ON CONSISTENT ESTIMATES OF THE SPECTRUM OF A STATIONARY TIME SERIES[1, 2]

By Emanuel Parzen[3]

Columbia University, Hudson Laboratories, Dobbs Ferry, N. Y.

Summary. This paper is concerned with the spectral analysis of wide sense stationary time series which possess a spectral density function and whose fourth moment functions satisfy an integrability condition (which includes Gaussian processes). Consistent estimates are obtained for the spectral density function as well as for the spectral distribution function and a general class of spectral averages. Optimum consistent estimates are chosen on the basis of criteria involving the notions of order of consistency and asymptotic variance. The problem of interpolating the estimated spectral density, so that only a finite number of quantities need be computed to determine the entire graph, is also discussed. Both continuous and discrete time series are treated.

1. Introduction. A stochastic process is a family of random variables $x(t)$, where t varies in some set T. If the set T is the infinite real line, then $x(t)$ is called a random function, and if $T = \{0, \pm 1, \pm 2, \cdots\}$, then $x(t)$ is called a random sequence. If the parameter t is interpreted as denoting time, then the stochastic process is called a time series, with the adjectives continuous or discrete being used to indicate whether it is a random function or a random sequence.

Let us suppose that we have observed a sample of length T of a (continuous or discrete) time series $x(t)$. The general problem of time series analysis is to infer the statistical characteristics of $x(t)$ from the observed sample. Now in order to perform a statistical analysis of $x(t)$, one has to assume a model for $x(t)$ which is completely specified except for the values of certain parameters which one proceeds to estimate on the basis of the observed sample.

A widely adopted model for $x(t)$ (see Grenander and Rosenblatt [4], [5]) is the following. It is assumed that $x(t)$ may be written as a sum of a mean value function $m(t)$ and a fluctuation function $y(t)$:

$$x(t) = m(t) + y(t). \tag{1.1}$$

The domain T of the variable t is to be taken as the infinite real line, $-\infty < t < \infty$, in the continuous case, and as the set of integers $0, \pm 1, \pm 2, \cdots$ in the discrete case. We seek to treat simultaneously both discrete and continuous time

Received September 13, 1955; revised October 15, 1956.

[1] This work was supported by the Office of Naval Research under Contract N6-ONR-27135. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Hudson Laboratories Contribution No. 15. A more detailed version of this paper may be found in Hudson Laboratories Technical Report No. 33.

[3] Now at Stanford University.

series. Most equations will hold for both cases, with the proper interpretation, which will be explained as we proceed.

It is assumed that the function $m(t)$ is nonrandom, and that there is a fixed number K of known functions $\varphi_1(t), \cdots, \varphi_K(t)$ such that $m(t)$ may be written as a linear combination of the $\varphi_j(t)$:

$$m(t) = m_1\varphi_1(t) + \cdots + m_K\varphi_K(t). \tag{1.2}$$

The constants m_j (for $j = 1, \cdots, K$) are unknown, and are to be estimated from the sample.

The fluctuation function $y(t)$ is a stochastic process, whose mean value function $Ey(t)$ vanishes identically in t. It is assumed that it possesses a finite second moment $E \mid y(t) \mid^2$, and that it is wide sense stationary, which means that the product moment $Ey(t)\, y(t + v)$ is independent of t, and depends only on v. One then defines the covariance function

$$R(v) = Ey(t)\, y(t + v). \tag{1.3}$$

In the case of random functions, it is assumed that $R(v)$ is continuous. Then, $R(v)$ possesses a representation as a Fourier-Stieltjes integral:

$$R(v) = \int e^{iwv}\, dF(w), \tag{1.4}$$

where $F(w)$ is a bounded non-decreasing function, called the spectral distribution function of the process. The domain of the variable v is the same as that of t, and the domain of the variable w is $-\infty$ to ∞ in the continuous case, and $-\pi$ to π in the discrete case. The domain of integration of an integral involving w is to be taken as the whole domain of w, in cases where it is not otherwise specified.

It is assumed next that $R(v)$ is summable. It then follows that the spectral distribution function $F(w)$ possesses a continuous density function $f(w)$, called the spectral density function of the time series $x(t)$. The following relations hold:

$$R(v) = \int e^{iwv} f(w)\, dw, \tag{1.5}$$

$$f(w) = \frac{1}{2\pi} \int e^{-iwv} R(v)\, dv \tag{1.6c}$$

$$= \frac{1}{2\pi} \sum e^{-iwv} R(v). \tag{1.6d}$$

In cases where the limits of integration (or summation) of an integral (or sum) involving the variables u or v are omitted, they are to be assumed to be $-\infty$ to $+\infty$. Henceforth, we write equations of the type of (1.6) only once, for the continuous case, with the understanding that for every such equation a corresponding equation may be written for the discrete case by replacing the integral by a sum. For certain important equations, we will write, without further explanation, two equations, with a suffix d for the discrete case and a suffix c for the continuous case.

The model for the process $x(t)$ which has just been described assumes only a knowledge of the first and second moments of the process, and assumes no knowledge of the probability distribution. The moments are assumed to be completely specified by the constants $m_1, \cdots, m_K$, and the covariance function $R(v)$, or equivalently the spectral density function $f(w)$. By analysis of an observed time series is meant the estimation of the value of these quantities on the basis of observed samples. The estimation of the constants $m_1, \cdots, m_K$ is called regression analysis, and the estimation of the spectral functions is called spectral analysis.

A basic requirement for an estimate is that it be *consistent* in quadratic mean. Let m be an unknown parameter of a time series $x(t)$, and let $x(t)$ for $0 \leqq t \leqq T$ (or $t = 1, \cdots, T$ in the discrete case) be an observed sample of the time series. An estimate m_T of m, formed on the basis of the sample, is said to be *consistent* in quadratic mean if the mean square error $E \mid m_T - m \mid^2$ tends to zero, as $T \to \infty$, where the expected value is taken under the assumption that m is the true parameter value. If an estimate is consistent, it is then asymptotically *unbiased*, which means that $Em_T \to m$ as $T \to \infty$.

However, we shall be interested in estimates which are consistent and asymptotically unbiased at certain prescribed rates. Let α be a positive number. We define an estimate to be asymptotically unbiased of the order of T^α if, for some finite constant β,

$$\lim_{T \to \infty} T^\alpha (Em_T - m) = \beta. \tag{1.7}$$

We say that an estimate possesses an asymptotic variance σ^2 of the order of $T^{2\alpha}$ if σ is positive and

$$\lim_{T \to \infty} T^\alpha \sigma[m_T] = \sigma, \tag{1.8}$$

where $\sigma^2 [m_T] = E \mid m_T - Em_T \mid^2$ is the variance of m_T. The importance of these notions derives from the central limit theorem, for dependent random variables, from which one may hope to obtain conditions that the normalized random variable $(m_T - Em_T)/\sigma[m_T]$ tends to a normal distribution. We define an estimate to be consistent of the order of $T^{2\alpha}$, with asymptotic bias β and asymptotic variance σ^2, if (1.7) and (1.8) hold. If such an estimate obeys the Central Limit Theorem, then the random variable $T^\alpha (m_T - m)$ tends to a normal distribution with mean β and variance σ^2. Many estimates that one encounters are consistent of the order of T; however, we will encounter below estimates which are consistent of a lower order.

A knowledge of the order of consistency, the asymptotic bias, and the asymptotic variance of an estimate is valuable on several counts, as will be shown in detail in a later paper [8].

The problem of regression analysis has been extensively treated by Grenander and Rosenblatt in several excellent papers (see [5]), in which they obtained expressions for the asymptotic variances (of order T) of various estimates of the constants m_j in the model given above, and obtained conditions that the least squares estimate and the best linear unbiased estimates have the same asymptotic variance. We mention regression analysis here only to point out that the results of this paper remain valid if in estimating the spectrum one uses the deviations of the observed values of $x(t)$ from the sample mean value function formed by inserting into (1.2) the least squares estimates of the constants m_j. As far as detailed considerations are concerned, we consider only the case where $m(t) = m$, an unknown constant.

The problem of outstanding interest at the present time in the analysis of time series is that of estimating the spectral density function, and it is this problem that is treated in this paper. In view of Eq. (1.6), the obvious way to estimate $f(w)$ is to form the Fourier transform $f_T(w)$ of the least squares estimate $R_T(v)$ of the covariance. The sample spectral density function $f_T(w)$ so obtained is essentially what has been studied by various authors under the name of the *periodogram*. However, as is well known, it turns out that $f_T(w)$ is not a consistent estimate of $f(w)$.

Rather, to begin with, we are only able to estimate what may be called *spectral averages*; that is, averages of the spectral density function of the form

$$J(A) = \int A(w)f(w)\,dw, \tag{1.9}$$

where $A(w)$ is a suitably chosen function. On the one hand $A(w)$ may be chosen to be a unit step function, $A(w) = 1$ or 0 according as $w < w_0$ or $w \geqq w_0$. Then $J(A)$ represents the spectral distribution function $f(w_0)$. On the other hand, $A(w)$ may be a function highly peaked about a center frequency w_0.

In Section 5, we obtain a class of consistent estimates of the spectral density function at a point w_0. However, the order of consistency of these estimates will be $T^{2\alpha}$, where $0 < \alpha < \frac{1}{2}$. Expressions are obtained for the asymptotic variance and bias of such estimates, so that the means are at hand for choosing among the large class of estimates presented. In Section 6, consistent estimates of the spectral density function, asymptotically optimum within the family of estimates considered, are discussed. In Section 7, we use the ideas leading to consistent estimates of the spectral density to obtain alternative estimates of the spectral averages. In Section 8, we treat the problem of interpolating the spectral density.

2. Assumptions on the fourth moments. Some additional assumptions are required in addition to the assumptions we have already stated. We assume that the fluctuation function $y(t)$ is wide sense stationary of order 4, in the sense that $E\,|\,y(t)\,|^4$ exists for all t, and the fourth moment function

$$P(v_1, v_2, v_3) = Ey(t)\,y(t + v_1)\,y(t + v_2)\,y(t + v_3) \tag{2.1}$$

is a function only of the time differences v_1, v_2, v_3, and not of the initial time t.

Now if the process $y(t)$ were normally distributed, then $P(v_1, v_2, v_3)$ could be expressed in terms of the covariance function $R(v)$ as follows:

$$P_G(v_1, v_2, v_3) = R(v_1)R(v_2 - v_3) + R(v_2)R(v_3 - v_1) + R(v_3)R(v_1 - v_2). \tag{2.2}$$

We introduce the function

$$Q(v_1, v_2, v_3) = P(v_1, v_2, v_3) - P_G(v_1, v_2, v_3), \tag{2.3}$$

which is the difference between the actual fourth moment function of $y(t)$, and what it would be if $y(t)$ were Gaussian. We refer to $Q(v_1, v_2, v_3)$ as the non-Gaussian part of the fourth moment function of $y(t)$; it is the same as the fourth cumulant function.

We assume that $Q(v_1, v_2, v_3)$ is absolutely summable (and, in the continuous parameter case, continuous) over all of (v_1, v_2, v_3) space.

We will find in many instances that $Q(v_1, v_2, v_3)$ admits of a representation as a Fourier integral:

$$Q(v_1, v_2, v_3) = \iiint \exp [i(w_1 v_1 + w_2 v_2 + w_3 v_3)] g(w_1, w_2, w_3)\, dw_1 dw_2 dw_3, \tag{2.4}$$

where the function $g(w_1, w_2, w_3)$ is absolutely integrable over all of (w_1, w_2, w_3) space. We may have also the relation

$$\int du Q(v_1, u, u + v_2) = 2\pi \iint dw_1 dw_2 g(w_1, -w_2, w_2) \exp [i(w_1 v_1 + w_2 v_2)]. \tag{2.5}$$

We will assume these relations to be valid, since they simplify the writing of some of the results. It should be pointed out that the notion of the Fourier transform of the non-Gaussian part of the fourth moment function has previously been considered by Magness [6] where some examples may be found.

In the continuous parameter case we assume also that the stochastic process $x(t, \omega)$, where ω varies in a space Ω on which the basic probability measure P is defined, is measurable jointly in t and ω. Then the random integrals, such as $\int_0^T x(t)\, dt$, which are employed exist with probability one, by virtue of the Fubini theorem (see Doob [9]). Alternatively, the random integrals employed may be defined as limits in quadratic mean (see Loève [10]).

3. The sample covariance and spectral density functions. The estimates of the spectrum that we shall consider will be defined in terms of two functions, the sample covariance function and the sample spectral density function, which are defined in this section. Given a sample of observed values of $x(t)$ for $0 \leqq t \leqq T$ (or for $t = 1, \cdots, T$), let m_T be the least squares estimate of m, and consider the function $Y_T(t)$, defined by

$$\begin{aligned} Y_T(t) &= x(t) - m_T \qquad && \text{for } 0 \leqq t \leqq T, \\ &= 0 && \text{otherwise.} \end{aligned} \tag{3.1}$$

Define now the function

$$f_T(w) = \frac{1}{T} \left| \int_0^T Y_T(t) e^{-itw}\, dt \right|^2 \tag{3.2c}$$

$$= \frac{1}{T} \left| \sum_{t=1}^T Y_T(t) e^{-itw} \right|^2, \tag{3.2d}$$

which may be regarded as the notion of the "periodogram" extended to the case of time series with an unknown mean value. We call $f_T(w)$ the sample spectral density function, because its Fourier integral

$$R_T(v) = \int e^{iwv} f_T(w)\, dw \tag{3.3}$$

is a consistent estimate of the covariance function. We call $R_T(v)$ the sample covariance function. It vanishes for $|v| \geqq T$, and for $|v| < T$,

$$R_T(v) = \frac{1}{T} \int_0^{T-|v|} Y_T(t) Y_T(t + |v|)\, dt \tag{3.4c}$$

$$= \frac{1}{T} \sum_{t=1}^{T-|v|} Y_T(t) Y_T(t + |v|). \tag{3.4d}$$

We may invert (3.3) to obtain

$$f_T(w) = \frac{1}{2\pi} \int_{-T}^{T} e^{-ivw} R_T(v)\, dv \tag{3.5c}$$

$$= \frac{1}{2\pi} \sum_{|v| \leqq T} e^{-ivw} R_T(v). \tag{3.5d}$$

In the continuous parameter case, the interval of integration in (3.3) is infinite, and to establish that $f_T(w)$ is summable, one needs to employ a standard argument involving Plancherel's theorem.

An important role in the sequel will be played by the following representation of $R_T(v)$, for $|v| \leqq T$:

$$R_T(v) = D_T(v) + b_T(v) + R(v)\left(1 - \frac{|v|}{T}\right), \tag{3.6}$$

where

$$D_T(v) = \frac{1}{T} \int_0^{T-|v|} dt \{y(t) y(t + |v|) - R(v)\} \tag{3.7}$$

and $b_T(v)$ is defined so as to make Eq. (3.6) correct.

The term $b_T(v)$ represents the bias arising from the fact that the sample covariances are computed using $Y_T(t)$, the deviations of the observations from the sample mean. That it may be essentially ignored in our calculations will follow from the fact that there is a constant K such that

$$T^2 E\, |b_T(v)|^2 \leqq K^2 \tag{3.8}$$

for any v and T. To establish (3.8), it suffices to show that there is a constant K' such that, for any choice of numbers T and T_1, T_2, T_3, T_4 satisfying $0 \leqq T_1, T_2, T_3, T_4 \leqq T$,

$$E \left| \int_{T_1}^{T_2} \int_{T_3}^{T_4} y(t_1) y(t_2)\, dt_1\, dt_2 \right|^2 \leqq K' T^2, \tag{3.9}$$

which follows from the fact that the expected value in (3.9) is less than

$$3\left\{T\int |R(u)|\,du\right\}^2 + T\iiint |Q(u_1, u_2, u_3)|\,du_1\,du_2\,du_3.$$

We next evaluate the covariance of $D_T(v)$. We obtain that, for any non-negative numbers v_1 and v_2,

$$\begin{aligned}TED_T(v_1)D_T(v_2) &= \int_{-T}^{T} duU_T(u, v_1, v_2)\{Q(v_1, u, u+v_2) \\ &\quad + R(u)R(u+v_1-v_2) + R(u+v_1)R(u-v_2)\},\end{aligned} \tag{3.10}$$

where $U_T(u, v_1, v_2)$ is a function with values between 0 and 1 defined as follows:

$$\begin{aligned}U_T(u, v_1, v_2) &= 0 && u \leqq -T + v^2 \\ &= 1 - \frac{v_2 + u}{T} && -T + v_2 \leqq u \leqq \min(0, v_2 - v_1) \\ &= 1 - \frac{\max(v_1, v_2)}{T} && \\ & && \min(0, v_2 - v_1) \leqq u \leqq \max(0, v_2 - v_1) \\ &= 1 - \frac{v_1 + u}{T} && \max(0, v_2 - v_1) \leqq u \leqq T - v_1 \\ &= 0 && T - v_1 \leqq u\end{aligned} \tag{3.11}$$

To establish (3.10) one makes the change of variable $u = t_2 - t_1$, $v = t_2$ in the expression

$$\begin{aligned}T^2ED_T(v_1)D_T(v_2) &= \int_0^{T-v_1}\int_0^{T-v_2} dt_1\,dt_2 \\ &\{Q(v_1, t_2 - t_1, t_2 - t_1 + v_2) + R(t_2 - t_1)\,R(t_2 - t_1 + v_2 - v_1) \\ &+ R(t_2 - t_1 + v_2)\,R(t_2 - t_1 - v_1)\}.\end{aligned}$$

As a first consequence of (3.8) and (3.10, we obtain the following theorem.

THEOREM 3: *For any non-negative numbers v, v_1, and v_2,*

$$\lim_{T\to\infty} T^{1/2}\,|ER_T(v) - R(v)| = 0, \tag{3.12}$$

$$\begin{aligned}&\lim_{T\to\infty} T\,\mathrm{Cov}[R_T(v_1), R_T(v_2)] \\ &= \int du\,\{Q(v_1, u, u+v_2) + R(u)R(u+v_1-v_2) + R(u+v_1)R(u-v_2)\}\end{aligned} \tag{3.13}$$

$$\begin{aligned}&= 2\pi\left\{\iint dw_1\,dw_2\,\exp[i(w_1v_1 + w_2v_2)]g(w_1, -w_2, w_2)\right. \\ &\quad \left. + \int dwf^2(w)\exp[iw(v_1 - v_2)] + \int dwf^2(w)\exp[iw(v_1 + v_2)]\right\}.\end{aligned} \tag{3.14}$$

4. Estimation of spectral averages. To estimate the spectral average $J(A)$ there are two methods available, which may be called the method of filtering and the method of covariance averages. In the method of filtering, one estimates the variance (zero lag covariance) of a new time series obtained by filtering the observed series. In the method of covariance averages, one defines a sample spectral average $J_T(A)$, which may be expressed as an average with respect to the sample spectral density function or with respect to the sample covariances. This latter form is the more convenient for computations. Only the method of covariance averages is discussed here.

Spectral averaging functions: A function $A(w)$ will be called a spectral averaging function if it is a real valued function which is both absolutely integrable and square integrable. Its Fourier transform

$$a(v) = \frac{1}{2\pi} \int e^{-iwv} A(w)\, dw \tag{4.1}$$

is then bounded and square integrable. We call $a(v)$ a covariance averaging function. We assume finally that

$$| a(v) | = o(| v |^{-r}) \qquad \text{for some } r > \tfrac{1}{2}. \tag{4.2}$$

If $A(w)$ has finite total variation (and also, in the continuous parameter case, vanishes at infinity), then $| a(v) | = O(| v |^{-1})$. From (4.2) it follows that, for some constant K_1 and some $\epsilon > 0$,

$$\int_{-T}^{T} | a(v) |\, dv \leqq K_1 T^{1/2-\epsilon} \tag{4.3}$$

for all T, and also that

$$\int | v |^{1/2} | a(v)R(v) |\, dv < \infty. \tag{4.4}$$

A lemma: Of frequent use in the sequel will be the following lemma.

LEMMA 4: *Let $q \geqq 0$ and $s > 0$. Let M_T be a sequence of constants tending to ∞ as $T \to \infty$. Suppose that*

$$\int | v |^{q} | a(v)R(v) |\, dv < \infty. \tag{4.5}$$

Then, as $T \to \infty$,

$$M_T^q \int_{|v| \geqq M_T} | a(v)R(v) |\, dv \to 0, \tag{4.6}$$

$$\frac{1}{M_T^s} \int_{|v| \leqq M_T} | v |^{q+s} | a(v)R(v) |\, dv \to 0. \tag{4.7}$$

Sample spectral averages: The spectral average $J(A)$ may be defined in terms of either the spectral density or the covariance function by

$$J(A) = \int A(w)f(w)\,dw = \int a(v)R(v)\,dv. \tag{4.8}$$

Accordingly, we define the sample spectral average $J_T(A)$ by

$$J_T(A) = \int A(w)f_T(w)\,dw = \int_{-T}^{T} a(v)R_T(v)\,dv. \tag{4.9}$$

The properties of $J_T(A)$ as an estimate of $J(A)$ are given in the following theorem:

THEOREM 4: *For any spectral averaging functions* $A(w)$, $A_1(w)$, *and* $A_2(w)$,

$$\lim_{T\to\infty} T^{1/2} \mid EJ_T(A) - J(A) \mid = 0, \tag{4.10}$$

$$\begin{aligned}\lim_{T\to\infty} T \operatorname{Cov}[J_T(A_1), J_T(A_2)] &= 4\pi \int dw f^2(w)A_1^0(w)A_2^0(w)\\ &\quad + 2\pi \iint dw_1\,dw_2\, g(w_1, -w_2, w_2)A_1^0(w_1)A_2^0(w_2),\end{aligned} \tag{4.11}$$

where

$$A^0(w) = \tfrac{1}{2}\{A(w) + A(-w)\}. \tag{4.12}$$

PROOF. Omitted, since it is similar to the proofs of Theorems 5A and 5B.

5. Estimation of the spectral density. Various authors have pointed out that the sample spectral density function, or periodogram, $f_T(w)$ is not a consistent estimate of the spectral density function $f(w)$. The suggestion has been made to estimate $f(w)$ at a point w_0 by averaging the values of $f_T(w)$ in the neighborhood of w_0. However, this yields a consistent estimate not of $f(w_0)$, but rather of the spectral average in the neighborhood of w_0. To eliminate this bias, one needs to narrow the neighborhood averaged over as the sample size is increased. The manner in which this is to be done is examined in this section. A similar method of obtaining a consistent estimate of the spectral density is that of Bartlett (see [1]), who has suggested a modified form of the periodogram. More general methods of constructing consistent estimates of the spectral density have been given by Grenander [3] and Tukey [7]. In this section these methods are generalized somewhat further. A noteworthy feature of the general method of constructing consistent estimates of the spectral density which is discussed in this section is that one may construct estimates which are consistent of any prescribed order $T^{2\alpha}$, where $0 < \alpha < \frac{1}{2}$.

Covariance averaging kernels: A function $k(z)$, defined for all real z, will be called a covariance averaging kernel if it fulfills the following conditions. It is even, bounded, square integrable, and normalized so that $k(0) = 1$. Its Fourier transform $K(w)$ is defined (as a limit in quadratic mean) to satisfy

$$k(z) = \int e^{-iwz}K(w)\,dw. \tag{5.1}$$

It is assumed that there is a constant K_1 and an $\epsilon > 0$ such that

$$B \int_{-T}^{T} | k(Bv) | \, dv \leqq K_1(BT)^{1/2-\epsilon}, \tag{5.2c}$$

$$B \sum_{|v| \leqq T} | k(Bv) | \leqq K_1(BT)^{1/2-\epsilon}, \tag{5.2d}$$

for every B and T. A sufficient condition for (5.2) to hold is that $k(z)$ satisfy (4.2).

Given a kernel $k(z)$, and a positive number r, define

$$k^{(r)} = \lim_{z \to 0} \frac{1 - k(z)}{| z |^r}. \tag{5.3}$$

We assume next that there is a largest number r, called the *characteristic exponent* of the kernel $k(z)$, such that $k^{(r)}$ exists and is finite (nonzero). If the limit in (5.3) exists for every positive r, then the kernel is said to have characteristic exponent ∞

Estimates of the spectral density: Let $k(z)$ be a covariance averaging kernel and let B_T be a sequence of constants tending to 0, as $T \to \infty$, in such a way that $B_T T \to \infty$. As an estimate of the spectral density function we define the even function

$$f_T^*(w) = \frac{1}{2\pi} \int_{-T}^{T} e^{-ivw} k(B_T v) R_T(v) \, dv \tag{5.4c}$$

$$= \frac{1}{2\pi} \sum_{|v| \leqq T} e^{-ivw} k(B_T v) R_T(v). \tag{5.4d}$$

The constant B_T may be called the *bandwidth* of the estimate. In terms of the sample spectral density, one may write

$$f_T^*(w) = \frac{1}{B_T} \int_{-\infty}^{\infty} K\left(\frac{\lambda - w}{B_T}\right) f_T(\lambda) \, d\lambda, \tag{5.5}$$

where $f_T(\lambda)$ is to be defined as a periodic function in the discrete parameter case. Alternate ways in which $f_T^*(w)$ may be written in the discrete parameter case are

$$f_T^*(w) = \int_{-\pi}^{\pi} d\lambda f_T(\lambda) \frac{1}{B_T} \sum_{n=-\infty}^{\infty} K\left(\frac{\lambda - w - 2\pi n}{B_T}\right) \tag{5.5'}$$

and

$$f_T^*(w) = \int_{-\pi}^{\pi} d\lambda f_T(\lambda) K_T(\lambda - w), \tag{5.5''}$$

where $K_T(w)$ is defined so that

$$k(B_T v) = \int_{-\pi}^{\pi} e^{-ivw} K_T(w) \, dw.$$

Various estimates of the spectral density which have been proposed (see Bartlett [1], [2], Grenander [3], Tukey [7]) may be obtained as instances of (5.4).

By choosing $k(z) = 1 - |z|$ if $|z| \leqq 1$ and 0 otherwise and letting $B_T = (1/M)$, where M is an integer less than T, one has a modified form of Bartlett's estimate:

$$\frac{1}{2\pi} \sum_{v=-M}^{M} e^{-ivw} \left(1 - \frac{|v|}{M}\right) R_T(v).$$

By choosing $k(z) = \sin z/z$ and letting $B_T = h$, one has Daniell's estimate:

$$\frac{1}{2\pi} \sum_{v=-T}^{T} e^{-ivw} \frac{\sin(hv)}{hv} R_T(v) = \frac{1}{2h} \int_{-h}^{h} f_T(\lambda - w)\, d\lambda.$$

By choosing $k(z) = 1$ if $|z| < 1$ and 0 otherwise, and letting $B_T = (1/M)$, one has the truncated estimate

$$\frac{1}{2\pi} \sum_{v=-M}^{M} e^{-ivw} R_T(v),$$

which, in view of the fact that the Fourier transform $K(w)$ of $k(z)$ is not non-negative, has the possibly undesirable property that it is not necessarily non-negative.

The properties of the estimate $f_T^*(w)$ are embodied in the following theorems.

THEOREM 5A. *The asymptotic covariance of the estimate $f_T^*(w)$ defined by* (5.4) *satisfies, for any non-negative frequencies w_1 and w_2,*

$$\lim_{T\to\infty} TB_T \operatorname{Cov}[f_T^*(w_1), f_T^*(w_2)] = f^2(w) \int k^2(z)\, dz\{1 + \delta(0, w_1)\}\delta(w_1, w_2), \tag{5.6}$$

where $\delta(w_1, w_2) = 1$ if $w_1 = w_2$ and 0 if $w_1 \neq w_2$. Further, for any $\epsilon > 0$, the limit in (5.6) *is uniform in w_1 and w_2 such that $w_1 \geqq \epsilon$ and $w_2 \geqq \epsilon$.*

REMARK. The integral in (5.6) is *not* to be replaced by a summation in the discrete parameter case.

THEOREM 5B. *Let $q > 0$ be such that*

$$\int |v|^q |R(v)|\, dv < \infty, \tag{5.7c}$$

$$\sum |v|^q R(v) < \infty. \tag{5.7d}$$

Define the generalized qth spectral derivative $f^{(q)}(w)$ by

$$f^{(q)}(w) = \frac{1}{2\pi} \int e^{-ivw} |v|^q R(v)\, dv \tag{5.8c}$$

$$= \frac{1}{2\pi} \sum e^{-ivw} |v|^q R(v). \tag{5.8d}$$

Let the covariance averaging kernel $k(v)$ have characteristic exponent $r \geqq q$. Let the constants B_T be chosen so that

$$0 < \lim_{T\to\infty} TB_T^{1+2q} < \infty. \tag{5.9}$$

Then, uniformly in w,

$$\lim_{T\to\infty} B_T^{-2q} \mid Ef_T^*(w) - f(w) \mid = \mid k^{(q)} f^{(q)}(w) \mid^2 \qquad \text{if } r = q \tag{5.10}$$
$$= 0 \qquad \text{if } r > q.$$

PROOFS. We first show that the term $b_T(v)$, defined by (3.6), has no effect by showing that uniformly in w,

$$\lim_{T\to\infty} TB_T E \left| \int_{-T}^{T} dve^{-ivw} k(B_T v) b_T(v) \right|^2 = 0. \tag{5.11}$$

By Minkowski's inequality, (3.8), and (5.2), the square root of the quantity in (5.11) whose limit is being taken is less than

$$K(TB_T)^{-1/2} B_T \int_{-T}^{T} dv \mid k(B_T v) \mid \leqq KK_1 T^{-\epsilon},$$

which tends to 0 as $T \to \infty$.

We next establish Theorem 5B. By (3.6), we write

$$2\pi f_T^*(w) = \int_{-T}^{T} dve^{-ivw} k(B_T v) \left\{ D_T(v) + b_T(v) + R(v) \left(1 - \frac{|v|}{T}\right)\right\}. \tag{5.12}$$

Therefore

$$\begin{aligned} 2\pi B_T^{-q}(Ef_T^*(w) - f(w)) = EB_T^{-q} & \int_{-T}^{T} dve^{-ivw} k(B_T v) b_T(v) \\ & - B_T^{-q} \int_{-T}^{T} dve^{-ivw}(1 - k(B_T v))R(v) \\ & - \frac{1}{T} B_T^{-q} \int_{-T}^{T} dve^{-ivw} \mid v \mid k(B_T v) R(v) \\ & - B_T^{-q} \int_{|v| \geqq T} dve^{-ivw} R(v). \end{aligned} \tag{5.13}$$

We now show that the first, third, and fourth terms in (5.13) tend to 0, as $T \to \infty$, uniformly in w. From (5.9) and (5.11), it follows that, uniformly in w,

$$\lim_{T\to\infty} E \left| \frac{1}{B_T^q} \int_{-T}^{T} dve^{-ivw} k(B_T v) b_T(v) \right|^2 = 0.$$

Next, if M is an upper bound for $\mid k(v) \mid$, the third term in (5.13) tends to 0 by (5.7) and (5.9) if $q \geqq 1$, and if $q < 1$ it is in absolute value less than

$$\frac{M}{(TB_T)^q} \frac{1}{T^{1-q}} \int_{-T}^{T} dv \mid v \mid \mid R(v) \mid,$$

which tends to 0 by (5.7) and Lemma 4. Similarly, the fourth term in (5.13), which is in absolute value less than

$$\frac{1}{(TB_T)^q} T^q \int_{|v| \geq T} dv \mid R(v) \mid,$$

tends to 0.

Consequently, (5.10) is proved for the case that the kernel has characteristic exponent $r = q$, since the second term in (5.13) then tends, uniformly in w, to $-2\pi\, k^{(q)} f^{(q)}(w)$. Next, to prove (5.10) for the case that $r > q$, it suffices to show that then

$$\lim_{T\to\infty} B_T^{-q} \int_{-T}^{T} \mid 1 - k(B_T v) \mid \mid R(v) \mid dv = 0. \tag{5.14}$$

Let M, M_1, and D be positive constants such that $\mid k(v) \mid \leq M$ for all v, and

$$\mid 1 - k(v) \mid \leq M_1 \mid v \mid^r \qquad \text{for } \mid v \mid \leq D. \tag{5.15}$$

If the characteristic exponent is infinite, we may take any exponent $r > q$ in (5.15). Let $s = r - q$. Then the quantity in (5.14) whose limit is being taken is less than

$$M_1 B_T^s \int_{|v| \leq DB_T^{-1}} \mid v \mid^{q+s} R(v)\, dv + MB_T^{-q} \int_{|v| \geq DB_T^{-1}} R(v)\, dv,$$

which tends to 0 in view of (5.7) and Lemma 4.

We next establish Theorem 5A. In view of the foregoing, it follows that the desired asymptotic covariance in (5.6) is given by the limit, as $T \to \infty$, of

$$\frac{4}{4\pi^2} TB_T \int_0^T \int_0^T dv_1 dv_2 \cos w_1 v_1 \cos w_2 v_2 k(B_T v_1) k(B_T v_2)\, ED_T(v_1) D_T(v_2). \tag{5.16}$$

We may write (5.16) as a sum of three 3-fold integrals, by replacing $TED_T(v_1) D_T(v_2)$ by its value (3.10). The term in this sum which involves $Q(v_1, u, u + v_2)$ clearly vanishes in the limit, uniformly in w_1 and w_2.

Next we show that the term involving $R(u + v_1)\, R(u - v_2)$ also vanishes in the limit, uniformly in w_1 and w_2. For this term is less than

$$B_T \int_0^T dv_2 \int_0^T dv_1 \int_{-T}^{T} du \mid k(B_T v_1) k(B_T v_2) R(u + v_1) R(u - v_2) \mid$$

Making the change of variable $v_1 = z_1 - v_2$, $u = z + v_2$, this becomes

$$B_T \int_0^T dv_2 \int_{v_2}^{T+v_2} dz_1 \int_{-T-v_2}^{T-v_2} dz \mid k(B_T v_2) k(B_T v_2 - B_T z_1) R(z) R(z + z_1) \mid.$$

Making the change of variable $z_2 = B_T v_2$, this becomes

$$\int_0^{B_T T} dz_2 \int_{(z_2/B_T)}^{T+(z_2/B_T)} dz_1 \int_{-T[1+(z_2/TB_T)]}^{T[1-(z_2/TB_T)]} dz \mid k(z_2) k(z_2 - B_T z_1) R(z) R(z + z_1) \mid,$$

which tends to 0 as $T \to \infty$, since the region of integration over the z_1 variable tends to infinity.

The value of (5.6) is then given by the limit of

$$\frac{1}{\pi^2} B_T \int_0^T \int_0^T dv_1 dv_2 \cos w_1 v_1 \cos w_2 v_2 k(B_T v_1) k(B_T v_2) \int_{-T}^{T} du U_T(u, v_1, v_2) R(u) R(u + v_1 - v_2). \tag{5.17}$$

By the change of variables $u_1 = v_1 - v_2$, $u_2 = v_2$ this becomes

$$\frac{1}{\pi^2} B_T \int_0^T du_2 \int_{-u_2}^{T-u_2} du_1 \cos w_1(u_1 + u_2) \cos w_2 u_2 \; k(B_T u_2) k(B_T u_2 + B_T u_1) \int_{-T}^{T} du U_T(u, u_1 + u_2, u_2) R(u) R(u + u_1). \tag{5.18}$$

By the change of variable $z = B_T u_2$, and the formula $2 \cos A \cos B = \cos (A + B) + \cos (A - B)$, one obtains that (5.18) is equal to

$$\frac{1}{2\pi^2} \int_0^{B_T T} dz \int_{-(z/B_T)}^{T[1-(z/TB_T)]} du_1 \left\{ \cos \left[z \left(\frac{w_1 - w_2}{B_T} \right) + u_1 w_1 \right] + \cos \left[z \left(\frac{w_1 + w_2}{B_T} \right) + u_1 w_1 \right] \right\} k(z) k(z + B_T u_1) \int_{-T}^{T} du U_T\left(u, u_1 + \frac{z}{B_T}, \frac{z}{B_T}\right) R(u) R(u + u_1). \tag{5.19}$$

By referring to (3.11), it may be verified that, as

$$T \to \infty, \qquad U_T\left(u, u_1 + \frac{z}{B_T}, \frac{z}{B_T}\right) \to 1.$$

Now to evaluate (5.19), one may distinguish three cases: case I, $w_1 \neq w_2$; case II, $w_1 = w_2 = w \neq 0$; case III, $w_1 = w_2 = 0$. In view of the Riemann-Lebesque Lemma, the first term in (5.19) vanishes in the limit if $w_1 - w_2 \neq 0$, and the second term vanishes in the limit if $w_1 + w_2 \neq 0$. Further, for any $\epsilon > 0$, the convergence to 0 is uniform in w_1 and w_2 such that $w_1 \geqq \epsilon$ and $w_2 \geqq \epsilon$. Thus one obtains that, in the limit, the value of (5.19) is 0 in case I; in case II, it is equal to

$$\frac{1}{2\pi^2} \int_0^\infty k^2(z)\, dz \int_{-\infty}^{\infty} du_1 \cos wu_1 \int_{-\infty}^{\infty} du\, R(u) R(u + u_1); \tag{5.20}$$

and, in case III, it is equal to twice (5.20). It is easily verified that (5.20) and (5.6) are equal.

To adapt the foregoing argument to the discrete parameter case requires some care in the phase of the argument following (5.19). The integration in (5.19) involving the variable z should be replaced by a summation over the lattice points $z_j = jB_T$, where $j = 1, \cdots, T$. As $T \to \infty$, the distance between the lattice points tends to 0, and the highest lattice point tends to infinity, so that the summation may be regarded as approaching the integral $\int_0^\infty k^2(z)\,dz$, as above.

6. Optimum consistent estimates of the spectral density. In view of Theorems 5A and 5B, the means are now at hand for choosing that estimate $f_T^*(w)$, of the form of (5.4), which is optimum in the sense that it possesses an order of consistency not less than that of any other such estimate. We obtain the following theorem.

THEOREM 6: *Suppose that* (5.7) *holds. Let the constants B_T be chosen so that, for some finite positive number B,*

$$\lim_{T\to\infty} T^{(1/1+2q)} B_T = B. \tag{6.1}$$

Let

$$\alpha = \frac{q}{1+2q}. \tag{6.2}$$

Then for any covariance averaging kernel $k(v)$ with characteristic exponent $r \geqq q$, the corresponding estimate $f_T^*(w)$ possesses an asymptotic mean square error given by

$$\lim_{T\to\infty} T^{2\alpha} E\,|\,f_T^*(w) - f(w)\,|^2 = \frac{f^2(w)}{B}\int k^2(z)\,dz\{1 + \delta(0, w)\} + B^{2q}\,|\,k^{(q)} f^{(q)}(w)\,|^2. \tag{6.3}$$

REMARK. If $q < r$, then $k^{(q)} = 0$.

Now, as q increases, the exponent α, as defined by (6.2), increases from 0 to $\frac{1}{2}$. Thus the factor which determines the highest order of consistency which may be attained, is the largest positive number q such that (5.7) holds. For want of a better name, we call this largest q the *exponent of uncorrelation* of the time series whose covariance function is $R(v)$, since the larger q is, the faster $R(v)$ falls off as $v \to \infty$, and the less correlated are successive observations of the time series. If (5.7) holds for all finite values of q, as is the case if $R(v)$ decreases exponentially, we define the exponent of uncorrelation to be infinite.

For computational convenience, the kernel with characteristic exponent r that we prefer is

$$k_r(z) = \begin{cases} 1 - |\,z\,|^r & \text{if } |\,z\,| < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.4}$$

Such a kernel leads to an estimate which does not require the computation of all the sample covariances. With this choice of kernel, $f_T^*(w)$ may be written letting $M_T = (1/B_T) \leqq T$,

$$f_T^*(w) = \frac{1}{2\pi}\sum_{|v| \leqq M_T} e^{-ivw}\left\{1 - \left(\frac{|\,v\,|}{M_T}\right)^r\right\} R_T(v). \tag{6.5}$$

The foregoing results may be interpreted from two points of view, emphasizing either the choice of kernel $k(z)$ or the choice of constants $M_T = 1/B_T$ (which, in the case of a kernel vanishing for $|z| > 1$, represent the number of sample covariances included in the estimate).

Let a kernel $k(z)$ be chosen whose characteristic exponent is r. Then the order of consistency of the corresponding estimate cannot be greater than $T^{2\alpha(r)}$, where $\alpha(r) = r/(1 + 2r)$, and will be $T^{2\alpha}$, where $\alpha \leqq \alpha(r)$, if the constants M_T satisfy the relation for some finite positive number M,

$$\lim_{T\to\infty} \frac{M_T}{T^{1-2\alpha}} = M \tag{6.6}$$

and if (5.7) holds for $q = \alpha/(1 - 2\alpha)$.

Therefore, if Bartlett's modified periodogram (which is (6.5) with $r = 1$) is used as the estimate, its order of consistency cannot be greater than $T^{2/3}$, and will be $T^{2\alpha}$ (where $\alpha \leqq \frac{1}{3}$) if the number of sample covariances included in the estimate is $T^{1-2\alpha}$. If the truncated periodogram (which is (6.5) with $r = \infty$) is used as the estimate, its order of consistency will be $T^{2\alpha}$ (where $\alpha < \frac{1}{2}$) if $M_T = T^{1-2\alpha}$, *and* if (5.7) holds for $q = \alpha/(1 - 2\alpha)$, which would be the case if the exponent of uncorrelation is infinite.

On the other hand, let the constants M_T be chosen so that (6.6) holds for some α between 0 and $\frac{1}{2}$. Then the order of consistency of the corresponding estimate $f_T^*(w)$ is $T^{2\alpha}$, no matter what the value of the characteristic exponent r of the kernel used so long as $r \geqq q(\alpha) = \alpha/(1 - 2\alpha)$, and (5.7) holds for $q = q(\alpha)$.

7. Alternative estimates of the spectral averages. In our study of the consistent estimates of the spectral density, we were led to consider estimates, such as Bartlett's modified periodogram, which had the property of only requiring the computation, on the basis of an observed sample of length T, of the sample covariances $R_T(v)$ for $|v|$ less than some root of T. In this section we show that for the spectral averages $J(A)$, one may define estimates $J_T^*(A)$, alternative to the previously given estimates $J_T(A)$, which have the same order of consistency and asymptotic variance as the latter, and require the computation of fewer sample covariances.

Let $A(w)$ be a spectral averaging function, with Fourier transform $a(v)$. Let $k(z)$ be a covariance averaging kernel, with Fourier transform $K(w)$. Let B_T be a sequence of constants tending to 0. Let $f_T^*(w)$ be defined by (5.4). Define

$$J_T^*(A) = \int f_T^*(w)A(w)\, dw. \tag{7.1}$$

One may write $J_T^*(A)$ in terms of the sample spectral density function by

$$J_T^*(A) = \int f_T(w)A_T(w)\, dw, \tag{7.2}$$

where

$$A_T(w) = \frac{1}{B_T}\int K\left(\frac{w-\lambda}{B_T}\right)A(\lambda)\, d\lambda.$$

In terms of the sample covariance functions, one may write

$$J_T^*(A) = \int_{-T}^{T} a(v)k(B_T v)R_T(v)\,dv \tag{7.3c}$$

$$= \sum_{|v| \leqq T} a(v)k(B_T v)R_T(v). \tag{7.3d}$$

The properties of the estimate $J_T^*(A)$ are embodied in the following two theorems, whose proofs are omitted.

THEOREM 7A. *For any two spectral averaging functions* $A_1(w)$ *and* $A_2(w)$, *the covariance* $\operatorname{Cov}[J_T^*(A_1), J_T^*(A_2)]$ *satisfies* (4.11).

THEOREM 7B. *Let* $a(v)$ *be a covariance averaging function. Let* $q > \frac{1}{2}$ *be such that*

$$\int | v |^q \, | a(v)R(v) | \, dv < \infty. \tag{7.4}$$

Let $k(z)$ *be a covariance averaging kernel with characteristic exponent* $r \geqq q$. *Let the positive constants* B_T *be chosen so that*

$$\limsup_{T\to\infty} T^{1/2}B_T^q \begin{cases} = 0 & \text{if} \quad r = q, \\ < \infty & \text{if} \quad r > q. \end{cases} \tag{7.5}$$

Then the bias $E\, J_T^*(A) - J(A)$ *satisfies* (4.10).

Optimum Estimates: The estimates $J_T^*(A)$ are all equivalent from the point of view of their order of consistency and asymptotic variance. If one desires to choose between them, the only basis is computational convenience. It is with this in mind that the following remarks are made. For the covariance averaging kernel, we choose $k_r(z)$. Then (7.3d) becomes, letting $M_T = (1/B_T)$

$$J_T^*(A) = \sum_{|v| \leqq M_T} a(v) \left\{ 1 - \left(\frac{| v |}{M_T} \right)^r \right\} R_T(v). \tag{7.7}$$

We choose B_T to be of the form $B_T = T^{-m}$, where the positive exponent m is to be chosen as small as possible, so that the number of terms in (7.7) will be as small as possible. Let q be the largest positive number such that (7.4) holds. Assuming q to be finite, choose $r \geqq q$. Then $J_T^*(A)$ will give a consistent estimate of $J(A)$, involving the calculation of a minimum number of sample covariances, if m is chosen as near to the lower bound as possible in the inequalities

$$\begin{aligned} m &> \frac{1}{2q} \qquad \text{if } r = q, \\ m &\geqq \frac{1}{2q} \qquad \text{if } r > q. \end{aligned} \tag{7.8}$$

8. Interpolating the spectral density. In order to obtain an estimate of the complete graph of the spectral density function $f(w)$ by means of the estimates $f_T^*(w)$ discussed in the foregoing, one needs to compute the estimate at all values of w. In this section, estimates $f_T^{**}(w)$ are constructed, which are equivalent to $f_T^*(w)$ from the point of view of order of consistency and asymptotic variance, and which require the computation of only a finite number of quantities in order to obtain the entire graph. Only the discrete parameter case is discussed in detail.

To begin with, define

$$w_m(T) = \frac{2\pi m}{2T+1} \qquad \text{for } m = 0, \pm 1, \cdots, \pm T, \tag{8.1}$$
$$= 0 \qquad \text{otherwise.}$$

We now show that $f_T^*(w)$, as defined by (5.4), may be expressed in terms of its values at the above $(2T + 1)$ lattice points by the formula

$$f_T^*(w) = \sum_{m=-T}^{T} c_m(w; T) f_T^*(w_m(T)), \tag{8.2}$$

where

$$c_m(w; T) = \frac{1}{2T+1} \sum_{v=-T}^{T} \exp\left[-iv(w - w_m(T))\right] \tag{8.3}$$
$$= \frac{\sin\left[(1/2)(2T+1)(w - w_m(T))\right]}{(2T+1)\sin\left[(1/2)(w - w_m(T))\right]}$$

To prove (8.2), we note that, for $v = 0, \pm 1, \cdots, \pm T$ and any w,

$$e^{-iwv} = \sum_{m=-T}^{T} c_m(w; T) \exp\left[-ivw_m(T)\right], \tag{8.4}$$

which may be verified by expanding the right-hand side. It is now easy to obtain (8.2) by substituting (8.4) into (5.4).

If $f_T^*(w)$ is given by (6.5), then it is determined by its value at even a fewer number of points, namely the lattice points $w_m(M_T)$, since by the same argument as above we may write

$$f_T^*(w) = \sum_{|m| \leqq M_T} c_m(w; M_T) f_T^*(w_m(M_T)). \tag{8.5}$$

Thus it is seen that it suffices to compute $f_T^*(w)$ at a finite number of points in order to know it on the entire interval $0 \leqq w \leqq \pi$. In view of the peaked nature of the $c_m(w; T)$ functions for large T, it might be thought that an adequate approximation to $f_T^*(w)$ would be $f_T^*(w_m(T))$, where $w_m(T)$ is the lattice point nearest to w. The problem which is raised by the representations (8.2) and (8.5) is when is such an approximation valid. From a statistical point of view, the justification of such an approximation must be in terms of its providing an estimate which has the proper order of consistency and asymptotic variance. It is from this point of view that we now consider the problem of using the value of $f_T^*(w)$ at a finite number of points to obtain estimates of its value at all points.

Let d_T be a sequence of constants tending to 0, and define the function, for $w \geqq 0$,

$$W_T(w) = \left[\frac{w}{d_T}\right] \cdot d_T, \tag{8.6}$$

where $[x]$ denotes the largest integer smaller than x. Consider the following estimate of the spectral density function

$$f_T^{**}(w) = f_T^*(W_T(w)), \tag{8.7}$$

where $f_T^*(w)$ is defined by (5.4). This estimate clearly has only a finite number of distinct values. The properties of this estimate are embodied in the following theorem.

THEOREM 8: *Assume that the conditions of Theorem 6 are fulfilled, so that the estimate $f_T^*(w)$ is consistent of order $T^{2\alpha}$, with asymptotic variance given by (6.3). Let*

$$\beta = \frac{\alpha}{q} = \frac{1}{1+2q} = 1 - 2\alpha. \tag{8.8}$$

Let the positive constants d_T be chosen such that

$$\limsup_{T\to\infty} T^\beta\, d_T < \infty \quad \text{if } 0 < q < 1, \quad \text{whence } 0 < \alpha < \tfrac{1}{3}, \tag{8.9}$$

$$\lim_{T\to\infty} T^\alpha\, d_T = 0 \quad \text{if } q \geqq 1, \quad \text{whence } \tfrac{1}{3} \leqq \alpha < \tfrac{1}{2}. \tag{8.10}$$

Then the estimate $f_T^{**}(w)$ is consistent of order $T^{2\alpha}$, with the same asymptotic bias and asymptotic variance as $f_T^*(w)$.

PROOF. One may suppose $w > 0$, since $f_T^{**}(0) = f_T^*(0)$. Now $w - d_T \leqq W_T(w) \leqq w$, so that $W_T(w) \to w$. In view of the uniform convergence in (5.6), it follows that the asymptotic variance of $f_T^{**}(w)$ is the same as that of $f_T^*(w)$. Next, in view of the uniform convergence in (5.10), to establish that $f_T^{**}(w)$ has the same asymptotic bias as that of $f_T^*(w)$ it suffices to show that

$$\lim_{T\to\infty} T^\alpha \,|\, f(W_T(w)) - f(w) \,|\, = 0. \tag{8.11}$$

Now the quantity in (8.11) whose limit is being taken is less than

$$T^\alpha d_T \sum_{|v| \leqq T^\beta} |\, v \,|\, R(v) + 2T^\alpha \sum_{|v| \geqq T^\beta} |\, R(v) \,|\, .$$

The second term tends to 0, since it is less than

$$2T^{\alpha - q} \sum_{|v| \geqq T^\beta} |v|^{\,q} \,|\, R(v) \,|\, .$$

The first term also tends to 0; by (8.10) if $q \geqq 1$, and by Lemma 4 and (8.9), if $q < 1$, since the term may be written

$$T^{\alpha - \beta q} T^\beta d_T \left(\frac{1}{T^\beta}\right)^{1-q} \sum_{|v| \leqq T^\beta} |\, v \,|\, |\, R(v) \,|\, .$$

If d_T is chosen by

$$d_T = \pi T^{-1/2}, \tag{8.12}$$

then (8.9) and (8.10) will be satisfied for $\frac{1}{4} \leqq \alpha < \frac{1}{2}$ (which corresponds to $\frac{1}{2} \leqq q < \infty$). It would seem that (8.12) provides a safe universal choice of the spacing of the lattice points.

If it is desired that the estimate of the spectral density be a continuous function, without jumps, then one may use the estimate

$$f_T^{**}(w) = a_T f_T^*(W_T(w)) + (1 - a_T) f_T^*(W_T(w) + d_T), \tag{8.13}$$

where a_T is a sequence of constants between 0 and 1, not in general approaching a limit. It may be verified that Theorem 8 holds for the estimate given by (8.13), provided that in (8.9) it is required that the limit be 0. Then it follows that $(d_T/B_T) \to 0$, and

$$\lim_{T\to\infty} T^{2\alpha} \operatorname{Cov}[f_T^*(W_T(w)), f_T^*(W_T(w) + d_T)] = \lim_{T\to\infty} T^{2\alpha} \operatorname{Var}[f_T^*(w)].$$

REFERENCES

[1] M. S. Bartlett, "Periodogram analysis and continuous spectra," *Biometrika*, Vol. 37 (1950), pp. 1–16.

[2] M. S. Bartlett, *An Introduction to Stochastic Processes*, Cambridge, 1955.

[3] U. Grenander, "On empirical spectral analysis of stochastic processes," *Ark. Mat.*, Vol. 1 (1951), pp. 503–531.

[4] U. Grenander and M. Rosenblatt, "Statistical spectral analysis of time series arising from stationary stochastic processes," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 537–559.

[5] U. Grenander and M. Rosenblatt, "Regression analysis of time series with stationary residuals," *Proc. Nat. Acad. Sci.*, Vol. 40 (1954), pp. 812–816.

[6] T. A. Magness, "Spectral response of a quadratic device to non-Gaussian noise," *J. Appl. Phys.*, Vol. 25 (1954), pp. 1357–1365.

[7] J. Tukey, "Measuring Noise Color," an unpublished manuscript prepared for distribution at a meeting of the Institute of Radio Engineers, November 7, 1951.

[8] E. Parzen, *Optimum Methods of Spectral Analysis of Finite Noise Samples*, Columbia University, Hudson Laboratories Technical Report No. 39, issued Summer 1956.

[9] J. L. Doob, *Stochastic Processes*, New York, 1953.

[10] M. Loève, *Probability Theory*, New York, 1955.