Proceedings of the
Twelfth Biennial Seminar of the
Canadian Mathematical Congress

on

Time Series and Stochastic Processes; Convexity and Combinatorics

---

With contributions by
R. M. Blumenthal, E. D. Bolker, D. R. Brillinger, K. L. Chung, Ky Fan, J. Gani, R. K. Getoor, B. Grünbaum, G. J. Leppink, R. M. Loynes, E. Parzen, A. Rényi, M. Rosenblatt, H. Ruben, J. Sacks, D. Ylvisaker

Edited by Ronald Pyke

Canadian Mathematical Congress
Société Mathématique du Canada
Montreal, Canada
1970

The frequency analysis of relations between stationary spatial series

David R. Brillinger

The London School of Economics
and Political Science, and University of California, Berkeley

1. Introduction. Let R^p denote real Euclidean p-space. The intention of this work is to develop several aspects of the frequency analysis of vector-valued spatial series $\mathbf{X}(t)$, $t \in R^p$. The results obtained will be applicable to the analysis of data recorded at a succession of time points, date recorded at points in some region of a surface and date recorded at points with both space and time coordinates. Data of the first sort is common in electrical engineering, Lee (1960), geophysics, Tukey (1965) and economics, Granger and Hatanaka (1964). Data of the second sort is common in optics, Goodman (1968), forestry, Matérn (1960) and geology, Harbaugh and Preston (1968). Data of the final sort appears in turbulence theory, Batchelor (1953), oceanography, Pierson and Tick (1955) and seismology, Robinson (1967).

General discussions of stationary spatial series have been given in: Jaglom (1957, 1961, 1962), Bartlett (1966), Blanc-Lapierre (1963) and Whittle (1964). When one examines this literature it becomes apparent that the terminology of the subject is not at all standard. Spatial series have been called: *random fields, spatial random functions, multidimensional processes, random functions of several variables, multiparameter time series* and *random surfaces.* We prefer the term spatial series because of its immediate relation with the term time series. Other references to the theory of spatial series include: Wiener (1930), Kampé de Fériet (1948), Bochner (1955, 1960), Grenander and Rosenblatt (1957), Cramér and Leadbetter (1967), Birendranath (1951) and Hannan (1965).

The principal conditions that we will set down for our analysis of spatial series are an invariance of the probabilistic structure of the series under translations of the space argument and a weak statistical dependence of values of the series that are far apart. These requirements seem reasonable for many series that occur in practice. We define a variety of parameters for the series and then develop statistical properties of estimates of the parameters. The main mathematical tool required in our work is the Fourier analysis of functions of several variables. This last is described in Bochner and Chandrasekharan (1949) and Bochner (1960) for example.

Some additional references to the applications of spatial series include: Whittle (1954, 1962), Embree et al. (1963), Fail and Grau (1963), Hoffman (1964), Rosenblatt (1957), Lin (1961), Tatarski (1961), Bracewell (1956), Preston (1965), O'Neill (1956), Heine (1955), Kovasnay and Arman (1957), Vander Lugt (1964), Hannan (1966), Zurflueh (1967), Curry (1969), Granger (1968), Rao (1969), Gazdag (1970), Robinson (1970).

Before commencing our analysis we set down some notation. R^p will denote real Euclidean p-space. $\lambda = (\lambda_1, \ldots, \lambda_p)$ and $t = (t_1, \ldots, t_p)$ are elements in R^p. For scalars α, β, $\alpha\lambda + \beta t = (\alpha\lambda_1 + \beta t_1, \ldots, \alpha\lambda_p + \beta t_p)$. (λ, t) denotes $\lambda_1 t_1 + \ldots + \lambda_p t_p$. $|t|$ denotes $(t, t)^{1/2}$. $i = \sqrt{-1}$. If z is complex-valued, then $\bar{z}$ is its complex conjugate. We write $z = \operatorname{Re} z + i \operatorname{Im} z$ to display its real and imaginary parts. Bold-face letters $\mathbf{A}$, $\mathbf{X}$ will denote matrices. If $\mathbf{X}$ is an $r \times s$ matrix, then its entries will be denoted $X_{jk}, j = 1, \ldots, r, k = 1, \ldots, s$. $\mathbf{X}^\tau$ denotes the transpose of $\mathbf{X}$. $\|\mathbf{X}\|$ denotes $(\text{trace } \mathbf{X}\bar{\mathbf{X}}^\tau)^{1/2}$. For $t \in R^p$, $\delta\{t\}$ denotes the Kronecker delta $\delta\{t\} = 1$ if $t = 0$ and $= 0$ otherwise, while $\delta(t)$ denotes the Dirac delta function. vec $\mathbf{A}$ will denote the column vector obtained from a matrix $\mathbf{A}$ by placing its columns under one another successively. $\mathbf{A} \otimes \mathbf{B}$ will denote the Kronecker product of $\mathbf{A}$ with $\mathbf{B}$. It is the block matrix $[\mathbf{A}B_{jk}]$.

If $(X_1, \ldots, X_r)$ is an r-variate random variable, then $\operatorname{cum}\{X_1, \ldots, X_r\}$ will denote the joint cumulant of $X_1, \ldots, X_r$. It is the coefficient of $t_1 \ldots t_r$ in the expansion of the cumulant generating function,

$$\log E \exp \sum X_j t_j. \tag{1.1}$$

Important properties of cumulants include:

(i) $\operatorname{cum}\{X_1, \ldots, X_r\}$ is symmetric in its arguments.

(ii) $\operatorname{cum}\{X_1, \ldots, X_r\}$ is multilinear in its arguments. For example

$$\operatorname{cum}\{a_1 X_1, \ldots, a_r X_r\} = a_1 \ldots a_r \operatorname{cum}\{X_1, \ldots, X_r\} \tag{1.2}$$

for constants $a_1, \ldots, a_r$.

Also

$$\operatorname{cum}\{X_1 + Y_1, X_2, \ldots, X_r\} = \operatorname{cum}\{X_1, \ldots, X_r\} + \operatorname{cum}\{Y_1, X_2, \ldots, X_r\}. \tag{1.3}$$

(iii) If μ is a constant and $r = 2, 3, \ldots$, then

$$\text{cum}\{X_1 + \mu, X_2, \ldots, X_r\} = \text{cum}\{X_1, X_2, \ldots, X_r\}. \tag{1.4}$$

(iv) If any proper subset of $X_1, \ldots, X_r$ is statistically independent of the set of remaining X's, then

$$\text{cum}\{X_1, \ldots, X_r\} = 0. \tag{1.5}$$

(v) If the random variates $(X_1, \ldots, X_r)$, and $(Y_1, \ldots, Y_r)$ are statistically independent, then

$$\begin{aligned}\text{cum}\{X_1 + Y_1, \ldots, X_r + Y_r\} &= \text{cum}\{X_1, \ldots, X_r\} \\ &\quad + \text{cum}\{Y_1, \ldots, Y_r\}.\end{aligned} \tag{1.6}$$

In our work, property (iv), will be especially important. It suggests how joint cumulants measure the statistical dependence of variates. Properties (i)—(v) may all be deduced directly from the definition of $\text{cum}\{X_1, \ldots, X_r\}$ through the cumulant generating function. An alternate definition is provided by

$$\text{cum}\{X_1, \ldots, X_r\} = \sum (-1)^{p-1}(p-1)!\left\{E\prod_{j\in v_1} X_j \cdots E\prod_{j\in v_p} X_j\right\} \tag{1.7}$$

where the summation extends over all partitions $(v_1, \ldots, v_p)$, $p = 1, \ldots, r$ of the integers $(1, \ldots, r)$. This definition may be employed in the case that some of the X_j are complex-valued. The inverse relation to (1.7) is

$$EX_1 \ldots X_r = \sum (\text{cum}\{X_j; j \in v_1\}) \ldots (\text{cum}\{X_j; j \in v_p\}), \tag{1.8}$$

An important property of cumulants is that one can set down expressions for the joint cumulants of polynomials in random variates whose joint cumulants are given, see Leonov and Širjaev (1959). $N_r(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ will denote an r vector-valued normal variate with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $N_r^C(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ will denote a complex r vector-valued normal variate. If $\mathbf{X}$ is $N_r^C(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$\begin{bmatrix}\text{Re}\,\mathbf{X} \\ \text{Im}\,\mathbf{X}\end{bmatrix} = N_{2r}\left(\begin{bmatrix}\text{Re}\,\boldsymbol{\mu}, \\ \text{Im}\,\boldsymbol{\mu},\end{bmatrix}\frac{1}{2}\begin{bmatrix}\text{Re}\,\boldsymbol{\Sigma} & -\text{Im}\,\boldsymbol{\Sigma} \\ \text{Im}\,\boldsymbol{\Sigma} & \text{Re}\,\boldsymbol{\Sigma}\end{bmatrix}\right). \tag{1.9}$$

We note that

$$E[(\mathbf{X} - \boldsymbol{\mu})(\overline{\mathbf{X} - \boldsymbol{\mu}})^\tau] = \boldsymbol{\Sigma} \tag{1.10}$$

while

$$E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\tau] = \mathbf{0} \tag{1.11}$$

for such a variate.

$W_r(n, \Sigma)$ denotes an $r \times r$ matrix-valued Wishart variate with n degrees of freedom. $W_r^C(n, \Sigma)$ denotes a complex $r \times r$ matrix-valued Wishart variate with n degrees of freedom. (See Goodman (1963).)

If $\mathbf{Y}$, $\mathbf{Z}$ are complex vector-valued random variables, we define their cross-covariance matrix by

$$\text{cov}\{\mathbf{Y}, \mathbf{Z}\} = E[(\mathbf{Y} - E\mathbf{Y})(\overline{\mathbf{Z} - E\mathbf{Z}})^{\tau}]. \tag{1.12}$$

2. Stationary spatial series. Let $(\Omega, \mathcal{F}, \mu)$, a probability space and R^p be given. By an *r vector-valued spatial series* we shall mean an r vector-valued function $\mathbf{X}(t, \omega)$, $t \in R^p$, $\omega \in \Omega$, with real-valued components which for every fixed t are measurable functions of ω. As little confusion seems possible in the present context, we shall not indicate explicitly the dependence on ω and write $\mathbf{X}(t)$ instead of $\mathbf{X}(t, \omega)$.

We shall say that the r vector-valued spatial series $\mathbf{X}(t)$ is *strictly stationary* if the probability distribution of any finite collection of values

$$X_{a_1}(t^{(1)} + u), \ldots, X_{a_k}(t^{(k)} + u), \tag{2.1}$$

$t^{(1)}, \ldots, t^{(k)}, u \in R^p$; $a_1, \ldots, a_k = 1, \ldots, r$; $k = 1, 2, \ldots$ does not depend on u. We will suppose that moments of all orders exist for $X_j(t)$, $j = 1, \ldots, r$, and define:

$$\begin{aligned} &c_{a_1 \ldots a_k}(u^{(1)}, \ldots, u^{(k-1)}) \\ &\quad = \text{cum}\{X_{a_1}(u^{(1)} + t), \ldots, X_{a_{k-1}}(u^{(k-1)} + t), X_{a_k}(t)\}, \end{aligned} \tag{2.2}$$

$u^{(1)}, \ldots, u^{(k-1)}$, $t \in R^p$; $a_1, \ldots, a_k = 1, \ldots, r$; $k = 1, 2, \ldots$. This *cumulant function* does not depend on t in view of the stationarity condition on $\mathbf{X}(t)$. We now set down our principal assumption concerning $\mathbf{X}(t)$.

ASSUMPTION 2.1. *$\mathbf{X}(t)$, $t \in R^p$ is a strictly stationary r vector-valued spatial series all of whose moments exist and satisfy*:

$$\begin{aligned} &\int \cdots \int \{|u^{(1)}| + \ldots + |u^{(k-1)}|\} \\ &\qquad \cdot |c_{a_1 \ldots a_k}(u^{(1)}, \ldots, u^{(k-1)})| \, du^{(1)} \ldots du^{(k-1)} < \infty \end{aligned} \tag{2.3}$$

for $u^{(1)}, \ldots, u^{(k-1)} \in R^p$; $a_1, \ldots, a_k = 1, \ldots, r$; $k = 1, 2, \ldots$.

In the case of a Gaussian series $\mathbf{X}(t)$, because cumulants of order greater than 2 are identically 0, (2.3) has the simple form

$$\int |u| \, |c_{ab}(u)| \, du < \infty, \qquad a, b = 1, \ldots, r. \tag{2.4}$$

We have seen that cumulants measure the extent of dependence of random variables. (2.3) is seen to be a requirement that well separated

values of the series are at most weakly statistically dependent. As such it is a form of mixing condition.

If the series $\mathbf{X}(t)$ satisfies Assumption 2.1, then we may define the *cumulant spectra of order k* by

$$f_{a_1 \ldots a_k}(\lambda^{(1)}, \ldots, \lambda^{(k-1)})$$

$$= (2\pi)^{-p(k-1)} \int \cdots \int c_{a_1 \ldots a_k}(u^{(1)}, \ldots, u^{(k-1)}) \cdot \exp\left\{-i \sum_{1}^{k-1} (\lambda^{(j)}, u^{(j)})\right\} du^{(1)} \ldots du^{(k-1)} \tag{2.5}$$

$\lambda^{(1)}, \ldots, \lambda^{(k-1)} \in R^p$; $a_1, \ldots, a_k = 1, \ldots, r$. In view of (2.3) they will exist and have uniformly bounded first derivatives. The cumulant spectra of order 2 will be especially important in our work. They are given by

$$f_{ab}(\lambda) = (2\pi)^{-p} \int c_{ab}(u) \exp\{-i(\lambda, u)\}\, du, \tag{2.6}$$

$\lambda \in R^p$; $a, b = 1, \ldots, r$. The argument λ of $f_{ab}(\lambda)$ is called the *spatial frequency*. $f_{aa}(\lambda)$ is called the *power spectrum* of $X_a(t)$. It may be seen to be nonnegative and symmetric about 0. $f_{ab}(\lambda)$, $a \neq b$, is called the *cross-spectrum* of $X_a(t)$ with $X_b(t)$. It will be convenient to arrange the second-order spectra into the *spectral density matrix*, $\mathbf{f}_{XX}(\lambda)$. The entry in the ath row and bth column is $f_{ab}(\lambda)$, $a, b = 1, \ldots, r$. $\mathbf{f}_{XX}(\lambda)$ may be seen to be a Hermitian nonnegative definite matrix. Theorem 2.21 of Bochner (1960) then implies that the entries of $\mathbf{f}_{XX}(\lambda)$ are absolutely integrable. This in turn implies that the series $\mathbf{X}(t)$ is continuous in mean of all orders.

Processes that satisfy a stationarity condition and for which $c_{aa}(u)$ is continuous at $u = 0$, $a = 1, \ldots, r$, have a Cramér representation given by

$$\mathbf{X}(t) = \int_{-\infty}^{\infty} \exp\{i(t, \lambda)\}\, d\mathbf{Z}_X(\lambda). \tag{2.7}$$

Here $\mathbf{Z}_X(\lambda)$, $\lambda \in R^p$, is an r vector-valued random function with orthogonal increments. (See Jaglom (1957), Cramér and Leadbetter (1967) for example.) We will indicate the source of the representation (2.7) later. It is of use in the investigation of the effect of operations on the series.

It is important to be able to understand the effect of certain manipulatory operations on a spatial series of interest. An important class of operations consists of those that are *linear and space invariant*. Consider an operation $\mathscr{A}$ whose domain is r vector-valued series $\mathbf{X}(t)$, $t \in R^p$, and whose range is s vector-valued series $\mathbf{Y}(t)$, $t \in R^p$. The operation is linear if

$$\mathscr{A}[\alpha_1 \mathbf{X}_1 + \alpha_2 \mathbf{X}_2](t) = \alpha_1 \mathscr{A}[\mathbf{X}_1](t) + \alpha_2 \mathscr{A}[\mathbf{X}_2](t), \tag{2.8}$$

when $\mathbf{X}_1(t), \mathbf{X}_2(t)$ are in the domain of $\mathscr{A}$. Let $T^u\mathbf{X}(t) = \mathbf{X}(t+u)$, $t, u \in R^p$, denote the translation operation. The operation $\mathscr{A}$ is space invariant if

$$\mathscr{A}[T^u\mathbf{X}](t) = \mathscr{A}[\mathbf{X}](t+u). \tag{2.9}$$

We may extend the domain of $\mathscr{A}$ to be $r \times r$ matrix-valued functions $\mathbf{U}(t)$, $t \in R^p$.

$$\mathscr{A}[\mathbf{U}](t) = \left[\mathscr{A}[\mathbf{U}_1](t) \ldots \mathscr{A}[\mathbf{U}_r](t)\right] \tag{2.10}$$

where $\mathbf{U}_1(t), \ldots, \mathbf{U}_r(t)$ are the columns of $\mathbf{U}(t)$. Its range is then $s \times r$ matrix-valued functions.

We now have the important lemma.

LEMMA 2.1. *Let $\mathscr{A}$ be a linear space invariant operation whose domain includes $\mathbf{e}(t) = \exp\{i(\lambda, t)\}\mathbf{I}$, for $\lambda, t \in R^p$, where $\mathbf{I}$ is the $r \times r$ identity. There is an $s \times r$ matrix function $\mathbf{A}(\lambda)$ such that*

$$\mathscr{A}[\mathbf{e}](t) = \mathbf{A}(\lambda)\exp\{i(\lambda, t)\}. \tag{2.11}$$

In other words a linear space invariant operation carries complex exponentials of spatial frequency λ over into complex exponentials of the same spatial frequency λ. The function $\mathbf{A}(\lambda)$ is called the *transfer function* of the operation. If $\mathbf{X}(t)$ has the Cramér representation (2.7) and $\mathbf{Y}(t) = \mathscr{A}[\mathbf{X}](t)$, then under regularity conditions

$$\mathbf{Y}(t) = \int \exp\{i(t, \lambda)\}\mathbf{A}(\lambda)\, d\mathbf{Z}_X(\lambda). \tag{2.12}$$

A common class of linear space invariant operations has the form

$$\mathbf{Y}(t) = \int \mathbf{a}(t-u)\mathbf{X}(u)\, du, \tag{2.13}$$

where $\mathbf{a}(t)$ satisfies some integrability condition. The function $\mathbf{a}(t)$ is called the *impulse response* of the operation. The transfer function of this operation is seen to be

$$\mathbf{A}(\lambda) = \int \mathbf{a}(u)\exp\{-i(\lambda, u)\}\, du. \tag{2.14}$$

One can show that if $\mathbf{X}(t)$ satisfies Assumption 2.1, and $\mathbf{a}(t)$ satisfies

$$\int |u|\,|\mathbf{a}(u)|\, du < \infty, \tag{2.15}$$

then $\mathbf{Y}(t)$ given by (2.13) also satisfies Assumption 2.1. If $\mathbf{f}_{YY}(\lambda)$ denotes the spectral density matrix of $\mathbf{Y}(t)$, then one has

$$\mathbf{f}_{YY}(\lambda) = \mathbf{A}(\lambda)\mathbf{f}_{XX}(\lambda)\overline{\mathbf{A}(\lambda)}^{\mathrm{T}}, \tag{2.16}$$

for example. Similar results hold for the higher order cumulant spectra, see Širjaev (1960).

Another important class of operations consists of those that are *polynomial and space invariant.* These have the property that for scalars α_1, α_2, $\mathscr{A}[\alpha_1\mathbf{X}_1 + \alpha_2\mathbf{X}_2](t)$ is a polynomial of some degree in α_1 and α_2 and the additional property (2.9). An example of such an operation is provided by

$$Y(t) = \int\int a(t-u, t-v)X(u)X(v)\,du\,dv. \tag{2.17}$$

Operations of this sort are discussed in Brillinger (1970).

We conclude the section with a definition. Under our assumptions, the second order cumulants

$$\operatorname{cov}\{X_a(u+t), X_b(t)\} = c_{ab}(u), \tag{2.18}$$

$t, u \in R^p$; $a, b = 1, \ldots, r$; depend only on u. On occasion it may prove reasonable to assume that they depend only on $|u|^2$, that is they depend only on the distance between the space points. Such a series is said to be *isotropic.* See Jaglom (1961), Itô (1956). In this case it may be shown that the second-order spectra depend only on $|\lambda|^2$, see Lord (1954), Blanc-Lapierre (1963), Bartlett (1966), Matérn (1960), for example.

3. Finite Fourier transforms. We now turn to the investigation of statistics of use in the estimation of the cumulant spectra of a series $\mathbf{X}(t)$, $t \in R^p$. We will suppose that the values of $\mathbf{X}(t)$ are available for $t \in \mathscr{R}_T$ where $\mathscr{R}_T$, $T = 1, 2, \ldots$, is a sequence of expanding regions of R^p. In the case $p = 1$, $\mathscr{R}_T$ often has the form of an interval say $[0, T]$. However in the case $p > 1$ its form can be quite complicated resulting from an experimenter's inability to record the values of $\mathbf{X}(t)$ in an elementary manner. In simple cases it may have the form of a rectangle,

$$0 \leqq t_j \leqq \alpha_j T \tag{3.1}$$

for some constants $\alpha_1, \ldots, \alpha_p$ or the form of a sphere,

$$0 \leqq t_1^2 + \ldots + t_p^2 \leqq T^2. \tag{3.2}$$

We shall build the region of observation into our results through a real-valued function $h(t)$. In this connection we set down,

ASSUMPTION 3.1. *$h(t)$, $t \in R^p$, is measurable in t, bounded, zero for $|t| > 1$, $\int h(t)^2\,dt \neq 0$ and there exists a finite K such that*

$$\int |h(t+u) - h(t)|\, dt < K|u|, \tag{3.3}$$

for all $u \in R^p$.

Given $T > 0$, $\mathscr{R}_T$ the region of observation will be taken as the set of t's in which $h(t/T)$ is nonzero. An example of an $h(t)$ of interest is

$$\begin{aligned} h(t) &= 1, \quad t \in A, \\ &= 0, \quad t \notin A, \end{aligned} \tag{3.4}$$

for some set A. We will see later that we may reduce the bias of our estimates by taking an $h(t)$ of the general form of (3.4), but without the discontinuities. Condition (3.3) is a form of integrated Lipschitz condition. It is satisfied by the characteristic functions of Jordan measurable sets, by functions $h(t)$ with uniformly bounded first derivatives and by $h(t)$ of bounded variation.

We define

$$H_k(\lambda) = \int h(t)^k \exp\{-i(\lambda, t)\}\, dt. \tag{3.5}$$

Because $h(t)$ is bounded and of compact support, $H_k(\lambda)$ exists. Because of the Riemann-Lebesgue theorem $H_k(\lambda)$ tends to 0 as $|\lambda| \to \infty$. Examples of $h(t)$ of interest and the associated $H_1(\lambda)$ are given in Table 3.1. The first integrals may be found in Bochner (1960) and Shapiro (1964).

TABLE 3.1

$h(t)$	$H_1(\lambda)$
1 for $\|t\| \leqq 1$	$(2\pi)^{-p/2} J_{p/2}(\|\lambda\|)\|\lambda\|^{-p/2}$
$(1 - \|t\|^2)^\alpha$ for $\|t\| \leqq 1$	$(2\pi)^{-p/2} 2^\alpha \Gamma(\alpha+1) J_{p/2+\alpha}(\|\lambda\|)\|\lambda\|^{-(p/2+\alpha)}$
1 for $\|t_j\| \leqq 1$	$\prod_j \{(2 \sin \lambda_j)/\lambda_j\}$
$\prod_j (1 - \|t_j\|)$ for $\|t_j\| \leqq 1$	$\prod_j (\sin \lambda_j/2)^2/(\lambda_j/2)^2$
$\exp\{\|t\|^2/(\|t\|^2 - 1)\}$ for $\|t\| < 1$	

For given $T > 0$, because of its typical shape, the function $h(t/T)$ is called a taper, see Tukey (1967). The basic data entering our frequency analysis will be the tapered values,

$$h(t/T)\mathbf{X}(t), \tag{3.6}$$

$t \in R^p$. These values are 0 for $|t| > T$.

The r vector-valued statistic,

$$\mathbf{d}_X^{(T)}(\lambda) = \int h(t/T)\mathbf{X}(t)\exp\{-i(\lambda, t)\}\,dt, \tag{3.7}$$

$\lambda \in R^p$, is called the *finite Fourier transform* of the series of tapered values. If $\mathbf{X}(t)$ has the Cramér representation (2.7), then

$$\mathbf{d}_X^{(T)}(\lambda) = T^p \int H_1(T[\lambda - \alpha])\,d\mathbf{Z}_X(\alpha). \tag{3.8}$$

In practice we would want $\mathbf{d}_X^{(T)}(\lambda)$ to be nearly proportional to $d\mathbf{Z}_X(\lambda)$; therefore $h(t)$ should be selected so that $H_1(\alpha)$ is a smooth weight function with most of its mass concentrated near $\alpha = 0$.

If

$$\mathbf{Y}(t) = \int \mathbf{a}(t - u)\mathbf{X}(u)\,du, \tag{3.9}$$

as in (2.13), then

$$\mathbf{d}_Y^{(T)}(\lambda) \doteq \mathbf{A}(\lambda)\,\mathbf{d}_X^{(T)}(\lambda) \tag{3.10}$$

with $\mathbf{A}(\lambda)$ given by (2.14). (The approximation (3.10) will be made precise in Lemma 5.1.)

A principal result concerning the finite Fourier transform of a stationary series is indicated in the following theorem.

THEOREM 3.1. *Let* $\mathbf{X}(t)$, $t \in R^p$, *satisfy Assumption* 2.1. *Let* $h(t)$ *satisfy Assumption* 3.1. *Then as* $T \to \infty$,

$$\begin{aligned}\operatorname{cum}\{d_{a_1}^{(T)}(\lambda^{(1)}), \ldots, d_{a_k}^{(T)}(\lambda^{(k)})\} \\ = T^p H_k\left(T\sum_1^k \lambda^{(j)}\right)(2\pi)^{p(k-1)} f_{a_1 \ldots a_k}(\lambda^{(1)}, \ldots, \lambda^{(k-1)}) + O(T^{p-1})\end{aligned} \tag{3.11}$$

for $a_1, \ldots, a_k = 1, \ldots, r$; $k = 1, 2, \ldots$. *The* $O(T^{p-1})$ *term is uniformly bounded in* $\lambda^{(1)}, \ldots, \lambda^{(k)}$.

The finite Fourier transform is seen to provide information on the various cumulant spectra. In particular we see from (3.11) that if $E\mathbf{X}(t) = 0$, then

$$E\mathbf{d}_X^{(T)}(\lambda)\overline{\mathbf{d}_X^{(T)}(\lambda)}^{\tau} = T^p H_2(0)(2\pi)^p \mathbf{f}_{XX}(\lambda) + O(T^{p-1}), \tag{3.12}$$

and so

$$\mathbf{I}_{XX}^{(T)}(\lambda) = (2\pi T)^{-p} H_2(0)^{-1} \mathbf{d}_X^{(T)}(\lambda)\overline{\mathbf{d}_X^{(T)}(\lambda)}^{\tau} \tag{3.13}$$

is a possible estimate of $\mathbf{f}_{XX}(\lambda)$ in this case. We call $\mathbf{I}_{XX}^{(T)}(\lambda)$ the *matrix of*

second-order periodograms. It will be investigated further in the next section.

Expression (3.11) indicates a weak statistical dependence of finite Fourier transforms evaluated at frequencies with $\Sigma\lambda^{(j)}$ not near 0. The statistical dependence is also affected by the rate at which $H_k(\lambda)$ falls off to 0 as $|\lambda|$ increases.

We may develop the following theorem from Theorem 3.1.

THEOREM 3.2. *Suppose the conditions of Theorem* 3.1 *are satisfied. Suppose* $E\mathbf{X}(t) = \mathbf{0}$ *and* $\mathbf{f}_{XX}(\lambda)$ *is nonsingular. Let* $\lambda^{(j)}(T)$, $\lambda \in R^p$, *with* $\lambda^{(j)}(T) \to \lambda$ *and* $T|\lambda^{(j)}(T)| \to \infty$, $T|\lambda^{(i)}(T) \pm \lambda^{(j)}(T)| \to \infty$, $1 \leqq i < j \leqq k$, *as* $T \to \infty$. *Then* $\mathbf{d}_X^{(T)}(\lambda^{(1)}(T)), \ldots, \mathbf{d}_X^{(T)}(\lambda^{(k)}(T))$ *are asymptotically independent.*

$$N_r^C(\mathbf{0}, (2\pi T)^p H_2(0)\mathbf{f}_{XX}(\lambda)) \tag{3.14}$$

variates.

The finite Fourier transforms are seen to satisfy a form of central limit theorem. This theorem will be of use in determining the asymptotic distributions of several statistics of interest. We note the reference Kawata (1966).

On occasion it may be helpful to have some idea of how irregular the behavior of $\mathbf{d}_X^{(T)}(\lambda)$ may be as a function of λ and T. In order to investigate this problem we set down the following assumption.

ASSUMPTION 3.2. *The series* $\mathbf{X}(t)$ *is uniformly bounded and satisfies Assumption* 3.1. *Also if* C_k *is given by*

$$C_k = \sup_{a_1 \ldots a_k} \int \cdots \int \sum |u_j| |c_{a_1 \ldots a_k}(u^{(1)}, \ldots, u^{(k-1)})| \, du^{(1)} \ldots du^{(k-1)}, \tag{3.15}$$

then

$$\sum_k C_k z^k / k! < \infty, \tag{3.16}$$

for $z \in R$ *in a neighborhood of* 0.

We may now state,

THEOREM 3.3. *Let* $X(t)$, *be real-valued, satisfy Assumption* 3.2 *and have mean* 0. *Let* $h(t)$ *satisfy Assumption* 3.1. *Let* Λ *be a finite nonnegative number. Then*

$$\overline{\lim_{T \to \infty}} \sup_{|\lambda| \leqq \Lambda} |d_X^{(T)}(\lambda)| / (T^p \log T)^{1/2} \leqq \left\{ 2(p^2 + p + 1)(2\pi)^p H_2(0) \sup_{|\lambda| \leqq \Lambda} f_{XX}(\lambda) \right\}^{1/2} \tag{3.17}$$

with probability 1.

This result may be compared instructively with the law of the iterated logarithm. In the present context this would correspond to the case $p = 1$ and a consideration of the Fourier transform at a single frequency, rather than a supremum over a set of frequencies. It would lead to a rate of growth of order $(T \log \log T)^{1/2}$ rather than $(T \log T)^{1/2}$. We note that the assumed boundedness of $X(t)$ leads to the elementary inequality $|d_X^{(T)}(\lambda)| T^p$ for some finite M.

In terms of the second-order periodogram, (3.17) has the form

$$\overline{\lim_{T\to\infty}} \sup_{|\lambda| \leqq \Lambda} I_{XX}^{(T)}(\lambda)/(\log T)^{1/2} \leqq \left\{2(p^2 + p + 1) \sup_{|\lambda| \leqq \Lambda} f_{XX}(\lambda)\right\}^{1/2} \tag{3.18}$$

with probability 1. Results related to (3.18) are given in Salem and Zygmund (1956), Whittle (1959) and Kahane (1968). Finer arguments than those presented here allow the reduction of the constant on the right-hand side of (3.18). We may conclude from (3.17) that

$$\sup_{|\lambda| \leqq \Lambda} T^{-p} |d_X^{(T)}(\lambda)| \to 0 \tag{3.19}$$

almost surely as $T \to \infty$. Wiener and Wintner (1941) investigate similar results. The strong law of large numbers for the series $X(t)$ is a particular case of (3.19).

Theorem 3.1 now leads us to a development of the Cramér representation, (2.7), for series $\mathbf{X}(t)$ satisfying Assumption 2.1. For such an $\mathbf{X}(t)$ take

$$\begin{aligned} h(u) &= 1, \qquad |u_j| \leqq 1, \\ &= 0, \qquad \text{otherwise}, \end{aligned} \tag{3.20}$$

and set

$$(2\pi)^p \mathbf{Z}(\lambda) = \mathbf{Z}^{(T)}(\lambda) = \int_0^{\lambda} \mathbf{d}_X^{(T)}(\alpha)\, d\alpha = \int_{-T}^{T} \prod_1^p \left[\frac{1 - \exp\{-i\lambda_j t_j\}}{i t_j}\right] \mathbf{X}(t)\, dt. \tag{3.21}$$

We then have,

THEOREM 3.4. *Let* $\mathbf{X}(t)$, $t \in R^p$, *satisfy Assumption* 2.1 *and let* $\mathbf{Z}^{(T)}(\lambda)$, $\lambda \in R^p$, *be given by* (3.21). *Then there exists* $\mathbf{Z}_X(\lambda)$ *such that* $\mathbf{Z}^{(T)}(\lambda) \to \mathbf{Z}_X(\lambda)$ *in mean of order* v *for any* $v > 0$ *and where*

$$\begin{aligned} &\operatorname{cum}\{Z_{a_1}(\lambda^{(1)}), \ldots, Z_{a_k}(\lambda^{(k)})\} \\ &= \int_0^{\lambda^{(1)}} \cdots \int_0^{\lambda^{(k)}} \delta\left(\sum_1^k \alpha^{(j)}\right) f_{a_1 \ldots a_k}(\alpha^{(1)}, \ldots, \alpha^{(k-1)})\, d\alpha^{(1)} \ldots d\alpha^{(k)}. \end{aligned} \tag{3.22}$$

Also

$$\mathbf{X}(t) = \int \exp\{i(t, \lambda)\}\, d\mathbf{Z}_X(\lambda). \tag{3.23}$$

with probability one.

The function $\mathbf{Z}^{(T)}(\lambda)$, $\lambda \in R^p$, is seen to be providing information on the spectral measures of the series $\mathbf{X}(t)$.

We conclude this section with alternate versions of Assumption 3.1 and Theorem 3.1.

ASSUMPTION 3.3. *$h(t), t \in R^p$, is measurable in t, bounded, zero for $|t| > 1$, $\int h(t)^2\, dt \neq 0$ and if*

$$H_1(\lambda) = \int h(t) \exp\{-i(\lambda, t)\}\, dt, \tag{3.24}$$

then

$$|H_1(\lambda)| < K(1 + |\lambda|)^{-\nu} \tag{3.25}$$

for some finite K where $\nu > p + 1$.

An example of such an $h(t)$ is provided by the second entry of Table 3.1 with $\alpha > p/2 + 1$. (3.25) will hold, for given ν, if $h(t)$ has sufficient continuous derivatives. It holds for all ν for the last entry of Table 3.1. We have

THEOREM 3.5. *Let $\mathbf{X}(t)$, $t \in R^p$, satisfy Assumption 2.1. Let $h(t)$ satisfy Assumption 3.3. Then as $T \to \infty$,*

$$\begin{aligned} &\operatorname{cum}\{d_{a_1}^{(T)}(\lambda^{(1)}), \ldots, d_{a_k}^{(T)}(\lambda^{(k)})\} \\ &\quad = T^p H_k\left(T \sum_1^k \lambda^{(j)}\right)(2\pi)^{p(k-1)} f_{a_1 \ldots a_k}(\lambda^{(1)}, \ldots, \lambda^{(k-1)}) \\ &\quad\quad + O\left[T^{p-1}\left(1 + T\left|\sum_1^k \lambda^{(j)}\right|\right)^{-\nu+1}\right] \end{aligned} \tag{3.26}$$

for $a_1, \ldots, a_k = 1, \ldots, r$; $k = 1, 2, \ldots$ with the error term uniform in the $\lambda^{(j)}$.

This version gives further information on the error term of (3.11). We see that it is of additionally lower order if the $\lambda^{(j)}$ have sum bounded away from 0. Also one can adjust the rate of fall-off by choice of ν.

If $h(t)$ is taken to be a quasi-analytic function (see Mandelbrojt (1935)), then one can obtain still more rapidly decreasing error terms.

4. Estimation of second-order spectra. We now turn to the problem of constructing an estimate of $\mathbf{f}_{XX}(0)$, the spectral density matrix of a series

$\mathbf{X}(t)$. In the previous section we defined the statistic,

$$\mathbf{I}_{XX}^{(T)}(\lambda) = (2\pi T)^{-p} H_2(0)^{-1} \mathbf{d}_X^{(T)}(\lambda) \overline{\mathbf{d}_X^{(T)}(\lambda)}^{\mathrm{T}}. \tag{4.1}$$

In connection with it we may prove

THEOREM 4.1. *Let* $\mathbf{X}(t)$, $t \in R^p$, *satisfy Assumption* 2.1 *and have mean* $\mathbf{0}$. *Let* $h(t)$ *satisfy Assumption* 3.1. *Then*

$$\begin{aligned} E\mathbf{I}_{XX}^{(T)}(\lambda) &= \left\{ \int \mathbf{f}_{XX}(\lambda - \alpha/T) |H_1(\alpha)|^2 \, d\alpha \right\} \left\{ \int |H_1(\alpha)|^2 \, d\alpha \right\}^{-1} \\ &= \mathbf{f}_{XX}(\lambda) + O(T^{-1}). \end{aligned} \tag{4.2}$$

From the first expression of (4.2) we see that the mean of $\mathbf{I}_{XX}^{(T)}(\lambda)$ is a weighted average of the spectral density matrix of interest with weight concentrated in a neighborhood of λ. $\mathbf{I}_{XX}^{(T)}(\lambda)$ therefore appears to be a plausible estimate of $\mathbf{f}_{XX}(\lambda)$ if T is sufficiently large. It is asymptotically unbiased. We note that the taper $h(t/T)$ employed affects the relative weighting of the values of $\mathbf{f}_{XX}(\cdot)$ appearing in (4.2). We are led to investigate further distributional aspects of $\mathbf{I}_{XX}^{(T)}(\lambda)$ and find

THEOREM 4.2. *Let* $\mathbf{X}(t)$, $t \in R^p$, *satisfy Assumption* 2.1 *and have mean* 0. *Let* $h(t)$ *satisfy Assumption* 3.1. *Then* $\mathbf{I}_{XX}^{(T)}(\lambda)$ *is asymptotically* $W_r^C(1, \mathbf{f}_{XX}(\lambda))$ *if* $\lambda \neq 0$ *and asymptotically* $W_r(1, \mathbf{f}_{XX}(0))$ *if* $\lambda = 0$.

We see that typically the statistic $\mathbf{I}_{XX}^{(T)}(\lambda)$ is not converging to $\mathbf{f}_{XX}(\lambda)$ as $T \to \infty$ and so it is not an appropriate estimate for most purposes. However under our assumptions $\mathbf{f}_{XX}(\alpha)$ is a smooth function of α and so it may be the case that an average of values of $\mathbf{I}_{XX}^{(T)}(\alpha)$ for α near λ will provide a reasonable estimate. Consider therefore

$$\begin{aligned} \mathbf{f}_{XX}^{(T)}(\lambda) &= k^{-1} \sum_{j=1}^{k} \mathbf{I}_{XX}^{(T)}(\lambda^{(j)}(T)) && \text{if } \lambda \neq 0, \\ &= (2k)^{-1} \sum_{j=1}^{k} [\mathbf{I}_{XX}^{(T)}(\lambda^{(j)}(T)) + \mathbf{I}_{XX}^{(T)}(-\lambda^{(j)}(T))] && \text{if } \lambda = 0. \end{aligned} \tag{4.3}$$

We may prove,

THEOREM 4.3. *Under the conditions of Theorem* 3.2, $\mathbf{f}_{XX}^{(T)}(\lambda)$ *is asymptotically* $k^{-1} W_r^C(k, \mathbf{f}_{XX}(\lambda))$ *if* $\lambda \neq 0$ *and asymptotically* $(2k)^{-1} W_r(2k, \mathbf{f}_{XX}(0))$ *if* $\lambda = 0$.

In the case $p = 1$, $\mathbf{f}_{XX}^{(T)}(\lambda)$ is seen to be asymptotically $\mathbf{f}_{XX}(\lambda)\chi_{2k}^2/(2k)$. (Wahba (1969) is also concerned with a complex Wishart approximation.)

The theorem suggests that as k is increased the values of $\mathbf{f}_{XX}^{(T)}(\lambda)$ should become more concentrated about $\mathbf{f}_{XX}(\lambda)$. However if k is held fixed the estimate (4.3) is not a consistent estimate of $\mathbf{f}_{XX}(\lambda)$ as $T \to \infty$. We are therefore led to consider a class of estimates for which, in effect, $k \to \infty$. These estimates involve a weight function satisfying

ASSUMPTION 4.1. *$W(\alpha)$, $\alpha \in R^p$, is real valued, even, absolutely integrable.*

$$\int W(\alpha)\, d\alpha = 1, \tag{4.4}$$

and if

$$w(u) = \int W(\alpha) \exp\{-i(u, \alpha)\}\, d\alpha, \tag{4.5}$$

then $w(u)$ is absolutely integrable.

Let $B_T > 0$, $T = 1, 2, \ldots$, be a sequence of constants. Define

$$W^{(T)}(\alpha) = B_T^{-p} W(B_T^{-1}\alpha). \tag{4.6}$$

In the case of $E\mathbf{X}(t) = 0$, we now take as our estimate of $\mathbf{f}_{XX}(\lambda)$,

$$\mathbf{f}_{XX}^{(T)}(\lambda) = \int W^{(T)}(\lambda - \alpha)\mathbf{I}_{XX}^{(T)}(\alpha)\, d\alpha. \tag{4.7}$$

If we define

$$\mathbf{m}_{XX}^{(T)}(u) = T^{-p} \int h([t + u]/T)\mathbf{X}(t + u)h(t/T)\mathbf{X}(t)^{\mathrm{t}}\, dt, \tag{4.8}$$

then (4.7) may be written

$$\mathbf{f}_{XX}^{(T)}(\lambda) = (2\pi)^{-p} \int w(B_T u)\mathbf{m}_{XX}^{(T)}(u) \exp\{-i(\lambda, u)\}\, du. \tag{4.9}$$

We may prove

THEOREM 4.4. *Let $\mathbf{X}(t)$, $t \in R^p$, satisfy Assumption* 2.1 *and have mean* 0. *Let $h(t)$ satisfy Assumption* 3.1. *Let $W(\alpha)$ satisfy Assumption* 4.1. *Let $\mathbf{f}_{XX}^{(T)}(\lambda)$ be given by* (4.8). *Then*

$$E\mathbf{f}_{XX}^{(T)}(\lambda) = \int W^{(T)}(\lambda - \alpha)\mathbf{f}_{XX}(\alpha)\, d\alpha + O(T^{-1}), \tag{4.10}$$

$$\begin{aligned}\operatorname{cov}\{f_{jk}^{(T)}(\lambda), f_{mn}^{(T)}(\mu)\} &= (2\pi/T)^p H_4(0)H_2(0)^{-2} \\ &\quad \cdot \left\{\int W^{(T)}(\lambda-\alpha)W^{(T)}(\mu-\alpha)f_{jm}(\alpha)f_{kn}(-\alpha)\,d\alpha\right. \\ &\quad \left. + \int W^{(T)}(\lambda-\alpha)W^{(T)}(\mu+\alpha)f_{jn}(\alpha)f_{km}(-\alpha)\,d\alpha\right\} \\ &\quad + O(B_T^{-p-1}T^{-p-1}),\end{aligned} \tag{4.11}$$

$j, k, m, n = 1, \ldots, r$. Also $\mathbf{f}_{XX}^{(T)}(\lambda^{(1)}), \ldots, \mathbf{f}_{XX}^{(T)}(\lambda^{(k)})$ are asymptotically jointly normal with covariance structure given by (4.11) *if $B_T T \to \infty$ as $T \to \infty$.*

We see, from (4.10), that the choice of $W(\alpha)$ affects the extent to which we are learning about $\mathbf{f}_{XX}(\lambda)$ from $\mathbf{f}_{XX}^{(T)}(\lambda)$. We could consider the entries $h(t)$ of Table 3.1 as possible choices for $W(t)$. The work of Elkins (1968), in a different context, suggests that the first entry of Table 3.1 might be a good choice. We have $0 < H_2(0)^2 \leqq H_4(0)$ and so we see, from (4.11), that the effect of tapering is to increase the variance. This increase in variance must be weighted against the improvement in resolution. The theorem has the following corollary.

COROLLARY. *Under the conditions of the theorem and if $B_T \to 0, B_T T \to \infty$ as $T \to \infty$,*

$$\lim_{T\to\infty} E\mathbf{f}_{XX}^{(T)}(\lambda) = \mathbf{f}_{XX}(\lambda), \tag{4.12}$$

$$\begin{aligned}&\lim_{T\to\infty} B_T^p T^p \operatorname{cov}\{f_{jk}^{(T)}(\lambda), f_{mn}^{(T)}(\mu)\} \\ &\quad = (2\pi)^p H_4(0)H_2(0)^{-2}\{\delta\{\lambda-\mu\}f_{jm}(\lambda)f_{kn}(-\lambda) \\ &\qquad + \delta\{\lambda+\mu\}f_{jn}(\lambda)f_{km}(-\lambda)\}\int W(\alpha)^2\,d\alpha\end{aligned} \tag{4.13}$$

for $j, k, m, n = 1, \ldots, r$.

We see that under the stated conditions $\mathbf{f}_{XX}^{(T)}(\lambda)$ is an asymptotically unbiased and consistent estimate of $\mathbf{f}_{XX}^{(T)}(\lambda)$. Grenander and Rosenblatt (1956) develop formulas for the mean and variance of spectral estimates of the form (4.7) in the case that $p = 2$ and that $W(\alpha) = 1, 0 \leqq \alpha_1, \alpha_2 \leqq 1$; $W(\alpha) = 0$ elsewhere. Another reference is Priestley (1964). Parzen (1957) and Rosenblatt (1959) consider the case $p = 1$.

On occasion we may wish a measure of the extent to which $\mathbf{f}_{XX}^{(T)}(\lambda)$ may deviate from $E\mathbf{f}_{XX}^{(T)}(\lambda)$ as a function of λ and T. We prove

THEOREM 4.5. *Let $\mathbf{X}(t)$ satisfy Assumption* 3.2 *and have mean* 0. *Let $h(t)$ satisfy Assumption* 3.1. *Let $W(\alpha)$ satisfy Assumption* 4.1 *and suppose $w(u)$*

has compact support. Let Λ *be a finite nonnegative number. Let* $\eta > 0$ *and* B_T *be such that* $\sum_T B_T^\eta < \infty$. *Let* $B_T T \to \infty$, *then*

$$\overline{\lim_{T\to\infty}} \sup_{|\lambda| \leqq \Lambda} |f_{jk}^{(T)}(\lambda) - Ef_{jk}^{(T)}(\lambda)|(B_T^p T^p/\log 1/B_T)^{1/2}$$

(4.14)

$$\leqq \left((2\pi)^p H_4(0)H_2(0)^{-2}4(p^2 + p + \eta)\int W(\alpha)^2\, d\alpha \sup_{|\lambda|\leqq\Lambda}\{f_{jj}(\lambda)f_{kk}(\lambda)\}\right)^{1/2}$$

with probability 1 *for* $j, k = 1, \ldots, r$.

A related reference to this last is Woodroofe and Van Ness (1967). In developing the estimate (4.7) we assumed that $E\mathbf{X}(t) = 0$. If $E\mathbf{X}(t) \neq 0$, then we would evaluate $\mathbf{X}^{(T)}$, the mean of the tapered values and base $\mathbf{f}_{XX}^{(T)}(\lambda)$ on $\mathbf{I}_{X-X^{(T)},X-X^{(T)}}^{(T)}(\lambda)$, rather than $\mathbf{I}_{XX}^{(T)}(\lambda)$. In the case that $\mathbf{X}(t)$ is isotropic we are led to modify the estimate to take advantage of the fact that $\mathbf{f}_{XX}(\lambda)$ depends only on $|\lambda|$. For example, we can consider the estimate

$$\begin{aligned}\mathbf{g}_{XX}^{(T)}(\lambda) &= \int \mathbf{f}_{XX}^{(T)}(|\lambda|\beta)\, d\Gamma(\beta) \\ &= \int_\alpha \mathbf{I}_{XX}^{(T)}(\alpha)\left\{\int_\beta W^{(T)}(|\lambda|\beta - \alpha)\, d\Gamma(\beta)\right\} d\alpha,\end{aligned} \tag{4.15}$$

where $\Gamma(\beta)$ denotes the invariant measure on the unit p-sphere. From (4.9) we see that (4.15) may be written

$$\mathbf{g}_{XX}^{(T)}(\lambda) = (2\pi)^{-\frac{p}{2}}\int w(B_T u)\mathbf{m}_{XX}^{(T)}(u)J_{p/2-1}(|\lambda||u|)\, du, \tag{4.16}$$

where $J_\nu(x)$ is the Bessel function and $\mathbf{m}_{XX}^{(T)}(u)$ is given by (4.8).

In connection with the estimate (4.15) we may prove

THEOREM 4.6. *Let* $\mathbf{X}(t)$, $t \in R^p$, $p > 1$, *be isotropic. Suppose the conditions of Theorem* 4.4 *are satisfied. Let* $\mathbf{g}_{XX}^{(T)}(\lambda)$ *be given by* (4.15). *If* $B_T \to 0$, $B_T T \to \infty$ *as* $T \to \infty$, *then*

$$\lim_{T\to\infty} E\mathbf{g}_{XX}^{(T)}(\lambda) = \mathbf{f}_{XX}(\lambda), \tag{4.17}$$

$$\lim_{T\to\infty} B_T^{\,p} T^p \operatorname{cov}\{g_{jk}^{(T)}(\lambda), g_{mn}^{(T)}(\mu)\}$$

$$= H_4(0)H_2(0)^{-2}\delta\{|\lambda| - |\mu|\}[f_{jm}(\lambda)f_{kn}(-\lambda) + f_{jn}(\lambda)f_{km}(-\lambda)] \tag{4.18}$$

$$\cdot |\lambda|^{-1} \cdot 2^{p-2}\pi^{-1}\Gamma\left(\frac{p}{2}\right)^2 \int |u|^{-p+1} w(u)^2\, du$$

Also $\mathbf{g}_{XX}^{(T)}(\lambda^{(1)}), \ldots, \mathbf{g}_{XX}^{(T)}(\lambda^{(k)})$ *are asymptotically jointly normal with covariance structure given by* (4.18).

As would be expected, the asymptotic variance of $\mathbf{g}_{XX}^{(T)}(\lambda)$ is of smaller order of magnitude than that of $\mathbf{f}_{XX}^{(T)}(\lambda)$. We also note how it falls off as $|\lambda|$ increases. A related reference is Tick (1960).

We conclude this chapter by noting that we can proceed as in Brillinger (1969a) to develop the asymptotic behaviour of the sample autocovariance function $\mathbf{m}_{XX}^{(T)}(u)$ and the sample spectral measure,

$$\int_0^{\lambda} \mathbf{I}_{XX}^{(T)}(\alpha)\, d\alpha. \tag{4.19}$$

5. Relations between a stochastic series and a fixed series. We next turn to a topic of some importance—the frequency analysis of linear space invariant relations between series. Let $\mathbf{Y}(t), \boldsymbol{\varepsilon}(t), t \in R^p$, denote s vector-valued spatial series. Suppose that $\boldsymbol{\varepsilon}(t)$ satisfies Assumption 2.1 and has mean $\mathbf{0}$. Let $\mathbf{X}(t)$, $t \in R^p$, denote an r vector-valued fixed function. We assume that the tapered values $h(t/T)\mathbf{Y}(t)$, $h(t/T)\mathbf{X}(t)$, $t \in R^p$, are available. Suppose that the series satisfy the following relation,

$$\mathbf{Y}(t) = \int \mathbf{a}(t-u)\mathbf{X}(u)\, du + \boldsymbol{\varepsilon}(t), \tag{5.1}$$

where the $s \times r$ matrix-valued $\mathbf{a}(t)$ satisfies (2.15). We are interested in constructing estimates of $\mathbf{a}(t)$, $\mathbf{A}(\lambda)$ given by (2.14) and $\mathbf{f}_{\varepsilon\varepsilon}(\lambda)$ the spectral density matrix of $\boldsymbol{\varepsilon}(t)$.

We set down the following assumption concerning the function $\mathbf{X}(t)$:

ASSUMPTION 5.1. $\mathbf{X}(t)$, $t \in R^p$, *is a function possessing the properties possessed with probability one by a spatial series satisfying Assumption* 3.2, *having mean* $\mathbf{0}$ *and nonsingular spectral density matrix.*

If we define

$$\mathbf{R}(t) = \int \mathbf{a}(t-u)\mathbf{X}(u)\, du, \tag{5.2}$$

then we have

LEMMA 5.1. *Suppose* $\mathbf{X}(t)$, $t \in R^p$, *satisfies Assumption* 5.1. *Suppose* $\mathbf{a}(t)$ *satisfies* (2.15). *Then there is a finite* K *such that*

$$|\mathbf{d}_R^{(T)}(\beta) - \mathbf{A}(\beta)\mathbf{d}_X^{(T)}(\beta)| < K(T^{p-2}\log T)^{1/2} \tag{5.3}$$

for all β *in a finite region.*

The relations (5.1) and (5.3) now give

$$\mathbf{d}_Y^{(T)}(\lambda) = \mathbf{A}(\lambda)\mathbf{d}_X^{(T)}(\lambda) + \mathbf{d}_\varepsilon^{(T)}(\lambda) + o(T^{p/2}). \tag{5.4}$$

Suppose $\lambda^{(j)}(T)$ is near $\lambda \in R^p$, $j = 1, \ldots, k$, then from (5.4)

$$\mathbf{d}_Y^{(T)}(\lambda^{(j)}(T)) \doteq \mathbf{A}(\lambda)\mathbf{d}_X^{(T)}(\lambda^{(j)}(T)) + \mathbf{d}_\varepsilon^{(T)}(\lambda^{(j)}(T)), \tag{5.5}$$

$j = 1, \ldots, k$. Following Theorem 3.2, the values $\mathbf{d}_\varepsilon^{(T)}(\lambda^{(j)}(T))$ are approximately independent, zero mean, variance $(2\pi T)^p H_2(0)\mathbf{f}_{XX}(\lambda)$ normal variates. The model (5.5) has therefore the approximate form of a standard regression model. We define

$$\mathbf{I}_{YX}^{(T)}(\lambda) = (2\pi T)^{-p}\mathbf{d}_Y^{(T)}(\lambda)\overline{\mathbf{d}_X^{(T)}(\lambda)}^{\mathrm{T}} H_2(0)^{-1} \tag{5.6}$$

with similar definitions for $\mathbf{I}_{YY}^{(T)}(\lambda)$, $\mathbf{I}_{XX}^{(T)}(\lambda)$. We next define

$$\begin{aligned} \mathbf{f}_{YX}^{(T)}(\lambda) &= k^{-1}\sum_{j=1}^{k} \mathbf{I}_{YX}^{(T)}(\lambda^{(j)}(T)) && \text{if } \lambda \neq 0, \\ &= (2k)^{-1}\sum_{j=1}^{k} [\mathbf{I}_{YX}^{(T)}(\lambda^{(j)}(T)) + \mathbf{I}_{YX}^{(T)}(-\lambda^{(j)}(T))] && \text{if } \lambda = 0 \end{aligned} \tag{5.7}$$

with similar definitions for $\mathbf{f}_{YY}^{(T)}(\lambda)$, $\mathbf{f}_{XX}^{(T)}(\lambda)$. We now estimate $\mathbf{A}(\lambda)$ by

$$\mathbf{A}^{(T)}(\lambda) = \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}, \tag{5.8}$$

and $\mathbf{f}_{\varepsilon\varepsilon}(\lambda)$ by

$$\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = [\mathbf{f}_{YY}^{(T)}(\lambda) - \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}\mathbf{f}_{XY}^{(T)}(\lambda)]k/(k-r). \tag{5.9}$$

THEOREM 5.1. *Let $\boldsymbol{\varepsilon}(t)$, $t \in R^p$, satisfy Assumption* 2.1 *and have mean* $\mathbf{0}$. *Let $\mathbf{X}(t)$, $t \in R^p$, be uniformly bounded. Let $\mathbf{Y}(t)$, $t \in R^p$, be given by* (5.1) *where $\mathbf{a}(t)$ satisfies* (2.13). *Let $\lambda^{(j)}(T) - \lambda = O(\log T)^{-1/2}$, $T\lambda^{(j)}(T) \to \infty$ and $T|\lambda^{(i)}(T) \pm \lambda^{(j)}(T)| \to \infty$, $1 \leqq i < j \leqq k$ as $T \to \infty$. Suppose $\mathbf{f}_{XX}^{(T)}(\lambda)$ is nonsingular for T sufficiently large. Then $\mathrm{vec}\{\mathbf{A}^{(T)}(\lambda) - \mathbf{A}(\lambda)\}$ is asymptotically*

$$N_{rs}^C(0, k^{-1}\mathbf{f}_{\varepsilon\varepsilon}(\lambda) \otimes \mathbf{f}_{XX}^{(T)}(\lambda)^{-1}) \quad \textit{if} \quad \lambda \neq 0$$

and asymptotically

$$N_{rs}(0, (2k)^{-1}\mathbf{f}_{\varepsilon\varepsilon}(0) \otimes \mathbf{f}_{XX}^{(T)}(0)^{-1}) \quad \textit{if} \quad \lambda = 0.$$

Also $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda)$ tends to

$$(k-r)^{-1}W_s^C(k-r, \mathbf{f}_{\varepsilon\varepsilon}(\lambda)) \quad \textit{if} \quad y \neq 0$$

and to

$$(2k-2r)^{-1}W_s^C(2k-2r, \mathbf{f}_{\varepsilon\varepsilon}(0)) \quad \textit{if} \quad \lambda = 0.$$

The limiting normal and Wishart distributions are independent.

We note that the tapering function $h(t)$ does not appear in this theorem except through the definition (5.6).

The estimates are not generally consistent under the assumptions of this theorem. However we may construct consistent estimates by the procedure of § 4. Specifically, take a weight function $W(\alpha)$, take as sequence B_T, form $W^{(T)}(\alpha)$ as in (4.6). We define

$$\mathbf{f}_{YX}^{(T)}(\lambda) = \int W^{(T)}(\lambda - \alpha)\mathbf{I}_{YX}^{(T)}(\alpha)\,d\alpha \tag{5.10}$$

with similar definitions for $\mathbf{f}_{YY}^{(T)}(\lambda)$, $\mathbf{f}_{XX}^{(T)}(\lambda)$. As estimates of $\mathbf{A}(\lambda)$, $\mathbf{f}_{\varepsilon\varepsilon}(\lambda)$ now take

$$\mathbf{A}^{(T)}(\lambda) = \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}, \tag{5.11}$$

$$\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = \mathbf{f}_{YY}^{(T)}(\lambda) - \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}\mathbf{f}_{XY}^{(T)}(\lambda). \tag{5.12}$$

We will require the following assumption,

ASSUMPTION 5.2. *$W(\alpha)$, $\alpha \in R^p$, is bounded, even, of compact support and*

$$\int W(\alpha)\,d\alpha = 1. \tag{5.13}$$

A new weight function is provided by

$$W^{(T)}(\alpha)^2\left\{\int W^{(T)}(\alpha)^2\,d\alpha\right\}^{-1}. \tag{5.14}$$

Let

$$\mathbf{h}_{XX}^{(T)}(\lambda) = \left\{\int W^{(T)}(\lambda - \alpha)^2\mathbf{I}_{XX}^{(T)}(\alpha)\,d\alpha\right\}\left\{\int W^{(T)}(\alpha)^2\,d\alpha\right\}^{-1}. \tag{5.15}$$

We may now prove

THEOREM 5.2. *Let $\varepsilon(t)$, $t \in R^p$, satisfy Assumption* 2.1 *and have mean* $\mathbf{0}$. *Let $\mathbf{X}(t)$, $t \in R^p$, satisfy Assumption* 5.1. *Let $\mathbf{Y}(t)$, $t \in R^p$, be given by* (5.1) *where $\mathbf{a}(t)$ satisfies* (2.15). *Let $h(t)$ satisfy Assumption* 3.1. *Let $W(\alpha)$ satisfy Assumption* 5.2. *Let $\mathbf{A}^{(T)}(\lambda)$, $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda)$ be given by* (5.11), (5.12). *Let $B_T \to 0$, $B_T T \to \infty$ as $T \to \infty$, then*

$$E\mathbf{A}^{(T)}(\lambda) = \left\{\int W^{(T)}(\lambda - \alpha)\mathbf{A}(\alpha)\mathbf{I}_{XX}^{(T)}(\alpha)\,d\alpha\right\}\left\{\int W^{(T)}(\lambda - \alpha)\mathbf{I}_{XX}^{(T)}(\alpha)\,d\alpha\right\}^{-1}$$

$$+ O(T^{-1}\log T) \tag{5.16}$$

$$= \mathbf{A}(\lambda) + O(B_T) + O(T^{-1}\log T),$$

$$E\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = \mathbf{f}_{\varepsilon\varepsilon}(\lambda) + O(B_T^{-p}T^{-p}\log B_T^{-1}) + O(B_T) + O(T^{-1}\log T), \tag{5.17}$$

$$\begin{aligned}\operatorname{cov}\{\operatorname{vec}\mathbf{A}^{(T)}(\lambda), \operatorname{vec}\mathbf{A}^{(T)}(\mu)\} &= \delta\{\lambda-\mu\}B_T^{-p}T^{-p}(2\pi)^p\int W(\alpha)^2\,d\alpha \\ &\quad\times \mathbf{f}_{\varepsilon\varepsilon}(\lambda)\otimes[\mathbf{f}_{XX}^{(T)}(\lambda)\mathbf{h}_{XX}^{(T)}(\lambda)^{-1}\mathbf{f}_{XX}^{(T)}(\lambda)] + o(B_T^{-p}T^{-p}),\end{aligned} \tag{5.18}$$

$$\operatorname{cov}\{\operatorname{vec}\mathbf{A}^{(T)}(\lambda), g_{jk}^{(T)}(\mu)\} = o(B_T^{-p}T^{-p}), \tag{5.19}$$

$$\begin{aligned}\operatorname{cov}\{g_{jk}^{(T)}(\lambda), g_{mn}^{(T)}(\mu)\} &= B_T^{-p}T^{-p}(2\pi)^pH_4(0)H_2(0)^{-2}\int W(\alpha)^2\,d\alpha \\ &\quad\times(\delta\{\lambda-\mu\}[\mathbf{f}_{\varepsilon\varepsilon}(\lambda)]_{jm}[\mathbf{f}_{\varepsilon\varepsilon}(-\lambda)]_{kn} + \delta\{\lambda+\mu\}[\mathbf{f}_{\varepsilon\varepsilon}(\lambda)]_{jn}[\mathbf{f}_{\varepsilon\varepsilon}(-\lambda)]_{km}) \\ &\quad + o(B_T^{-p}T^{-p})\end{aligned} \tag{5.20}$$

for $j, k, m, n = 1, \ldots, s$. Also the variates $\mathbf{A}^{(T)}(\lambda)$, $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\mu)$ are asymptotically jointly normal with the above covariance structure.

We see that the estimates $\mathbf{A}^{(T)}(\lambda)$, $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\mu)$ are asymptotically unbiased and consistent under the conditions of the theorem. They are asymptotically independent of each other.

Let us now turn to the problem of estimating $\mathbf{a}(t)$. We may invert the relation (2.14) to obtain

$$\mathbf{a}(u) = (2\pi)^{-p}\int \mathbf{A}(\lambda)\exp\{i(u,\lambda)\}\,d\lambda. \tag{5.21}$$

Let $\mathbf{A}^{(T)}(\lambda)$ be an estimate of the form just considered. Suppose $\mathbf{A}(\lambda)$ vanishes for $|\lambda_j| > \pi$, $j = 1, \ldots, p$. Let $Q_T = 2R_T + 1$ be a sequence of odd integers tending to ∞ with T. As an estimate of $\mathbf{a}(u)$ we consider

$$\mathbf{a}^{(T)}(u) = Q_T^{-p}\sum_q \mathbf{A}^{(T)}(\pi q/R_T)\exp\{i\pi(q,u)/R_T\} \tag{5.22}$$

with the summation extending over $|q_j| \leqq R_T$, $j = 1, \ldots, p$. We have

THEOREM 5.3. *Suppose the conditions of Theorem 5.2 are satisfied. Suppose $\mathbf{A}(\lambda) = 0$ for $|\lambda_j| > \pi$, $j = 1, \ldots, p$. If $B_T \to 0$, $B_T \leqq Q_T^{-1}$, $Q_T \to \infty$, $Q_T B_T T \to \infty$ as $T \to \infty$,*

$$\begin{aligned}E\mathbf{a}^{(T)}(u) &= Q_T^{-p}\sum_q \mathbf{A}(\pi q/R_T)\exp\{i\pi(q,u)/R_T\} + O(B_T) + O(T^{-1}\log T) \\ &= \mathbf{a}(u) + \sum_{j\neq 0}\mathbf{a}(u+jQ_T) + O(B_T) + O(T^{-1}\log T),\end{aligned} \tag{5.23}$$

$$\begin{aligned}\operatorname{cov}\{\operatorname{vec}\mathbf{a}^{(T)}(u), \operatorname{vec}\mathbf{a}^{(T)}(v)\} &= Q_T^{-p}B_T^{-p}T^{-p}(2\pi)^p H_4(0)H_2(0)^{-2}\int W(\alpha)^2\,d\alpha \\ &\times Q_T^{-p}\sum_q \exp\{i\pi(q, u-v)/R_T\}\mathbf{f}_{\varepsilon\varepsilon}(\pi q/R_T) \\ &\otimes \mathbf{f}_{XX}^{(T)}(\pi q/R_T)\mathbf{h}_{XX}^{(T)}(\pi q/R_T)^{-1}\mathbf{f}_{XX}(\pi q/R_T) + o(Q_T^{-p}B_T^{-p}T^{-p}).\end{aligned} \tag{5.24}$$

Also $\mathbf{a}^{(T)}(u^{(1)}), \ldots, \mathbf{a}^{(T)}(u^{(k)})$ *are asymptotically jointly normal with covariance structure given by* (5.24).

6. Relations between two stochastic series. Suppose that

$$\begin{bmatrix}\mathbf{X}(t)\\ \mathbf{Y}(t)\end{bmatrix}, \tag{6.1}$$

$t \in R^p$, is an $(r+s)$ vector-valued stationary series with $\mathbf{X}(t)$ r vector-valued and $\mathbf{Y}(t)$ s vector-valued. Suppose $\boldsymbol{\varepsilon}(t)$ is an s vector-valued series that is statistically independent of $\mathbf{X}(t)$ at all lags. We again consider the model

$$\mathbf{Y}(t) = \int \mathbf{a}(t-u)\mathbf{X}(u)\,du + \boldsymbol{\varepsilon}(t), \tag{6.2}$$

however we now take $\mathbf{X}(t)$ to be stochastic.

Let

$$\begin{bmatrix}\mathbf{f}_{XX}(\lambda) & \mathbf{f}_{XY}(\lambda)\\ \mathbf{f}_{YX}(\lambda) & \mathbf{f}_{YY}(\lambda)\end{bmatrix} \tag{6.3}$$

denote the spectral density matrix of the series (6.1) and $\mathbf{A}(\lambda)$ denote the transfer function of $\mathbf{a}(t)$. Then (6.2) gives

$$\mathbf{f}_{YX}(\lambda) = \mathbf{A}(\lambda)\mathbf{f}_{XX}(\lambda). \tag{6.4}$$

We are led to consider

$$\mathbf{A}^{(T)}(\lambda) = \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1} \tag{6.5}$$

as an estimate of $\mathbf{A}(\lambda)$ where

$$\begin{bmatrix}\mathbf{f}_{XX}^{(T)}(\lambda) & \mathbf{f}_{XY}^{(T)}(\lambda)\\ \mathbf{f}_{YX}^{(T)}(\lambda) & \mathbf{f}_{YY}^{(T)}(\lambda)\end{bmatrix} \tag{6.6}$$

is an estimate of (6.3). We also have from (6.2)

$$\begin{aligned}\mathbf{f}_{YY}(\lambda) &= \mathbf{A}(\lambda)\mathbf{f}_{XX}(\lambda)\overline{\mathbf{A}(\lambda)}^{\tau} + \mathbf{f}_{\varepsilon\varepsilon}(\lambda) \\ &= \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda) + \mathbf{f}_{\varepsilon\varepsilon}(\lambda),\end{aligned} \tag{6.7}$$

and are led to consider

$$\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = \mathbf{f}_{YY}^{(T)}(\lambda) - \mathbf{f}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}\mathbf{f}_{XY}^{(T)}(\lambda) \tag{6.8}$$

as an estimate of $\mathbf{f}_{\varepsilon\varepsilon}(\lambda)$. These estimates parallel those of the previous section. We may prove

THEOREM 6.1. *Let the series* (6.1) *satisfy Assumption* 2.1 *and have mean* $\mathbf{0}$. *Suppose* $\mathbf{f}_{XX}(\lambda)$ *is nonsingular. Let* $\mathbf{\varepsilon}(t)$, $t \in R^p$, *satisfy Assumption* 2.1, *have mean* $\mathbf{0}$ *and be statistically independent of* $\mathbf{X}(t)$ *at all lags. Let* $\mathbf{a}(t)$ *satisfy* (2.15). *Let* $h(t)$ *satisfy Assumption* 3.1. *Let* $W(\alpha)$ *satisfy Assumption* 4.1. *Let* (6.6) *be constructed in the manner of* (4.7). *Then if* $B_T \to 0$, $B_T T \to \infty$ *as* $T \to \infty$

$$\lim_{T\to\infty} \overrightarrow{\text{ave}}\, \mathbf{A}^{(T)}(\lambda) = \mathbf{A}(\lambda), \tag{6.9}$$

$$\lim_{T\to\infty} B_T^p T^p \overrightarrow{\text{cov}}\{\text{vec}\, \mathbf{A}^{(T)}(\lambda), \text{vec}\, \mathbf{A}^{(T)}(\mu)\}$$

$$= (2\pi)^p H_4(0) H_2(0)^{-2} \int W(\alpha)^2\, d\alpha \delta\{\lambda - u\} \mathbf{f}_{\varepsilon\varepsilon}(\lambda) \otimes \mathbf{f}_{XX}(\lambda)^{-1}, \tag{6.10}$$

$$\lim_{T\to\infty} \overrightarrow{\text{ave}}\, \mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) \pm \mathbf{f}_{\varepsilon\varepsilon}(\lambda), \tag{6.11}$$

$$\lim_{T\to\infty} B_T^p T^p \overrightarrow{\text{cov}}\{g_{jk}^{(T)}(\lambda), g_{mn}^{(T)}(\mu)\} = (2\pi)^p H_4(0) H_2(0)^{-2}$$

$$\cdot(\delta\{\lambda - \mu\}[\mathbf{f}_{\varepsilon\varepsilon}(\lambda)]_{jm}[\mathbf{f}_{\varepsilon\varepsilon}(-\lambda)]_{kn}$$

$$+ \delta\{\lambda + \mu\}[\mathbf{f}_{\varepsilon\varepsilon}(\lambda)]_{jn}[\mathbf{f}_{\varepsilon\varepsilon}(-\lambda)]_{km}) \int W(\alpha)^2\, d\alpha, \tag{6.12}$$

$$\lim_{T\to\infty} B_T^p T^p \overrightarrow{\text{cov}}\{\text{vec}\, \mathbf{A}^{(T)}(\lambda), g_{jk}^{(T)}(\mu)\} = 0 \tag{6.13}$$

for $j, k, m, n = 1, \ldots, s$. *Also the variates* $\mathbf{A}^{(T)}(\lambda)$, $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\mu)$ *are asymptotically jointly normal with the above covariance structure.*

(In this theorem $\overrightarrow{\text{ave}}$ and $\overrightarrow{\text{cov}}$ have technical definitions allowing the use of the delta method. See Brillinger and Tukey (1964).)

Suppose now that $\mathbf{f}_{YX}(\lambda)$ vanishes for $|\lambda_j| > \pi$, $j = 1, \ldots, p$. Let $Q_T = 2R_T + 1$ be a sequence of odd positive integers tending to ∞ with T. As an estimate of $\mathbf{a}(u)$ we consider

$$\mathbf{a}^{(T)}(u) = Q_T^{-p} \sum_q \mathbf{A}^{(T)}(\pi q/R_T) \exp\{i\pi(q, u)/R_T\} \tag{6.14}$$

with the summation extending over $|q_j| \leqq R_T$, $j = 1, \ldots, p$. We may prove

THEOREM 6.2. *Suppose the conditions of Theorem* 6.1 *are satisfied. Then if* $B_T \to 0$, $B_T \leqq Q_T^{-1}$, $Q_T \to \infty$, $Q_T B_T T \to \infty$ *as* $T \to \infty$,

$$\lim_{T\to\infty} \overrightarrow{\text{ave}}\, \mathbf{a}^{(T)}(u) = \mathbf{a}(u), \tag{6.15}$$

$$\lim_{T\to\infty} Q_T^p B_T^p T^p \overrightarrow{\text{cov}}\{\text{vec } \mathbf{a}^{(T)}(u), \text{vec } \mathbf{a}^{(T)}(v)\}$$

$$= (2\pi)^p H_4(0) H_2(0)^{-2} \int W(\alpha)^2\, d\alpha \int_{-\pi}^{\pi} \exp\{i(u-v)\lambda\}\mathbf{f}_{\varepsilon\varepsilon}(\lambda) \otimes \mathbf{f}_{XX}(\lambda)^{-1}\, d\lambda. \tag{6.16}$$

Also $\mathbf{a}^{(T)}(u^{(1)}), \ldots, \mathbf{a}^{(T)}(u^{(k)})$ *are asymptotically jointly normal with covariance structure given by* (6.16).

Hannan (1967) considers the model (6.2) in the discrete time case and with $\mathbf{a}(u) = 0$ for $|u|$ sufficiently large.

On occasion an alternate form of asymptotic distribution may be more relevant than that of Theorem 6.1. We suppose that the estimate (6.6) has the form of (4.3) rather than (4.7) and have

THEOREM 6.3. *Let the series* (6.1) *satisfy Assumption* 2.1 *and have mean* $\mathbf{0}$. *Let* $\boldsymbol{\varepsilon}(t)$, $t \in R^p$, *satisfy Assumption* 2.1, *have mean* $\mathbf{0}$ *and be statistically independent of* $\mathbf{X}(t)$ *at all lags. Let* $\mathbf{a}(t)$ *satisfy* (2.15) *and* $h(t)$ *satisfy Assumption* 3.1. *Let* (6.6) *be constructed in the manner of* (4.3) *with* $\lambda^{(j)}(T) \to \lambda^{(j)}$, $T|\lambda^{(j)}(T)| \to \infty$, $T|\lambda^{(i)}(T) \pm \lambda^{(j)}(T)| \to \infty$, $1 \leqq i < j \leqq k$, *as* $T \to \infty$. *Then* (6.6) *is asymptotically,*

$$k^{-1} W_r^{\mathbb{C}}\left(k, \begin{bmatrix} \mathbf{f}_{XX}(\lambda) & \mathbf{f}_{XY}(\lambda) \\ \mathbf{f}_{YX}(\lambda) & \mathbf{f}_{YY}(\lambda) \end{bmatrix}\right), \tag{6.17}$$

if $\lambda \neq 0$, *and asymptotically*

$$(2k)^{-1} W_r\left(2k, \begin{bmatrix} \mathbf{f}_{XX}(0) & \mathbf{f}_{XY}(0) \\ \mathbf{f}_{YX}(0) & \mathbf{f}_{YY}(0) \end{bmatrix}\right), \tag{6.18}$$

if $\lambda = 0$. *In consequence the limiting distribution of* $\mathbf{A}^{(T)}(\lambda)$, *given by* (6.5), *and* $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda)$, *given by* (5.9) *will be that of* $\mathbf{W}_{YX}\mathbf{W}_{XX}^{-1}$,

$$(\mathbf{W}_{YY} - \mathbf{W}_{YX}\mathbf{W}_{XX}^{-1}\mathbf{W}_{XY})k/(k-r)$$

respectively where

$$\begin{bmatrix} \mathbf{W}_{XX} & \mathbf{W}_{XY} \\ \mathbf{W}_{YX} & \mathbf{W}_{YY} \end{bmatrix} \tag{6.19}$$

has the distribution (6.17) *if* $\lambda \neq 0$ *and* (6.18) *if* $\lambda = 0$.

For example, the asymptotic distribution of $\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda)$ will be

$$(k-r)^{-1}W_s^C(k-r, \mathbf{f}_{\varepsilon\varepsilon}(\lambda)), \tag{6.20}$$

if $\lambda \neq 0$.

7. Canonical analyses. Once again consider the $(r+s)$ vector-valued series (6.1) and suppose that it has mean 0 and spectral density matrix (6.3). Suppose that $q \leqq r$ channels are available for the transmission of filtered values of the $\mathbf{X}(t)$ series and that on receipt of the transmitted values one wishes to form a series near $\mathbf{Y}(t)$. Specifically suppose that one transmits the q vector-valued series

$$\zeta(t) = \int \mathbf{b}(t-u)\mathbf{X}(u)\,du \tag{7.1}$$

for some $q \times r$ matrix-valued $\mathbf{b}(t)$ and on receipt of $\zeta(t)$ one forms

$$\begin{aligned}\hat{\mathbf{Y}}(t) &= \int \mathbf{d}(t-u)\zeta(u)\,du \\ &= \int \mathbf{a}(t-u)\mathbf{X}(u)\,du\end{aligned} \tag{7.2}$$

for $\mathbf{d}(t)$ an $s \times q$ matrix-valued function and with $\mathbf{a}(t)$ denoting the convolution of $\mathbf{d}(t)$ with $\mathbf{b}(t)$. The question arises, how should one choose $\mathbf{b}(t)$, $\mathbf{d}(t)$, or equivalently $\mathbf{a}(t)$, so that $\hat{\mathbf{Y}}(t)$ is near $\mathbf{Y}(t)$? Let

$$\begin{aligned}\mathbf{A}(\lambda) &= \int \mathbf{a}(u)\exp\{-i(\lambda, u)\}\,du \\ &= \mathbf{D}(\lambda)\mathbf{B}(\lambda),\end{aligned} \tag{7.3}$$

$\lambda \in R^p$, where $\mathbf{D}(\lambda)$, $\mathbf{B}(\lambda)$ are the Fourier transforms of $\mathbf{d}(t)$, $\mathbf{b}(t)$ respectively. We may prove

THEOREM 7.1. *Let* (6.1) *satisfy Assumption* 2.1, *have mean* $\mathbf{0}$ *and spectral density matrix* (6.3). *Suppose* $\mathbf{f}_{XX}(\lambda)$ *is nonsingular for all* λ. *Then*

$$E\{\mathbf{Y}(t) - \hat{\mathbf{Y}}(t)\}^{\tau}\{\mathbf{Y}(t) - \hat{\mathbf{Y}}(t)\} \tag{7.4}$$

is minimized for $\mathbf{A}(\lambda)$ *of the form* (7.3) *by*

$$\mathbf{A}(\lambda) = \sum_{j=1}^{q} \mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^{\tau}\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}. \tag{7.5}$$

Here $\mathbf{V}_j(\lambda)$ *is the jth latent vector of* $\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)$. *If* $\mu_j(\lambda)$ *is the associated latent root, then the minimum achieved in* (7.4) *is*

$$\int \operatorname{tr}\{\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)\}\, d\lambda + \int \left\{ \sum_{j>q} \mu_j(\lambda) \right\} d\lambda. \tag{7.6}$$

The spectral density matrix of the residual series $\mathbf{Y}(t) - \hat{\mathbf{Y}}(t)$ *for the optimum* $\hat{\mathbf{Y}}(t)$ *is*

$$\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda) + \sum_{j>q} \mu_j(\lambda)\mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^{\tau}. \tag{7.7}$$

The theorem has several important corollaries.

COROLLARY 7.1. *Under the conditions of the theorem and if* $q = r$,

$$\mathbf{A}(\lambda) = \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}. \tag{7.8}$$

The spectral density matrix of the residual series is

$$\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda). \tag{7.9}$$

The expressions (7.7) and (7.8) are the same as expressions (6.4) and (6.7) and so we have two methods of developing these parameters.

COROLLARY 7.2. *Under the conditions of the theorem and if* $s = r$, $\mathbf{Y}(t) = \mathbf{X}(t)$ *for all t, one has*

$$\mathbf{A}(\lambda) = \sum_{j=1}^{q} \mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^{\tau}. \tag{7.10}$$

Here $\mathbf{V}_j(\lambda)$ *is the jth latent vector of* $\mathbf{f}_{XX}(\lambda)$. *The spectral density matrix of the residual series is*

$$\sum_{j>q} \mu_j(\lambda)\mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^{\tau}. \tag{7.11}$$

We are here interested in predicting the series $\mathbf{X}(t)$ from a linearly degenerate version of itself. This corollary provides a spatial series generalization of the classical multivariate technique of principal component analysis.

An invariant procedure that is a generalization of the classical technique of canonical correlation analysis is obtained if one takes, in the context of the theorem,

$$\mathbf{A}(\lambda) = \mathbf{f}_{YY}(\lambda)^{-1/2}\left\{ \sum_{j=1}^{q} \mathbf{U}_j(\lambda)\overline{\mathbf{U}_j(\lambda)}^{\tau} \right\}\mathbf{f}_{YY}(\lambda)^{1/2}\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}, \tag{7.12}$$

where $\mathbf{U}_j(\lambda)$ is the jth latent vector of

$$\mathbf{f}_{YY}(\lambda)^{-1/2}\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)\mathbf{f}_{YY}(\lambda)^{-1/2}. \tag{7.13}$$

If stretches of the series $\mathbf{X}(t)$, $\mathbf{Y}(t)$ are available then one would base an estimate of $\mathbf{A}(\lambda)$ on the matrix (6.6) rather than (6.3). Various statistical properties of the principal component analysis in the case $p = 1$, t discrete, are presented in Brillinger (1969b). The principal alternation to those results for the case of spatial series is that under the assumptions of Theorem 6.1 one multiplies the asymptotic covariances presented in that reference by $B_T^{-p}T^{-p}H_4(0)H_2(0)^{-2}$ rather than $B_T^{-1}T^{-1}$. ($H(\alpha)$ there corresponds to $W(\alpha)$ here.) Under the assumptions of Theorem 6.3 the asymptotic distributions of Brillinger (1969b) remain unaltered.

8. Higher order cumulant spectra. We next turn to the construction of estimates of the cumulant spectrum,

$$f_{a_1\ldots a_k}(\lambda^{(1)},\ldots,\lambda^{(k-1)}), \tag{8.1}$$

defined by (2.5). We will assume that the tapering function $h(t)$ satisfies Assumption 3.2 for all $\nu > p + 1$, although this requirement may be relaxed to some extent.

An examination of expression (3.26) suggests the consideration of

$$I^{(T)}_{a_1\ldots a_k}(\lambda^{(1)},\ldots,\lambda^{(k)}) = T^{-p}H_k(0)^{-1}(2\pi)^{-p(k-1)}\prod_1^k d^{(T)}_{a_j}(\lambda^{(j)}) \tag{8.2}$$

as an estimate of (8.1) in the case that $\sum_1^k \lambda^{(j)} = 0$, but no proper subset of the $\lambda^{(j)}$ have sum 0. This statistic is seen to be a natural generalization of the second-order periodogram considered in § 4. Unfortunately in the case $k > 2$, (8.2) suffers from the same nonconvergence properties as in the case $k = 2$ (see Brillinger and Rosenblatt (1967)) and so it must be modified somewhat to produce a useful estimate. We set down

ASSUMPTION 8.1. *$W(\alpha^{(1)},\ldots,\alpha^{(k)})$ vanishes unless $\sum_1^k \alpha^{(j)} = 0$. It is bounded, has uniformly bounded first derivatives and vanishes for $\sum_1^k |\alpha^{(j)}|$ sufficiently large. Also*

$$\int\cdots\int W(\alpha^{(1)},\ldots,\alpha^{(k)})\delta\left(\sum_1^k \alpha^{(j)}\right)d\alpha^{(1)}\ldots d\alpha^{(k)} = 1. \tag{8.3}$$

Given the sequence of nonnegative values B_T, $T = 1, 2, \ldots$, we set

$$W^{(T)}(\alpha^{(1)},\ldots,\alpha^{(k)}) = B_T^{-p(k-1)}W(B_T^{-1}\alpha^{(1)},\ldots,B_T^{-1}\alpha^{(k)}). \tag{8.4}$$

We also define $\Psi(\alpha^{(1)}, \ldots, \alpha^{(k)}) = 1$ if $\sum_1^k \alpha^{(j)} = 0$, but no subset of the $\alpha^{(j)}$ has sum 0 and $\Psi = 0$ otherwise.

As our estimate of (8.1) we take

$$f^{(T)}_{a_1 \ldots a_k}(\lambda^{(1)}, \ldots, \lambda^{(k-1)})$$

$$= \Delta_T^{p(k-1)} \sum_{s^{(1)}} \cdots \sum_{s^{(k)}} W^{(T)}(\lambda^{(1)} - s^{(1)}\Delta_T, \ldots, \lambda^{(k)} - s^{(k)}\Delta_T) \tag{8.5}$$

$$\times \Psi(s^{(1)}, \ldots, s^{(k)}) I^{(T)}_{a_1 \ldots a_k}(s^{(1)}\Delta_T, \ldots, s^{(k)}\Delta_T)$$

where $\Delta_T > 0$ and the $s^{(j)}$ are integers.

We see that (8.5) has the form of a weighted average of the function

$$I^{(T)}_{a_1 \ldots a_k}(\alpha^{(1)}, \ldots, \alpha^{(k)}) \tag{8.6}$$

on the hyperplane $\sum_1^k \alpha^{(j)} = 0$, but avoiding points such that some subset of the $\alpha^{(j)}$ has sum 0. The averaging is concentrated in the neighborhood of spatial frequencies of interest. The parameter Δ_T affects how close one comes to $\alpha^{(j)}$ with a subset summing to 0.

In the case $k = 2$ expression (8.5) is an average of the second-order periodogram avoiding wave numbers in the neighborhood of 0. This estimate has the advantage, over the estimate of § 4, of not requiring that $E\mathbf{X}(t) = 0$.

In connection with the estimate (8.5) we may prove

THEOREM 8.1. *Let* $\mathbf{X}(t)$, $t \in R^p$, *satisfy Assumption* 2.1. *Let* $h(t)$ *satisfy Assumption* 3.3 *for all* $v > p + 1$. *Then if* $\Delta_T B_T^{-1} \to 0$, $T\Delta_T \to \infty$ *as* $T \to \infty$.

$$= \int \cdots \int W^{(T)}(\lambda^{(1)} - \alpha^{(1)}, \ldots, \lambda^{(k)} - \alpha^{(k)}) \delta\left(\sum_1^k \alpha^{(j)}\right) \tag{8.7}$$

$$\times f_{a_1 \ldots a_k}(\alpha^{(1)}, \ldots, \alpha^{(k-1)})\, d\alpha^{(1)} \ldots d\alpha^{(k)}$$

$$+ O(T^{-1}) + O(T^{-v}\Delta_T^{-v}) + O(\Delta_T B_T^{-1})$$

for all v.

We see that the estimate (8.5) is asymptotically unbiased under the conditions of the theorem.

Turning to second-order moments, we have

THEOREM 8.2. *Suppose the conditions of Theorem* 8.1 *are satisfied and*

in addition $T^{k-2}(\Delta_T B_T^{-1})^{k-1} \to 0$ *as* $T \to \infty$, *then*

$$\lim_{T\to\infty} T^{-p(k-2)}\Delta_T^{-p(k-1)}B_T^{(k-1)} \times \operatorname{cov}\{f^{(T)}_{a_1\ldots a_k}(\lambda^{(1)},\ldots,\lambda^{(k-1)}), f^{(T)}_{b_1\ldots b_k}(\mu^{(1)},\ldots,\mu^{(k-1)})\}$$

$$= H_k(0)^{-2}H_2(0)^k(2\pi)^{-p(k-2)}\sum_P \delta\{\lambda^{(1)}-\mu^{(P1)}\}\ldots\delta\{\lambda^{(k)}-\mu^{(Pk)}\} \times \prod_1^k [f_{a_j b_{P_j}}(\lambda^{(j)})]\int\ldots\int W(\alpha^{(1)},\ldots,\alpha^{(k)})W(\alpha^{(P1)},\ldots,\alpha^{(Pk)}) \times \delta\left(\sum_1^k \alpha^{(j)}\right)d\alpha^{(1)}\ldots d\alpha^{(k)}. \tag{8.8}$$

We see that the proposed estimate is consistent under the assumptions of this theorem. Continuing we have

THEOREM 8.3. *Under the conditions of Theorem* 8.2, *the estimate* (8.5) *is asymptotically Gaussian with covariance structure given by* (8.8).

One can also prove that collections of these spectral estimates are asymptotically normal. In particular it turns out that estimates of different orders are asymptotically independent.

We have seen that second-order spectra are of use in investigating linear systems. Higher-order cumulant spectra are of use in investigating polynomial systems. For example, suppose

$$Y(t) = \int\ldots\int a(t-u^{(1)},\ldots,t-u^{(k)})X(u^{(1)})\ldots X(u^{(k)})\,du^{(1)}\ldots du^{(k)} + \varepsilon(t) \tag{8.9}$$

with $Y(t)$, $X(t)$, $\varepsilon(t)$, $t \in R^P$, real-valued, $\varepsilon(t)$ independent of $X(t)$ at all lags and $X(t)$ Gaussian. Suppose $a(u^{(1)},\ldots,u^{(k)})$ is symmetric in $u^{(1)},\ldots,u^{(k)}$. Let

$$A(\lambda^{(1)},\ldots,\lambda^{(k)}) \int\ldots\int a(u^{(1)},\ldots,u^{(k)})\exp\left\{-i\sum_1^k(\lambda^{(j)},u^{(j)})\right\}du^{(1)}\ldots du^{(k)}, \tag{8.10}$$

$\lambda^{(j)} \in R^P$, denote the transfer function of the system. Then, see Brillinger (1970),

$$A(\lambda^{(1)},\ldots,\lambda^{(k)}) = f_{X\ldots XY}(-\lambda^{(1)},\ldots,-\lambda^{(k)})/[k!f_{XX}(\lambda^{(1)})\ldots f_{XX}(\lambda^{(k)})], \tag{8.11}$$

and so the function may be estimated from an estimate of a cumulant spectrum of order $k+1$ and an estimate of a power spectrum.

The model (8.9) may be generalized to include terms of different orders k. See Brillinger (1970). It turns out that if one carries out the expansion in

terms of Itô multiple Wiener integrals, then the transfer functions of various orders are given by expressions of the form (8.11).

9. Proofs.

PROOF OF LEMMA 2.1. We have

$$\mathscr{A}[T^u\mathbf{e}](t) = e^{i(\lambda,u)}\mathscr{A}[\mathbf{e}](t)$$
$$= \mathscr{A}[\mathbf{e}](t+u),$$

using the properties of linearity and space invariance. Setting $t = 0$ it follows that

$$\mathscr{A}[\mathbf{e}](u) = e^{i(\lambda,u)}\mathscr{A}[\mathbf{e}](0),$$

and we have (2.11) with $\mathbf{A}(\lambda) = \mathscr{A}[\mathbf{e}](0)$.

PROOF OF THEOREM 3.1. The cumulant in question is given by

$$\int\cdots\int h(t^{(1)}/T)\ldots h(t^{(k)}/T)c_{a_1\ldots a_k}(t^{(1)}-t^{(k)},\ldots,t^{(k-1)}-t^{(k)})$$
$$\cdot\exp\left\{-i\sum_1^k(\lambda^{(j)},t^{(j)})\right\}dt^{(1)}\ldots dt^{(k)}.$$

If we set $u^{(j)} = t^{(j)} - t^{(k)}$, $j = 1,\ldots,k-1$ and $t = t^{(k)}$, then this may be written

$$\int\cdots\int c_{a_1\ldots a_k}(u^{(1)},\ldots,u^{(k-1)})\exp\left\{-i\sum_1^{k-1}(\lambda^{(j)},u^{(j)})\right\}$$
$$\cdot\left[\int h(t/T+u^{(1)}/T)\ldots h(t/T+u^{(k-1)}/T)h(t/T)\right.$$
$$\left.\cdot\exp\left\{-i\sum_i^k(\lambda^{(j)},t)\right\}dt\right]du^{(1)}\ldots du^{(k-1)}.$$

It follows from Assumption 3.1 that

$$\int|h(t/T+u^{(1)}/T)\ldots h(t/T+u^{(k-1)}/T)h(t/T)-h(t/T)^k|\,dt$$
$$<C^{k-1}T^{p-1}\sum_1^{k-1}|u^{(j)}|,$$

for some finite C. Also

$$\int h(t/T)^k\exp\left\{-i\sum_1^k(\lambda^{(j)},t)\right\}dt = T^pH_k\left(T\sum_1^k\lambda^{(j)}\right).$$

Expression (3.11) now follows from the above expressions and (2.5).

PROOF OF THEOREM 3.2. We prove this theorem by demonstrating that the cumulants of the Fourier transforms when standardized appropriately

tend to the cumulants of the normal distribution (3.14). We consider only the case $\lambda \neq 0$. The other case follows in the same manner. We begin by noting

$$ET^{-p/2}\mathbf{d}_X^{(T)}(\lambda^{(j)}(T)) = \mathbf{0},$$

and so the first cumulant has the required value. Next from Theorem 3.1

$$ET^{-p}\mathbf{d}_X^{(T)}(\lambda^{(j)}(T))\overline{\mathbf{d}_X^{(T)}(\lambda^{(j)}(T))}^{\tau}$$
$$= H_2(0)(2\pi)^p \mathbf{f}_{XX}(\lambda^{(j)}(T)) + O(T^{-1}) \to H_2(0)(2\pi)^p \mathbf{f}_{XX}(\lambda),$$

as $T \to \infty$. Next

$$ET^{-p}\mathbf{d}_X^{(T)}(\lambda^{(j)}(T))\mathbf{d}_X^{(T)}(\lambda^{(j)}(T))^{\tau}$$
$$= H_2(2T\lambda^{(j)}(T))(2\pi)^p \mathbf{f}_{XX}(\lambda^{(j)}(T)) + O(T^{-1}) \to 0,$$

as $T \to \infty$ by the Riemann-Lebesgue theorem. Finally if $i \neq j$, then

$$ET^{-p}\mathbf{d}_X^{(T)}(\lambda^{(i)}(T))\mathbf{d}_X^{(T)}(\pm\lambda^{(j)}(T))^{\tau}$$

is also seen to tend to 0 by the Riemann-Lebesgue theorem. These limiting second-order cumulants are those required in Theorem 3.2.

Again using Theorem 3.1 we see that

$$T^{-p/2}\operatorname{cum}\{d_{a_1}^{(T)}(\mu^{(1)}(T)), \ldots, d_{a_k}^{(T)}(\mu^{(k)}(T))\}$$
$$= T^{-pk/2}T^p H_k\left(T\sum_1^k \mu^{(j)}(T)\right) f_{a_1 \ldots a_k}(\mu^{(1)}(T), \ldots, \mu^{(k-1)}(T))(2\pi)^{p(k-1)}$$
$$+ O(T^{p-1-pk/2})$$

where $\mu^{(j)}(T)$ is one of $\pm\lambda^{(1)}(T), \ldots, \pm\lambda^{(k)}(T)$. Because $H_k(\alpha)$ is bounded this standardized cumulant tends to 0 if $k > 2$. The proof of the theorem is now complete.

PROOF OF THEOREM 3.3. We proceed via a sequence of lemmas. Let

$$\gamma = (2\pi T)^p H_2(0)/2,$$

where for convenience we suppress the dependence of γ on T.

LEMMA 9.1. *For given λ, ε and α sufficiently small*

$$E\exp\{\alpha \operatorname{Re} d_X^{(T)}(\lambda)/(f_{XX}(\lambda))^{1/2}\} \leqq \exp\{\alpha^2\gamma(1+\varepsilon)/2\}.$$

PROOF. From Theorem 3.1 and its proof we have

$$|\operatorname{cum}\{\operatorname{Re} d_X^{(T)}(\lambda), \ldots, \operatorname{Re} d_X^{(T)}(\lambda)\}| \leqq LT^p C_k + T^{p-1}D^k$$

for finite L, D. Therefore

$$|\log E \exp\{\alpha \operatorname{Re} d_X^{(T)}(\lambda)/(f_{XX}(\lambda))^{1/2}\} - \alpha^2\gamma/2|$$
$$\leqq T^{p-1}D^2|\alpha|^2/2 + \sum_3^\infty (LT^pC_k + T^{p-1}D^k)|\alpha|^k/k!.$$

We see that this expression is $O(T^p)|\alpha|^2$ for $|\alpha|$ sufficiently small in view of (3.16).

COROLLARY 9.1. *Under the conditions of the lemma*

$$E \exp\{\alpha|\operatorname{Re} d_X^{(T)}(\lambda)|/(f_{XX}(\lambda))^{1/2}\} \leqq 2 \exp\{\alpha^2\gamma(1 + \varepsilon)/2\}.$$

LEMMA 9.2. *Because $X(t)$ is bounded,*

$$|\operatorname{Re} d_X^{(T)}(\lambda) - \operatorname{Re} d_X^{(T)}(\mu)| \leqq KT^{p+1}|\lambda - \mu|,$$

for some finite K.

PROOF.

$$|\operatorname{Re} d_X^{(T)}(\lambda) - \operatorname{Re} d_X^{(T)}(\mu)| = \left|\int h(t/T)X(t)[\exp\{-i(\lambda, t)\} - \exp\{-i(\mu, t)\}]\, dt\right|$$
$$\leqq M \int |h(t/T)|\,|\exp\{i(\mu - \lambda, t)\} - 1|\, dt,$$

giving the stated result.

LEMMA 9.3. *Under the conditions of the theorem there are finite M, L such that*

$$E \exp\left\{\alpha \sup_{|\lambda| \leqq \Lambda} |\operatorname{Re} d_X^{(T)}(\lambda)|/(f_{XX}(\lambda))^{1/2}\right\}$$
$$\leqq M \exp\{LT^{p+1}/R + p \log R + \alpha^2\gamma(1 + \varepsilon)/2\}$$

for $\alpha > 0$ sufficiently small and given $R > 0$.

PROOF. Consider points $\lambda(r)$ of the form $\Lambda r/R$ where r is a lattice point in the hypercube $|r_j| \leqq R$, $j = 1, \ldots, p$. Using Lemma 9.2 we see

$$E \exp\left\{\alpha \sup_{|\lambda| \leqq \Lambda} |\operatorname{Re} d_X^{(T)}(\lambda)|/(f_{XX}(\lambda))^{1/2}\right\}$$
$$\leqq E \exp\left\{\alpha \sup_r |\operatorname{Re} d_X^{(T)}(\lambda(r))| + LT^{p+1}/R\right\}$$
$$\leqq \sum_r 2 \exp\{\alpha^2\gamma(1 + \varepsilon)/2 + LT^{p+1}/R\},$$

giving the indicated result as the summation runs over $O(R^p) = O(\exp\{p \log R\})$ points.

LEMMA 9.4. *Given $\varepsilon, \delta > 0$, let $a^2 = 2(p^2 + p + 1 + \delta)[\log T]\gamma(1 + \varepsilon)$.*

Under the conditions of the theorem

$$\text{Prob}\left[\sup_{|\lambda| \leqq \Lambda} |\text{Re } d_X^{(T)}(\lambda)|/(f_{XX}(\lambda))^{1/2} \geqq a\right] \leqq NT^{-1-\delta}$$

for some finite N.

PROOF. From Lemma 9.3 the probability in question is, for $\alpha > 0$,

$$\leqq \exp\{-\alpha a\} M \exp\{LT^{p+1}/R + p \log R + \alpha^2 \gamma(1 + \varepsilon)/2\}$$

for all R and α sufficiently small. This expression is minimized if one sets $R = LT^{p+1}/p$, $\alpha = a/(\gamma(1 + \varepsilon))$. With these choices we see that the probability is $\leqq N \exp\{-(1 + \delta) \log T\}$ giving the result.

PROOF OF THEOREM. The theorem now follows from the convergent series criterion, the fact that ε and δ are arbitrary, and the corresponding result for $\text{Im } d_X^{(T)}(\lambda)$.

PROOF OF THEOREM 3.4. Let $0 < \tau \leqq T$. Arguing as in the proof of Theorem 3.1 one sees that all cumulants of $\mathbf{Z}^{(T)}(\lambda) - \mathbf{Z}^{(\tau)}(\lambda)$ tend to 0 as $T, \tau \to \infty$. From this it follows that

$$E[Z_j^{(T)}(\lambda) - Z_j^{(\tau)}(\lambda)]^{2k} \to 0$$

as $T, \tau \to \infty$ for $j = 1, \ldots, r$. Therefore there exists a random variable $Z_j(\lambda)$ to which $Z_j^{(T)}(\lambda)$ converges in mean of order $2k$ (and hence of any lower order ν). The cumulants of this $\mathbf{Z}_X(\lambda)$ are given by

$$\lim_{T \to \infty} \text{cum}\{Z_{a_1}^{(T)}(\lambda^{(1)}), \ldots, Z_{a_k}^{(T)}(\lambda^{(k)})\},$$

which may be seen to be (3.22). Finally one can see that all cumulants of $\mathbf{X}(t) - \int \exp\{i(t, \lambda)\}\, d\mathbf{Z}^{(T)}(\lambda)$ tend to 0 as $T \to \infty$. This gives (3.23).

LEMMA 9.5. *Suppose* $|H(\alpha)| \leqq K(1 + |\alpha|)^{-\nu}$, $\nu > p + 1$, *then*

$$\int \cdots \int |\alpha^{(j)}|\, |H(\alpha^{(1)}) \cdots H(\alpha^{(k-1)})| \left| H\left(\beta - \sum_1^{k-1} \alpha^{(j)}\right)\right| d\alpha^{(1)} \cdots d\alpha^{(k-1)}$$
$$\leqq L(1 + |\beta|)^{-\nu+1}.$$

PROOF. It is enough for us to prove

$$\int (1 + |\alpha|)^{-\nu} |\alpha| (1 + |\beta - \alpha|)^{-\nu}\, d\alpha < M(1 + |\beta|)^{-\nu+1}.$$

Split the integral into an integral over the region $|\alpha| > \frac{1}{2}|\beta|$ and an integral over the region $|\alpha| < \frac{1}{2}|\beta|$. In the first region

$$|\alpha|(1 + |\alpha|)^{-\nu} \leqq (|\beta| + |\beta - \alpha|)(1 + \tfrac{1}{2}|\beta|)^{-\nu}.$$

In the second region

$$(1 + |\beta - \alpha|)^{-\nu} \leqq (1 + \tfrac{1}{2}|\beta|)^{-\nu}.$$

The integrals over the two regions are seen to be $O[(1 + |\beta|)^{-\nu}]$. This gives the indicated result.

PROOF OF THEOREM 3.5. If we examine the proof of Theorem 3.1, we see that we must consider

$$\int [h^{(T)}(t + u^{(1)}) \dots h^{(T)}(t + u^{(k-1)})h^{(T)}(t) - h^{(T)}(t)^k] \exp\{-i(\Lambda, t)\}\, dt$$

$$= T^p(2\pi)^{pk} \int \dots \int \left[\exp\left\{i \sum_1^{k-1} (u^{(j)}, \alpha^{(j)})/T\right\} - 1\right]$$

$$\cdot H(\alpha^{(1)}) \dots H(\alpha^{(k-1)})\, H\left(T\Lambda - \sum_1^{k-1} \alpha^{(j)}\right) d\alpha^{(1)} \dots d\alpha^{(k-1)}.$$

Now $|\exp\{i\rho\} - 1| \leqq |\rho|$ and we see the above integral is bounded by

$$KT^p(1 + T|\Lambda|)^{-\nu} \sum_1^k |u^{(j)}|$$

and the proof follows as did the proof of Theorem 3.1.

PROOF OF THEOREM 4.1. The first expression follows from (3.8), the second from (3.12).

PROOF OF THEOREM 4.2. This follows directly from Theorem 3.2 and a theorem of Slutsky (see Mann and Wald (1943)).

PROOF OF THEOREM 4.3. As for Theorem 4.2.

PROOF OF THEOREM 4.4. Expression (4.10) follows from (4.2).

For (4.11) and the indicated asymptotic normality we evaluate joint cumulants and in particular show that the standardized cumulants of order greater than 2 tend to 0. Let $w^{(T)}(u) = w(B_T u)$, $h^{(T)}(t) = h(t/T)$. Consider

$$\operatorname{cum}\{f^{(T)}_{j_1 k_1}(\lambda^{(1)}), \dots, f^{(T)}_{j_n k_n}(\lambda^{(n)})\}$$

$$(*) \qquad = (2\pi)^{-pn} \int \dots \int w^{(T)}(u^{(1)}) \dots w^{(T)}(u^{(n)}) \exp\left\{-i \sum_1^n (\lambda^{(j)}, u^{(j)})\right\}$$

$$\cdot \operatorname{cum}\{m^{(T)}_{j_1 k_1}(u^{(1)}), \dots, m^{(T)}_{j_n k_n}(u^{(n)})\}\, du^{(1)} \dots du^{(n)}.$$

Now

$$\operatorname{cum}\{m_{j_1k_1}^{(T)}(u^{(1)}), \ldots, m_{j_nk_n}^{(T)}(u^{(n)})\}$$

$$= T^{-pn} \int \cdots \int h^{(T)}(t^{(1)} + u^{(1)})h^{(T)}(t^{(1)}) \cdots h^{(T)}(t^{(n)} + u^{(n)})h^{(T)}(t^{(n)})$$

$$\cdot \operatorname{cum}\{X_{j_1}(t^{(1)} + u^{(1)})X_{k_1}(t^{(1)}), \ldots X_{j_n}(t^{(n)} + u^{(n)})X_{k_n}(t^{(n)})\}\, dt^{(1)} \cdots dt^{(n)}.$$

The cumulant on the right-hand side of this last expression is a sum of terms

$$c_{\nu_1}(s^{(1)}, \ldots, s^{(N_1)})c_{\nu_2}(s^{(N_1+1)}, \ldots, s^{(N_2)}) \cdots c_{\nu_p}(s^{(N_{p-1}+1)}, \ldots, s^{(N_p)}),$$

where ν_j is a collection of 1's, ..., r's; $s^{(j)} = t^{(\alpha)} - t^{(\beta)} + \varepsilon_j u^{(\alpha)} - \eta_j u^{(\beta)}$ with α equal one of $1, \ldots, n$ as does β and $\varepsilon_j, \eta_j = 1$ or 0. These terms correspond to indecomposable partitions and we have $n \leqq N_p \leqq 2n - 1$. Consider the transformation of variables taking

$$(t^{(1)}, \ldots, t^{(n)}, u^{(1)}, \ldots, u^{(N-n+1)}, u^{(N-n+2)}, \ldots, u^{(n)})$$

into

$$(t^{(1)}, s^{(1)}, \ldots, s^{(n-1)}, s^{(n)}, \ldots, s^{(N)}, u^{(N-n+2)}, \ldots, u^{(n)}).$$

The region of integration of these variables is bounded by a multiple of T. If $n > 2$ we bound (*) by

$$MT^{-pn}T^p(B_T^p)^{-2n+N-1}.$$

This last tends to 0 when divided by $(B_T T)^{pn/2}$.

For (4.11) we evaluate the integral after the change of variables. The above form of argument is employed in Brillinger (1968).

PROOF OF THEOREM 4.5. We indicate lemmas paralleling those of the proof of Theorem 3.3. Let

$$\hat{f}_{jk}^{(T)}(\lambda) = f_{jk}^{(T)}(\lambda) - Ef_{jk}^{(T)}(\lambda),$$

$$\rho = \operatorname{var}\{\operatorname{Re} f_{jk}^{(T)}(\lambda)\}.$$

LEMMA 9.6. *For given λ, ε and α sufficiently small*

$$E \exp\{\alpha \operatorname{Re} \hat{f}_{jk}^{(T)}(\lambda)\} \leqq \exp\{\alpha^2\rho(1 + \varepsilon)/2\}.$$

COROLLARY.

$$E \exp\{\alpha|\operatorname{Re} \hat{f}_{jk}^{(T)}(\lambda)|\} \leqq 2 \exp\{\alpha^2\rho(1 + \varepsilon)/2\}.$$

LEMMA 9.7.

$$|\operatorname{Re} \hat{f}_{jk}^{(T)}(\lambda) - \operatorname{Re} \hat{f}_{jk}^{(T)}(\mu)| \leqq KB_T^{-p-1}|\lambda - \mu|.$$

LEMMA 9.8. *There are finite M, L such that*

$$E \exp\left\{\alpha \sup_{|\lambda| \leqq \Lambda} |\mathrm{Re}\, \hat{f}_{jk}^{(T)}(\lambda)|\right\} \leqq M \exp\{LB_T^{-p-1}/R + p \log R + \alpha^2 \rho(1+\varepsilon)/2\}.$$

LEMMA 9.9. *Given $\varepsilon, \delta > 0$, let $a^2 = 2(p^2 + p + \eta)[\log 1/B_T]\rho(1+\varepsilon)$, then*

$$\mathrm{Prob}\left[\sup_{|\lambda| \leqq \Lambda} |\mathrm{Re}\, \hat{f}_{jk}^{(T)}(\lambda)| \geqq a\right] \leqq NB_T^{\eta}.$$

The indicated theorem now follows from these lemmas and similar lemmas for the variate $\mathrm{Im}\, \hat{f}_{jk}^{(T)}(\lambda)$.

PROOF OF THEOREM 4.6. Expression (4.17) follows directly from (4.10). We develop (4.18) from (4.11). The covariance is given by

$$(2\pi/T)^p H_4(0) H_2(0)^{-2} \left\{ \int\int\int W^{(T)}(|\lambda|\beta - \alpha) W^{(T)}(|\mu|\gamma - \alpha) \cdot f_{jm}(\alpha) f_{kn}(-\alpha)\, d\alpha\, d\Gamma(\beta)\, d\Gamma(\gamma) \right\}$$

$$+ \text{a similar term.}$$

Making the change of variable $\alpha = |\lambda|\beta - B_T\delta$ and noting that the second-order spectra are rotationally symmetric we replace $f_{jm}(\alpha)$ by $f_{jm}(\lambda) + O(B_T)$ for example.

We see that we must now consider

$$\int_{\alpha}\int_{\beta}\int_{\gamma} W^{(T)}(|\lambda|\beta - \alpha) W^{(T)}(|\mu| - \alpha)\, d\Gamma(\beta)\, d\Gamma(\gamma)\, d\alpha$$

$$= (2\pi)^{-p} \int J_{p/2-1}(|\lambda|\,|u|) J_{p/2-1}(|\mu|\,|u|) w^{(T)}(u)^2\, du$$

$$= (2\pi)^{-p} B_T^{-p} \int J_{p/2-1}(B_T^{-1}|\lambda|\,|v|) J_{p/2-1}(B_T^{-1}|\mu|\,|v|) w(v)^2\, dv$$

$$\doteq (2\pi)^{-p} B_T^{-p} \int B_T |\lambda|^{-1} |v|^{-1} w(v)^2\, dv\, \delta\{|\lambda| - |\mu|\}$$

if we use the following asymptotic expression for the Bessel function

$$J_n(x) \doteq x^{-1/2} \cos(x - \pi/4 - n\pi/2)$$

for large x and the Riemann-Lebesgue theorem.

Turning to the asymptotic normality, in the proof of Theorem 4.4 we saw that the cumulants of order n of the $\mathbf{f}_{XX}^{(T)}(\lambda)$ were of order

$$O(B_T^p T^p)^{-n+1}.$$

The joint cumulants of the $\mathbf{g}_{XX}^{(T)}(\lambda)$ will also be of this order. When standardized by $(B_T^{p-1}T^p)^{n/2}$ we see that the joint cumulants tend to 0 giving the required result.

PROOF OF LEMMA 5.1. Suppose $\mathbf{X}(t)$, $t \in R^p$, is a sample path of a zero mean series satisfying Assumption 3.2. Consider the variate $\mathbf{d}_R^{(T)}(\beta) - \mathbf{A}(\beta)\mathbf{d}_X^{(T)}(\beta)$. It has mean $\mathbf{0}$ and covariance matrix

$$T^{2p}\int H_1(T[\beta-\alpha])^2\{\mathbf{A}(\alpha)-\mathbf{A}(\beta)\}\mathbf{f}_{XX}(\alpha)\overline{\{\mathbf{A}(\alpha)-\mathbf{A}(\beta)\}}^{\tau}\,d\alpha < KT^{p-2},$$

for some finite K.

One can evaluate similar bounds for the higher-order cumulants. Arguments of the form of those of the proof of Theorem 3.3, then give, for finite Λ,

$$\sup_{|\beta|<\Lambda} |\mathbf{d}_R^{(T)}(\beta) - \mathbf{A}(\beta)\mathbf{d}_X^{(T)}(\beta)| < K(T^{p-2}\log T)^{1/2}$$

with probability one. This gives (5.3).

PROOF OF THEOREM 5.1. From Theorem 3.3 we see

$$\mathbf{d}_X^{(T)}(\beta) = O(T^p \log T)^{1/2}.$$

This together with the fact that $\mathbf{A}(\beta)$ has a uniformly bounded first derivative gives

$$\mathbf{A}(\lambda^{(j)}(T))\mathbf{d}_X^{(T)}(\lambda^{(j)}(T)) = \mathbf{A}(\lambda)\mathbf{d}_X^{(T)}(\lambda^{(j)}(T)) + o(T^{p/2}),$$

as $|\lambda^{(j)}(T) - \lambda| = o(\log T)^{-1/2}$. We see, from (5.4), that

$$\mathbf{d}_Y^{(T)}(\lambda^{(j)}(T)) = \mathbf{A}(\lambda)\mathbf{d}_X^{(T)}(\lambda^{(j)}(T)) + \mathbf{d}_\varepsilon^{(T)}(\lambda^{(j)}(T)) + o(T^{p/2}).$$

The remainder of the proof follows the proof of Theorem 2.1 in Brillinger (1969c).

PROOF OF THEOREM 5.2. We begin with a lemma.

LEMMA 9.10. *Under the indicated conditions*:

(i)
$$\mathbf{f}_{RX}^{(T)}(\lambda) = \int W^{(T)}(\lambda-\alpha)\mathbf{A}(\alpha)\mathbf{I}_{XX}^{(T)}(\alpha)\,d\alpha + O(T^{-1}\log T)$$
$$= \mathbf{A}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda) + O(B_T) + O(T^{-1}\log T),$$

(ii)
$$\mathbf{f}_{RR}^{(T)}(\lambda) = \mathbf{A}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)\overline{\mathbf{A}(\lambda)}^{\tau} + O(B_T) + O(T^{-1}\log T),$$

(iii) $$\mathbf{f}_{R\varepsilon}^{(T)}(\lambda) = \mathbf{A}(\lambda)\mathbf{f}_{X\varepsilon}^{(T)}(\lambda) + O(B_T) + O(T^{-1}\log T)$$

with probability one,

(iv) $$\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\lambda) + O(B_T^{-p}T^{-p}\log B_T^{-1}) + O(B_T) + O(T^{-1}\log T)$$

with probability one.

PROOF. We have

$$\mathbf{f}_{RX}^{(T)}(\lambda) = \int W^{(T)}(\lambda - \alpha)(2\pi T)^{-p}H_2(0)^{-1}\mathbf{d}_R^{(T)}(\alpha)\overline{\mathbf{d}_X^{(T)}}(\alpha)^{\mathrm{T}}\,d\alpha.$$

The first part of (i) follows from this in view of (5.3) and the fact that

$$\mathbf{d}_X^{(T)}(\alpha) = O(T^p\log T)^{1/2}.$$

The second part follows because $\mathbf{A}(\alpha)$ has a uniformly bounded first derivative.

Expression (ii) follows in a similar manner.

Expression (iii) follows directly as did expression (i) once one notes that

$$\mathbf{d}_\varepsilon^{(T)}(\alpha) = O(T^p\log T)^{1/2}$$

with probability one.

Turning to (iv), we have

$$\begin{aligned}\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) &= \mathbf{f}_{R\varepsilon}^{(T)}(\lambda) + \mathbf{f}_{R\varepsilon}^{(T)}(\lambda) + \mathbf{f}_{\varepsilon R}^{(T)}(\lambda)\\ &\quad + \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\lambda) - [\mathbf{f}_{RX}^{(T)}(\lambda) + \mathbf{f}_{\varepsilon X}^{(T)}(\lambda)]\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}[\mathbf{f}_{XR}^{(T)}(\lambda) + \mathbf{f}_{X\varepsilon}^{(T)}(\lambda)]\\ &= \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\lambda) - \mathbf{f}_{\varepsilon X}^{(T)}(\lambda)\mathbf{f}_{XX}^{(T)}(\lambda)^{-1}\mathbf{f}_{X\varepsilon}^{(T)}(\lambda) + O(B_T) + O(T^{-1}\log T)\end{aligned}$$

with probability one using (i)–(iii). The second term here is

$$O(B_T^{-p}T^{-p}\log B_T^{-1})$$

with probability one from Theorem 4.5.

PROOF OF (5.16). This follows directly from (i) of the previous lemma and the fact that $E\boldsymbol{\varepsilon}(t) = 0$.

PROOF OF (5.17). This follows from (iv) above and (4.10).

LEMMA 9.11.

$$\begin{aligned}&\operatorname{cov}\{\mathbf{f}_{\varepsilon X}^{(T)}(\lambda)_{jk}, \mathbf{f}_{\varepsilon X}^{(T)}(\mu)_{mn}\}\\ &= \mathbf{f}_{\varepsilon\varepsilon}(\lambda)_{jm}(2\pi)^pT^{-p}\int W^{(T)}(\lambda - \alpha)W^{(T)}(\mu - \alpha)[\mathbf{I}_{XX}^{(T)}(\alpha)]_{kn}\,d\alpha + o(B_T^{-p}T^{-p}).\end{aligned}$$

PROOF. The covariance is given by

$$(2\pi T)^{-2p}H_2(0)^{-2}\int\int W^{(T)}(\lambda-\alpha)W^{(T)}(\mu-\beta)\overline{\mathbf{d}_X^{(T)}(\alpha)_k}$$

$$\cdot \mathbf{d}_X^{(T)}(\beta)_n \operatorname{cov}\{\mathbf{d}_\varepsilon^{(T)}(\alpha)_j, \mathbf{d}_\varepsilon^{(T)}(\beta)_m\}\, d\alpha\, d\beta$$

$$= (2\pi)^p T^{-p}\int W^{(T)}(\lambda-\alpha)W^{(T)}(\mu-\alpha)\mathbf{I}_{XX}^{(T)}(\alpha)_{kn}\mathbf{f}_{\varepsilon\varepsilon}(\alpha)_{jn}\, d\alpha$$

$$+ \text{ a remainder of smaller magnitude.}$$

from Theorem 3.1. The lemma now follows.

PROOF OF (5.18). This result follows directly from the previous lemma.

LEMMA 9.12.

$$\operatorname{cov}\{\mathbf{f}_{\varepsilon X}^{(T)}(\lambda)_{jk}, \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\mu)\} = o(B_T^{-p}T^{-p}).$$

PROOF. This follows in the manner of Lemma 9.11.

PROOF OF (5.19). From the previous lemma.

PROOF OF (5.20). From (iv) and (4.13).

LEMMA 9.13. *Under the assumptions of the theorem,*

$$\operatorname{cum}\{\mathbf{f}_{\varepsilon X}^{(T)}(\lambda^{(1)})_{j_1k_1}, \ldots, \mathbf{f}_{\varepsilon X}^{(T)}(\lambda^{(m)})_{j_mk_m}, \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\lambda^{(m+1)})_{j_{m+1},k_{m+1}}, \ldots, \mathbf{f}_{\varepsilon\varepsilon}^{(T)}(\lambda^{(n)})_{j_nk_n}\}$$
$$= o(B_T^{-np/2}T^{-np/2}).$$

PROOF. One simply extends the argument of Lemma 9.11.

PROOF OF ASYMPTOTIC NORMALITY. The standardized joint cumulants of order greater than 2 are seen to tend to 0 from the previous lemma.

PROOF OF THEOREM 5.3. This follows from the results of Theorem 5.2 and the lemmas developed in the course of the proof of that theorem.

PROOF OF THEOREM 6.1. Let $\hat{\mathbf{f}}_{YX}^{(T)}(\lambda) = \mathbf{f}_{YX}^{(T)}(\lambda) - \mathbf{f}_{YX}(\lambda)$ with a similar notation for $\hat{\mathbf{f}}_{YY}^{(T)}(\lambda)$, $\hat{\mathbf{f}}_{XX}^{(T)}(\lambda)$. We have the Taylor expansions

$$\mathbf{A}^{(T)}(\lambda) = \mathbf{A}(\lambda) + \hat{\mathbf{f}}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1} - \mathbf{f}_{YX}(\lambda)\hat{\mathbf{f}}_{XX}^{(T)}(\lambda)\mathbf{f}_{XX}(\lambda)^{-2} + \ldots,$$

$$\mathbf{g}_{\varepsilon\varepsilon}^{(T)}(\lambda) = \mathbf{f}_{\varepsilon\varepsilon}(\lambda) + \hat{\mathbf{f}}_{YY}^{(T)}(\lambda) - \hat{\mathbf{f}}_{YX}^{(T)}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)$$
$$- \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\hat{\mathbf{f}}_{XY}^{(T)}(\lambda) + \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\hat{\mathbf{f}}_{XX}^{(T)}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda) + \ldots.$$

Now the indicated first- and second-order moments follow from (4.10), (4.11) using a theorem of Brillinger and Tukey (1964). The asymptotic normality follows from Mann and Wald (1943).

PROOF OF THEOREM 6.2. This follows from the results of Theorem 6.1

PROOF OF THEOREM 6.3. The first part follows as did Theorem 4.3. The second part follows from a theorem of Slutsky (Mann and Wald (1943))

PROOF OF THEOREM 7.1. A direct generalization of Theorem 2 of Brillinger (1969b).

PROOF OF THEOREM 8.1. Using (1.8) and (3.26) we see that the expected value of interest is given by

$$= \Delta_T^{p(k-1)} \sum_{s^{(1)}} \cdots \sum_{s^{(k)}} W^{(T)}(\lambda^{(1)} - s^{(1)}\Delta_T, \ldots, \lambda^{(k)} - s^{(k)}\Delta_T)\Psi(s^{(1)}, \ldots, s^{(k)})$$

$$\cdot f_{a_1 \ldots a_k}(s^{(1)}\Delta_T, \ldots, s^{(k-1)}\Delta_T) + O(T^{-1}) + O(T^{-v}\Delta_T^{-v}),$$

for v sufficiently large.

The error involved in replacing the sum by an integral is

$$O(\Delta_T B_T^{-1}).$$

PROOF OF THEOREM 8.2. One argues as in Brillinger and Rosenblatt (1967) using (3.26) and the rules of Leonov and Širjaev (1959).

PROOF OF THEOREM 8.3. One proves that the standardized joint cumulants of order greater than 2 tend to 0. This was done in Brillinger and Rosenblatt (1967).

REFERENCES

Avadhani, T. V. (1969), *Sampling errors of covariances and correlations for K-dimensional stationary stochastic processes*, Sankhyā.

Bartlett, M. S. (1964), *The spectral analysis of two-dimensional point processes*, Biometrika **51**, 299–311. MR **30** #5439.

——— (1966), *An introduction to stochastic processes, with special reference to methods and applications*, Cambridge Univ. Press, New York. MR **35** #7385.

Batchelor, G. K. (1953), *The theory of homogeneous turbulence*, Cambridge Monographs on Mechanics and Appl. Math., Cambridge Univ. Press, New York. MR **14**, 597.

Birendranath, Gh. (1951), *On multidimensional stochastic fields*, Bull. Calcutta Statist. Soc. **4**, 29–32.

Blanc-Lapierre, A. (1963), *Modèles statistiques pour l'étude de phénomères de fluctuations*, Masson, Paris.

Bochner, S. (1956), *Stationarity, boundedness, almost periodicity of random-valued functions*, Proc. Third Berkeley Sympos. Math. Statist. and Prob. (Berkeley, 1954/55) vol. II, Univ. of California Press, Berkeley, Calif., pp. 7–27. MR **18**, 940.

——— (1960), *Harmonic analysis and the theory of probability*, Univ. of California Press, Berkeley, Calif.

Bochner, S. and Chandrasekharan, K. (1949), *Fourier transforms*, Ann. of Math. Studies, no. 19, Princeton Univ. Press, Princeton, N.J. MR **11**, 173.

Bracewell, R. N. (1956), *Two-dimensional aerial smoothing in radio astronomy*, Austral. J. Phys. **9**, 297–314. MR **18**, 365.

Brillinger, D. R. (1968), *Estimation of the cross-spectrum of a stationary bivariate Gaussian process from its zeros*, J. Roy. Statist. Soc. Ser. B **30**, 145–159.

——— (1969a), *Asymptotic properties of spectral estimates of second-order*, Biometrika **56**, 375–390.

Brillinger, D. R. (1969b), *The canonical analysis of stationary time series*. Multivariate Analysis II (P. R. Krishnaiah, editor) Academic Press, New York, pp. 331–350.

——— (1969c), *A search for a relationship between monthly sunspot numbers and certain climatic series*, Bull. Inst. Internat. Statist. **43**, 293–306.

——— (1970), *The identification of polynomial systems by means of higher order spectra*, J. Sound and Vibration.

Brillinger, D. R. and Rosenblatt, M. (1967), *Asymptotic theory of estimates of k-th order spectra*, Proc. Advanced Sem. Spectral Analysis of Time Series (Madison, Wis., 1966) Wiley, New York, pp. 153–188. MR **35** # 1168.

Brillinger, D. R. and Tukey, J. W. (1964), *Asymptotic variances, moments, cumulants and other average values* (unpublished manuscript).

Cox, C. and Munk, W. (1954), *Measurements of the roughness of the sea surface from photographs of the sun's glitter*, J. Opt. Soc. Amer. **44**, 838–850.

Cramér, H. and Leadbetter, M. R. (1967), *Stationary and related stochastic processes. Sample function properties and their applications*, Wiley, New York. MR **36** # 949.

Curry, L. (1969), *Univariate spatial forecasting*, London Invitational Conference Commission on Quantitative Methods, International Geographical Union.

Czan, C.-P. (1957), *On the linear extrapolation of a discrete homogeneous random field*, Dokl. Akad. Nauk SSSR **112**, 207–210.

Dudley, R. M. (1965), *Gaussian processes on several parameters*, Ann. Math. Statist. **36**, 771–788. MR **33** #796.

Elkins, T. A. (1968), *Cubical and spherical estimation of multivariate probability density*, J. Amer. Statist. Assoc. **63**, 1495–1513.

Embree, P., Burg, J. P. and Backus, M. M. (1963), *Wide band velocity filtering*, Geophysics **28**, 948–974.

Fail, M. D. and Grau, G. (1963), *Les filters en eventail*, Geophysical Prospecting **II**, 131–163.

Fortus, M. I. (1963), *Extrapolation of a random field satisfying the wave equation*, Teor. Verojatnost. i Primenen. **8** (1963), 220–223 = Theor. Probability Appl. **8**, 204–207. MR **27** #5291.

Gazdag, J. (1970), *Bubble-chamber picture processing through digital filtering*, J. Computational Phys. **5**, 34–55.

Gel'fand, I. M. and Vilenkin, N. Ja. (1964), *Generalized functions*. Vol. 4: *Some applications of harmonic analysis*, Fizmatgiz, Moscow, 1961; English transl., Academic Press, New York. MR **26** #4173; MR **30** #4152.

Ghosh, B. (1951), *Some exponential forms for topographic correlation*, Sankhyā **11**, 29–36. MR **13**, 141.

Goodman, J. W. (1968), *Introduction to Fourier optics*, McGraw-Hill, New York.

Goodman, N. R. (1963), *Statistical analysis based on a certain multivariate complex Gaussian distribution. (An introduction)*, Ann. Math. Statist. **34**, 152–177. MR **26** #3148a.

Granger, C. W. J. (1968), "Spatial data and time series analysis," in *London studies in regional science* (A. J. Scott, editor), Pion, London.

Granger, C. W. J. and Hatanaka, M. (1964), *Spectral analysis of economic time series*, Princeton Studies in Math. Economics, no. 1, Princeton Univ. Press, Princeton, N.J. MR **30** #4368.

Grenander, U. and Rosenblatt, M. (1956), *Some problems in estimating the spectrum of a time series*, Proc. Third Berkeley Sympos. Math. Statist. and Prob. (Berkeley, Calif., 1954/55) vol. 1, Univ. of California Press, Berkeley, Calif., pp. 77–93. MR **18**, 946.

Grenander, U. and Rosenblatt, M. (1957), *Statistical analysis of stationary time series*, Wiley, New York. MR **18**, 959.

Hannan, E. J. (1963), "Regression for time series," in *Time series analysis*, (M. Rosenblatt, editor) Wiley, New York, pp. 17–37.

——— (1965), *Group representations and applied probability*, J. Appl. Probability **2**, 1–68, MR **35** # 2307a.

——— (1965), *Aliasing*, Report No. 49(638)-1302, Dept. of Statistics, Johns Hopkins University.

——— (1966), *Spectral analysis for geophysical data*, Geophys. J. R. Astr. Soc. **11**, 225–236.

——— (1967), *The estimation of a lagged regression relation*, Biometrika **54**, 409–418. MR **36** #4759.

Harbaugh, J. W. and Preston, F. W. (1968), "Fourier series analysis in geology," in *Spatial analysis: A reader in statistical geography* (B. J. L. Berry and S. F. Marble, editors) Prentice-Hall, Englewood Cliffs, N.J., pp. 218–238.

Heine, V. (1955), *Models for two-dimensional stationary stochastic processes*, Biometrika **42**, 170–178. MR **17**, 167.

Hoffman, W. C. (1964), *Wave propagation in a general random continuous medium*, Proc. Sympos. Appl. Math., vol. 16, Amer. Math. Soc., Providence, R. I., pp. 117–144. MR **28** #5686.

Itô, K. (1956), *Isotropic random current*, Proc. Third Berkeley Sympos. Math. Statist. and Prob. (Berkeley, Calif., 1954/55) vol. 2, Univ. of California Press, Berkeley, Calif., pp. 125–132. MR **18**, 941.

Jaglom, A. M. (1957), *Certain types of random fields in n-dimensional space similar to stationary stochastic processes*, Teor. Verojatnost. i Primenen. **2** (1957), 292–338 = Theor. Probability Appl. **2**, 273–320. MR **20** #1353.

——— (1961), *Second-order homogeneous random fields*, Proc. Fourth Berkeley Sympos. Math. Statist. and Prob., vol. 2, Univ. of California Press, Berkeley, Calif., pp. 593–622. MR **26** #4399.

——— (1962), *An introduction to the theory of stationary random functions*, Uspehi Mat. Nauk **7** (1952), no. 5 (51), 3–168; English transl., Prentice-Hall, Englewood Cliffs, N.J. MR **14**, 485; MR **32** #1762.

Jajte, R. (1967), *On stochastic processes over an abelian locally compact group*, Colloq. Math. **17**, 351–355. MR **36** #6001.

Jones, R. H. (1963), *Stochastic processes on a sphere*, Ann. Math. Statist. **34**, 213–218. MR **30** #616.

Kahane, J. (1968), *Some random series of functions*, Heath, Lexington, Ky.

Kampé de Fériet, J. (1948), *Analyse harmonique des fonctions aléatoires stationnaires d'ordre 2 définies sur un group abélian localement compact*, C. R. Acad. Sci. Paris **226**, 868–870. MR **9**, 520.

Kawata, T. (1966), *On the Fourier series of a stationary stochastic process*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **6**, 224–245. MR **35** #6200.

Knopov, P. S. (1965), *On estimating the spectrum of a homogeneous and isotropic Gaussian field*, Kibernetika (Kiev) **1965**, no. 6, 59–64. (Russian) MR **35** #2361.

——— (1966), *Some remarks on hypothesis testing for random fields*, Kibernetika (Kiev) **1966**, no. 6, 51–52. (Russian) MR **35** #3725.

Kovasznay, L. S. G. and Arman, A. (1957), *Optical autocorrelation measurement of two-dimensional random patterns*, Rev. Sci. Instr. **28**, 793–797.

Lee, Y. W. (1960), *Statistical theory of communication*, Wiley, New York. MR **24** #B441.

Leonov, V. P. and Širjaev, A. N. (1959), *Sur le calcul des semi-invariants*, Te
Verojatnost. i Primenen. **4** (1959), 342–355 = Theor. Probability Appl. **4**, 319–3
MR **22** #5058.
Lin, C. C. (1959), *Statistical theories of turbulence. Turbulent flows and heat transf*
High Speed Aerodynamics and Jet Propulsion, vol. 5, Princeton Univ. Pre
Princeton, N.J. MR **21** #6174.
Longuett-Higgins, M. S. (1957), *The statistical analysis of a random, moving surfa*
Philos. Trans. Roy. Soc. London Ser. A **249**, 321–387. MR **19**, 328.
——— (1962), *The statistical geometry of random surfaces*, Proc. Sympos. Ap
Math., vol. 13, Amer. Math. Soc., Providence, R. I., pp. 105–143. MR **25** #359
Lord, R. D. (1954), *The use of the Hankel transform in statistics*. I, II, Biometrika
44–55; II, 344–350. MR **15**, 885; MR **16**, 382.
Mandelbrojt, M. (1935), *Séries de Fourier et classes quasianalytiques de fonctio*
Gauthier-Villars, Paris.
Mann, H. B. and Wald, A. (1943), *On stochastic limit and order relationships*, An
Math. Statist. **14**, 217–226. MR **5**, 125.
Matérn, B. (1960), *Spatial variation: Stochastic models and their application to son*
problems in forest surveys and other sampling investigations, Almqvist, Stockholr
Moore, C. N. (1966), *Summable series and convergence factors*, Dover, New Yor
MR **34** #1743.
Neumann, P. (1965), *Spektralschätzungen stationärer zufällinger Felder*, Teo
Verojatnost. i Primenen. **10** (1965), 140–143 = Theor. Probability Appl. **1**
128–132. MR **30** #5374.
O'Neill, E. L. (1956), *Spatial filtering in optics*, IEEE Trans. Information Theor
IT-2, 56.
Paley, R. E. A. C. and Wiener, N. (1934), *Fourier transforms in the complex domai*
Amer. Math. Soc. Colloq. Publ., vol. 19, Amer. Math. Soc., Providence, R.I
reprint, 1967.
Parzen, E. (1957), *On consistent estimates of the spectrum of a stationary time serie*
Ann. Math. Statist. **28**, 329–348. MR **19**, 587.
Petersen, D. P. and Middleton, D. (1962), *Sampling and reconstruction of wave*
number-limited functions in N-dimensional Euclidean spaces, Information an
Control **5**, 279–323. MR **27** #1317.
Pierson, W. J. and Tick, L. J. (1955), *Stationary random processes in meteorolog*
and oceanography, Bull. Inst. Internat. Statist. **35**, 271–281.
Pinsker, M. S. (1957), *Extrapolation of homogeneous stochastic fields and amount o*
information about a Gaussian field contained in another Gaussian stochastic field
Dokl. Akad. Nauk SSSR **112**, 815–818. (Russian) MR **19**, 470.
Popov, Ju. D. (1967), *Some problems of linear prediction for homogeneous random*
fields, Dopovīdī Akad. Nauk Ukraïn. RSR Ser. A **1967**, 1076–1078. (Ukrainian
MR **36** #4632.
Priestley, M. B. (1964), *The analysis of two-dimensional stationary processes wit*
discontinuous spectra, Biometrika **51**, 195–217. MR **30** #3539.
Preston, K., Jr. (1965), "Use of the Fourier transformable properties of lenses for
signal spectrum analysis," in *Optical and electro-optical information processing*
(J. T. Tippett, editor) M.I.T. Press, Cambridge, Mass., BD 75.
Rao, M. M. (1969), *Representation theory of multidimensional generalized random*
fields. Multivariate Analysis-II (P. R. Krishnaiah, editor), Academic Press,
New York, pp. 411–436.
Robinson, E. A. (1967), *Statistical communication and detection*, Griffin, London.

Robinson, J. E. (1970), *Spatial filtering of geological data*, Rev. Inst. Internat. Statist. **38**, 21–32.

Rosenblatt, M. (1957), *A random model of the sea surface generated by a hurricane*, J. Math. Mech. **6**, 235–246. MR **19**, 87.

——— (1959), *Statistical analysis of stochastic processes with stationary residuals. Probability and statistics: The Harald Cramér volume*, Wiley, New York, pp. 246–275. MR **21** #6675.

Salem, R. and Zygmund, A. (1956), *A note on random trigonometric polynomials*, Proc. Third Berkeley Sympos. Math. Statist. and Prob., vol. 2, Univ. of California Press, Berkeley, Calif., pp. 243–246. MR **18**, 891.

Shapiro, V. L. (1964), *Fourier series in several variables*, Bull. Amer. Math. Soc. **70**, 48–93. MR **28** #1448.

Širjaev, A. N. (1960), *Some problems in the spectral theory of higher-order moments*. I, Teor. Verojatnost. i Primenen. **5** (1960), 293–313 = Theor. Probability Appl. **5**, 265–284. MR **24** #A2988.

Tatarski, V. I. (1961), *Wave propagation in a turbulent medium*, Izdat. Akad. Nauk SSSR, Moscow, 1959; English transl., McGraw-Hill, New York. MR **23** # B716.

Tick, L. J. (1960), *Contributions to the theory and applications of stationary random processes in fluid mechanics*, Ph.D. Thesis, Columbia University, New York.

Tobler, W. R. (1966), *Spectral analysis of spatial series*, Proc. Fourth Annual Conf. Urban Plan. Inf. Syst. Prog., Univ. of California Press, Berkeley, Calif., pp. 179–186.

Tukey, J. W. (1965), *Uses of numerical spectrum analysis in geophysics*, Bul. I.S.I., 35th Session, pp. 267–307.

——— (1967), *An introduction to the calculations of numerical spectrum analysis*, Proc. Advanced Sem. Spectral Analysis Time Series (Madison, Wis., 1966) Wiley, New York, pp. 25–46. MR**35** #3842.

Vander Lugt, A. B. (1964), *Signal detection by complex spatial filtering*, IEEE Trans. Information Theory **IT-10**, 2.

Wahba, G. (1968), *On the distribution of some statistics useful in the analysis of jointly stationary time series*, Ann. Math. Statist. **39**, 1849–1862. MR **38** #5363.

Whittle, P. (1954), *On stationary processes in the plane*, Biometrika **41**, 434–449. MR **16**, 731.

——— (1959), *Sur la distribution du maximum d'un polynôme trigonométrique à coefficients aléatoires*, Le calcul de probabilités et ses applications (Paris, 1958) Colloques Internationaux du Centre National de la Recherche Scientifique, 87, Centre Nationale de la Recherche Scientifique, Paris, pp. 173–184. MR **21** #4472.

——— (1962), *Topographic correlation, power-law covariance functions, and diffusion*, Biometrika **49**, 305–314. MR **31** #5305.

——— (1963), *Stochastic processes in several dimensions*, Bull. Inst. Internat. Statist. **40**, 974–994. MR **30** #3500.

Wiener, N. (1930), *Generalized harmonic analysis*, Acta. Math. **55**, 117–258.

Wiener, N. and Wintner, A. (1941), *On the ergodic dynamics of almost periodic systems*, Amer. J. Math. **63**, 794–824. MR **4**, 15.

Woodroofe, M. B. and Van Ness, J. W. (1967), *The maximum deviation of sample spectral densities*, Ann. Math. Statist. **38**, 1558–1569. MR **35** #7546.

Zurflueh, E. G. (1967), *Applications of two-dimensional linear wavelength filtering*, Geophysics **32**, 1015–1035.