

THE ESTIMATION OF COHERENCE

by

Nigel Nettheim

TECHNICAL REPORT NO. 5

May 16, 1966

PREPARED UNDER CONTRACT Nonr-225(80) (NR-042-234) FOR
OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

ACKNOWLEDGMENTS

I wish to express my gratitude to Professor Emanuel Parzen for his helpful supervision and for having created an atmosphere in which this thesis could be written. This work was partially supported by the Office of Naval Research under Contract Nonr 225(80) and was begun at the U. S. Bureau of the Census, Washington, D. C., for which opportunity I am indebted to Dr. Harry M. Rosenblatt. I also thank Miss Beverly Ralphs for her excellent typing.

TABLE OF CONTENTS

	<u>Page</u>
Acknowledgments	iii
Chapter 1. Theory and Assumptions for a Pair of Time Series	
1.1 Time Series	1
1.2 Stationary Time Series	2
1.3 Spectra	3
1.4 Coherence	6
1.5 Phase and Gain	10
1.6 The Aims of This Thesis	11
Chapter 2. A Critical Study of the Usual Estimation Procedures	
2.1 Estimation of Autospectra	12
(i) The periodogram	12
(ii) Windowed periodograms	13
(iii) Prewhitening	15
2.2 Estimation of Cross-spectra	17
(i) The cross-periodogram	17
(ii) Windowed cross-periodograms	18
(iii) Preflattening	19
2.3 Estimation of Coherence	20
Chapter 3. Estimation of Coherence by Cross-regressive Filtering and Related Methods	
3.1 Cross-regressive Filtering	23
3.2 Asymptotic Theory of Moments of Estimated Spectra	27
(i) Lemma on diagonal summation	27
(ii) Lemma on means of estimated spectra	28
(iii) General central moments of estimated spectra	30
3.3 Mean and Variance of the Estimates of Coherence	43
3.4 Comparison with Goodman's Distribution	50
3.5 Testing for Zero Coherence	54
3.6 Shifting the Lag Window by the Estimated Phase Derivative	58
3.7 Best Linear Pre-flattening	60
(i) A single transformation to serve for estimation at all frequencies	60
(ii) A separate transformation for each frequency	63
(iii) Separate transformations for the cospectrum and quadrature spectrum at each frequency	65
3.8 Alternative Estimating Functions for Coherence	67
3.9 Summary of Recommended Procedure	69
References	70

THE ESTIMATION OF COHERENCE

By

Nigel Nettheim

1. Theory and Assumptions for a Pair of Time Series

1.1 Time Series

This thesis deals with the problem of estimating certain properties of a pair of time series whose values at time t are given by the real random sequence

$$\{X_j(t), X_k(t)\}, \quad t = \dots, -1, 0, 1, \dots$$

from a sample of observations at times $t = 1, \dots, T$. We refer to t as time, although it could for instance be a spacial coordinate. The same symbols are used for random variables as for observed values since the context will indicate the meaning.

More general index sets could arise: for example, one might have T_j observations on $X_j(t)$ and T_k on $X_k(t)$; or one series might be quarterly while the other is monthly; or again the sample might be taken continuously. Such cases will not be considered here, although the population from which the sample is taken can always be regarded as defined in continuous time. It is hoped in the future to extend the present methods and results to problems involving three or more series and higher order spectra.

1.2 Stationary Time Series

The series considered in this work are always assumed to have expected value zero:

$$E[X_j(t)] \equiv 0 .$$

In most practical cases the mean value will be a non-zero function of time but it can usually be removed readily by methods, known as detrending, which have been discussed by Parzen (1964).

The series are further assumed to be jointly stationary in the sense that their covariances exist and are functions only of the time lag v :

$$E[X_j(t) X_k(t+v)] = R_{jk}(v), \quad v = \dots, -1, 0, 1, \dots .$$

When $j = k$ we speak of the autocovariance functions and when $j \neq k$ of the cross-covariance functions; note that $R_{jk}(v) = R_{kj}(-v)$. It is emphasized that stationarity of each individual series does not imply joint stationarity. The stationarity assumption has sometimes been relaxed, for instance by Priestley (1965). However, it is generally felt that stationary spectral methods are quite robust, and experience indicates that meaningful conclusions can be drawn even when they are applied to highly non-stationary series; in any case they are the natural first approach.

Moments of order higher than the second will also be assumed to exist when needed.

1.3 Spectra

It is well known (Yaglom, 1962, p. 55) that stationary covariance functions have a spectral representation:

$$R_{jk}(v) = \int_{-\pi}^{\pi} e^{iv\lambda} dF_{jk}(\lambda)$$

where the functions $F_{jk}(\lambda)$ are called spectral distribution functions.

Here we assume that each $F_{jk}(\lambda)$ is absolutely continuous (for which a sufficient condition is $\sum_{v=-\infty}^{\infty} |R_{jk}(v)| < \infty$) so that

$$R_{jk}(v) = \int_{-\pi}^{\pi} e^{iv\lambda} f_{jk}(\lambda) d\lambda .$$

The functions $f_{jk}(\lambda)$ are called autospectra when $j = k$ and cross-spectra otherwise; they are defined explicitly by

$$f_{jk}(\omega) = \frac{1}{2\pi} \sum_{v=-\infty}^{\infty} R_{jk}(v) e^{-iv\omega}, \quad -\pi \leq \omega \leq \pi .$$

These functions will always be assumed to be uniformly continuous. Note that

$$f_{jk}(\omega) = f_{kj}(-\omega) = \overline{f_{kj}(\omega)} ,$$

the bar denoting complex conjugation. Thus autospectra are always real, while cross-spectra are in general complex-valued. It is easy to show that the matrix $\underline{f}(\omega)$ of spectral densities is non-negative definite, so that

$$f_{jj}(\omega) \geq 0, \quad f_{jj}(\omega) f_{kk}(\omega) - |f_{jk}(\omega)|^2 \geq 0 .$$

In interpreting much of spectral analysis it is useful to note the famous spectral representation theorem (Yaglom, 1962, p. 54) which is

stated here in a simplified form: any real stationary random sequence with zero mean value function and absolutely continuous spectral distribution function has the canonical representation

$$X_j(t) = \int_{-\pi}^{\pi} \cos \lambda t dU_j(\lambda) + \int_{-\pi}^{\pi} \sin \lambda t dV_j(\lambda)$$

where $dU_j(\lambda)$ and $dV_j(\lambda)$ are real random variables indexed by the parameter λ , the integral is defined as the limit in mean square of the natural approximating sum, components at frequencies λ, μ are uncorrelated unless $|\lambda| = |\mu|$, and

$$\begin{aligned} E[dU_j(\lambda)] &= E[dV_j(\lambda)] = 0 \\ E[(dU_j(\lambda))^2] &= E[(dV_j(\lambda))^2] = \frac{1}{2} f_{jj}(\lambda) d\lambda \end{aligned}$$

while for two series $X_j(t)$ and $X_k(t)$ we have, assuming $F_{jk}(\lambda)$ is absolutely continuous,

$$\begin{aligned} E[dU_j(\lambda) dU_k(\lambda)] &= E[dV_j(\lambda) dV_k(\lambda)] = \frac{1}{2} c_{jk}(\lambda) d\lambda \\ E[dU_j(\lambda) dV_k(\lambda)] &= -E[dV_j(\lambda) dU_k(\lambda)] = \frac{1}{2} q_{jk}(\lambda) d\lambda \end{aligned}$$

where

$$f_{jk}(\omega) = c_{jk}(\omega) + iq_{jk}(\omega), \quad j \neq k;$$

similar relations hold at pairs of frequencies $(\lambda, -\lambda)$. More concisely, the representation

$$X_j(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ_j(\lambda)$$

holds, where

$$\begin{aligned} E[|dZ_j(\lambda)|^2] &= f_{jj}(\lambda) d\lambda \\ E[dZ_j(\lambda) \overline{dZ_k(\lambda)}] &= f_{jk}(\lambda) d\lambda \end{aligned}$$

It is clear from the above representation that the autospectrum $f_{jj}(\omega)$ indicates the importance of the basic trigonometric components of given frequency ω for the series $X_j(t)$; for example, a sequence of independent random variables with zero mean and constant variance σ^2 has a constant spectrum

$$f_{jj}(\omega) \equiv \sigma^2/2\pi ;$$

such a process is called white noise. The cross-spectrum indicates, by its real part, whether the frequency ω is important for both series $X_j(t)$ and $X_k(t)$ in so far as their "in phase" components are concerned, and by its imaginary part, in so far as their "in quadrature" components are concerned. The quantities $c_{jk}(\omega)$ and $q_{jk}(\omega)$ are called the cospectrum and quadrature spectrum, respectively.

Another important approach to the interpretation of spectra is the filter approach. If $j \neq k$ and

$$X_j(t) = \sum_{s=-\infty}^{\infty} b_s X_k(t-s) ,$$

then the process $X_j(t)$ is regarded as the output, or response, of the linear time-invariant filter defined by the sequence $\{b_s\}$, when the input is $X_k(t)$; often b_s is zero for $s \leq 0$. The spectra of $X_j(t)$ and $X_k(t)$ can be shown to satisfy the relations

$$f_{jj}(\omega) = |B_{jk}(\omega)|^2 f_{kk}(\omega)$$

$$f_{jk}(\omega) = B_{jk}(\omega) f_{kk}(\omega)$$

where the frequency response function $B_{jk}(\omega)$ is defined by

$$B_{jk}(\omega) = \sum_{s=-\infty}^{\infty} b_s e^{is\omega} .$$

If we had not assumed the mean values of each series to be identically zero and the spectral distribution functions to be absolutely continuous, then it would be possible for some or all of the functions $F_{jk}(\omega)$ to contain one or more jumps; in that case, special problems arise in the study of pairs of time series which are treated elsewhere (Wahba, 1966). We pause only to note that if both autospectral distribution functions are absolutely continuous then it follows from the non-negative definiteness of $f(\omega)$ that the cross spectral distribution function is also absolutely continuous. On the other hand, it is quite possible for both autospectral distribution functions to contain a jump while the cross spectral distribution function is absolutely continuous, as in the following case:

$$\left. \begin{aligned} X_j(t) &= \cos(\alpha t + \Theta) \\ X_k(t) &= \cos(\beta t + \Psi) \end{aligned} \right\} t = 1, 2, \dots$$

where α and β are fixed while Θ and Ψ are independently distributed uniformly on $[-\pi, \pi]$; then $F_{jj}(\omega)$ and $F_{kk}(\omega)$ are unit step functions at α and β , respectively, while $F_{jk}(\omega) \equiv 0$.

1.4 Coherence

The idea of coherence was probably first recognized early in the present century in the study of sound waves when the relative position of their source is varied; rays of light were studied in the same way. Its debut as a quantity of statistical interest was made in an important paper by Wiener (1930, p. 194) where it was defined by the formula

$$\frac{|f_{jk}(\omega)|}{\sqrt{f_{jj}(\omega) f_{kk}(\omega)}} .$$

Some authors (see Tukey, 1965, p. 45) remove the modulus sign from the above definition while others, including the present author, use its square. We thus define

$$W_{jk}(\omega) = \frac{|f_{jk}(\omega)|^2}{f_{jj}(\omega) f_{kk}(\omega)} , \quad -\pi \leq \omega \leq \pi .$$

It follows that

$$W_{jk}(\omega) = W_{kj}(\omega) = W_{jk}(-\omega) .$$

If the denominator is zero, then so is the numerator (since $f(\omega)$ is non-negative definite) and we define

$$W_{jk}(\omega) = \lim_{\lambda \rightarrow \omega} \frac{|f_{jk}(\lambda)|^2}{f_{jj}(\lambda) f_{kk}(\lambda)} .$$

Koopmans (1964, p. 532) prefers to set the coherence equal to zero when the denominator is zero, but this might not be considered entirely natural; for instance, if $X_j(t) \equiv X_k(t)$ one would like the coherence to be identically one even though the autospectra may vanish at some frequencies.

In terms of the spectral representation

$$X_j(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ_j(\lambda)$$

we have

$$W_{jk}(\lambda) = \frac{|E[dZ_j(\lambda) \overline{dZ_k(\lambda)}]|^2}{E[|dZ_j(\lambda)|^2] E[|dZ_k(\lambda)|^2]} ,$$

showing that the form of the coherence is analogous to that of the square of the classical correlation coefficient for real random variables; however, this analogy is not complete and it does not seem to be particularly useful.

A more useful interpretation of coherence relates it to a linear filter with additive (but not necessarily white) noise. Suppose that

$$X_j(t) = \sum_{s=-\infty}^{\infty} b_s X_k(t-s) + \eta(t)$$

where $\eta(t)$ is uncorrelated with $X_k(t-s)$ for all s . Let us write this in suggestive notation as

$$X_j(t) = X_{j|k}(t) + X_{j\perp k}(t),$$

where $X_{j|k}(t) = \sum_{s=-\infty}^{\infty} b_s X_k(t-s)$ and $X_{j\perp k}(t) = \eta(t)$. Then

$$\begin{aligned} W_{jk}(\omega) &= \frac{|B_{jk}(\omega)|^2 f_{kk}(\omega)}{|B_{jk}(\omega)|^2 f_{kk}(\omega) + f_{\eta\eta}(\omega)} \\ &= \frac{1}{1 + \frac{f_{j\perp k}(\omega)}{f_{j|k}(\omega)}} \end{aligned}$$

where the f 's are the autospectra of the series indicated by subscripts.

This relation shows that

$$0 \leq W_{jk}(\omega) \leq 1$$

for all ω , which also follows from the non-negative definiteness of $\underline{f}(\omega)$.

The interpretation of $W_{jk}(\omega)$ is now clear: if there is a precise linear filter relation between $X_j(t)$ and $X_k(t)$, then

$W_{jk}(\omega) \equiv 1$; if the linear filter model is satisfied with $b_s \equiv 0$, then $W_{jk}(\omega) \equiv 0$; in intermediate cases the proportion of power at each frequency ω explained by the filter will be reflected in the value of $W_{jk}(\omega)$, for we have the relations

$$W_{jk}(\omega) = \frac{f_{j|k}(\omega)}{f_{jj}(\omega)} = \frac{f_{k|j}(\omega)}{f_{kk}(\omega)} .$$

These and other properties of coherence have been derived in an analytic setting by Koopmans (1964). The naturalness of the above interpretative formulae is the main reason for preferring our definition of coherence to the others mentioned; the extension to the multiple coherence between three or more series provides further justification, by analogy with the classical multiple correlation coefficient usually denoted R^2 .

A simple function of coherence has been shown by Gelfand and Yaglom (1959, p. 229) to be a natural measure of the amount of information contained in either of two processes with respect to the other. The measure, which is derived from Shannon's definition of information for random variables, is

$$\begin{aligned} \mathcal{J}_{jk} &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(1 - W_{jk}(\lambda)) d\lambda \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left(1 + \frac{f_{j|k}(\lambda)}{f_{j\perp k}(\lambda)} \right) d\lambda . \end{aligned}$$

In view of the above remarks about the nature of coherence it is not surprising that it has been found, particularly since the second world war, to be a very useful object of study in many scientific fields including meteorology, oceanography, geophysics, engineering, communications, optics and economics. In some branches of each of these

fields and others coherence is now an established tool of research; in fact, the statistical terminology is derived from some of these fields.

However, it appears that the statistical properties of estimates of coherence from sample observations have not been studied quite sufficiently; Chapter 2 is designed to explain why the usual procedures, established in the last fifteen years, are in general not satisfactory and can lead the user seriously astray.

1.5 Phase and Gain

It is customary to consider, along with coherence, the so called phase angle

$$\varphi_{jk}(\omega) = \arctan(q_{jk}(\omega)/c_{jk}(\omega))$$

and the gain

$$G_{jk}(\omega) = |f_{jk}(\omega)|/f_{kk}(\omega) ;$$

in the notation of Section 1.3 above; an arbitrary rule is needed to avoid the non-uniqueness of the arctangent. We then have

$$f_{jk}(\omega) = f_{kk}(\omega) G_{jk}(\omega) e^{i\varphi_{jk}(\omega)} .$$

In simple cases such as a lagged series $X_j(t) = X_k(t-\tau)$ the phase may be easy to interpret, but it seems to be a useful quantity only in a few special cases; in general, positive and negative lags can occur between two series simultaneously, in which case the author is not convinced of the relevance of phase as defined above (see also Granger and Hatanaka, 1964, p. 103). The gain may be more easily interpreted than phase, but since coherence is a normalized function of gain it is generally

the more important quantity. For these reasons this thesis is restricted to the study of the estimation of coherence.

1.6 The Aims of This Thesis

In Chapter 2 the usual procedure for estimating coherence is stated and discussed; it is explained that this procedure needs to be improved. With this improvement in mind, the so called cross regressive filter model is proposed in Section 3.1. In spite of the naturalness of estimates of coherence obtained in this way, they have apparently not been used in the past. In testing whether the proposed model is satisfied one needs to test the hypothesis of identically zero coherence; a natural test is proposed in Section 3.5, requiring a knowledge of the mean and variance of the estimated coherence under the null hypothesis.

Approximate values of these moments are obtained in Sections 3.2 and 3.3 and the opportunity is taken to derive formulae for joint moments of any order of the estimated spectra. These formulae (see Theorems 1 and 2) are believed to be new. In deriving these formulae no attempt has been made to achieve mathematical rigor. Thus, precise smoothness conditions on the spectra are not given, and the behavior of delta-like functions is treated in an intuitive way. To avoid this feature, one would like to work in the time domain rather than the frequency domain, but this proved impracticable. The lack of rigor will not affect the application of the results, as is indicated in Section 3.4 by comparison with an approximate distribution for coherence obtained by N. Goodman assuming, among other things, a rectangular spectral window.

In Sections 3.6, 3.7, 3.8 some miscellaneous alternative procedures are briefly considered, but none of these is recommended.

2. A Critical Study of the Usual Estimation Procedures

2.1 Estimation of Autospectra

(i) The periodogram. The autospectrum is the Fourier transform of the autocovariance function, so the first attempt at estimating the autospectrum was to obtain the Fourier transform of the estimated autocovariance function, called the sample periodogram:

$$\begin{aligned} I_{jj}(\omega) &= \frac{1}{2\pi} \sum_{|v| < T} \cos v\omega \hat{R}_{jj}(v) \\ &\equiv \frac{1}{2\pi T} \left| \sum_{t=1}^T X_j(t) e^{-i\omega t} \right|^2 \end{aligned}$$

where

$$\hat{R}_{jj}(v) = \frac{1}{T} \sum_{t=1}^{T-|v|} X_j(t) X_j(t+v) .$$

In the definition of $\hat{R}_{jj}(v)$ the divisor $T - |v|$ is sometimes used instead of T ; this matter, which has been discussed at length in the literature (e.g., Tukey, 1961, pp. 211-213), will be mentioned briefly in Section 2.1 (ii).

Although

$$E[I_{jj}(\omega)] \rightarrow f_{jj}(\omega) \quad \text{as } T \rightarrow \infty ,$$

it turns out that

$$\begin{aligned} \text{Var}[I_{jj}(\omega)] &\rightarrow f_{jj}^2(\omega), & |\omega| \neq 0, \pi \\ &2f_{jj}^2(\omega), & |\omega| = 0, \pi . \end{aligned}$$

Hence the periodogram is not consistent and, therefore, not satisfactory as an estimate of $f_{jj}(\omega)$.

(ii) Windowed periodograms. In order to achieve a consistent estimate of $f_{jj}(\omega)$ an average can be taken of the periodogram at frequencies near ω . We, therefore, define

$$\begin{aligned}\hat{f}_{jj}(\omega) &= \int_{-\pi}^{\pi} I_{jj}(\lambda) K_M(\lambda-\omega) d\lambda \\ &\equiv \frac{1}{2\pi} \sum_{|v| \leq M} \hat{R}_{jj}(v) k\left(\frac{v}{M}\right) \cos v\omega\end{aligned}$$

where $K_M(\lambda) \equiv \frac{1}{2\pi} \sum_{|v| \leq M} k\left(\frac{v}{M}\right) \cos v\lambda$.

There is some freedom in the choice of the weights $K_M(\lambda)$ in the frequency domain, or equivalently, $k\left(\frac{v}{M}\right)$ in the time domain. The extensive literature on this choice will not be repeated here but the weights will be assumed to satisfy the following properties.

The spectral window $K_M(\lambda)$ is real valued, continuous, symmetrical about $\lambda = 0$, and has period 2π ; on $[-\pi, \pi]$ it takes its maximum value at $\lambda = 0$, decreases monotonically as $|\lambda|$ increases, and is positive for small $|\lambda|$ (preferably positive for all λ). The bandwidth (any function measuring the degree of concentration of $K_M(\lambda)$ about $\lambda = 0$) depends on the number M ; as M increases the spectral window approaches a delta function at $\lambda = 0$.

The lag window $k\left(\frac{v}{M}\right)$ is real valued, continuous, takes its maximum value of 1 at $v = 0$, is symmetrical about $v = 0$, decreases monotonically as $|v|$ increases, is positive for $|v| \leq M$ and vanishes otherwise. Hence, M is called the truncation point.

A useful approximation introduced by Parzen is obtained by defining

$$K(\lambda) = \frac{1}{2\pi} \int_{-1}^1 k(u) \cos u\lambda \, du, \quad -\infty < \lambda < \infty$$

so that

$$K_M(\lambda) \doteq MK(M\lambda).$$

Two popular choices are the Parzen kernel

$$K(\lambda) = \frac{3}{8\pi} \left\{ \frac{\sin(\lambda/4)}{\lambda/4} \right\}^4$$

and the Tukey kernel

$$K(\lambda) = \frac{1}{4\pi} \left\{ 2 \frac{\sin \lambda}{\lambda} + \frac{\sin(\lambda+\pi)}{\lambda+\pi} + \frac{\sin(\lambda-\pi)}{\lambda-\pi} \right\}.$$

The Parzen kernel has the advantage that, if $\hat{R}_{jj}(v)$ is defined with divisor T , then $\hat{f}_{jj}(\omega)$ is necessarily positive. Other kernels are given in Table 5, Section 3.4.

Consistent estimates of $f_{jj}(\omega)$ are achieved by letting $T/M \rightarrow \infty$ and $M \rightarrow \infty$. The mean and variance of $\hat{f}_{jj}(\omega)$ will be studied in Sections 3.2 (ii), (iii). In ensuring that the variance approaches zero some bias has been accepted; the bias also approaches zero and the procedure is a good one in principle. It is necessary, however, to guard against introducing a serious bias when using particular finite values of T and M . It is clear from the definition of $\hat{f}_{jj}(\omega)$ in terms of $\hat{R}_{jj}(v)$ that a large bias will be avoided if and only if $k(\frac{v}{M})$ is never small when $R_{jj}(v)$ is significantly different from zero. The equivalent requirement in the frequency domain is that the spectral window should be sufficiently concentrated at its central value to pick up significant oscillations in $f_{jj}(\omega)$ and to avoid "leakage" from nearby frequencies.

Hence if the estimate $\hat{f}_{jj}(\omega)$ is to be satisfactory, it is essential that the truncation point M be chosen large enough to avoid the danger mentioned. This may be feasible in cases where the greatest lag at which $R_{jj}(v)$ is significant, say $v = \bar{v}$, is known in advance and such cases apparently arise quite often in economics and some other fields. It is also conceivable that knowledge of \bar{v} could be inferred from a study of the estimated autocovariance function (see Schaerf, 1963, p. 25). However, no such procedure is yet known which would be suited to routine application to any series which might be read into a computer.

(iii) Prewhitening. To deal with the difficulty mentioned above it is often thought desirable to make a preliminary transformation of the given series, so that the covariance function of the transformed series is close to zero at lags other than zero. If this is achieved, then the choice of truncation point is no longer so vital. The natural transformation to make is suggested by the linear filter or autoregression model

$$X_j(t) - \sum_{s=-\infty}^{\infty} b_s X_j(t-s) = \epsilon_{jj}(t)$$

where $\epsilon_{jj}(t)$ is white noise.

Again the problem arises of how many lags to use, that is, which values of s to choose; but now reliable procedures are available to test whether enough lags have been used. Valid tests can be based on stagewise least squares estimation formulae for the \hat{b}_s , in terms of the capacity of new regressors to reduce the residual variance; these formulae have been given by Efroymsen (1960), Parzen (1966) and others. Alternatively they can be based on functions of the estimated

autocovariances as in Diananda (1953), Quenouille (1947) and others.

In the present context the stagewise tests are the more convenient.

The usual practice has been to set $b_s = 0$ for $s \geq 0$. The author sees no good reason for this, since the aim here is not prediction but rather interpolation; he accordingly suggests that the \hat{b}_s be formed for all significant lags s , positive and negative. Since the estimates of b_s will be more reliable for small lags $|s|$ than for large lags, it can only be advantageous to allow the possibility of lags of both signs.

The autoregression model is here nothing more than a tool which may be useful: it need not be believed without reservation. A test for the model can be based on the estimated spectrum of the residuals, which should be approximately that of white noise. A formal procedure such as the Kolmogorov-Smirnov test could be used; otherwise, a visual test may be satisfactory. In any case it should be realized that the appearance of a flat windowed spectrum does not imply that the residual series is approximately white noise, for it might retain a periodicity of higher frequency than can be picked up by the window used. (A single value of much different order of magnitude from the values at other time points, which may be caused by a recording error, will also give rise to a flat spectrum.) All these considerations reflect the fact that an experimenter must be able to make at least some minimal assumptions about the nature of his series, for instance a bound \bar{M} on the greatest significant lag \bar{v} , before a meaningful statistical analysis can be performed. (Such a bound exists since we are assuming

$\sum_{s=-\infty}^{\infty} b_s^2 < \infty$, or in the words of Schaerf (1963, p. 31),

"... we are confident that the remote past will not influence very much the variable to be estimated.")

The final estimate of the spectral density can be derived from the fitted autoregression by one of two methods:

- (a) assuming the fitted model is valid, one forms

$$\hat{f}_{jj}^{(a)}(\omega) = \frac{\hat{\sigma}_e^2}{2\pi} / \{1 - |\sum_s b_s e^{-i\omega s}|^2\};$$

- (b) allowing for the possibility that the residuals are not precisely white noise, one forms the recolored estimate

$$\hat{f}_{jj}^{(r)}(\omega) = \hat{f}_{e jj}(\omega) / \{1 - |\sum_s b_s e^{-i\omega s}|^2\}.$$

These two methods will normally give similar results.

One might ask whether the desired transformation to a situation where the spectrum is relatively flat could be achieved by an operation in the frequency domain, rather than in the time domain via the filter model described above. In principle it is possible to do so, but the difficulty of analysing the situation in that way is, in the author's opinion, very great; this point will be discussed further in connection with the estimation of coherence in Section 3.1.

2.2 Estimation of Cross-spectra

- (i) The cross-periodogram. The problem of estimating the cross-spectrum is of course similar to that of estimating the autospectra.

We define the cross-periodogram:

$$\begin{aligned}
I_{jk}(\omega) &= \frac{1}{2\pi} \sum_{|v| < T} e^{-iv\omega} \hat{R}_{jk}(v) \\
&\equiv \frac{1}{2\pi} \sum_{t=1}^T X_j(t) e^{-i\omega t} \sum_{t=1}^T X_k(t) e^{i\omega t} .
\end{aligned}$$

The question of the divisor for $\hat{R}_{jk}(v)$, here taken to be T , will be mentioned in Section 2.3 and again in Section 3.1. The cross-periodogram, like the periodogram, is asymptotically unbiased but is not consistent.

(ii) Windowed cross-periodograms. Consistency can again be achieved by the use of a window:

$$\begin{aligned}
\hat{f}_{jk}(\omega) &= \int_{-\pi}^{\pi} I_{jk}(\lambda) K_M(\lambda - \omega) d\lambda \\
&\equiv \frac{1}{2\pi} \sum_{|v| \leq T} \hat{R}_{jk}(v) k\left(\frac{v}{M}\right) e^{-iv\omega} .
\end{aligned}$$

Usually the same window is used for estimating the cross-spectrum as for estimating the autospectra, but there is now less reason. For example, the symmetry of the autocovariance function means that the window is most naturally taken to be symmetrical, but this is no longer true of the cross-covariance function; hence, an asymmetrical window might be more appropriate in general and one might better speak of two truncation points, an upper and a lower one. It is also now possible that a complex-valued window would be preferable. Nevertheless, windows having the properties listed in Section 1 (ii) of this chapter will be found adequate for cross-spectral estimation provided a preliminary transformation (to be discussed) is first carried out. The relevance of these windows in the absence of such a transformation is very doubtful.

In any case it is necessary, as before, to check that the bias

introduced by the window is tolerable. To see the danger which exists one need only consider any pair of series for which $|R_{jk}(v)|$ is significantly large for a value of v at which $k(v/M)$ is small, possibly zero. An extreme example would be

$$X_k(t) = X_j(t-L) = \text{white noise with variance } \sigma^2$$

and $M = L$; then $f_{jk}(\omega) \equiv \sigma^2$, but $\hat{f}_{jk}(\omega) \equiv 0$. The problem of bias is possibly more difficult for the cross-spectrum than for autospectra because there may be less a priori knowledge of the cross-covariance structure than of the autocovariance structure; in economics, for instance, autocovariance diagrams are widespread but cross-covariance diagrams seem to be studied only rarely. However, it is fairly clear that this problem has not usually been faced at all: the formulae obtained for autospectral estimates have been applied unchanged, even as to truncation point, to the problem of cross-spectral estimation (e.g., Granger and Hatanaka, 1964 pp. 78, 101; Rosenblatt, 1965, p. 48).

(iii) Preflattening. The danger of a serious bias in the cross-spectral estimates described above has not generally been recognized. Therefore, it is not surprising that a preliminary transformation analogous to prewhitening has not generally been sought. The analogous transformation may be called "preflattening," since whereas the autospectra were made relatively flat by prewhitening, the cross-spectrum is now to be made flat so that the application of a window will no longer have its former danger. Suggested methods for achieving this effect will be presented in Chapter 3.

A few writers have recognized the above problem. The first were

Dalzell and Yamanouchi (1958).

Tick (1963, p. 203) wrote:

"In any discussion of spectral estimation we should keep to the fore the fact that estimates of spectra are biased - and may be seriously so. In fact, contrary to what I hear in mathematical statistical circles, I believe this is the controlling factor in many situations."

Akaike and Yamanouchi (1962, p. 55) sought but did not find an appropriate transformation:

"... our new window Q is better suited...fairly large bias due to smoothing still remains...we believe that the use of well designed and properly shifted window will eventually lead to successful results."

Pierson and Dalzell (1960) and Parzen (1965, pp. 46-48) gave a method which is appropriate for dealing with cases where the significant cross-covariances are concentrated in a neighborhood of one lag, say $v = v^*$, but which is not appropriate otherwise (see Section 3.6).

In summary, the problem of serious bias which can occur in the estimated cross-spectrum has not generally been recognized, and when recognized has not been solved satisfactorily. One aim of Chapter 3 is to solve this problem in so far as it affects estimates of coherence.

2.3 Estimation of Coherence

The coherence is a function of the auto- and cross-spectra; it has usually been estimated by the same function of the estimated auto- and cross-spectra, modifications sometimes being needed to ensure that the estimate lies between 0 and 1. Thus, the usual estimate is

$$\hat{W}_{jk}(\omega) = \frac{|\hat{f}_{jk}(\omega)|^2}{\hat{f}_{jj}(\omega) \hat{f}_{kk}(\omega)} .$$

This will lie between 0 and 1 if the same truncation point is used with the Parzen window and the divisor T for each of $\hat{f}_{jk}(\omega)$, $\hat{f}_{jj}(\omega)$ and $\hat{f}_{kk}(\omega)$. However, as indicated in the previous section, the same truncation point will in general not be appropriate for all three estimates as a result of the different autocovariance and cross-covariance structures.

Of course, a function $g(\hat{f})$ is not necessarily a good estimate of the function $g(f)$, even if \hat{f} is a good estimate of f . In the present case a positive bias will be found to occur, unless $W_{jk}(\omega) = 1$; the possibility of using a different function of \hat{f} to estimate coherence will be considered in Section 3.8. But the main reasons why the author claims that estimates of coherence have often been poor are contained in the previous sections, namely, the criticism of the estimates of the cross-spectrum and, to a lesser extent, of the autospectra. The bias in $\hat{W}_{jk}(\omega)$ arising from an inappropriate truncation point (or window) for the cross-spectrum will be seen to be negative. It is possible that the two sources of bias mentioned will roughly cancel out in some cases, but such an accident is hardly to be relied upon.

The unsatisfactory nature of the usual estimates of coherence is not only theoretical but has been observed in practice; this can best be indicated by some views stated in the literature, of which several are given below.

Pierson and Dalzell (1960, p. 20): "The loss of coherency when computed by standard techniques from samples of vector Gaussian processes where, theoretically, the coherency ought to be one, is largely explained by the lack of resolution in the cross spectra and the effect of the convolving filter made necessary by the nature of the finite sample. High coherencies can be regained by modifying the experimental design so as to obtain less rapidly varying cross spectra and by increasing the resolution."

Akaike and Yamanouchi (1962, p. 23): "The method described by Goodman was a direct application of the method of estimation of the spectral density to that of the cross-spectral density, but some experimenters who applied this kind of method to their numerical data experienced the very low coherency of their estimates."

Granger (1965, p. 232): "Monte Carlo studies currently under way show that the estimates of coherence for short series are badly biased towards the value $\frac{1}{2}$. This suggests that new methods of estimating cross-spectra need to be investigated."

Priestley (1965, p. 235): "...it seems clear that, in common with other branches of spectral analysis, the final word on estimation procedures has yet to be written."

3. Estimation of Coherence by Cross-regressive Filtering and Related Methods

3.1 Cross-regressive Filtering

In Chapter II we established the need for a preliminary transformation of the given series in an attempt to make the cross-spectrum relatively flat. The natural method for achieving this is the analog of the autoregressive procedure. We therefore consider the model

$$X_j(t) = \sum_{s=-\infty}^{\infty} b_s X_k(t-s) + \eta(t)$$

where $\eta(t)$ has zero mean and is independent of (or at least uncorrelated with) $X_k(t-s)$ for all s ; that is, $W_{\eta k}(\omega) \equiv 0$ where $W_{\eta k}(\omega)$ is the coherence between $\eta(t)$ and $X_k(t)$. It is not now reasonable in general to assume that $\eta(t)$ is a white noise process, whereas this was the natural assumption in the autoregression model. For instance, if $X_j(t)$ and $X_k(t)$ are uncorrelated, then all the b_s are zero so that $\eta(t)$ will have the same spectrum as has $X_j(t)$ itself, which might be quite different from white noise.

The method of estimating the coefficients b_s will again be the stagewise least-squares procedure, using the formulae of Efroymsen (1960) or Parzen (1966). Only those lags s will be utilized which lead to a significant reduction in the variance of the residuals; of course s can take values $0, \pm 1, \pm 2, \dots$.

Assuming the model has been satisfactorily estimated by coefficients \hat{b}_s one would estimate the cross-spectrum by

$$\hat{f}_{jk}^{(c)}(\omega) = \sum_s \hat{b}_s e^{i\omega s} \hat{f}_{kk}^{(a)}(\omega),$$

the superscript (c) indicating cross-regression. This estimate is

appropriate because if the model is satisfied we have

$$\begin{aligned} R_{jk}(\nu) &= E[\{\sum_s b_s X_k(t-s) + \eta(t)\} X_k(t+\nu)] \\ &= \sum_s b_s R_{kk}(\nu+s) . \end{aligned}$$

If it is desired to take into account the presumably small amount of correlation between $\hat{\eta}(t)$ and $X_j(t-s)$ for each s , one would form the estimate analogous to the recolored autospectral estimate:

$$\hat{f}_{jk}^{(r)}(\omega) = \sum_s \hat{b}_s e^{i\omega s} \hat{f}_{kk}^{(r)}(\omega) + \hat{f}_{\eta k}(\omega) ,$$

the superscript (r) indicating recoloring. Again there will be little difference between $\hat{f}_{jk}^{(c)}(\omega)$ and $\hat{f}_{jk}^{(r)}(\omega)$. Note that a separate model is used for the purpose of estimating the autospectrum of $X_k(t)$, as described in Section 2.1 (iii). Of course the two models could be combined to give, say,

$$X_j(t) = \sum_r g_r X_k(t-r) + \sum_r h_r \epsilon(t-r) + \eta(t)$$

but it seems better to treat the two models separately.

Our main interest here is the estimation of coherence and we propose the following estimates:

$$\begin{aligned} \hat{W}_{jk}^{(ac)}(\omega) &= \frac{|\hat{f}_{jk}^{(c)}(\omega)|^2}{\hat{f}_{jj}^{(a)}(\omega) \hat{f}_{kk}^{(a)}(\omega)} \\ \hat{W}_{jk}^{(r)}(\omega) &= \frac{|\hat{f}_{jk}^{(r)}(\omega)|^2}{\hat{f}_{jj}^{(r)}(\omega) \hat{f}_{kk}^{(r)}(\omega)} . \end{aligned}$$

If we write the model as

$$X_j(t) = X_{j|k}(t) + X_{j\perp k}(t) ,$$

then we have

$$\hat{W}_{jk}^{(ac)}(\omega) = \frac{1}{1 + \frac{\hat{\hat{f}}_{j\perp k}(\omega)}{\hat{\hat{f}}_{j|k}(\omega)}}$$

where the double hats indicate the usual estimates of spectra of the estimated component series, say $\hat{X}_{j|k}(t)$, $\hat{X}_{j\perp k}(t)$ (compare the corresponding equation in Section 1.4). We see that the estimate will lie between zero and one if the Parzen window is used with divisor T , even though different lags may be used in estimating the cross-spectra and each of the autospectra.

One might ask whether the two series, $X_j(t)$ and $X_k(t)$, should be individually filtered before estimating the cross-regression model. However, it is easy to see that such auto-filtering would have no effect on the outcome, since relations within one series have no bearing on its relations to another series. As a result of the symmetry of $W_{jk}(\omega)$, either series can be chosen as input ($X_k(t)$) when using the cross-regression model. The possibility of achieving satisfactory estimates of coherence by operations in the frequency domain alone is no more at hand, in the author's view, than in the autospectral case, as noted at the end of Section 2.1 (iii).

At this point mention must be made of a paper by Hannan (1963) in which the point of view taken is the opposite one to that adopted here. Hannan seeks to estimate the same model as we do (Hannan, 1963, pp 34-35)

by spectral methods alone. However, the determination of the number of lags to be used cannot, in our view, be handled satisfactorily in the frequency domain alone; while this problem is not of great mathematical interest it is very important in the use of spectral methods. Hannan's description seems, in comparison with the stagewise least squares procedure, to be rather vague: "...the number of lags is to be determined by an examination of the data...in any case, lags are included until the two methods of estimating $f_{ee}(\lambda)$ give the same results effectively." (Hannan, 1963, p. 35). Apparently no allowance is made here for a lag which might have been significant when first introduced but not significant after some subsequent lag was utilized. In any case we wish to state our view that time domain methods and frequency domain methods should be used in conjunction wherever possible, as a test and safeguard of the validity of either approach alone.

In testing the fitted cross-regression model we are no longer concerned with the spectrum of the residuals, but only with their cross-spectrum, or coherence, with the series $X_k(t)$. We wish to test the null hypothesis

$$W_{\eta k}(\omega) \equiv 0 .$$

For this purpose we will need the distribution of the estimated coherence $\hat{W}_{jk}(\omega)$ for two series (in the application j will refer to $\eta(t)$ and k to $X_k(t)$) under the null hypothesis and under relevant alternatives. A statistic such as

$$\int_{-\pi}^{\pi} \hat{W}_{jk}(\lambda) d\lambda$$

could then be used. The next three sections are devoted to a study of some aspects of the distribution of the sample coherence, after which we will return to the problem of estimating coherence and testing the cross-regression model.

3.2 Asymptotic Theory of Moments of Estimated Spectra

(i) Lemma on diagonal summation.

Lemma 1 Let g be any function of $n - 1$ real variables such that

$$\sum_{\tau_1=-T}^T \cdots \sum_{\tau_{n-1}=-T}^T g(\tau_1, \dots, \tau_{n-1})$$

converges absolutely as $T \rightarrow \infty$. Then for large T and for $n = 2, 3, \dots$

$$\frac{1}{T} \sum_{t_1=1}^{T-v_1} \cdots \sum_{t_n=1}^{T-v_n} g(t_1-t_n, \dots, t_{n-1}-t_n) \doteq \sum_{\tau_1=-T}^T \cdots \sum_{\tau_{n-1}=-T}^T g(\tau_1, \dots, \tau_{n-1})$$

where the v_i are fixed non-negative integers.

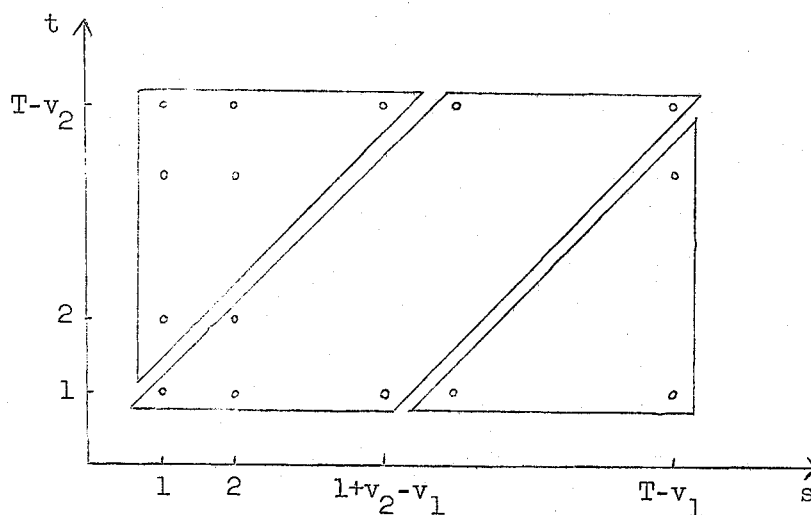
Proof: We begin with the case $n = 2$. For $T > \max v_i$ we have

$$\frac{1}{T} \sum_{t_1=1}^{T-v_1} \sum_{t_2=1}^{T-v_2} g(t_1-t_2) \equiv \sum_{\tau=-T+1}^{T-1} U_T(\tau; v_1, v_2) g(\tau)$$

where

$$U_T(\tau; v_1, v_2) = \begin{cases} 1 - \frac{v_2 - \tau}{T} & -T + v_2 + 1 \leq \tau \leq \min(0, v_2 - v_1) - 1 \\ 1 - \frac{\max(v_1, v_2)}{T} & \min(0, v_2 - v_1) \leq \tau \leq \max(0, v_2 - v_1) \\ 1 - \frac{v_1 + \tau}{T} & \max(0, v_2 - v_1) + 1 \leq \tau \leq T - v_1 - 1 \\ 0 & \text{otherwise} \end{cases}$$

The diagram illustrates the case $v_1 < v_2$; since $|U_T(\tau; v_1, v_2)| \leq 1$ we can interchange the summation and limit as $T \rightarrow \infty$, yielding the result.



The general case is treated by the same method, but the exact expression for $U_T(\tau_1, \dots, \tau_{n-1}; v_1, \dots, v_n)$ is complicated since it takes different non-zero values on each of N regions of an n -dimensional lattice, where N equals 3 raised to the power $\binom{n}{2}$. We give explicitly only the special case

$$U_T(\tau_1, \dots, \tau_{n-1}; 0, \dots, 0) = 1 - \frac{\max(\tau_1, \dots, \tau_{n-1})}{T}, \quad -T+1 \leq \tau_i \leq T-1.$$

In any case

$$|U_T(\tau_1, \dots, \tau_{n-1}; v_1, \dots, v_n)| \leq 1$$

which yields the general result, as in the case $n = 2$.

Note: The case $n = 2$ was given by Parzen (1957, p. 335) for the case of continuous time; but note that the sign of u in the second line of the definition of $U(u, v_1, v_2)$ is incorrect.

(ii) Lemma on means of estimated spectra.

Lemma 2 For sufficiently large values of T and M we have, for

jointly stationary series $X_j(t)$, $X_k(t)$, having uniformly continuous spectra,

$$E[\hat{f}_{jk}(\omega)] = f_{jk}(\omega), \quad |\omega| \leq \pi.$$

The result holds for $j = k$ and for $j \neq k$.

Proof:

$$\begin{aligned} E[\hat{f}_{jk}(\omega)] &= \frac{1}{2\pi} \sum_{|v| \leq M} E \frac{1}{T} \sum_{t=1}^{T-|v|} X_j(t) X_k(t+v) e^{-iv\omega} k\left(\frac{v}{M}\right) \\ &= \frac{1}{2\pi} \sum_{|v| \leq M} R_{jk}(v) e^{-iv\omega} k\left(\frac{v}{M}\right), \end{aligned}$$

for sufficiently large T

$$= \frac{1}{2\pi} \sum_{|v| \leq M} \int_{-\pi}^{\pi} f_{jk}(\lambda) e^{iv\lambda} d\lambda e^{-iv\omega} k\left(\frac{v}{M}\right)$$

$$\equiv \int_{-\pi}^{\pi} f_{jk}(\lambda) K_M(\lambda - \omega) d\lambda$$

$$\doteq f_{jk}(\omega) \int_{-\pi}^{\pi} K_M(\lambda - \omega) d\lambda, \quad \text{for sufficiently large } M, \\ \text{using the uniform continuity} \\ \text{of } f_{jk}(\lambda)$$

$$= f_{jk}(\omega).$$

Note: Further terms can be obtained under appropriate conditions by expanding the spectral density function in a Taylor series, leading to the result

$$\begin{aligned} E[\hat{f}_{jk}(\omega)] &\doteq \sum_{s=0}^{\infty} \frac{1}{(2s)!} \frac{1}{M^{2s}} f_{jk}^{(2s)}(\omega) \int_{-\infty}^{\infty} \lambda^{2s} K(\lambda) d\lambda \\ &= \sum_{s=0}^{\infty} (-1)^s \frac{1}{(2s)!} \frac{1}{M^{2s}} k^{(2s)}(0) f_{jk}^{(2s)}(\omega), \quad |\omega| \leq \pi. \end{aligned}$$

Thus if $f_{jk}(\omega) = 0$ but $f_{jk}''(\omega) \neq 0$ we have

$$E[\hat{f}_{jk}(\omega)] \doteq -\frac{1}{2M^2} k''(0) f_{jk}''(\omega).$$

The odd terms of the Taylor series do not appear since $K(\lambda)$ is an even function.

(iii) General central moments of estimated spectra. In the sequel we will need to know joint moments of various orders of estimated spectra. All the cases we need could be computed separately as they arise, but this is very tedious and some symmetries, which are useful in checking the formulae, would be lost. The general formula is quite easy to apply in any particular case although when stated in symbols it appears rather formidable. We will give a statement of the general result, but for a clear understanding of it a reading of the main steps in the proof is recommended before close scrutiny of the statement itself. Since the formulae obtained may be useful in various investigations involving spectral estimates, we have taken the opportunity to give a little more detail and completeness than is required for the present purposes.

Theorem 1 Let $\mu_{12, \dots, 2n-1, 2n} \equiv E[(\hat{f}_{12}(\omega) - f_{12}(\omega)) \dots (\hat{f}_{2n-1, 2n}(\omega) - f_{2n-1, 2n}(\omega))]$ for $n = 1, 2, \dots$. If the series are jointly stationary and jointly normal and have sufficiently smooth spectra, then for large enough values of T and M we have, recursively,

$$\mu_{12, \dots, 2n-1, 2n} = \sum_{\mathcal{P}: 2 \leq m \leq \bar{m}} \mu_{i_1 j_1, \dots, i_m j_m} \cdot \mu_{i_{m+1} j_{m+1}, \dots, i_n j_n} + \left(\frac{2\pi M}{T}\right)^{n-1} \int_{-\infty}^{\infty} K^n(\lambda) d\lambda \sum_{\mathcal{P}: m=n} f_{i_1 j_1} \dots f_{i_n j_n} \quad \text{if } |\omega| \neq 0, \pi$$

where

$$\{i_1, \dots, i_n\} = \mathcal{P}\{1, 3, \dots, 2n-1\}$$

$$\{j_1, \dots, j_n\} = \mathcal{P}\{2, 4, \dots, 2n\}$$

where \mathcal{P} means "a permutation of";

$$\begin{aligned}
 m &\equiv m\{i_1, j_1, \dots, i_n, j_n\} \\
 &= \min_{m' \in \{1, \dots, n\}} (m' \text{ such that } \{i_{k_1}, j_{k_1}, \dots, i_{k_{m'}}, j_{k_{m'}}\} \\
 &= \mathcal{P}\{i_1, i_1+1, \dots, i_{m'}, i_{m'}+1\});
 \end{aligned}$$

and

$$\bar{m} = \max(2, [n/2]).$$

If $|\omega| = 0, \pi$, replace $\mu_{i_1 j_1, \dots, i_m j_m}$ by $\mu_{l_1 l_2, \dots, l_{2m-1} l_{2m}}$, $\mu_{i_{m+1} j_{m+1}, \dots, i_n j_n}$ by $\mu_{l_{2m+1} l_{2m+2}, \dots, l_{2n-1} l_{2n}}$, and $f_{i_1 j_1} \dots f_{i_n j_n}$ by $f_{l_1 l_2, \dots, l_{2n-1} l_{2n}}$, where $\{l_1, l_2, \dots, l_{2n}\} = \mathcal{P}\{1, 2, \dots, 2n\}$. Also, the

uncentered moment

$$\mu_{l_2, \dots, l_{2n-1} l_{2n}} \equiv E[\hat{f}_{l_2}(\omega) \dots \hat{f}_{l_{2n-1} l_{2n}}(\omega)]$$

is given by the same recursion relation, with $\sum_{\mathcal{P}: m=2}^{\bar{m}}$ replaced by $\sum_{\mathcal{P}: m=1}^{\bar{m}}$.

The time series identified by integer subscripts may or may not be all different.

Note. The assumption of joint normality is not essential; any joint moment of a process can be expressed as the sum of two terms: a Gaussian part, being the value of the moment for a normal process, and by subtraction a non-Gaussian part. In many cases it is possible to give the non-Gaussian part a Fourier representation and to show that its contribution is asymptotically negligible. This concept, introduced by Magness (1954), is applied by Parzen (1957) in the case of fourth moments, but the details will not be spelled out in the present more complicated case.

Proof: The proof for general n can be given most clearly by first writing out in detail the cases $n = 2, 3, 4$ (the case $n = 1$ is Lemma 2).

The case $n = 2$. We have

$$\begin{aligned} & E[(\hat{f}_{12}(\omega) - f_{12}(\omega))(\hat{f}_{34}(\omega) - f_{34}(\omega))] \\ &= \left(\frac{1}{2\pi}\right)^2 \sum_{u,v} \frac{1}{T^2} \sum_{t_1, t_2} E[X_1(t_1) X_2(t_1+u) X_3(t_2) X_4(t_2+v)] e^{-i(u+v)\omega} k\left(\frac{u}{M}\right) k\left(\frac{v}{M}\right) \\ & \quad - f_{12}(\omega) f_{34}(\omega) \end{aligned}$$

where we have used Lemma 2 in subtracting off the means. To evaluate the expectation we apply Isserlis's formula (1918), obtaining

$$\begin{aligned} & E[X_1(t_1) X_2(t_1+u) X_3(t_2) X_4(t_2+v)] \\ &= R_{12}(u) R_{34}(v) + R_{13}(t_2 - t_1) R_{24}(t_2 + v - t_1 - u) + R_{14}(t_2 + v - t_1) R_{23}(t_2 - t_1 - u) . \end{aligned}$$

Substituting

$$R_{jk}(v) = \int_{-\pi}^{\pi} f_{jk}(\lambda) e^{iv\lambda} d\lambda$$

and using Lemma 1, the central moment is approximated for large T by

$$\begin{aligned} & \frac{1}{T} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} 2\pi D_T(\lambda+\mu) \{f_{13}(\lambda) f_{24}(\mu) K_M(\mu-\omega) K_M(\mu+\omega) \\ & \quad + f_{14}(\lambda) f_{23}(\mu) K_M(\lambda-\omega) K_M(\mu+\omega)\} d\lambda d\mu \\ & \quad + \int_{-\pi}^{\pi} f_{12}(\lambda) K_M(\lambda-\omega) d\lambda \int_{-\pi}^{\pi} f_{34}(\mu) K_M(\mu-\omega) d\mu - f_{12}(\omega) f_{34}(\omega) \end{aligned}$$

where $D_T(\theta) = \frac{1}{2\pi} \sum_{\tau=-T}^T e^{i\tau\theta}$. Now, $D_T(\lambda+\mu)$ is the dominating function in the double integrand, since $D_T(0) = O(T)$ while $K_M(0) = O(M)$. The function $D_T(\lambda+\mu)$ is a delta function at $\mu = -\lambda$, since

$$\int_{-\pi}^{\pi} D_T(\lambda+\mu) d\mu = 1 .$$

If $T \rightarrow \infty$ "faster" than M , the central moment is approximately equal to

$$\begin{aligned} \frac{2\pi}{T} \int_{-\pi}^{\pi} \{f_{13}(\lambda) f_{24}(-\lambda) K_M(-\lambda-\omega) K_M(-\lambda+\omega) + f_{14}(\lambda) f_{23}(-\lambda) K_M(\lambda-\omega) K_M(-\lambda+\omega)\} d\lambda \\ + \int_{-\pi}^{\pi} f_{12}(\lambda) K_M(\lambda-\omega) d\lambda \int_{-\pi}^{\pi} f_{34}(\mu) K_M(\mu-\omega) d\mu - f_{12}(\omega) f_{34}(\omega) . \end{aligned}$$

But since also $M \rightarrow \infty$; then $K_M(\varphi)$ is a delta function at $\varphi = 0$, so that the mean correction terms cancel and the limiting value is

$$\frac{2\pi}{T} f_{14}(\omega) f_{23}(-\omega) \int_{-\pi}^{\pi} K_M^2(\lambda-\omega) d\lambda \quad |\omega| \neq 0, \pi ,$$

$$\frac{2\pi}{T} \{f_{14}(\omega) f_{23}(-\omega) + f_{13}(\omega) f_{24}(-\omega)\} \int_{-\pi}^{\pi} K_M^2(\lambda-\omega) d\lambda \quad |\omega| = 0, \pi .$$

Since

$$\int_{-\pi}^{\pi} K_M^n(\lambda-\omega) d\lambda \doteq M^{n-1} \int_{-\infty}^{\infty} K^n(\lambda) d\lambda, \quad n = 1, 2, \dots$$

we can write finally

$$\begin{aligned} E[(\hat{f}_{12}(\omega) - f_{12}(\omega))(\hat{f}_{34}(\omega) - f_{34}(\omega))] / \frac{2\pi M}{T} \int_{-\infty}^{\infty} K^2(\lambda) d\lambda \\ \doteq f_{14}(\omega) f_{32}(\omega) \quad |\omega| \neq 0, \pi \\ \doteq f_{14}(\omega) f_{32}(\omega) + f_{13}(\omega) f_{24}(\omega) \quad |\omega| = 0, \pi . \end{aligned}$$

These formulae agree with those given by Rosenblatt (1959, p. 256).

The case $n = 3$. We begin by noting the identity

$$\mu_{ijk}^i \equiv \mu_{ijk}^i + \mu_i^i \mu_{jk} + \mu_j^i \mu_{ik} + \mu_k^i \mu_{ij} + \mu_i^i \mu_j^i \mu_k^i$$

where μ^i is the ordinary and μ the central moment of random variables

indicated by the subscripts.

Applying Isserlis's formula as before we obtain

$$\begin{aligned}
& \mu_{12,34,56} \\
& \doteq \left(\frac{1}{2\pi}\right)^3 \sum_{|u| \leq M} \sum_{|v| \leq M} \sum_{|s| \leq M} e^{-i(u+v+s)\omega} k\left(\frac{u}{M}\right) k\left(\frac{v}{M}\right) k\left(\frac{s}{M}\right) \frac{1}{T^3} \sum_{t_1=1}^{T-|u|} \sum_{t_2=1}^{T-|v|} \sum_{t_3=1}^{T-|s|} \\
& \left\{ R_{12}(u) [R_{34}(v) R_{56}(s) + R_{35}(t_3-t_2) R_{46}(t_3+s-t_2-v) \right. \\
& \quad \left. + R_{36}(t_3+s-t_2) R_{45}(t_3-t_2-v)] \right. \\
& + R_{13}(t_2-t_1) [R_{24}(t_2+v-t_1-u) R_{56}(s) + R_{25}(t_3-t_1-u) R_{46}(t_3+s-t_2-v) \\
& \quad \left. + R_{26}(t_3+s-t_1-u) R_{45}(t_3-t_2-v)] \right. \\
& + R_{14}(t_2+v-t_1) [R_{23}(t_2-t_1-u) R_{56}(s) + R_{25}(t_3-t_1-u) R_{36}(t_3+s-t_2) \\
& \quad \left. + R_{26}(t_3+s-t_1-u) R_{35}(t_3-t_2)] \right. \\
& + R_{15}(t_3-t_1) [R_{23}(t_2-t_1-u) R_{46}(t_3+s-t_2-v) + R_{24}(t_2+v-t_1-u) R_{36}(t_3+s-t_2) \\
& \quad \left. + R_{26}(t_3+s-t_2-u) R_{34}(v)] \right. \\
& \left. + R_{16}(t_3+s-t_1) [R_{23}(t_2-t_1-u) R_{45}(t_3-t_2-v) + R_{24}(t_2+v-t_1-u) R_{35}(t_3-t_2) \right. \\
& \quad \left. + R_{25}(t_3-t_1-u) R_{34}(v)] \right\} \\
& - \mu_{12}^i \mu_{34,56} - \mu_{34}^i \mu_{12,56} - \mu_{56}^i \mu_{12,34} - \mu_{12}^i \mu_{34,56} .
\end{aligned}$$

On substituting

$$R_{jk}(v) = \int_{-\pi}^{\pi} f_{jk}(\lambda) e^{iv\lambda} d\lambda$$

and using Lemma 1 for the terms where its condition of validity is satisfied we obtain the approximation for large T

$$\begin{aligned}
\mu_{12,34,56} &\doteq \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d\lambda d\mu dv \\
&\left\{ f_{12}(\lambda) K_M(\lambda-\omega) [f_{34}(\mu) K_M(\mu-\omega) f_{56}(v) K_M(v-\omega) \right. \\
&\quad + f_{35}(\mu) f_{46}(v) K_M(v+\omega) K_M(v-\omega) \frac{2\pi}{T} D_T(\mu+v) \\
&\quad + f_{36}(\mu) K_M(\mu-\omega) f_{45}(v) K_M(v+\omega) \frac{2\pi}{T} D_T(\mu+v)] \\
&+ f_{13}(\lambda) K_M(\mu+\omega) \frac{2\pi}{T} D_T(\lambda+\mu) [f_{24}(\mu) K_M(\mu-\omega) f_{56}(v) K_M(v-\omega) \frac{2\pi}{T} D_T(\lambda+\mu) \\
&\quad + f_{25}(\mu) f_{46}(v) K_M(v+\omega) K_M(v-\omega) \frac{2\pi}{T} D_T(v-\lambda) \\
&\quad + f_{26}(\mu) K_M(\mu-\omega) f_{45}(v) K_M(v+\omega) \frac{2\pi}{T} D_T(v-\lambda)] \\
&+ f_{14}(\lambda) K_M(\lambda-\omega) K_M(\mu+\omega) \frac{2\pi}{T} D_T(\lambda+\mu) [f_{23}(\mu) f_{56}(v) K_M(v-\omega) \\
&\quad + f_{25}(\mu) f_{36}(v) K_M(v-\omega) \frac{2\pi}{T} D_T(v-\lambda) \\
&\quad + f_{26}(\mu) K_M(\mu-\omega) f_{35}(v) \frac{2\pi}{T} D_T(v-\lambda)] \\
&+ f_{15}(\lambda) K_M(\mu+\omega) K_M(v-\omega) \frac{2\pi}{T} D_T(\lambda+\mu) [f_{23}(\mu) f_{46}(v) K_M(v+\omega) \frac{2\pi}{T} D_T(\mu-v) \\
&\quad + f_{24}(\mu) K_M(\mu-\omega) f_{36}(v) \frac{2\pi}{T} D_T(\mu-v) \\
&\quad + f_{26}(\mu) K_M(\mu-\omega) k_{34}(v)] \\
&+ f_{16}(\lambda) K_M(\lambda-\omega) K_M(\mu+\omega) \frac{2\pi}{T} D_T(\lambda+\mu) [f_{23}(\mu) f_{45}(v) K_M(v+\omega) \frac{2\pi}{T} D_T(\mu-v) \\
&\quad + f_{24}(\mu) K_M(\mu-\omega) f_{35}(v) \frac{2\pi}{T} D_T(\mu-v) \\
&\quad + f_{25}(\mu) f_{34}(v) K_M(v-\omega)] \left. \right\} \\
&- \mu_{12}^i \mu_{34,56} - \mu_{34}^i \mu_{12,56} - \mu_{56}^i \mu_{12,34} - \mu_{12}^i \mu_{34}^i \mu_{56}^i .
\end{aligned}$$

A. Assume $|\omega| \neq 0, \pi$.

Of the 15 terms in braces, consider the 1st, 3rd, 7th and 15th.

Using the result for $n = 1, 2$, these terms are seen to be asymptotically

equal to

$$\mu_{12}^i \mu_{34}^i \mu_{56}^i + \mu_{12}^i \mu_{34,56}^i + \mu_{56}^i \mu_{12,34}^i + \mu_{34}^i \mu_{12,56}^i$$

and are therefore cancelled by the mean correction terms outside the braces. Next consider the 9 terms numbered 2, 4, 5, 6, 9, 10, 11, 12, 14. Each of these terms contains a pair of kernels such as $K_M(\nu+\omega) K_M(\nu-\omega)$ and is therefore asymptotically negligible if $|\omega| \neq 0, \pi$. Consider finally the 8th and 13th terms, which are

$$\left(\frac{2\pi}{T}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d\lambda \, d\mu \, d\nu \left\{ f_{14}(\lambda) f_{25}(\mu) f_{36}(\nu) D_T(\lambda+\mu) D_T(\nu-\lambda) K_M(\lambda-\omega) K_M(\mu+\omega) K_M(\nu-\omega) \right. \\ \left. + f_{16}(\lambda) f_{23}(\mu) f_{45}(\nu) D_T(\lambda+\mu) D_T(\mu-\nu) K_M(\lambda-\omega) K_M(\mu+\omega) K_M(\nu+\omega) \right\}.$$

We evaluate the first of these by the following procedure: the kernel $D_T(\lambda+\mu)$ picks out $\mu = -\lambda$, on integration with respect to μ , giving the approximation

$$\left(\frac{2\pi}{T}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{14}(\lambda) f_{25}(-\lambda) f_{36}(\nu) D_T(\nu-\lambda) K_M(\lambda-\omega) K_M(-\lambda+\omega) K_M(\nu-\omega) d\lambda \, d\nu ;$$

the kernel $D_T(\nu-\lambda)$ picks out $\nu = \lambda$, giving

$$\left(\frac{2\pi}{T}\right)^2 \int_{-\pi}^{\pi} f_{14}(\lambda) f_{25}(-\lambda) f_{36}(\lambda) K_M(\lambda-\omega) K_M(-\lambda+\omega) K_M(\lambda-\omega) d\lambda ;$$

the kernel $K_M(\lambda-\omega)$, which is symmetrical about ω , picks out $\lambda = \omega$, giving

$$\left(\frac{2\pi}{T}\right)^2 f_{14}(\omega) f_{25}(-\omega) f_{36}(\omega) \int_{-\pi}^{\pi} K_M^3(\lambda-\omega) d\lambda \\ \doteq f_{14}(\omega) f_{52}(\omega) f_{36}(\omega) \left(\frac{2\pi M}{T}\right)^2 \int_{-\infty}^{\infty} K^3(\lambda) d\lambda .$$

The second term is similarly found to give

$$\begin{aligned}
& \left(\frac{2\pi}{T}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f_{16}(\lambda) f_{23}(\mu) f_{45}(\mu) D_T(\lambda+\mu) K_M(\lambda-\omega) K_M(\mu+\omega) K_M(\mu+\omega) d\lambda d\mu \\
& \doteq \left(\frac{2\pi}{T}\right)^2 \int_{-\pi}^{\pi} f_{16}(\lambda) f_{23}(-\lambda) f_{45}(-\lambda) K_M(\lambda-\omega) K_M(-\lambda+\omega) K_M(-\lambda+\omega) d\lambda \\
& \doteq \left(\frac{2\pi}{T}\right)^2 f_{16}(\omega) f_{23}(-\omega) f_{45}(-\omega) \int_{-\pi}^{\pi} K_M^3(\lambda-\omega) d\lambda \\
& \doteq f_{16}(\omega) f_{32}(\omega) f_{54}(\omega) \left(\frac{2\pi M}{T}\right)^2 \int_{-\infty}^{\infty} K^3(\lambda) d\lambda .
\end{aligned}$$

B. Assume $|\omega| = 0$ or π .

The first term in braces still gives $\mu_{12}^i \mu_{34}^i \mu_{56}^i$, while the 2nd and 3rd together give $\mu_{12}^i \mu_{34,56}^i$, the 4th and 7th give $\mu_{56}^i \mu_{12,34}^i$ and the 12th and 15th give $\mu_{34}^i \mu_{12,56}^i$. These terms cancel with terms outside the braces. The remaining 8 terms all contribute to the result, which can be written down, remembering that

$$f_{jk}(\omega) = f_{jk}(-\omega) \quad \text{if } |\omega| = 0, \pi .$$

We have thus proved that

$$\begin{aligned}
& E[(\hat{f}_{12}(\omega) - f_{12}(\omega))(\hat{f}_{34}(\omega) - f_{34}(\omega))(\hat{f}_{56}(\omega) - f_{56}(\omega))] / \left(\frac{2\pi M}{T}\right)^2 \int_{-\infty}^{\infty} K^3(\lambda) d\lambda \\
& \doteq f_{14}(\omega) f_{52}(\omega) f_{36}(\omega) + f_{16}(\omega) f_{32}(\omega) f_{54}(\omega) \quad |\omega| \neq 0, \pi \\
& \doteq f_{13}(\omega) f_{25}(\omega) f_{46}(\omega) + f_{13}(\omega) f_{26}(\omega) f_{45}(\omega) \\
& + f_{14}(\omega) f_{25}(\omega) f_{36}(\omega) + f_{14}(\omega) f_{26}(\omega) f_{35}(\omega) \\
& + f_{15}(\omega) f_{23}(\omega) f_{46}(\omega) + f_{15}(\omega) f_{24}(\omega) f_{36}(\omega) \\
& + f_{16}(\omega) f_{23}(\omega) f_{45}(\omega) + f_{16}(\omega) f_{24}(\omega) f_{35}(\omega) \quad |\omega| = 0, \pi .
\end{aligned}$$

The case $n = 4$. We do not write out this case fully for the pattern of derivation is already apparent. The moment identity becomes

$$\begin{aligned} \mu'_{ijkl} \equiv & \mu_{ijkl} + (\mu'_1 \mu'_{jkl} + \dots + \mu'_l \mu'_{ijk}) \\ & + (\mu'_1 \mu'_j \mu'_{kl} + \dots + \mu'_k \mu'_l \mu'_{ij}) + \mu'_i \mu'_j \mu'_k \mu'_l . \end{aligned}$$

Isserlis's formula now has 105 terms each being a product of 4 covariances. We imagine the expression in braces for $n = 3$ being written for $n = 4$ with the same convention for the order of terms, and evaluate it in the same manner.

A. Assume $|\omega| \neq 0, \pi$.

The terms appearing in the following places are cancelled by mean correction terms outside the braces: 1, 3, ..., 103 (15 terms). The following terms approach zero: 2, ..., 104 (81 terms). The following terms have a non-zero limit: 33, ..., 105 (9 terms).

Care is needed in evaluating these nine terms. In the first one, the quadruple integral factors into a product of two double integrals each of which is evaluated as in the case $n = 2$:

$$\begin{aligned} & \left(\frac{1}{2\pi T}\right)^2 T(2\pi)^3 \iint f_{14}(\lambda) f_{23}(\mu) D_T(\lambda+\mu) K_M(\lambda-\omega) K_M(\mu+\omega) d\lambda d\mu \\ & \cdot \left(\frac{1}{2\pi T}\right)^2 T(2\pi)^3 \iint f_{58}(\xi) f_{67}(\theta) D_T(\xi+\theta) K_M(\xi-\omega) K_M(\theta+\omega) d\xi d\theta \\ & \doteq \left(\frac{2\pi M}{T} \int K^2(\lambda) d\lambda\right)^2 f_{14}(\omega) f_{32}(\omega) f_{58}(\omega) f_{76}(\omega) . \end{aligned}$$

The fifth and ninth are of the same type. The remainder are like the second one:

$$\begin{aligned}
& \left(\frac{1}{2\pi T}\right)^4 T(2\pi)^7 \iiint f_{14}(\lambda) f_{25}(\mu) f_{38}(\xi) f_{67}(\theta) D_T(\lambda-\xi) D_T(\mu-\theta) D_T(\xi+\theta) \\
& \quad K_M(\lambda-\omega) K_M(\mu+\omega) K_M(\xi-\omega) K_M(\theta+\omega) d\lambda d\mu d\xi d\theta \\
& \doteq \left(\frac{2\pi M}{T}\right)^3 \int K^4(\lambda) d\lambda f_{14}(\omega) f_{52}(\omega) f_{38}(\omega) f_{76}(\omega) .
\end{aligned}$$

B. Assume $|\omega| = 0$ or π .

The following terms are cancelled by mean correction terms:

1, ..., 104 (45 terms). The following terms have a non-zero limit:

17, ..., 105 (60 terms). The 60 terms are evaluated by the method used above. We obtain finally (suppressing the frequency ω)

$$\begin{aligned}
& E[(\hat{f}_{12} - f_{12})(\hat{f}_{34} - f_{34})(\hat{f}_{56} - f_{56})(\hat{f}_{78} - f_{78})] \\
& \doteq \left(\frac{2\pi M}{T}\right)^2 \int K^2(\lambda) d\lambda \{f_{14}f_{32}f_{58}f_{76} + f_{16}f_{52}f_{38}f_{74} + f_{18}f_{72}f_{36}f_{54}\} \\
& \quad + \left(\frac{2\pi M}{T}\right)^3 \int K^4(\lambda) d\lambda \{f_{14}f_{52}f_{38}f_{76} + f_{14}f_{72}f_{36}f_{58} + f_{16}f_{32}f_{74}f_{58} \\
& \quad \quad \quad + f_{16}f_{72}f_{38}f_{54} + f_{18}f_{32}f_{54}f_{76} + f_{18}f_{52}f_{36}f_{74}\} \\
& \quad \quad \quad \text{if } |\omega| \neq 0, \pi \\
& \doteq \left(\frac{2\pi M}{T}\right)^2 \int K^2(\lambda) d\lambda \{(f_{13}f_{24} + f_{14}f_{23})(f_{57}f_{68} + f_{58}f_{67}) \\
& \quad \quad \quad + (f_{15}f_{26} + f_{16}f_{25})(f_{37}f_{48} + f_{38}f_{47}) \\
& \quad \quad \quad + (f_{17}f_{28} + f_{18}f_{27})(f_{35}f_{46} + f_{36}f_{45})\} \\
& \quad + \left(\frac{2\pi M}{T}\right)^3 \int K^4(\lambda) d\lambda \{f_{13}f_{25}f_{47}f_{68} + \dots + f_{18}f_{26}f_{37}f_{45}\} \\
& \quad \quad \quad \text{if } |\omega| = 0, \pi .
\end{aligned}$$

The number of terms in the last braces is 48. For comparison of these results with the general recursion formula in the statement of this theorem, note that, for example, the term in $\left(\frac{2\pi M}{T}\right)^2 \int K^2(\lambda) d\lambda$ for $|\omega| \neq 0, \pi$ is equal to

$$\mu_{12,34}\mu_{56,78} + \mu_{12,56}\mu_{34,78} + \mu_{12,78}\mu_{34,56} .$$

The cases $n = 5, 6, 7$. Proceeding in the same way we obtain the following results, suppressing the frequency ω :

$$\begin{aligned}
 E[(\hat{f}_{12}^{-f_{12}}) \dots (\hat{f}_9 10^{-f_9} 10)] &= \left(\frac{2\pi M}{T}\right)^3 \int K^2(\lambda) d\lambda \int K^3(\lambda) d\lambda \\
 &\quad \{f_{14}^{f_{32}} (f_{58}^{f_{96}} f_7 10^{+f_5} 10^{f_{76}} f_{98}) \\
 &\quad + f_{16}^{f_{52}} (f_{38}^{f_{94}} f_7 10^{+f_3} 10^{f_{74}} f_{98}) \\
 &\quad + f_{18}^{f_{72}} (f_{36}^{f_{94}} f_5 10^{+f_3} 10^{f_{54}} f_{96}) \\
 &\quad + f_1 10^{f_{92}} (f_{36}^{f_{74}} f_{58}^{+f_{38}} f_{54} f_{76})\} \\
 &\quad + o\left(\left(\frac{M}{T}\right)^4\right) \quad \text{if } |\omega| \neq 0, \pi.
 \end{aligned}$$

$$\begin{aligned}
 E[(\hat{f}_{12}^{-f_{12}}) \dots (\hat{f}_{11} 12^{-f_{11}} 12)] &= \left(\frac{2\pi M}{T}\right)^3 \int K^2(\lambda) d\lambda \\
 &\quad \{f_{14}^{f_{32}} (f_{58}^{f_{76}} f_9 12^{f_{11}} 10^{+f_5} 10^{f_{96}} f_7 12^{f_{11}} 8^{+f_5} 12^{f_{11}} 6^f 7 10^{f_{98}}) \\
 &\quad + f_{16}^{f_{52}} (f_{38}^{f_{74}} f_9 12^{f_{11}} 10^{+f_3} 10^{f_{94}} f_7 12^{f_{11}} 8^{+f_3} 12^{f_{11}} 4^f 7 10^{f_{98}}) \\
 &\quad + f_{18}^{f_{72}} (f_{36}^{f_{54}} f_9 12^{f_{11}} 10^{+f_3} 10^{f_{94}} f_5 12^{f_{11}} 6^{+f_3} 12^{f_{11}} 4^f 5 10^{f_{96}}) \\
 &\quad + f_1 10^{f_{92}} (f_{36}^{f_{54}} f_7 12^{f_{11}} 8^{+f_{38}} f_{74} f_5 12^{f_{11}} 6^{+f_3} 12^{f_{11}} 4^f 58^f 76) \\
 &\quad + f_1 12^{f_{11}} 2 (f_{36}^{f_{54}} f_7 10^{f_{98}} f_{38}^{f_{74}} f_5 10^{f_{96}} f_3 10^{f_{94}} f_{58} f_{76})\} \\
 &\quad + o\left(\left(\frac{M}{T}\right)^4\right) \quad \text{if } |\omega| \neq 0, \pi.
 \end{aligned}$$

$$E[(\hat{f}_{12}^{-f_{12}}) \dots (\hat{f}_{13} 14^{-f_{13}} 14)] = o\left(\left(\frac{M}{T}\right)^4\right), \quad |\omega| \leq \pi.$$

General n. A study of the derivation for $n = 3$ and $n = 4$ shows the origin of each term in the resulting formula; the formula for general n can then be written down as in the statement of the theorem.

The identity with which we began the proof for $n = 3, 4$ becomes

$$\mu_{1 \dots n}' \equiv \mu_{1 \dots n} + \mu_1' \dots \mu_n' + \sum_{r=1}^{n-2} \mu_{i_1}' \dots \mu_{i_r}' \mu_{i_{r+1}} \dots \mu_{i_n}$$

where $\{i_1, \dots, i_n\}$ is a permutation of $\{1, \dots, n\}$.

Isserlis's formula has $1 \cdot 3 \cdot 5 \dots (2n-1) = \frac{(2n)!}{n! 2^n}$ terms. The number of terms contributing to the general formula is

$$\sum_{a=0}^n (-1)^a \binom{n}{a} (n-a)! \quad |\omega| \neq 0, \pi$$

$$\sum_{a=0}^n (-1)^a \binom{n}{a} \frac{(2(n-a))!}{(n-a)! 2^{n-a}} \quad |\omega| = 0, \pi.$$

The expression for $|\omega| = 0, \pi$ is obtained as the number of ways of pairing the integers $\{1, \dots, 2n\}$ so that no pair contains an odd integer together with the integer exceeding it by 1, and for $|\omega| \neq 0, \pi$ so that in addition each pair contains one even and one odd integer. Given the first round draw for a competition, these expressions give the number of possible second round draws in a round-robin tournament ($|\omega|=0, \pi$) and a match between two teams ($|\omega| \neq 0, \pi$). The values for $n = 1, \dots, 6$ are given in the following table.

Table 1

Number of terms in formula for $\mu_{12, \dots, 2n-1, 2n}$

n	$ \omega \neq 0, \pi$	$ \omega = 0, \pi$
1	0	0
2	1	2
3	2	8
4	9	60
5	44	544
6	265	6040

We note that

$$E[(\hat{f}_{12}(\omega) - f_{12}(\omega)) \dots (\hat{f}_{2n-1,2n}(\omega) - f_{2n-1,2n}(\omega))] = O((M/T)^x)$$

where

$$x = \left[\frac{n+1}{2} \right], \quad n = 2, 3, \dots$$

As a first application of Theorem 1 we can immediately write down moments of the estimated autospectral density, assuming $f_{jj}(\omega) \neq 0$.

Table 2

$$E[(\hat{f}_{jj}(\omega) - f_{jj}(\omega))^m] / f_{jj}^m(\omega)$$

m	$ \omega \neq 0, \pi$	$ \omega = 0, \pi$
1	0	0
2	$\frac{2\pi M}{T} \int K^2(\lambda) d\lambda$	$2 \frac{2\pi M}{T} \int K^2(\lambda) d\lambda$
3	$2 \left(\frac{2\pi M}{T}\right)^2 \int K^3(\lambda) d\lambda$	$8 \left(\frac{2\pi M}{T}\right)^2 \int K^3(\lambda) d\lambda$
4	$3 \left(\frac{2\pi M}{T}\right)^2 \int K^2(\lambda) d\lambda)^2$ $+ 6 \left(\frac{2\pi M}{T}\right)^3 \int K^4(\lambda) d\lambda$	$12 \left(\frac{2\pi M}{T}\right)^2 \int K^2(\lambda) d\lambda)^2$ $+ 48 \left(\frac{2\pi M}{T}\right)^3 \int K^4(\lambda) d\lambda$

Corollary. The case when some spectra are zero. If k of the n spectra appearing in any term of the right-hand side of the general formula for $E \left[\prod_{j=1}^n (\hat{f}_{2j-1,2j}(\omega) - f_{2j-1,2j}(\omega)) \right]$ are zero at ω and the spectra are sufficiently smooth, we proceed to improve these terms by analogy with the note to Lemma 2. The following changes are made:

- (a) Each of the k spectra $f(\omega)$ is replaced by $f''(\omega)$.

(b) The term is multiplied by $\left(\frac{1}{2M^2}\right)^k$.

(c) $\int_{-\infty}^{\infty} K^n(\lambda) d\lambda$ is replaced by $\int_{-\infty}^{\infty} \lambda^{2k} K^n(\lambda) d\lambda$.

(d) Further expansion of other terms in this manner may be needed to preserve the order of approximation.

In the case that ℓ of these k second derivatives are zero, we make the following further changes:

(a') Each of the ℓ second derivatives $f''(\omega)$ is replaced by $f^{(iv)}(\omega)$.

(b') The factor $\left(\frac{1}{2M^2}\right)^k$ in (b) is replaced by $\left(\frac{1}{2M^2}\right)^{k-\ell} \left(\frac{1}{4!M^4}\right)^\ell$.

(c') The factor $\int_{-\infty}^{\infty} \lambda^{2k} K^n(\lambda) d\lambda$ is replaced by $\int_{-\infty}^{\infty} \lambda^{2(k-\ell)+4\ell} K^n(\lambda) d\lambda$.

(d') = (d).

Higher order zeros of the spectra can be treated similarly.

If any of the spectra appearing on the left-hand side of the formula is zero, the formula is unchanged, except that the mean correction is now $E[\hat{f}_{2j-1, 2j}(\omega)]$ rather than $f_{2j-1, 2j}(\omega)$ for the mean correction terms are modified in the same manner when they appear positively and negatively and, therefore, still cancel out.

3.3 Mean and Variance of the Estimates of Coherence

Theorem 2 Let

$$\frac{1}{n} = \frac{2\pi M}{T} \int_{-\infty}^{\infty} K^2(\lambda) d\lambda$$

and let

$$K_0 = \frac{\int_{-\infty}^{\infty} K^3(\lambda) d\lambda}{\left(\int_{-\infty}^{\infty} K^2(\lambda) d\lambda\right)^2} - 1.$$

Then for stationary normal processes with sufficiently smooth spectra we have to order n^{-3} , that is, to order $(M/T)^3$: if $|\omega| \neq 0, \pi$

$$E[\hat{W}] = W + \frac{1}{n} (1-W)^2 + \frac{2}{n^2} (1-W)^2 (W-K_0)$$

$$\text{Var}[\hat{W}] = \frac{2}{n} W(1-W)^2 + \frac{1}{n^2} (1-W)^2 (1-2W(1+2K_0)+13W^2)$$

and in particular

$$E[\hat{W}|W=0] = \frac{1}{n} - \frac{2}{n^2} K_0$$

$$\text{Var}[\hat{W}|W=0] = \frac{1}{n^2}$$

while if $|\omega| = 0, \pi$

$$E[\hat{W}] = W + \frac{1}{n} (1-W)(1-2W) + \frac{2}{n^2} (1-W)(-2K_0+(3+4K_0)W-4W^2)$$

$$\text{Var}[\hat{W}] = \frac{4}{n} W(1-W)^2 + \frac{2}{n^2} (1-W)(1-2W(13+8K_0)+2W^2(25+8K_0)-52W^3)$$

and in particular

$$E[\hat{W}|W=0] = \frac{1}{n} - \frac{4}{n^2} K_0$$

$$\text{Var}[\hat{W}|W=0] = \frac{2}{n^2}$$

Notes:

1. For the interpretation of the quantities n and K_0 see Section 4.
2. The normality assumption is not essential; see the note to Theorem 1.
3. The different orders of magnitude of the variance of \hat{W} when $W = 0$ and when $W \neq 0$ has a counterpart in the estimation of the ordinary multiple correlation coefficient R^2 ; see Kendall and Stuart (1961, p. 341).

Proof: We use the Taylor series expansion (suppressing the frequency ω)

$$\begin{aligned} E[\hat{W}] &= E[\hat{f}_{jk} \hat{f}_{kj} / \hat{f}_{jj} \hat{f}_{kk}] \\ &= W E[(1+g_{jk})(1+g_{kj})(1-g_{jj}+g_{jj}^2-\dots)(1-g_{kk}+g_{kk}^2-\dots)] \end{aligned}$$

where

$$g_{jk} \equiv (\hat{f}_{jk} - f_{jk}) / f_{jk},$$

valid if $W \neq 0$. Since we wish to retain terms of order $(M/T)^2$, we retain terms of the fourth order in the g 's. Therefore, write

$$\begin{aligned} E[\hat{W}] &= W E[1 + g_{jk} + g_{kj} - g_{jj} - g_{kk} \\ &\quad + g_{jj}^2 + g_{jj}g_{kk} + g_{kk}^2 - (g_{jk}+g_{kj})(g_{jj}+g_{kk}) + g_{jk}g_{kj} \\ &\quad - (g_{jj}^3 + g_{jj}^2g_{kk} + g_{jj}g_{kk}^2 + g_{kk}^3) + (g_{jk}+g_{kj})(g_{jj}^2 + g_{jj}g_{kk} + g_{kk}^2) \\ &\quad - (g_{jj}+g_{kk})g_{jk}g_{kj} \\ &\quad + g_{jj}^4 + g_{jj}^3g_{kk} + g_{jj}^2g_{kk}^2 + g_{jj}g_{kk}^3 + g_{kk}^4 \\ &\quad - (g_{jk}+g_{kj})(g_{jj}^3 + g_{jj}^2g_{kk} + g_{jj}g_{kk}^2 + g_{kk}^3) \\ &\quad + g_{jk}g_{kj}(g_{jj}^2 + g_{jj}g_{kk} + g_{kk}^2)] + O((M/T)^3). \end{aligned}$$

The expectations needed here can be read off from Theorem 1, in particular from the special cases stated fully in the course of the proof of that theorem. These expectations are listed in Table 3 for $|\omega| \neq 0, \pi$ and Table 4 for $|\omega| = 0, \pi$. We use the notation

$$\begin{aligned} N_1 &\equiv \frac{2\pi M}{T} \int_{-\infty}^{\infty} K^2(\lambda) d\lambda \\ N_2 &\equiv \left(\frac{2\pi M}{T}\right)^2 \int_{-\infty}^{\infty} K^3(\lambda) d\lambda. \end{aligned}$$

Thus we have the relations

$$\frac{1}{n} = N_1$$

$$K_0 = (N_2/N_1^2) - 1 .$$

As an example of the way in which Tables 3 and 4 were constructed, consider

$$E[g_{jk}g_{kj}] = E\left[\frac{\hat{f}_{jk} - f_{jk}}{f_{jk}} \cdot \frac{\hat{f}_{kj} - f_{kj}}{f_{kj}}\right]$$

$$= \mu_{jk,kj} / f_{jk} f_{kj}$$

$$= N_1 f_{jj} f_{kk} / f_{jk} f_{kj} \quad \text{if } |\omega| \neq 0, \pi$$

$$= N_1 / W ,$$

assuming $W \neq 0$. On inserting these expectations in the Taylor series expansion we obtain

$$E[\hat{W}] = W + N_1(1-W)^2 - 2N_2(1-W)^2$$

$$+ 2N_1^2(1-W)^2(1+W) + o((M/T)^3), \quad |\omega| \neq 0, \pi$$

$$= W + N_1(1-W)(1-2W) - 4N_2(1-W)(1-2W)$$

$$+ 2N_1^2(1-W)(2-W-4W^2) + o((M/T)^3), \quad |\omega| = 0, \pi ,$$

from which the mean values stated in the theorem follow. If $W = 0$, the Taylor series expansion becomes

$$E[\hat{W}] = (1/f_{jj} f_{kk}) E[\hat{f}_{jk} \hat{f}_{kj} (1-g_{jj} + g_{jj}^2 - \dots)(1-g_{kk} + g_{kk}^2 - \dots)]$$

but the results for this case can be obtained simply by substituting $W = 0$ in the results obtained for $W \neq 0$.

Table 3. Expected Values of Functions of Spectra, $|\omega| \neq 0, \pi$

Function	Expectation
$\varepsilon_{jj}, \varepsilon_{jk}$	0
$\varepsilon_{jj}^2, \varepsilon_{jj}\varepsilon_{jk}, \varepsilon_{jk}^2$	N_1
$\varepsilon_{jj}\varepsilon_{kk}$	$N_1 W$
$\varepsilon_{jk}\varepsilon_{kj}$	N_1/W
$\varepsilon_{jj}^3, \varepsilon_{jj}^2\varepsilon_{jk}, \varepsilon_{jj}\varepsilon_{jk}^2$	$2N_2$
$\varepsilon_{jj}^2\varepsilon_{kk}$	$2N_2 W$
$\varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}$	$N_2(1+W)$
$\varepsilon_{jj}\varepsilon_{jk}\varepsilon_{kj}$	$N_2(1+1/W)$
$\varepsilon_{jk}^2\varepsilon_{kj}$	$2N_2/W$
$\varepsilon_{jj}^4, \varepsilon_{jj}^3\varepsilon_{jk}, \varepsilon_{jj}^2\varepsilon_{jk}^2, \varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}\varepsilon_{kj}$	$3N_1^2$
$\varepsilon_{jj}^3\varepsilon_{kk}$	$3N_1^2 W$
$\varepsilon_{jj}^2\varepsilon_{kk}^2$	$N_1^2(1+2W^2)$
$\varepsilon_{jj}^2\varepsilon_{kk}\varepsilon_{jk}, \varepsilon_{jj}^2\varepsilon_{kk}\varepsilon_{kj}$	$N_1^2(1+2W)$
$\varepsilon_{jj}^2\varepsilon_{jk}\varepsilon_{kj}$	$N_1^2(2+1/W)$
$\varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}^2$	$N_1^2(2+W)$
$\varepsilon_{jj}\varepsilon_{jk}\varepsilon_{kj}^2, \varepsilon_{jj}\varepsilon_{jk}\varepsilon_{kj}^2$	$N_1^2(1+2/W)$
$\varepsilon_{jk}^2\varepsilon_{kj}^2$	$N_1^2(1+2/W^2)$

Table 4. Expected Values of Functions of Spectra, $|\omega| = 0, \pi$.

Function	Expectation
$\varepsilon_{jj}, \varepsilon_{jk}$	0
$\varepsilon_{jj}^2, \varepsilon_{jj}\varepsilon_{jk}$	$2N_1$
$\varepsilon_{jk}^2, \varepsilon_{jk}\varepsilon_{kj}$	$N_1(1+1/W)$
$\varepsilon_{jj}\varepsilon_{kk}$	$2N_1W$
$\varepsilon_{jj}^3, \varepsilon_{jj}^2\varepsilon_{jk}$	$8N_2$
$\varepsilon_{jj}\varepsilon_{jk}^2, \varepsilon_{jj}\varepsilon_{jk}\varepsilon_{kj}$	$2N_2(3+1/W)$
$\varepsilon_{jj}^2\varepsilon_{kk}$	$8N_2W$
$\varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}$	$4N_2(1+W)$
$\varepsilon_{jk}^2\varepsilon_{kj}$	$2N_2(1+3/W)$
$\varepsilon_{jj}^4, \varepsilon_{jj}^3\varepsilon_{jk}$	$12N_1^2$
$\varepsilon_{jj}^3\varepsilon_{kk}$	$12N_1^2W$
$\varepsilon_{jj}^2\varepsilon_{kk}^2$	$4N_1^2(1+2W^2)$
$\varepsilon_{jj}^2\varepsilon_{jk}, \varepsilon_{jj}^2\varepsilon_{jk}\varepsilon_{kj}$	$2N_1^2(5+1/W)$
$\varepsilon_{jj}^2\varepsilon_{kk}\varepsilon_{jk}, \varepsilon_{jj}^2\varepsilon_{kk}\varepsilon_{kj}$	$4N_1^2(1+2W)$
$\varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}\varepsilon_{kj}, \varepsilon_{jj}\varepsilon_{kk}\varepsilon_{jk}^2$	$2N_1^2(5+W)$
$\varepsilon_{jj}\varepsilon_{jk}^2\varepsilon_{kj}, \varepsilon_{jj}\varepsilon_{jk}\varepsilon_{kj}^2$	$6N_1^2(1+1/W)$
$\varepsilon_{jk}^2\varepsilon_{kj}^2$	$3N_1^2(1 + 2/W + 1/W^2)$

For $E[\hat{W}^2]$ we use the square of the Taylor expansion given above, again retaining terms of the fourth order in the g 's. We have

$$\begin{aligned}
E[\hat{W}^2] = W^2 E[& 1 + 2(g_{jk} + g_{kj} - g_{jj} - g_{kk}) \\
& + 3g_{jj}^2 + 4g_{jj}g_{kk} + 3g_{kk}^2 - 4(g_{jk} + g_{kj})(g_{jj} + g_{kk}) \\
& + g_{jk}^2 + 4g_{jk}g_{kj} + g_{kj}^2 \\
& - 2(2g_{jj}^2 + 3g_{jj}^2g_{kk} + 3g_{jj}g_{kk}^2 + 2g_{kk}^3) \\
& + 2(g_{jk} + g_{kj})(3g_{jj}^2 + 4g_{jj}g_{kk} + 3g_{kk}^2) \\
& - 2(g_{jj} + g_{kk})(g_{jk}^2 + 4g_{jk}g_{kj} + g_{kj}^2) \\
& + 2(g_{jk}g_{kj} + g_{jk}g_{kj}^2) \\
& + 5g_{jj}^4 + 8g_{jj}^3g_{kk} + 9g_{jj}^2g_{kk}^2 + 8g_{jj}g_{kk}^3 + 5g_{kk}^4 \\
& - 4(g_{jk} + g_{kj})(2g_{jj}^3 + 3g_{jj}^2g_{kk} + 3g_{jj}g_{kk}^2 + 2g_{kk}^3) \\
& + (g_{jk}^2 + 4g_{jk}g_{kj} + g_{kj}^2)(3g_{jj}^2 + 4g_{jj}g_{kk} + 3g_{kk}^2) \\
& - 4(g_{jj} + g_{kk})(g_{jk}g_{kj} + g_{jk}g_{kj}^2) + g_{jk}^2g_{kj}^2] + o((M/T)^3).
\end{aligned}$$

By the same method we obtain

$$\begin{aligned}
E[\hat{W}^2] = W^2 + 4N_1W(1-W)^2 - 8N_2W(1-W)^2 \\
+ 2N_1^2(1-W)^2(1-2W+9W^2) + o((M/T)^3), \quad |\omega| \neq 0, \pi \\
= W^2 + 2N_1W(1-W)(3-4W) - 8N_2W(1-W)(3-4W) \\
+ 3N_1^2(1-W)(1-5W+24W^2-24W^3) + o((M/T)^3), \quad |\omega| = 0, \pi,
\end{aligned}$$

from which the variances stated in the theorem follow. The proof of Theorem 2 is now complete. It may be noted that in taking the Taylor series expansion for $E[\hat{W}]$ and $E[\hat{W}^2]$, allowance was made not only for

the variance of the numerator and denominator of \hat{W} , but also for the correlation between numerator and denominator.

3.4 Comparison with Goodman's Distribution

The quantity n in Theorem 2 is known as the number of degrees of freedom for the estimates; see, for example, Amos and Koopmans (1963, p. 26). It is not restricted to integral values.

The quantity K_0 reflects the extent to which the spectral window departs from being rectangular; its value is zero for a rectangular window. Some common windows are listed in Tables 5 and 6 with their formulae. This list corresponds to those given by Jenkins (1961, p. 146) and Parzen (1961, p. 178). The value of K_0 , together with the second, third and fourth moments, is given in Table 7 for each of these windows.

In a valuable thesis Goodman (1957) derived the approximate asymptotic distribution of the estimated spectral density matrix $\hat{f}(\omega)$, and from it obtained probability distributions for several functions of the matrix including the square root of coherence (p. 125), whose probability density is:

$$dp_{\sqrt{\hat{W}}}(z) = \frac{2(1-W)^n}{\Gamma(n)\Gamma(n-1)} z(1-z^2)^{n-2} \sum_{k=0}^{\infty} \frac{W^k \Gamma^2(n+k)}{\Gamma^2(k+1)} z^{2k} dz, \quad 0 \leq z \leq 1.$$

Table 5. Some Common Lag Windows.

Identification number	Name	Formula for $k(u)$, $ u \leq 1$
1	Bartlett (1)	1
2	Bartlett (2)	$1 - u $
3	Hamming	$\frac{1}{2}(1 + \cos \pi u)$
4	Hamming	$.54 + .46 \cos \pi u$
5	Parzen (1)	$1 - u^2$
6	Parzen (2)	$1 - 6u^2 + 6 u ^3$, $ u \leq \frac{1}{2}$ $2(1 - u)^3$, $\frac{1}{2} \leq u \leq 1$
7	Daniell	$\frac{\sin \pi u}{\pi u}$

Table 6. Spectral Window Generators.

Identification number	Formula for $K(\lambda)$, $-\infty < \lambda < \infty$
1	$(\sin \lambda)/\pi\lambda$
2	$2(\sin \frac{1}{2} \lambda)^2/\pi\lambda^2$
3	$\{.50(\sin \lambda)/\lambda + .25(\sin(\lambda+\pi))/(\lambda+\pi) + .25(\sin(\lambda-\pi))/(\lambda-\pi)\}/\pi$
4	$\{.54(\sin \lambda)/\lambda + .23(\sin(\lambda+\pi))/(\lambda+\pi) + .23(\sin(\lambda-\pi))/(\lambda-\pi)\}/\pi$
5	$2\{(\sin \lambda)/\lambda^3 - (\cos \lambda)/\lambda^2\}/\pi$
6	$3\{(\sin(\lambda/4))/(\lambda/4)\}^4/8\pi$
7	$1/2hM$, $-hM < \lambda < hM$ 0, otherwise

Table 7. Moments of Spectral Window Generators.

Identification number	$\int K^2(\lambda) d\lambda$	$\int K^3(\lambda) d\lambda$	$\int K^4(\lambda) d\lambda$	$K_0 \equiv$
				$\frac{\int K^3(\lambda) d\lambda}{(\int K^2(\lambda) d\lambda)^2} - 1$
1	.318126	.075991	.021501	-.249130
2	.106103	.013932	.001933	.237542
3	.119366	.015723	.002182	.103492
4	.126488	.017946	.002686	.121681
5	.169764	.029763	.005527	.032727
6	.085830	.008419	.000873	.142832
7	$1/2hM$	$(1/2hM)^2$	$(1/2hM)^3$	0

In the course of his derivation Goodman made many approximations, the key one being the replacement of the spectral window by a rectangular window whose height and width were chosen appropriately (see Amos and Koopmans, 1963, pp 24-25; in this reference extensive tables are given for the distribution of $\sqrt{\hat{W}_{jk}(\omega)}$). As a result of the numerous approximations and heuristic derivation, Goodman's distribution has not been fully trusted (see, for example, Granger and Hatanaka, 1964, pp 80, 295; note that the distribution function and table given by Granger on page 79 of this reference, purporting to refer to $\hat{W}_{jk}(\omega)$, actually refers to $\sqrt{\hat{W}_{jk}(\omega)}$ due to a misuse of the tables of Amos and Koopmans (1963)). We may also quote here from Jenkins (1963, p. 272): "...any confidence intervals for spectra should at best be regarded as rough guides because of the approximations and assumptions made in their derivation... ." However, Goodman's heuristic derivation has now been made rigorous by Wahba (1966), while we have in our Theorem 2 the means to examine the effect of the rectangular window assumption.

In a case of chief interest, namely, the case $W_{jk} \equiv 0$, Goodman's distribution becomes

$$dp_{W=0}(\omega) = (n-1)(1-\omega)^{n-2} d\omega, \quad 0 \leq \omega \leq 1$$

where we have made the transformation from the square root quantity to coherence as defined in Section 1.4; that is, $\hat{W}_{jk}(\omega)/(n-1)$ has the beta distribution $B(1, n-1)$ when $W_{jk}(\omega) \equiv 0$. From this distribution we derive

$$E[\hat{W}|W=0] = \frac{1}{n}$$

$$\text{Var}[\hat{W}|W=0] = (n-1)/n^2(n+1).$$

These values agree well with those obtained in Theorem 2 for n exceeding about 5 (depending on the window used). We conclude that Goodman's rectangular window assumption has led to reasonable results, in the case $W_{jk}(\omega) \equiv 0$. For very small values of n further accuracy may be desired and can be obtained by taking further terms in the Taylor series expansion in Theorem 2.

It should be mentioned that a first approximation (again using, in effect, the rectangular window assumption) to the variance of the estimated coherence has been given by Jenkins (1963, p. 275); however, his definition corresponds to our $\sqrt{W_{jk}(\omega)}$, so that no direct comparison is possible.

3.5 Testing for Zero Coherence.

The first comment that must be made in considering the application of Theorem 2 is that it can be useful only if the truncation point M is so large that

$$\hat{f}_{jk}(\omega) \doteq \sum_{|v| \leq M} R_{jk}(v) e^{-iv\omega}, \quad j = 1, 2; k = 1, 2.$$

This requirement applies to each of the autospectra as well as to the cross-spectrum. It is assumed that the same value of M , and the same spectral window, is used for all three estimates

$$\hat{f}_{jk}(\omega), \quad \hat{f}_{jj}(\omega), \quad \hat{f}_{kk}(\omega).$$

This will as a rule be wise only if both autospectra have been pre-whitened and the cross-spectrum has been preflattened. If this has been attempted, one can test the null hypothesis that $W_{jk}(\omega) \equiv 0$ by means of the test statistic

$$\mathcal{W}_{jk} = \frac{1}{Q} \left\{ \sum_{j=1}^{Q-1} \hat{W}_{jk}(\pi j/Q) + \frac{1}{2}(\hat{W}_{jk}(0) + \hat{W}_{jk}(\pi)) \right\}$$

where $Q + 1$ is the number of points at which estimates have been made.

The estimates $\hat{W}_{jk}(\pi j/Q)$ are not in general entirely independent at different frequencies. They are, however, asymptotically independent since kernels which have peaks at different frequencies almost annihilate each other on being multiplied together. For the test, which will be proposed, to be valid it is necessary that the kernels at neighboring frequencies do not overlap to any great extent; we assume that the chosen value of Q satisfies this requirement. Then according to the central limit theorem we have asymptotically

$$\left\{ Q/\text{Var}[\hat{W}_{jk}(\omega) | W_{jk}(\omega) \equiv 0] \right\}^{\frac{1}{2}} (\mathcal{W}_{jk} - E[\hat{W}_{jk}(\omega) | W_{jk}(\omega) \equiv 0]) \sim \mathcal{N}(0,1),$$

the effect of the different contributions from frequencies $0, \pm \pi$ being negligible in the limit. It should be noted that this asymptotic distribution applies when $T/Q \rightarrow \infty$ and $Q \rightarrow \infty$ simultaneously.

We, therefore, propose the test procedure: reject $H_0: W_{jk}(\omega) \equiv 0$ if

$$\mathcal{W}_{jk} > \frac{1}{n} - \frac{2}{n^2} K_0 + d_\alpha/n\sqrt{Q}$$

and accept H_0 otherwise, where d_α is the upper one-tail α percentage point of the standard normal distribution. The critical region then depends on the window $K(\cdot)$, the number n of degrees of freedom, the number $Q + 1$ of estimation points, and the level of significance α . The factors governing the choice of the significance level, which is at the disposal of the statistician, are the same as arise in the general

classical hypothesis testing problem.

To illustrate the test procedure a pair of incoherent white noise series was generated on a computer. The estimated coherence is shown in Figure 1. The details of the test procedure are as follows.

$$T = 200, \quad M = 40, \quad Q = 40, \quad K_0 = 0.1428 \text{ (Parzen window)}$$

$$n = 200 / (40 \times 2\pi \times 0.086) = 9.4$$

$$\hat{W}_{jk} = 0.1063.$$

Hence under the null hypothesis we have, for an observation on $\mathcal{N}(0,1)$, the number

$$\left\{ 0.1063 - \left(\frac{1}{9.4} - \frac{2}{(9.4)^2} \times 0.1428 \right) \right\} \times 9.4 \times \sqrt{40}$$

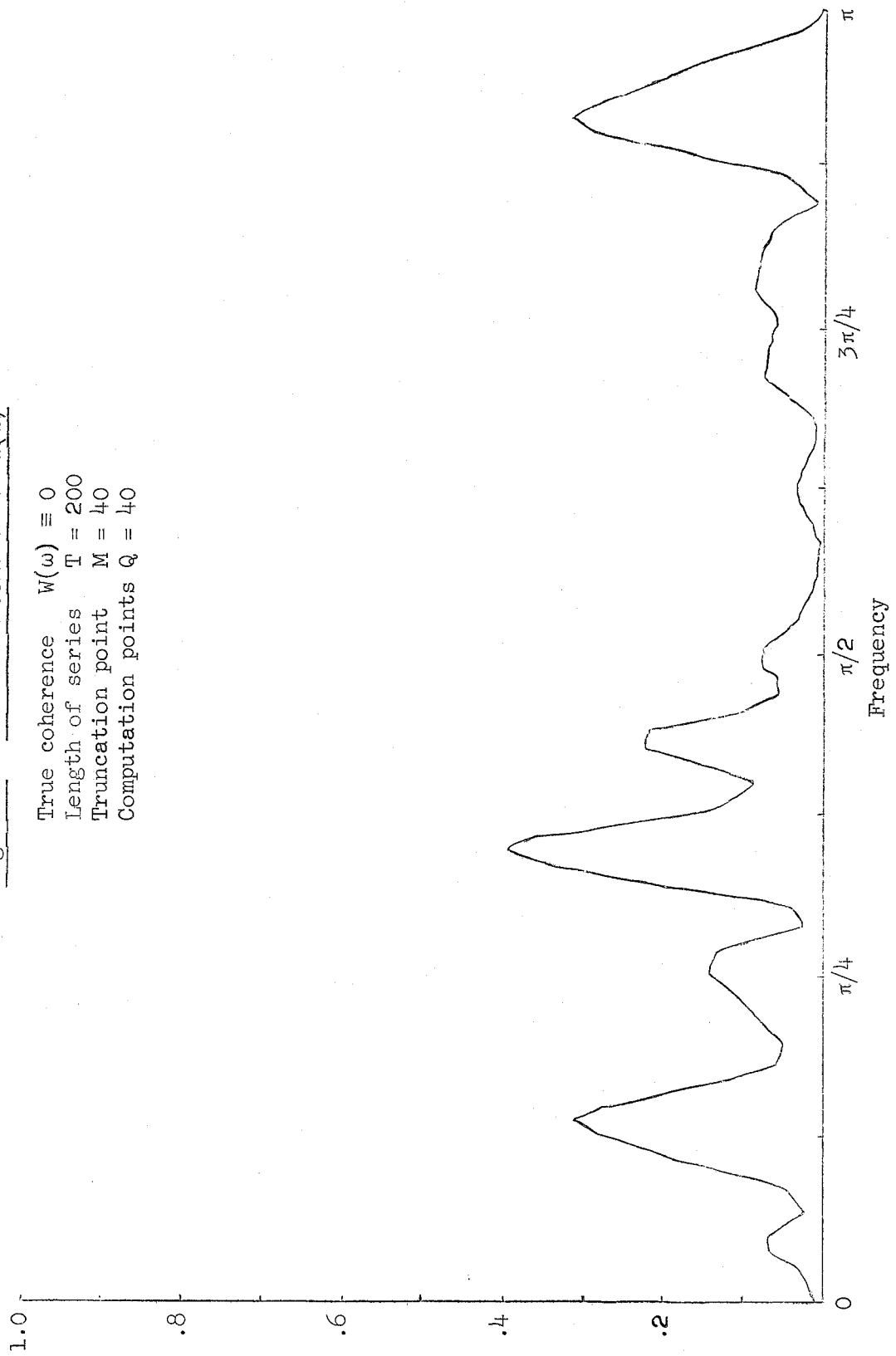
$$= 0.1843 .$$

Therefore, the null hypothesis is not rejected.

It is no doubt true that the measure of information \mathcal{I}_{jk} proposed by Gelfand and Yaglom and defined in Section 1.4 would provide a good test statistic for the hypothesis of zero coherence, but since its distribution is not known we have preferred the statistic \hat{W}_{jk} , which also seems natural and satisfactory.

It is also possible, of course, to test other null hypotheses about the true coherence; we have given the results in Theorem 2 for any values of $W_{jk}(\omega)$ partly to enable the power of the test of H_0 against specified alternatives to be computed and partly because if a bound \bar{M} on the greatest significant lag in $R_{jj}(v)$, $R_{kk}(v)$ and $R_{jk}(v)$ simultaneously is known (see Section 2.1 (iii)), then prewhitening and preflattening might on some occasions be dispensed with, so that

Figure 1. Estimated Coherence $\hat{W}(\omega)$



True coherence $W(\omega) \equiv 0$
Length of series $T = 200$
Truncation point $M = 40$
Computation points $Q = 40$

non zero true values of $W_{jk}(\omega)$ will be of interest.

Confidence intervals could be constructed for $\hat{W}_{jk}(\omega)$, $-\pi \leq \omega \leq \pi$, but the difficulty of interpreting a joint confidence interval applying to all frequencies in an interval, or to all of a grid of frequencies, is so great that we regard the use of such intervals with disfavor; this is an area where research might be considered desirable, to provide a statistical tool which can be used not only in time series analysis but also in other situations calling for a joint confidence band around a large set of estimates. A forthcoming paper by Van Ness and Woodroffe (1966) contains a partial result in this direction.

3.6 Shifting the Lag Window by the Estimated Phase Derivative

The method of cross-regressive filtering may not always be the best method of estimating the cross-spectrum and coherence, for it depends on the validity of the model stated in Section 3.1. It might be of value to have a procedure which is not based on any such model.

A different approach has been suggested by Parzen (1965, pp 46-48). The idea is that, since the central problem of cross-spectral estimation arises from the possibility that the true cross-spectrum is not flat compared to the spectral window used, or, expressing this in the time domain, that there is significant cross-correlation at lags v receiving a small weight $k(v/M)$, one would like to center the lag window $k(v/M)$ (see Section 2.2 (ii)) at the point v^* at which $|R_{jk}(v)|$ takes its maximum value. This will clearly be a very good procedure if all the significant relationships occur at lags close to v^* . Parzen suggested as a first approximation that v^* might be taken to be the estimate of

the derivative of the phase angle, using an argument occurring in an electrical engineering situation (Mason and Zimmerman, 1960, p. 367).

Parzen's (and Mason and Zimmerman's) analysis is based on a particular choice of an asymptotic expansion of a function of a complex variable (Parzen, 1965, p. 45):

$$\begin{aligned} f_{jk}(\lambda) &= f_{kk}(\omega) G_{j;k}(\omega) \exp[-i\{\varphi_{j;k}(\omega) + (\lambda-\omega) \varphi'_{j;k}(\omega)\}] \\ &= f_{jk}(\omega) \exp[-i(\lambda-\omega) \varphi'_{j;k}(\omega)] ; \end{aligned}$$

the notation is defined in Section 1.5. The limitation of this analysis is that this particular choice of expansion is relevant only if all significant relationships occur at lags close to v^* , for only then will the gain be relatively constant and the phase angle the chief source of variation of $f_{jk}(\omega)$ with respect to ω . This situation may arise frequently in electrical engineering and other fields, but it can hardly be considered to be the general case. The simplest case which illustrates our point (though this example is not chosen for realism) is

$$\begin{aligned} X_j(t) &= \epsilon(t) \\ X_k(t) &= \epsilon(t-L) + \epsilon(t+L) \end{aligned}$$

where $\epsilon(t)$ is white noise. In this case the phase and the phase derivative are both identically zero, so that no shifting will improve matters, and the bias in $\hat{f}_{jk}(\omega)$ will be $(k(L/M)-1) \times 100\%$ which is -100% if $L \geq M$.

However, the idea of shifting the lag window could contain the germ of a workable method, which we attempt to develop in the next section by

taking into account not only the most important lag v^* but all significant lags.

3.7 Best Linear Preflattening

In an attempt to obtain a general method of cross-spectral estimation which does not depend on the validity of the cross-regression model, we consider making a linear transformation of one of the two given series, in which the coefficients will be chosen so that the new cross-spectrum will be as flat as possible in an appropriate mean square sense. This is clearly a generalization of the shift method described in Section 3.6; in that case only one coefficient was allowed to be non-zero.

Let the series after transformation be

$$Y_j(t) = X_j(t)$$
$$Y_k(t) = \sum_{s=-p}^q A(s) X_k(t-s) .$$

The range of s should cover all significant lags of the cross-correlation function; since we no longer have the cross-regression model, the appropriate values of p and q will have to be judged by examining a graph of $\hat{R}_{jk}(v)$, in the absence of prior information; alternatively all lags available from the given length of the observed series could be included.

(i) A single transformation to serve for estimation at all frequencies.

We can attempt to make the cross-spectrum between $Y_j(t)$ and $Y_k(t)$ as flat as possible, in the sense that the variance is a minimum, by minimizing the quantity

$$\Phi = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{f}_{jk}^y(\lambda) - b|^2 d\lambda$$

where $b \neq 0$ is a real constant, upon which the solution will not depend; in this section superscripts will indicate the X or the Y series.

Thus

$$\Phi = \sum_v \left\{ \frac{1}{2\pi} \sum_s A(s) \hat{R}_{jk}^x(v-s) k\left(\frac{v}{M}\right) \right\}^2 - 2b \frac{1}{2\pi} \sum_s A(s) \hat{R}_{jk}^x(-s) + b^2 .$$

The minimum is achieved by differentiation, giving the linear system of equations

$$\sum_{s=-p}^q A(s) \sum_v \hat{R}_{jk}^x(v-s) \hat{R}_{jk}^x(v-u) k^2\left(\frac{v}{M}\right) = 2\pi b \hat{R}_{jk}^x(-u) .$$

Since the solution $A'(s)$ depends on b only to the extent of a multiplicative constant, we can take for the transformation

$$\hat{A}(s) = A'(s) / \sum_r |A'(r)|, \quad s = -p, \dots, q .$$

The preflattening transformation defined by the weights $\{\hat{A}(s)\}$ is designed to play the role which was played by cross-regressive filtering in Section 3.1. The estimated cross-spectral density function of the X series will then be

$$\tilde{f}_{jk}^x(\omega) = \hat{f}_{jk}^y(\omega) / \sum_s A(s) e^{-is\omega}, \quad |\omega| \leq \pi .$$

It may be noted that in the case of lagged white noise, $X_k(t) = X_j(t-L)$, if we replace the estimated cross-covariances by their true values, $R_{jk}^x(v) = \delta_{v,L}$, the solution is $\hat{A}(u) = \delta_{-u,L}$ so that the indicated transformation is

$$Y_k(t) = X_k(t+L) ,$$

as expected.

As another illustration, we compute the appropriate linear transformation for the example given in Section 3.6, namely,

$$\begin{aligned} X_j(t) &= \epsilon(t) \\ X_k(t) &= \epsilon(t-L) + \epsilon(t+L) \end{aligned}$$

where $\epsilon(t)$ is white noise with variance σ^2 . In this case

$$\begin{aligned} R_{jk}^x(v) &= \sigma^2 (\delta_{v-L} + \delta_{v+L}) \\ f_{jk}^x(\omega) &= \frac{\sigma^2}{\pi} \cos L\omega. \end{aligned}$$

The linear system of equations becomes

$$\begin{aligned} A(u) \{k^2((u+L)/M) + k^2((u-L)/M)\} + A(u-2L) k^2((u-L)/M) \\ + A(u+2L) k^2((u+L)/M) = (2\pi b/\sigma^2) \delta_{|u|-L}, \quad u = -p, \dots, q. \end{aligned}$$

The solution will depend on the window $k(\cdot)$ used, but the nature of the solution can be seen by replacing the window by unity. The system of equations then becomes

$$2A(u) + A(u-2L) + A(u+2L) = c \delta_{|u|-L}.$$

If $L = 1$ and $p = q = 2$, this becomes, in matrix form,

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} A(-2) \\ A(-1) \\ A(0) \\ A(1) \\ A(2) \end{bmatrix} = \begin{bmatrix} 0 \\ c \\ 0 \\ c \\ 0 \end{bmatrix}.$$

The solution is

$$A'(-1) = A'(1) = c/3, \quad A'(-2) = A'(0) = A'(2) = 0 ;$$

the general solution is therefore

$$A(-L) = A(L) = \frac{1}{2}, \quad A(s) = 0 \quad \text{if } |s| \neq L .$$

Thus the series after transformation are

$$\begin{aligned} Y_j(t) &= X_j(t) \\ Y_k(t) &= \frac{1}{2} X_k(t-L) + \frac{1}{2} X_k(t+L) , \end{aligned}$$

from which we compute

$$\begin{aligned} R_{jk}^Y(v) &= (\sigma^2/2)(\delta_{v-2L} + 2\delta_v + \delta_{v+2L}) \\ f_{jk}^Y(\omega) &= \frac{\sigma^2}{2\pi} (1 + \cos 2L\omega) . \end{aligned}$$

The new cross-correlation function is more favorable than the old (see Figure 2) leading to a smoother cross-spectrum (Figure 3).

(ii) A separate transformation for each frequency. To improve the estimates derived above we now restrict attention to the estimation problem at a given frequency ω . We again make a preflattening transformation of the given series, but we now improve the choice of the weights, say $A_\omega(s)$.

The contribution to the bias in the estimate $\hat{f}_{jk}^Y(\omega)$ resulting from all the frequencies ν , $-\pi \leq \nu \leq \pi$, over which the spectral window operates (see Section 2.2 (ii)) can now be minimized in a mean square sense by minimizing the quantity

$$\Psi_\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{f}_{jk}^Y(\nu) K(\omega-\nu) - bK(\omega-\nu)|^2 d\nu .$$

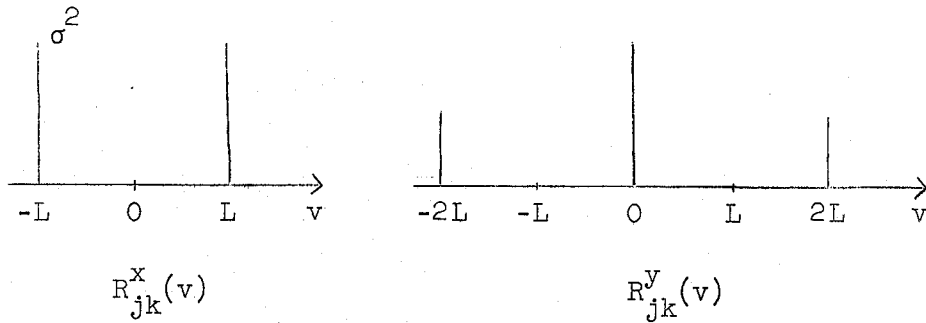


Figure 2.

Original and Transformed Cross-covariance Function

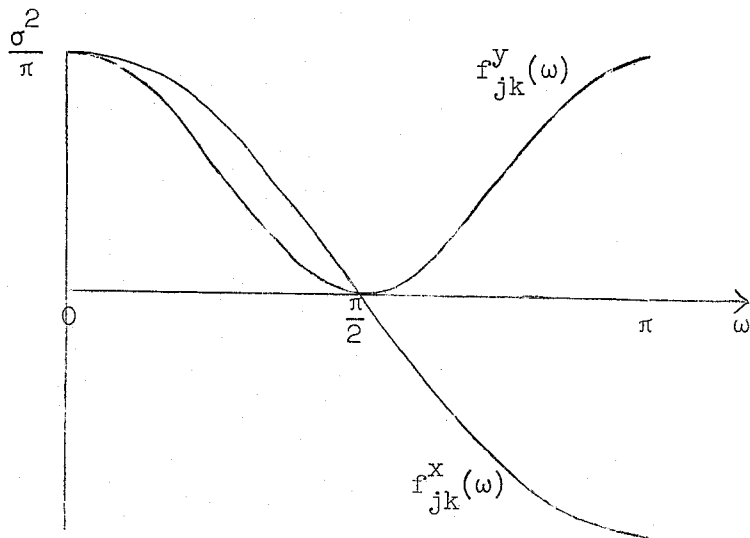


Figure 3.

Original and Transformed Cross-spectrum

Note that frequencies ν which are close to the given frequency ω are now given more weight than those distant from ω , in accordance with the weight function $K(\omega-\nu)$; this is appropriate since we desire that the spectral density be level in the region where the window is near its peak, while oscillations which are damped out by the window are not so important.

If the transformed series are now written

$$Y_j(t) = X_j(t)$$

$$Y_k(t) = \sum_s A_\omega(s) X_k(t-s),$$

we obtain in the same way as before the linear system of equations

$$\sum_s A_\omega(s) \left\{ \sum_{\nu \geq \nu'} L_\omega(\nu-\nu') \hat{R}_{jk}^x(\nu-s) \hat{R}_{jk}^x(\nu'-u) k^2\left(\frac{\nu}{M}\right) \right\}$$

$$= 2\pi b_\omega \sum_\nu \hat{R}_{jk}^x(\nu-u) L_\omega(\nu) k\left(\frac{\nu}{M}\right),$$

where $L_\omega(\nu) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(\omega-\mu) \cos \nu\mu \, d\mu$. If the solution to this system is $\{A'_\omega(s)\}$, then we take

$$\hat{A}_\omega(s) = A'_\omega(s) / \sum_r |A'_\omega(r)|.$$

Again in the (trivial) case of lagged white noise we obtain the expected solution $\hat{A}_\omega(u) = \delta_{-u,L}$, though in general the solution will depend on ω , and fairly extensive computations are required.

(iii) Separate transformations for the cospectrum and quadrature spectrum at each frequency. Instead of minimizing the mean squared modulus Ψ_ω , greater accuracy may be obtained by estimating the cospectrum and quadrature spectrum from two different transformations.

Let

$$Y_j(t) = X_j(t)$$

$$Y_k(t) = \sum_s A_{\omega,c}(s) X_k(t-s)$$

and let

$$Z_j(t) = X_j(t)$$

$$Z_k(t) = \sum_s A_{\omega,q}(s) X_k(t-s)$$

for $|\omega| < \pi$, and define

$$\Psi_{\omega,c} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\operatorname{Re} \hat{f}_{jk}^y(v) K(\omega-v) - b_{\omega,c} K(\omega-v)\}^2 dv$$

$$\Psi_{\omega,q} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\operatorname{Im} \hat{f}_{jk}^z(v) K(\omega-v) - b_{\omega,q} K(\omega-v)\}^2 dv .$$

To minimize $\Psi_{\omega,c}$ we have the linear system of equations

$$\begin{aligned} & \sum_s A_{\omega,c}(s) \left\{ \sum_{v \geq v'} \hat{R}_{jk}^x(v-s) \hat{R}_{jk}^x(v'-u) L_{\omega,c}(v, v') \right\} \\ & = 2\pi b_{\omega,c} \sum_v \hat{R}_{jk}^x(v-u) L_{\omega,c}(v) \end{aligned}$$

where

$$L_{\omega,c}(v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(\omega-\mu) \cos v\mu \, d\mu$$

and

$$L_{\omega,c}(v, v') = \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(\omega-\mu) \cos v\mu \cos v'\mu \, d\mu .$$

To minimize $\Psi_{\omega,q}$, set $A_{\omega,c}(-s^*) = 1$, where

$$|\hat{R}_{jk}^x(s^*)| = \max_s |\hat{R}_{jk}^x(s)| ;$$

this avoids the trivial solution $A_{\omega,q}(s) \equiv 0$. Then we have the linear

system

$$\sum_{s \neq s^*} A_{\omega, q}(s) \left\{ \sum_{v \geq v'} \hat{R}_{jk}^x(v-s) \hat{R}_{jk}^x(v'-u) L_{\omega, q}(v, v') \right\}$$

$$= 2\pi b_{\omega, q} \sum_{v \geq v'} \hat{R}_{jk}^x(v-s^*) \hat{R}_{jk}^x(v'-u) L_{\omega, q}(v, v')$$

where

$$L_{\omega, q}(v, v') = \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(\omega - \mu) \sin v\mu \sin v'\mu \, d\mu .$$

The weights $\hat{A}_{\omega, c}(s)$, $\hat{A}_{\omega, q}(s)$ are obtained by normalization as before.

One now computes

$$\hat{c}_{jk}^y(\omega) = \text{Re } \hat{r}_{jk}^y(\omega)$$

$$\hat{q}_{jk}^z(\omega) = \text{Im } \hat{r}_{jk}^z(\omega) ,$$

after which the following two equations yield estimates $\tilde{c}_{jk}^x(\omega)$, $\tilde{q}_{jk}^x(\omega)$; these equations involve a slight generalization of the concept of the transfer function:

$$\hat{c}_{jk}^y(\omega) = \sum_s \hat{A}_{\omega, c}(s) \cos \omega s \tilde{c}_{jk}^x(\omega) - \sum_s \hat{A}_{\omega, c}(s) \sin \omega s \tilde{q}_{jk}^x(\omega)$$

$$\hat{q}_{jk}^z(\omega) = \sum_s \hat{A}_{\omega, q}(s) \sin \omega s \tilde{c}_{jk}^x(\omega) + \sum_s \hat{A}_{\omega, q}(s) \cos \omega s \tilde{q}_{jk}^x(\omega) .$$

We have spelled out the details of this approach not because we seriously recommend it for the given problem, but to see where the shift method of the previous section leads when extended in an apparently logical way.

3.8 Alternative Estimating Functions for Coherence

The application of classical and decision-theoretic doctrines to the present problem is virtually excluded by the complexity of the

distribution of $\hat{W}(\omega)$ (see Section 3.4). In any case, the desirability of the relation

$$0 \leq \hat{W}(\omega) \leq 1$$

means that the natural estimator, which possesses this property, is unlikely to be given up readily.

It is possible that a transformation of the same nature as Fischer's transformation of the classical correlation coefficient,

$$Z = \frac{1}{2} \log\left(\frac{1+r}{1-r}\right) = \text{arc tanh } r$$

would be helpful in centering the distribution more closely about its mean value; but we will here merely mention the possible estimator

$$\hat{\hat{W}} = \hat{W} - \frac{1}{n} (1-\hat{W})^2 .$$

To order $\frac{1}{n^2}$ we have

$$E[\hat{\hat{W}}] = W - \frac{2}{n} W(1-W)^2$$

$$\text{Var}[\hat{\hat{W}}] = \frac{2}{n} W(1-W)^2 .$$

The bias in $\hat{\hat{W}}$ is smaller than that in \hat{W} for all values of W while the variance is unchanged, but $\hat{\hat{W}}$ can assume negative values; of course, one could replace negative estimates by zero, but this would increase the bias and complicate the distribution theory. This situation is similar to that for the ordinary multiple correlation coefficient; see Kendall and Stuart (1961, p. 342).

We conclude that the usual estimating function

$$\hat{W}(\omega) = \frac{|\hat{f}_{jk}(\omega)|^2}{\hat{f}_{jj}(\omega) \hat{f}_{kk}(\omega)}$$

is on the whole preferable.

3.9 Summary of Recommended Procedure

A single procedure should probably not be laid down in advance for all problems of estimating coherence but, to the extent that this is possible, and by way of a summary, we list the following steps.

1. Eliminate non-stationarity - Section 1.2.
2. Prewhiten each series by autoregressive filtering and test by examining the spectrum of the residuals (we assume that this test is satisfied) - Section 2.1 (iii).
3. Estimate the cross-regression model by stagewise least squares, using truncation points believed to be larger than the greatest significant lags (by a priori knowledge or after examining $\hat{R}_{jk}(v)$) - Section 3.1.
4. Test the fitted cross-regression model - Section 3.5.
5. If the test for the cross-regression model is satisfied, estimate the spectrum of the residuals in that model - Section 2.1 (iii), and estimate coherence by the formula in Section 3.1.
6. If the test for the cross-regression model is not satisfied, the method of best linear preflattening could be resorted to - Section 3.7.

References

- [1] Akaike, H. and Yamanouchi, Y. (1962). On the statistical estimation of frequency response function. Ann. Inst. Statist. Math. XIV, No. 1, Tokyo, 23-56.
- [2] Amos, D. E. and Koopmans, L. H. (1963). Tables of the coefficient of coherence for stationary bivariate Gaussian processes. Sandia Corporation (text 50 pp.; tables 275 pp.).
- [3] Dalzell, J. F. and Yamanouchi, Y. (1958). The analysis of model test results in irregular head seas to determine the amplitude and phase relations to the waves. Stevens Institute of Technology Report 708, Hoboken, New Jersey.
- [4] Diananda, P. H. (1953). Some probability limit theorems with statistical applications. Proc. Cambridge Philos. Soc. 49, 239-246.
- [5] Efroymsen, M. A. (1960). Multiple Regression Analysis, in Mathematical Methods for Digital Computers. Ralston, A. and Wilf, H. S., editors, Wiley, Chapter 17, 191-203.
- [6] Gelfand, I. M. and Yaglom, A. M. (1959). Calculation of the amount of information about a random function contained in another such function. Amer. Math. Soc. Transl. 12, Series 2, 199-246.
- [7] Goodman, N. R. (1957). On the joint estimation of the spectra, cospectrum and quadrature spectrum of a two-dimensional stationary Gaussian process. Scientific Paper No. 10, Engineering Statistics Laboratory, College of Engineering, New York University, (168 pp.).
- [8] Granger, C. W. J. (1965). Discussion of paper by Priestly (1965, p. 232).
- [9] Granger, C. W. J. and Hatanaka, M. (1964). Spectral analysis of economic time series. Princeton University Press, (299 pp.).

- [10] Hannan, E. J. (1963). Regression for time series. In Time Series Analysis, edited by M. Rosenblatt, Wiley, New York, 17-37.
- [11] Isserlis, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika 12, 134-139.
- [12] Jenkins, G. M. (1963). Cross-spectral analysis and the estimation of linear open loop transfer functions. Chapter 18 of Time Series Analysis, edited by M. Rosenblatt, Wiley, New York, 267-276.
- [13] Kendall, M. G. and Stuart, A. (1961). The Advanced Theory of Statistics. Griffin, London, Vol. 2 (676 pp.).
- [14] Koopmans, L. H. (1964). On the coefficient of coherence for weakly stationary stochastic processes. Ann. Math. Statist. 35, 532-549.
- [15] Magness, T. A. (1954). Spectral response of a quadratic device to non-Gaussian noise. J. Appl. Phys. 25, No. 11, 1357-1365.
- [16] Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. Ann. Math. Statist. 28, 329-348.
- [17] Parzen, E. (1964). An approach to empirical time series analysis. Radio Science 68D, No. 9, 937-951. (Also published as Technical Report No. 10, Stanford University, January 24, 1964.
- [18] Parzen, E. (1965). On empirical multiple time series analysis. Technical Report No. 13, Stanford University, July 13, 1965.
- [19] Parzen, E. (1966). Empirical analysis of time series and stochastic processes. Unpublished manuscript.
- [20] Pierson, W. J. and Dalzell, J. F. (1960). The apparent loss of coherency in vector Gaussian processes due to computational procedures with application to ship motions and random seas. Technical Report, Department of Meteorology and Oceanography, College of Engineering, Research Division, New York University (21 pp.).

- [21] Priestley, M. B. (1965). Evolutionary spectra and non-stationary processes. J. Roy. Statist. Soc. Ser. B, 27, No. 2, 204-229; discussion 229-237.
- [22] Quenouille, M. H. (1947). A large sample test for the goodness of fit of autoregressive schemes. J. Roy. Statist. Soc. Ser. A, 110, 123-129.
- [23] Rosenblatt, H. M. (1965). Spectral analysis and parametric methods for seasonal adjustment of economic time series. Working Paper No. 23, Bureau of the Census, U. S. Department of Commerce (pp. 52).
- [24] Rosenblatt, M. (1959). Statistical analysis of stochastic processes with stationary residuals. In Probability and Statistics, edited by U. Grenander, Wiley, 246-275.
- [25] Schaerf, M. C. (1963). Estimation of the covariance and autoregressive structure of a stationary time series. Ph.D. thesis, Stanford University (pp. 72).
- [26] Scheinok, P. A. (1960). The error on using the asymptotic variance and bias of spectrograph estimates for finite observation time. Ph.D. thesis, Indiana University (pp. 74).
- [27] Tick, L. J. (1963). Conditional spectra, linear systems, and coherency. In Time Series Analysis, edited by M. Rosenblatt, Wiley, New York, 197-203.
- [28] Tukey, J. W. (1961). Discussion, emphasizing the connection between analysis of variance and spectrum analysis. Technometrics 3, No. 2, 191-220.
- [29] Tukey, J. W. (1965). Uses of numerical spectrum analysis in geophysics. Prepared for presentation at the meetings of the International Association for Statistics in the Physical Sciences and the International Statistical Institute, Belgrad, Yugoslavia, September, 1965 (pp. 110, including extensive bibliography).

- [30] Van Ness, J. W. and Woodroffe, M. (1966). The stability of the maximum deviation of sample spectral densities. Technical Report, Stanford University (to appear).
- [31] Wahba, G. G. (1966). Cross spectral distribution theory for mixed spectra and estimation of prediction filter coefficients. Ph.D. thesis, Stanford University.
- [32] Wiener, N. (1930). Generalized harmonic analysis. Acta Math. 55, 117-258.
- [33] Wiener, N. (1949). Extrapolation, Interpolation and Smoothing of Stationary Time Series. Wiley, New York, republished by M.I.T. Press, Cambridge, Massachusetts, 1964 (pp. 163).
- [34] Yaglom, A. M. (1962). An Introduction to the Theory of Stationary Random Functions. Prentice-Hall (235 pp.).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D		
<i>(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)</i>		
1. ORIGINATING ACTIVITY (Corporate author) Stanford University Department of Statistics Stanford, California		2a. REPORT SECURITY CLASSIFICATION Unclassified
		2b. GROUP
3. REPORT TITLE THE ESTIMATION OF COHERENCE		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report, May 1966		
5. AUTHOR(S) (Last name, first name, initial) Nettheim, Nigel		
6. REPORT DATE May 16, 1966	7a. TOTAL NO. OF PAGES 73	7b. NO. OF REFS 34
8a. CONTRACT OR GRANT NO. Nonr 225(80)	9a. ORIGINATOR'S REPORT NUMBER(S) Report No. 5	
b. PROJECT NO. NR-042-234	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
c.		
d.		
10. AVAILABILITY/LIMITATION NOTICES Releasable without limitations on dissemination.		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D. C. 20360	
13. ABSTRACT Coherence is the frequency domain analog of the classical correlation coefficient. In recent years investigators have become aware that the usual procedure for estimating coherence may often be seriously inadequate; in particular, low estimates have appeared in cases where the true coherence is known to be high. A critical discussion of the usual estimates leads to the proposal of the so-called cross-regressive filter model. In spite of the naturalness of estimates of coherence obtained in this way, they have apparently not been used in the past. In testing whether the proposed model is satisfied one needs to test the hypothesis of identically zero coherence; a natural test is proposed, requiring a knowledge of the mean and variance of the estimated coherence under the null hypothesis. Approximate values of these moments are obtained in terms of the number of degrees of freedom $n = 1 / \left\{ \frac{2\pi M}{T} \int_{-\infty}^{\infty} K^2(\lambda) d\lambda \right\},$ where T is the length of the two observed series, M is the truncation point beyond which cross-correlations are not computed, and K(λ) is the spectral window generator; together with a quantity (abstract continued)		

DD FORM 1473
1 JAN 64

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Mathematical Statistics Time series analysis Spectral analysis Multivariate analysis						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year, or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

Abstract (continued)

$$K_0 = \frac{\int_{-\infty}^{\infty} K^3(\lambda) d\lambda}{\left(\int_{-\infty}^{\infty} K^2(\lambda) d\lambda \right)^2} - 1$$

reflecting the extent to which the spectral window departs from being rectangular.

The opportunity is also taken to derive formulae for joint moments of any order of the estimated spectra. These formulae are believed to be new. In deriving these formulae a high level of mathematical rigor has not been achieved; in particular, the behavior of several delta-like functions appearing in tandem is treated in an intuitive way. To avoid this feature one would like to work in the time domain rather than the frequency domain but this proved impracticable. The lack of rigor will not affect the application of the results, as is indicated by comparison with an approximate distribution for coherence obtained by N. Goodman assuming, among other things, a rectangular spectral window.

Finally some miscellaneous alternative procedures are briefly considered but none of these is recommended.

RECENT REPORTS ON TIME SERIES ANALYSIS

Research in time series analysis under the direction of Professor Emanuel Parzen is supported by the following agencies:

Office of Naval Research - Contract Nonr 225(80)

Army Research Office - Grant ARO(D)-31-124 G726

National Science Foundation - Grant NSF GP-4265

Since January 1, 1965, the following technical reports have been issued.

<u>Author</u>	<u>Title</u>	<u>Date</u>	<u>Agency</u>	<u>Report No.</u>
J. W. Van Ness	Asymptotic Normality of Bispectral Estimates	4-9-65	ARO	14
J. W. Van Ness B. B. Mandelbrot	Fractional Brownian Motion, Self-similar Processes, and Applications	3-1-66	ARO	15
J. W. Van Ness M. B. Woodrooffe	The Maximum Deviations of Sample Spectral Densities	4-27-66	ARO	16
G. R. Hext	A New Approach to Time Series with Mixed Spectra	5-16-66	ARO	17
E. Parzen	On Empirical Multiple Time Series Analysis	7-19-65	NSF	13
G. R. Hext	A New Approach to Time Series with Mixed Spectra	5-16-66	NSF	14
G. Wahba	Cross Spectral Distribution Theory for Mixed Spectra and Estimation of Prediction Filter Coefficients	5-16-66	NSF	15
E. Parzen	The Role of Spectral Analysis in Time Series Analysis	7-12-65	ONR	2
E. Parzen	On Empirical Multiple Time Series Analysis	7-19-65	ONR	3
E. Parzen	Analysis and Synthesis of Linear Models for Time Series	4-18-66	ONR	4
N. Nettheim	On Estimation of Coherence	5-16-66	ONR	5