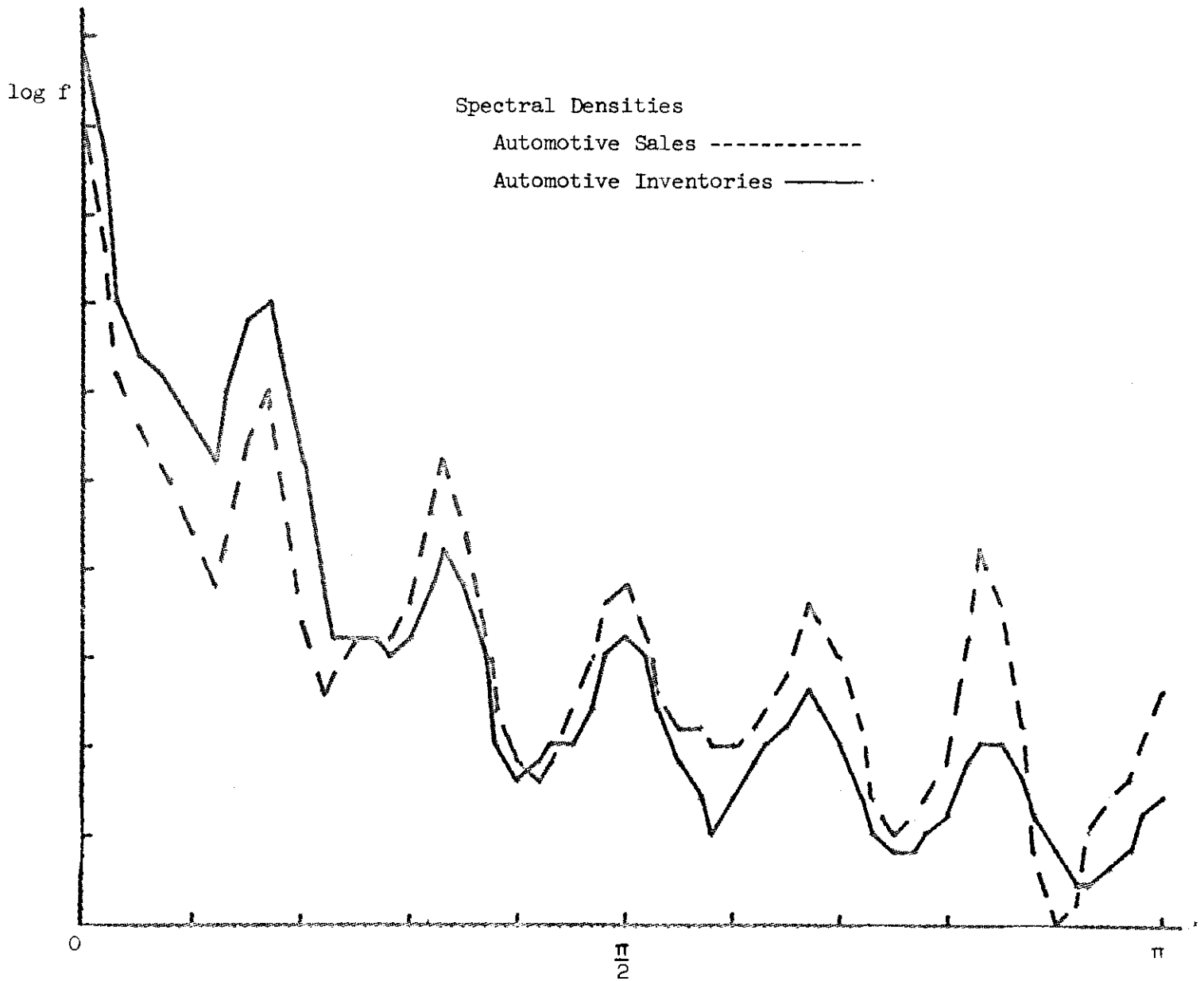

log f
Spectral Densities
Automotive Sales ----------
Automotive Inventories ———
0
π/2
π

FIGURE 1

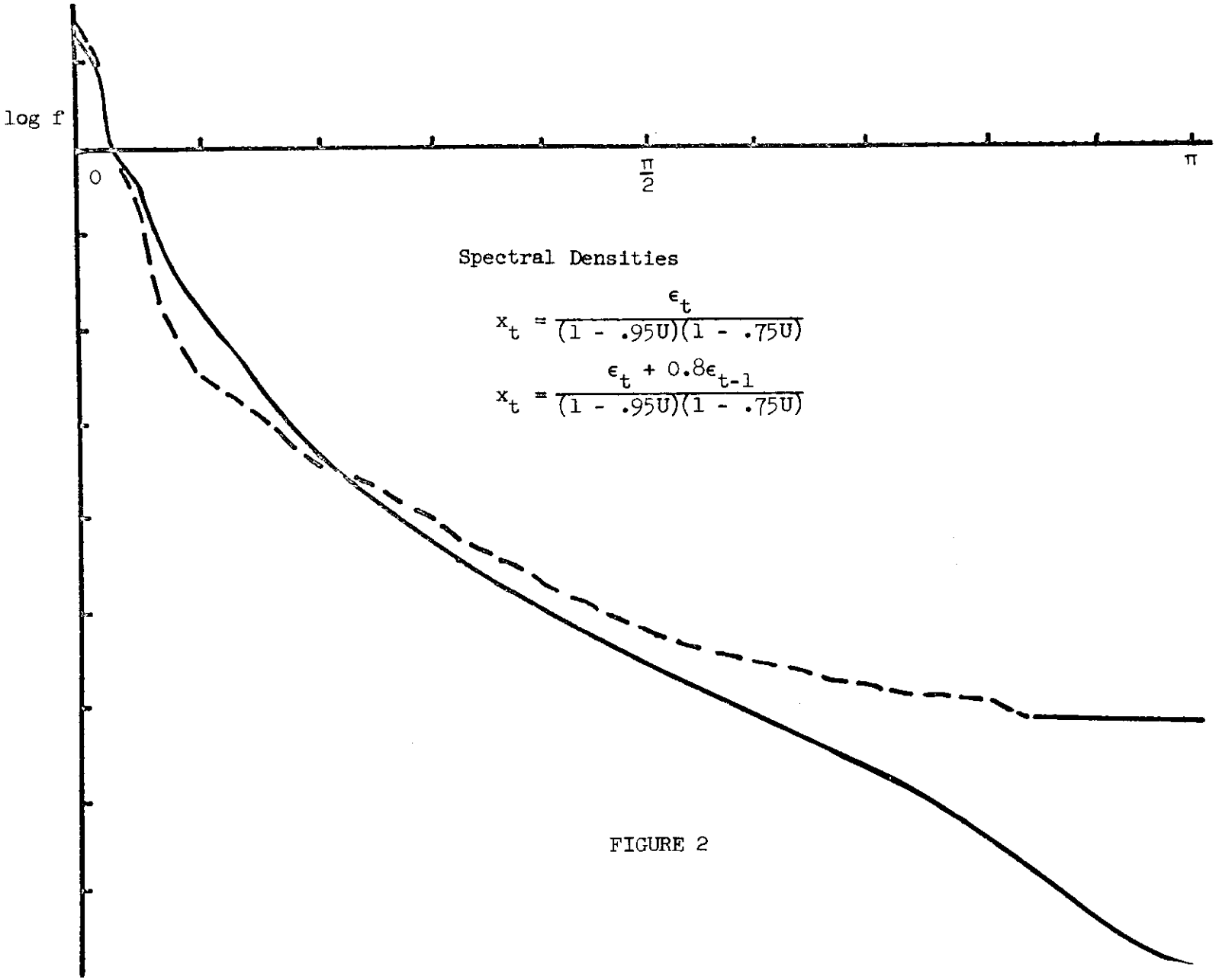
found in economic time series. First, there are peaks at each of the so-called seasonal frequencies (with monthly data these frequencies correspond to 1, 2, 3, 4, 5, and 6 cycles per year). Second, apart from the seasonal peaks, the power spectra are generally decreasing with frequency, having substantially more power at the frequencies near zero than anywhere else.

In our analysis, we decided to use a three component model with components corresponding to the traditional "trend-cycle", "seasonal", and "irregular". In choosing a model, we sought one which was simple, i.e., had only a few parameters, and one which had a spectral density with characteristics which could be associated with particular components of the sort assumed.

The "trend-cycle" should be a series, the time pattern of which is dominated by gradual cumulative movements and does not have any prominent short term regularities. In terms of its spectral density, it should have maximum power at the origin and decreasing power throughout its range. Many low-order autoregressive processes have spectra of the type described. Figure 2 shows the spectral densities of

$$(4.1)\qquad \begin{cases} X_t = \dfrac{\epsilon_t}{(1 - .95U)(1 - .75U)}\,, & \{\epsilon_t\} \text{ white noise} \\[2ex] T_t = \dfrac{\epsilon_t + 0.8\,\epsilon_{t-1}}{(1 - .95U)(1 - .75U)} \end{cases}$$

while both models meet the general requirements mentioned, the second



Spectral Densities

$$x_t = \frac{\epsilon_t}{(1 - .95U)(1 - .75U)}$$

$$x_t = \frac{\epsilon_t + 0.8\epsilon_{t-1}}{(1 - .95U)(1 - .75U)}$$

FIGURE 2

model was chosen since its spectrum declines somewhat faster than the first at the high frequencies. This allows almost all of the power at the higher frequencies to be contributed by the seasonal and irregular components.

The simplest process with peaks at each of the seasonal frequencies is

$$z_t = \frac{v_t}{(1 - \alpha U^L)}, \quad \{v_t\} \text{ white noise},$$

where L is the number of observations per year. A process of this type amounts to specifying L independent processes each of which appears as a first-order autoregression when observed at annual intervals. Figure 3 shows the spectrum of z_t ($\alpha = 0.9$, $L = 12$); note that the peaks at each seasonal frequency are the same height, which is not characteristic of real economic time series. To produce smaller peaks at the higher seasonal frequencies the following model was used:

$$S_t = \frac{v_t + 0.6v_{t-1}}{1 - 0.9U^{12}}. \tag{4.2}$$

The spectral density of this process is graphed in Figure 4.

The irregular component used is a white-noise sequence. Although the irregular component need not be serially uncorrelated, e.g., major strikes or international incidents may have effects which last for several months, we nonetheless chose an uncorrelated sequence for simplicity. In principle we could have used

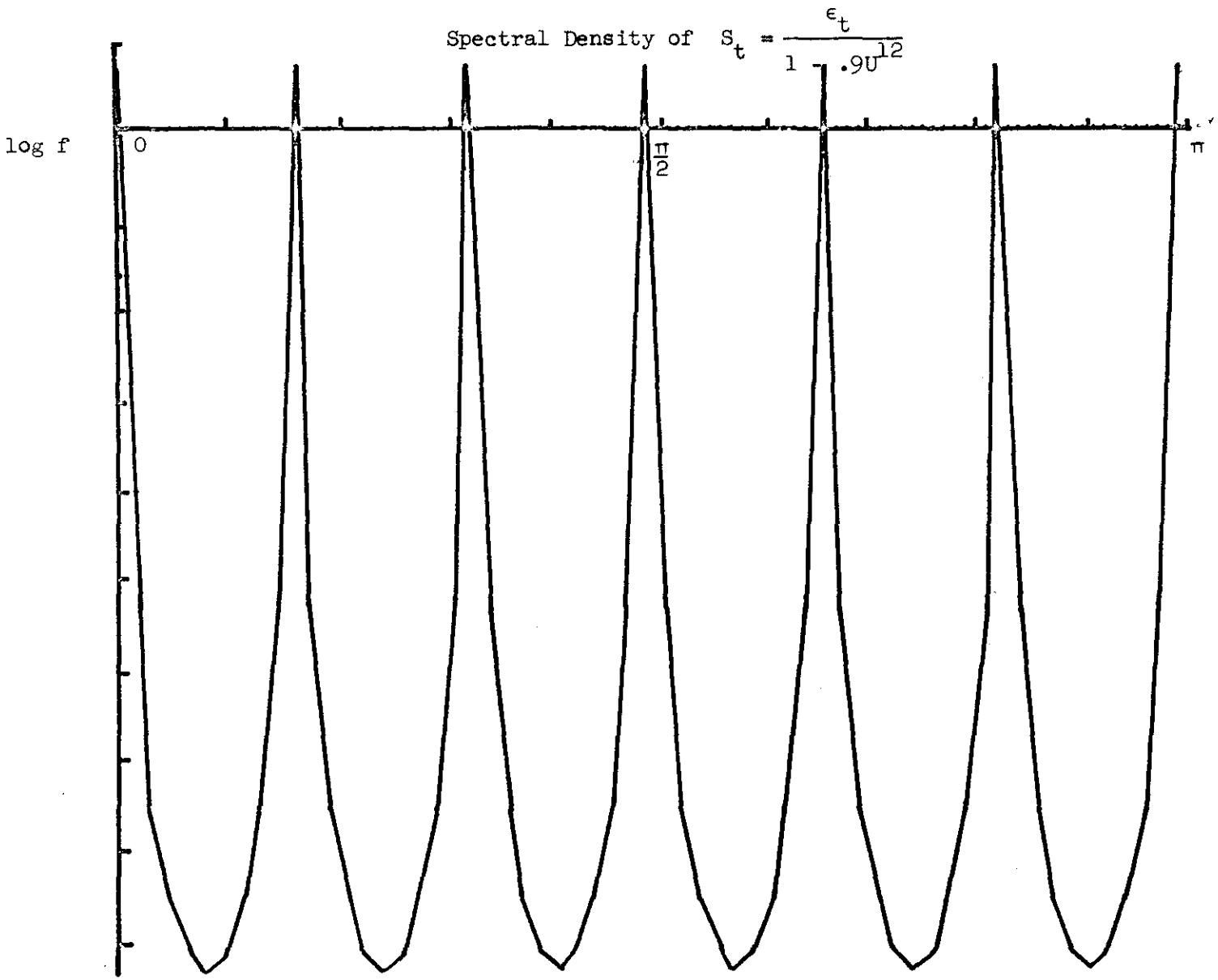

Spectral Density of $S_t = \frac{\epsilon_t}{1 - .9U^{12}}$
log f
0
$\frac{\pi}{2}$
π

FIGURE 3

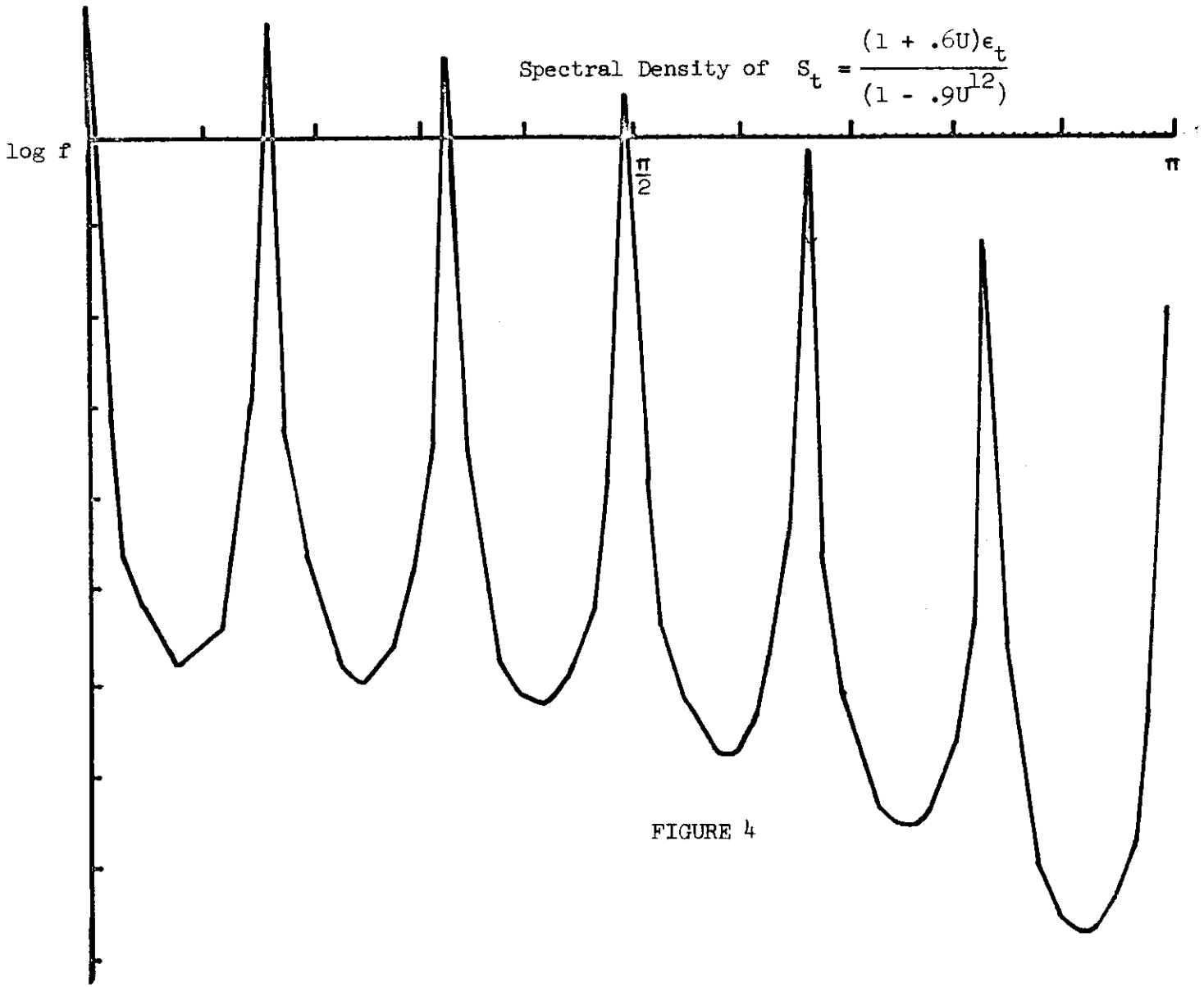

Spectral Density of $S_t = \frac{(1 + .6U)\epsilon_t}{(1 - .9U^{12})}$
log f
$\frac{\pi}{2}$
π

FIGURE 4

$$I_t = \sum_{i=0}^{L-1} \alpha_i \eta_{t-i} \ , \quad \{\eta_t\} \text{ white noise,}$$

where L is the number of observations per year.[12] Unless the contribution of the irregular component to the overall variance of the time series is substantial, varying the model chosen for it has little effect on the spectral characteristics of the series assumed to be observed; hence, we used the simplest specification. Thus

$$(4.3) \qquad X_t = T_t + S_t + I_t$$

$$= \frac{\epsilon_t + 0.8\,\epsilon_{t-1}}{(1 - .95U)(1 - .75U)} + \frac{v_t + 0.6v_{t-1}}{(1 - .9U^{12})} + \eta_t \ ,$$

$\{\epsilon_t\}$, $\{v_t\}$, $\{\eta_t\}$ mutually uncorrelated white noise sequences. The variances of the noise sequences were chosen so that 85 percent of the variance of X_t is accounted for by the trend-cycle, 10 percent by the seasonal component, and 5 percent by the irregular. Figure 5 gives the spectral density of X_t . While we attempted to pick a model whose spectral shape was realistic, it should be noted that we did not attempt to duplicate the spectral density of any particular economic time series. The spectral shape in Figure 1 is merely intended to illustrate the characteristic sought. In practice, estimated spectra are quite irregular in appearance even when the true spectral densities are highly regular. The model used is a realistic one in the sense that its second moments, or spectral

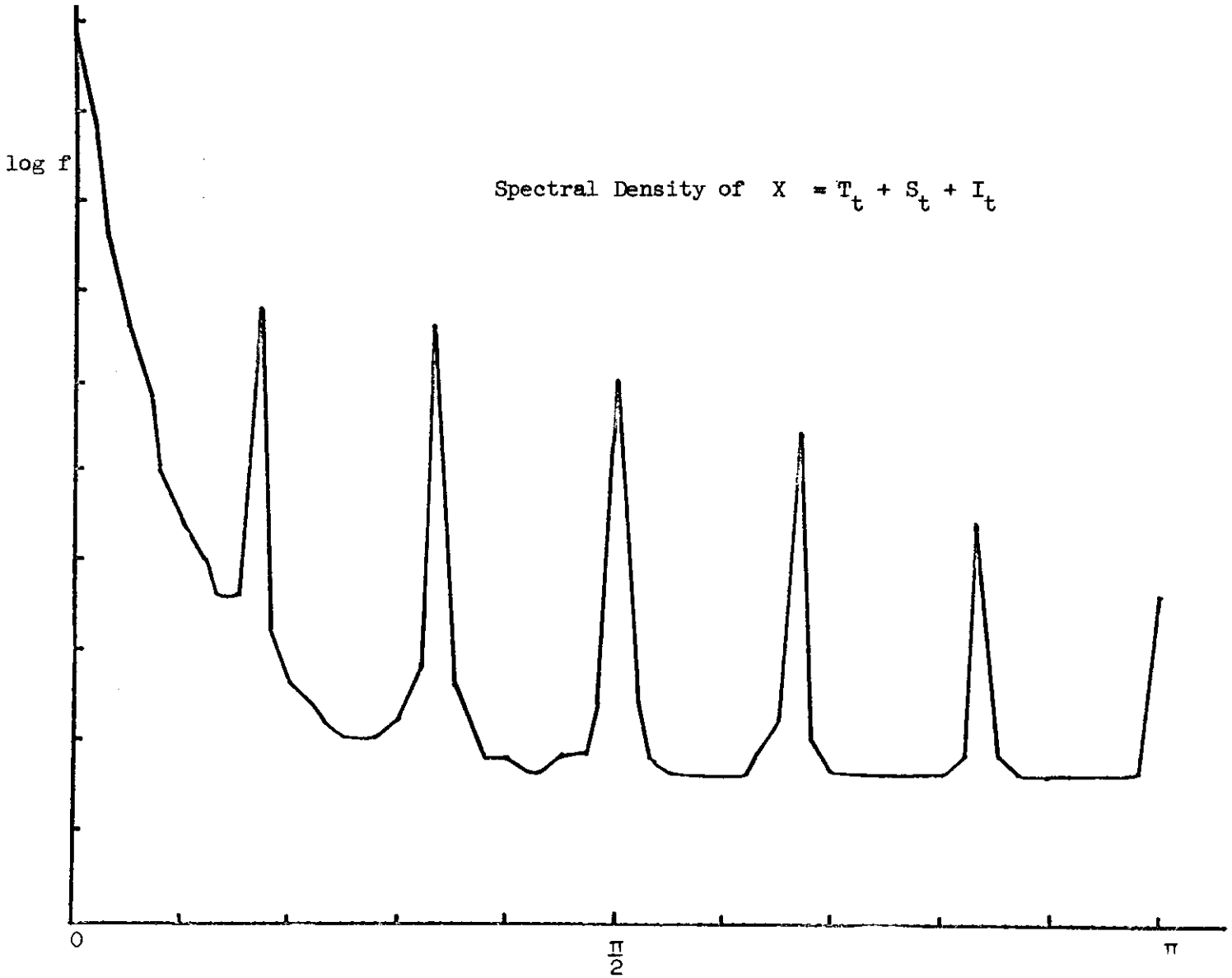


FIGURE 5

density, exhibits characteristics similar to those of many real economic time series.

The series analyzed were simulated by generating normal random variables, with the appropriate variances, to be used as inputs to the several components of the model. Each component was initialized by picking a random variable from a population with variance equal to the variance of the component; 480 observations were then generated, of which the first 120 were discarded. Figure 6 shows that last 100 observations of the simulated time series.

To apply the theory outlined in the previous section, we need the canonical factorization of the covariance generating function of X_t. Since the individual components are orthogonal,

$$(4.4)\quad g_{XX}(z) = g_{TT}(z) + g_{SS}(z) + g_{II}(z)$$

$$= \frac{\sigma_\epsilon^2(1+.8z)(1+.8z^{-1})}{(1-.95z)(1-.75z)(1-.95z^{-1})(1-.75z^{-1})}$$

$$+ \frac{\sigma_v^2(1+.6z)(1+.6z^{-1})}{(1-.9z^{12})(1-.9z^{-12})} + \sigma_\eta^2$$

$$= \sigma^2 \frac{P(z)P(z^{-1})}{(1-.95z)(1-.75z)(1-.9z^{12})(1-.95z^{-1})(1-.75z^{-1})(1-.9z^{-12})}$$

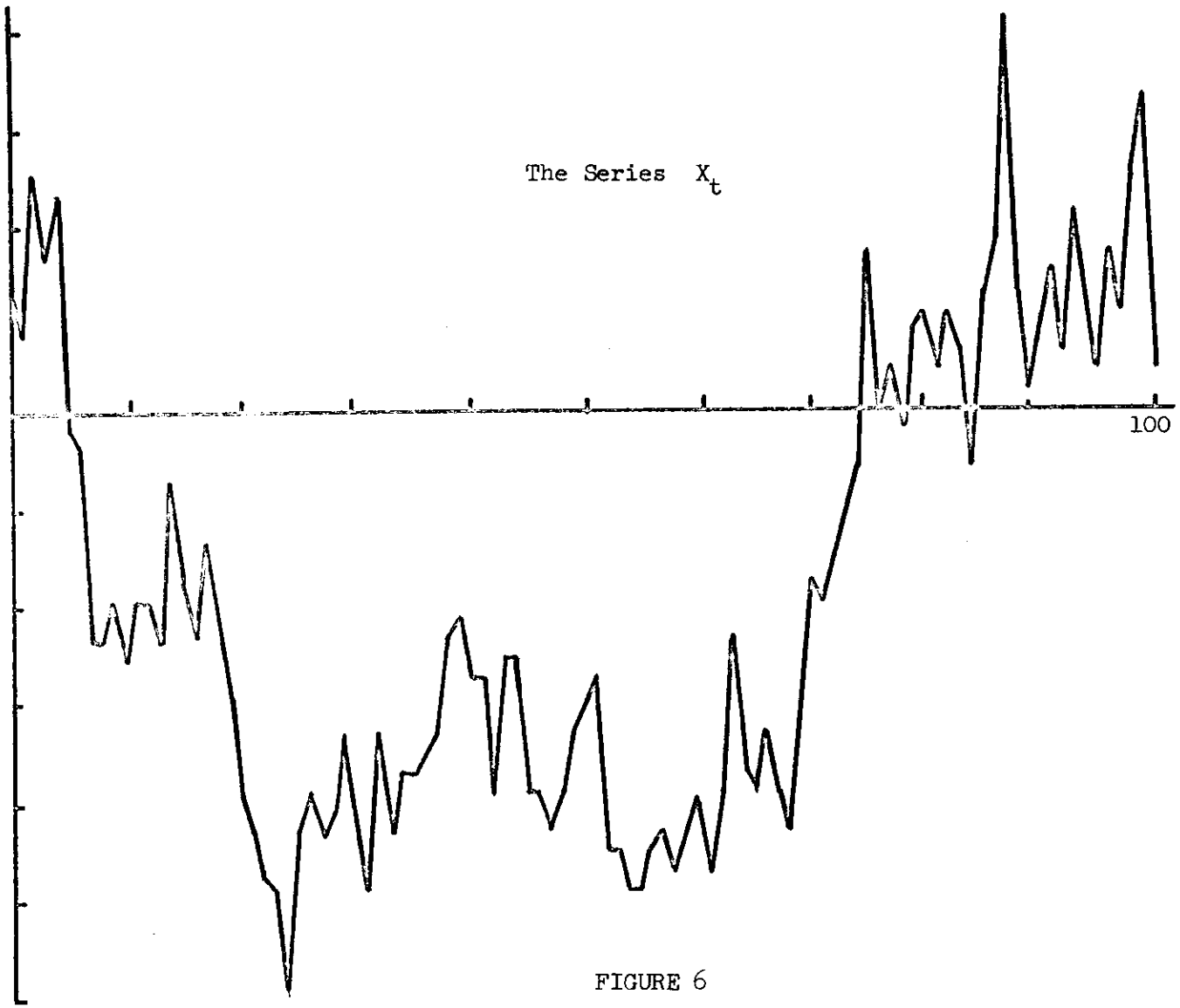


FIGURE 6

where

$$\sigma^2 P(z)P(z^{-1}) = \sigma_\epsilon^2(1+.8z)(1+.8z^{-1})(1-.9z^{12})(1-.9z^{-12})$$

$$+ \sigma_\nu^2(1+.6z)(1+.6z^{-1})(1-.95z)(1-.75z)(1-.95z^{-1})(1-.75z^{-1})$$

$$+ \sigma_\eta^2(1-.95z)(1-.75z)(1-.9z^{12})(1-.95z^{-1})(1-.75z^{-1})(1-.9z^{-12}).$$

σ^2 is chosen so that the constant in $P(z)$ is one.[13]

Once the polynomial $P(z)$ is obtained the predictions and extractions can be calculated using (3.30). As an example, consider the problem of finding the least squares estimate of $T_{t+\nu}$ at time t (ν may be any integer). Since the components are uncorrelated with one another

$$g_{TX}(z) = g_{TT}(z) \; ; \tag{4.5}$$

hence, the generating function of the sequence of weights is given by

$$\gamma(z) = \frac{(1-.95z)(1-.75z)(1-.9z^{12})}{P(z)}\left[\frac{(1+.8z)(1+.8z^{-1})(1-.95z^{-1})(1-.75z^{-1})(1-.9z^{-12}}{(1-.95z)(1-.95z^{-1})(1-.75z)(1-.75z^{-1})P(z^{-1})z^{\nu}}\right]_+$$

$$= \frac{(1-.95z)(1-.75z)(1-.9z^{12})}{P(z)}\left[\frac{(1+.8z)(1+.8z^{-1})(1-.9z^{-12})}{(1-.95z)(1-.75z)P(z^{-1})z^{\nu}}\right]_+ \tag{4.6}$$

Generating functions for estimates of the other components and for predictions can be obtained by evaluating similar expressions. Since

$$(4.7) \qquad \hat{X}_{t+\nu,t} = \hat{T}_{t+\nu,t} + \hat{S}_{t+\nu,t} + \hat{I}_{t+\nu,t} \,, \quad {}^{14}$$

it is not necessary to evaluate generating functions for each of the four cases explicitly. This is especially helpful if ν is negative since in that case the best predictor of $X_{t+\nu}$ as of time t is clearly $X_{t+\nu}$.

The calculations needed to evaluate (4.6) can be easily handled by using the following theorem of Whittle [26, p. 93]:

Let $Q(z)$ be a function of z analytic in $\rho < |z| < \rho^{-1}$, and let θ be a number such that $|\theta| < 1$. Then

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

where $\Pi_p(z)$ is a polynomial in z of degree p^{-1}, so chosen that the differential coefficients of orders $0, 1, \ldots, p^{-1}$ of $R(z)$ are respectively equal to those of $Q(z)$ at $z = \theta^{-1}$

i.e. $\Pi_p(z) = \sum_0^{p-1} \frac{Q^{(j)} - (\theta^{-1})}{j!} (z - \theta^{-1})^j$ where

$$Q_-(z) = [Q(z)]_-$$

Identifying $\frac{(1+.8z)(1+.8z^{-1})(1-.9z^{-12})}{P(z^{-1})z^{\nu}}$ in (4.6) with $Q(z)$ and expanding $(1-.95z)^{-1}(1-.75z)^{-1}$ by partial fractions, we may apply Whittle's theorem to each term in the sum and obtain $\gamma(z)$ as a rational function of z.

5. Simulation Results

In this section the results of using the various methods of "optimal" seasonal adjustment on our artificial series are reported.

Consider the special case of extracting the trend-cycle component when the entire history of X_t is known. The generating transform of the optimal weights is

$$\gamma(z) = \frac{g_{XT}(z)}{g_{XX}(z)} = \frac{g_{TT}(z)}{g_{XX}(z)} \,. \tag{5.1}$$

[15]

The spectral density of the estimate is

$$\begin{aligned} f_{\hat{T}\hat{T}}(\lambda) &= \gamma(e^{i\lambda})\gamma(e^{-i\lambda})f_{XX}(\lambda) \\ &= \frac{f_{TT}(\lambda)}{f_{XX}(\lambda)/f_{TT}(\lambda)} \\ &= \frac{f_{TT}(\lambda)}{1 + \dfrac{f_{SS}(\lambda)}{f_{TT}(\lambda)} + \dfrac{f_{II}(\lambda)}{f_{TT}(\lambda)}} \end{aligned} \tag{5.2}$$

The ratios $f_{SS}(\lambda)/f_{TT}(\lambda)$ will generally be large near the seasonal frequencies. Thus "optimal" adjustment will produce dips in the spectral density of the adjusted series at those frequencies. (See, for example, Figure 7.) Such dips do not represent "overadjustment," as Nerlove [14] thought, but rather are characteristic of this sort of "optimal" adjustment.

Since the filter used is linear and symmetric, the coherence will be one at every frequency and the phase shift zero; however, the spectral density of the adjusted series will show a loss of power at every frequency. Note that not only will there be a loss of power relative to the spectrum of the observed series X_t, but from (5.2) it is clear that there will be a loss of power relative to the spectrum of the trend-cycle itself. Figure 8 shows that the same phenomena can be produced with filters which are only moderately two-sided. That the dips are not consequences of over adjustment can be seen from Figure 9 which shows the estimates, $\hat{T}_t$, and the true values, T_t. The "adjusted" series does well in tracking the trend cycle, however. (See Figure 15 below.)[16]

Using one-sided filters only does not lead to dips at the seasonal frequencies (one-sided filters occur when forecasting or estimating the current value of a component). As Figures 10-14 show, however, the other effects noted in [14] are still present. Relative to the observed series all methods produce "distortions" at all frequencies and generally lead to a loss of power throughout the entire range.

In the case of one-sided filters, the filters are not symmetric, and the phase angle is thus generally different from zero, although this does not imply that the method of adjustment used is improper. The coherence between the adjusted and the unadjusted series is unity at every frequency because the filters used were linear.

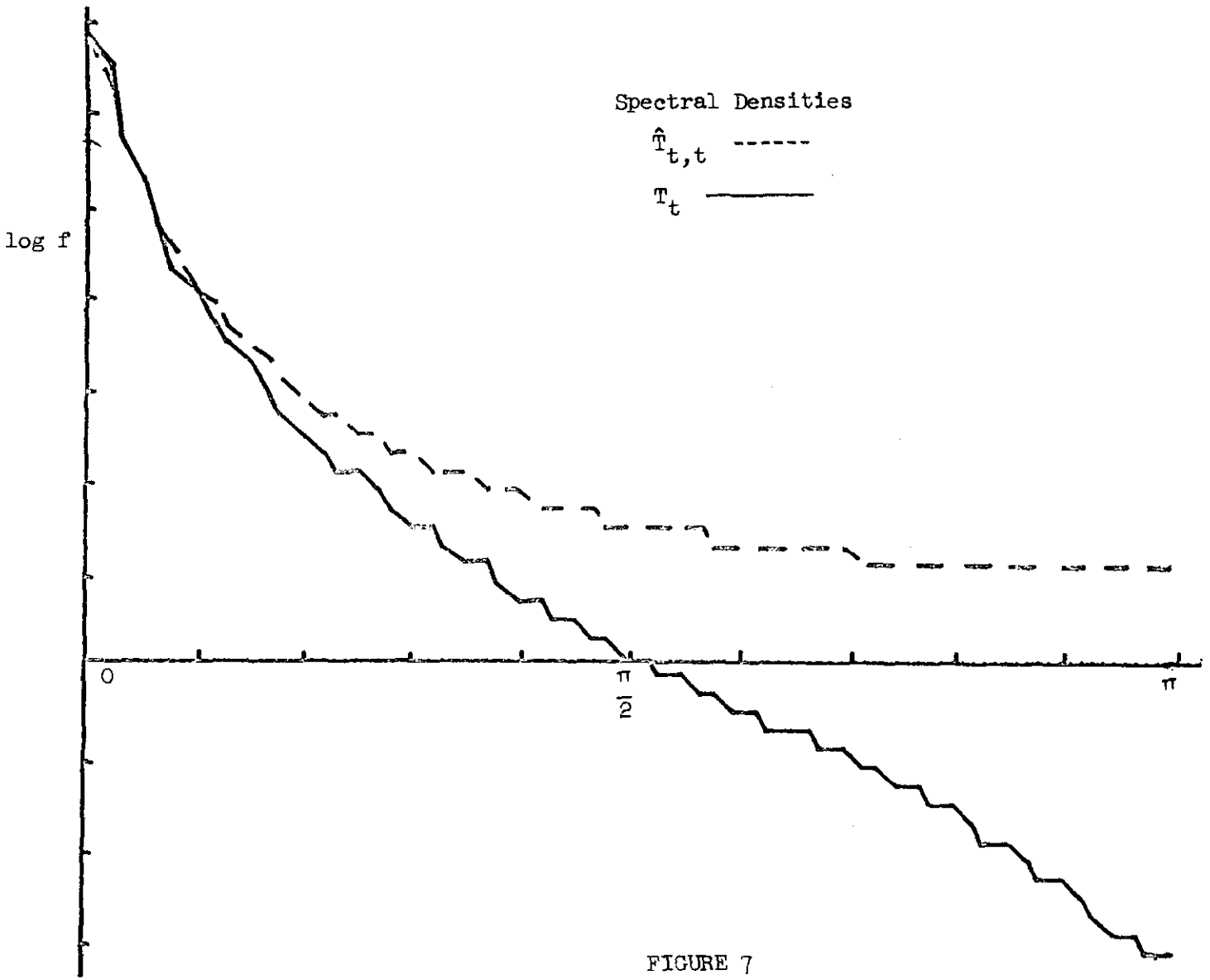


FIGURE 7

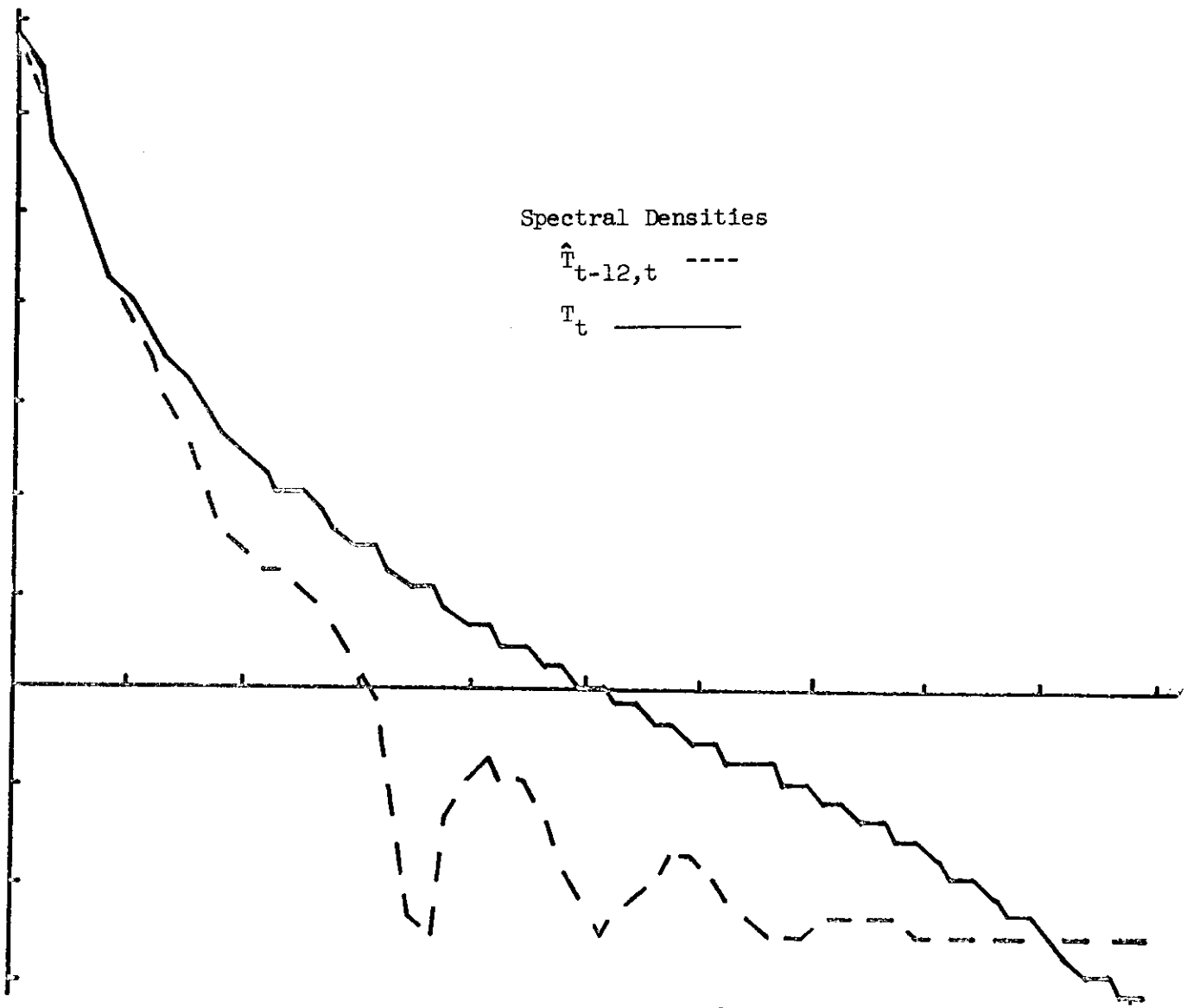


FIGURE 8

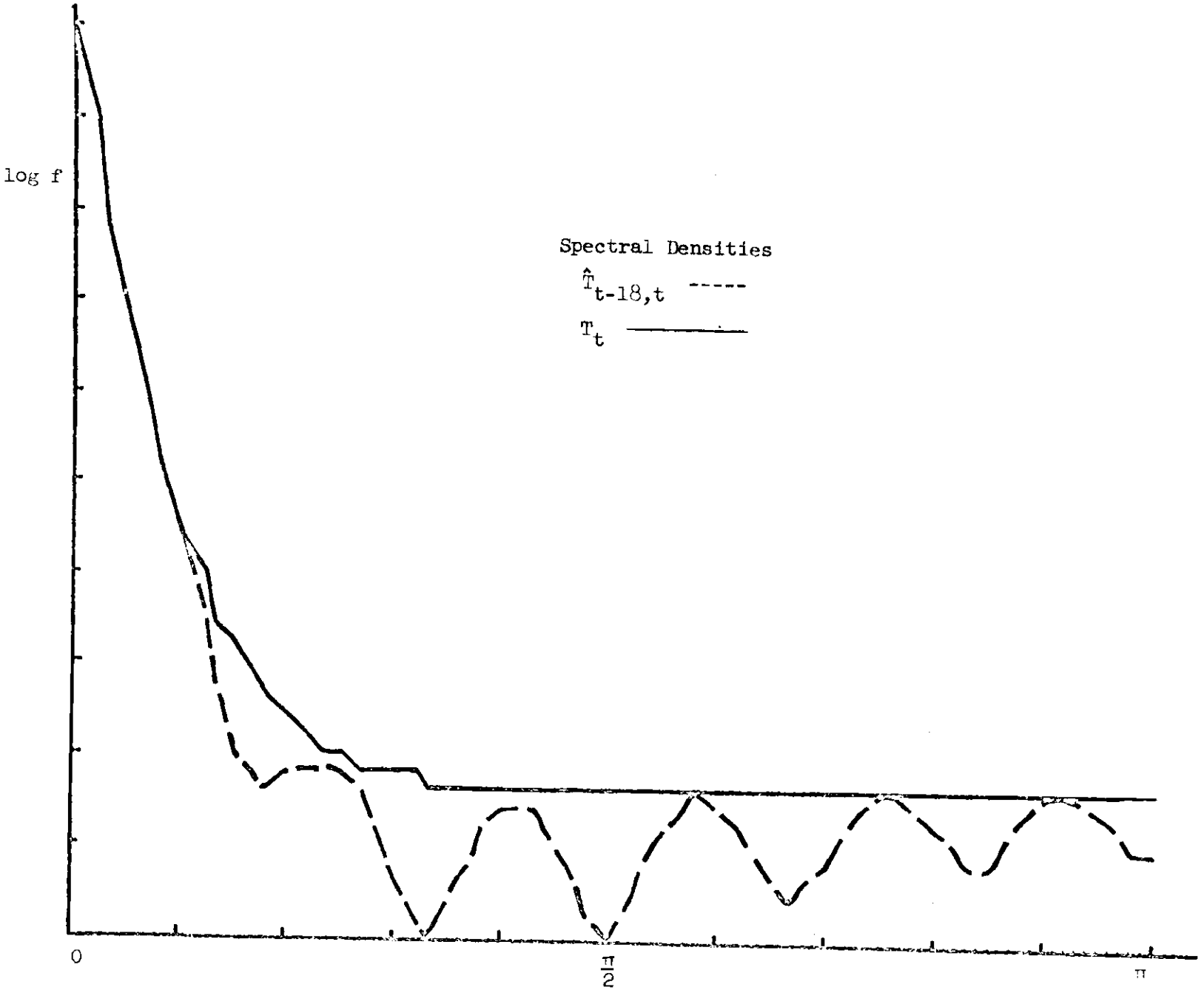


FIGURE 9

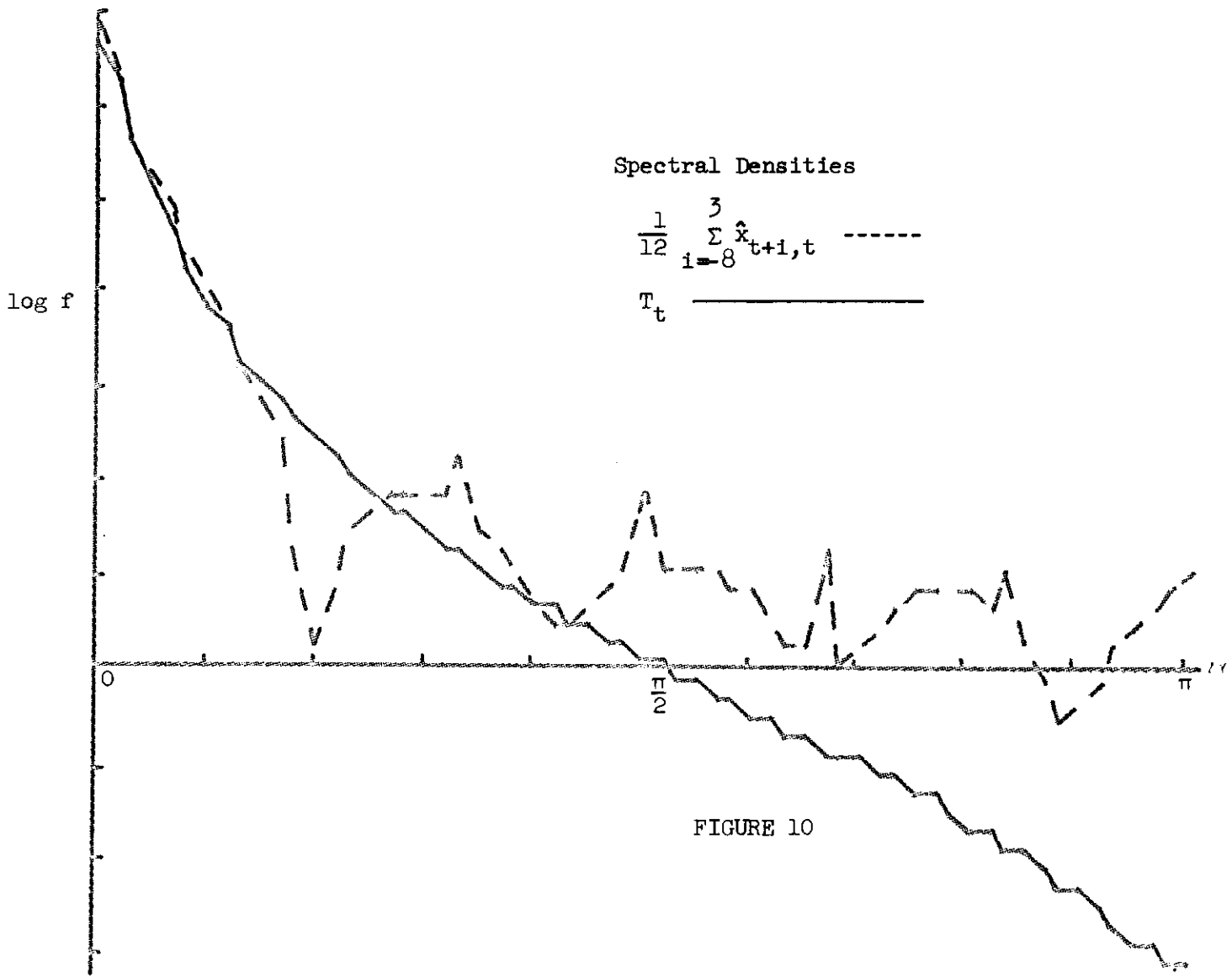


FIGURE 10

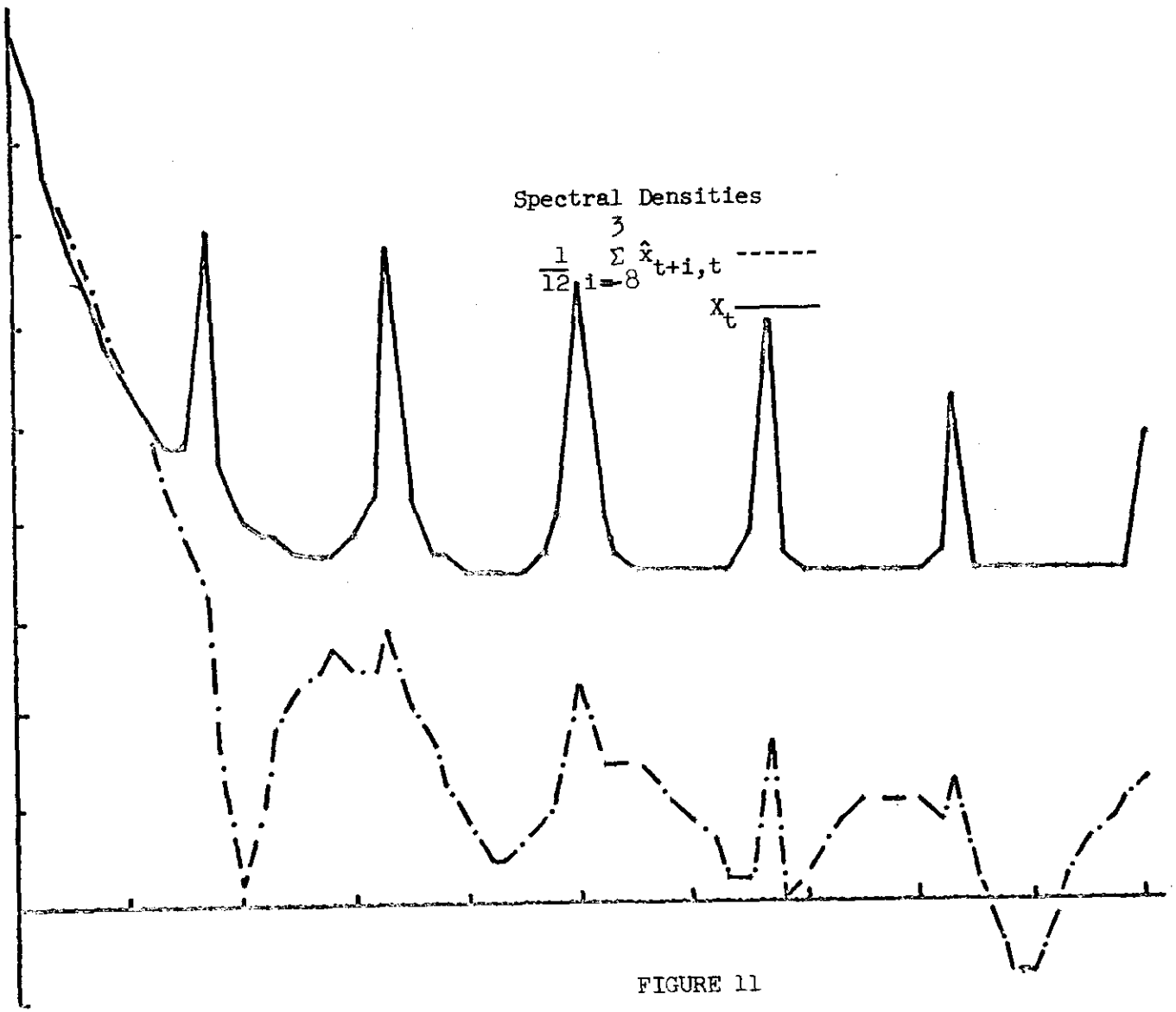


FIGURE 11

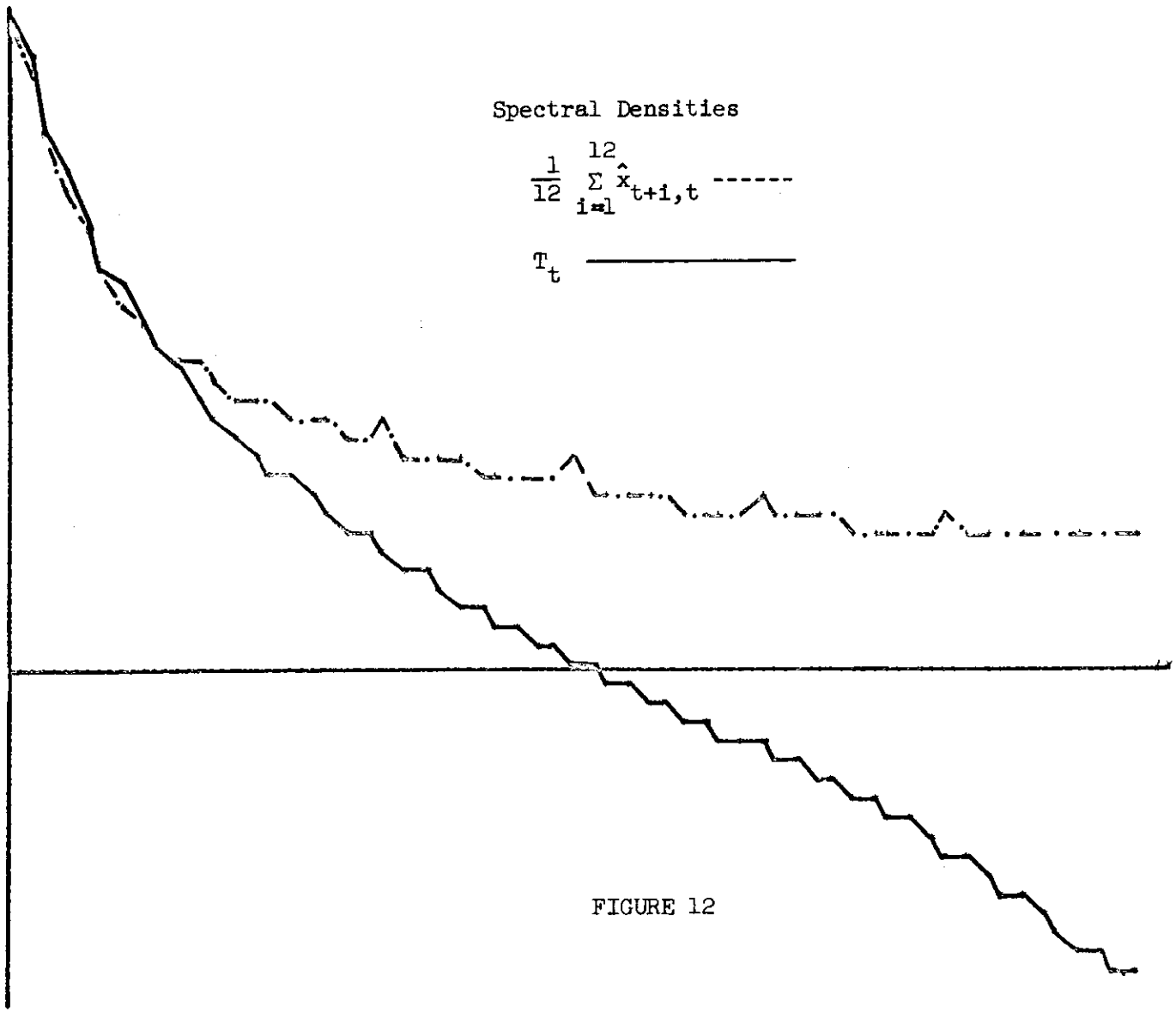


FIGURE 12

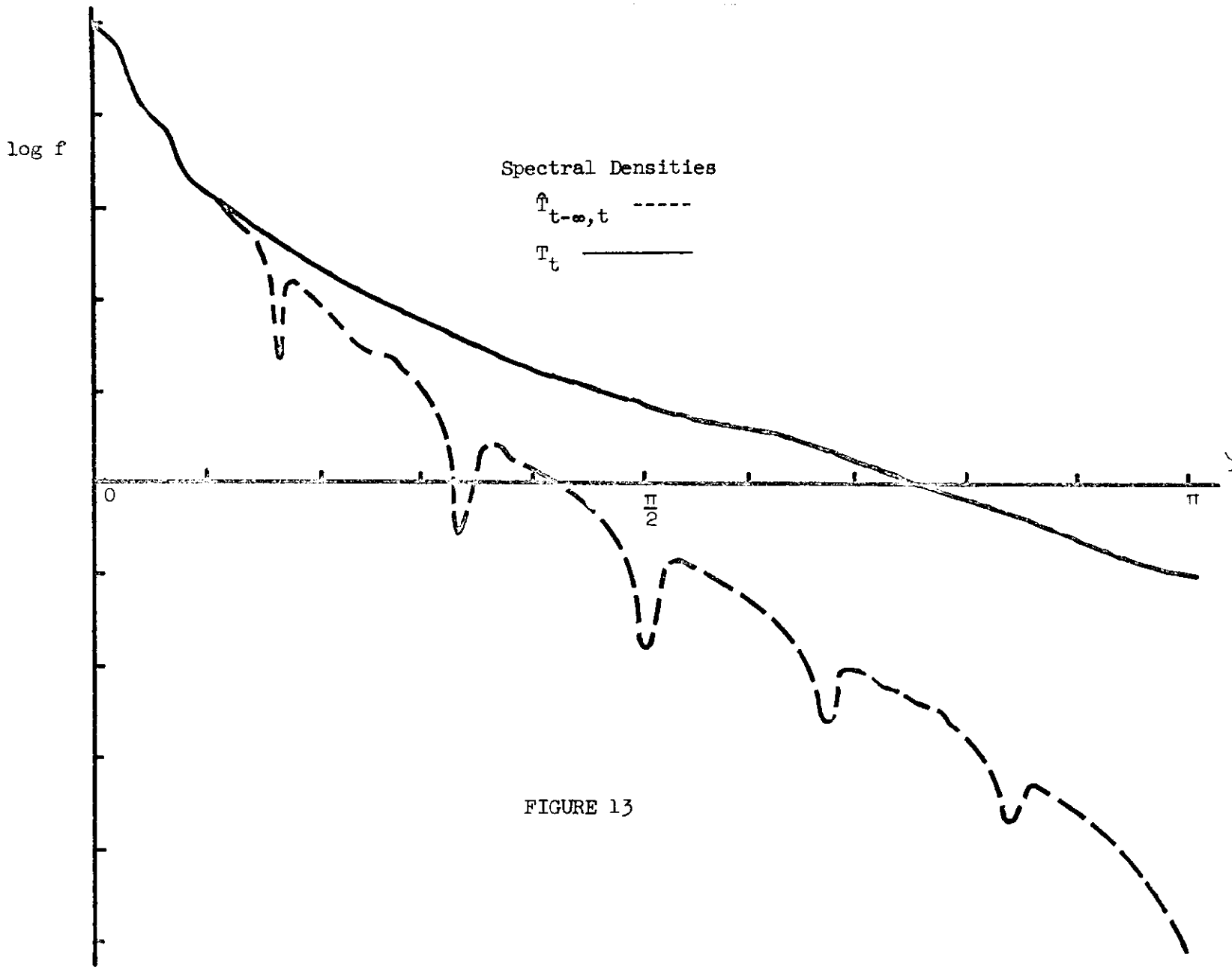


FIGURE 13

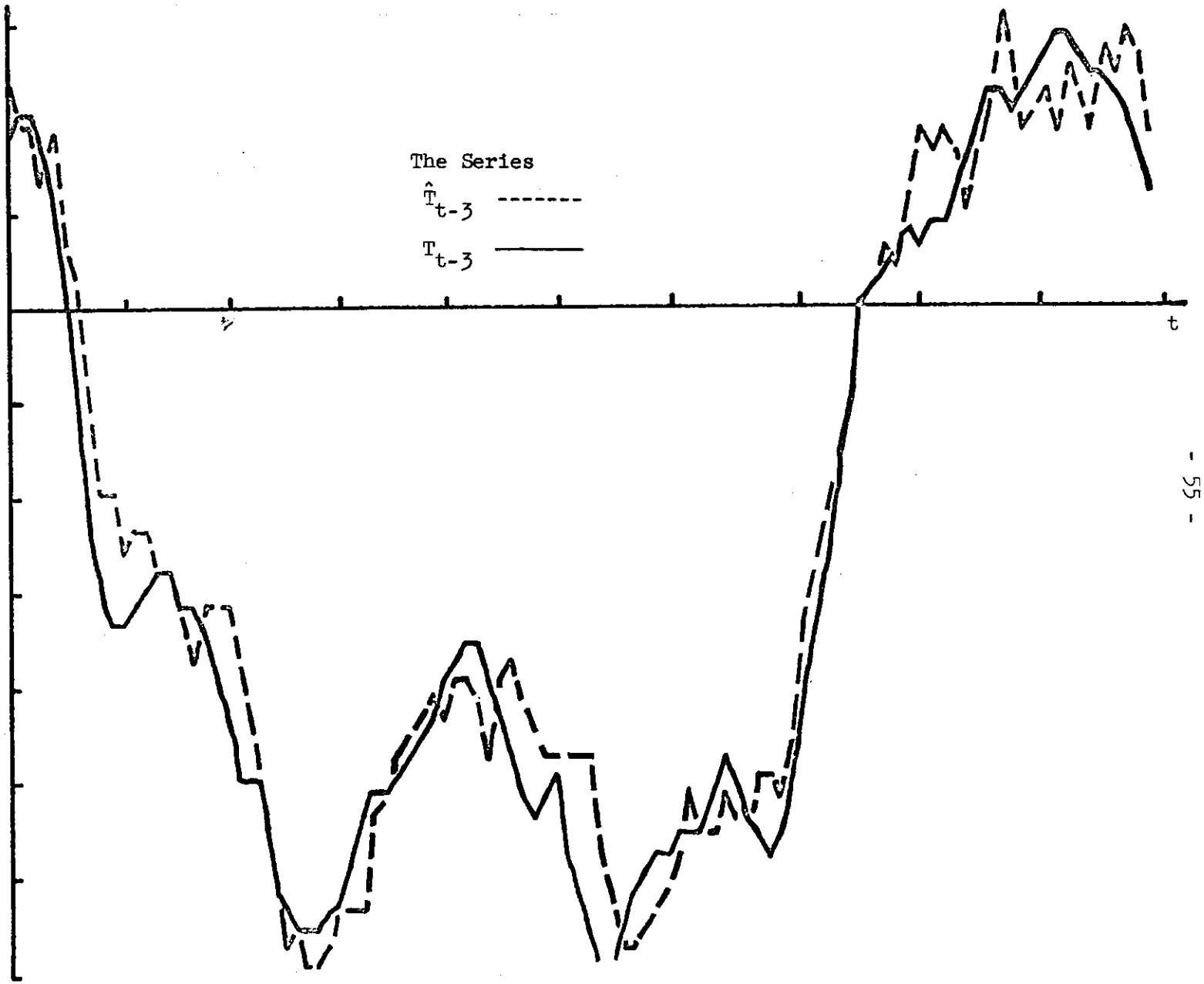


FIGURE 14

However, if one were to estimate the coherence between the series using the simulated forecasts and the series X_t the estimates might be substantially below one. When such coherences were estimated this result was indeed obtained. It should be pointed out that low coherence in this case derives only from sampling errors and truncation of the forecasting or extraction operator.

Our simulation results clearly suggest that all of the undesirable features noted in the spectral comparisons of the unadjusted and BLS seasonally adjusted unemployment series in [14] are reproduced by the three sorts of "optimal" adjustments here considered. Since the criteria of optimality are rather plausible, it can only be concluded that the spectral criteria suggested in [14] leave much to be desired. Furthermore, the results obtained do not depend on sampling problems or the difficulties of too short series; they are logical consequences of the alternative assumptions we have made concerning the objectives of seasonal adjustment. To the extent that the BLS and Census Methods of adjustment have succeeded in matching the characteristics of these "optimal" procedures, which, please note, are applied under ideal circumstances, they represent a truely remarkable achievement of trial and error. The achievement is especially noteworthy in that it is based on no formal model and little understanding of the nature of seasonality.

6. A Conclusion

The effects, desirable or undesirable, of a particular method of seasonal adjustment can only be assessed properly in the time domain and only in relation to the objectives of such adjustment. In making an assessment we are greatly handicapped by the inadequate attention which practitioners of seasonal adjustment have paid to the purposes of such adjustment and the lack of a clearly formulated conception of the nature of seasonality. A sound basis for the use of unobserved component models in economic analysis has never, apparently, been developed; such models were introduced uncritically by early nineteenth century economic statisticians who also worked with meteorological and astronomical data. In [14] seasonality was defined as that characteristic of an economic time series which gives rise to peaks at seasonal frequencies and an attempt was made to develop informal, but generally applicable, criteria for "good" seasonal adjustment in spectral terms. Any criteria should reflect time-domain effects even if couched in frequency terms. Lack of clear objectives precluded formulation of appropriate criteria and spectral criteria were developed in the hope that they might serve as second best.

While we reserve judgment on the general question of the applicability of unobserved component models in economic analysis, we do here adopt such an approach as a working hypothesis. In terms of the traditional simple three-component model, modified by the addition of appropriate stochastic assumptions, we formulate three plausible

objectives of "seasonal adjustment." We show how the minimum mean-square error criterion may be used to obtain "optimal" methods of seasonal adjustment, neglecting sampling problems. We show both empirically and theoretically that such "optimal" methods of adjustment reproduce many of the features of the relation between seasonally adjusted and unadjusted series noted in [14].

We conclude, not that spectral comparisons are useless, but rather that such comparisons must be interpreted with great care. Clearly, the criteria suggested in [14] were naive. Further research must emphasize objectives and models. Whether these are formulated in frequency terms or in the time domain is of secondary importance. What is relevant is the empirical validity of the models and a precise statement of objectives. A more careful analysis by economic policy makers and others concerned with the uses of economic statistics of the question of why they prefer seasonally adjusted series (or whether they really do!) should aid immeasurably this endeavor. We must, as well, reexamine our passive acceptance, now of 150 years standing, of the validity of unobserved-components models in the analysis of economic time series.

REFERENCES

[1] Babbage, Charles, "Analysis of the Statistics of the Clearing House during the Year 1839," Statistical Journal, XIX (1858).

[2] Bell, G.M., The Country Banks and the Currency: An Examination of the Evidence on Bank of Issue, Given before a Select Committee of the House of Commons in 1841 (Longman, Brown, Green, and Longman, London, 1842).

[3] Buys Ballot, C.H.D., Les Changements Périodiques de Température (Utrecht: 1847).

[4] Couts, D., D. Grether, and M. Nerlove, "Forecasting Non-Stationary Economic Time Series," Management Science, 13: 1-21 (September 1966).

[5] Forbes, James D., "On the Horary Oscillations of the Barometer near Edinburgh, deduced from 4410 Observations; with an Inquiry into the Laws of the Geographical Distribution of the Phenomenon," Edinburgh Journal of Science, XII, New Series (1832).

[6] Gilbart, J.W., "The Laws of the Currency as Exemplified in the Circulation of Country Bank Notes in England since the Passing of the Act of 1844," Statistical Journal, XVII (1854).

[7] _____, "The Laws of the Currency in Scotland," Statistical Journal, XIX (1856).

[8] Grether, D., "Application of Signal Extraction Techniques in the Study of Economic Time Series," Cowles Foundation Discussion Paper, No. 202 (February 25, 1966).

[9] _____, "Studies in the Analysis of Economic Time Series," unpublished Ph.D. dissertation, Stanford University (January 1969).

[10] Hannan, E.J., "The Estimation of Seasonal Variation in Economic Time Series," Jour. Amer. Stat. Assoc., 58: 31-44 (March 1963).

[11] Jevons, W.S., Investigation in Currency and Finance (London: Macmillan and Co., Ltd., 1909).

[12] Moore, G.H., ed., Business Cycle Indicators, Vol. I (Princeton: Princeton University Press, 1961).

[13] Muller, D., "A Method for Solving Algebraic Equations Using an Automatic Computer," Mathematical Tables and Other Aids to Computation, X (October 1956).

[14] Nerlove, M., "Spectral Analysis of Seasonal Adjustment Procedures," Econometrica, 32: 241-86 (July 1964).

[15] _____, "A Comparison of a Modified 'Hannan' and the BLS Seasonal Adjustment Filters," Jour. Amer. Stat. Assoc., 60: 442-91 (June 1965).

[16] _____, "Distributed Lags and Unobserved Components in Economic Time Series," in W. Fellner, et al., Ten Essays in Economic Theory (New York: John Wiley, 1967).

[17] Nettheim, Nigel F., "A Spectral Study of 'Overadjustment' for Seasonality," Working Paper No. 21, U.S. Bureau of the Census (March 1965).

[18] Organization for Economic Co-operation and Development, Seasonal Adjustment on Electronic Computers (Paris: O.E.C.D., 1961).

[19] Pannekoek, A., A History of Astronomy, (New York: Interscience, 1961).

[20] President's Committee to Appraise Employment and Unemployment Statistics, Measuring Employment and Unemployment (Washington: U.S. Government Printing Office, 1962).

[21] Rosenblatt, Harry M., "Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series," Working Paper No. 23, U.S. Bureau of the Census (November 1965).

[22] _____, "Spectral Evaluation of BLS and Census Revised Seasonal Adjustment Procedures," paper presented at the meeting of the American Statistical Association, Los Angeles, California (August 17, 1966).

[23] _____, "Some Research on the Evaluation of Time Series Methods," memorandum presented to the Census Advisory Committee of the American Statistical Association (April 1, 1967).

[24] _____, "Spectral Evaluation of BLS and Census Revised Seasonal Adjustment Procedures," Jour. Amer. Stat. Assoc., 63: 472-501 (June 1968).

[25] Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series (New York: MIT Technology Press and John Wiley and Sons, 1950).

[26] Whittle, D., Prediction and Regulation by Linear Least-Square Methods (London: The English Universities Press, Ltd., 1963).

[27] Wold, H., A Study in the Analysis of Stationary Time Series (Stockholm: Almqvist and Wiksell, 1938).

FOOTNOTES

[1]For true coherence equal to zero, the phase angle is approximately uniformly distributed in the interval $[-\pi/2, \pi/2]$.

[2]Rosenblatt [23, pp. 4-5] puts the matter felicitously as follows: "Like the physician's stethoscope and electrocardiograph, the spectrum is a highly sensitive instrument. Not only will these instruments display the readily recognizable characteristics of a patient (time series) which are not too difficult to interpret, but they will also point to much finer effects which at first may not be readily understood, but whose meaning may become clearer through research, experimentation, and experience. Often, difficulty in the interpretation of the spectrum will be a reflection of the degree of deviation from the spectral criteria. This process of examination is analogous to the procedure followed by a physician in examining a patient. He compares his findings with his standards for good health; he may classify one individual as being in better health than another, yet, unless deviations from norm are extreme, it is difficult for him to say that his patient will not live a full and fruitful life."

[3]See Nerlove [14, pp. 259-60].

[4]See Pannekoek [19, p. 280].

[5]A more complete and documented history of the idea of unobserved components and the basis for seasonal adjustment is contained in Chapter II of Grether [9].

[6]For example, in his paper in [12], Julius Shiskin writes (p. 530): "Cyclical movements are shown more accurately and stand out more clearly in data that are seasonally adjusted. ... seasonally adjusted data not only avoid some of the biases to which same-month-year- ago comparisons are subject but also often reveal cyclical changes several months earlier. Seasonal adjustments, therefore, help the business statistician to make more accurate and more prompt diagnoses of the current economic situation .... As a general purpose aid, both in historical studies of the business cycle and in studies of current economic trends, seasonal adjustments rank second only to the provision of the raw observations themselves."

In a careful investigation of all aspects of the U.S. employment and unemployment statistics, a committee chaired by R.A. Gordon wrote [20, p. 165]: "... the purpose of seasonal adjustment ... is to bring out the effects of the less transitory factors that affect the series.

Let us imagine that all the underlying economic and social factors that influence a series were held constant for an entire year. The average level of the series over that year would then reflect the impact of these underlying factors, free of transitory seasonal influences."

[7] E.g., $L = 4$ for quarterly data, $L = 12$ for monthly data, and $L = 52$ for weekly data.

[8] This section is based on Grether [9] and Nerlove [16, sec. II] and is introduced here largely to make the expositions self-contained. The basic theory is due to Wiener [25].

[9] I.e., have means which do not depend on the time index t and autocovariances which depend only on the lag involved and not on the absolute value of the time index.

[10] Note the normalization $b_0 = 1$ has been imposed. We could equally well have imposed the alternative normalization $\sigma^2 = 1$.

[11] Because of the nature of $g_{XX}(z)$, these coefficients cannot be zero.

[12] Shiskin [18] states that the irregular component should appear uncorrelated when examined at intervals over one year in length.

[13] The method used to obtain the roots of $P(z)P(z^{-1})$ was a variant of Muller's method [13]. Since the roots of $P(z)P(z^{-1})$ come in reciprocal pairs, if r is any complex root then $\bar{r}$, r^{-1}, $\bar{r}^{-1}$ are also roots. Muller's method was adapted by taking all four numbers as roots when any complex root was found. This considerably reduces the computational load. The 28-degree polynomial could be completely factored by finding only 7 to 14 roots. (Exactly how many depends upon the number of real roots found.) Most procedures for finding roots of polynomials have difficulty if there are multiple roots or clusters of roots which are close together. The method was altered to guard against the latter possibility: If after a fixed number of steps the $P(z)P(z^{-1})$ was not sufficiently close to zero, or if the step size became small, the search was resumed in the neighborhood of the reciprocal of the point at which trouble was encountered. The idea was that even if r_1 and r_2 are close together, $1/r_1$ and $1/r_2$ may be quite far apart and could be more easily distinguished.

[14] The generating function for the ν-step predictor of X_t may be written

$$\gamma(z) = \frac{1}{B(z)}\left[\frac{g_{XX}(z)}{B(z^{-1})z^{\nu}}\right]_+ = \frac{1}{B(z)}\left[\frac{g_{TT}(z) + g_{SS}(z) + g_{II}(z)}{B(z^{-1})z^{\nu}}\right]_+$$

$$= \frac{1}{B(z)}\left[\frac{g_{TT}(z)}{B(z^{-1})z^{\nu}}\right]_+ + \frac{1}{B(z)}\left[\frac{g_{SS}(z)}{B(z^{-1})z^{\nu}}\right]_+ + \frac{1}{B(z)}\left[\frac{g_{II}(z)}{B(z^{-1})z^{\nu}}\right]_+$$

[15] Let $Q = E\{[T_t - \sum_{j=\infty}^{\infty} \gamma_j X_{t-j}]^2\}$. Differentiating with respect to the γ_k gives

$$E(T_t X_{t-k}) = E(T_t T_{t-k}) = \sum_j \gamma_j E(X_{t-j} X_{t-k}) \qquad k = 0, \pm 1, \pm 2, \ldots.$$

Taking z transforms of both sides yields

$$g_{TT}(z) = \gamma(z) g_{XX}(z) .$$

[16] The generating functions derived above apply only if the entire past of the series X_t is known. In calculating the estimates shown in Figure 9, the filters were simply truncated at the earliest observation. Thus one estimate uses 360 observations, one uses 359, and so on.