The Annals of Statistics
1982, Vol. 10, No. 1. 132–153

A CENTRAL LIMIT THEOREM FOR STATIONARY PROCESSES AND THE PARAMETER ESTIMATION OF LINEAR PROCESSES

BY YUZO HOSOYA[1] AND MASANOBU TANIGUCHI

Tohoku University and Hiroshima University

A central limit theorem is proved for the sample covariances of a linear process. The sufficient conditions for the theorem are described by more natural ones than usual. We apply this theorem to the parameter estimation of a fitted spectral model, which does not necessarily include the true spectral density of the linear process. We also deal with estimation problems for an autoregressive signal plus white noise. A general result is given for efficiency of Newton-Raphson iterations of the likelihood equation

1. Introduction. In this paper, we deal with estimation of linear processes and related problems. Most of the models which pertain to applications of time-series analysis can be reduced to a class of stationary processes which are usually termed linear processes, and a variety of investigations have been concerned with statistical inference based on these processes. Whittle (1952, 1962) gave the first systematic study with respect to the parameter estimation for a scalar-valued linear process $x(n) = \sum_{j=0}^{\infty} \alpha_j(\theta)e(n-j)$ with $\alpha_0(\theta) = 1$, where the $e(j)$ are i.i.d. random variables with mean zero and the innovation variance $\operatorname{Var}\{e(n)\} = \sigma^2$ does not depend on θ.

In order to estimate θ, Whittle proposed as an estimate of θ a value $\hat{\theta}$ which minimizes the quantity $\int_{-\pi}^{\pi} I_X(\omega)/f_\theta(\omega)\, d\omega$, where $I_X(\omega)$ is the periodogram calculated from a partial realization $x(1), \cdots, x(N)$ and $f_\theta(\omega)$ is the spectral density of the process; for an estimate of σ^2 he proposed $\hat{\sigma}^2 = \int_{-\pi}^{\pi} I_X(\omega)/g_{\hat{\theta}}(\omega)\, d\omega$ where $g_\theta(\omega) = |\sum_{j=0}^{\infty} \alpha_j(\theta)e^{i\omega j}|^2$.

Whittle suggested an asymptotic theory pertaining to these estimates and Walker (1964) and Hannan (1973) later expounded the theory in a more rigorous fashion. Finding that there are cases where the innovation variance σ^2 depends upon θ, Hosoya (1974) proposed to minimize $\int_{-\pi}^{\pi} \{\log f_\theta(\omega) + I_X(\omega)/f_\theta(\omega)\}\, d\omega$ instead of $\int_{-\pi}^{\pi} I_X(\omega)/f_\theta(\omega)\, d\omega$ in order to find an estimate of θ, and he gave the asymptotic distribution of the estimate under regularity conditions similar to those of Walker's paper. In particular, Hosoya noted that while the asymptotic covariance of the estimate does not depend on the fourth cumulant of the innovation in Whittle's model, it does in his model. For a vector-valued linear process $x(n) = \sum_{j=0}^{\infty} A_j(\theta)e(n-j)$, where the $x(n)$ and $e(n)$ are mean-zero vector-valued processes and the coefficients $A_j(\theta)$ are matrices, Dunsmuir and Hannan (1976) and Dunsmuir (1979) consider estimation of θ by the minimization of the quantity $\log \det K(\theta) + \int_{-\pi}^{\pi} \operatorname{tr}\{f_\theta(\omega)^{-1} I_X(\omega)\}\, d\omega$, where $K(\theta)$ is the covariance matrix of the $e(n)$ and f_θ and I_X are now spectral density and periodogram matrices of $x(n)$.

The former paper assumed that $K(\theta)$ and the $A_j(\theta)$'s are separately parameterized; the latter removed that assumption. Besides the regularity conditions on f_θ, these papers assumed that $\{e(n)\}$ is a strictly stationary ergodic process such that (i) $E\{e(n) \mid \mathscr{F}_{n-1}\} = 0$ a.e., (ii) $E\{e_a(n)e_b(n) \mid \mathscr{F}_{n-1}\} = K_{ab}$ a.e., (iii) $E\{e_a(n)e_b(n)e_c(n) \mid \mathscr{F}_{n-1}\} = Q_{abc}$ a.e., (iv) $E\{e_a(n)e_b(n)e_c(n)e_d(n)\} < \infty$, where the subscripts denote respective components of $e(n)$. Though condition (i) is natural in view of the definition of innovation process and (ii) is

Received January 1981; revised September 1981.

[1] Research partially supported by the U.S. Office of Naval Research contract number N00014-75-C-0442 in the Department of Statistics, Stanford University.

AMS 1980 *subject classifications.* Primary 60F05, 62M15; secondary 60G10, 60G35.

Key words and phrases. Stationary processes, central limit theorem, linear processes, spectral density, periodogram, Gaussian maximum likelihood estimate, robustness, autoregressive signal with white noise, Newton-Raphson iteration.

natural for the purpose of a central limit theorem for covariances, conditions (ii) and (iii) seem a little too artificial. We would replace these conditions by more natural ones with respect to the second and fourth conditional moments and replace strict stationarity by the fourth-order stationarity, at the same time dispensing with the explicit assumption of ergodicity. It is to be noted, however, that the above previously-mentioned papers established the central limit theorem and the strong consistency of the estimate under fairly general regularity conditions on f_θ.

In the following discussion, we aim at further development of those previous works on linear processes by establishing some general results which might have independent interest in themselves. To be more specific, we prove two central limit theorems in Section 2. The basic assumption known so far for a central limit theorem to hold for stationary processes is the strongly mixing condition for strictly stationary processes (see Rozanov, 1967, and Ibragimov and Linnik, 1971). We propose a new condition in Theorem 2.1 and discuss its relationships with various mixing conditions in Remark 2.1. Theorem 2.1 is applied to establish Theorem 2.2 which extends Hannan's central limit theorem for serial covariances (see Hannan, 1976).

In Section 3, we deal with the estimation problem. By regarding the quantity $\int_{-\pi}^{\pi} \{\log \det f_1(\omega) + \operatorname{tr} f_1(\omega)^{-1} f_2(\omega)\}\, d\omega$ as a measure of divergence of a spectral density matrix f_1 from another f_2 when f_1 and f_2 are of the same dimension, we can interpret the minimization of the integral $\int_{-\pi}^{\pi} \{\log \det f_\theta(\omega) + \operatorname{tr} f_\theta(\omega)^{-1} I_X(\omega)\}\, d\omega$ with respect to $\theta \in \Theta$ as the way to choose the least diverged spectral density in a fitted model $\{f_\theta;\, \theta \in \Theta\}$ from the observed spectral density, namely I_X. In Section 3 we give an asymptotic theory for the estimate without the assumption that the true spectral density belongs to $\{f_\theta;\, \theta \in \Theta\}$. In that respect, the above measure of divergence will play an important role.

In Remark 3.1, we give an example which illustrates that this extension is not a formal one and that the usual asymptotic properties do not necessarily hold for our case. The asymptotic theory we expound in Section 3 is derived as an application of results in Section 2. Furthermore, by way of Proposition (c) in Lemma 3.1, we show, as an extension of Lemma 1.7 of Hosoya (1974), that in general the quantity $\int_{-\pi}^{\pi} \{\log \det f_\theta(\omega) + \operatorname{tr} f_\theta(\omega)^{-1} f_{\theta_0}(\omega)\}\, d\omega$ is minimized at $\theta = \theta_0$. The result, incidentally, was treated as an assumption in Dunsmuir and Hannan (1976). The results in Section 3 also are an extension of those of Taniguchi (1979) for non-Gaussian vector processes.

Section 4 is for a model of an autoregressive signal plus white noise. There we suggest a Newton-Raphson iterative computation procedure for constructing an estimate.

In Section 5 we give a justification for that procedure. Fisher (1925) was the first to note the fact that an estimate which is obtained as the first step of the Newton-Raphson iteration for solving a likelihood equation is in general equivalent to the maximum likelihood estimate up to probability order $O_p(1/\sqrt{N})$ when observations are i.i.d. and sample size is N. We can say more. The second iteration will produce an estimate which is equivalent to the maximum likelihood estimate up to order $O_p(1/N)$. We establish the fact in a very general framework in Theorem 5.1.

As for notations used in this paper, we denote the set of all integers by J, and denote Kronecker's-delta by $\delta(m, n)$.

2. Some limit theorems for stationary processes.

Let $\{z(n);\, n \in J\}$ be a vector-valued linear process generated as

$$z(n) = \sum_{j=0}^{\infty} G(j) e(n-j), \qquad n \in J, \tag{2.1}$$

where the $z(n)$'s have s components and the $e(n)$'s are p-vectors such that $E\{e(n)\} = 0$ and $E\{e(m)e(n)'\} = \delta(m, n)K$, with K a nonsingular p by p matrix; the $G(j)$'s are s by p matrices; and the components of z, e and G are all real. If $\sum_{j=0}^{\infty} \operatorname{tr} G(j)KG(j)' < \infty$ (this condition is assumed throughout), the process $\{z(n)\}$ is a second-order stationary process and has a spectral density matrix $f(\omega)$ which is representable as

(2.2) $$f(\omega) = \frac{1}{2\pi} k(\omega) K k(\omega)^*, \qquad -\pi \le \omega \le \pi,$$

where $k(\omega) = \sum_{j=0}^{\infty} G(j) e^{i\omega j}$. Denote by $C_z(s)$ and $I_z(\omega)$ respectively, the serial covariance and the periodogram matrices are constructed from a partial realization $\{z(1), \cdots, z(N)\}$; namely,

$$C_z(s) = \frac{1}{N} \sum_{m=1}^{N-s} z(m) z(m+s)', \qquad 0 \le s \le N-1,$$

and $C_z(s) = C_z'(-s)$ for $-N+1 \le s < 0$; and

$$I_z(\omega) = F_z(\omega) F_z(\omega)^*, \qquad \text{where} \quad F_z(\omega) = \frac{1}{\sqrt{2\pi N}} \sum_{n=1}^{N} z(n) e^{in\omega}, \qquad -\pi \le \omega \le \pi.$$

Denote the (α, β) component of $G(j)$, C_z and I_z by $G_{\alpha\beta}(j)$, $C^z_{\alpha\beta}$ and $I^z_{\alpha\beta}$ respectively, and denote the αth component of $z(n)$ and $e(n)$ by $z_\alpha(n)$ and $e_\alpha(n)$. Assuming that $\{e(n)\}$ is fourth-order stationary, let $Q^e_{\alpha_1 \cdots \alpha_4}(t_1, t_2, t_3)$ be the joint fourth cumulant of $e_{\alpha_1}(t)$, $e_{\alpha_2}(t+t_1)$, $e_{\alpha_3}(t+t_2)$, $e_{\alpha_4}(t+t_3)$ and assume that

$$\sum_{t_1, t_2, t_3 = -\infty}^{\infty} | Q^e_{\alpha_1 \cdots \alpha_4}(t_1, t_2, t_3) | < \infty \ (1 \le \alpha_1, \cdots, \alpha_4 \le p);$$

then the process $\{e(n)\}$ has a fourth-order spectral density $\tilde{Q}^e_{\alpha_1 \cdots \alpha_4}(\omega_1, \omega_2, \omega_3)$ such that

(2.3) $$\tilde{Q}^e_{\alpha_1 \cdots \alpha_4}(\omega_1, \omega_2, \omega_3) = \frac{1}{(2\pi)^3} \sum_{t_1, t_2, t_3 = -\infty}^{\infty} \exp\{-i(\omega_1 t_1 + \omega_2 t_2 + \omega_3 t_3)\} Q^e_{\alpha_1 \cdots \alpha_4}(t_1, t_2, t_3).$$

Denote by $Q^z_{q_1 \cdots q_4}$ and $\tilde{Q}^z_{q_1 \cdots q_4}$, respectively, the fourth-order cumulant and spectral density of the process $\{z(n)\}$, $1 \le q_1, \cdots, q_4 \le s$.

We now set down a central limit theorem for a second-order stationary process.

THEOREM 2.1. *Let a zero-mean vector-valued second-order stationary process* $\{x(t) = (x_1(t), \cdots, x_p(t))' : t \in J\}$ *be such that, for a positive constant* ε, (i) $\mathrm{Var}\{E(x_\alpha(t+\tau) | \mathcal{F}_t)\} = O(\tau^{-2-\varepsilon})$ *uniformly in t, for* $\alpha = 1, \cdots, p$, (ii) *for a positive constant* η,

$$E | E\{x_\alpha(\ell) x_\beta(m) | \mathcal{F}_t\} - E\{x_\alpha(\ell) x_\beta(m)\} | = O[\{\min(|\ell - t|, |m - t|)\}^{-1-\eta}]$$

uniformly in t, for $\alpha, \beta = 1, \cdots, p$ *and* ℓ, m *both greater than t,* (iii) $\{x(t)\}$ *has a spectral density matrix* $f(\omega) = \{f_{\alpha\beta}(\omega); \alpha, \beta = 1, \cdots, p\}$ *such that each element is continuous at the origin and* $f(0)$ *is non-degenerate, then* $\xi_N = N^{-1/2} \sum_{n=1}^{N} x(n)$ *is asymptotically normally distributed with mean zero and covariance matrix* $2\pi f(0)$, *where* $\mathcal{F}_t$ *is the* σ*-field generated by the set of random vectors* $\{x(n); n \le t\}$.

The proof is given in Section 6, together with the proofs of the other theorems and lemmas in Sections 2 and 3.

REMARK 2.1. The relationship between the previous conditions (i) and (ii) and the strongly mixing condition seems not to be a straightforward one. The mixing condition is for strictly stationary process, whereas the previous conditions do not require that restriction. Moreover, the former is concerned with a property between events belonging to $\mathcal{F}^t_{-\infty}$ and $\mathcal{F}^{\infty}_{t+\tau}$, whereas the latter conditions concern only the relation between $\mathcal{F}^t_{-\infty}$ and $\mathcal{F}_{t+\tau}$, where $\mathcal{F}^t_{-\infty}$, $\mathcal{F}^{\infty}_{t+\tau}$ and $\mathcal{F}_{t+\tau}$ are σ-fields generated respectively by $\{x(s); s \le t\}$, $\{x(s); s \ge t + \tau\}$ and $\{x(t+\tau)\}$. On the other hand, conditions (i) and (ii) seem to be more strict with respect to the upper bound on the difference $|P(A)P(B) - P(A \cap B)|$ $(A \in \mathcal{F}^t_{-\infty}, B \in \mathcal{F}_{t+\tau})$ than the strongly mixing condition does. However, if a process $\{x(t)\}$ is a strictly stationary uniformly mixing process such that

$$\sup_{A \in \mathcal{F}^t_{-\infty}, B \in \mathcal{F}^{\infty}_{t+\tau}} | P(A \cap B) - P(A)P(B) | / P(A) = O(\tau^{-2-\varepsilon})$$

and $Ex(t)^{2+\eta} < \infty$ for certain positive ε and η, then the process satisfies the conditions (i) and (ii). On the other hand, if a process is absolutely mixing with finite second-order moment and if there exist ε and $\lambda > 0$ such that

$$E \sup_{B \in \mathcal{F}_{t+\tau}^{\infty}} |P(B) - P(B \mid \mathcal{F}_{-\infty}^{t})| = O(\tau^{-2-\varepsilon-2\lambda})$$

and as c tends to infinity

$$\int_{|x|>c} x^2 \, dP = O\left(\frac{1}{c^{(2+\varepsilon)/\lambda}}\right)$$

(this latter condition always holds for normal distribution), then (i) and (ii) follow; see Ibragimov and Solev (1969) for absolute regularity.

The first proposition can be shown as follows. First, $x(n)$ can be approximated by a simple function $\tilde{x}(n) = \sum \alpha_j I(\omega \in A_j)$ such that $A_j \in \mathcal{F}_n^{\infty}$ and $E|x(n) - \tilde{x}(n)|^2 < \tau^{-2-\varepsilon}$ for a fixed τ. Since

$$|E\{\tilde{x}(n)\} - E\{\tilde{x}(n) \mid \mathcal{F}_{-\infty}^{n-\tau}\}|^2 = |\textstyle\sum_j \alpha_j\{P(A_j) - P(A_j \mid \mathcal{F}_{-\infty}^{n-\tau})\}|^2$$

$$\leq [\textstyle\sum_j |\alpha_j|^2\{P(A_j) + P(A_j \mid \mathcal{F}_{-\infty}^{n-\tau})\}]\{\sum_j |P(A_j) - P(A_j \mid \mathcal{F}_{-\infty}^{n-\tau})|\}$$

and ess sup $\sum_j |P(A_j) - P(A_j \mid \mathcal{F}_{-\infty}^{n-\tau})| = O(\tau^{-2-\varepsilon})$ uniformly for any decomposition $\{A_j\}$ (Ibragimov and Linnik, 1971, page 308) it follows that

$$\operatorname{Var}[E\{\tilde{x}(n)\} - E\{\tilde{x}(n) \mid \mathcal{F}_{-\infty}^{n-\tau}\}] = O(\tau^{-2-\varepsilon}).$$

Since $E|x(n) - \tilde{x}(n)|^2 < \tau^{-2-\varepsilon}$, $\{x(t)\}$ is seen to satisfy the condition (i). Condition (ii) follows from a slightly modified argument by means of Hölder's inequality. As for the absolutely mixing case, let y be the truncated version of $x(n)$ such that $y = x(n)$ if $|x(n)| \leq \tau^{\lambda}$ and $y = \operatorname{sign}(x(n)) \cdot \tau^{\lambda}$ if $|x(n)| > \tau^{\lambda}$, and let $\hat{x}(n)$ be a simple function $\hat{x}(n) = \sum \beta_j I(\omega \in C_j)$, $C_j \in \mathcal{F}_n^{\infty}$, such that $\sup|\hat{x}(n) - y| < \tau^{-2-\varepsilon}$. In the relation

$$|E\{\hat{x}(n)\} - E\{\hat{x}(n) \mid \mathcal{F}_{-\infty}^{n-\tau}\}|^2$$

$$\leq [\textstyle\sum_j |\beta_j|^2\{P(C_j) + P(C_j \mid \mathcal{F}_{-\infty}^{n-\tau})\}] \times \sum_j |P(C_j) - P(C_j \mid \mathcal{F}_{-\infty}^{n-\tau})|,$$

the first term on the right-hand side is less than $2\tau^{2\lambda}$ and

$$E\{\textstyle\sum_j |P(C_j) - P(C_j \mid \mathcal{F}_{-\infty}^{n-\tau})|\} \leq 2E\{\sup C \in \mathcal{F}_n^{\infty} |P(C) - P(C \mid \mathcal{F}_{-\infty}^{n-\tau})|\} = O(\tau^{-2-\varepsilon-2\lambda}),$$

$$\operatorname{Var}\{E(\hat{x}(n)) - E\hat{x}(n) \mid \mathcal{F}_{-\infty}^{n-\tau})\} = O\left(\frac{1}{\tau^{2+\varepsilon}}\right).$$

Now condition (i) follows from the fact that $E|\hat{x}(n) - x(n)|^2 = O(\tau^{-2-\varepsilon})$. Condition (ii) holds in the same way.

Let $\{z(n); n \in J\}$ be the linear process introduced in the first paragraph of this section; that is, $z(n) = \sum_{j=0}^{\infty} G(j)e(n-j)$, $n \in J$. Denote by $\mathcal{B}(t)$ the σ-field generated by $\{e(n); n \leq t\}$.

THEOREM 2.2. *Suppose that*

(i) *for each* β_1, β_2 *and* m,

$$\operatorname{Var}[E\{e_{\beta_1}(n)e_{\beta_2}(n+m) \mid \mathcal{B}(n-\tau)\} - \delta(m,0)K_{\beta_1\beta_2}] = O(\tau^{-2-\varepsilon}), \qquad \varepsilon > 0,$$

uniformly in n;

(ii) $E|E\{e_{\beta_1}(n_1)e_{\beta_2}(n_2)e_{\beta_3}(n_3)e_{\beta_4}(n_4) \mid \mathcal{B}(n_1 - \tau)\}$

$$- E\{e_{\beta_1}(n_1)e_{\beta_2}(n_2)e_{\beta_3}(n_3)e_{\beta_4}(n_4)\}| = O(\tau^{-1-\eta}),$$

uniformly in n_1, *where* $n_1 \leq n_2 \leq n_3 \leq n_4$ *and* $\eta > 0$;

(iii) *the spectral densities* $f_{\beta\beta}(\beta = 1, \cdots, s)$ *are square-integrable*;

(iv) $\sum_{j_1,j_2,j_3=-\infty}^{\infty} |Q_{\beta_1\cdots\beta_4}^{e}(j_1, j_2, j_3)| < \infty$.

Then $\sqrt{N}\{C^z_{\alpha_1\alpha_2}(m) - \gamma^z_{\alpha_1\alpha_2}(m)\}(\alpha_1, \alpha_2 = 1, \cdots, s, 0 \le m \le L)$ *have a joint asymptotic normal distribution whose mean is zero and the asymptotic covariance between* $\sqrt{N}\{C^z_{\alpha_1\alpha_2}(m_1) - \gamma^z_{\alpha_1\alpha_2}(m_1)\}$ *and* $\sqrt{N}\{C^z_{\alpha_3\alpha_4}(m_2) - \gamma^z_{\alpha_3\alpha_4}(m_2)\}$ *is given as*

$$2\pi\int_{-\pi}^{\pi}[f_{\alpha_1\alpha_3}(\omega)\overline{f_{\alpha_2\alpha_4}(\omega)}\exp\{-i(m_2-m_1)\omega\} + f_{\alpha_1\alpha_4}(\omega)\overline{f_{\alpha_2\alpha_3}(\omega)}\exp\{i(m_1+m_2)\omega\}]\,d\omega$$

$$+ 2\pi\sum_{\beta_1,\ldots,\beta_4=1}^{p}\int\int_{-\pi}^{\pi}\exp\{im_1\omega_1 + im_2\omega_2\}k_{\alpha_1\beta_1}(\omega_1)$$

$$\cdot k_{\alpha_2\beta_2}(-\omega_1)k_{\alpha_3\beta_3}(\omega_2)k_{\alpha_4\beta_4}(-\omega_2)\tilde{Q}^e_{\beta_1\ldots\beta_4}(\omega_1, -\omega_2, \omega_2)\,d\omega_1 d\omega_2.$$

3. Asymptotic properties of quasi-Gaussian maximum likelihood estimates for a linear process. In this section we shall apply the results of the previous section for estimation of parameters of the spectral density fitted for the process (2.1). Let $\{z(n); n \in J\}$ be the linear process defined by (2.1) with spectral density $f(\omega)$ given by (2.2). Fitting a certain parametric spectral density model $f_\theta(\omega)$, $\theta \in \Theta \subset R^q$, for this process we shall estimate θ. Let $\mathscr{P}$ denote the set of all spectral density matrices of linear processes whose coefficients satisfy

$$\sum_{j=0}^{\infty}\operatorname{tr}\{G(j)KG(j)'\} < \infty.$$

A functional T defined on $\mathscr{P}$ is determined by the requirement that for a parametric family of spectral density matrices $\{f_\theta; \theta \in \Theta \subset R^q\}$, there exists a unique $T(f)$ in Θ for every $f \in \mathscr{P}$ such that

$$D(f_{T(f)}, f) = \min_{t\in\Theta} D(f_t, f), \tag{3.1}$$

where

$$D(f_t, f) = \int_{-\pi}^{\pi}[\log\det f_t(\omega) + \operatorname{tr}\{f_t(\omega)^{-1}f(\omega)\}]\,d\omega.$$

Define a convergence on $\mathscr{P}$ as follows. If, for every continuous $s \times s$ matrix-valued function $\psi(\omega)$,

$$\int_{-\pi}^{\pi}\operatorname{tr}\{\psi(\omega)f_N(\omega)\}\,d\omega \to \int_{-\pi}^{\pi}\operatorname{tr}\{\psi(\omega)f(\omega)\}\,d\omega \quad \text{as} \quad N\to\infty, \qquad f_N \in \mathscr{P},$$

then we say that f_N converges to f weakly, denoted by $f_N \to_w f$. To ensure the existence of $T(f)$, some assumptions are needed on the parametric family $\{f_\theta; \theta \in \Theta\}$.

LEMMA 3.1. *Suppose that Θ is a compact subset of R^q, that $\theta_1 \ne \theta_2$ implies $f_{\theta_1} \ne f_{\theta_2}$ on a set of positive Lebesgue measure, that $f_\theta(\omega)$ is positive definite, and also that every component of $f_\theta(\omega)$ is continuous in θ and ω. Then* (a) *for every $f \in \mathscr{P}$, there exists a value $T(f) \in \Theta$ satisfying* (3.1); (b) *if $T(f)$ is unique and if $f_N \to_w f$, $T(f_N) \to T(f)$ as $N \to \infty$;* (c) *$T(f_\theta) = \theta$ for every $\theta \in \Theta$.*

For an estimate of $T(f)$ we propose $T\{I_z(\omega)\}$, which is to be called a quasi-Gaussian maximum likelihood estimate under the model f_θ. For simplicity hereafter, denote $T\{I_z(\omega)\}$ by $T(I_z)$. We assume that every component of $f_\theta(\omega)$ is a twice continuously differentiable function of $\theta \in \Theta$, where Θ is compact such that it has a non-empty open subset, and that the second derivatives of these components are continuous in $\omega \in [-\pi, \pi]$. Then we have

THEOREM 3.1. *Suppose that $T(f)$ exists uniquely and lies in* IntΘ, *and that*

$$M_f = \int_{-\pi}^{\pi} \left[\frac{\partial^2}{\partial\theta\partial\theta'} \operatorname{tr}\{ f_\theta(\omega)^{-1} f(\omega)\} + \frac{\partial^2}{\partial\theta\partial\theta'} \log\det f_\theta(\omega) \right]_{\theta=T(f)} d\omega$$

is a nonsingular matrix. Further we assume the conditions (i)–(iv) *in Theorem* 2.2, *and* (v) $f(\omega) \in \mathrm{Lip}(\alpha)$, *the Lipschitz class of degree* α, $\alpha > 1/2$.

Then

$$p\text{-}\lim_{N\to\infty} T(I_z) = T(f),$$

and the distribution of the vector $\sqrt{N}\{T(I_z) - T(f)\}$ *under* f, *as* $N \to \infty$, *tends to the normal distribution with mean zero and covariance matrix* $M_f^{-1}\tilde{V}M_f^{-1}$, *where* $\tilde{V} = \{\tilde{V}_{j\ell}\}$ *is a* $q \times q$ *matrix such that*

$$\tilde{V}_{j\ell} = 4\pi \int_{-\pi}^{\pi} \operatorname{tr}\left[f(\omega) \frac{\partial}{\partial\theta_j} \{f_\theta(\omega)\}^{-1} f(\omega) \frac{\partial}{\partial\theta_\ell} \{f_\theta(\omega)\}^{-1} \right]_{\theta=T(f)} d\omega$$

$$+ 2\pi \textstyle\sum_{r,t,u,v=1}^{s} \displaystyle\iint_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial\theta_j} f_\theta^{(r,t)}(\omega_1) \cdot \frac{\partial}{\partial\theta_\ell} f_\theta^{(u,v)}(\omega_2) \right\}_{\theta=T(f)} \tilde{Q}^z_{rtuv}(-\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2,$$

and $f_\theta^{(r,t)}(\omega)$ *is the* (r, t)-th *element of* $\{f_\theta(\omega)\}^{-1}$.

COROLLARY 3.1. *The above* $\tilde{V}_{j\ell}$ *can be expressed as*

$$\tilde{V}_{j\ell} = 4\pi \int_{-\pi}^{\pi} \operatorname{tr}\left[f(\omega) \frac{\partial}{\partial\theta_j} \{f_\theta(\omega)\}^{-1} f(\omega) \frac{\partial}{\partial\theta_\ell} \{f_\theta(\omega)\}^{-1} \right]_{\theta=T(f)} d\omega$$

$$+ 2\pi \textstyle\sum_{a,b,c,d=1}^{p} \sum_{r,t,u,v=1}^{s} \displaystyle\iint_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial\theta_j} f_\theta^{(r,t)}(\omega_1) \cdot \frac{\partial}{\partial\theta_\ell} f_\theta^{(u,v)}(\omega_2) \right\}_{\theta=T(f)}$$

$$\times k_{ra}(-\omega_1) k_{tb}(\omega_1) k_{uc}(-\omega_2) k_{vd}(\omega_2) \tilde{Q}^e_{abcd}(-\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2.$$

Note that in the above theorem the sufficient conditions for the central limit theorem are described only by moment conditions without the ergodic or mixing properties which have been used in much literature up to now.

Now we have the following proposition.

PROPOSITION 3.1. *If*

$$\operatorname{cum}\{e_a(n_1), e_b(n_2), e_c(n_3), e_d(n_4)\} = \begin{cases} \kappa_{abcd} & \text{if } n_1 = n_2 = n_3 = n_4, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

then the asymptotic covariance matrix of the quasi-Gaussian maximum likelihood estimates in Theorem 3.1 *is equal to* $M_f^{-1}UM_f^{-1}$, *where* U *has* (j, ℓ) *element*

$$U_{j\ell} = 4\pi \int_{-\pi}^{\pi} \operatorname{tr}\left[f(\omega) \frac{\partial}{\partial\theta_j} \{f_\theta(\omega)\}^{-1} f(\omega) \frac{\partial}{\partial\theta_\ell} \{f_\theta(\omega)\}^{-1} \right]_{\theta=T(f)} d\omega$$

$$+ \textstyle\sum_{a,b,c,d=1}^{s} \kappa_{abcd} \displaystyle\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\omega) \frac{\partial}{\partial\theta_j} \{f_\theta(\omega)\}^{-1} k(\omega)\, d\omega \right]_{ab}$$

$$\times \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\omega) \frac{\partial}{\partial\theta_\ell} \{f_\theta(\omega)\}^{-1} k(\omega)\, d\omega \right]_{cd} \Bigg|_{\theta=T(f)},$$

with $[\ ]_{ab}$ *denoting the* (a, b) *element of the matrix in the bracket.*

In general, we say that an estimate $\hat{\theta}$ of parameters of a process is robust against the fourth cumulant if the asymptotic distribution of $\hat{\theta}$ is independent of the fourth cumulants of the process. In the case of the estimation of innovation-free parameters, namely when the relationship

$$\frac{\partial}{\partial \theta} \int_{-\pi}^{\pi} \operatorname{tr}[\{f_\theta(\omega)\}^{-1} f(\omega)]\, d\omega \,|_{\theta = T(f)} = 0 \tag{3.3}$$

holds, the following remarks about the above proposition and robustness are pertinent.

REMARK 3.1. Suppose that $G(j)$ are square matrices, i.e. $s = p$, and that $\det\{\sum_{j=0}^{\infty} G(j) z^j\} \neq 0$ for $|z| \leq 1$. Then, in the case where $f(\omega) = f_\theta(\omega)$, $(k(\omega) = k_\theta(\omega))$, and when θ is the innovation-free parameter, it follows that

$$\int_{-\pi}^{\pi} k_\theta^*(\omega) \left[\frac{\partial}{\partial \theta_j} \{f_\theta(\omega)\}^{-1} \right] k_\theta(\omega)\, d\omega$$

$$= -\int_{-\pi}^{\pi} \left[\frac{\partial}{\partial \theta_j} k_\theta^*(\omega) \{k_\theta^*(\omega)\}^{-1} K^{-1} + K^{-1} \{k_\theta(\omega)\}^{-1} \frac{\partial}{\partial \theta_j} k_\theta(\omega) \right] d\omega = 0.$$

Thus the quasi-Gaussian maximum likelihood estimates for the innovation-free parameters are robust against the fourth cumulant under the conditions (3.2) and $f(\omega) = f_\theta(\omega)$. However, in the case of $f(\omega) \neq f_\theta(\omega)$, the relation

$$\int_{-\pi}^{\pi} k^*(\omega) \left[\frac{\partial}{\partial \theta_j} \{f_\theta(\omega)\}^{-1} \right]_{\theta = T(f)} k(\omega)\, d\omega = 0$$

is not satisfied generally. For example, let the true spectral density matrix be that for the $MA(1)$ process, namely

$$f(\omega) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & ½ \\ -½ & 1 \end{pmatrix} e^{i\omega} \right\} \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & ½ \\ -½ & 1 \end{pmatrix} e^{i\omega} \right\}^*,$$

and let the fitted spectral model be that for the AR(1) process with

$$\{f_\theta(\omega)\}^{-1} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \theta_1 & \theta_4 \\ \theta_3 & \theta_2 \end{pmatrix} e^{i\omega} \right\} \Sigma \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \theta_1 & \theta_4 \\ \theta_3 & \theta_2 \end{pmatrix} e^{-i\omega} \right\}^*,$$

where $\sum = \operatorname{diag}(\theta_5, \theta_6)$.
Then it is not difficult to show that

$$\int_{-\pi}^{\pi} k^*(\omega) \left[\frac{\partial}{\partial \theta_3} \{f_\theta(\omega)\}^{-1} \right]_{\theta = T(f)} k(\omega)\, d\omega = \begin{pmatrix} 24\pi/61 & 32\pi/61 \\ 32\pi/61 & -24\pi/61 \end{pmatrix} \neq 0.$$

This implies that, in the case of $f(\omega) \neq f_\theta(\omega)$, even if (3.2) is satisfied, the quasi-Gaussian maximum likelihood estimates for the innovation-free parameters are generally not robust against the fourth cumulant.

However if $s = 1$, i.e., the process concerned is scalar-valued, we have the following unified result.

REMARK 3.2. Consider the case $s = 1$, and assume that

$$\operatorname{cum}\{e(n_1), e(n_2), e(n_3), e(n_4)\} = \begin{cases} \kappa_4 & \text{if } n_1 = n_2 = n_3 = n_4, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

is satisfied. Then the quasi-Gaussian maximum likelihood estimates for the innovation-free parameters are robust against fourth cumulant even if $f(\omega) \neq f_\theta(\omega)$. In fact if θ is the innovation-free parameter, noting (3.3), we have

$$\int_{-\pi}^{\pi} k^*(\omega)\left[\frac{\partial}{\partial\theta}\{f_\theta(\omega)\}^{-1}\right]_{\theta=T(f)} k(\omega)\,d\omega = \frac{2\pi}{K}\int_{-\pi}^{\pi}\left[\frac{\partial}{\partial\theta}\{f_\theta(\omega)\}^{-1}f(\omega)\right]_{\theta=T(f)} d\omega = 0.$$

For example, consider the h-step ahead linear predictor with length q for the scalar case ($s = 1$). Then we construct a linear combination $\mathbf{a}_1 z(n-1) + \cdots + \mathbf{a}_q z(n-q)$ which satisfies

$$(3.5)\qquad \begin{aligned} &E\,|z(n+h) - \mathbf{a}_1 z(n-1) - \cdots - \mathbf{a}_q z(n-q)|^2 \\ &\qquad = \min_{b_1,\dots,b_q} E\,|z(n+h) - b_1 z(n-1) - \cdots - b_q z(n-q)|^2. \end{aligned}$$

However, this problem is equivalent to that of fitting

$$f_\theta(\omega) = |1 - \theta_1 e^{i(h+1)\omega} - \cdots - \theta_q e^{i(h+q)\omega}|^{-2}$$

by the criterion $D(f_\theta, f)$. In fact, since θ is independent of the innovation parameter, we have

$$\int_{-\pi}^{\pi} \frac{\partial}{\partial\theta}\{f_\theta(\omega)\}^{-1} f(\omega)\,d\omega\,|_{\theta=T(f)} = 0,$$

which implies $T(f) = (\mathbf{a}_1, \cdots, \mathbf{a}_q)'$. Putting $T(I_z) = (\hat{\mathbf{a}}_1, \cdots, \hat{\mathbf{a}}_q)'$, we can see that the distribution of the vector $\sqrt{N}(\hat{\mathbf{a}}_1 - \mathbf{a}_1), \cdots, \sqrt{N}(\hat{\mathbf{a}}_q - \mathbf{a}_q)$ tends to the normal distribution with zero mean vector and covariance matrix $(2R_q)^{-1}V(2R_q)^{-1}$, where R_q has (j, k) element $\gamma_z(j-k)$, $1 \le j, k \le q$, with $\gamma_z(j) = E\{z(n)z(n+j)\}$ and $\gamma_z(-j) = \gamma_z(j)$,

$$V = 4\pi\int_{-\pi}^{\pi} f(\omega)^2\left[\frac{\partial}{\partial\theta}\{f_\theta(\omega)\}^{-1}\frac{\partial}{\partial\theta'}\{f_\theta(\omega)\}^{-1}\right]_{\theta=T(f)} d\omega, \qquad T(f) = R_q^{-1}\begin{pmatrix}\gamma_z(h+1)\\ \vdots \\ \gamma_z(h+q)\end{pmatrix}.$$

Evidently the asymptotic distribution of $\sqrt{N}(\hat{\mathbf{a}}_1 - \mathbf{a}_1), \cdots, \sqrt{N}(\hat{\mathbf{a}}_q - \mathbf{a}_q)$ is robust against the fourth cumulant.

4. An autoregressive signal with white noise. As an application of the preceding asymptotic theory, there is the problem of estimation of an autoregressive signal which is observed when superimposed with white noise. The model with which we deal in this section is as follows. Suppose that a signal $\{s(t);\, t \in J\}$ is generated by a scalar-valued autoregressive process

$$(4.1)\qquad \textstyle\sum_{j=1}^{q} \theta_j s(t-j) = \eta(t), \qquad t \in J,\ \theta_0 \equiv 1,$$

where all zeroes of $\sum \theta_j z^j$ are assumed to be outside the unit circle and $E\{\eta(t)\} = 0$, $E\{\eta(t)\eta(s)\} = \theta_{q+1}\delta(t, s)$. Suppose then that the observed process $\{X(t);\, t \in J\}$ is given by

$$(4.2)\qquad X(t) = s(t) + e(t)$$

where $\{e(t)\}$ is a scalar-valued white noise such that $E\{e(t)\} = 0$, $E\{e(t)e(s)\} = \theta_{q+2}\delta(t, s)$, and $E\{e(t)\eta(s)\} = 0$ for all t and s.

For the estimation of the θ_j's in the model (4.1) and (4.2), based on a partial realization $X(1), \cdots, X(N)$, Hosoya (1974) suggested the minimization of

$$(4.3)\qquad D(f_\theta, I_X) = \int_{-\pi}^{\pi} \{\log f_\theta(\omega) + I_X(\omega)/f_\theta(\omega)\}\,d\omega,$$

where f_θ is the spectral density of $\{X(t)\}$ given by

$$(4.4)\qquad f_\theta(\omega) = \frac{1}{2\pi}\left(\frac{\theta_{q+1}}{|\sum_{j=0}^{q}\theta_j e^{ij\omega}|^2} + \theta_{q+2}\right),$$

and $I_X(\omega)$ is the periodogram $I_X(\omega) = \dfrac{1}{2\pi N}|\sum_{t=1}^{N} X(t)e^{i\omega t}|^2$. Hosoya's reason for the use of

(4.3) is as follows. According to the Fejér-Riesz theorem (see, e.g., Achiezer, 1956, page 152), if $g(\omega) = \sum_{k=-q}^{q} \alpha_k e^{ik\omega}$ and $g(\omega)$ is real and nonnegative, then there exists an $h(\omega)$ such that $g(\omega) = |h(\omega)|^2$ and $h(\omega) = \sum_{k=0}^{p} \beta_k e^{ik\omega}$. Thus, f_θ is representable as

$$f_\theta(\omega) = \frac{\sigma^2}{2\pi} \left| \frac{\sum_{j=0}^{q} \psi_j e^{ij\omega}}{\sum_{j=0}^{q} \theta_j e^{ij\omega}} \right|^2, \qquad \psi_0 = 1, \tag{4.5}$$

where σ^2 and ψ are functions of θ. Then we cannot omit the term $\sigma^2 = 2\pi \exp\{(1/2\pi)\cdot \int_{-\pi}^{\pi} \log f_\theta(\omega)\, d\omega\}$ in the approximation of the Gaussian likelihood function as Whittle (1952) did in defining the least-squares estimate. As in the previous section, denote by $T(I_X)$ the value of θ minimizing $D(f_\theta, I_X)$. Hosoya (1974) derived the asymptotic distribution of $T(I_X)$ for the case where $\{e(t)\}$ and $\{\eta(t)\}$ are Gaussian. The next result is more general and can be obtained as an immediate corollary of Theorem 3.1 and Corollary 3.1.

PROPOSITION 4.1. *For the model represented by* (4.1) *and* (4.2), *assume that*
(a) $\{e(t)\}$ *and* $\{\eta(t)\}$ *are fourth-order stationary processes such that the vector-valued process* $\{e(t), \eta(t)\}$ *satisfies conditions* (i), (ii), (iii), (iv) *of Theorem* 2.2 *and the fourth-order spectral density is denoted by* $\tilde{Q}^{e,\eta}_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, \omega_3)$, *where* $\alpha_1, \cdots, \alpha_4 = 1$ *or* 2.
(b) *Let* θ^0 *be the true value of* θ; *then* $\theta^0 = (\theta_1^0, \cdots, \theta_{q+2}^0) \in B \times K_1 \times K_2$, *where* B *is a compact subset of* R^q *such that for* $(\theta_1, \cdots, \theta_q) \in B$, *all zeroes of* $\sum_{j=0}^{q} \theta_j z^j$ *are outside the unit circle, and* K_1 *and* K_2 *are, respectively, compact subsets of* R^+. *Then* $\sqrt{N}\{T(I_X) - \theta^0\}$ *is asymptotically normally distributed with mean* 0 *and with covariance matrix* $M_f^{-1} V M_f^{-1}$, *where* $V = \{V_{j\ell}\}$, $j, \ell = 1, \cdots, q+2$ *such that*

$$V_{j\ell} = 4\pi \int_{-\pi}^{\pi} \{f_{\theta^0}(\omega)\}^2 \frac{\partial}{\partial\theta_j} \{f_{\theta^0}(\omega)\}^{-1} \frac{\partial}{\partial\theta_\ell} \{f_{\theta^0}(\omega)\}^{-1}\, d\omega$$

$$+ 2\pi \sum_{\alpha_1,\cdots,\alpha_4=1}^{2} \iint_{-\pi}^{\pi} \frac{\partial}{\partial\theta_j} \{f_{\theta^0}(\omega_1)\}^{-1} \frac{\partial}{\partial\theta_\ell} \{f_{\theta^0}(\omega_2)\}^{-1} k_{\alpha_1}(-\omega_1) k_{\alpha_2}(\omega_1) k_{\alpha_3}(-\omega_2)$$

$$\times k_{\alpha_4}(\omega_2) \tilde{Q}^{e,\eta}_{\alpha_1\cdots\alpha_4}(-\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2,$$

where $k_1(\omega) = 1/(\sum_{j=0}^{q} \theta_j^0 e^{i\omega j})$ *and* $k_2(\omega) = 1$.

For the numerical computation, it would be more appropriate to use, instead of (4.3),

$$A(\theta, X) = \sum_{j=0}^{N-1} \log f_\theta(\omega_j) + \sum_{j=0}^{N-1} \frac{I_X(\omega_j)}{f_\theta(\omega_j)}, \tag{4.6}$$

where

$$\omega_j = 2\pi j/N, \qquad j = 0, \cdots, N-1 \quad \text{and} \quad I_X(\omega_j) = \frac{1}{2\pi N} |\textstyle\sum_t X(t) e^{i\omega_j t}|^2.$$

Since the first derivative of $A(\theta, X)$ is nonlinear with respect to θ, a certain approximation is required for the solution of $\partial A(\theta, X)/\partial\theta = 0$. For that purpose, the Newton-Raphson iteration procedure starting with a consistent estimate of θ seems appropriate for the reason that follows. Then a possible procedure is:

A1: Solve $\sum_{j=0}^{q} \theta_j \{\sum_{t=q+\ell+1}^{N} X(t-j) X(t-q-\ell)/(N-q-\ell)\} = 0$, $\ell = 1, \cdots, q$, for θ. Let the solution be $\tilde{\theta}_1, \cdots, \tilde{\theta}_q$.

A2: Calculate $\tilde{\theta}_{q+1}$ and $\tilde{\theta}_{q+2}$ as

$$\tilde{\theta}_{q+1} = \frac{\frac{1}{N}\sum_j 2\pi I_X(\omega_j) \{g(\omega_j | \tilde{\theta}_1, \cdots, \tilde{\theta}_a)\}^{-1} - \frac{1}{N}\sum_j 2\pi I_X(\omega_j) \frac{1}{N}\sum_j \{g(\omega_j | \tilde{\theta}_1, \cdots, \tilde{\theta}_q)\}^{-1}}{\frac{1}{N}\sum_j \{g(\omega_j | \tilde{\theta}_1, \cdots, \tilde{\theta}_q)\}^{-2} - \left[\frac{1}{N}\sum_j \{g(\omega_j | \tilde{\theta}_1, \cdots, \tilde{\theta}_q)\}^{-1}\right]^2}$$

$$\tilde{\theta}_{q+2} = \frac{1}{N}\sum_j 2\pi I_X(\omega_j) - \tilde{\theta}_{q+1} \frac{1}{N}\sum_j \{g(\omega_j | \tilde{\theta}_1, \cdots, \tilde{\theta}_q)\}^{-1},$$

where $g(\omega | \theta_1, \cdots, \theta_q) = |\sum_{k=0}^{q} \theta_k e^{i\omega k}|^2$.

A3: Let $\theta^{(1)} = \tilde{\theta}$ and apply the iteration formula

$$\theta^{(k)} = \theta^{(k-1)} - \left[\frac{\partial^2 A(\theta^{(k-1)}, X)}{\partial\theta\partial\theta'}\right]^{-1} \frac{\partial A(\theta^{(k-1)}, X)}{\partial\theta}, \qquad k = 2, 3, \cdots.$$

Another class of models in which we get an ARMA model whose innovation variance depends on the other parameters is when a continuous differential equation is sampled discretely. This case is considered by Robinson (1980) which gives an algorithm similar to A1–A3, except that periodogram averages are used in place of autocovariances in A1.

5. Newton-Raphson iterative method. Under the assumptions of Proposition 4.1, we can easily see that the estimate $\tilde{\theta}$ calculated in the steps A1–A2 is consistent in the sense that $\tilde{\theta} \to \theta^0$ in probability. Moreover, it turns out that $\theta^{(3)}$ is equivalent to $T(I_X)$ up to the order $O_p(N^{-1})$; that is, $N\{\theta^{(3)} - T(I_X)\}$ tends to 0 in probability. This fact can be established in a more general framework, as we show in the next paragraph.

Let $L_N(\theta)$ be a function of an unknown parameter $\theta = \{\theta_1, \cdots, \theta_{q'}\}'$ and observations $X(1), X(2), \cdots, X(N)$ and let $N_\delta(\theta^0) = \{\theta: \| \theta - \theta^0 \| < \delta\}$ where θ^0 is the true value of θ. Now assume the following.

B1: $\log L_N(\theta)$ is third-order differentiable with respect to θ_i, $i = 1, \cdots, q'$, for $\theta \in N_\delta(\theta^0)$, and $\ell_{ij} = p\text{-}\lim_{N\to\infty} N^{-1}\partial^2 \log L_N(\theta^0)/\partial\theta_i\partial\theta_j$, $i, j = 1, \cdots, q'$, exist such that the matrix $\{\ell_{ij}\}$ is nonsingular.

B2: $N^{-1}\partial^3 \log L_N(\theta)/\partial\theta_i\partial\theta_j\partial\theta_k$ is bounded in probability uniformly in $\theta \in N_\delta(\theta^0)$.

B3: There exists a consistent estimate θ^1, i.e., $\theta^1 \to \theta^0$ in probability as $N \to \infty$, such that $\sqrt{N}(\theta^1 - \theta^0)$ has a limiting distribution with a finite covariance matrix.

B4: $\hat{\theta}$ is a solution of the equation $\partial \log L_N(\theta)/\partial\theta = 0$ which is consistent; and $\sqrt{N}(\hat{\theta} - \theta^0)$ has a finite asymptotic covariance matrix.

Let $\Gamma_N(\theta)$ be the $q' \times q'$ matrix whose (i,j) element is $\partial^2 \log L_N(\theta)/\partial\theta_i\partial\theta_j$ and let $\gamma_N(\theta)$ be the q'-vector whose ith element is $\partial \log L_N(\theta)/\partial\theta_i$. Now define

$$\theta^2 = \theta^1 - \Gamma_N(\theta^1)^{-1}\gamma_N(\theta^1) \tag{5.1}$$

and

$$\theta^3 = \theta^2 - \Gamma_N(\theta^2)^{-1}\gamma_N(\theta^2). \tag{5.2}$$

THEOREM 5.1. *If* B1 *through* B4 *hold, then* $\sqrt{N}(\theta^2 - \tilde{\theta})$ *tends to* 0 *in probability; in other words,* $\sqrt{N}(\theta^2 - \theta^0)$ *has the same limiting distribution as* $\sqrt{N}(\hat{\theta} - \theta^0)$. *Furthermore, under the same conditions,* $N(\theta^3 - \hat{\theta})$ *tends to* 0 *in probability.*

PROOF. By the Taylor expansion of $\partial \log L_N(\hat{\theta})/\partial\theta_i = 0$ around θ^{-1},

$$\partial \log L_N(\theta^1)/\partial\theta_i + \textstyle\sum_j (\hat{\theta}_j - \theta_j^1)\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j + \sum_j\sum_k (\hat{\theta}_j - \theta_j^1)(\hat{\theta}_k - \theta_k^1)\partial^3 \log L_N(\theta^*)/\partial\theta_i\partial\theta_j\partial\theta_k = 0, \qquad i = 1, \cdots, q', \tag{5.3}$$

where θ_i^* is between θ_i^1 and $\hat{\theta}_i$ for $i = 1, \cdots, q'$. In (5.3) above,

$$\begin{aligned} &\textstyle\sum_j (\hat{\theta}_j - \theta_j^1)\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j \\ &\quad = \textstyle\sum_j (\hat{\theta}_j - \theta_j^2)\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j + \sum_j (\theta_j^2 - \theta_j^1)\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j \\ &\quad = \textstyle\sum_j (\hat{\theta}_j - \theta_j^2)\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j - \partial \log L_N(\theta^1)/\partial\theta_i, \end{aligned} \tag{5.4}$$

by (5.1). From (5.3) and (5.4), it follows that

$$\textstyle\sum_j \sqrt{N}(\hat{\theta}_j - \theta_j^2) N^{-1}\partial^2 \log L_N(\theta^1)/\partial\theta_i\partial\theta_j = -\sum_j\sum_k \sqrt{N}(\hat{\theta}_j - \theta_j^1)(\hat{\theta}_k - \theta_k^1) N^{-1}\partial^3 \log L_N(\theta^*)/\partial\theta_i\partial\theta_j\partial\theta_k. \tag{5.5}$$

Writing the term on the right-hand side above as

$$\sqrt{N}(\hat{\theta}_j - \theta_j^1) N^{\varepsilon}(\hat{\theta}_k - \theta_k^1)\partial^3 \log L_N(\theta^*)/\{N^{1+\varepsilon}\partial\theta_i\partial\theta_j\partial\theta_k\},$$

we see that, for $0 < \varepsilon < ½$, both $\partial^3 \log L_N(\theta^*)/\{N^{1+\varepsilon}\partial\theta_i\partial\theta_j\partial\theta_k\}$ and $N^{\varepsilon}(\hat{\theta}_k - \theta_k^1)$ converge to 0 in probability and $\sqrt{N}(\hat{\theta}_j - \theta_j^1)$ is asymptotically of finite variance by assumption. Thus the whole quantity on the right of (5.5) converges to 0 in probability. It is easy to see that $N^{-1}\partial^2\log L_N(\theta^1)/\partial\theta_i\partial\theta_j$ converges to ℓ_{ij} defined in B1. By assumption, the matrix $\{\ell_{ij}\}$ is nonsingular so that each $\sqrt{N}(\hat{\theta}_j - \theta_j^2)$ tends to 0 in probability. In order to prove the second assertion of the theorem, note the following equation:

$$N(\theta^3 - \hat{\theta}) = [I - \{\Gamma_N(\theta^2)\}^{-1}\Gamma_N(\hat{\theta})]N(\theta^2 - \hat{\theta}) - \tfrac{1}{2}\{\Gamma_N(\theta^2)\}^{-1}\{\textstyle\sum_{k=1}^{q'} (\theta_k^2 - \hat{\theta}_k)\partial\Gamma_N(\theta^{**})/\partial\theta_k\} N(\theta^2 - \hat{\theta}), \tag{5.6}$$

where θ^{**} is a vector such that θ_j^{**} is between $\hat{\theta}_j$ and $\theta_j^2, j = 1, \cdots, q'$, $\partial\Gamma_N(\theta)/\partial\theta_k$ is a $q' \times q'$ matrix with $\partial^3\log L_N(\theta)/\partial\theta_i\partial\theta_j\partial\theta_k$ as its (i, j) element and I is the $q' \times q'$ identity matrix. Then if $N(\theta^2 - \hat{\theta})$ is bounded in probability, the first term on the right-hand side of (5.6) converges to 0 in probability since $\{\Gamma_N(\theta^2)\}^{-1}\Gamma_N(\hat{\theta})$ converges to the identity matrix, and the second term tends to 0 since $\{\Gamma_N(\theta^2)\}^{-1}\partial\Gamma_N(\theta^{**})/\partial\theta_k$ is asymptotically bounded. The fact that $N(\theta^2 - \hat{\theta})$ is bounded in probability is evident in view of the equation

$$N(\theta^2 - \hat{\theta}) = \sqrt{N}[I - \{\Gamma_N(\theta^1)\}^{-1}\Gamma_N(\hat{\theta})]\sqrt{N}(\theta^1 - \hat{\theta}) - \tfrac{1}{2}\{\Gamma_N(\theta^1)\}^{-1}[\textstyle\sum_k \{\partial\Gamma_N(\tilde{\theta})/\partial\theta_k\} \sqrt{N}(\theta_k^1 - \hat{\theta}_k)]\sqrt{N}(\theta^1 - \hat{\theta}) \tag{5.7}$$

where $\tilde{\theta}_j$ is between θ_j^1 and $\theta_j, j = 1, \cdots, q'$, since

$$\sqrt{N}[I - \{\Gamma_N(\theta^1)\}^{-1}\Gamma_N(\hat{\theta})] = \textstyle\sum_j \{\partial\Gamma_N(\tilde{\tilde{\theta}})/N\partial\theta_j\} \sqrt{N}(\hat{\theta}_j - \theta_j^1)$$

where $\tilde{\tilde{\theta}}_j$ is between θ_j^1 and $\hat{\theta}_j$, and $\partial\Gamma_N(\tilde{\tilde{\theta}})/N\partial\theta_j$ is bounded in probability.

REMARK 5.1. More generally, let θ^k be the kth step estimate obtained by repeated use of formulae (5.1) and (5.2). Then, in view of the above proof, it is easily seen that $N^{(k-1)/2}(\theta^k - \hat{\theta})$ tends to 0 in probability under the same conditions of Theorem 5.1.

REMARK 5.2. Denote by $\tilde{\Gamma}_N(\theta)$ the matrix whose (i, j) element is $E\left(\dfrac{\partial^2\log L(\theta)}{\partial\theta_i\partial\theta_j}\right)$. Then the estimates of the scoring method are obtained as $\phi^2 = \phi^1 - \{\tilde{\Gamma}_N(\theta^1)\}^{-1}\gamma_N(\theta^1)$ and $\phi^3 = \phi^2 - \{\tilde{\Gamma}_N(\phi^2)\}^{-1}\gamma_N(\phi^2)$. For these estimates, instead of (5.6) and (5.7) we have

$$N(\phi^3 - \hat{\theta}) = [I - \{\tilde{\Gamma}_N(\phi^2)\}^{-1}\Gamma_N(\hat{\theta})]N(\phi^2 - \hat{\theta}) + o_p(1), \tag{5.8}$$

and

$$N(\phi^2 - \hat{\theta}) = \sqrt{N}[I - \{\tilde{\Gamma}_N(\theta^1)\}^{-1}\Gamma_N(\hat{\theta})]\sqrt{N}(\theta^1 - \hat{\theta}) + O_p(1). \tag{5.9}$$

Therefore when $N^{-1}\tilde{\Gamma}_N(\theta)$ tends to ℓ_{ij}, which is generally true, it follows from (5.9) that $N(\phi^2 - \hat{\theta}) = O_p(1)$ and thus that $N(\phi^3 - \hat{\theta})$ tends to 0 in probability, since in (5.8) $\{\tilde{\Gamma}_N(\phi^2)\}^{-1}\Gamma_N(\hat{\theta})$ tends to the identity matrix as in Theorem 5.1.

6. Proofs of theorems. In this section we give the proofs of the theorems and lemmas stated in Sections 2 and 3.

6.1 *Proofs for Section 2.* We first set down some lemmas that will be needed in the proofs.

LEMMA A2.1. *If* $\sum_{j=0}^{\infty} |G_{\alpha\beta}(j)|^2 < \infty$ *for each* α, β *and if* $\sum_{j_1,\ldots,j_3=-\infty}^{\infty} |Q^e_{\alpha_1\cdots\alpha_4}(j_1, j_2, j_3)| < \infty$, *then the process* $\{z(n)\}$ *has a fourth-order spectral density* $\tilde{Q}^z_{q_1\cdots q_4}(\omega_1, \omega_2, \omega_3)$ *such that*

$$\tilde{Q}^z_{q_1\cdots q_4}(\omega_1, \omega_2, \omega_3) = \textstyle\sum_{\alpha_1,\ldots,\alpha_4=1}^{p} k_{q_1\alpha_1}(\omega_1 + \omega_2 + \omega_3)k_{q_2\alpha_2}(-\omega_1)k_{q_3\alpha_3}(-\omega_2)k_{q_4\alpha_4}(-\omega_3) \times \tilde{Q}^e_{\alpha_1\cdots\alpha_4}(\omega_1 + \omega_2 + \omega_3, \omega_2, \omega_3) \quad \text{a.e.} \tag{6.1}$$

PROOF. Define $G_{\alpha\beta}(j) = 0$ for $j < 0$. Then the fourth cumulant $Q^z_{q_1\cdots q_4}(n_1, n_2, n_3)$ is given as an absolutely convergent series such that

$$Q^z_{q_1\cdots q_4}(n_1, n_2, n_3) = \sum_{\alpha_1,\cdots,\alpha_4=1}^{p} \sum_{j_1,\cdots,j_4=-\infty}^{\infty} G_{q_1\alpha_1}(j_1)\cdots G_{q_4\alpha_4}(j_4) \times Q^e_{\alpha_1\cdots\alpha_4}(n_1 + j_1 - j_2, n_2 + j_1 - j_3, n_3 + j_1 - j_4). \tag{6.2}$$

In what follows in this proof, $Q^z_{q_1\cdots q_4}(n_1, n_2, n_3)$ is abbreviated simply by Q^z. Define

$$h_2(\omega \mid n_2 + j_1 - j_3, n_3 + j_1 - j_4) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} e^{-i\omega j} Q^e_{\alpha_1\cdots\alpha_4}(j, n_2 + j_1 - j_3, n_3 + j_1 - j_4).$$

Then in view of the fact that h_2 and $k_{q_2\alpha_2}$ are square-integrable, it follows that

$$\sum_{j_2=-\infty}^{\infty} G_{q_2\alpha_2}(j_2) Q^e_{\alpha_1\cdots\alpha_4}(n_1 + j_1 - j_2, n_2 + j_1 - j_3, n_3 + j_1 - j_4) = \int_{-\pi}^{\pi} k_{q_2\alpha_2}(-\omega_1) e^{i(n_1+j_1)\omega_1} h_2(\omega_1 \mid n_2 + j_1 - j_3, n_3 + j_1 - j_4)\, d\omega. \tag{6.3}$$

Since for a certain positive constant c_1,

$$\sum_{j_3} \int_{-\pi}^{\pi} | G_{q_3\alpha_3}(j_3) k_{q_2\alpha_2}(\omega) h_2(\omega \mid n_2 + j_1 - j_3, n_3 + j_1 - j_4)|\, d\omega \le c_1 \int_{-\pi}^{\pi} | k_{q_2\alpha_2}(\omega)|\, d\omega \sum_{j_1,j_2,j_3=-\infty}^{\infty} | Q^e_{\alpha_1\cdots\alpha_4}(j_1, j_2, j_3)| \tag{6.4}$$

so that the left-hand side of (6.4) is finite, it follows from the exchange of the summation and integration that

$$\begin{aligned} Q^z &= \sum_{\alpha_1\cdots\alpha_4=1}^{p} \sum_{j_1,j_4=-\infty}^{\infty} G_{q_1\alpha_1}(j_1) G_{q_4\alpha_4}(j_4) \int_{-\pi}^{\pi} \sum_{j_3=-\infty}^{\infty} G_{q_3\alpha_3}(j_3) \\ &\quad \times k_{q_2\alpha_2}(-\omega_1) e^{i(n_1+j_1)\omega_1} h_2(\omega_1 \mid n_2 + j_1 - j_3, n_3 + j_1 - j_4)\, d\omega_1 \\ &= \sum_{\alpha_1,\cdots,\alpha_4=1}^{p} \sum_{j_1,j_4=-\infty}^{\infty} G_{q_1\alpha_1}(j_1) G_{q_4\alpha_4}(j_4) \iint_{-\pi}^{\pi} k_{q_2\alpha_2}(-\omega_1) k_{q_3\alpha_3}(-\omega_2) \\ &\quad \times \exp\{i(n_1 + j_1)\omega_1 + i(n_2 + j_1)\omega_2\} h_3(\omega_1, \omega_2 \mid n_3 + j_1 - j_4)\, d\omega_1\, d\omega_2, \end{aligned} \tag{6.5}$$

where

$$h_3(\omega_1, \omega_2 \mid n_3 + j_1 - j_4) = \frac{1}{2\pi} \sum_{j_2,j_3=-\infty}^{\infty} e^{-i\omega_1 j_2 - i\omega_2 j_3} \times Q^e_{\alpha_1\cdots\alpha_4}(j_2, j_3, n_3 + j_1 - j_4)$$

and the second equation is due to the Parseval equality. By means of repeated use of a similar argument,

$$Q^z = \sum_{\alpha_1,\cdots,\alpha_4=1}^{p} \sum_{j_1=-\infty}^{\infty} G_{q_1\alpha_1}(j_1) \iiint_{-\pi}^{\pi} k_{q_2\alpha_2}(-\omega_1) k_{q_3\alpha_3}(-\omega_2) k_{q_4\alpha_4}(-\omega_3) \times \left[\prod_{\ell=1}^{3} \exp\{i(n_\ell + j_1)\omega_\ell\}\right] \tilde{Q}^e_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, \omega_2)\, d\omega_1\, d\omega_2\, d\omega_3. \tag{6.6}$$

Since again by the Parseval equality

$$\sum_{j_1=-\infty}^{\infty} G_{q_1\alpha_1}(j_1) \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\{i(\omega_1 + \omega_2 + \omega_3)j_1\} k_{q_2\alpha_2}(-\omega_1)(2\pi) \tilde{Q}^e_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, \omega_3)\, d\omega_1 = \int_{-\pi}^{\pi} k_{q_1\alpha_1}(\omega_1) k_{q_2\alpha_2}(-\omega_1 + \omega_2 + \omega_3) \tilde{Q}^e_{\alpha_1\cdots\alpha_4}(\omega_1 - \omega_2 - \omega_3, \omega_2, \omega_3)\, d\omega_1,$$

it follows from (6.6) that

$$Q^z = \sum_{\alpha_1,\cdots,\alpha_4=1}^{p} \iiint_{-\pi}^{\pi} \left[\prod_{\ell=1}^{3} \exp\{in_\ell\omega_\ell\}\right] k_{q_1\alpha_1}(\omega_1+\omega_2+\omega_3)\, k_{q_2\alpha_2}(-\omega_1)$$

$$\times\, k_{q_3\alpha_3}(-\omega_2)\, k_{q_4\alpha_4}(-\omega_3)\, \tilde{Q}^e_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, \omega_3)\, d\omega_1\, d\omega_2\, d\omega_3.$$

Accordingly, the relationship (6.1) holds.

LEMMA A2.2. *Assume* $\sum_{j_1,j_2,j_3=-\infty}^{\infty} | Q^z_{\alpha_1\cdots\alpha_4}(j_1, j_2, j_3)| < \infty$. *For any square-integrable functions* W_1 *and* W_2 *defined on* $[-\pi, \pi]$,

$$\lim_{N\to\infty} N \operatorname{Cov}\left\{\int_{-\pi}^{\pi} W_1(\omega) I^z_{\alpha_1\alpha_2}(\omega)\, d\omega, \int_{-\pi}^{\pi} W_2(\omega) I^z_{\alpha_3\alpha_4}(\omega)\, d\omega\right\}$$

$$(6.7)\quad = 2\pi \int_{-\pi}^{\pi} W_1(\omega)\, \overline{W_2(\omega)}\, f_{\alpha_1\alpha_3}(\omega)\, \overline{f_{\alpha_2\alpha_4}(\omega)}\, d\omega + 2\pi \int_{-\pi}^{\pi} W_1(\omega)\, \overline{W_2(-\omega)}\, f_{\alpha_1\alpha_4}(\omega)\, \overline{f_{\alpha_2\alpha_3}(\omega)}\, d\omega$$

$$+ 2\pi \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} W_1(\omega_1)\, W_2(-\omega_2)\, \tilde{Q}^z_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2.$$

PROOF. Let

$$\tilde{W}_1(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_1(\omega) e^{in\omega}\, d\omega, \qquad \tilde{W}_2(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_2(\omega) e^{in\omega}\, d\omega,$$

then

$$N \operatorname{Cov}\left\{\int_{-\pi}^{\pi} W_1(\omega) I^z_{\alpha_1\alpha_2}(\omega)\, d\omega, \int_{-\pi}^{\pi} W_2(\omega) I^z_{\alpha_3\alpha_4}(\omega)\, d\omega\right\}$$

$$= \frac{1}{N} \sum_{n_1,\cdots,n_4=1}^{N} \tilde{W}_1(n_1-n_2)\, \tilde{W}_2(n_3-n_4)\{\gamma^z_{\alpha_1\alpha_3}(n_3-n_1)\gamma^z_{\alpha_2\alpha_4}(n_4-n_2)$$

$$+ \gamma^z_{\alpha_1\alpha_4}(n_4-n_1)\gamma^z_{\alpha_2\alpha_3}(n_3-n_2) + Q^z_{\alpha_1\cdots\alpha_4}(n_2-n_1, n_3-n_1, n_4-n_1)\},$$

where the terms containing covariances converge to the first two terms in the right-hand side of (6.7); see Grenander-Rosenblatt (1957). Now let

$$B_N = \frac{1}{N} \sum_{n_1,\cdots,n_4=1}^{N} \tilde{W}_1(n_1-n_2)\, \tilde{W}_2(n_3-n_4)\, Q^z_{\alpha_1\cdots\alpha_4}(n_2-n_1, n_3-n_1, n_4-n_1),$$

and write $\ell_1 = n_1$, $\ell_2 = n_2 - n_1$, $\ell_3 = n_3 - n_1$, $\ell_4 = n_4 - n_1$. Then

$$(6.8)\quad B_N = \frac{1}{N} \sum_{\ell_2,\ell_3,\ell_4=-N+1}^{N-1} \{N - S(\ell_2, \ell_3, \ell_4)\}\, \tilde{W}_1(-\ell_2)\, \tilde{W}_2(\ell_3-\ell_4)\, Q^z_{\alpha_1\cdots\alpha_4}(\ell_2, \ell_3, \ell_4),$$

where

$$S(\ell_2, \ell_3, \ell_4) = \begin{cases} \max(|\ell_2|, |\ell_3|, |\ell_4|) & \text{if sign } \ell_2 = \text{sign } \ell_3 = \text{sign } \ell_4, \\ \max(|\ell_i|, |\ell_j|) + |\ell_k| & \text{if sign } \ell_i = \text{sign } \ell_j = -\text{sign } \ell_k. \end{cases}$$

However, since, for a certain positive constant c_2,

$$\frac{1}{N} \left| \sum_{\ell_2,\ell_3,\ell_4=-N+1}^{N-1} S(\ell_2, \ell_3, \ell_4)\, \tilde{W}_1(-\ell_2)\, \tilde{W}_2(\ell_3-\ell_4)\, Q^z_{\alpha_1\cdots\alpha_4}(\ell_2, \ell_3, \ell_4)\right|$$

$$\le \frac{c_2}{N} \sum_{\ell_2,\ell_3,\ell_4=-N+1}^{N-1} (|\ell_2| + |\ell_3| + |\ell_4|)\, |Q^z_{\alpha_1\cdots\alpha_4}(\ell_2, \ell_3, \ell_4)|$$

and also since the terms

$$\sum_{\ell_2,\ell_3,\ell_4} \frac{|\ell_j|}{N} |Q^z_{\alpha_1\cdots\alpha_4}(\ell_2, \ell_3, \ell_4)|, j = 2, 3, 4,$$

converge to 0 as $N \to \infty$, it follows that

(6.9) $$\lim_{N\to\infty} B_N = \sum_{\ell_2,\ell_3,\ell_4=-\infty}^{\infty} \tilde{W}_1(-\ell_2)\,\tilde{W}_2(\ell_3 - \ell_4)\,Q^z_{\alpha_1\cdots\alpha_4}(\ell_2, \ell_3, \ell_4).$$

Then, by repeated application of the Parseval equality,

(6.10) $$\lim_{N\to\infty} B_N = 2\pi \iint_{-\pi}^{\pi} W_1(\omega_1)\,W_2(-\omega_2)\,\tilde{Q}^z_{\alpha_1\cdots\alpha_4}(\omega_1, \omega_2, -\omega_2)\,d\omega_1\,d\omega_2.$$

LEMMA A2.3. *If* $\sum_j |G_{\alpha\beta}(j)|^2 < \infty$ *for each* α, β *and if*

$$\sum_{j_1,j_2,j_3} |Q^e_{\alpha_1\cdots\alpha_4}(j_1, j_2, j_3)| < \infty,$$

then for any α_1, α_2, α_3, α_4, we have

(6.11) $$\begin{aligned}\lim_{N\to\infty} N\operatorname{Cov}\{C^z_{\alpha_1\alpha_2}(m), C^z_{\alpha_3\alpha_4}(n)\} &= 2\pi \int_{-\pi}^{\pi} \{f_{\alpha_1\alpha_3}(\omega)\overline{f_{\alpha_2\alpha_4}(\omega)}e^{-i(n-m)\omega} + f_{\alpha_1\alpha_4}(\omega)\,\overline{f_{\alpha_2\alpha_3}(\omega)}e^{i(n+m)\omega}\}\,d\omega\\ &+ 2\pi \sum_{\beta_1,\ldots,\beta_4=1}^{p} \iint_{-\pi}^{\pi} \exp(im\omega_1 + in\omega_2)k_{\alpha_1\beta_1}(\omega_1)k_{\alpha_2\beta_2}(-\omega_1)k_{\alpha_3\beta_3}(\omega_2)k_{\alpha_4\beta_4}(-\omega_2)\\ &\times \tilde{Q}^e_{\beta_1\cdots\beta_4}(\omega_1, -\omega_2, \omega_2)\,d\omega_1\,d\omega_2.\end{aligned}$$

PROOF. In the relation

(6.12) $$\begin{aligned}N\operatorname{Cov}\{C^z_{\alpha_1\alpha_2}(m), C^z_{\alpha_3\alpha_4}(n)\} &= \sum_{u=-N+1}^{N-1}\left(1 - \frac{|u|}{N}\right)\{\gamma^z_{\alpha_1\alpha_3}(n)\gamma^z_{\alpha_2\alpha_4}(u + n - m) + \gamma^z_{\alpha_1\alpha_4}(u + n)\,\gamma^z_{\alpha_2\alpha_3}(u - m)\}\\ &+ \sum_{u=-N+1}^{N-1}\left(1 - \frac{|u|}{N}\right)Q^z_{\alpha_1\cdots\alpha_4}(m, u, u + n),\end{aligned}$$

the first sum converges to the first integral in (6.11); see, for example, Hannan (1976). Denote by D_N the second sum in the right-hand side of (6.12), and denote by $L_N(\lambda)$ the Fejér kernel; then

(6.13) $$\begin{aligned}D_N = 2\pi \sum_{\beta_1,\ldots,\beta_4=1}^{p} \iint_{-\pi}^{\pi} &\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} L_N(\omega_2 + \omega_3)\exp(im\omega_1 + in\omega_3)k_{\alpha_1\beta_1}(\omega_1 + \omega_2 + \omega_3)\right.\\ &\left.\times k_{\alpha_3\beta_3}(-\omega_2)\tilde{Q}^e_{\beta_1\cdots\beta_4}(\omega_1 + \omega_2 + \omega_3, \omega_2, \omega_3)\,d\omega_2\right\}k_{\alpha_2\beta_2}(-\omega_1)k_{\alpha_4\beta_4}(-\omega_3)\,d\omega_1\,d\omega_3.\end{aligned}$$

Let

$$H_N(\omega_1, \omega_2, \omega_3) = k_{\alpha_1\beta_1}(\omega_1 + \omega_2 + \omega_3)k_{\alpha_3\beta_3}(-\omega_2)\tilde{Q}^e_{\beta_1\cdots\beta_4}(\omega_1 + \omega_2 + \omega_3, \omega_2, \omega_3).$$

Then,

(6.14) $$\begin{aligned}\Bigg|\iiint_{-\pi}^{\pi} &\left\{\frac{1}{2\pi}L_N(\omega_2 + \omega_3)H(\omega_1, \omega_2, \omega_3) - H(\omega_1, -\omega_3, \omega_3)\right\}\\ &\times k_{\alpha_2\beta_2}(-\omega_1)k_{\alpha_4\beta_4}(-\omega_3)\exp(im\omega_1 + in\omega_3)\,d\omega_1\,d\omega_2\,d\omega_3\Bigg|\end{aligned}$$

$$\leq \left\{ \iint_{-\pi}^{\pi} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} L_N(\omega_2 + \omega_3) H(\omega_1, \omega_2, \omega_3)\, d\omega_2 - H(\omega_1, -\omega_3, \omega_3) \right|^2 d\omega_1\, d\omega_3 \right\}^{1/2}$$

$$\times \left\{ \int_{-\pi}^{\pi} | k_{\alpha_2\beta_2}(-\omega_1)|^2\, d\omega_1 \int_{-\pi}^{\pi} | k_{\alpha_4\beta_4}(-\omega_3)|^2\, d\omega_3 \right\}^{1/2}.$$

By a slight extension of the known result of L^2-convergence of a Cesaro sum, and the application of it to the first factor in the right-hand side of (6.14), it follows that this first factor converges to 0 as $N \to \infty$. Therefore,

$$\lim_{N\to\infty} D_N = \sum_{\beta_1,\cdots,\beta_4=1}^{p} 2\pi \iint_{-\pi}^{\pi} \exp(im\omega_1 + in\omega_3) k_{\alpha_1\beta_1}(\omega_1) k_{\alpha_2\beta_2}(-\omega_1)$$

$$\times\, k_{\alpha_3\beta_3}(\omega_3) k_{\alpha_4\beta_4}(-\omega_3) \tilde{Q}^{e}_{\beta_1\cdots\beta_4}(\omega_1, -\omega_3, \omega_3)\, d\omega_1\, d\omega_3. \qquad \square$$

Let $\{w_m(n), \mathscr{F}_m(n);\ n = 0, 1, 2, \cdots, n(m)\}$, $m = 1, 2, \cdots$, be a zero-mean square integrable martingale for each m where $\{\mathscr{F}_m(n);\ n = 0, 1, 2, \cdots, n(m)\}$ is a sequence of increasing σ-fields and $n(m) \to \infty$ as $m \to \infty$. The next lemma is due to Brown (1971), with a slight modification, and is essential for the theorem which follows it. Let $w_m(0) = 0$, $u_m(1) = w_m(1)$, and $u_m(k) = w_m(k) - w_m(k-1)$.

LEMMA A2.4. *Suppose that*

(i) $$\lim_{N\to\infty} \frac{1}{n(m)} \sum_{k=1}^{n(m)} E[u_m(k)^2 I\{| u_m(k)| \geq \varepsilon n(m)\}] = 0$$

for any $\varepsilon > 0$, *with I the indicator function, and*

(ii) $$[\textstyle\sum_{k=1}^{n(m)} E\{u_m(k)^2 \mid \mathscr{F}_m(k-1)\}]/[\sum_{k=1}^{n(m)} E\{u_m(k)^2\}] \to 1$$

in probability as $m \to \infty$. *Then*

$$\textstyle\sum_{k=1}^{n(m)} u_m(k)/[\sum_{k=1}^{n(m)} E\{u_m(k)^2\}]^{1/2}$$

is asymptotically normally distributed with mean 0 *and variance* 1.

Using the above lemmas, we shall now prove Theorem 2.1.

PROOF OF THEOREM 2.1. It suffices to show that the result holds for the case $p = 1$, since the proof for general vector-valued process is reduced to it by considering linear combinations $\sum_{\alpha=1}^{p} c_\alpha x_\alpha(n)$. Hence assume $p = 1$ in what follows. Let (s_k, t_k), $k = 1, \cdots, M$ be a sequence of pairs of integers such that $1 \leq s_k, t_k \leq N$, $s_k - t_k = \tau$, $t_k - s_{k-1} = \tau'$, and $t_1 = 1$. Without loss of generality it is assumed that $N = S_{M-1}$. The integers τ and τ' are chosen such that

$$\frac{(M-\tau)}{(\tau+\tau')} \to \infty, \qquad \tau' \to \infty, \qquad \frac{\tau'}{\tau \to 0} \quad \text{and} \quad \frac{M}{(\tau')^{1+\varepsilon/2}} \to 0.$$

(In order to see that such a choice is possible, refer to Rozanov, 1967, page 195). Then it follows that $\tau/(\tau')^{1+\varepsilon/2} \to 0$. Represents ξ_N in the form

$$\xi_N = (\tau/N)^{1/2} \textstyle\sum_{k=1}^{M} \eta_k + (\tau'/N)^{1/2} \sum_{k=1}^{M-1} n'_k,$$

where $\eta_k = \sum_{t_k \leq \ell < s_k} x(\ell)/\sqrt{\tau}$ and $\eta'_k = \sum_{s_k \leq \ell < t_{k+1}} x(\ell)/\sqrt{\tau'}$. In the inequality

(6.15) $$E\left\{\left(\frac{\tau'}{N}\right)^{1/2} \textstyle\sum_{k=1}^{M-1} \eta'_k\right\}^2 \leq \frac{\tau'}{N}(M-1)\mathrm{Var}(\eta'_1) + \frac{2\tau'}{N} \sum_{k=1}^{M-1} \sum_{\ell=k+1}^{M-1} | \mathrm{Cov}(\eta'_k, \eta'_\ell)|,$$

the term $(\tau'/N)(M-1)\mathrm{Var}(\eta'_1)$ tends to 0 since $\mathrm{Var}(\eta'_1)$ is bounded in view of the fact that, under the assumption of the theorem, $\lim_{\tau'\to\infty}\mathrm{Var}(\eta'_1) = 2\pi f(0)$, whereas, with respect to the second term, since

$$\mathrm{Var}\{E(\eta'_{k+j} \mid \mathscr{F}'_k)\} \le \frac{\{(\tau')^2 d_2\}}{\{\tau'(\tau j)^{2+\varepsilon}\}}, \quad j = 1, 2, \cdots,$$

for a positive constant d_2,

$$\begin{aligned}\frac{2\tau'}{N}\sum_{k=1}^{M-1}\sum_{\ell=k+1}^{M-1} |\mathrm{Cov}(\eta'_k, \eta'_\ell)| &\le \frac{2\tau'}{N}\sum_k\sum_\ell \{\mathrm{Var}(n'_k)\}^{1/2}[\mathrm{Var}\{E(\eta'_\ell \mid \mathscr{F}'_k)\}]^{1/2}\\ &\le \frac{2\tau' d_1}{N}\sum_{k=1}^{M-1}\sum_{j=1}^{M-1-k} j^{-1-\varepsilon/2}\cdot\tau'^{1/2}\cdot\tau^{-1-\varepsilon/2}\\ &\le \frac{2\tau'(M-1)\, d_1}{N}\sum_{j=1}^{\infty} j^{-1-\varepsilon/2}\cdot\tau'^{1/2}\cdot\tau^{-1-\varepsilon/2},\end{aligned} \tag{6.16}$$

where $\mathscr{F}'_k$ is the σ-field generated by $\{x(t); t \le t_k - 1)$ and d_1 is a positive constant. Since the term in the right-hand side in the third inequality of (6.16) is of order $O(\tau'^{3/2}\tau^{-2-\varepsilon/2})$, in view of the fact that $M/N = O(\tau^{-1})$, it is seen to converge to 0. Hence

$$\lim_{N\to\infty} E\left\{\left(\frac{\tau'}{N}\right)^{1/2}\sum_{k=1}^{M-1}\eta'_k\right\}^2 = 0.$$

Furthermore,

$$\mathrm{Var}\left\{\left(\frac{\tau}{N}\right)^{1/2}\sum_{k=1}^{M} E(\eta_k \mid \mathscr{F}^*_{k-1})\right\} \le d_2(M\tau)(\tau')^{-2-\varepsilon} \tag{6.17}$$

for a positive constant d_2, where $\mathscr{F}^*_{k-1}$ is the σ-field generated by $\{x(t); t < s_{k-1}\}$ and $\mathscr{F}^*_0$ is the σ-field generated by $\{x(t); t \le -\tau'\}$. Hence

$$\mathrm{Var}\left\{\sqrt{\frac{\tau}{N}}\sum_{k=1}^{M} E(\eta_k \mid \mathscr{F}^*_{k-1})\right\} \to 0,$$

and consequently ξ_N has the same limiting distribution as $\xi'_N = M^{-1/2}\sum_{\ell=1}^{M}\{\eta_\ell - E(\eta_\ell|\mathscr{F}^*_{\ell-1})\}$. It is evident that $\{\eta_\ell - E(\eta_\ell|\mathscr{F}^*_{\ell-1}); \mathscr{F}^*_{\ell-1}; \ell = 1, \cdots, M\}$ is a martingale difference. What remains to be shown is that the sequence $\{\eta_\ell - E(\eta_\ell|\mathscr{F}^*_{\ell-1}); \ell = 1, \cdots, M\}$ satisfies the conditions (i) and (ii) of Lemma A2.4. As for (i), this is seen to be satisfied in view of the inequality

$$\begin{aligned}&E[\{\eta_\ell - E(\eta_\ell \mid \mathscr{F}^*_{\ell-1})\}^2 I\{|\eta_\ell - E(\eta_\ell \mid \mathscr{F}^*_{\ell-1})| \ge \varepsilon M\}]\\ &\qquad \le \mathrm{Var}\{\eta_\ell - E(\eta_\ell \mid \mathscr{F}^*_{\ell-1})\}^{3/2}(\varepsilon M)^{-2}\end{aligned} \tag{6.18}$$

and the relation $\mathrm{Var}\{\eta_\ell - E(\eta_\ell|\mathscr{F}^*_{\ell-1})\} \le 2\,\mathrm{Var}(\eta_\ell) < d_4$ for some constant d_4. Moreover, $\mathrm{Var}(\eta_\ell) = \mathrm{Var}(\eta_1)$ for $l = 2, \cdots, M$, $\lim_{\tau\to\infty}\mathrm{Var}(\eta_\ell) = 2\pi f(0)$ and $E\{E(\eta_\ell|\mathscr{F}^*_{\ell-1})\}^2 = O\{\tau(\tau')^{-2-\varepsilon}\}$, so that it follows that

$$\lim_{N\to\infty} M^{-1}\sum_{\ell=1}^{M}\mathrm{Var}\{\eta_\ell - E(\eta_\ell \mid \mathscr{F}^*_{\ell-1})\} = 2\pi f(0). \tag{6.19}$$

On the other hand,

$$M^{-1}\sum_{\ell=1}^{M} E(\eta_\ell^2 \mid \mathscr{F}^*_{\ell-1}) = \sum_{1\le j,k\le\tau}\tau^{-1}E(x(j)x(k)) + M^{-1}\sum_{\ell=1}^{M} E\{\eta_\ell^2 - E(\eta_\ell^2)| \mathscr{F}^*_{\ell-1}\}$$

where the first sum in the right-hand side converges to $2\pi f(0)$ as $N \to \infty$, whereas

$$\begin{aligned}&E\,|M^{-1}\sum_{\ell=1}^{M} E\{\eta_\ell^2 - E(\eta_\ell^2)| \mathscr{F}^*_{\ell-1}\}|\\ &\qquad \le M^{-1}\sum_{\ell=1}^{M}\{\tau^{-1}\sum_{t_\ell\le j,k<s_\ell} E\,|E\{x(j)x(k) - E(x(j)x(k))| \mathscr{F}^*_{\ell-1}\}| \le d_5\tau(\tau')^{-1-\eta},\end{aligned}$$

for a positive constant d_5. Since it can be assumed that $\tau(\tau')^{-1-\eta} \to 0$, $M^{-1}\sum_{\ell=1}^{M} E(\eta_\ell^2 \mid \mathscr{F}^*_{\ell-1})$

converges in probability to $2\pi f(0)$. Finally, since

$$p\text{-}\lim_{N\to\infty} M^{-1} \sum_{\ell=1}^{M} E[\{\eta_\ell - E(\eta_\ell \mid \mathscr{F}_{\ell-1}^*)\}^2 \mid \mathscr{F}_{\ell-1}^*] = p\text{-}\lim_{N\to\infty} M^{-1} \sum_{\ell=1}^{M} E(\eta_\ell^2 \mid \mathscr{F}_{\ell-1}^*),$$

it is seen that the condition (ii) of Lemma A2.4 is satisfied.

PROOF OF THEOREM 2.2. Represent $z(n)$ as $z(n) = z^{(1)}(n) + z^{(2)}(n)$, where $z^{(1)}(n) = \sum_{j=1}^{M} G(j)e(n-j)$ and $z^{(2)}(n) = \sum_{j=M+1}^{\infty} G(j)e(n-j)$, and let $C_{\alpha\beta}^{(k,j)}(m) = N^{-1}\sum_{s=1}^{N-m} z_\alpha^{(k)}(s) z_\beta^{(j)}(s+m)$, $k, j = 1, 2$. For fixed M, each of $\sqrt{N}\{C_{\alpha\beta}^{(1,1)}(m) - E(C_{\alpha\beta}^{(1,1)}(m))\}$ $\cdot$ $(\alpha, \beta = 1, \cdots, s;\ m = 0, \cdots, L)$ is a finite linear combination of the terms

$$N^{-1/2}\{\sum_{n=1}^{N} e_\alpha(n) e_\beta(n+m) - \delta(m, 0) K_{\alpha\beta}\},$$

plus a term which converges to 0 in probability. Set

$$y_{\alpha\beta,m}(n) = e_\alpha(n) e_\beta(n+m) - \delta(m, 0) K_{\alpha\beta}$$

and let $y(n)$ be the vector whose elements are the $y_{\alpha\beta,m}(n)$ suitably ordered. Then $\{y(n)\}$ is a zero-mean second-order stationary process. In view of the assumption (i), $\mathrm{Var}\{E(y_{\alpha\beta,m}(n) \mid \mathscr{B}(n-\tau))\} = O(\tau^{-2-\varepsilon})$. Also, it is easy to derive the relation

$$E \mid E\{y_{\alpha_1\beta_1,m_1}(n_1) y_{\alpha_2\beta_2,m_2}(n_2) \mid \mathscr{B}(n_1-\tau)\} - E\{y_{\alpha_1\beta_1,m_1}(n_1) y_{\alpha_2\beta_2,m_2}(n_2)\} \mid = O(\tau^{-1-\eta}),\ n_1 < n_2,$$

from the assumptions (i) and (ii). Furthermore, the process $\{y(n)\}$ has a continuous spectral density, since

$$\begin{aligned} &\mathrm{Cov}\{y_{\alpha_1\beta_1,m_1}(n_1), y_{\alpha_2\beta_2,m_2}(n_2)\} \\ &= \int_{-\pi}^{\pi} \exp\{i(n_2-n_1)\omega\}\Big[\frac{1}{2\pi} K_{\alpha_1\alpha_2} K_{\beta_1\beta_2} \exp\{i(m_2-m_1)\omega\} \\ &\qquad + \frac{1}{2\pi} K_{\alpha_1\beta_2} K_{\beta_1\alpha_2} \exp\{i(m_2-m_1)\omega\} + h(\omega)\Big] d\omega, \end{aligned}$$

where

$$h(\omega) = \iint_{-\pi}^{\pi} \exp\{i\omega_1 m_1 + i\omega_2 m_2\} \tilde{Q}^e_{\alpha_1\beta_1\alpha_2\beta_2}(\omega_1, -\omega_2+\omega, \omega_2)\, d\omega_1\, d\omega_2,$$

and $h(\omega)$ is continuous. Consequently the process $\{y(n)\}$ satisfies the conditions of Theorem 2.1 and $N^{-1/2}\sum_{n=1}^{N} y(n)$ has an asymptotic multivariate normal distribution. In order to complete the proof, in view of Theorem 2.1, it suffices to show that the asymptotic variance of $\sqrt{N}\, C_{\alpha\beta}^{(k,j)}(m)$ converges to 0 as $M \to \infty$ if k and j are not both 1. The asymptotic variance of $\sqrt{N}\, C_{\alpha\beta}^{(k,j)}(m)$ is given as

$$\begin{aligned} \lim_{N\to\infty}\mathrm{Var}\{\sqrt{N} C_{\alpha\beta}^{(k,j)}(m)\} &= 2\pi \int_{-\pi}^{\pi} \{f_{\alpha\alpha}^{(k,k)}(\omega) f_{\beta\beta}^{(j,j)}(\omega) + f_{\alpha\beta}^{(k,j)}(\omega)\overline{f_{\beta\alpha}^{(j,k)}(\omega)} e^{i2m\omega}\}\, d\omega \\ &\quad + 2\pi \sum_{\beta_1,\dots,\beta_4=1}^{p} \iint_{-\pi}^{\pi} \exp\{im(\omega_1+\omega_2)\} \\ &\quad \cdot k_{\alpha\beta_1}^{(k)}(\omega_1) k_{\beta\beta_2}^{(j)}(-\omega_1) k_{\alpha\beta_3}^{(k)}(\omega_2) k_{\beta\beta_4}^{(j)}(-\omega_2) \tilde{Q}^e_{\beta_1\dots\beta_4}(\omega_1, -\omega_2, \omega_2)\, d\omega_1\, d\omega_2, \end{aligned} \tag{6.20}$$

where

$$k_{\alpha\beta}^{(1)}(\omega) = \sum_{j=0}^{M} G_{\alpha\beta}(j) e^{i\omega j} \quad \text{and} \quad k_{\alpha\beta}^{(2)}(\omega) = \sum_{j=M+1}^{\infty} G_{\alpha\beta}(j) e^{i\omega j}.$$

Hannan (1976) showed that the first integral in the right-hand side above can be made arbitrarily small for sufficiently large M for k, j not both 1. Whereas in view of the relationship

$$\left| 2\pi \sum_{\beta_1,\cdots,\beta_4=1}^{p} \iint_{-\pi}^{\pi} \exp\{i(\omega_1+\omega_2)m\} k_{\alpha\beta_1}^{(k)}(\omega_1) \cdots k_{\beta\beta_4}^{(j)}(-\omega_2) \tilde{Q}_{\beta_1\cdots\beta_4}^{e}(\omega_1, -\omega_2, \omega_2)\, d\omega_1\, d\omega_2 \right|$$

$$\leq d \sum_{\beta_1,\cdots,\beta_4=1}^{p} \left\{ \int_{-\pi}^{\pi} |k_{\alpha\beta_1}^{(k)}(\omega)|^4\, d\omega \int_{-\pi}^{\pi} |k_{\beta\beta_2}^{(j)}(\omega)|^4\, d\omega \int_{-\pi}^{\pi} |k_{\alpha\beta_3}^{(k)}(\omega)|^4\, d\omega \int_{-\pi}^{\pi} |k_{\beta\beta_4}^{(j)}(\omega)|^4\, d\omega \right\}^{1/4},$$

with d a positive constant, if k and j are not both 1, the second term in (6.20) converges to 0 as $M \to \infty$, since each of $k_{\alpha\beta}(\omega)$ is 4th order integrable and thus $\int_{-\pi}^{\pi} | k_{\alpha\beta}^{(k)}(\omega)|^4\, d\omega$ tends to 0 if $k = 2$.

6.2. *Proofs for Section 3.*

PROOF OF LEMMA 3.1. The proofs of (a) and (b) follow from Theorem 1 of Taniguchi (1979).

In the following we shall prove (c).

For almost every $\omega \in [-\pi, \pi]$ we have

$$\begin{aligned} &\log \det f_t(\omega) + \operatorname{tr}\{f_t(\omega)\}^{-1} f_\theta(\omega) \\ &\quad = \log \det f_\theta(\omega) - \log \det f_\theta(\omega)\{f_t(\omega)\}^{-1} + \operatorname{tr}\{f_t(\omega)\}^{-1} f_\theta(\omega) \\ &\quad = \log \det f_\theta(\omega) + \textstyle\sum_{j=1}^{s} \{\lambda_j(\omega) - \log \lambda_j(\omega) - 1\} + s, \end{aligned}$$

where $\lambda_j(\omega)$ denotes the jth latent root of $\{f_t(\omega)\}^{-1} f_\theta(\omega)$. Now $\lambda_j(\omega) - \log \lambda_j(\omega) - 1 \geq 0$ and the equality holds if and only if $\lambda_j(\omega) = 1, j = 1, \cdots, s$ (i.e., $f_t(\omega) = f_\theta(\omega)$). This implies

$$\int_{-\pi}^{\pi} \{\log \det f_t(\omega) + \operatorname{tr}\{f_t(\omega)\}^{-1} f_\theta(\omega)\}\, d\omega \geq \int_{-\pi}^{\pi} \{\log \det f_\theta(\omega) + s\}\, d\omega,$$

and equality holds if and only if $f_t(\omega) = f_\theta(\omega)$ a.s. □

The following lemma is established in the same way as in Theorem 2 of Taniguchi (1979).

LEMMA A3.2. *Suppose that $T(f)$ exists uniquely and lies in* Int Θ, *and that*

$$M_f = \int_{-\pi}^{\pi} \left[\frac{\partial^2}{\partial\theta\partial\theta'} \operatorname{tr}\{f_\theta(\omega)\}^{-1} f(\omega) + \frac{\partial^2}{\partial\theta\partial\theta'} \log \det f_\theta(\omega) \right]_{\theta = T(f)} d\omega$$

is a nonsingular matrix. Then for every sequence of spectral density matrices $\{f_N\}$ satisfying $f_N \to_w f$, we have

$$T(f_N) = T(f) - \int_{-\pi}^{\pi} M_f^{-1} \cdot \frac{\partial}{\partial\theta} (\operatorname{tr}[\{f_\theta(\omega)\}^{-1} \cdot \{f_N(\omega) - f(\omega)\}])_{\theta = T(f)}\, d\omega$$

$$+ a_N \int_{-\pi}^{\pi} \frac{\partial}{\partial\theta} (\operatorname{tr}[\{f_\theta(\omega)\}^{-1} \cdot \{f_N(\omega) - f(\omega)\}])_{\theta = T(f)}\, d\omega,$$

where a_N is a $q \times q$ matrix which tends to zero as $N \to \infty$.

For the purpose of estimating $T(f)$, the next lemma is necessary.

LEMMA A.3.3. *Assume that* $\{z(n)\}$ *defined in* (2.1) *satisfies the conditions* (i)–(iv) *in Theorem* 2.2, *and* (v) *in Theorem* 3.1. *Let* $\phi_j(\omega)$, $j = 1, \cdots, q$, *be* $s \times s$ *matrix-valued continuous functions on* $[-\pi, \pi]$ *such that* $\phi_j(\omega) = \phi_j(\omega)^*$. *Then*

(1) $p\text{-}\lim_{N\to\infty} \int_{-\pi}^{\pi} \mathrm{tr}\{I_z(\omega)\phi_j(\omega)\}\, d\omega = \int_{-\pi}^{\pi} \mathrm{tr}\{f(\omega)\phi_j(\omega)\}\, d\omega$,

(2) *the quantities*

$$\sqrt{N} \int_{-\pi}^{\pi} \mathrm{tr}[\{I_z(\omega) - f(\omega)\}\phi_j(\omega)]\, d\omega, \qquad j = 1, \cdots, q,$$

have, asymptotically, a normal distribution with zero mean vector and covariance matrix V *whose* (j, ℓ) *element is*

$$4\pi \int_{-\pi}^{\pi} \mathrm{tr}\{f(\omega)\phi_j(\omega)f(\omega)\phi_\ell(\omega)\}\, d\omega + 2\pi \sum_{r,t,u,v=1}^{s} \iint_{-\pi}^{\pi} \phi_{rt}^{(j)}(\omega_1)\phi_{uv}^{(\ell)}(\omega_2)\tilde{Q}_{rtuv}^{z}(-\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2,$$

where $\phi_{rt}^{(j)}(\omega)$ *is the* (r, t)*-th element of* $\phi_j(\omega)$.

PROOF. We can see that

$$\begin{aligned} Var \int_{-\pi}^{\pi} \mathrm{tr}\, [\{I_z(\omega) - EI_z(\omega)\}\phi_j(\omega)]\, d\omega &= \mathrm{Var} \sum_{\alpha,\beta=1}^{s} \int_{-\pi}^{\pi} [\{I_{\alpha\beta}^z(\omega) - EI_{\alpha\beta}^z(\omega)\}\phi_{\beta\alpha}^{(j)}(\omega)]\, d\omega \\ &\le 2s^2 \sum_{\alpha,\beta=1}^{s} \mathrm{Var} \int_{-\pi}^{\pi} \{I_{\alpha\beta}^z(\omega) - EI_{\alpha\beta}^z(\omega)\}\phi_{\beta\alpha}^{(j)}(\omega)\, d\omega. \end{aligned}$$

In view of Lemma A2.2, the last expression is at most of order $O(N^{-1})$. Now we shall evaluate the bias term. Noting that $EI_z(\omega)$ is the Cesaro sum of $f(\omega)$ and that $f(\omega) \in \mathrm{Lip}(\alpha)$, $\alpha > 1/2$, we have $\mathrm{tr}\{EI_{\gamma\beta}^z(\omega) - f_{\gamma\beta}(\omega)\} = O(N^{-\alpha})$, uniformly in ω (e.g., Hannan, 1970, page 513). Thus we have

$$\begin{aligned} \int_{-\pi}^{\pi} \left| \sqrt{N} \int_{-\pi}^{\pi} \mathrm{tr}[\{EI_z(\omega) - f(\omega)\}\phi_j(\omega)]\, d\omega \right| &= \max_{(\beta,\gamma)}\sup_\omega |\, \phi_{\beta\gamma}^{(j)}(\omega)| \sum_{\gamma,\beta=1}^{s} \{\sqrt{N}\, |\, EI_{\gamma\beta}^z(\omega) - f_{\gamma\beta}(\omega)|\} \\ &= O(N^{1/2-\alpha}), \end{aligned} \tag{6.21}$$

which converges to zero as $N \to \infty$. This completes the proof of (1). Part (2) is proved as follows. For any $\varepsilon > 0$, taking M sufficiently large, we have the Cesaro sum

$$\phi_{(r,t),M}^{(j)}(\omega) = \frac{1}{2\pi} \sum_{n=-M+1}^{M-1} \phi_{(r,t),M}^{(j)}(n)\left(1 - \frac{|\, n\,|}{M}\right)\exp(-in\omega)$$

such that

$$\max_{r,t=1,\ldots,s} \sup_{\omega \in [-\pi,\pi]} | \phi_{rt}^{(j)}(\omega) - \phi_{(r,t),M}^{(j)}(\omega)| < \varepsilon.$$

Hereafter let $\phi_{j,M}(\omega) = \{\phi_{(r,t),M}^{(j)}(\omega)\}$ and $\phi_{j,M}(n) = \{\phi_{(r,t),M}^{(j)}(n)\}$. In the integral

$$\text{(6.22)} \qquad \sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}[\{I_z(\omega) - EI_z(\omega)\}\phi_j(\omega)]\, d\omega,$$

$\phi_j(\omega)$ may be replaced by $\phi_{j,M}(\omega)$. For that purpose, put

$$\delta_M(\omega) = \phi_j(\omega) - \phi_{j,M}(\omega) = \{\delta_{(\alpha,\beta),M}(\omega)\},$$

and evaluate

$$\text{(6.23)} \qquad \begin{aligned} &\operatorname{Var}\left(\sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}[\{I_z(\omega) - EI_z(\omega)\}\delta_M(\omega)]\, d\omega\right) \\ &= \operatorname{Var}\left[\textstyle\sum_{\alpha,\beta=1}^{s} \sqrt{N} \displaystyle\int_{-\pi}^{\pi} \{I_{\alpha\beta}^{z}(\omega) - EI_{\alpha\beta}^{z}(\omega)\}\delta_{(\beta,\alpha),M}(\omega)\, d\omega\right]. \end{aligned}$$

In view of Lemma A2.2, (6.23) is dominated by

$$\text{(6.24)} \qquad \begin{aligned} 2s^2 \textstyle\sum_{\alpha,\beta=1}^{s} \Bigg\{ &2\pi \displaystyle\int_{-\pi}^{\pi} | \delta_{(\beta,\alpha),M}(\omega)|^2 f_{\alpha\alpha}(\omega) f_{\beta\beta}(\omega)\, d\omega \\ &+ 2\pi \int_{-\pi}^{\pi} \delta_{(\beta,\alpha),M}(\omega)\overline{\delta_{(\alpha,\beta),M}(\omega)} f_{\alpha\beta}(\omega)\overline{f_{\beta\alpha}(\omega)}\, d\omega \\ &+ 2\pi \iint_{-\pi}^{\pi} \delta_{(\beta,\alpha),M}(\omega_1)\delta_{(\beta,\alpha),M}(-\omega_2)\tilde{Q}_{\alpha\beta\alpha\beta}^{z}(\omega_1, \omega_2, -\omega_2)\, d\omega_1\, d\omega_2 \Bigg\}, \end{aligned}$$

which tends to zero as $M \to \infty$. By Bernstein's lemma (e.g., Hannan, 1970, page 242), the asymptotic normality of $\sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}[\{I_z(\omega) - f(\omega)\}\phi_j(\omega)]\, d\omega$ is equivalent to that of

$$\text{(6.25)} \qquad \sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}[\{I_z(\omega) - EI_z(\omega)\}\phi_{j,M}(\omega)]\, d\omega,$$

for each M. The above (6.25) is equal to

$$\text{(6.26)} \qquad \frac{1}{2\pi} \operatorname{tr} \textstyle\sum_{n=-M+1}^{M-1} \left(1 - \dfrac{|n|}{M}\right) \sqrt{N} \left\{C_z(-n) - \left(1 - \dfrac{|n|}{N}\right)\gamma_z(-n)\right\}\phi_{j,M}(n).$$

The asymptotic normality for (6.26) follows from Theorem 2.2. Now evaluate the asymptotic covariance

$$\text{(6.27)} \qquad \begin{aligned} &\operatorname{Cov}\left[\sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}\{I_z(\omega)\phi_j(\omega)\}\, d\omega, \quad \sqrt{N} \int_{-\pi}^{\pi} \operatorname{tr}\{I_z(\omega)\phi_{\ell}(\omega)\}\, d\omega\right] \\ &= N \operatorname{Cov}\left\{\int_{-\pi}^{\pi} \textstyle\sum_{\alpha_1,\alpha_2=1}^{s} I_{\alpha_1\alpha_2}^{z}(\omega)\phi_{\alpha_2\alpha_1}^{(j)}(\omega)\, d\omega, \quad \displaystyle\int_{-\pi}^{\pi} \textstyle\sum_{\alpha_3,\alpha_4=1}^{s} I_{\alpha_3\alpha_4}^{z}(\omega)\phi_{\alpha_4\alpha_3}^{(\ell)}(\omega)\, d\omega\right\}. \end{aligned}$$

By Lemma A2.2, (6.27) converges to

$$\sum_{\alpha_1,\dots,\alpha_4=1}^{s}\Big\{2\pi\int_{-\pi}^{\pi}\phi^{(j)}_{\alpha_2\alpha_1}(\omega)\bar{\phi}^{(\ell)}_{\alpha_4\alpha_3}(\omega)f_{\alpha_1\alpha_3}(\omega)\overline{f_{\alpha_2\alpha_4}(\omega)}\,d\omega$$

$$+2\pi\int_{-\pi}^{\pi}\phi^{(j)}_{\alpha_2\alpha_1}(\omega)\bar{\phi}^{(\ell)}_{\alpha_4\alpha_3}(-\omega)f_{\alpha_1\alpha_4}(\omega)\overline{f_{\alpha_2\alpha_3}(\omega)}\,d\omega$$

$$+2\pi\iint_{-\pi}^{\pi}\phi^{(j)}_{\alpha_2\alpha_1}(\omega_1)\phi^{(\ell)}_{\alpha_4\alpha_3}(-\omega_2)\tilde{Q}^{z}_{\alpha_1\cdots\alpha_4}(\omega_1,\omega_2,-\omega_2)\,d\omega_1\,d\omega_2\Big\}.$$

Noting that

$$\bar{\phi}^{(j)}_{\alpha\beta}(\omega)=\phi^{(j)}_{\beta\alpha}(\omega),\qquad \phi^{(\ell)}_{\alpha\beta}(-\omega)=\phi^{(\ell)}_{\beta\alpha}(\omega)\quad\text{and}\quad\overline{f_{\alpha\beta}(\omega)}=f_{\beta\alpha}(\omega),$$

we have the desired results.

PROOF OF THEOREM 3.1. Lemma A3.3 implies that $I_z(\omega)\to_w f(\omega)$, in probability. Thus Theorem 3.1 follows from Lemmas A3.2 and A3.3.

PROOF OF PROPOSITION 3.1. Because of (3.2), we can see $\tilde{Q}^{e}_{abcd}(\omega_1,\omega_2,\omega_3)=\kappa_{abcd}/(2\pi)^3$. By Lemma A2.1, we have

$$\tilde{Q}^{z}_{rtuv}(-\omega_1,\omega_2,-\omega_2)=\sum_{a,b,c,d=1}^{p}k_{ra}(-\omega_1)k_{tb}(\omega_1)k_{uc}(-\omega_2)k_{vd}(\omega_2)\kappa_{abcd}/(2\pi)^3.$$

Thus we have

$$2\pi\sum_{r,t,u,v=1}^{s}\iint_{-\pi}^{\pi}\frac{\partial}{\partial\theta_j}f_\theta^{(r,t)}(\omega_1)\frac{\partial}{\partial\theta_\ell}f_\theta^{(u,v)}(\omega_2)\big|_{\theta=T(f)}\times\tilde{Q}^{z}_{rtuv}(-\omega_1,\omega_2,-\omega_2)\,d\omega_1\,d\omega_2$$

$$=\sum_{a,b,c,d=1}^{s}\kappa_{abcd}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}k^*(\omega)\frac{\partial}{\partial\theta_j}\{f_\theta(\omega)\}^{-1}k(\omega)\,d\omega\right]_{ab}$$

$$\times\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}k^*(\omega)\frac{\partial}{\partial\theta_\ell}\{f_\theta(\omega)\}^{-1}k(\omega)\,d\omega\right]_{cd}\Bigg|_{\theta=T(f)}.$$

REFERENCES

ACHIESER, N. I. (1956). *Theory of Approximation.* Unger, New York.

BROWN, B. M. (1971). Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.

DUNSMUIR, W. and HANNAN, E. J. (1976). Vector linear time series models. *Adv. in Appl. Probab.* **8** 339–364.

DUNSMUIR, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Ann. Statist.* **7** 490–506.

FISHER, R. A. (1925). Theory of statistical estimation. *Proc. Cambridge Philos. Soc.* **22** 700–725.

GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York.

HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

HANNAN, E. J. (1973). The asymptotic theory of linear time series models. *J. Appl. Probab.* **10** 130–145.

HANNAN, E. J. (1976). The asymptotic distribution of serial covariances. *Ann. Statist.* **4** 396–399.

HANNAN, E. J. and ROBINSON, P. M. (1973). Lagged regression with unknown lags. *J. Roy. Statist. Soc.* **35** 252–267.

HOSOYA, Y. (1974). *Estimation problems on stationary time-series models.* Ph.D. thesis, Yale Univ.

IBRAGIMOV, I. A. and LINNIK, YU. (1971). *Independent Stationary Sequences of Random Variables.* Wolters Noordhoff, Groningen.

IBRAGIMOV, I. A. and SOLEV, V. N. (1969). A condition for regularity of a Gaussian stationary process. *Soviet Math. Dokl.* **10** 371–375.

ROBINSON, P. M. (1980). Continuous model fitting from discrete data. *Directions in Time Series.* Inst. Math. Statist. Bull. 263–278.

ROZANOV, YU. A. (1967). *Stationary Random Processes.* Holden-Day, San Francisco.

TANIGUCHI, M. (1979). On estimation of parameters of Gaussian stationary processes. *J. Appl. Prob.* **16** 575–591.

WALKER, A. M. (1964). Asymptotic properties of least-square estimates of parameters of the spectrum of a stationary nondeterministic time series. *J. Austral. Math. Soc.* **4** 363–384.

WHITTLE, P. (1952). Some result in time series analysis. *Skandivanisk Aktuarietidskrift.* **35** 48–60.

WHITTLE, P. (1962). Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist.* **39-2** 105–129.

FACULTY OF ECONOMICS
TOHOKU UNIVERSITY
KAWAUCHI, SENDAI 980
JAPAN

DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ENGINEERING
HIROSHIMA UNIVERSITY
SENDAMACHI-3, HIROSHIMA 730
JAPAN