Pp. 331-350 in Multivariate Analysis-II. (Ed. P.R. Krishnaiah) Academic. (1969)

The Canonical Analysis of Stationary Time Series[1]

DAVID R. BRILLINGER

DEPARTMENT OF STATISTICS[2]
THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
ALDWYCH, LONDON

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

1. SUMMARY

Given a vector-valued stationary series with finite second-order moments, this article considers a class of problems leading to the calculation of the latent values and vectors of matrices based on the spectral density matrix of the series. Certain statistical properties of these characteristics are considered for the case in which they are based on an estimated spectral density matrix. Calculations,[3] of the type discussed in this article, are carried out for a series consisting of the mean monthly temperatures recorded at fourteen European cities over a 170 year period.

2. INTRODUCTION AND NOTATION

Throughout the article, matrices will be denoted by boldface letters, **A**, **a**, for example. The term **I** will denote the identity matrix and **0** the matrix with all entries 0. The term $\mathbf{A}^\tau$ will denote the transpose of **A**, and $\overline{\mathbf{A}}$ will denote the matrix whose entries are the complex conjugates of those of **A**. The term A_{jk} will denote the entry in the *j*th row and *k*th column of **A**, and we will write

[1] This research was partially supported by the National Science Foundation Grant NSF-7454.

[2] Present address. The author was on leave of absence from The London School of Economics when this paper was written.

[3] These calculations were carried out at the Bell Telephone Laboratories, Murray Hill, New Jersey, on their GE 645 computer while the author was a summer visitor in 1967. The original data was provided by J. M. Craddock of the Meteorological Office, England.

$\mathbf{A} = [A_{jk}]$. If $\mathbf{A} = [A_{jk}]$, $|\mathbf{A}|$ will denote $[|A_{jk}|]$. The trace of $\mathbf{A}$ will be denoted by $\mathrm{tr}\,\mathbf{A}$. If the Hermitian matrix $\mathbf{A}$ is nonnegative definite, we will write $\mathbf{A} \geq \mathbf{0}$. The latent values $\mu_1, \mu_2, \ldots$, of a Hermitian $\mathbf{A}$ will be ordered so that $\mu_1 \geq \mu_2 \geq, \ldots$. If $\mathbf{a}(t) = [a_{jk}(t)]$, $t = 0, \pm 1, \ldots$ is an $s \times q$ matrix-valued function and $\mathbf{b}(t) = [b_{jk}(t)]$ is a $q \times r$ matrix-valued function satisfying

$$\sum_{t=-\infty}^{\infty} |a_{jk}(t)| < \infty \tag{2.1}$$

$$\sum_{t=-\infty}^{\infty} |b_{jk}(t)| < \infty, \tag{2.2}$$

we will write $\mathbf{a} * \mathbf{b}(t)$ for the convolution

$$\sum_{u=-\infty}^{\infty} \mathbf{a}(t-u)\mathbf{b}(u). \tag{2.3}$$

If $\mathbf{Z} = [Z_j]$ is an r vector-valued random variable, and $E|Z_j|^r < \infty$, $j = 1, \ldots, r$, we denote its joint cumulant of order r by

$$\mathrm{cum}[Z_1, \ldots, Z_r]. \tag{2.4}$$

Let $\mathbf{W}(t)$, $t = 0, \pm 1, \ldots$ be an m vector-valued time series with real-valued components. Suppose the second order moments of $\mathbf{W}(t)$ are finite, and

$$E\mathbf{W}(t) = \mathbf{c}_W, \qquad E\{\mathbf{W}(t+u) - \mathbf{c}_W\}\{\mathbf{W}(t) - \mathbf{c}_W\}^\tau = \mathbf{c}_{WW}(u) \tag{2.5}$$

for $t, u = 0, \pm 1, \ldots$.

Definition 2.1. *$H_W{}^n$ is the closure, in the norm, $E\{\mathrm{tr}\,\mathbf{ff}^\tau\}$, of the space of n vector-valued linear combinations, $\mathbf{f}$, of the form*

$$\mathbf{f} = \sum_t \mathbf{a}(t)\{\mathbf{W}(t) - \mathbf{c}_W\} \tag{2.6}$$

where $\mathbf{a}(t)$ is $n \times m$ matrix-valued and such that only a finite number of the $\mathbf{a}(t)$ are nonzero.

The space $H_W{}^n$ is the space of n vector-values of the series $\mathbf{W}(t)$. An $n \times n$ matrix valued inner product, $\langle \mathbf{f}, \mathbf{g} \rangle$ may be introduced in $H_W{}^n$ by

$$\langle \mathbf{f}, \mathbf{g} \rangle = E\mathbf{fg}^\tau \tag{2.7}$$

and, $\mathbf{f}, \mathbf{g} \in H_W{}^n$. Now $H_W{}^n$ becomes an LVH space in the terminology of Loynes [13]. Related spaces are considered by Masani [15] and Rozanov [23].

Let

$$\mathbf{W}(t) - \mathbf{c}_W = \int_{-\pi}^{\pi} \exp\{i\lambda t\}\, d\mathbf{Z}_W(\lambda) \tag{2.8}$$

be the Cramér representation of $\mathbf{W}(t)$, and

$$\mathbf{c}_{WW}(u) = \int_{-\pi}^{\pi} \exp\{i\lambda u\}\, d\mathbf{F}_{WW}(\lambda) \tag{2.9}$$

be the Bochner representation of $\mathbf{c}_{WW}(u)$. (See Rozanov [23] for a discussion of these representations.) In (2.8), $\mathbf{Z}_W(\lambda)$ is an m vector-valued process with orthogonal increments. In (2.9), $\mathbf{F}_{WW}(\lambda)$ is an $m \times m$ matrix-valued, bounded, Hermitian, nondecreasing function.

Definition 2.2. *$L_2^{n \times m}(\mathbf{F}_{WW})$ is the space of $n \times m$ matrix-valued functions $\mathbf{A}(\lambda)$ (with complex-valued entries), such that*

$$\int_{-\pi}^{\pi} \mathrm{tr}\{\mathbf{A}(\lambda)\, d\mathbf{F}_{WW}(\lambda)\overline{\mathbf{A}(\lambda)}^\tau\} < \infty. \tag{2.10}$$

See Rosenberg [22] for the specific meaning of (2.10). We have:

Theorem 2.1. *If $\mathbf{f} \in H_W{}^m$, then there exists $\mathbf{A}(\lambda) \in L_2^{n \times m}(\mathbf{F}_{WW})$ such that*

$$\mathbf{f} = \int_{-\pi}^{\pi} \mathbf{A}(\lambda)\, d\mathbf{Z}_W(\lambda). \tag{2.11}$$

This result is proved by Rosenberg [22].

Suppose now that $\mathbf{W}(t)$ has the form,

$$\mathbf{W}(t) = \begin{bmatrix} \mathbf{X}(t) \\ \mathbf{Y}(t) \end{bmatrix} \tag{2.12}$$

with $\mathbf{X}(t)$ r vector-valued, $\mathbf{Y}(t)$ s vector-valued. In obvious notation, we write

$$\mathbf{c}_W = \begin{bmatrix} \mathbf{c}_X \\ \mathbf{c}_Y \end{bmatrix} \tag{2.13}$$

$$\mathbf{c}_{WW}(u) = \begin{bmatrix} \mathbf{c}_{XX}(u) & \mathbf{c}_{XY}(u) \\ \mathbf{c}_{YX}(u) & \mathbf{c}_{YY}(u) \end{bmatrix} \tag{2.14}$$

$$\mathbf{F}_{WW}(\lambda) = \begin{bmatrix} \mathbf{F}_{XX}(\lambda) & \mathbf{F}_{XY}(\lambda) \\ \mathbf{F}_{YX}(\lambda) & \mathbf{F}_{YY}(\lambda) \end{bmatrix} \tag{2.15}$$

$$\mathbf{Z}_W(\lambda) = \begin{bmatrix} \mathbf{Z}_X(\lambda) \\ \mathbf{Z}_Y(\lambda) \end{bmatrix}. \tag{2.16}$$

Consider $H_W{}^s$. We note that $\mathbf{Y}(t) - \mathbf{c}_Y \in H_W{}^s$. We also note that

$$\sum_t \mathbf{d} * \mathbf{b}(t)\{\mathbf{X}(t) - \mathbf{c}_X\} \tag{2.17}$$

is in $H_W{}^s$ if $\mathbf{d}(t)$ is $s \times q$ matrix-valued, $\mathbf{b}(t)$ is $q \times r$ matrix valued, $q \leq s, r$, and only a finite number of the coefficients are nonzero.

Definition 2.3. *J_X^q is the closure in H_W^s of the set of s vectors of the form* (2.17).

Theorem 2.2. *If $\mathbf{f} \in J_X^q$, then*

$$\mathbf{f} = \int_{-\pi}^{\pi} \mathbf{A}(\lambda)\, d\mathbf{Z}_X(\lambda) \tag{2.18}$$

where $\mathbf{A}(\lambda) \in L_2^{s\times r}(\mathbf{F}_{XX})$ has rank $\le q$ for almost all λ.[4]

Intuitively, J_X^q may be thought of as containing s vectors derived from $\mathbf{X}(t)$ by first filtering $\mathbf{X}(t)$ to have dimension q and then filtering this result to have dimension s.

We note that $\mathbf{A}(\lambda)$, of the theorem, may be written $\mathbf{D}(\lambda)\mathbf{B}(\lambda)$ where $\mathbf{D}(\lambda)$ is $s \times q$ and $\mathbf{B}(\lambda)$ is $q \times r$ matrix-valued.

3. CANONICAL VARIATES FOR TIME SERIES

Suppose that $\mathbf{X}(t)$ is an r vector-valued series related to the s vector-valued series $\mathbf{Y}(t)$. Suppose that $q \le r$ channels are available for the transmission of values derived from the $\mathbf{X}(t)$ series and that on receipt of these values one desires to form a series near $\mathbf{Y}(t)$. One can imagine forming the series

$$\zeta(t) = \sum_u \mathbf{b}(t-u)\{\mathbf{X}(u) - \mathbf{c}_X\}, \tag{3.1}$$

with $\mathbf{b}(u)$ $q \times r$ matrix-valued, and transmitting this series over the q available channels. On receipt one can form

$$\mathbf{m} + \sum_u \mathbf{d}(t-u)\zeta(u) \tag{3.2}$$

where $\mathbf{d}(u)$ is $s \times q$ matrix-valued and $\mathbf{m}$ an s vector. Consider the problem of choosing $\mathbf{m}$, $\mathbf{b}(u)$, $\mathbf{d}(u)$ so that the value of (3.2) is near that of $\mathbf{Y}(t)$. From the form of (3.2), we can consider the problem of finding $\mathbf{m} + \hat{\mathbf{Y}}(t)$, $\hat{\mathbf{Y}}(t) \in J_X^q$, with a value that is near $\mathbf{Y}(t)$. We have:

Theorem 3.1. *Let $\mathbf{W}(t)$ $t = 0, \pm 1, \ldots$ be an $(r+s)$ vector-valued time series of the form* (2.2). *Suppose*

$$d\mathbf{F}_{WW}(\lambda) = \begin{bmatrix} \mathbf{f}_{XX}(\lambda) & \mathbf{f}_{XY}(\lambda) \\ \mathbf{f}_{YX}(\lambda) & \mathbf{f}_{YY}(\lambda) \end{bmatrix} d\Phi(\lambda) \tag{3.3}$$

for some[5] *nonnegative measure $\Phi(\lambda)$. Suppose $\mathbf{f}_{XX}(\lambda)$ is nonsingular for almost all λ, then*

$$E\{\mathbf{Y}(t) - \mathbf{m} - \hat{\mathbf{Y}}(t)\}^\tau\{\mathbf{Y}(t) - \mathbf{m} - \hat{\mathbf{Y}}(t)\} \tag{3.4}$$

[4] The proof of this, and other theorems, may be found in Section 7.
[5] Such a measure always exists. Take $\Phi(\lambda)$ proportional to tr $\mathbf{F}_{WW}(\lambda)$, for example.

is minimized, for $\hat{\mathbf{Y}}(t) \in J_X^q$, by

$$\mathbf{m} = \mathbf{c}_Y \tag{3.5}$$

$$\hat{\mathbf{Y}}(t) = \int_{-\pi}^{\pi} \exp\{i\lambda t\}\mathbf{A}(\lambda)\, d\mathbf{Z}_X(\lambda) \tag{3.6}$$

where

$$\mathbf{A}(\lambda) = \sum_{j=1}^{q} \mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^\tau \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}. \tag{3.7}$$

Here, $\mathbf{V}_j(\lambda)$ is the jth latent vector of $\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)$. Letting $\mu_j(\lambda)$ denote the corresponding latent root, the minimum achieved in (3.4) *is*

$$\int_{-\pi}^{\pi} \operatorname{tr}\{\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)\}\, d\Phi(\lambda)$$
$$+ \int_{-\pi}^{\pi} \{\mu_{q+1}(\lambda) + \cdots + \mu_Q(\lambda)\}\, d\Phi(\lambda) \tag{3.8}$$

with $Q = \min(r, s)$. The spectral density matrix of the residuals $\mathbf{Y}(t) - \mathbf{c}_W - \hat{\mathbf{Y}}(t)$ is

$$\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda) + \sum_{j=q+1}^{Q} \mu_j(\lambda)V_j(\lambda)\overline{V_j(\lambda)}^\tau \tag{3.9}$$

with respect to the measure $\Phi(\lambda)$.

This theorem was presented by Brillinger [1]. For the case of vector-valued random variables see Rao [20, 21] (see also Kramer and Matthews [12]).

Corollary 3.1.1. *Under the conditions of the theorem, if $q = r$,*

$$\mathbf{A}(\lambda) = \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}. \tag{3.10}$$

The spectral density matrix of the residuals is

$$\mathbf{f}_{YY}(\lambda) - \mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda) \tag{3.11}$$

with respect to the measure $\Phi(\lambda)$.

In the case $s = 1$, this corollary provides the result of the usual multiple linear time invariant regression of time series (see Koopmans [11], Rozanov [23], Masani [15], Parzen [17]).

Corollary 3.1.2. *Let $\mathbf{X}(t)$ $t = 0, \pm 1, \ldots$ be an r vector-valued time series having spectral density matrix $\mathbf{f}_{XX}(\lambda)$ with respect to the nonnegative measure $\Phi(\lambda)$. For $\hat{\mathbf{X}}(t) \in J_X^q$,*

$$E\{\mathbf{X}(t) - \mathbf{m} - \hat{\mathbf{X}}(t)\}^\tau\{\mathbf{X}(t) - \mathbf{m} - \hat{\mathbf{X}}(t)\} \tag{3.12}$$

is minimized by

$$\mathbf{m} = \mathbf{c}_X \tag{3.13}$$

$$\mathbf{X}(t) = \int_{-\pi}^{\pi} \exp\{i\lambda t\}\mathbf{A}(\lambda)\, d\mathbf{Z}_X(\lambda) \tag{3.14}$$

where

$$\mathbf{A}(\lambda) = \sum_{j=1}^{q} \mathbf{V}_j(\lambda)\overline{\mathbf{V}_j(\lambda)}^\tau. \tag{3.15}$$

Here, $\mathbf{V}_j(\lambda)$ *is the jth latent vector of* $\mathbf{f}_{XX}(\lambda)$, *and* $\mu_j(\lambda)$ *the corresponding latent root. If* $\mathbf{V}(\lambda) = [\mathbf{V}_1(\lambda), \ldots, \mathbf{V}_r(\lambda)]$, *the spectral density matrix of* $\int_{-\pi}^{\pi} \exp\{i\lambda t\} \overline{\mathbf{V}(\lambda)}^\tau d\mathbf{Z}_X(\lambda)$ *is* $\operatorname{diag}\{\mu_1(\lambda), \ldots, \mu_r(\lambda)\}$ *with respect to* $\Phi(\lambda)$.

In Section 5, we provide a worked example of this corollary. The latent roots and vectors of spectral density matrices appear in the work of Wiener [26], Whittle [25], and Pinsker [18].

Notice that the criterion (3.4) is not invariant with respect to linear time invariant transformations (filterings) of $\mathbf{Y}(t)$. For this reason, we may sometimes take, for an $s \times s$ positive definite $\mathbf{\Gamma}(\lambda)$,

$$\mathbf{A}(\lambda) = \mathbf{\Gamma}(\lambda)^{-1/2}\left\{\sum_{j=1}^{q} \mathbf{U}_j(\lambda)\overline{\mathbf{U}_j(\lambda)}^\tau\right\}\mathbf{\Gamma}(\lambda)^{1/2}\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1} \tag{3.16}$$

where $\mathbf{U}_j(\lambda)$ is the jth latent vector of $\mathbf{\Gamma}(\lambda)^{1/2}\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)\mathbf{\Gamma}(\lambda)^{1/2}$. The choice $\mathbf{\Gamma}(\lambda) = \mathbf{f}_{YY}(\lambda)^{-1}$ gives an invariant procedure that is a generalization of Hotelling's canonical correlation analysis.

In certain cases one can obtain expressions for $\hat{\mathbf{Y}}(t)$ of (3.6) in terms of $\ldots, \mathbf{X}(-1) - \mathbf{c}_X, \mathbf{X}(0) - c_X, \mathbf{X}(1) - \mathbf{c}_X, \ldots$. Suppose that $\Phi(\lambda)$ of Theorem 3.1 is Lebesgue measure on $(-\pi, \pi)$ and that the entries of $\mathbf{f}_{XX}(\lambda)$ are bounded. In this case, if $\mathbf{A}(\lambda) \in L_2^{s\times r}(\mathbf{F}_{XX})$, it results that

$$\int_{-\pi}^{\pi} \operatorname{tr}\{\mathbf{A}(\lambda)\overline{\mathbf{A}(\lambda)}^\tau\}\, d\lambda < \infty \tag{3.17}$$

and so for $\mathbf{a}(u)$ given by

$$(2\pi)^{-1} \int_{-\pi}^{\pi} \exp\{i\lambda u\}\mathbf{A}(\lambda)\, d\lambda, \tag{3.18}$$

one has

$$\sum_{u=-\infty}^{\infty} \operatorname{tr}\{\mathbf{a}(u)\mathbf{a}(u)^\tau\} < \infty. \tag{3.19}$$

We therefore have:

Theorem 3.2. *Under the conditions of Theorem* 3.1, *with* $\Phi(\lambda)$ *Lebesgue measure and the entries of* $\mathbf{f}_{XX}(\lambda)$ *bounded,*

$$\hat{\mathbf{Y}}(t) = \sum_{u=-\infty}^{\infty} \mathbf{a}(t-u)\{\mathbf{X}(u) - \mathbf{c}_X\} \tag{3.20}$$

where $\mathbf{a}(u)$ *is given by* (3.18) *and satisfies* (3.19).

Unfortunately the convolution form of $\mathbf{a}(u)$, suggested by (3.1) and (3.2), is apparently not deducible under the conditions of the theorem. It does follow under alternate assumptions. Let $\alpha(u)$ be a positive function defined for $u = 0, \pm 1, \ldots$, and such that

$$\alpha(u+v) \le K\alpha(u)\alpha(v) \tag{3.21}$$

for some K, and $u, v = 0, \pm 1, \ldots$. (An example of such a function is $\alpha(u) = 1 + |u|^l, l \ge 0$.) We have:

Theorem 3.3. *Suppose the conditions of Theorem* 3.2 *are satisfied. Suppose also that the entries of* $\sum_{u=-\infty}^{\infty} \alpha(u)\,|\mathbf{c}_{WW}(u)|$ *are bounded and that the latent roots of* $\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)$ *are simple for all* λ. *Then*

$$\hat{\mathbf{Y}}(t) = \sum_{u=-\infty}^{\infty} \mathbf{d} * \mathbf{b}(t-u)\{\mathbf{X}(u) - \mathbf{c}_X\} \tag{3.22}$$

with $\mathbf{d}(u)$ $s \times q$, $\mathbf{b}(u)$ $q \times r$ *and*

$$\sum_{u=-\infty}^{\infty} \alpha(u)\,|\mathbf{d}(u)| \tag{3.23}$$

$$\sum_{u=-\infty}^{\infty} \alpha(u)\,|\mathbf{b}(u)| \tag{3.24}$$

both being finite.

We see that the manner in which $\mathbf{d}(u)$ and $\mathbf{b}(u)$ fall off as $|u| \to \infty$ is directly related to the falloff of $\mathbf{c}_{WW}(u)$. In fact if $\mathbf{W}(t)$ is an m dependent process, we see that $\mathbf{d}(u), \mathbf{b}(u) = \mathbf{0}$ for $|u| > m$.

Other examples of canonical variates, for time series, appear in the work of Hannan [9] and Yaglom [28]. Further results of the type in this section are given by Brillinger [2]. Principal component analyses, in the time domain, occur in the work of Stone [24] and Craddock [5].

4. STATISTICAL PROPERTIES

Suppose one has available a stretch, $\mathbf{X}(t)$, $t = 0, \ldots, T-1$ of the r vector-valued stationary series $\mathbf{X}(t)$, and one wishes to construct estimates of the $\mu_j(\lambda)$, $\mathbf{V}_j(\lambda)$, $j = 1, \ldots, r$, and $\mathbf{A}(\lambda)$, for a particular q, of Corollary 3.1.2. One means of doing this is to construct an estimate, $\mathbf{f}_{XX}^{(T)}(\lambda)$, of $\mathbf{f}_{XX}(\lambda)$, and then to determine the latent values, $\mu_j^{(T)}(\lambda)$ and latent vectors, $\mathbf{V}_j^{(T)}(\lambda)$, of this matrix. We turn to a variety of statistical properties of a class of such estimates:

Let

$$\mathbf{d}_X^{(T)}(\lambda) = \sum_{t=0}^{T-1} \exp\{-i\lambda t\}\mathbf{X}(t) \tag{4.1}$$

$$\mathbf{I}_{XX}^{(T)}(\lambda) = (2\pi T)^{-1}\, \mathbf{d}_X^{(T)}(\lambda)\, \overline{\mathbf{d}_X^{(T)}(\lambda)}^{\tau} \tag{4.2}$$

$$\mathbf{f}_{XX}^{(T)}(\lambda) = 2\pi T^{-1} \sum_{s=1}^{T-1} H^{(T)}[\lambda - (2\pi s/T)]\mathbf{I}_{XX}^{(T)}(2\pi s/T) \tag{4.3}$$

$-\infty < \lambda < \infty$. In (4.3), $H^{(T)}(\alpha)$ is determined as follows: There is an $H(\alpha)$ satisfying:

Assumption 4.1. *$H(\alpha)$, $-\infty < \alpha < \infty$, is real-valued, continuously differentiable, $H(\alpha) = H(-\alpha)$, $\int_{-\infty}^{\infty} H(\alpha)\, d\alpha = 1$, and there exist finite $K, \varepsilon > 0$ such that*

$$|\alpha H(\alpha)|, \qquad \left|\frac{dH(\alpha)}{d\alpha}\right| \le K(1+|\alpha|)^{-1-\varepsilon}. \tag{4.4}$$

For $B_T > 0$, one then sets

$$H^{(T)}(\alpha) = B_T^{-1} \sum_{j=-\infty}^{\infty} H(B_T^{-1}[\alpha + 2\pi j]). \tag{4.5}$$

Omitting the term with $s = 0$, from (4.3), has the effect of removing the sample mean from the data. Before turning to asymptotic properties of the latent roots and vectors of $\mathbf{f}_{XX}^{(T)}(\lambda)$, we set down:

Assumption 4.2. *$\mathbf{X}(t) = [X_j(t)]$, $t = 0, \pm 1, \ldots$, is a strictly stationary, r vector-valued series, all of whose moments exist and with*

$$c_{a_1,\ldots,a_k}(u_1, \ldots, u_{k-1}) = \operatorname{cum}\{X_{a_1}(t+u_1), \ldots X_{a_{k-1}}(t+u_{k-1}), X_{a_k}(t)\} \tag{4.6}$$

$$\sum_{u_1,\ldots,u_{k-1}=-\infty}^{\infty} |c_{a_1,\ldots,a_k}(u_1, \ldots, u_{k-1})| < \infty \tag{4.7}$$

for $a_1, \ldots, a_k = 1, \ldots, r$, and $k = 2, 3, \ldots$.

This kind of assumption is made by Brillinger and Rosenblatt [3].

We will sometimes require:

Assumption 4.3 (*P*). *$\mathbf{X}(t)$ satisfies Assumption 4.2 and for $P > 0$*

$$\sum_{u=-\infty}^{\infty} |u|^P |c_{a_1,a_2}(u)| < \infty \tag{4.8}$$

$a_1, a_2 = 1, \ldots, r$.

Denoting the latent roots and vectors of

$$\int_0^{2\pi} H^{(T)}(\lambda - \alpha)\mathbf{f}_{XX}(\alpha)\, d\alpha \tag{4.9}$$

by $\nu_j^{(T)}(\lambda)$, $\mathbf{W}_j^{(T)}(\lambda)$, $j = 1, \ldots, r$, we have:

Theorem 4.1. *Let $\mathbf{X}(t)$ be an r vector-valued series satisfying Assumption 4.3(1). Let $\mu_j^{(T)}(\lambda)$, $\mathbf{V}_j^{(T)}(\lambda)$, $j = 1, \ldots, r$ be the latent roots and vectors of $\mathbf{f}_{XX}^{(T)}(\lambda)$ where $\mathbf{f}_{XX}^{(T)}(\lambda)$ has been constructed in the manner of (4.3) with $H(\alpha)$ satisfying Assumption 4.1. If $B_T \to 0$, $B_T T \to \infty$ as $T \to \infty$, then*

$$E\mu_j^{(T)}(\lambda) = \nu_j^{(T)}(\lambda) + O(B_T T)^{-1/2}. \tag{4.10}$$

If the latent roots of $\mathbf{f}_{XX}(\lambda)$ are simple, then

$$\overrightarrow{\text{ave}}\, \mu_j^{(T)}(\lambda) = \nu_j^{(T)}(\lambda) + O(B_T T)^{-1} \tag{4.11}$$

$$\overrightarrow{\text{ave}}\, \mathbf{V}_j^{(T)}(\lambda) = \mathbf{W}_j^{(T)}(\lambda) + O(B_T T)^{-1}. \tag{4.12}$$

(In this theorem, $\overrightarrow{\text{ave}}$ has a technical definition allowing the use of the Δ-method. See Brillinger and Tukey [4].)

We have:

Corollary 4.1.1. *Under the conditions of the theorem,*

$$E\mu_j^{(T)}(\lambda), \overrightarrow{\text{ave}}\, \mu_j^{(T)}(\lambda) \to \mu_j(\lambda) \qquad \text{and} \qquad \overrightarrow{\text{ave}}\, \mathbf{V}_j^{(T)}(\lambda) \to \mathbf{V}_j(\lambda).$$

In connection with $\nu_j^{(T)}(\lambda)$, $\mathbf{W}_j^{(T)}(\lambda)$, we have:

Theorem 4.2. *Let $H(\alpha)$ satisfy Assumption 4.1 and suppose $H(\alpha) \ge 0$, then*

$$0 \le \nu_j^{(T)}(\lambda) \le \int_0^{2\pi} H^{(T)}(\lambda - \alpha)\mu_j(\alpha)\, d\alpha. \tag{4.13}$$

Let $H(\alpha)$ satisfy Assumption 4.1, and let

$$H_2 = \int_{-\infty}^{\infty} \alpha^2 H(\alpha)\, d\alpha \tag{4.14}$$

with $\int_{-\infty}^{\infty} |\alpha|^3 H(\alpha)\, d\alpha < \infty$. Let (4.8) be satisfied with $P = 3$. Suppose the $\mu_j(\lambda)$, $j = 1, \ldots, r$ are distinct, then if $B_T \to 0$ as $T \to \infty$,

$$\nu_j^{(T)}(\lambda) = \mu_j(\lambda) + \tfrac{1}{2}B_T^2 H_2\, \overline{\mathbf{V}_j(\lambda)}^{\tau} \frac{d^2\mathbf{f}_{XX}(\lambda)}{d\lambda^2} \mathbf{V}_j(\lambda) + O(B_T^3) \tag{4.15}$$

and

$$\mathbf{W}_j^{(T)}(\lambda) = \mathbf{V}_j(\lambda) + \tfrac{1}{2}B_T^2 H_2 \sum_{k \ne j} \{\overline{\mathbf{V}_k(\lambda)}^{\tau} \frac{d^2\mathbf{f}_{XX}(\lambda)}{d\lambda^2} \mathbf{V}_j(\lambda)\}\mathbf{V}_k(\lambda)/\{\mu_j(\lambda) - \mu_k(\lambda)\} + O(B_T^3). \tag{4.16}$$

This theorem, together with Theorem 4.1, indicates that the asymptotic biases of the $\mu_j^{(T)}(\lambda)$, $\mathbf{V}_j^{(T)}(\lambda)$ depend directly on the bandwidth, B_T, employed and the smoothness, with respect to λ, of the entries of $\mathbf{f}_{XX}(\lambda)$. The importance of prefiltering is apparent. Continuing we have:

Theorem 4.3. *Let* $\mathbf{X}(t)$ *be r vector-valued, satisfying Assumption 4.2, and such that the latent roots of* $\mathbf{f}_{XX}(\lambda)$ *are simple for* $\lambda = \lambda_1, \lambda_2$. *Let* $\mathbf{f}_{XX}^{(T)}(\lambda)$ *be constructed in the manner of (4.3) with* $H(\alpha)$ *satisfying Assumption 4.1. If* $B_T \to 0$, $B_T T \to \infty$ *as* $T \to \infty$, *then the variate*

$$\{\mu_j^{(T)}(\lambda_1), \mathbf{V}_j^{(T)}(\lambda_1), \mu_j^{(T)}(\lambda_2), \mathbf{V}_j^{(T)}(\lambda_2);\ \ j = 1, \ldots, r\}$$

is asymptotically Gaussian with

$$\{\mu_j^{(T)}(\lambda_1), \mu_j^{(T)}(\lambda_2);\ \ j = 1, \ldots, r\}$$

asymptotically independent of

$$\{\mathbf{V}_j^{(T)}(\lambda_1), \mathbf{V}_j^{(T)}(\lambda_2);\ \ j = 1, \ldots, r\}.$$

The asymptotic covariance structure is given by

$$\lim_{T\to\infty} B_T T \overrightarrow{\operatorname{cov}} \{\mu_j^{(T)}(\lambda_1), \mu_k^{(T)}(\lambda_2)\}$$
$$= \delta_{jk} 2\pi\left[\int H^2(\alpha)\,d\alpha\right][\eta\{\lambda_1 - \lambda_2\} + \eta\{\lambda_1 + \lambda_2\}][\mu_j(\lambda_1)]^2 \tag{4.17}$$

$$\lim_{T\to\infty} B_T T \overrightarrow{\operatorname{cov}} \{V_{pj}^{(T)}(\lambda_1), V_{qk}^{(T)}(\lambda_2)\}$$
$$= 2\pi\left[\int H^2(\alpha)\,d\alpha\right]\left[\eta\{\lambda_1 - \lambda_2\}\delta_{jk}\mu_j(\lambda_1)\right.$$
$$\cdot\left\{\sum_{l\neq j} \mu_l(\lambda_1)(\mu_j(\lambda_1) - \mu_l(\lambda_1))^{-2}\overline{V}_{pl}(\lambda_1)V_{ql}(\lambda_1)\right\} + \eta\{\lambda_1 - \lambda_2\}$$
$$\left.\cdot\{1 - \delta_{jk}\}\mu_k(\lambda_1)\mu_j(\lambda_1)(\mu_j(\lambda_1) - \mu_k(\lambda_1))^{-2}\overline{V}_{pj}(\lambda_1)V_{qk}(\lambda_1)\right]. \tag{4.18}$$

(In (4.17), (4.18); $\delta_{jk} = 1$ if $j = k$ and $=0$ otherwise, $\eta\{\lambda\} = 1$ if $\lambda \equiv 0$ (mod 2π) and $=0$ otherwise, the covariance of complex-valued random variables X_1, X_2 is taken to be $E\{X_1 - EX_1\}\{\overline{X_2 - EX_2}\}$.)

We see that (4.17) implies

$$\lim_{T\to\infty} B_T T \overrightarrow{\operatorname{var}} \log \mu_j^{(T)}(\lambda) = [1 + \eta\{2\lambda\}]2\pi \int H^2(\alpha)\,d\alpha. \tag{4.19}$$

This is the same result as that obtained for the asymptotic variance of a power spectral estimate (see Parzen [16], for example).

On occasion, an alternate form of asymptotic distribution may prove relevant. Suppose we estimate $\mathbf{f}_{XX}(\lambda)$ by a simple average of periodograms. For example, with $s(T)$, m integers and $2\pi s(T)/T$ near λ, consider

$$(2m+1)^{-1} \sum_{s=-m}^{m} \mathbf{I}_{XX}^{(T)}(2\pi[s(T) + s]/T) \tag{4.20}$$

if $\lambda \not\equiv 0$ (mod π). Consider

$$(2m+2)^{-1}\left[\mathbf{I}_{XX}^{(T)}(\pi) + \sum_{s=-m}^{m} \mathbf{I}_{XX}^{(T)}(\pi + 2\pi s/T)\right] \tag{4.21}$$

if $\lambda = \pm\pi, \pm 3\pi, \ldots,$ and consider

$$(2m)^{-1}\left\{\sum_{s=-m}^{-1} + \sum_{s=1}^{m}\right\}\mathbf{I}_{XX}^{(T)}(2\pi s/T) \tag{4.22}$$

if $\lambda \equiv 0$ (mod 2π). We have:

Theorem 4.4. *Let* $\mathbf{X}(t)$ *be an r vector-valued series satisfying Assumption 4.2. Let m be fixed and* $2\pi s(T)/T \to \lambda$ *as* $T \to \infty$. *If* $\lambda \not\equiv 0$ (mod π), (4.20) *tends, in distribution, to* $(2m+1)^{-1}W_r^C(2m+1, \mathbf{f}_{XX}(\lambda))$. *If* $\lambda = \pm\pi, \pm 3\pi, \ldots,$ (4.21) *tends to* $(2m+1)^{-1}W_r(2m+1, \mathbf{f}_{XX}(\lambda))$. *If* $\lambda = 0, \pm 2\pi, \pm 4\pi, \ldots,$ (4.22) *tends to* $(2m)^{-1}W_r(2m, \mathbf{f}_{XX}(\lambda))$.

(In this theorem $W_r(\nu, \mathbf{\Sigma})$ denotes a real Wishart variable of dimension r, degrees of freedom ν and var-covar matrix $\mathbf{\Sigma}$. See Rao [20]. Whereas $W_r^C(\nu, \mathbf{\Sigma})$ denotes a complex Wishart variable of dimension r, degrees of freedom ν and var-covar matrix $\mathbf{\Sigma}$. See Goodman [8].)

Corollary 4.4.1. *Under the conditions of the theorem, the latent roots and vectors of* (4.20)–(4.22) *tend, in distribution, to the latent roots and vectors of the limiting Wishart distributions of the theorem.*

The distributions of the latent roots of matrices, with real and complex Wishart distributions, are given by James [10].

The estimates (4.20) and (4.21) correspond, approximately, to taking $H(\alpha) = (2\pi)^{-1}$ for $|\alpha| \le \pi$ and 0 otherwise in (4.3), implying $m \doteq B_T T/2$. This leads us to approximate the distribution of $\mathbf{f}_{XX}^{(T)}(\lambda)$, constructed in the manner of (4.3) with general $H(\alpha)$, by a $\nu^{-1}W_r^C(\nu, \mathbf{f}_{XX}(\lambda))$ or $\nu^{-1}W_r(\nu, \mathbf{f}_{XX}(\lambda))$ variate having

$$\nu = B_T T \Big/ \left\{2\pi \int H^2(\alpha)\,d\alpha\right\}. \tag{4.23}$$

This was suggested by Goodman [8]. One can then approximate the distributions of the $\mu_j^{(T)}(\lambda)$, $\mathbf{V}_j^{(T)}(\lambda)$ by the distributions of corresponding variates based on a Wishart distribution. Notice that as $m \to \infty$, one is led back to the approximation of Theorem 4.3.

If one takes $m = T$ in (4.22), that is, one is smoothing the matrix of periodograms across the whole frequency domain, then (4.22) becomes $(2\pi)^{-1}\mathbf{c}_{XX}^{(T)}(0)$. The suggested analysis is seen to reduce to a standard principal component analysis of the estimated variance-covariance matrix of the variate $\mathbf{X}(t)$.

Analogs of Theorems 4.1 to 4.4 exist for the variates of Theorem 3.1. These are given by Brillinger [2]. There are also direct extensions to the case of a continuous time parameter.

5. A WORKED EXAMPLE

In this section, we provide a preliminary report of the empirical analysis of a vector-valued series along the lines suggested by Corollary 3.1.2 and Section 4. The components of the series analyzed are the monthly mean temperatures, in °C, recorded at fourteen European stations over a period of approximately 170 years. The stations and the periods of available data are listed in Table I.

TABLE I

City	Period available	City	Period available
Basle	1755–1957	Greenwich	1763–1962
Berlin	1769–1950	New Haven	1780–1950
Breslau	1792–1950	Prague	1775–1939
Budapest	1780–1947	Stockholm	1756–1960
Copenhagen	1798–1950	Trondheim	1761–1946
De Bilt	1711–1960	Vienna	1775–1950
Edinburgh	1764–1959	Vilna	1781–1938

In the analysis, the series were first seasonally adjusted, then $\mathbf{d}^{(T)}(\lambda)$ of (4.1) was calculated for $\lambda = 2\pi k/2048$, $k = 0, \ldots, 2047$ using a fast Fourier transform algorithm. These values were then used to estimate $\mathbf{f}_{XX}(\lambda)$, $\lambda = \pi k/64$, $k = 0, \ldots, 64$ in the manner of (4.20) and (4.21) with $m = 57$. The logarithms (base 10) of the resulting power spectral estimates are given in Fig. 1 for the various components. The bandwidth of these estimates $\doteq 0.11\pi$ and the asymptotic standard error $\doteq 0.09$ if $\lambda \neq 0, \pi$. At these last points it is 0.18.

Next the latent values, $\mu_j^{(T)}(\lambda)$, and latent vectors, $\mathbf{V}_j^{(T)}(\lambda)$, of the estimated spectral density matrix, $\mathbf{f}_{XX}^{(T)}(\lambda)$, were calculated for $\lambda = \pi k/64$, $k = 0, \ldots, 64$, and $j = 1, \ldots, 14$. This was done by associating a real matrix with $\mathbf{f}_{XX}^{(T)}(\lambda)$ in the manner of Wilkinson [27, p. 174]. $\text{Log}_{10}\, \mu_j^{(T)}(\lambda)$ is plotted in Fig. 2 for $j = 1, \ldots, 14$. The asymptotic standard error of these estimates, following (4.19), is approximately 0.09 if $\lambda \neq 0, \pi$. At these last points it is 0.18.

A comparison of Figs. 1 and 2 indicates that adjacent estimates of the latent roots are apparently less correlated than are the corresponding power spectral estimates. This and (4.19) suggest that one smooth the plots of the logarithms.

Plots of the latent vectors are given by Brillinger [2]. A preliminary examination of the first latent vector suggests that it corresponds, in large part, with a simple average of the fourteen series.

6. OPEN QUESTIONS

A number of interesting problems remain. How are the results of this article affected if one requires that the filters employed be realizable? In this case $J_X{}^q$ of Section 2 would be replaced by $J_X{}^q(t)$, the closure of s vectors of the form

$$\sum_{u \le t} \mathbf{d} * \mathbf{b}(u)\{\mathbf{X}(u) - \mathbf{c}_X\}, \tag{6.1}$$

and one would seek $\hat{\mathbf{Y}}(t)$ in this space. If one wishes a realtime procedure, as in the construction of a vocoder, or if one is interested in prediction, these results are important.

No one appears to have derived the exact distribution of the latent vectors of a real or complex Wishart matrix. James [10] has derived the distribution of the latent roots.

Finally, the discovery of techniques of rotating the factor series,

$$\int_{-\pi}^{\pi} \exp\{i\lambda t\}\overline{\mathbf{V}_j(\lambda)}^\tau\, d\mathbf{Z}_X(\lambda), \qquad j = 1, \ldots, q$$

of Corollary 3.1.2, to ease their interpretation, seems especially important. The difficulty of direct interpretation is augmented by the fact that if $\mathbf{V}_j(\lambda)$ is a latent vector of $\mathbf{f}_{XX}(\lambda)$, then so is $\pm i\mathbf{V}_j(\lambda)$.

7. PROOFS

Generally the proofs are only sketched. More extensive results are given by Brillinger [2].

Theorem 2.2. Equation (2.15) has the form (2.16) with $\mathbf{A}(\lambda) = \mathbf{D}(\lambda)\mathbf{B}(\lambda)$, $\mathbf{D}(\lambda)$, $\mathbf{B}(\lambda)$ being the Fourier transforms of $\mathbf{d}(u)$, $\mathbf{b}(u)$. This $\mathbf{A}(\lambda)$ has rank $\le q$ and so $\mathbf{A}(\lambda)\overline{\mathbf{A}(\lambda)}^\tau$ has at most q nonzero latent roots. Consider a convergent sequence of $\mathbf{A}(\lambda)$'s in $L_2^{s \times r}(F_{XX})$. (Rosenberg [22] proved that this space is complete.) The latent roots are continuous functions of the entries of a matrix and so the limit of the sequence of $\mathbf{A}(\lambda)$'s can have at most q nonzero latent roots. The result follows.

Theorem 3.1.

$$E\{\mathbf{Y}(t) - \mathbf{m} - \hat{\mathbf{Y}}(t)\}^\tau\{\mathbf{Y}(t) - \mathbf{m} - \hat{\mathbf{Y}}(t)\}$$
$$= E\{\mathbf{Y}(t) - \mathbf{c}_Y - \hat{\mathbf{Y}}(t)\}^\tau\{\mathbf{Y}(t) - \mathbf{c}_Y - \hat{\mathbf{Y}}(t)\} + E\{\mathbf{c}_Y - \mathbf{m}\}^\tau\{\mathbf{c}_Y - \mathbf{m}\}$$

and (3.5) follows. Now

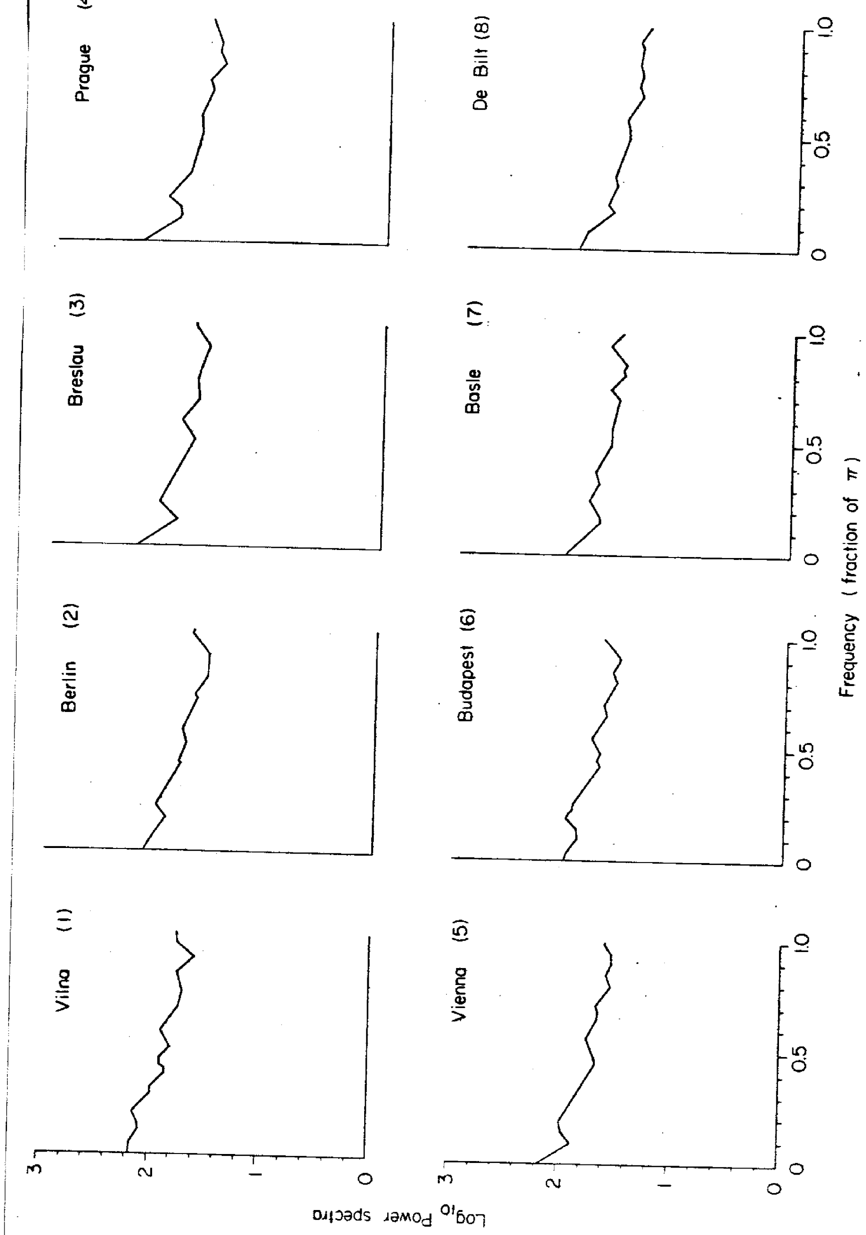

Fig. 1.

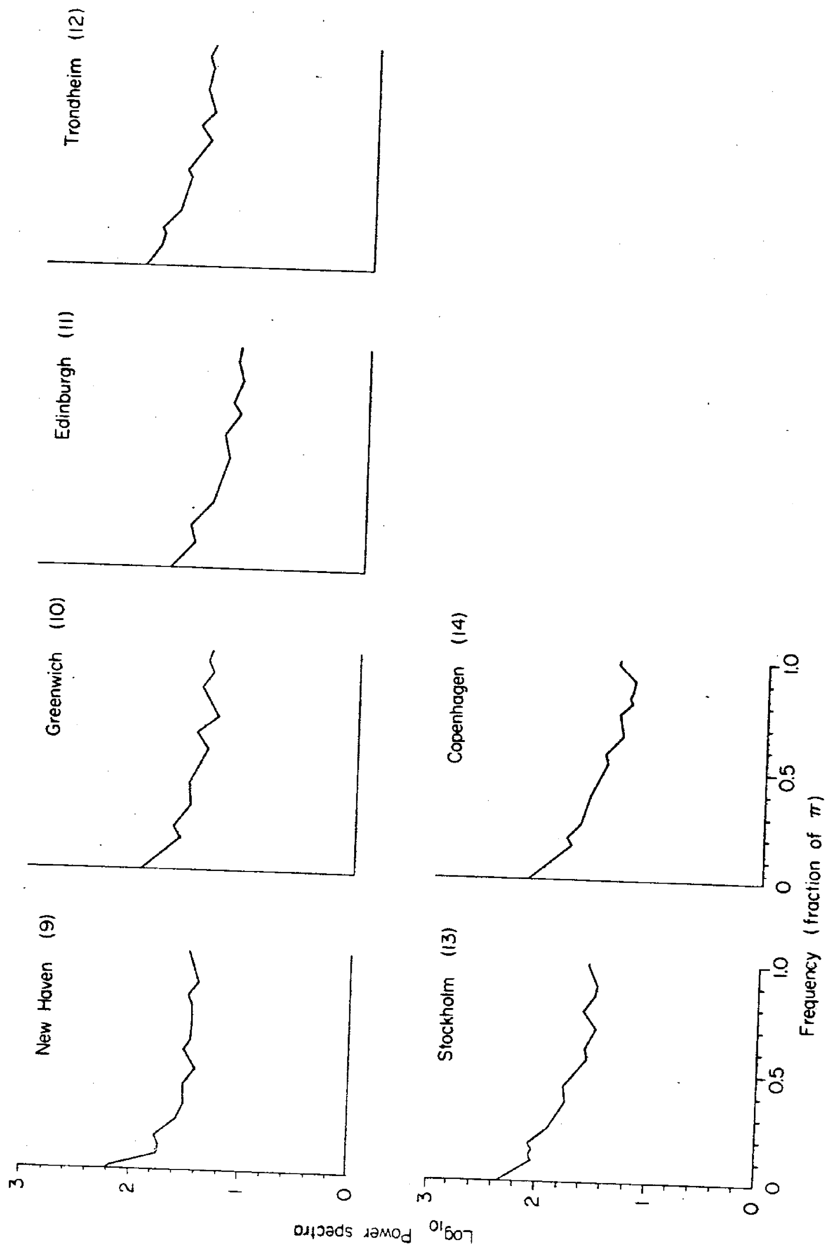

Fig. 1 (continued).

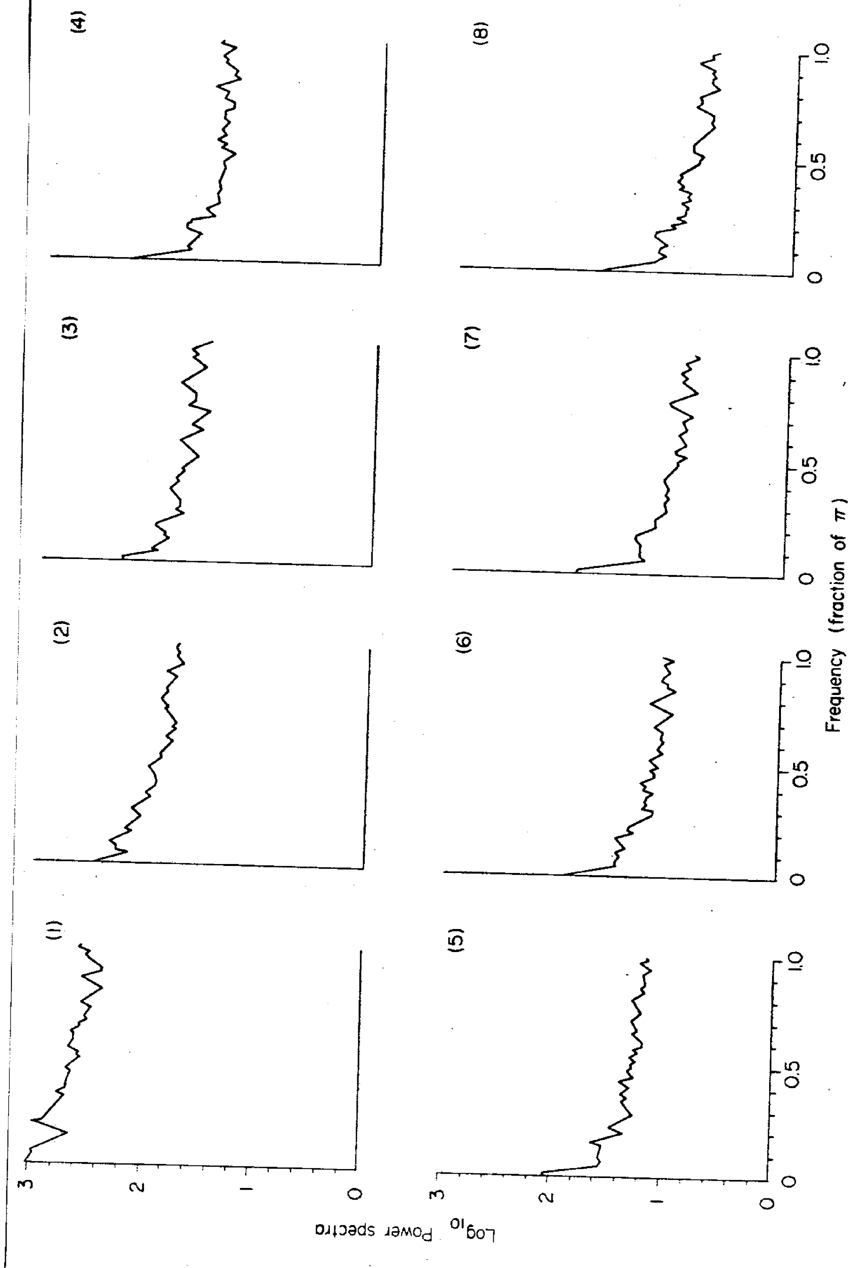

Fig. 2.

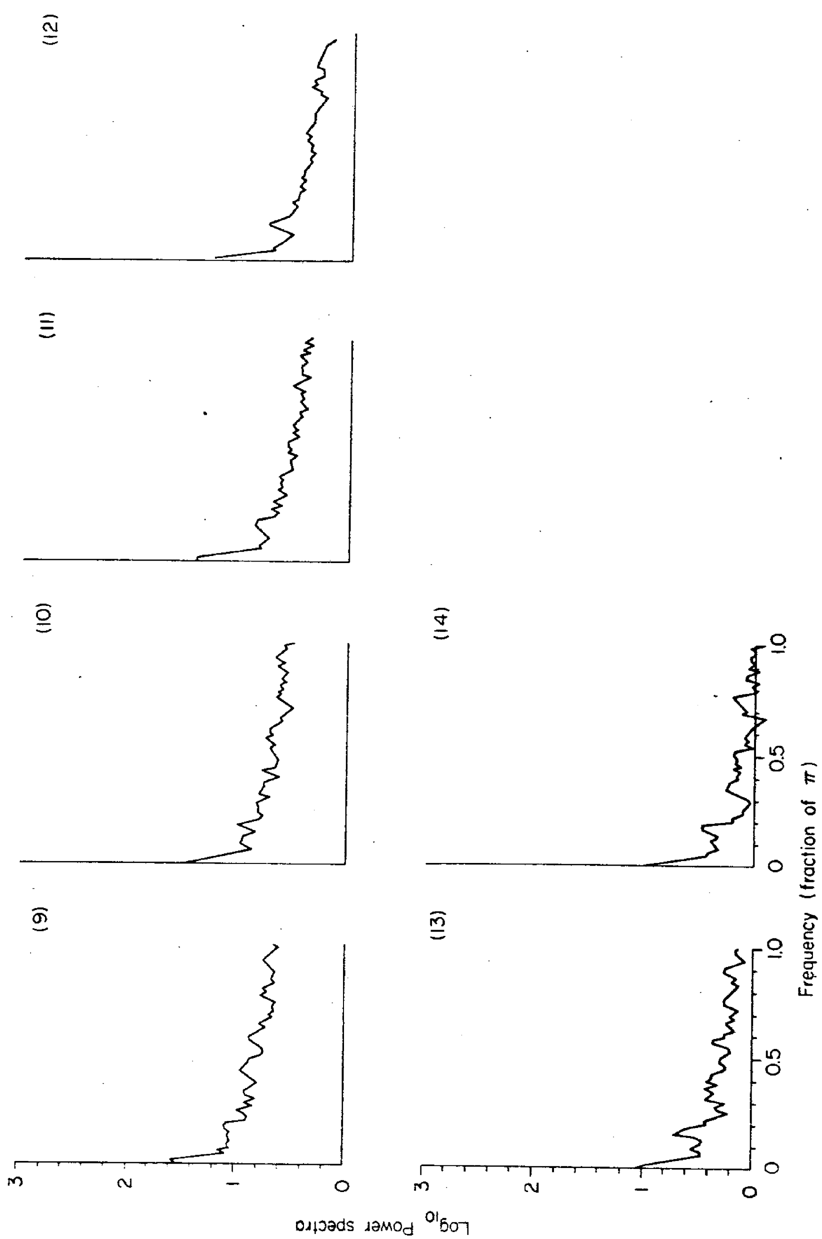

Fig. 2 (continued).

$$E\{\mathbf{Y}(t)-\mathbf{c}_Y-\hat{\mathbf{Y}}(t)\}^\tau\{\mathbf{Y}(t)-\mathbf{c}_Y-\hat{\mathbf{Y}}(t)\}$$
$$=\int_{-\pi}^{\pi}\operatorname{tr}\{\mathbf{f}_{YY}(\lambda)-\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\mathbf{f}_{XY}(\lambda)\}\,d\Phi(\lambda)$$
$$+\int_{-\pi}^{\pi}\operatorname{tr}\{\mathbf{A}(\lambda)-\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}\}$$
$$\times\,\mathbf{f}_{XX}(\lambda)\{\overline{\mathbf{A}(\lambda)-f_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}}\}^\tau\,d\Phi(\lambda).$$

As $\mathbf{A}(\lambda)$ has rank $\leq q$, the question becomes one of approximating one matrix by another of smaller rank. The indicated result follows from a complex generalization of a theorem of Eckart and Young [6]. Related real variate results may be found in Rao [20, 21].

Theorem 3.3. We begin by noting that the latent vectors of a matrix with simple latent roots are real-analytic functions of the entries. (This follows from Portman [19].) Next we take

$$\mathbf{D}(\lambda)=[\mathbf{V}_1(\lambda),\ldots,\mathbf{V}_q(\lambda)],\qquad \mathbf{B}(\lambda)=\overline{\mathbf{D}(\lambda)}^\tau\mathbf{f}_{YX}(\lambda)\mathbf{f}_{XX}(\lambda)^{-1}.$$

Both of these are real-analytic functions of the elements of $\mathbf{f}_{WW}(\lambda)$. The theorem now follows from the Wiener-Levy Theorem on functions that operate on Fourier transforms (see Gelfand *et al.* [7]).

Theorem 4.1. From the Wielandt-Hoffman theorem (see Wilkinson [27]),

$$\sum_{j=1}^{r}\{\mu_j^{(T)}(\lambda)-\nu_j^{(T)}(\lambda)\}^2\leq\sum_{j,k=1}^{r}|f_{jk}^{(T)}(\lambda)-\int_0^{2\pi}H^{(T)}(\lambda-\alpha)f_{jk}(\alpha)\,d\alpha|^2.$$

Now

$$E|f_{jk}^{(T)}(\lambda)-\int_0^{2\pi}H^{(T)}(\lambda-\alpha)f_{jk}(\alpha)\,d\alpha|^2=O(B_T T)^{-1}$$

under the stated conditions (see Brillinger and Rosenblatt [3]) giving (4.10).

Following Wilkinson [27, p. 68], we have the following Taylor series expansions:

$$\mu_j^{(T)}(\lambda)=\nu_j^{(T)}(\lambda)+\overline{\mathbf{W}_j^{(T)}(\lambda)}^\tau\Big\{\mathbf{f}_{XX}^{(T)}(\lambda)-\int_0^{2\pi}H^{(T)}(\lambda-\alpha)$$
$$\cdot\,\mathbf{f}_{XX}(\alpha)\,d\alpha\Big\}\,\mathbf{W}_j^{(T)}(\lambda)+\text{higher order terms.}$$

and

$$\mathbf{V}_j^{(T)}(\lambda)=\mathbf{W}_j^{(T)}(\lambda)+\sum_{k\neq j}\Big[\overline{\mathbf{W}_k(\lambda)}^\tau\Big\{\mathbf{f}_{XX}^{(T)}(\lambda)-\int_0^{2\pi}H^{(T)}(\lambda-\alpha)$$
$$\cdot\,\mathbf{f}_{XX}(\alpha)d\alpha\Big\}\mathbf{W}_j(\lambda)\Big]\mathbf{W}_k(\lambda)/\{\nu_j^{(T)}(\lambda)-\nu_k^{(T)}(\lambda)\}+\text{higher order terms.}$$

Under the stated conditions, Eqs. (4.11) and (4.12) follow by taking expected values in these expressions (see Brillinger and Tukey [4]).

Theorem 4.2. Equations (4.15) and (4.16) follow from perturbation expansions of the type used in Theorem 4.1. Equation (4.16) follows from Weyl's characterization of latent values.

Theorem 4.3. It follows from the results of Brillinger and Rosenblatt [3] that the entries of $\mathbf{f}_{XX}^{(T)}(\lambda_1)$, $\mathbf{f}_{XX}^{(T)}(\lambda_2)$ are asymptotically jointly normal with covariance structure given by

$$\lim_{T\to\infty}B_T\,T\,\operatorname{cov}\{f_{a_1a_2}^{(T)}(\lambda_1),f_{b_1b_2}^{(T)}(\lambda_2)\}$$
$$=2\pi\int H(\alpha)^2\,d\alpha[\eta\{\lambda_1-\lambda_2\}f_{a_1b_1}(\lambda_1)f_{a_2b_2}(-\lambda_1)+\eta\{\lambda_1+\lambda_2\}$$
$$\cdot f_{a_1b_2}(\lambda_1)f_{a_2b_1}(-\lambda_1)],$$

for

$$a_1,a_2,b_1,b_2=1,\ldots,r.$$

The result of the theorem now follows from the perturbation exapansions of Theorem 4.1 and a theorem of Mann and Wald [14].

Theorem 4.4. This is proved by Brillinger [2]. The corollary follows from a theorem of Mann and Wald [14].

REFERENCES

1. Brillinger, D. R. (1964). The generalization of the techniques of factor analysis, canonical correlation and principal components to stationary time series. *Roy. Statist. Soc. Conf. Cardiff, Wales*, 1964.
2. Brillinger, D. R. (1970). "An Introduction to the Frequency Analysis of Vector-Valued Time Series." Holt, New York. To be published.
3. Brillinger, D. R., and Rosenblatt, M. (1967). Asymptotic theory of kth order spectra. *In* "Spectral Analysis of Time Series" (B. Harris, ed.), pp. 153–188. Wiley, New York.
4. Brillinger, D. R., and Tukey, J. W. (1964). "Asymptotic Variances, Moments, Cumulants and Other Average Values." Unpublished manuscript.
5. Craddock, J. M. (1965). The analysis of meteorological time series for use in forecasting. *Statistician* **15** 167–190.
6. Eckhart, C., and Young, G. (1936). On the approximation of one matrix by another of lower rank. *Psychometrika* **1** 211–218.
7. Gelfand, I., Raikov, D., and Shilov, G. (1964). "Commutative Normed Rings." Chelsea, New York.
8. Goodman, N. R. (1963). Statistical analysis based upon a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Statist.* **34** 152–177.
9. Hannan, E. J. (1961). The general theory of canonical correlation and its relation to functional analysis. *J. Austral. Math. Soc.* **2** 229–242.

10. JAMES, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.
11. KOOPMANS, L. H. (1964). On the multivariate analysis of weakly stationary stochastic processes. *Ann. Math. Statist.* **35** 1765–1780.
12. KRAMER, H. P., and MATTHEWS, M. V. (1956). A linear coding for transmitting a set of correlated signals. *IRE Trans. Information Theory* **2** 41–46.
13. LOYNES, R. M. (1965). On a generalization of second order stationarity. *Proc. London Math. Soc.* **15** 385–398.
14. MANN, H. B., and WALD, A. (1943). On stochastic limit and order relationships. *Ann. Math. Statist.* **14** 217–226
15. MASANI, P. (1966). Recent trends in multivariate prediction theory. *In* "Multivariate Analysis" (P. R. Krishnaiah, ed.), pp. 351–382. Academic Press, New York.
16. PARZEN, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.* **28** 329–348.
17. PARZEN, E. (1966). On empirical multiple time series analysis. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* (L. LeCam and J. Neyman, eds.), pp. 305–340. Univ. of California Press, Berkeley, California.
18. PINSKER, M. S. (1964). "Information and Information Stability of Random Variables and Processes." Holden-Day, San Francisco, California.
19. PORTMAN, W. O. (1960). Hausdorff-analytic functions of matrices. *Proc. Amer. Math. Soc.* **11** 97–101.
20. RAO, C. R. (1965). "Linear Statistical Inference and its Applications." Wiley, New York.
21. RAO, C. R. (1965). The use and interpretation of principal component analysis in applied research. *Sankhyā Ser. A* **26** 329–358.
22. ROSENBERG, M. (1964). The square-integrability of matrix-valued functions with respect to nonnegative hermitian measure. *Duke Math. J.* **31** 291–298.
23. ROZANOV, YU. A. (1967). "Stationary Random Processes." Holden-Day, San Francisco, California.
24. STONE, R. (1947). On the interdependence of blocks of transactions. *J. Roy. Statist. Soc. Ser. B* **9** 1–32.
25. WHITTLE, P. (1953). The analysis of multiple stationary time series. *J. Roy. Statist. Soc. Ser. B* **15** 125–139.
26. WIENER, N. (1930). Generalized harmonic analysis. *Acta Math.* **55** 117–258.
27. WILKINSON, J. H. (1965). "The Algebraic Eigenvalue Problem." Oxford Univ. Press (Clarendon), London and New York.
28. YAGLOM, A. M. (1965). Stationary Gaussian processes satisfying the strong mixing condition and best predictable functionals. *In* "Bernoulli, Bayes, Laplace" (J. Neyman and L. M. LeCam, eds.), pp. 241–252. Springer, New York.