*The Annals of Statistics*
1980, Vol. 8, No. 5, 1100–1109

# DISCRETE-TIME SPECTRAL ESTIMATION OF CONTINUOUS-TIME PROCESSES— THE ORTHOGONAL SERIES METHOD[1]

BY E. MASRY

*University of California, San Diego*

Let $\{X(t), -\infty < t < \infty\}$ be a stationary time series with spectral density function $\phi(\lambda)$. Let $\{t_n\}$ be a stationary Poisson point process on the real line. The existence of consistent estimates of $\phi(\lambda)$ based on the discrete-time observations $\{X(t_n)\}_{n=1}^{N}$, when the actual sampling times are *not* known, has been an open question (Beutler). Using an orthogonal series method, a class of spectral estimates is considered and its uniform and integrated uniform consistency in quadratic mean is established. Rates of convergence are established and are compared with the optimal rates of the available (Brillinger, Masry) kernel-type estimates based on the observations $\{X(t_n), t_n\}_{n=1}^{N}$.

**1. Introduction.** There is an extensive literature on the subject of spectral estimation of continuous and discrete parameters time series. Here we are concerned with the estimation of the spectral density function $\phi(\lambda)$ of a time series $X = \{X(t), -\infty < t < \infty\}$ based on the discrete-time observations $\{X(t_n), t_n\}$, where the sampling process $\{t_n\}$ constitutes a stationary point process on the real line. Brillinger (1972), in his fundamental work on the spectral analysis of stationary interval functions, discussed the consistency and asymptotic normality as $T \to \infty$ of kernel-type spectral estimates $\phi_T(\lambda)$ based on the modified periodogram

$$(1.1)\qquad I_T(\lambda) = \frac{1}{2\pi T}\left\{\left|\int_0^T e^{-it\lambda}X(t)dN(t)\right|^2 - \int_0^T X^2(t)dN(t)\right\}$$

where $N(t) = N((0, t])$ is the counting process associated with $\{t_n\}$. Considering $\phi_T(\lambda)$ as an estimate based on discrete-time observations, it clearly employs a random sample size $N(T)$. Assuming a deterministic sample size $N$, kernel-type spectral estimates $\hat{\psi}_N(\lambda)$ based on the observations $\{X(t_n), t_n\}_{n=1}^{N}$, where $N$ is a positive integer and $\{t_n\}$ is a stationary Poisson point process, have been considered by Masry (1978) where their pointwise consistency in quadratic mean as $N \to \infty$ is derived; $\hat{\psi}_N(\lambda)$ is based on the modified periodogram

$$(1.2)\qquad J_N(\lambda) = \frac{1}{2\pi N}\left\{\left|\sum_{n=1}^{N} e^{-it_n\lambda}X(t_n)\right|^2 - \sum_{n=1}^{N} X^2(t_n)\right\}.$$

For a Poisson sampling process $\{t_n\}$, the second order statistics of $J_N(\lambda)$ and $I_T(\lambda)$ are distinct but the second order statistics of the corresponding kernel-type estimates $\hat{\psi}_N(\lambda)$ and $\phi_T(\lambda)$ are virtually identical, as shown in [5].

---

Received October 1978; revised March 1979.

[1]This research was supported by the Office of Naval Research under Contract N00014-75-C-0652.

*AMS* 1970 *subject classifications.* Primary 62M15.

*Key words and phrases.* Spectral estimation, orthogonal series estimates, convergence rates.

The above estimates require the knowledge of the actual values of the sampling instants $\{t_n\}_{n=1}^N$ (to evaluate $e^{it_n\lambda}$) in addition to the data sequence $\{X(t_n)\}_{n=1}^N$. This paper investigates the estimation of $\phi(\lambda)$ on the basis of the data sequence $\{X(t_n)\}_{n=1}^N$ alone, i.e., when the actual sampling instants are *not* known. (Such a situation may arise in randomly-sampled data systems where the actual sampling times are not transmitted. See also Gaster and Roberts (1975) for examples in fluid mechanics.) Clearly, the periodogram approach (1.1)-(1.2) is no longer applicable. In fact, the existence of consistent estimates of $\phi(\lambda)$ under these circumstances has been an open question as noted in Beutler (1970). This paper resolves this question when the sampling process $\{t_n\}$ is Poisson: By employing an orthogonal series method, a class of spectral estimates $\hat{\phi}_N(\lambda)$ is introduced and its uniform consistency and integrated uniform consistency in quadratic mean is established. The approach is similar to the one used by Čencov (1962), Schwartz (1967), Watson (1969) and Rosenblatt (1971) for probability density estimation. However, unlike these works, the complete orthogonal set in $\mathfrak{L}_2$ used for the representation of $\phi(\lambda)$, cannot be arbitrary and is, in fact, generated by the statistics of the sampling process $\{t_n\}$.

In Section 2, the series representation for $\phi(\lambda)$ is introduced and some preliminary results are given. Conditions for the consistency of $\hat{\phi}_N(\lambda)$ and bounds on the two types of errors are established in Section 3. These results are compared in Section 4 with the optimal convergence properties of the estimates based on the periodogram (1.1) or (1.2). The comparison indicates that the orthogonal series estimate $\hat{\phi}_N(\lambda)$ has appreciably slower rates of convergence. The proofs are collected in Section 5. The question of whether these rates are the best possible remains open.

**2. Preliminaries.** Throughout this paper, $X = \{X(t), -\infty < t < \infty\}$ is a real stationary measurable fourth order process with mean zero, continuous covariance function $C(t) \in \mathfrak{L}_1$, spectral density $\phi(\lambda) = (2\pi)^{-1}\int_{-\infty}^{\infty}C(t)\exp[-it\lambda]dt$ and fourth order cumulant function $Q(u_1, u_2, u_3,)$. The sampling process $\{t_n\}$ is assumed to be a stationary Poisson point process on $[0, \infty)$, independent of $X$, with known mean intensity $\beta$, i.e., $t_0 = 0$ a.s., $t_n = t_{n-1} + T_n$, $n = 1, 2, \cdots$, where the $T_n$'s are independent identically distributed random variables with exponential distribution $F(x) = [1 - \exp(-\beta x)]$.

The basic idea of the paper is to employ an appropriate orthogonal series representation for $\phi(\lambda)$ whose coefficients can be estimated from the data sequence $\{X(t_n)\}_{n=1}^N$ *alone*. Note that if $\{U_n(\lambda)\}_{n=1}^\infty$ is an arbitrary complete orthonormal set in $\mathfrak{L}_2(-\infty, \infty)$, then $\phi(\lambda) = \Sigma_{n=1}^\infty b_n U_n(\lambda)$ in $\mathfrak{L}_2$, $b_n = \int_{-\infty}^{\infty}\phi(\lambda)U_n^*(\lambda)d\lambda$ and, as in probability density estimation by orthogonal series [3], [9], [12], one could estimate $b_n$ by $\hat{b}_n(N) = (1/\beta)\int_{-\infty}^{\infty}J_N(\lambda)U_n^*(\lambda)d\lambda$, where $J_N(\lambda)$ is given in (1.2). Then

$$\hat{b}_n(N) = (1/2\pi\beta N)\Sigma_{j,k=1;j\neq k}^N X(t_j)X(t_k)u_n^*(t_j - t_k)$$

where $u_n(t) = \int_{-\infty}^{\infty}\exp[it\lambda]U_n(\lambda)d\lambda$. Thus, $\hat{b}_n(N)$ requires the knowledge of the sampling instants $\{t_n\}_{n=1}^N$ in addition to the data sequence $\{X(t_k)\}_{k=1}^N$. Hence, the basis $\{U_n(\lambda)\}$ cannot be arbitrary for our purposes. Next we note (Shapiro and Silverman (1960)) that the discrete-parameter process $\{X(t_k)\}$ has mean zero and covariance sequence $\{c_n\}$

$$c_n = E[X(t_{k+n})X(t_k)] = \int_0^\infty C(t)f_n(t)dt, \qquad n = 1,2,\cdots, \tag{2.1}$$

where $f_n(t) = \beta[(\beta t)^{n-1}/(n-1)!]\exp(-\beta t)$, $n = 1,2,\cdots$. The set $\{f_n(t)\}$ is complete in $\mathfrak{L}_2(0,\infty)$ and its orthonormalization yields

$$g_n(t) = (2\beta)^{\frac{1}{2}}L_{n-1}(2\beta t)e^{-\beta t}1_{[0,\infty)}(t), \qquad n = 1,2,\cdots, \tag{2.2}$$

where $L_n(t)$ is the $n$th Laguerre polynomial. Note that [10] $g_n(t) = \sum_{k=1}^n \theta_{n,k}f_k(t)$ where

$$\theta_{n,k} = (2/\beta)^{\frac{1}{2}}(-2)^{k-1}\binom{n-1}{k-1}. \tag{2.3}$$

Since $C(t) \in \mathfrak{L}_2(-\infty,\infty)$, we have the $\mathfrak{L}_2$ expansion $C(t) = \sum_{n=1}^\infty a_n g_n(|t|)$ where

$$a_n = \int_0^\infty C(t)g_n(t)dt = \sum_{k=1}^n \theta_{n,k}c_k. \tag{2.4}$$

Hence

$$\phi(\lambda) = \sum_{n=1}^\infty a_n G_n(\lambda) \tag{2.5}$$

in $\mathfrak{L}_2(-\infty,\infty)$ where, by direct Fourier transform of (2.2),

$$G_n(\lambda) = \frac{1}{2\pi}\int_{-\infty}^\infty e^{-it\lambda}g_n(|t|)dt = -\frac{(2\beta)^{\frac{1}{2}}}{\pi}\frac{\cos[(2n-1)\tan^{-1}(\lambda/\beta)]}{(\lambda^2+\beta^2)^{\frac{1}{2}}}, \qquad n = 1,2,\cdots. \tag{2.6}$$

$\{G_n(\lambda)\}_{n=1}^\infty$ is complete and orthogonal in $\mathfrak{L}_2(-\infty,\infty)$ with respect to even functions on $(-\infty,\infty)$. The series expansion (2.5) has been considered by Shapiro and Silverman (1960) in connection with a concept of "alias-free" sampling.

The approach: given the data sequence $\{X(t_n)\}_{n=1}^N$ estimate $c_n$ by

$$\begin{aligned} \hat{c}_n(N) &= \frac{1}{N}\sum_{k=1}^{N-n}X(t_{k+n})X(t_k), && 1 \leqslant n < N \\ &= 0, && N \leqslant n. \end{aligned} \tag{2.7}$$

Then, via (2.4), estimate the expansion coefficient $a_n$ by

$$\hat{a}_n(N) = \sum_{k=1}^n \theta_{n,k}\hat{c}_k(N), \tag{2.8}$$

and finally, estimate $\phi(\lambda)$ by

$$\hat{\phi}_N(\lambda) = \sum_{n=1}^\infty \gamma_n(N)\hat{a}_n(N)G_n(\lambda) \tag{2.9}$$

where $\{\gamma_n(N)\}$ is an appropriate averaging sequence to be specified below.

A bound on the rate of decay of $\{a_n\}$ is needed and is given below; $AC^r[0,\infty)$ denotes the set of functions which are $r$-times absolutely continuous on $[0,\infty)$.

LEMMA 2.1. *For* $t \geqslant 0$, *let* $C(t) \in AC^{r-1}[0, \infty)$ *such that*

$$t^{r/2}C^{(k)}(t) \in \mathfrak{L}_2(0, \infty) \text{ for } k = 0, 1, \cdots, r. \tag{2.10}$$

*Then for* $n = 1, 2, \cdots$

$$|a_n| \leqslant A_1(r)n^{-r/2}$$

*where*

$$A_1(r) = (2\beta)^{-\frac{1}{2}}\|t^{r/2}e^{t/2}\frac{d^r}{dt^r}[C(t/2\beta)e^{-t/2}]\|_{\mathfrak{L}_2(0,\infty)}.$$

REMARK. The hypothesis of Lemma 2.1 does not require the differentiability of $C(t)$ at the origin. Thus, the spectral moments of $\phi(\lambda)$ need not exist. For example, $C(t) = e^{-|t|}$ satisfies the hypothesis of Lemma 2.1 fòr every integer $r \geqslant 1$. Similarly, if $\phi(\lambda)$ is a rational function in $\lambda$, then Lemma 2.1 holds for every integer $r \geqslant 1$.

In investigating the convergence in quadratic mean of the estimate (2.9) for a fixed $\lambda$, the pointwise convergence of the $\mathfrak{L}_2$ expansion (2.5) is needed. We have

LEMMA 2.2. *Assume* $C(t)$ *satisfies the hypothesis of Lemma* 2.1 *for some* $r > 2$. *Then*

$$\phi(\lambda) = \Sigma_{n=1}^{\infty} a_n G_n(\lambda) \tag{2.11}$$

*uniformly on* $(-\infty, \infty)$.

**3. Consistency and rates of convergence.** We first consider the consistency of the estimates $\hat{c}_n(N)$ and $\hat{a}_n(N)$. The following assumption is needed.

ASSUMPTION 3.1. The process $X$ satisfies

(i) $tC(t) \in \mathfrak{L}_2(-\infty, \infty)$

(ii) $|Q(u_1, u_2, u_3)| \leqslant h(u_1, u_2, u_3)$

where $h$ is even and nonincreasing on $[0, \infty)$ in each variable such that $\int_{-\infty}^{\infty} h(0, u, 0)du < \infty$.

THEOREM 3.1. *Under Assumption* 3.1, *the estimate* $\hat{c}_n(N)$ *is consistent in quadratic mean with*

(i) $E[\hat{c}_n(N)] = (1 - (n/N))c_n$

(ii) $\mathrm{Var}[\hat{c}_n(N)] \leqslant A_2/N$

*where* $A_2$ *is a constant independent of* $n$ *and* $N$.

Before considering the consistency of the estimate $\hat{a}_n(N)$, we note that the mapping (2.4) of $\{c_n\}$ to $\{a_n\}$ is unbounded in $l_\infty$ since $\Sigma_{k=1}^n \theta_{n,k} = (2/\beta)^{\frac{1}{2}}(-1)^{n-1}$ but $\Sigma_{k=1}^n |\theta_{n,k}| = (2/\beta)^{\frac{1}{2}}3^{n-1}$. Thus, since $\mathrm{Var}[\hat{a}_n(N)] = \Sigma_{k,l=1}^n \theta_{n,k}\theta_{n,l}$ $\mathrm{Cov}[\hat{c}_k(N), \hat{c}_l(N)]$, a small variability in the estimate $\hat{c}_n(N)$ is likely to produce a large variability in the estimate $\hat{a}_n(N)$ for large $n$. A simulation study in [4] appears to confirm this observation. We have

THEOREM 3.2. *Under Assumption* 3.1 *the estimate* $\hat{a}_n(N)$ *is consistent in quadratic mean with*

(i) $E[\hat{a}_n(N)] = a_n - (1/N)[na_n - (n-1)a_{n-1}]$
(ii) $\mathrm{Var}[\hat{a}_n(N)] \leqslant A_3(3^{2n}/N)$
*where* $A_3 = 2A_2/9\beta$.

Consider now the class of spectral estimates $\hat{\phi}_N(\lambda)$ defined by (2.9), where the averaging sequence $\{\gamma_n(N)\}$ is of the form

$$\gamma_n(N) = h(e^{\alpha n}/N^b), \quad \alpha > \ln 3, 0 < b < \frac{\alpha}{2\ln 3} \tag{3.1}$$

and $h(u)$ is a real even function on the real line satisfying

$$\begin{aligned} &\text{(i) } |h(u)| \leqslant h(0) = 1 && \text{for all } u, \\ &\text{(ii) } |uh(u)| \leqslant K_1 && \text{for } |u| \geqslant 1, \\ &\text{(iii) } 1 - h(u) \leqslant K_2|u| && \text{for all } u. \end{aligned} \tag{3.2}$$

The kernel $\{\gamma_n(N)\}$ is then of the exponential type as defined by Parzen (1958). The choice of exponential-type kernel (rather than algebraic) and the conditions on the parameters $\alpha$ and $b$ are necessitated by the variance expression of $\hat{a}_n(N)$.

THEOREM 3.3. *Assume X satisfies Assumption* 3.1 *and the hypothesis of Lemma* 2.1 *for some* $r > 1$. *Then the estimate* (2.9) *is integratedly uniformly consistent in quadratic mean with*

$$E\int_{-\infty}^{\infty}|\hat{\phi}_N(\lambda) - \phi(\lambda)|^2 d\lambda \leqslant B_1[1/\ln N]^{r-1}(1 + o(1))$$

*where*

$$B_1 = 4A_1^2(r)(2\alpha/b)^{r-1}/(r-1).$$

Next we have

THEOREM 3.4. *Assume X satisfies Assumption* 3.1 *and the hypothesis of Lemma* 2.1 *for some* $r > 2$. *Then the estimate* (2.9) *is uniformly consistent in quadratic mean with*

$$E|\hat{\phi}_N(\lambda) - \phi(\lambda)|^2 \leqslant B_2[1/\ln N]^{r-2}(1 + o(1))$$

*uniformly in* $\lambda$, *where*

$$B_2 = (32/\pi^2\beta)A_1^2(r)(2\alpha/b)^{r-2}(r-2)^{-2}.$$

**4. Discussion.** We first establish the optimal convergence properties of the kernel-type estimate $\hat{\psi}_N(\lambda)$ based on the periodogram (1.2). $\hat{\psi}_N(\lambda)$ is of the form [2][5]

$$\hat{\psi}_N(\lambda) = (1/\beta)\int_{-\infty}^{\infty}W_N(\lambda - \mu)J_N(\mu)d\mu \tag{4.1}$$

where $W_N(\lambda) = M_N K(M_N\lambda)$, $M_N \to \infty$ and $M_N/N \to 0$ as $N \to \infty$ and the kernel $K(\lambda)$ is a real even continuously differentiable function on $(-\infty, \infty)$ satisfying

$$\begin{aligned} &\text{(i) } \sup_{-\infty<\lambda<\infty}|K(\lambda)| < \infty && \text{(ii) } \sup_{-\infty<\lambda<\infty}|K'(\lambda)| < \infty \\ &\text{(iii) } \int_{-\infty}^{\infty}|K(\lambda)|d\lambda < \infty && \text{(iv) } \int_{-\infty}^{\infty}K(\lambda)d\lambda = 1. \end{aligned} \tag{4.2}$$

Aside from the conditions on the fourth order cumulant $Q(u_1, u_2, u_3)$ we have by [5], Theorems 1 and 3 (or [2], Theorem 4.3), that if $tC(t) \in \mathcal{L}_1(-\infty, \infty)$ then

$$E\big[\hat{\psi}_N(\lambda)\big] = \int_{-\infty}^{\infty} W_N(\lambda - \mu)\phi(\mu)\,d\mu + O(1/N)$$

$$\operatorname{Var}\big[\hat{\psi}_N(\lambda)\big] = 2\pi\beta(M_N/N)\left[\phi(\lambda) + \frac{C(0)}{2\pi\beta}\right]^2(1 + \delta_{0,\lambda})\int_{-\infty}^{\infty} K^2(u)\,du \times (1 + O(M_N/N)).$$

To find the optimal convergence properties of $\hat{\psi}_N(\lambda)$ we proceed as in Wahba (1975). Assume that for some integer $r \geqslant 1$, $K(\lambda)$ satisfies in addition

$$\text{(4.2)}\qquad \begin{aligned} &\text{(v)}\ \int_{-\infty}^{\infty} \lambda^j K(\lambda)\,d\lambda = 0, \qquad j = 1, \cdots, r-1 \\ &\text{(vi)}\ \int_{-\infty}^{\infty} |\lambda|^r |K(\lambda)|\,d\lambda < \infty \end{aligned}$$

and that

$$\text{(4.3)}\qquad t^r C(t) \in \mathcal{L}_1(-\infty, \infty).$$

Expanding $\phi(\lambda + \mu)$ in a Taylor series, and using (4.2) (iv)–(vi) in $\int_{-\infty}^{\infty} K(-u)$ $\phi(\lambda + u/M_N)du$, we find $b^2[\hat{\psi}_N(\lambda)] \leqslant B_3 M_N^{-2r}(1 + o(1))$ where

$$B_3 = \frac{1}{[2\pi r!]^2}\Big[\int_{-\infty}^{\infty} |u|^r |K(u)|\,du\Big]^2\Big[\int_{-\infty}^{\infty} |t|^r |C(t)|\,dt\Big]^2.$$

Thus, ignoring a factor $(1 + O(M_N/N))$ in the variance and $(1 + o(1))$ in the bias, we have

$$\text{(4.4)}\qquad E|\hat{\psi}_N(\lambda) - \phi(\lambda)|^2 \leqslant B_3 M_N^{-2r} + B_4(M_N/N)$$

where

$$B_4 = 4\pi\beta\left[\int_{-\infty}^{\infty} |C(t)|\,dt + \frac{C(0)}{2\pi\beta}\right]^2 \int_{-\infty}^{\infty} K^2(u)\,du.$$

Then choosing $M_N = [2rB_3/B_4]^{1/(2r+1)} N^{1/(2r+1)}$, which minimizes the right-hand side of (4.4), we have finally

$$\text{(4.5)}\qquad E|\hat{\psi}_N(\lambda) - \phi(\lambda)|^2 \leqslant B_5 N^{-2r/(2r+1)}(1 + o(1))$$

where

$$B_5 = \frac{(2r+1)}{(2r)^{2r/(2r+1)}}\big[B_3 B_4^{2r}\big]^{1/(2r+1)}.$$

The parameter $r$ of (4.3) and Lemma 2.1 represents a degree of smoothness of $\phi(\lambda)$. For the same value of $r$, the hypothesis of Lemma 2.1 is generally more restrictive then (4.3), even though (4.3) implies (2.10) with $k = 0$. For the same value of $r$, the mean square error of the estimate $\hat{\psi}_N(\lambda)$ is $O(1/N^{2r/(2r+1)})$ whereas for the series estimate $\hat{\phi}_N(\lambda)$ it is $O(1/[\ln N]^{r-2})$ by Theorem 3.4. Thus, $\hat{\psi}_N(\lambda)$ has an appreciably higher rate of convergence—this at the expense of requiring a record of the sampling instants.

Comparison on the basis of mean integrated square error cannot be made since the kernel-type estimate (4.1) has not so far been shown to be consistent in this sense (see [2][5]).

The proofs of Theorems 3.3 and 3.4 indicate that the logarithmic convergence rates of the series estimates $\hat{\phi}_N(\lambda)$ are due to the exponential growth in $n$ of $\mathrm{Var}[\hat{a}_n(N)]$ as given by the bound in Theorem 3.2. The discussion preceding Theorem 3.2 provides evidence for such a rapid growth. It remains an open question whether the convergence rates of the series estimate obtained here are the *best* possible.

Finally, we note that when the sampling process $\{t_n\}$ is not necessarily Poisson but "alias-free" in the sense of [1], [10], an orthogonal series estimate of $\phi(\lambda)$ of the form (2.9) can be considered in a similar fashion. However, the analysis becomes more complex since the basis $\{g_n(t)\}_{n=1}^{\infty}$ is generated by the statistics of the point process $\{t_n\}$. (In the Poisson case, we have the Laguerre functions (2.2), whose properties are well known.)

## 5. Proofs.

Proof of Lemma 2.1. The following relationship for generalized Laguerre functions is easily verified.

$$t^{\nu-1}L_{n-1}^{\nu-1}(t) = \frac{d}{dt}\left[\frac{t^{\nu}}{(n+\nu-1)}L_{n-1}^{\nu}(t)\right]. \tag{5.1}$$

By repeated substitution of (5.1) and integration by parts in the integral below, we have

(5.2)

$$\int_0^A C(t/2\beta)L_{n-1}(t)e^{-t/2}dt = \sum_{k=0}^{r-1}\frac{(-1)^k t^{k+1}L_{n-1}^{k+1}(t)}{n(n+1)\cdots(n+k)}\frac{d^k}{dt^k}\left[C(t/2\beta)e^{-t/2}\right]\Bigg|_{0+}^{A}$$
$$+\frac{(-1)^r}{n(n+1)\cdots(n+r-1)}\int_0^A\left\{\frac{d^r}{dt^r}\left[C(t/2\beta)e^{-t/2}\right]\right\}t^r L_{n-1}^r(t)dt.$$

Each term in the sum vanishes at infinity by the hypothesis of Lemma 2.1. Finally, by

$$\|t^{r/2}L_n^r(t)e^{-t/2}\|_{\mathcal{L}_2(0,\infty)} = \left[\frac{(n+r)!}{n!}\right]^{\frac{1}{2}},$$

the dominated convergence theorem and the Cauchy-Schwarz inequality

$$|a_n| \leqslant A_1(r)\left[\frac{(n-1)!}{(n+r-1)!}\right]^{\frac{1}{2}} \leqslant \frac{A_1(r)}{n^{r/2}}.$$

Proof of Lemma 2.2. By (2.6) and Lemma 2.1, $\sum_{n=1}^{\infty}|a_n|\,|G_n(\lambda)| \leqslant (2/\pi^2\beta)^{\frac{1}{2}}\sum_{n=1}^{\infty}|a_n| < \infty$. Since $G_n(\lambda)$ is continuous, the series in (2.11) converges uniformly to an even continuous function, say, $\psi(\lambda)$. Since $\sum_{n=1}^{\infty}|a_n|^2 < \infty$,

$\psi(\lambda) \in \mathcal{L}_2(-\infty, \infty)$ by the Riesz-Fischer theorem and hence $\int_{-\infty}^{\infty}[\phi(\lambda) - \psi(\lambda)]G_n(\lambda)d\lambda = 0$, $n = 1, 2, \cdots$. Finally, since $\{G_n(\lambda)\}_{n=1}^{\infty}$ is complete in $\mathcal{L}_2(-\infty, \infty)$ with respect to even functions, we have $\phi(\lambda) = \psi(\lambda)$ a.e. and the result follows by the continuity of $\phi(\lambda)$ and $\psi(\lambda)$.

PROOF OF THEOREM 3.1. (i) is clear. For (ii), we have

$$\operatorname{Var}[\hat{c}_n(N)] = \Sigma_{j=1}^{4} T_{n,j}(N) - E^2[\hat{c}_n(N)]$$

where

$$T_{n,1}(N) = \frac{1}{N^2}\Sigma_{k,l=1}^{N-n} E[C(t_{k+n} - t_k)C(t_{l+n} - t_l)],$$

$$T_{n,2}(N) = \frac{1}{N^2}\Sigma_{k,l=1}^{N-n} E[C(t_l - t_k)C(t_{l+n} - t_{k+n})],$$

$$T_{n,3}(N) = \frac{1}{N^2}\Sigma_{k,l=1}^{N-n} E[C(t_{l+n} - t_k)C(t_{k+n} - t_l)],$$

$$T_{n,4}(N) = \frac{1}{N^2}\Sigma_{k,l=1}^{N-n} E[Q(t_{k+n} - t_k, t_l - t_k, t_{l+n} - t_k)].$$

By a bounding argument similar to the one employed in [5], it can be shown that if $tC(t) \in \mathcal{L}_2(-\infty, \infty)$ then $|T_{n,1}(N) - E^2[\hat{c}_n(N)]| \leqslant (A_{2,1})/N$, if $C(t) \in \mathcal{L}_1(-\infty, \infty)$ then $|T_{n,j}(N)| \leqslant (A_{2,j})/N$, $j = 2, 3$, and if $Q(u_1, u_2, u_3)$ satisfies Assumption 3.1, then $|T_{n,4}(N)| \leqslant A_{2,4}/N$, where $A_{2,j}$, $j = 1, 2, 3, 4$ are constants independent of $n$ and $N$. The result (ii) follows.

PROOF OF THEOREM 3.2. (i) By (2.8), Theorem 3.1 and (2.4), $E[\hat{a}_n(N)] = a_n - (1/N)\Sigma_{k=1}^{n} k\theta_{n,k}c_k$. Thus the bias $b[\hat{a}_n(N)]$ is

$$b[\hat{a}_n(N)] = -(1/N)\textstyle\int_0^{\infty} C(t)q_n(t)dt \tag{5.3}$$

where

$$q_n(t) = \Sigma_{k=1}^{n} k\theta_{n,k}f_k(t) = (2\beta)^{\frac{1}{2}}\Sigma_{k=0}^{n-1}\frac{(k+1)}{k!}\binom{n-1}{k}(-2\beta t)^k e^{-\beta t}.$$

Using the properties of the Laguerre functions [7, page 299] it can be shown that $q_n(t) = ng_n(t) - (n-1)g_{n-1}(t)$ and the result follows by (5.3).

(ii) We have $\operatorname{Var}^{\frac{1}{2}}[\hat{a}_n(N)] \leqslant \Sigma_{k=1}^{n}|\theta_{n,k}|\operatorname{Var}^{\frac{1}{2}}[\hat{c}_k(N)]$ and the result follows by Theorem 3.1 and $\Sigma_{k=1}^{n}|\theta_{n,k}| = (2/\beta)^{\frac{1}{2}}3^{n-1}$.

PROOF OF THEOREM 3.3. We have

$$E\textstyle\int_{-\infty}^{\infty}|\hat{\phi}_N(\lambda) - \phi(\lambda)|^2 d\lambda = \int_{-\infty}^{\infty} b^2[\hat{\phi}_N(\lambda)]d\lambda + \int_{-\infty}^{\infty}\operatorname{Var}[\hat{\phi}_N(\lambda)]d\lambda$$

where by Theorem 3.2

$$\textstyle\int_{-\infty}^{\infty} b^2[\hat{\phi}_N(\lambda)]d\lambda = \frac{1}{\pi}\Sigma_{n=1}^{\infty}\left\{a_n[1 - \gamma_n(N)] + \frac{\gamma_n(N)}{N}[na_n - (n-1)a_{n-1}]\right\}^2, \tag{5.4}$$

and

$$(5.5)\qquad \int_{-\infty}^{\infty}\mathrm{Var}\big[\hat{\phi}_N(\lambda)\big]d\lambda = \frac{1}{\pi}\Sigma_{n=1}^{\infty}\gamma_n^2(N)\mathrm{Var}\big[\hat{a}_n(N)\big].$$

The truncated sum in (5.5) at $M$, the integer part of $(b/\alpha)\ \ln N$, is $O(N^{-p})$, $p = 1 - (2b/\alpha)\ln 3$, by Theorem 3.2 and (3.2i); and by Theorem 3.2 and (3.2ii) the tail sum is also $O(N^{-p})$. In fact,

$$(5.6)\qquad \int_{-\infty}^{\infty}\mathrm{Var}\big[\hat{\phi}_N(\lambda)\big]d\lambda \leqslant D_1N^{-p};\ D_1 = \frac{A_3k_1}{\pi(1 - e^{-2(\alpha - \ln 3)})} + \frac{9A_3}{8\pi}.$$

Next (5.4) is bounded by $(1/\pi)(S_1 + S_2 + S_3)^2$ where

$$S_1^2 = \Sigma_{n=1}^{\infty}a_n^2\big[1 - \gamma_n(N)\big]^2,\quad S_2^2 = \frac{1}{N^2}\Sigma_{n=1}^{\infty}n^2\gamma_n^2(N)a_n^2$$

$$S_3^2 = \frac{1}{N^2}\Sigma_{n=1}^{\infty}n^2\gamma_{n+1}^2(N)a_n^2.$$

$S_1$ is the dominant term in the integrated bias, for with $M$ as before, we have by (3.2i) and (3.2ii) and an argument similar to the one above that $S_k^2 \leqslant (\frac{1}{2})(b/\alpha)^2\|C\|_{\ell_2}(\ln N/N)^2(1 + o(1))$, $k = 2, 3$. Now for any integer $m > 1$, we have by (3.2i), (3.2iii) and Lemma 2.1 that

$$S_1^2 \leqslant \left(\tfrac{1}{2}\right)K_2^2\|C\|_{\ell_2}^2e^{2\alpha m}N^{-2b} + \frac{4A_1^2(r)}{r - 1}\cdot\frac{1}{m^{r-1}}.$$

The optimal $m$ which minimizes the right-hand side is then the solution of a transcendental equation and cannot be found explicitly. However, $m$ is essentially logarithmic in $N$ and upon choosing $m - 1$ to be the integer part of $(b/2\alpha)\ln N$, we have

$$(5.7)\qquad \int_{-\infty}^{\infty}b^2\big[\hat{\phi}_N(\lambda)\big]d\lambda \leqslant \frac{D_2}{(\ln N)^{r-1}}(1 + o(1)),\ D_2 = 4A_1^2(r)\frac{(2\alpha/b)^{r-1}}{(r - 1)}.$$

The result follows by (5.6) and (5.7).

Proof of Theorem 3.4. By Lemma 2.2 and Theorem 3.2, the bias of $\hat{\phi}_N(\lambda)$ is

$$b\big[\hat{\phi}_N(\lambda)\big] = -\Sigma_{n=1}^{\infty}a_n\big[1 - \gamma_n(N)\big]G_n(\lambda) - (1/N)\Sigma_{n=1}^{\infty}\gamma_n(N)\big[na_n - (n - 1)a_{n-1}\big]G_n(\lambda)$$

and by (2.6) $b[\hat{\phi}_N(\lambda)] \leqslant (2/\pi^2\beta)^{\frac{1}{2}}(Z_1 + Z_2)$ uniformly in $\lambda$, where

$$Z_1 = \Sigma_{n=1}^{\infty}|a_n|\big[1 - \gamma_n(N)\big],$$

$$Z_2 = (1/N)\Sigma_{n=1}^{\infty}n|a_n|\big[|\gamma_n(N)| + |\gamma_{n+1}(N)|\big].$$

$Z_1$ is the dominant term in the bias, for with $M$ the integer part of $(b/\alpha)\ln N$, the truncated series at $M$ in $Z_2$ is $O(\ln N/N)$ by (3.2i) and Lemma 2.1, and the tail sum is $O(N(\ln N)^{1-r/2})$ by (3.2iii). Next, for any integer $m > 1$, we have by (3.2i) (3.2iii) and Lemma 2.1 that

$$Z_1 \leqslant \frac{K_2 A_1(r) r}{(r-2)} \cdot N^{-b} e^{\alpha m} + \frac{4A_1(r)}{(r-2)} \cdot m^{1-(r/2)},$$

and, by an argument similar to the one employed for $S_1^2$ of Theorem 3.3, we find (5.8)

$$b\left[\hat{\phi}_N(\lambda)\right] \leqslant D_3(\ln N)^{(2-r)/2}(1 + o(1)); \; D_3 = 4A_1(r)(b/\alpha)^{1-(r/2)}/\,(r-2).$$

For the variance we have by the Cauchy-Schwarz inequality and (2.6) that

$$\operatorname{Var}\left[\hat{\phi}_N(\lambda)\right] \leqslant (2/\pi^2\beta)\left[\Sigma_{n=1}^{\infty}|\gamma_n(N)|\operatorname{Var}^{\frac{1}{2}}[\hat{a}_n(N)]\right]^2.$$

Again, with $M$ the integer part of $(b/\alpha)\ln N$, the truncated sum at $M$ is $O(N^{-p/2})$, $p = 1-(2b/\alpha)\ln 3$, by Theorem 3.2 and (3.2i); whereas the tail sum is $O(N^{-1/2})$ by Theorem 3.2 and (3.2ii). Hence uniformly in $\lambda$,

$$\operatorname{Var}\left[\hat{\phi}_N(\lambda)\right] = O(N^{-p}), p = 1 - (2b/\alpha)\ln 3. \tag{5.9}$$

The result follows by (5.8) and (5.9).

## REFERENCES


[1] Beutler, F. J. (1970). Alias free randomly timed sampling of stochastic processes. *IEEE Trans. Information Theory* **16** 147–152.

[2] Brillinger, D. R. (1972). The spectral analysis of stationary interval functions. *Proc. Sixth Berkeley Symp. Math. Statist. Probability* 483–513.

[3] Čencov, N. N. (1962). Evaluation of an unknown distribution density from observations. *Soviet Math.* **3** 1559–1562.

[4] Gaster, M. and Roberts, J. B. (1975). Spectral analysis of randomly sampled signals. *J. Inst. Math. Appl.* **15** 195–216.

[5] Masry, E. (1978). Poisson sampling and spectral estimation of continuous-time processes. *IEEE Trans. Information Theory* **24** 173–183.

[6] Parzen, E. (1958). On asymptotically efficient consistent estimates of the spectral density function of a stationary time series. *J. Roy. Statist. Soc. Ser. B.* **20** 303–322.

[7] Rosenblatt, M. (1971). Curve estimates. *Ann. Math. Statist.* **42** 1815–1842.

[8] Sansone, G. (1959). *Orthogonal Functions*. Interscience, New York.

[9] Schwartz, S. C. (1967). Estimation of probability density by an orthogonal series. *Ann. Math. Statist.* **38** 1261–1265.

[10] Shapiro, H. S. and Silverman, R. A. (1960). Alias-free sampling of random noise. *J. Soc. Indust. Appl. Math.* **8** 225–248.

[11] Wahba, G. (1975). Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation. *Ann. Statist.* **3** 15–29.

[12] Watson, G. S. (1969). Density estimation by orthogonal series. *Ann. Math. Statist.* **40** 1496–1498.



Department of Applied Physics
and Information Science
University of California, San Diego
La Jolla, California 92093