CNA RESEARCH CONTRIBUTION NO. 130

NAVAL WARFARE ANALYSIS GROUP

CENTER FOR NAVAL ANALYSES

INTRODUCTION TO FOURIER ANALYSIS OF DATA

By M. Hinich

November 1969

Work conducted under contract N00014-68-A-0091

Enclosure (1) to
(NWG)45-70
Dated 13 March 1970

INTRODUCTION TO FOURIER ANALYSIS OF DATA

The purpose of this paper is to introduce to economists and data processors the basic elements of Fourier analysis of a time series which is a sum of deterministic components and a stationary random process. Techniques are discussed for Fourier decomposition of the deterministic components of a time series and for estimating the power spectrum of the stationary random component. The various mathematical concepts relating to Fourier analysis are presented in a basically intuitive manner. The goal of this paper is a reasonably clear exposition of a rather complicated body of material which has been developed primarily for physical science and engineering problems. Mathematical rigor and completeness have been sacrificed for the sake of intuitive understanding. Several theorems are stated and proved, but the proofs are instructive in nature and are not rigorous. The reader is required to understand the statistical and mathematical concepts of multiple regression and the linear model.

Fourier analysis techniques can be an important tool for economic time series analysis if the concepts and methods are properly understood. It is easy to misuse and misapply the techniques.

1. LINEAR MODELS AND PERIODIC FUNCTIONS OF TIME

The trigonometric functions $\cos \omega t$ and $\sin \omega t$ play a fundamental role in the analysis of periodic functions of time and stationary random time series. The parameter ω is called the angular frequency (figure 1). A function f(t) is called periodic with period T if for every t,

$$f(t + T) = f(t) \ . \qquad (1)$$

It is easy to check that $\cos \omega t$ and $\sin \omega t$ are periodic with period $2\pi / \omega$. If t is measured in discrete units of time, the unit of ω is radians per time unit, e.g., if $t = n\delta$ where n is an integer and $\delta = 1$ second, the unit of ω is rad/sec. For $t = n\delta$ where δ is a fixed observation interval (such as one second or one month), Fourier's classic theorem proves that if f(t) is periodic with period T, f(t) can be written as a trigonometric polynominal as follows: Let $\omega_k = \frac{2\pi k}{T}$. Then

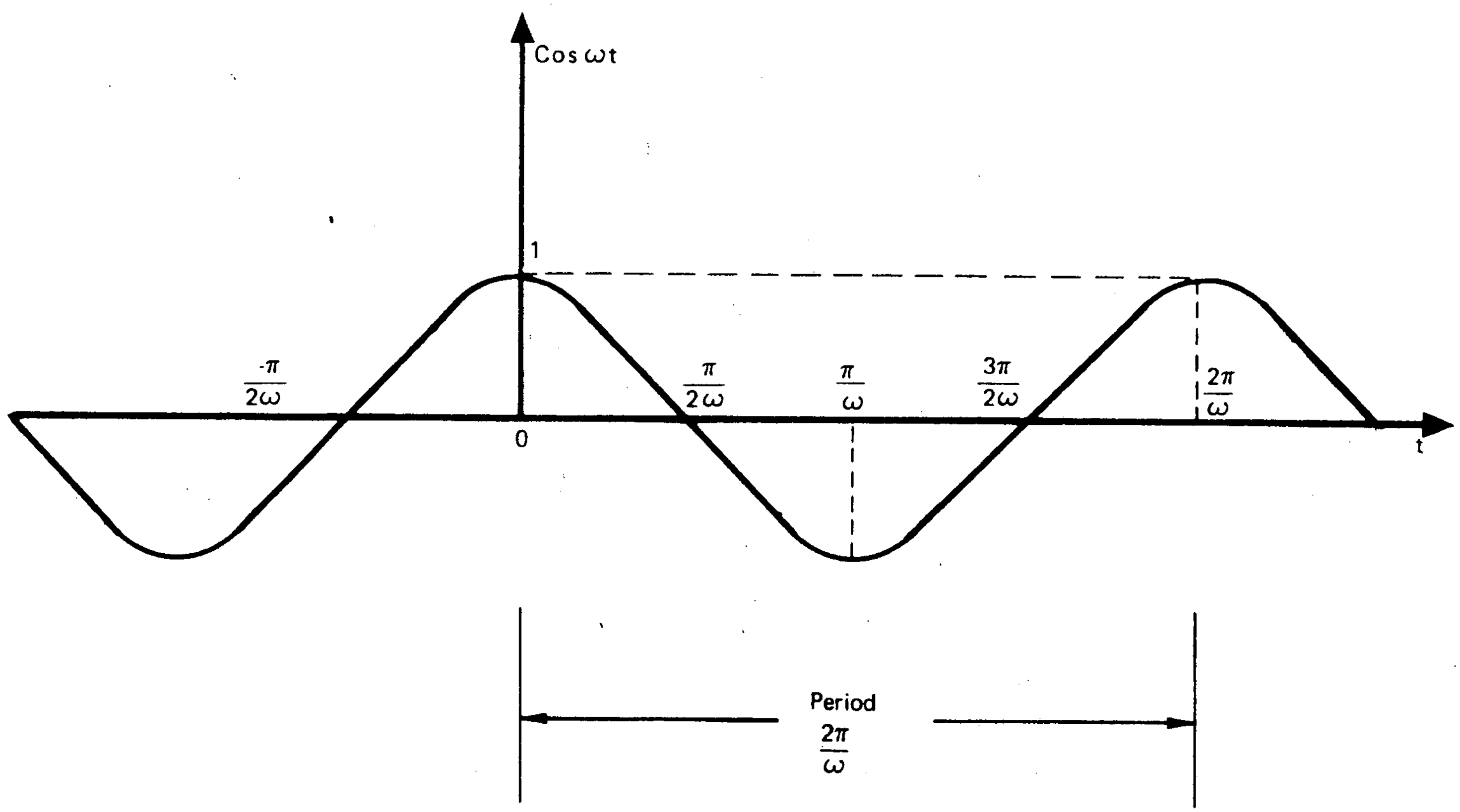

FIG. 1: COSINE FUNCTION OF PERIOD $\frac{2\pi}{\omega}$

$$f(t) = a(0) + 2 \sum_{k=1}^{(T-1)/2\delta} \left[a(\omega_k) \cos \omega_k t + b(\omega_k) \sin \omega_k t \right] \tag{2a}$$

If T/δ is odd, and for even T/δ

$$f(t) = a(0) + 2 \sum_{k=1}^{T/2\delta - 1} \left[a(\omega_k) \cos \omega_k t + b(\omega_k) \sin \omega_k t \right] + a(\omega_{T/2\delta}) (-1)^t \tag{2b}$$

The coefficients $a(\omega_k)$ and $b(\omega_k)$ are given by

$$a(\omega_k) = \frac{\delta}{T} \sum_{t=0}^{T-\delta} f(t) \cos \omega_k t$$

$$b(\omega_k) = \frac{\delta}{T} \sum_{t=0}^{T-\delta} f(t) \sin \omega_k t \tag{3}$$

For notational simplicity let us select the time unit so that the observation interval δ equals one time unit, e.g., $\delta = 1$ month for monthly observed data.

The length of the day is a periodic function of time with a period of 12 months, or 365 days depending on the time unit chosen. The mean value of almost any economic time series has a yearly periodicity called a seasonal. However, there may be longer period cycles such as 17 year cycles in the series, or shorter period cycles such as a monthly cycle. Most seasonal cycles are not a simple sine or cosine function, but instead have flatter peaks and valleys than the sine or cosine. The higher frequency (shorter period) terms in equation 2, $\cos \omega_k t$ and $\sin \omega_k t$ for $k > 1$, are required to describe the seasonal fluctuations. Of course, if only quarterly observations are available, dummy variables provide an adequate model of the seasonal changes assuming that the period is 12 months and the peak and valley in the year falls in the middle of the quarter. Fourier analysis provides a more flexible method for finding periodicities in time series.

As an extreme example of a cycle which has flat peaks and valleys, consider, for $\delta = 1$ week, the following periodic function whose period is six months:

$$f(t) = \begin{cases} 1 & -6 + 26m \leq t \leq 6 + 26m \\ -1 & 7 + 26m \leq t \leq 19 + 26m \end{cases}$$

where $m = 0, \pm 1, \pm 2, \ldots$ (figure 2). Since $f(t) = f(-t)$ and $\sin \omega t = \sin \omega(-t)$ whereas $\cos \omega t = \cos \omega(-t)$, it follows from (2) that $b(\omega_k) = 0$ for each

$k = 0, \ldots, 13$. Moreover, $a(0) = 0$ since $\sum_{t=0}^{25} f(t) = 0$ and it can be shown that

$$a(\omega_k) = \begin{cases} \frac{1}{13} \left[1 - \frac{1}{\sin \pi k/26}\right] & \text{if k is odd} \\ \frac{1}{13} & \text{if k is even} \end{cases}$$

i.e., $a(\omega_1) = -0.56$, $a(\omega_2) = 0.08$, $a(\omega_3) = -0.14$,

$a(\omega_4) = 0.08$, $a(\omega_5) = -0.06$, $a(\omega_6) = 0.08$, $a(\omega_7) = -0.03 \ldots$.

In general if $f(t)$ is an even function, i.e., $f(t) = f(-t)$, then $b(\omega_k) = 0$ for all k and if $f(t)$ is an odd function, i.e., $f(t) = -f(-t)$, then $a(\omega_k) = 0$ for all k.

By very elementary use of complex variable theory, $f(t)$ can be expressed as a sum of trigonometric functions in a form which is algebraically simpler and computationally more efficient than (2). The following definition of $e^{i\phi}$ is the key element in the discussion:

Letting $i = \sqrt{-1}$, for any angle ϕ

$$\exp(i\phi) = \cos\phi + i \sin\phi \tag{4}$$

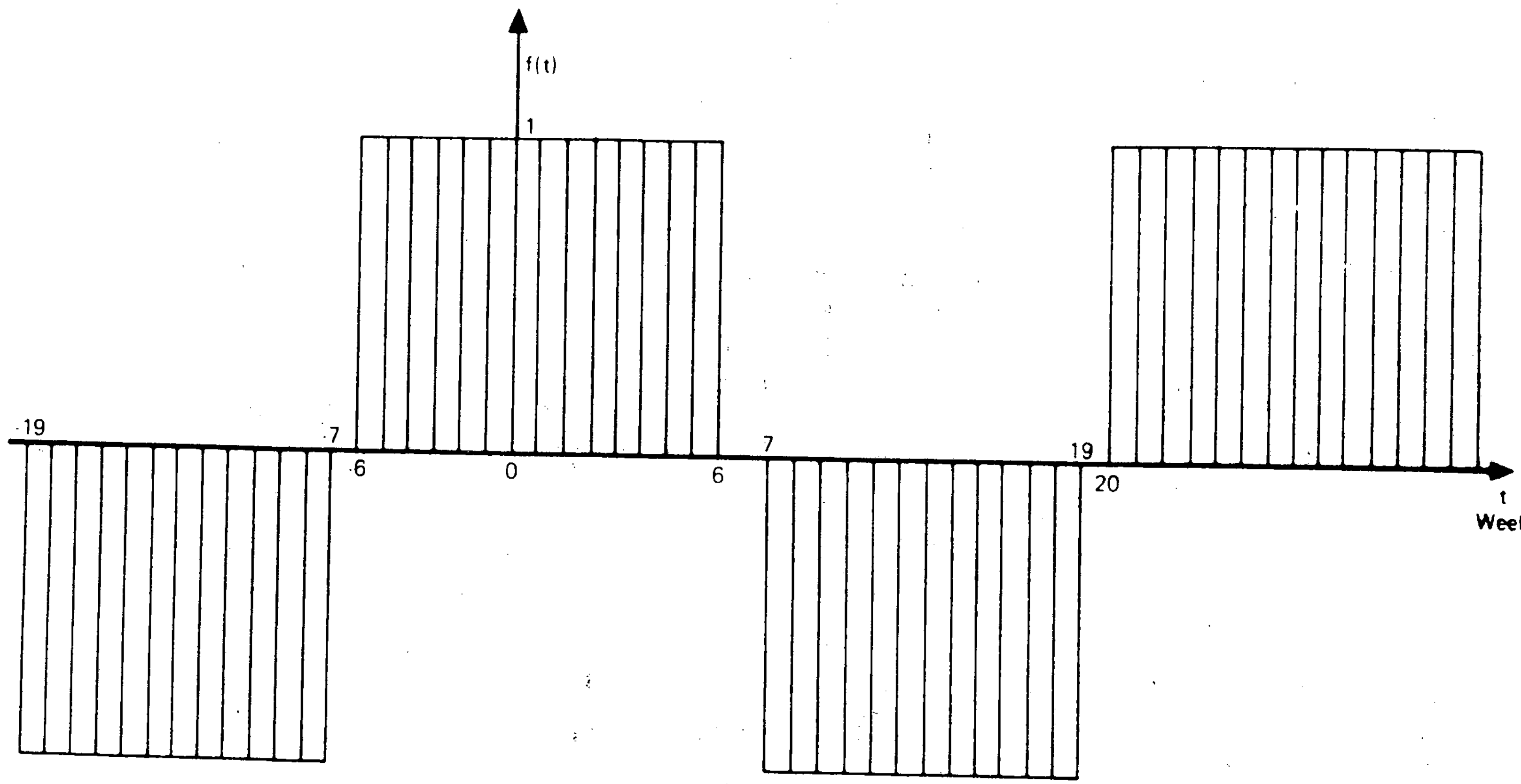

FIG. 2: EXAMPLE OF NON-SINUSOIDAL PERIODIC FUNCTION

It follows from (4) that for any integer t, $\exp(i2\pi t) - 1$. Now for any ω, define the complex variable

$$A(\omega) = \frac{1}{T} \sum_{t=0}^{T-1} f(t) \exp(i\omega t) . \tag{5}$$

Since $\exp(i\omega t + i2\pi t) = \exp(i\omega t)$ for any integer t, we have $A(\omega + 2\pi) = A(\omega)$ for any ω, i.e. $A(\omega)$ is a periodic function of ω with period 2π. Moreover $A(-\omega) = \overline{A(\omega)}$, where the bar denotes the complex conjugate, since

$$\exp(-i\omega t) = \overline{\exp(i\omega t)}$$

Thus for each $\omega_k = 2\pi \ k/T$,

$$A(\omega_{T-k}) = \overline{A(\omega_k)} \tag{6}$$

since $\omega_{T-k} = -\omega_k + 2\pi$. If T is even, $A(\omega_{T/2}) = \overline{A(\omega_{T/2})}$ from

and thus $A(\omega_{T/2})$ is real.

The following theorem essentially proves (2).

<u>THEOREM 1</u> $\quad f(t) = \sum_{k=0}^{T-1} A(\omega_k) \exp(-i\omega_k t)$

and thus from (2), (3), and (6) with $b(0) = b(\pi) = 0$

$$A(\omega_k) = a(\omega_k) + ib(\omega_k) \qquad k = 0, 1, \ldots, [T/2] \tag{7}$$

$$A(\omega_{T-k}) = a(\omega_k) - ib(\omega_k)$$

where $[T/2]$ denotes the closest integer less than or equal to $T/2$.

PROOF: For $z \neq 1$, $\sum_{k=0}^{T-1} z^k = \frac{1-z^T}{1-z}$. Setting $z = \exp(i2\pi m/T)$ for integer m, we have

$$\sum_{k=0}^{T-1} \exp\left(i2\pi \frac{m}{T} k\right) = 0 \qquad m \neq 0 \tag{8}$$

From (5),

$$\sum_{k=0}^{T-1} A(\omega_k) \exp(-i\omega_k t) = \frac{1}{T} \sum_{k=0}^{T-1} \sum_{s=1}^{T-1} f(s) \exp\left(i\omega_k(s-t)\right)$$

$$= \frac{1}{T} \sum_{s=0}^{T-1} \sum_{k=0}^{T-1} f(s) \exp\left(i2\pi \frac{s-t}{T} k\right)$$

$$= f(t)$$

since the sum over k is zero for $s \neq t$ by (8), and for s=t the sum is T.

Let us define $A_k = \sqrt{a^2(\omega_k) + b^2(\omega_k)}$ and $\theta_k = \tan^{-1} \frac{b(\omega_k)}{a(\omega_k)}$.

Thus $A(\omega_k) = A_k \exp(i\theta_k)$. From the theorem,

(Figure 3)

$$f(t) = \sum_{k=0}^{T-1} A_k \exp(-i(\omega_k t - \theta_k)) \tag{9a}$$

or alternatively due to the symmetry given by (6)

$$f(t) = a(0) + 2 \sum_{k=1}^{T/2-1} A_k \cos(\omega_k t - \theta_k) + a(\pi)(-1)^t \tag{9b}$$

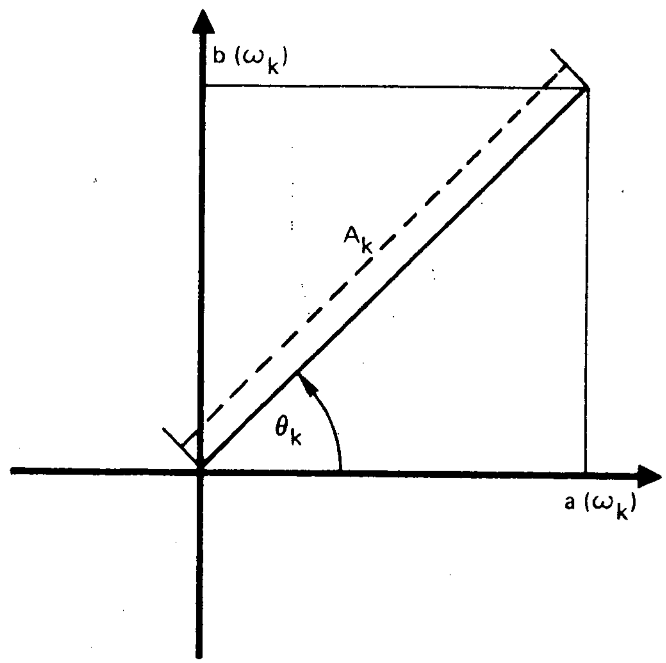

FIG. 3: AMPLITUDE AND PHASE OF THE kth FOURIER COEFFICIENT

For each frequency $\omega_k = 2\pi k/T$, A_k is called the amplitude and θ_k is called the phase. The frequency $\omega_1 = 2\pi/T$ is called the fundamental since the period of $\cos(\omega_1 t - \theta_1)$ is T, the period of f(t). Higher frequencies ω_k are integer multiples of ω_1 and are called the higher harmonies of the fundamental frequency. Given the observation interval $\delta = 1$ unit, the highest harmonic possible in the trigonometric sum representation of f(t) is π rad/unit.

Now consider the linear stochastic model

$$Y(t) = f(t) + \epsilon(t) \qquad t = 0, 1, \ldots, n-1 \tag{10}$$

where $\epsilon(t)$ is a stochastic disturbance with the usually assumed properties, i.e., $E\epsilon(t) = 0$, $E\epsilon^2(t) = \sigma^2$ for all t, and $E\epsilon(t)\epsilon(t') = 0$ for $t \neq t'$. From (2) it is clear that the independent variables in the model are $\cos\omega_k t$ and $\sin\omega_k t$ for $k = 0, 1, \ldots, [T/2]$. If n, the number of observation of Y(t), is an integral multiple of T, the independent variables are orthogonal, i.e.

$\sum_{t=0}^{n-1} \cos\omega_j t \cos\omega_k t, \sum_{t=0}^{n-1} \sin\omega_j t \sin\omega_k t = 0$ for $j \neq k$, and

$\sum_{t=0}^{n-1} \sin\omega_j t \cos\omega_k t = 0$ for all j, k. By solving the set of normal equations for the linear model, we obtain the following results.

THEOREM 2. Consider the estimator

$$\hat{A}(\omega) = \frac{1}{n} \sum_{t=0}^{n-1} Y(t) \exp(i\omega t)$$

For n=mT where m is an integer, the least-squares estimators of the coefficients $a(\omega_k)$ and $b(\omega_k)$ for $\omega_k = 0, 2\pi/T, \ldots, 2\pi[T/2]/T$ are

$$\hat{a}(\omega_k) = \text{Re}[\hat{A}(\omega_k)] = \frac{1}{n} \sum_{t=0}^{n-1} Y(t) \cos\omega_k t \tag{11}$$

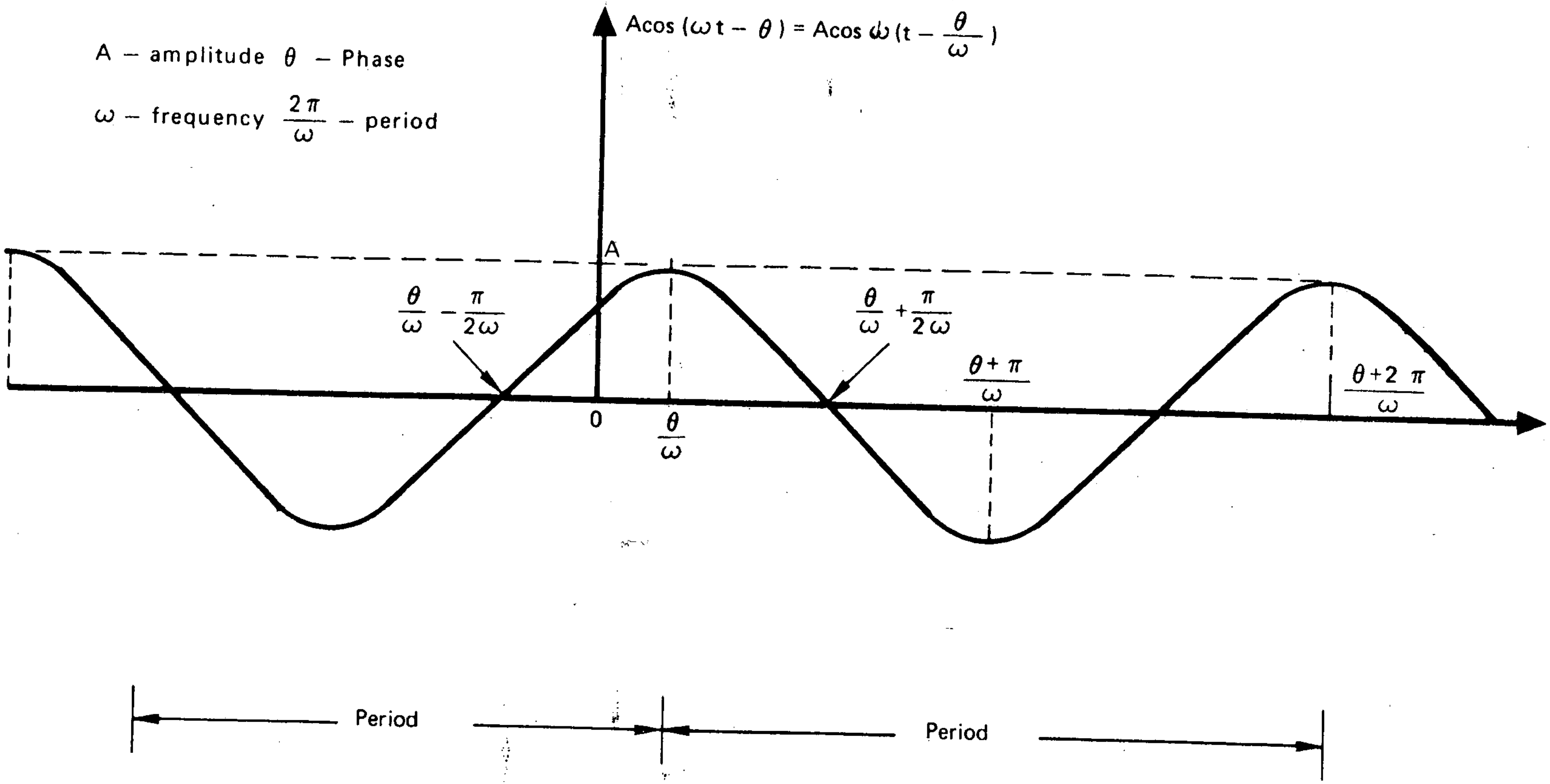

FIG. 4: SINUSOIDAL COMPONENT OF FREQUENCY ω

$$\hat{b}(\omega_k) = \text{Im}\left[\hat{A}(\omega_k)\right] = \frac{1}{n}\sum_{t=0}^{n-1} Y(t)\sin\omega_k t$$

Since these estimators are least-squares, they are unbiased, and thus $E\,\hat{A}(\omega_k) = A(\omega_k)$ since $\hat{A}(\omega) = \hat{a}(\omega) + i\hat{b}(\omega)$. The variances and covariances of $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ are

$$V\left[\hat{a}(\omega_k)\right] = V\left[\hat{b}(\omega_k)\right] = \frac{\sigma^2}{2n} \tag{12}$$

except that since $\hat{b}(0) \equiv 0$, its variance is zero, and

$$\text{COV}\left[\hat{a}(\omega_k),\ \hat{b}(\omega_k)\right] = 0$$

For $j \neq k$,

$$\text{COV}\left[\hat{a}(\omega_j),\ \hat{a}(\omega_k)\right] = \text{COV}\left[\hat{b}(\omega_j),\ \hat{b}(\omega_k)\right] = 0$$

Letting $\hat{f}(t)$ denote $f(t)$ with the a_k and b_k replaced by their estimators, the estimator of the variance σ^2 is just

$$\hat{\sigma}^2 = \frac{1}{n-T}\sum_{t=0}^{n-1}\left[Y(t) - \hat{f}(t)\right]^2 \tag{13}$$

It is not necessary to compute $\cos\omega_k t$ and $\sin\omega_k t$ for each k in order to compute $A(\omega_k)$. There exist several versions of an algorithm, called by Tukey and Cooley (1965) the fast Fourier transform, for fast and efficient computation of $\Sigma\, Y(t)\exp(i\,\omega_k t)$. The finite discrete Fourier transform, the algorithm for its computation, and its applications are discussed by Alsop and Nowroozi (1966), Bingham et al. (1967), Gentleman and Sande (1966), Hinich and Clay (1968), and Welch (1961), among others. The algorithm makes feasible the computation of $\hat{A}(\omega_k)$ for large n.

If n, the number of observations of Y, is not an integer multiple of the period T, the model is not orthogonal and the least-squares estimates are in general correlated across frequencies and between the sine and cosine variables for each frequency. As long as $n > T$, one can orthogonalize the model by merely dropping observations to make the sample size be an integer multiple of T. If the period is unknown, however, the situation is no longer simple. An operational method for detecting periodic components in an observed time series will be presented in Section 4. However, it is instructive to indicate how the estimator $\hat{A}(\omega)$, calculated at frequencies $\omega_j = 2\pi j/n$ for $j=0, \ldots, [n/2]$, can be used to estimate T.

For simplicity, assume that the function f(t) has no harmonics, i.e. f(t) is the sinusoidal periodic function

$$f(t) = 2a\cos\omega_o t + 2b\sin\omega_o t \qquad \omega_o = \frac{2\pi}{T}$$

Suppose that n is <u>not</u> an integer multiple of T. Let $\hat{a}(\omega_j)$ and $\hat{b}(\omega_j)$ be defined as in (11), but with $\omega_j = 2\pi j/n$ for $j=0, \ldots, [n/2]$. Then the expected value of $\hat{a}(\omega_j)$ is

$$E\ \hat{a}(\omega_j) = \frac{2a}{n}\sum_{t=0}^{n-1}\cos\omega_o t\ \cos\omega_j t + \frac{2b}{n}\sum_{t=0}^{n-1}\sin\omega_o t\ \cos\omega_j t$$

From Chapter 2 of Courant and Hilbert, Vol. 1 (1953) it can be shown that for sufficiently large n, there exists an integer sequence $\{j^*(n)\}$ such that $|\omega_{j^*} - \omega_o| < n^{-1}$ and

$$\frac{2}{n}\sum_{t=0}^{n-1}\cos\omega_o t\ \cos\omega_j t = \begin{cases} 1 + O\left(\frac{1}{n^2}\right) & \text{if } j=j^* \\ O\left(\frac{1}{n|\omega_j-\omega_o|}\right) & \text{if } j\neq j^* \end{cases} \tag{14}$$

and

$$\frac{1}{n}\sum_{t=0}^{n-1}\sin\omega_o t\ \cos\omega_j t = O\left(\frac{1}{n|\omega_j-\omega_o|}\right) \tag{15}$$

where $O(x^{-1})$ denotes that $\lim_{x\to\infty} x\,O(x^{-1}) = c$, a constant.

Thus for sufficiently large n, there exists a $j^*(n)$ such that $\omega_{j^*} = \omega_o + O(n^{-1})$ and

$$E\ \hat{a}(\omega_j) = \begin{cases} a + O(n^{-2}) & \text{if } j=j^* \\ O(n^{-1}) & \text{if } j \neq j^* \end{cases}$$

Since the variances of these estimators is $\sigma^2/2n$ and the variance of $\hat{\sigma}^2$ is of the order n^{-1}, the precision of the estimates is very high for large n. Thus the ω_j whose amplitude is significantly non-zero will be close to the true ω_o. If f(t) is a periodic function with harmonies, as $n \to \infty$ the fundamental frequency and the higher harmonics can be estimated with a probability arbitrarily near one, i.e., consistently. Moreover the amplitudes and phases of the sinusoidal components of f(t) can be estimated consistently. Note however that although the estimators $\hat{a}$ and $\hat{b}$, the natural estimators of the amplitude and phase, $\sqrt{\hat{a}^2(\omega) + \hat{b}^2(\omega)}$ and $\tan^{-1}(\hat{b}(\omega)/\hat{a}(\omega))$, are non-linear, the asymptotic properties of these estimators can easily be derived using standard large-sample techniques (Walker 1968).

2. SERIALLY CORRELATED DISTURBANCES

The problem of detecting hidden periodicities in data is relatively straightforward for the case of uncorrelated disturbances. If the disturbances are serially correlated, the variance of the ordinary least-squares estimator of $A(\omega_k)$ will depend on the frequency ω_k. In order to simplify the exposition, let us assume that the time series of disturbances is a jth order stationary Gaussian Markov process. To be explicit, consider the process $\{Y(t)\}$,

$$Y(t) = f(t) + u(t) \tag{17}$$

where the disturbance terms u(t) satisfy the stochastic difference equation

$$\sum_{s=o}^{\alpha} h(s)\, u(t-s) = \varepsilon(t) \tag{18}$$

with $h(0), \ldots, h(\alpha)$ as fixed constants, and for any integers t and τ, $\{\varepsilon(t), \ldots, \varepsilon(t+\tau-1)\}$ has the τ-variate normal distribution $N(0, \sigma^2 I)$.

The process $\{u(t)\}$ is called a αth order stationary normal or Gaussian Markov process. For example, consider the first order process with $h(0) = 1$ and $h(1) = \gamma \neq \pm 1$, i.e.

$$u(t) - \gamma\, u(t-1) = \varepsilon(t) \tag{19}$$

It is easy to check that the process

$$u(t) = \sum_{s=0}^{\infty} \gamma^{s} \varepsilon(t-s) \qquad (20a)$$

satisfies (19), provided that $|\gamma| < 1$. If $|\gamma| > 1$,

$$u(t) = \sum_{s=1}^{\infty} \gamma^{-s} \varepsilon(t+s) \qquad (20b)$$

satisfies (19). The process is 1st order Markov since

$$E\left[u(t) \mid u(s),\ s < t\right] = \gamma\, u(t-1)$$

Further $\{u(t), \ldots, u(t+\tau-1)\}$ has the τ-variate normal distribution $N(0, \Sigma)$ where Σ is the covariance matrix whose j, kth element is

$$\sigma_{j,k} = \sigma^2 \cdot \frac{\gamma^{|j-k|}}{1-\gamma^2} \qquad (21)$$

From (21) it is clear that as $\gamma \to \pm 1$, the variance of the process goes to infinity. For an αth order process, the variance will be infinity if any one of the α roots of the polynomial

$\sum_{s=0}^{\alpha} h(s) z^{s}$ is on the unit circle $z = e^{i\omega}$ in the complex plane.

Let us restrict the $h(s)$'s such that $\sum_{t=0}^{\alpha} h(t) \exp(i\omega t) \neq 0$ for all ω.

Not all stationary Gaussian random processes are Markovian. A process $\{u(t)\}$ is called stationary if the joint distribution of $\{u(t_1), \ldots, u(t_n)\}$ is the same as the joint distribution of $\{u(t_1+\tau), \ldots, u(t_n+\tau)\}$ for all τ, n, and points $t_1, \ldots, t_n$. Thus, the mean value $Eu(t)$ and the τth covariance $E\left[u(t+\tau)u(t)\right]$ do not depend on t if $\{u(t)\}$ is stationary. Unless otherwise specified we will assume that $Eu(t) = 0$.

The assumption of the stationarity for a process does not specify the form of the joint distribution of the $u(t)$ for various t's. A process is called Gaussian if the joint distribution of $\{u(t_1), \ldots, u(t_n)\}$ is normal for all n and $t_1, \ldots,$

t_n. Thus, if a process is both stationary and Gaussian, $\{u(t_1), \ldots, u(t_n)\}$ has the multivariate normal distribution $N(0, \Sigma)$ where the j, kth element of Σ, the covariance matrix, has the form

$$\sigma_{jk} = \sigma(|t_j - t_k|)$$

i.e., the elements on the diagonal or any off-diagonal of Σ are equal.

Now let us recall the estimator

$$\hat{A}(\omega) = \frac{1}{n}\sum_{t=0}^{n-1} Y(t) e^{i\omega t} \qquad (22)$$

from Theorem 2. Since $Eu(t) = 0$ for all t, for n=mT the estimators $\hat{a}(\omega_k) = \mathrm{Re}\,[\hat{A}(\omega_k)]$, the real part of $\hat{A}$, and $\hat{b}(\omega_k) = \mathrm{Im}\,[\hat{A}(\omega_k)]$, the imaginary part of $\hat{A}$, are the ordinary least-squares estimators of $a(\omega_k)$ and $b(\omega_k)$ respectively. However, these estimators are not minimum variance although they are unbiased. The generalized least-squares estimators, which are a function of the covariances of the disturbances, have the minimum variances of the ordinary l.s. estimators approaches the minimum variances as $n \to \infty$. The linear model in question is the one defined by (17) for $t = 0, \ldots, n-1$.

THEOREM 3. $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ are asymptotically minimum variance linear unbiased estimators of $a(\omega_k)$ and $b(\omega_k)$ respectively as $n \to \infty$. For large n these estimators are approximately independent normal random variables with means $a(\omega_k)$ and $b(\omega_k)$ and identical variances

$$V[\hat{a}(\omega_k)] = V[\hat{b}(\omega_k)] = \frac{\sigma^2}{2n|H(\omega_k)|^2}$$

where

$$H(\omega) = \sum_{t=0}^{\alpha} h(t) e^{i\omega t} \qquad (23)$$

$$|H(\omega)|^2 = [\mathrm{Re}\,H(\omega)]^2 + [\mathrm{Im}\,H(\omega)]^2$$

for $k = 0$ and $k = T/2$ where T is even

$$V[\hat{a}(0)] = \frac{\sigma^2}{n|H(0)|^2}, \qquad V[\hat{a}(\pi)] = \frac{\sigma^2}{n|H(\pi)|^2}$$

Moreover the estimates at different frequencies are asymptotically independent.

Proof: Consider the process $\{Z(t)\}$ where

$$Z(t) = \sum_{s=0}^{\alpha} h(s)\ Y(t-s)$$

From (17) and (18)

$$Z(t) = \sum_{s=0}^{\alpha} h(s)\ f(t-s) + \epsilon(t) \tag{24}$$

The function $\sum_{s} h(s)\ f(t-s)$ is also periodic with period T.

The kth complex Fourier coefficient of this function is $H(\omega_k)\ A(\omega_k)$ since from (5) and (23)

$$\frac{1}{T}\sum_{t=0}^{T-1}\left[\sum_{s=0}^{\alpha} h(s)\, f(t-s)\right] e^{i\omega t} = \sum_{s=0}^{\alpha} h(s)\, e^{i\omega s} \sum_{t=0}^{T-1} f(t-s)\, e^{i\omega(t-s)}$$

$$= H(\omega)\ A(\omega)$$

The ordinary l. s. estimators of the coefficients $H(\omega_k)\ A(\omega_k)$ are

$$\hat{B}(\omega_k) = \frac{1}{n}\sum_{t=0}^{n-1} Z(t)\, e^{i\omega_k t}$$

which is minimum variance since the disturbances in equation (24) are spherical. We will now show that for large n,

$\hat{A}(\omega_k) \approx \dfrac{\hat{B}(\omega_k)}{H(\omega_k)}$ and thus $\hat{A}(\omega_k)$ is asymptotically the minimum variance estimator of $A(\omega_k)$ in the sense that the real and imaginary parts of $\hat{A}(\omega_k)$ are asymptotically the minimum variance estimators of the real and imaginary parts of $A(\omega_k)$.

For the sake of exposition, let $\{u(t)\}$ be first order Markov with $h(0) = 1$ and $h(1) = -\gamma$, $|\gamma| < 1$, as before. The proof for $\alpha > 1$ is a straightforward extension of the proof for $\alpha = 1$. If ω is an integer multiple of $2\pi/n$,

$$\hat{B}(\omega) = \frac{1}{n}\sum_{t=0}^{n-1} \left[Y(t) - \gamma Y(t-1)\right]e^{i\omega t} \tag{25}$$

$$= \hat{A}(\omega) - \gamma e^{i\omega}\,\hat{A}(\omega) - \frac{\gamma}{n}\left[Y(-1) - Y(n-1)\right]$$

$$= H(\omega)\,\hat{A}(\omega) + \frac{\gamma}{n}\left[Y(n-1) - Y(-1)\right]$$

since from (23),

$$H(\omega) = 1 - \gamma e^{i\omega}$$

The variance of $Y(n-1) - Y(-1)$ is

$$E\left[u(n-1) - u(-1)\right]^2 = \frac{1-\gamma^n}{1-\gamma^2}\,2\sigma^2 \tag{26}$$

from (21). The covariance between $Y(n-1) - Y(-1)$ and $n\hat{A}(\omega)$ is less than $\frac{\sigma^2}{1-\gamma^2}$ since

$$\left|E\left[u(n-1) - u(-1)\right] n\hat{A}(\omega)\right| = \left|\sum_{t=0}^{n-1} \sigma^2\,\frac{\gamma^{n-1-t} - \gamma^{t+1}}{1-\gamma^2}\,e^{i\omega t}\right| \tag{27}$$

$$\leq \frac{1-\gamma^n}{1-\gamma^2}\,\sigma^2$$

The variance of the real and imaginary parts of $\hat{B}(\omega)$ is $\frac{\sigma^2}{2n}$ from Theorem 2. Dividing both sides of equation (25) by $H(\omega)$, it follows from (26) and (27) that

$$V\left[\hat{a}(\omega)\right] = \frac{\sigma^2}{2n\left|H(\omega)\right|^2} + O\left(\frac{\sigma^2}{n^2\left|H(\omega)\right|^2(1-\gamma^2)}\right)$$

and similarly for $\hat{b}(\omega)$. If $n \gg (1-\gamma)^{-1/2}$ the variances of $\hat{a}(\omega)$ and $\hat{b}(\omega)$ are equal to $\frac{\sigma^2}{2n\left|H(\omega)\right|^2} + O(\frac{1}{n^2})$. The variances are minimum since $\hat{B}$ is the minimum variance estimator of HA and $\hat{A} \approx \frac{\hat{B}}{H}$.

Since the real and imaginary parts of $\hat{B}$ are uncorrelated and have equal variances, the real and imaginary parts of $\frac{\hat{B}}{H}$ are uncorrelated. Thus, from (25) and (26), the covariance between the real and imaginary parts of $\hat{A}$ is of the order $O(\frac{1}{n^2})$. Similarly, the covariance between Fourier coefficients of different frequencies is $O(\frac{1}{n^2})$.

In most applications, the function $H(\omega)$ is unknown, as are the parameters $h(1), \ldots, h(\alpha)$ of the Markov disturbance. Note that similar to $A(\omega)$, $H(\omega+2\pi) = H(\omega)$ and $H(-\omega) = \overline{H(\omega)}$. In the next section we will discuss the problem of estimating $\sigma^2 |H(\omega)|^{-2}$.

3. POWER SPECTRA

The variances of the estimators $\hat{a}(2\pi k/T)$ and $\hat{b}(2\pi k/T)$ depend on $\sigma^2 \, H|(2\pi k/T)|^{-2}$. It is possible to obtain a consistent and asymptotically normal estimator of $\sigma^2 |H(2\pi k/T)|^{-2}$ for each k, based upon the same sample which is used to estimate the Fourier coefficients. Note, however, that if the disturbances are serially uncorrelated with $h(0) = 1$, then $\alpha = 0$ and $H(\omega) = 1$ for all ω. In that case the estimator of the variance is the normalized sum of squared residuals.

Suppose that we observe $Y(0), \ldots, Y(n-1)$ with $n = mT$ as before. From now on let $\omega_j = 2\pi j/n$ for $j = 0, \ldots, n-1$. With this notation, the harmonic frequencies of $f(t)$ are ω_{km}, $k = 0, \ldots, T-1$. Consider, now, the estimator $\hat{A}(\omega_j)$. From (8), (17), and (22),

$$E\,\hat{A}(\omega_j) = \frac{1}{n}\sum_{t=0}^{n-1}\sum_{k=0}^{T-1} A(\omega_{km}) \exp(-i\omega_{km}t) \exp(i\omega_j t)$$

$$= \frac{1}{n}\sum_{k=0}^{T-1} A(\omega_{km}) \sum_{t=0}^{n-1} \exp\left(i2\pi \frac{j-km}{n}\right)$$

$$= \begin{cases} A(\omega_{km}) & \text{if } j=km \\ 0 & \text{if } j \neq km \end{cases}$$

From Theorem 3 with $m=1$ and $n=T$, the variances of the real and imaginary parts of $\hat{A}$ are

$$V[\hat{a}(\omega_j)] = V[\hat{b}(\omega_j)] = \frac{\sigma^2}{2n|H(\omega_j)|^2} \tag{28}$$

except that since $\hat{b}(0) = 0$, its variance is zero. Since $\hat{a}(\omega_j)$ and $\hat{b}(\omega_j)$ are independent and normally distributed with zero means for $j \neq km$ and the above variances,

$$2n\sigma^{-2} \left|H(\omega_j)\ \hat{A}(\omega_j)\right|^2 \sim \chi_2^2 \qquad (29a)$$

for $j \neq km$, where χ_2^2 denotes a chi-squared variable with two degrees of freedom. For $j = km$ such that $k \neq 0$ or $T/2$

$$2n\sigma^{-2} \left|H(\omega_{km})\right|^2 \left|\hat{A}(\omega_{km}) - A(\omega_{km})\right|^2 \sim \chi_2^2 \qquad (29b)$$

For $\omega = 0$ or π, $n\sigma^{-2}\left|H(\omega)\right|^2 \left|\hat{a}(\omega) - a(\omega)\right|^2$ has a χ_1^2 distribution.

Let $S(\omega) = \sigma^2 \left|H(\omega)\right|^{-2}$ since $H(\omega) \neq 0$, $S(\omega)$ is defined for all ω. Note that $S(\omega + 2\pi) = S(\omega)$ for all ω, and $S(-\omega) = S(\omega)$. The following theorem presents the properties of a consistent estimator of $S(\omega_{km})$.

<u>THEOREM 4</u>. Given an even integer $M < m$ where $n = mT$, define the estimator $\hat{S}(\omega_{km})$ for $0 < \omega_{km} < \pi$ by

$$\hat{S}(\omega_{km}) = \frac{n}{M} \sum_{j=-M/2}^{M/2}{}' \left|\hat{A}(\omega_{km+j})\right|^2 \qquad (30)$$

where the prime denotes that the $j = 0$ term, $\left|\hat{A}(\omega_{km})\right|^2$, is not included in the sum.

Let m and $M \to \infty$ such that $M/m \to 0$. Then in the limit

$$\sqrt{M}\left[\hat{S}(2\pi k/T) - S(2\pi k/T)\right] \sim N(0, S^2(2\pi k/T)) \qquad (31)$$

where $S(\omega) = \sigma^2 \left|H(\omega)\right|^{-2}$. For $\omega_{km} = 0$ or π, (31) holds with M replaced by $M/4$ and

$$\hat{S}(0) = \frac{2n}{M} \sum_{j=1}^{M/2} \left|\hat{A}(\omega_j)\right|^2, \quad \hat{S}(\pi) = \frac{2n}{M} \sum_{j=1}^{M/2} \left|\hat{A}(\omega_{\frac{n}{2}-j})\right|^2$$

Proof: Since $E(\chi_2^2) = 2$, $V(\chi_2^2) = 4$, and since $\hat{A}(\omega_j)$ and $\hat{A}(\omega_\ell)$ are independent for $j \neq \ell$, from (29)

$$E\hat{S}(\omega_{km}) = \frac{\sigma^2}{M} \sum_{j=-M/2}^{M/2}{}' \left|H(\omega_{km} + \omega_j)\right|^{-2} \qquad (32)$$

since $\omega_{km+j} = \omega_{km} + \omega_j$, and

$$V[\hat{S}(\omega_{km})] = \frac{\sigma^4}{M^2} \sum_{j=-M/2}^{M/2}{}' |H(\omega_{km}) + \omega_j)|^{-4} \tag{33}$$

Since $|H(\omega)|^{-1}$ is a bounded continuous function for $-\pi \leq \omega \leq \pi$, as both $M \to \infty$ and $M/n \to 0$ we have

$$\lim_{\substack{M \to \infty \\ M/n \to 0}} E[\hat{S}(\omega_{km})] = \lim \frac{2}{2\pi M} \int_{-\pi M/n}^{\pi M/n} S(\omega_{km} + \omega)\, d\omega$$

$$= S(2\pi k/T)$$

By a similar argument for (33), the variance of $\hat{S}$ has the following property

$$\lim_{M \to \infty} M\, V[\hat{S}(\omega_{km})] = S^2\,(2\pi k/T)$$

The asymptotic normality of $\hat{S}$ follows from the central limit theorem.

The function $S(\omega) = \sigma^2 |H(\omega)|^{-2}$ is called the power spectrum of the process $\{u(t)\}$. In order to better understand the concept of a power spectrum, set $f(t) = 0$, i.e., let $Y(t) = u(t)$. From (28), $1/2\, S(\omega_j)$ is the variance of the real and imaginary parts of $\sqrt{n}\, \hat{A}(\omega_j)$. From Theorem 1 we have

$$u(t) = \sum_{j=0}^{n-1} \hat{A}(\omega_j) e^{-i\omega_j t} \tag{34}$$

where $\hat{A}(\omega_{n-j}) = \overline{\hat{A}(\omega_j)}$ and $\omega_j = 2\pi j/n$. Thus by the independence of $\hat{A}(\omega_j)$ and $\hat{A}(\omega_{j'})$, $j \neq j'$, from (34) we have for all t that

$$Eu^2(t) = \frac{1}{n} S(0) + \frac{2}{n} \sum_{j=1}^{n/2} S(\omega_j) \tag{35}$$

Letting $n \to \infty$ in (35), the variance of the disturbances is

$$Eu^2(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega)\, d\omega \tag{36}$$

by the definition of an integral as a limit of a sum. Thus, the power spectrum of a process indicates how its total variance is distributed over various frequency components of its Fourier Stieltjes integral representation.

$$u(t) = \int_{-\pi}^{\pi} e^{i\omega t} dU(\omega)$$

where for each point ω

$$dU(\omega) = \lim_{n \to \infty} \hat{A}(\omega_{j*}) \qquad \omega_{j*} = \frac{2\pi j^*(n)}{n}$$

and $\{j^*(n)\}$ is a sequence of integers such that $\omega_{j*} \to \omega$ as $n \to \infty$. Thus

$$E|dU(\omega)|^2 = \frac{1}{2\pi} S(\omega)\, d\omega$$

since $nE|\hat{A}(\omega_{j*})|^2 = S(\omega)$ by (28), and $\frac{2\pi}{n} \approx d\omega$.

If the disturbances are serially uncorrelated with $h(0) = 1$, $S(\omega) = \sigma^2$ for all ω. A process whose power spectrum is constant for all ω is called white noise.

Assuming that $Y(t) = u(t)$, it is possible to obtain a consistent estimator of $S(\omega)$ for any ω. Define

$$\hat{S}(\omega) = \frac{n}{M+1} \sum_{j=-M/2}^{M/2} |\hat{A}(\omega_{j*} + \omega_j)|^2$$

Letting M and $n \to \infty$ such that $M/n \to 0$, it can be shown from the proof of Theorem 4 that in the limit

$$\sqrt{M}\,[\hat{S}(\omega) - S(\omega)] \sim N(0, S^2(\omega)).$$

The fast Fourier transform algorithm has made practical the estimation of spectra by direct use of the computed Fourier coefficients. Until this algorithm was developed, the computation of the Fourier coefficients for large n required a great deal of computer time in order to obtain numerically accurate values. The spectral estimation method which has been widely used until now involves the computation of lagged products. Let us briefly discuss this standard method.

First let us present another expression for the power spectrum of $\{u(t)\}$. Since the process is stationary and $Eu(t) = 0$, the τth covariance

$$\hat{\rho}(\tau) = E[u(t+\tau)\, u(t)]$$

is independent of t. The natural unbiased estimator of $\rho(\tau)$ for $\tau > 0$ is the sample covariance

$$\hat{\rho}(\tau) = \frac{1}{n-\tau} \sum_{t=0}^{n-1-\tau} u(t+\tau)\, u(t)$$

By straightforward algebra, we have

$$n\left|\hat{A}(\omega)\right|^2 = \sum_{t=-n+1}^{n-1} \left(1 - \frac{|\tau|}{n}\right) \hat{\rho}(\tau)\, e^{-i\omega\tau} \tag{37}$$

where $\hat{\rho}(-\tau) = \hat{\rho}(\tau)$. Taking the expected value of equation (37) and using a result from Courant and Hilbert (1953), it follows that

$$nE\left|\hat{A}(\omega)\right|^2 = \sum_{\tau=n+1}^{n-1} \left(1 - \frac{|\tau|}{n}\right) \rho(\tau)\, e^{-i\omega\tau} \tag{38}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{S}(\omega-\omega') \frac{\sin^2 \frac{n\omega'}{2}}{n\sin^2 \frac{\omega'}{2}}\, d\omega'$$

where

$$\tilde{S}(\omega) = \sum_{\tau=-\infty}^{\infty} \rho(\tau)\, e^{-i\omega\tau} \tag{39}$$

$$= \rho(0) + 2 \sum_{t=1}^{\infty} \rho(\tau) \cos \omega\tau$$

The function $\frac{\sin^2 n\omega/2}{n\sin^2 \omega/2}$ is called the Fejer Kernel. Its integral is one for $\omega = 0$ and zero otherwise, and the right hand side of equation (38) goes to $\tilde{S}(\omega)$. However, $nE\left|\hat{A}(\omega)\right|^2$ is equal to $S(\omega)$ for ω an integer multiple of $2\pi/n$. Thus, $\tilde{S} \equiv S$ and equation (39) becomes an alternative definition of the power spectrum of $\{u(t)\}$ which holds for non-Markov stationary process.

The variance of the estimator $\hat{\rho}(\tau)$ increases with τ since there are fewer lagged products $u(t+\tau)\, u(t)$ to average over the sample. Let us modify equation (37) by cutting off the 'tails' of the $\hat{\rho}(\tau)$ function and define the following statistic:

$$\hat{P}(\omega_j) = \sum_{\tau=-m}^{m} \hat{\rho}(\tau)\, e^{-i\omega_j\tau} \qquad \omega_j = \frac{\pi j}{m}$$

for $m \ll n$ and $j = 0, \ldots, m$. The expected value of $\hat{P}(\omega)$ is

$$E\,\hat{P}(\omega) = \sum_{\tau=-m}^{m} \rho(\tau)\, e^{-i\omega\tau}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} S(\omega-\omega')\,\frac{\sin(m+1/2)\,\omega'}{\sin\frac{\omega'}{2}}\,d\omega'$$

The function $\dfrac{\sin(m+1/2)\,\omega'}{\sin\frac{\omega'}{2}}$ is called the Dirichlet Kernel. It is similar to the Fejer Kernel in its asymptotic properties and thus

$$\lim_{m\to\infty} E\,\hat{P}(\omega) = S(\omega)$$

Barlett (1955) has shown that if $m/n \to \infty$, the variance of $\hat{P}(\omega)$ is approximately

$$V\,\hat{P}(\omega) \approx \frac{2m}{n}\; S^2(\omega) .$$

In practice m is generally between 10 percent to 20 percent of n.

The precision spectral estimates based on $\hat{P}$ can be improved by smoothing the $\hat{P}(\omega_j)$ as follows:

$$c_{-1}\,\hat{P}(\omega_{j-1}) + c_0\,\hat{P}(\omega_j) + c_1\,\hat{P}(\omega_{j+1})$$

where $c_{-1} + c_0 + c_1 = 1$. If $c_0 = 1/2$ and $c_1 = c_{-1} = 1/4$, this smoothing operation is called Hanning by Blackman and Tukey (1959).

The estimates $\hat{P}(\omega_j)$ require the computation of only m Fourier coefficients instead of $n/2$, but the computation of the $\hat{\rho}(\tau)$ is time consuming. The direct method of estimation using the $\hat{A}(\omega_j)$ is now faster than the lag product method if n is large. In addition the computation of all the Fourier coefficients provides greater flexability in the analysis of a sample of a time series.

4. HIDDEN PERIODICITIES

Let us return to the model $Y(t) = f(t) + u(t)$ where $\{u(t)\}$ is Markov. If T, the period of $f(t)$ is known, we have shown that we can obtain precise estimates of the coefficients $a(2\pi k/T)$ and $b(2\pi k/T)$ if we have a sufficiently large number of consecutive observations of $\{Y(t)\}$. Moreover we can obtain precise estimates of the power spectrum $S(\omega)$ of the disturbance process. If the spectrum is slowly varying about the fundamental and harmonic frequencies of $f(t)$, we can derive independent F-tests for the null hypothesis that $A(\omega_{km}) = 0$ for $k=0, \ldots, [(T/2)]$. To be more precise, suppose that $|H(\omega)|$ is slowly varying in the frequency band $(\omega_{km} - \frac{\pi M}{n}, \omega_{km} + \frac{\pi M}{n})$ for each k. Then for ω_j in the band, $\sigma^{-2} |H(\omega_{km+j})|^2$ is approximately equal to $\frac{1}{S(\omega_{km})}$ and from (29), $2M \hat{S}(\omega_{km})/S(\omega_{km}) \sim \chi^2_{2M}$ for $\omega_{km} \neq 0$ or π. If $\omega_{km} = 0$ or π, the degrees of freedom is M. Since $S(\omega_{km})$ and $A(\omega_{km})$ are independent,

$$\frac{n|\hat{A}(\omega_{km})|^2}{\hat{S}(\omega_{km})} \sim \begin{cases} F_{2,2M}(\lambda_k) & \text{if } \omega_{km} \neq 0, \pi \\ F_{1,M}(\lambda_o) & \text{if } \omega_{km} = 0, \pi \end{cases} \tag{40}$$

where $F_{2,2M}(\lambda_k)$ is the non-central F distribution with 2, 2M degrees-of-freedom and non-centrality parameter

$$\lambda_k = \frac{n A_k^2}{S(\omega_{km})} \tag{41}$$

$$= \frac{|H(\omega_{km})|^2 n A_k^2}{\sigma^2}$$

Thus the larger the amplitude $A_k = |A(\omega_{km})|$ of the kth sinusoidal component of $f(t)$, the greater the probability of detecting it in the observed time series. The denominator degrees-of-freedom is only M for the cases of $\omega_{km} = 0$ and π, since the sum in the definition of $\hat{S}$ is taken over only M elements.

The F statistics for different k's are independent since $\hat{A}(\omega_j)$ and $\hat{A}(\omega_\ell)$ are independent for $j \neq \ell$.

Now suppose that we do not know the period of the function $f(t)$. For sufficiently large n we still can obtain precise estimates of the Fourier coefficients of $f(t)$. Even for a moderately sized number of conservative observations of $Y(t)$, we can determine a great deal about the structure of $f(t)$, provided the power spectrum of the disturbances is not too large in the vicinity of the fundamental frequency and the first few harmonics of $f(t)$. The power spectrum will be large if σ^2, the variance of the $\{\epsilon(t)\}$ disturbances, is large, or if $H(\omega)$ is near zero for ω in the bands of interest. We will discuss the problem of $H(\omega) \approx 0$ later.

It would be helpful in the discussion of hidden periodicities to consider an example. Suppose that we have ten years of weekly data for some variable Y(t). Thus the sampling interval is $\delta = 1$ week and $n = 520$. Let $Y(t) = f(t) + u(t)$ where f(t) is a seasonal periodic function. For sake of discussion, let us assume that we have no a-priori knowledge about the true T, the period of f(t). Of course, T = 52 weeks is a natural assumption. For $\tau = 1$ week, the Fourier sum representation of the seasonal is

$$f(t) = \sum_{k=0}^{51} A\left(\frac{2\pi k}{52}\right) \exp\left(-i \frac{2\pi kt}{52}\right) \tag{42}$$

$$= a(0) + 2\sum_{k=1}^{25} A_k \cos\left(\frac{2\pi kt}{52} - \theta_k\right) + a(\pi)(-1)^t$$

where

$$A(\omega_k) = A_k e^{i\theta_k} = a(\omega_k) + ib(\omega_k)$$

$$A(\omega_{26-k}) = A_k e^{-i\theta_k} = a(\omega_k) - ib(\omega_k)$$

for $k=0,\ldots, 25$ and $\omega_k = 2\pi k/52$ and $A(\omega_{26}) = a(\pi)$. Note that the zero frequency component $A(0) = a(0)$ is just the mean of $f(t)$ over the year, ie.,

$$a(0) = \frac{1}{52}\sum_{t=0}^{51} f(t)$$

Let us assume that $a(0) = 0$.

The highest harmonic component in the representation of $f(t)$ has the angular frequency π radians/week, or 1/2 cycles per week, i.e., it has a period of 14 days. The limit for the highest harmonic is a property of the sampling interval δ. The shorter the time between successive observations, the higher the frequency possible for the highest harmonic component of the representation of $f(t)$. However, suppose that $f(t)$ really is of the form

$$f(t) = A_1 \cos\left(\frac{2\pi t}{52} - \theta_1\right) + A_D \cos\left(\frac{2\pi t}{1/7} - \theta_D\right) \tag{43}$$

i. e., $f(t)$ consists of a yearly cycle plus a daily cycle. Since $\cos 14\pi t = 1$ and $\sin 14\pi t = 0$ for any integer t, equation (43) becomes

$$f(t) = A_1 \cos\left(\frac{2\pi t}{52} - \theta_1\right) + A_D \cos \theta_D \tag{44}$$

If $Y(t)$ is regressed on the functions $\cos\omega_k t$ and $\sin\omega_k t$ and the estimators $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ computed from (44) and Theorem 3, it follows that

$$E\,\hat{a}(0) = A_D \cos\theta_D$$

Thus for small σ^2 we will make the false inference that the mean of $Y(t)$ is non-zero (unless $\theta_D = \pi$ or $\frac{3\pi}{2}$ and we will be unable to detect the presence of the daily cycle because we have sampled the process insufficiently often.

Given any frequency ω_0, $\cos\omega_0 t = \cos(\omega_0 + 2\pi m)t$ and $\sin\omega_0 t = \sin(\omega_0 + 2\pi m)t$ for any integers m and t. The functions $\cos(\omega_0 + 2\pi m)t$ and $\sin(\omega_0 + 2\pi m)t$ are <u>aliases</u> of $\cos\omega_0 t$ and $\sin\omega_0 t$ respectively. In order to resolve the ambiguity, the sampling interval δ must be less than a half cycle of the frequency component, i. e., $\delta \le \frac{\pi}{\omega_0}$. For example, suppose we have quarterly observations of a series which contains a four week cycle. This cycle has a frequency of $2\pi\,\frac{13}{4}$ radians/quarter. Thus, the alias which is detected by Fourier analysis is the cycle whose frequency is $\frac{\pi}{2}$ radians/quarter, i. e., a cycle of a year.

Returning to our original example, suppose that $A(\frac{2\pi k}{52}) = 0$ for $6 \le k \le 26$ for sake of argument. The highest harmonic has a period of 10 2/5 weeks. The periods of the lower harmonics are 26, 17 1/3, and 13 weeks. Assume that we do not know the fundamental period is 52 weeks. If we hypothesize that $T = 52$ weeks, we can use the previous results to make independent F-tests for whether or not $A(\frac{2\pi k}{52}) = 0$ for $k = 1, \ldots, 5$. However, the use of F-tests restricts our attentions to the discrete frequencies we assume in our null hypothesis. For example, if we make a mistake and guess $T = 26$ weeks, and if the amplitudes $A_1, \ldots, A_5$ are not too small or if the power spectrum of the disturbances is not too large, we will detect the higher harmonics but we will miss the detectable fundamental periodicity. Moreover, if we guess that $T = 51$ weeks, it can be shown from (16) that we will accept the hypotheses that $A(\frac{2\pi k}{51}) \neq 0$ for $k = 1, \ldots, 5$ and, thus incorrectly infer that the fundamental is 51 weeks. Since the basic problem here is one of estimation, there is an uncertainty band around the estimate of the fundamental frequency for a fixed n. It is clear that we should look at all the Fourier coefficients we can estimate from the sample.

Suppose that we have computed the $\hat{A}(\frac{2\pi j}{520})$, $j=0,\ldots, 259$, from the sample of $\{Y(t)\}$ consisting of 520 consecutive weekly observations. From (29), for each j

$$E\,|\hat{A}(\omega_j)|^2 = \begin{cases} A_k^2 + \frac{1}{520} S\ (\frac{2\pi k}{52}) & \text{if } j=10k \quad k=1,\ldots, 5 \\ \frac{1}{520}\ S(\omega_j) & \text{otherwise} \end{cases}$$

and for all j

$$V(|\hat{A}(\omega_j)|^2) = [\frac{1}{520}\ S(\omega_j)]^2$$

Thus, if $A_k^2 >> \frac{1}{520}\ S(\omega_j)$, large values of $|\hat{A}(\omega_j)|^2$ indicate the ω_j corresponding to the fundamental and harmonic periods. Moreover, we can obtain an estimate of the power spectrum $S(\omega)$ of the disturbances for the frequencies $\omega = \omega_j$, using the statistic defined in Theorem 4.

The easiest method to spot large values of $|\hat{A}(\omega_j)|^2$ is to plot it as a function of ω_j. The plot of $|\hat{A}(\omega_j)|^2$ are independent random variables whose coefficients of variation are one, the periodogram will be jagged and have many false peaks which correspond to true hidden periodicities (Bartlett (1955), Priestly (1962a), and 1962b)).

For example, suppose we divide the frequency band $(0, \pi)$ into N equal parts or bands $B_1, \ldots, B_N$, and then average $|\hat{A}(\omega_j)|^2$ over the ω_j in the bands. The plot of the N averages yields a smoothed periodogram. In order to illustrate this simple type of smoothing, let $N=20$ and thus $B_k\ (\pi\frac{k-1}{20}, \pi\frac{k}{20})$ for $k=1,\ldots, 20$. There are 13 estimated Fourier coefficients in each piece. The smoothed periodogram ordinates are

$$\hat{S}_k = \frac{1}{13}\sum_{j=1}^{13} |\hat{A}\ (2\pi\ \frac{13\,(k-1)+j}{520})|^2 \tag{45}$$

From (39), $520\hat{S}_k = \hat{S}\ (\pi\ \frac{k-1/2}{20})$ with $M=12$. In general there are $\frac{n}{2N}$ ω_j in each band and thus if $\frac{n}{2N}$ were not odd, $n\hat{S}_k$ and $\hat{S}(\pi\frac{k-1/2}{N})$ would be slightly different.

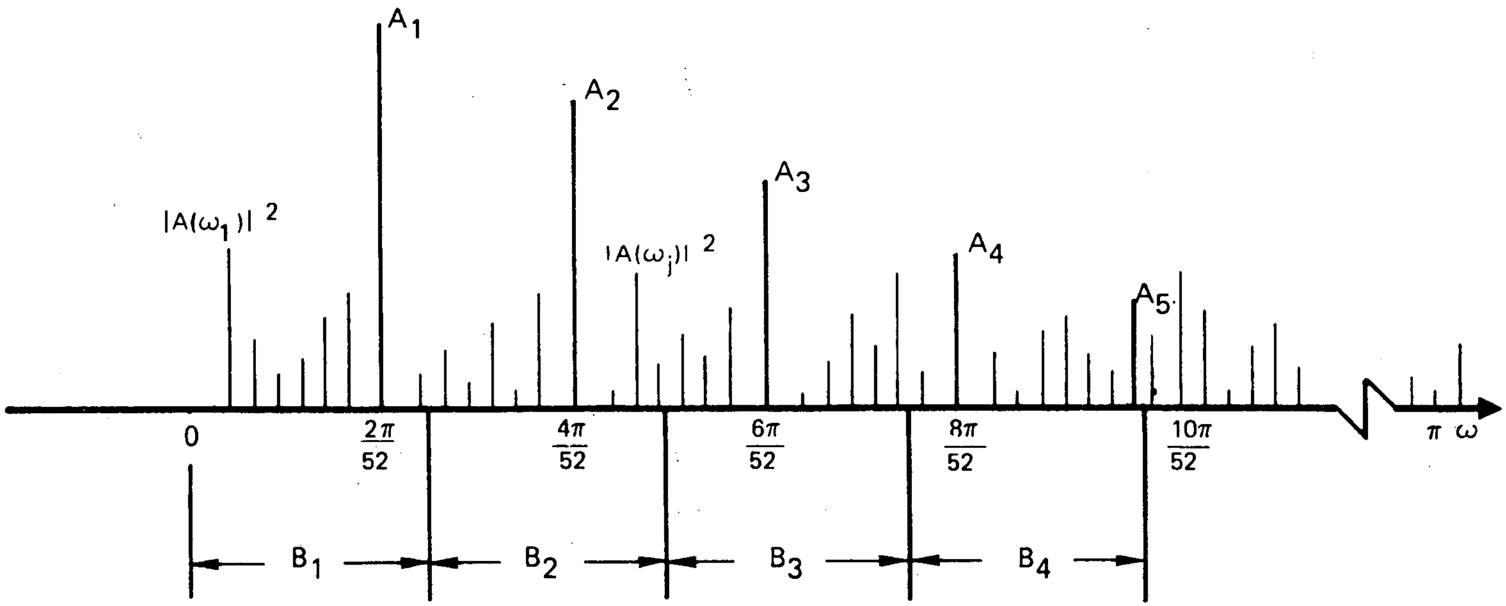

FIG. 5: PERIODOGRAM

The funadmental frequency is in B_1, the first harmonic $2\pi/26$ is in B_2, the second harmonic $2\pi/13$ is in B_3, and the third and fourth harmonics are in B_4 (figure 5). Assuming that the power spectrum $S(\omega)$ is slowly varying, from (29) and Theorem 4, we have

$$E\hat{S}_k \approx \begin{cases} \frac{1}{13} A_k^2 + \frac{1}{520} S\left(\pi \frac{k-1/2}{20}\right) & \text{if } k=1, 2, 3 \\ \frac{1}{13} (A_4^2 + A_5^2) + \frac{1}{520} S\left(\frac{7\pi}{40}\right) & k=4 \\ \frac{1}{520} S\left(\frac{k-1/2}{20}\right) & k=5, \ldots, 20 \end{cases}$$

and

$$V(\hat{S}_k) \approx \frac{1}{13} \left[\frac{1}{520} S\left(\pi \frac{k-1/2}{20}\right)\right]^2$$

Thus, the smoothing reduces the true peaks in the periodogram, but it also reduces the variance of periodogram ordinates.

It is wise to smooth using several different bandwidths since the amount of variation of the spectrum over different frequency bands is generally unknown. The analysis of a sample of a time series which is made up of deterministic and random components is not a one shot deal. One has to try several variations of the basic type of analysis. The $n/2+1$ random variables $\hat{A}(\frac{2\pi j}{n})$ for $j=0,\ldots,[n/2]$ are the fundamental elements in the analysis. However, it is often important to modify the dependent variables $Y(t)$ when trends and discontinuities are present in the series will be discussed in Section 6. First, let us discuss several misconceptions which are commonly held about Fourier analysis.

5. TRANSIENTS AND RANDOM PERIODICITIES

It is a common misconception that the only application of Fourier methods to time series analysis is for determining hidden periodicities. These periodicities are assumed to be periodic functions of time which repeat themselves indefinitely over time with a fixed period, such as the position of the earth around the sun or the height of tides. Fourier analysis, however, can be successfully applied to the analysis of transients in the mean of a time series.

Suppose that the series $Y(t) = f(t) + u(t)$ is observed at the n consecutive times $t_o, \ldots, t_o + (n-1)\delta$, where $f(t)$ is the function defined by equation 2a or 2b and $t_o = 0$ for convenience. All of the results in the previous sections are valid for the problem of estimating the Fourier coefficients $a(\omega_k)$ and $b(\omega_k)$ and the power spectrum of the disturbances. The question of whether or not

$f(t)$ repeated itself in the past $(t < t_o)$ or will repeat itself in the future $(t > t_o + n\delta)$ can not be answered by an analysis of the sample; the answer must be assumed in the extrapolation of the model. We can use the estimates of $a(\omega_k)$ and $b(\omega_k)$ to predict $f(t)$ in the future. The best predictor is just $\hat{f}(t)$. However if $f(t) + u(t)$ is not a reasonably accurate model for $Y(t)$ in the future, then $f(t)$ is not a useful predictor of the expected value of the dependent variable. If, on the other hand, there are time periods in the future when $f(t) + u(t)$ is appropriate, then $\hat{f}$ is again a good predictor of the expected value of Y.

For example, suppose that $Y(t)$ is the interest rate at time t and $f(t)$ is a characteristic transient fluctuation in the mean rate after a discrete change in the money supply. That is, if a change occurs at t_o, $f(t)$ for $t_o \leq t_o + n\delta$ describes the transient behavior of the mean rate as it moves from the old steady-state position to the new steady-state. If $f(t)$ is a sum of several damped sinusoidal functions, Fourier analysis of the sample of $Y(t)$ taken after t_o can yield good estimates of the periods of the damped oscillations and can even determine the rate of decay of the transient. The resulting estimate of $f(t)$ can be used to predict the transient response in the interest rate after a future change in the money supply. The subject of transient response analysis for a linear system is discussed in almost any book on linear systems in engineering.

Another common misconception is that any significant peak in the estimated power spectrum is due to the existence of a sinusoidal component in the time series which will repeat itself indefinitely with a fixed period. It is true that a large peak at ω_o in the estimated spectrum indicates that there is a sinusoidal variation in the data whose period is about $2\pi/\omega_o$. However, there is no way to infer that the period or the amplitude of the sinusoidal will remain constant in the future, even if the underlying parameters of time series remain stationary over time.

As an example, consider the Gaussian Markov process $\{u(t)\}$ where $\alpha = 2$ and $h(0) = (1-\Delta)^2$, $h(1) = 0$, $h(2) = 1$ where Δ is a small positive number. i.e., $(1-\Delta)^2 u(t) + u(t-2) = \epsilon(t)$.

Thus from (24)

$$H(\omega) = (1-\Delta)^2 + e^{i2\omega}$$

and

$$|H(\omega)|^2 = \left[\cos 2\omega + (1-\Delta)^2\right]^2 + \sin^2 2\omega.$$

Since $|H(\frac{\pi}{2})| = \Delta|\Delta-2|$, for small Δ the power spectrum $\sigma^2|H(\omega)|^{-2}$ of $\{u(t)\}$ has a large peak of height $\sigma^2/4\Delta^2$ at $\omega = \frac{\pi}{2}$. As $\Delta \to 0$, the height of the peak goes to infinity whereas for $\omega \neq \frac{\pi}{2}$, $S(\omega) \to \frac{\sigma^2}{2(1+\cos 2\omega)}$. Nonetheless, there is no sinusoidal component of period 4 time-units in the Markov process since $Eu(t) = 0$ for all t. Actually there are no deterministic sinusoidal components in any stationary random process.

There is an intuitive explanation of why the above process does not contain the deterministic sinusoidal component whose frequency is the point where the power spectrum has a large peak. Suppose that we have a non-overlapping set of samples of n consecutive observations of $\{u(t)\}$. Let t_p denote the starting point for the pth sample, with $t_1 < t_2 < \dots$. For $\omega_j = \frac{2\pi j}{n}$, the real and imaginary parts of

$$\hat{A}_p(\omega_j) = \frac{1}{n} \sum_{t=t_p}^{t_p+n-1} u(t) e^{i\omega_j t}$$

are independent identically distributed $N(0, \frac{1}{2n} S(\omega_j))$ variables, where $S(\omega) = \sigma^2 |H(\omega)|^{-2}$. From theorem 1 and (9b), we can write u(t) for $t_p \leq t < t_p + n$ as the sum

$$u(t) = \hat{a}_p(0) + 2 \sum_{j=1}^{n/2-1} |\hat{A}_p(\omega_j)| \cos(\omega_j t - \hat{\theta}_p(\omega_j)) + \hat{a}_p(\pi)(-1)^t \tag{47}$$

where $\hat{\theta}_p(\omega) = \tan^{-1} \frac{\mathrm{Im}[\hat{A}_p(\omega)]}{\mathrm{Re}[\hat{A}_p(\omega)]}$. For each frequency $\omega = \omega_j$, the phases $\hat{\theta}_p(\omega)$ have the uniform density $\frac{1}{2\pi}$ for $0 < \theta < 2\pi$, and the amplitudes $|\hat{A}_p(\omega)|$ have the expected value $\left[\frac{\pi}{4} \frac{S(\omega)}{n}\right]^{1/2}$ and variance $(1 - \frac{\pi}{4}) \frac{S(\omega)}{n}$ for large n. If we then let $\Delta = n^{-1}$, $S(\frac{\pi}{2}) \approx \frac{\sigma^2 n^2}{4}$ and thus from (47),

$$u(t) \approx \hat{A}_p \cos(\frac{\pi}{4} t - \hat{\theta}_p) \qquad t_p \leq t < t_p + n \tag{48}$$

where $$E(\hat{A}_p) = \frac{\sqrt{\pi}}{4} \sqrt{n\sigma^2} \quad \text{and} \quad V(\hat{A}_p) = (1 - \frac{\pi}{4}) \frac{n\sigma^2}{4}$$

If the t_p are sufficiently far apart, the phases θ_p are uncorrelated.

Over a time duration of several periods the process $\{u(t)\}$ resembles the sinusoidal periodic function $A \cos (\frac{\pi}{2} t - \theta)$. However, as the process is observed over a longer time, the amplitude and phase appear to vary in a purely random manner. Since the expected value and standard deviation of $\hat{A}_p$ are both approximately $\sqrt{n}\sigma$, the heights of the peaks of the sinusoid varies between zero and $\sqrt{n}\sigma$. The phase vary between zero and 2π radians, making the peaks and valleys move randomly with regards to the time origin $t = 0$. Thus it is only possible to predict $u(t)$ with any precision for a period or two in the future.

As an example of the concept of phase, consider an electric clock. If the clock is exact and is working properly, the motion of the hour hand around the dial is periodic with a period of exactly 12 hours. It is possible to accurately predict the position of the hand well into the future. Suppose, however, that the clock has a defect which causes the hour hand to speed up and slow down slightly in a purely random manner. Once this clock is set, it will show the correct hour for a certain period of time, but after a few days the position of the hour hand will be uncorrelated with the true time. The phase is the difference in radians between the true hour position and the position of the hour hand on the dial. The phase divided by the angular frequency is the time difference between the true hour and the clock time. If the phase is random and uniformly distributed over $(0, 2\pi)$, the time difference will vary in an incoherent manner over time.

A large peak, then, in the estimated power spectrum of a time series indicates that the sample has a sizable sinusoidal component with the indicated frequency. The only way to infer whether the amplitude and phase of the component are fixed or are random is to measure their variability over time. This requires a sample length which is many times the period of this sinusoid. A relatively small variance of the phase indicates that the component is deterministic. The phase variance would be zero if the component is a sinusoid with fixed amplitude and phase and if there were no random disturbances, i. e., $u(t) \equiv 0$. If the phase drifts between 0 and 2π in a random manner over time, the component is a random process. In many applications exogenous information leads the investigator to assume that a large spectral peak at frequency $\omega_k = 2\pi k/n$ indicates that the amplitude of the deterministic component $\cos(\omega_k t - \theta_k)$ in the linear model $y(t) = f(t) + u(t)$ is significantly different from zero. For example, for economic time series a large peak in the estimated spectrum at the frequency corresponding to a 12 month period indicates a deterministic seasonal cycle, based upon the exogenous information that the cyclic behavior is caused by the weather.

6. LINEAR TRENDS

Most economic time series contain trends. Suppose that the trend is linear and enters the model additively, i. e., $Y(t) = c + dt + f(t) + u(t)$. The presence of the trend will bias the estimates of the Fourier coefficients of $f(t)$ since c + dt is not orthogonal to the independent variables $\cos\omega_k t$ and $\sin\omega_k t$. The bias is greatest for the low frequency or long period components, and is negligible for the high frequency or short period components. It is important to reduce the bias effects of the trend as much as possible. Moreover, the slope of the trend is generally of interest. The simplest approach to estimating the trend and removing its effects from the series involves estimating c and d by regressing Y(t) on c + dt, and then computing the residual time series $Y(t) - \hat{c} - \hat{d}t$. The Fourier coefficients of f(t) are computed from the residuals (Durbin 1962). However, the estimates $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ are still biased since the trend function and the independent variables are not orthogonal. Nonetheless the biases of the estimates $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ computed from the residuals are smaller than the biases of the estimates computed directly from the Y(t). Bear in mind, however, that these $\hat{a}(\omega_k)$ and $\hat{b}(\omega_k)$ are <u>not</u> the least-squares estimators of $a(\omega_k)$ and $b(\omega_k)$ for the model with the linear trend.

If the disturbances are white, i.e., uncorrelated and homoskedastic, the least-squares estimators of d and $A(\omega_k) = a(\omega_k) + ib(\omega_k)$ are easily derived from the normal equations of the model $Y(t) = dt + f(t) + \epsilon(t)$. The intercept c is not necessary in the model since $a(0) = c$ and $b(0) = 0$. The least-squares estimators are:

$$\begin{pmatrix} \hat{d} \\ \hat{A}(0) \\ \hat{A}(\omega_1) \\ \vdots \\ \hat{A}(\omega_T) \end{pmatrix} = \begin{pmatrix} \Sigma t^2 & \Sigma t & \Sigma t e^{i\omega_1 t} & \cdots & \Sigma t e^{i\omega_T t} \\ \Sigma t & n & 0 & \cdots & 0 \\ \Sigma t e^{-i\omega_1 t} & 0 & n & 0 \cdots & 0 \\ \vdots & \vdots & 0 & n & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \Sigma t e^{-i\omega_T t} & 0 & 0 & \cdots 0 & n \end{pmatrix}^{-1} \begin{pmatrix} \Sigma t Y(t) \\ \Sigma Y(t) \\ \Sigma Y(t)e^{-i\omega_1 t} \\ \vdots \\ \Sigma Y(t) e^{-i\omega_T t} \end{pmatrix} \quad (49)$$

Note that all the sums in (49) can be computed by using the fast Fourier transform algorithm.

It should be clear that the two-stage technique for estimating the Fourier coefficients of f(t) can be extended to the case of an additive polynomial trend, or any trend of known functional form. For example, if the trend is $d_1 t + d_2 t^2$, d_1 and d_2 are estimated by regressing $Y(t)$ on $d_1 t + d_2 t^2$ and the residuals $Y(t) - \hat{d}_1 t - \hat{d}_2 t^2$ are then computed. The Fourier coefficients are computed from the residuals.

A multiplicative trend model, however, poses special problems. Suppose that we have the model

$$Y(t) = (c + dt)\, f(t)\, \varepsilon(t) \qquad 0 \leq t < n$$

We can use the above procedures for $\{\log Y(t)\}$ if we assume that $\{\log \varepsilon(t)\}$ is a white Gaussian process and that d/c is small enough so that

$$\log(c + dt) \approx \log c + \frac{d}{c}\, t$$

The presence of an undetected trend causes a specification error in the ordinary least-squares estimation of the Fourier coefficients of a periodic mean of a time series. In addition, a trend will produce a bias in the estimate of the power spectrum of the disturbance. The bias is most severe for the low frequency spectrum estimates.

For the case of estimating the power spectrum of a stationary random process, another source of bias is the presence of discontinuities in the beginning and end of the sample. The estimates of the low frequency part of the spectrum will be biased if there is an appreciable difference between the sample mean and the first few or last few observations in the sample. In order to reduce this effect, the beginning and end of the sample should be tapered. One of the simplest tapering methods is to multiply the first $n/20$ observations by $k/20n$ for $k = 1, \ldots, n/20$ and to multiply the last $n/20$ observations by $1 - k/20n$. This type of tapering effects 10 percent of the sample. The investigator, of course, is free to taper the sample more or less as he chooses. Tapering, however, produces a bias in the estimates of the Fourier coefficients of the periodic mean $f(t)$ if it is present in the data. This bias is less than the bias produced by the discontinuities at the ends of the record.

REFERENCES

1. Alsop, L. and A. Nowroozi, "Fast Fourier Analysis," J. Geophys. Res., 71, 5482-5483 1966

2. Bartlett, M., An Introduction to Stochastic Processes, 312 pp, Cambridge University Press, Cambridge, England, 1955

3. Bingham, C., M. Godfrey, and J. W. Tukey, "Modern Techniques of Power Spectrum Estimation," IEEE Trans. Audio Electro-Acoustics, AU-15, 55-66, Jun 1967

4. Blackman, R. and J. W. Tukey, The Measurement of Power Spectra, 190 pp, Dover, New York, 1959

5. Cooley, J. and J. W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," Math. Computation, 19, 297-301, 1965

6. Courant, R. and D. Hilbert, Methods of Mathematical Physics, Volume 1, Interscience, New York, 1953

7. Durbin, J. "Trend Elimination by Moving-Average and Variate-Difference Filters," Bull. Intern. Statist. Inst., 39

8. Gentleman, W. M., and G. Sande, "Fast Fourier Transforms - For Fun and Profit," in Proceedings of the Fall Joint Computer Conference, 563-578, Spartan Books, Washington, D.C., 1966

9. Grenander, U. and M. Rosenblatt, Statistical Analysis of Stationary Time Series, 300 pp, John Wiley, New York, 1957

10. Hannan, E. J., Time Series Analysis, 152 pp. Methen Statistical Monographs, London, 1960

11. Hinich, M. and C. S. Clay, "The Application of the Discrete Fourier Transform in the Estimation of Power Spectra, Coherence, and Bispectra of Geophysical Data," Reviews of Geophysics, 6, No. 3, Aug 1968

12. Jenkins, G. M. and D. G. Watts, Spectral Analysis and its Applications, 541 pp, Holden-Day, San Francisco, 1968

13. Parzen, E., "On Consistent Estimates of the Spectrum of a Stationary Time Series," Ann. Math. Statist. 28, 329-348, 1957

14. Priestly, M. B., "The Analysis of Stationary Processes with Mixed Spectra - I," J. Roy. Stat. Soc. B., 24, 215-233, 1962a.

15. ____________________, "The Analysis of Stationary Processes with Mixed Spectra - II," J. Roy. Stat. Soc. B, 24, 511-529, 1962b.

16. Walker, A.M., "Large-sample Properties of Least-squares Estimators of Harmonic Components in a Time Series with Stationary Residuals," Technical Report No. 23, Department of Statistics, Stanford University, Stanford, California, 1968.

17. Welch, P. D., "A Direct Digital Method of Power Spectrum Estimation," IBM J. Res. Development, 5, 141-156, 1961.