AN INTRODUCTION TO POLYSPECTRA[1]

By David R. Brillinger

The London School of Economics and Political Science

1. Introduction and summary. The subject of this paper is the higher-order spectra or polyspectra of multivariate stationary time series. The intent is to derive (i) certain mathematical properties of polyspectra, (ii) estimates of polyspectra based on an observed stretch of time series, (iii) certain statistical properties of the proposed estimates and (iv) several applications of the results obtained.

As might be expected, in lower order cases the polyspectrum reduces to spectra already considered. If one is considering a single time series, the first order polyspectrum is the usual power spectrum considered in [2], [14], [22], while the second order polyspectrum is the bispectrum considered in [12], [23], [28]. Also, if one is considering a pair of time series the first order polyspectrum is the cross-spectrum considered in [6], [10], [15].

For the case of a single time series the idea of a higher-order spectrum occurs in [3]. The idea has since been developed to a higher level of algebraic and analytic detail in [24]. Also in [24] the notion of considering a spectral representation for a cumulant rather than for a product moment occurs and is acknowledged to be due to Kolmogorov. Another related early paper is [18].

The present paper generalizes the definitions of these papers in the sense that k-dimensional time series are considered. Another contribution is a theorem indicating that for a broad class of processes one is wise to restrict consideration to cumulants rather than product moments.

Finally it should be noted that the term polyspectrum is due to J. W. Tukey. I have perhaps used the term in a more restricted sense than he would wish in that I have reserved it for the Fourier transform of a cumulant (at the expense of other functions of moments).

2. General motivation. In a heuristic sense the harmonic analysis of a time series $X(t)$ may be looked upon as the consideration of a representation of the series in the form,

$$X(t) = \sum R_k \exp [i(w_k t + \phi_k)]. \tag{2.1}$$

This consideration gains some validity from a theorem of Cramér's [9] to the effect that any covariance stationary time series $X(t)$ with mean 0, has a representation in the form

Received 13 July 1964; revised 2 April 1965.

[1] This research was commenced when the author had a joint appointment at the Bell Telephone Laboratories and Princeton University. At Princeton the research received the support of National Science Foundation Grant NSF-GS30 and Army Research Office Contract DA 36-034-ORD-2297.

$$X(t) = \int e^{iwt}\, dZ(w) \tag{2.2}$$

where $Z(w)$ is a stochastic function.

A further aspect of harmonic analysis is that one often acts as if the various terms, $R_k \exp [i(w_k t + \phi_k)]$, appearing in (2.1) are independent of one another. This simplifying thought is possibly instigated by the knowledge that $Z(w)$ appearing in (2.2) is such that

$$E\{dZ(w)\, dZ^*(w')\} = 0, \qquad w \neq w', \tag{2.3}$$

implying independence in the real-valued Gaussian case. Or perhaps it is due to the fact that one often imagines a series as being generated by a number of linear time invariant operations on a Wiener process and one knows that such operations do not mix up frequencies. (See [21], p. 83 for example.)

In practice however the frequency components of a time series do not always appear to be independent. In a study of ocean wave records, [12], Hasselmann, Munk and MacDonald have found empirically various wave components related to one another. In a study of the effect of introducing a signal into the eye [31] Van der Tweel has found that the responses at 5 c/s and 10 c/s are related to one another. Many economists have noted a seasonal effect in economic time series of persistent non-cosinusoidal shape. This finding perhaps indicates that the various harmonics of 1 cycle/year are in some form of fixed relation with one another.

A simple form of tying together of frequency components occurs if a number of independent frequency components, $R_k \exp [i(w_k t + \phi_k)]$, instead of simply adding together to produce a series $X(t)$, as in (2.1), add together and also multiply together in pairs to produce the series

$$\sum R_k \exp [i(w_k t + \phi_k)] + \sum A_{jk} \exp (i\alpha_{jk}) R_j R_k \exp [i(w_j + w_k) + i(\phi_j + \phi_k)]. \tag{2.4}$$

That is, we are moving away from an additive model to a model containing second order product interactions.

The reader will note that, in the expression (2.4), the correlation between the product of the components at frequencies w_j and w_k with the component at frequency $w_j + w_k$ is one, provided the sum of no other pair of frequencies present is $w_j + w_k$. This observation will later lead us directly to the polyspectrum.

Continuing to consider (2.4), a simple means of producing a time series containing terms such as those in (2.4) is to take a series $X(t)$ with a simple harmonic analysis and then to form the series

$$Y(t) = f[X(t)] \tag{2.5}$$

where f is a non-linear function. In a situation in which one is given the series $Y(t)$, one would like to find the function f in order to be able to remove the non-additivity that it has introduced. A coefficient will be proposed for this purpose in Section 6.

As a final point, people often introduce the power spectrum by noting the ease with which linear time invariant operations may be described in terms of it. If one wishes to describe easily the effect of a multilinear or polynomial (in the sense of [20]) operation, one finds oneself led to higher order spectra. Tukey in [30] has commenced the development of a calculus relating polynomial operations to higher order spectra.

3. Definitions. Various classes of stochastic processes have been introduced in order to deal with higher order spectral moments. Specifically the classes $T^{(k)}$, $S^{(k)}$, and $\Phi^{(k)}$ defined below have been discussed in [24]. However we shall require a new class. Before defining this new class $\Psi^{(k)}$, let us present the definitions of $T^{(k)}$, $S^{(k)}$, and $\Phi^{(k)}$ as they will also be needed in the paper.

Let $U(t)$ be a real measurable random process, $-\infty < t < \infty$.
Then

$T^{(k)}$ denotes the class of processes $U(t)$ for which

$$E\,|U(t)|^k \leqq C_k < \infty; \tag{3.1}$$

$S^{(k)}$ denotes the class of processes $U(t)$ belonging to $T^{(k)}$ and such that for $1 \leqq j \leqq k$, $-\infty < u < \infty$,

$$EU(t_1) \cdots U(t_j) = EU(t_1 + u) \cdots U(t_j + u); \tag{3.2}$$

$\Phi^{(k)}$ denotes the class of processes $U(t)$ belonging to $T^{(k)}$ and such that for $1 \leqq j \leqq k$, there exist functions $M^{(j)}(w_1, \cdots, w_j)$ of bounded variation such that

$$EU(t_1) \cdots U(t_j) = \int \cdots \int \exp\,[i(w_1t_1 + \cdots + w_jt_j)]\, dM^{(j)}(w_1, \cdots, w_j). \tag{3.3}$$

Before defining $\Psi^{(k)}$ the following notation will be required:

(i) $(v_1, \cdots, v_j)$ denotes a grouping of the integers $1, 2, \cdots, k$ into j groups $v_1, \cdots, v_j$;

(ii) $t_v = (t_{h_1}, \cdots, t_{h_n})$ when v corresponds to the grouping $(h_1, \cdots, h_n)$. For example if $v = (1, 8, 9)$ then $t_v = (t_1, t_8, t_9)$;

(iii) $\{X_1(t), \cdots, X_k(t)\}$ stands for a k-dimensional complex-valued stochastic process;

(iv) $m_{1\cdots k}(t_1, \cdots, t_k)$ denotes the kth order product moment $EX_1(t_1) \cdots X_k(t_k)$;

(v) $c_{1\cdots k}(t_1, \cdots, t_k)$ denotes the kth order cumulant

$$\sum (-1)^{p-1}(p-1)!\, m_{v_1}(t_{v_1}) \cdots m_{v_p}(t_{v_p}) \tag{3.4}$$

where the summation extends over all groupings of the integers $1, \cdots, k$.

$\Psi^{(k)}$ is now defined as the class of discrete or continuous time k-dimensional complex-valued processes $\{X_1(t), \cdots, X_k(t)\}$ such that

(a) for $1 \leqq j \leqq k$ and $1 \leqq h_1, \cdots, h_j \leqq k$, $m_{h_1\cdots h_j}(t_1, \cdots, t_j)$ exists,

(b) for $1 \leqq j \leqq k$ and $-\infty < u < \infty$ in the continuous case or $u = 0, \pm 1, \pm 2, \cdots$ in the discrete case

$$m_{h_1\cdots h_j}(t_1 + u, \cdots, t_j + u) = m_{h_1\cdots h_j}(t_1, \cdots, t_j),$$

(c) for $1 \leqq j \leqq k$ there exist measures $\delta(w_1 + \cdots + w_j)C_{h_1 \cdots h_j}(w_1, \cdots, w_j) \, dw_1 \cdots dw_j$ absolutely continuous with respect to Lebesgue measure on the plane $w_1 + \cdots + w_j = 0$ such that $c_{h_1 \cdots h_j}(t_1, \cdots, t_j)$ equals

$$(3.5) \quad \int \cdots \int \exp [i(w_1 t_1 + \cdots + w_j t_j)] \cdot \delta(w_1 + \cdots + w_j) C_{h_1 \cdots h_j}(w_1, \cdots, w_j) \, dw_1 \cdots dw_j .$$

(Throughout the paper $\delta(w)$ denotes the Dirac delta function. The reader may easily see that the condition (b) above implies that the measure must in fact have support on the plane $w_1 + \cdots + w_j = 0$.)

If a process $\{X_1(t), \cdots, X_k(t)\}$ belongs to $\Psi^{(k)}$ then $C_{1 \cdots k}(w_1, \cdots, w_k)$ is defined to be the $(k - 1)$th order *polyspectrum* of the process.

A number of comments may be made about $\Psi^{(k)}$.

(i) In the integrals above, the range of the arguments w is $-\pi \leqq w \leqq \pi$ in the discrete case and $-\infty < w < \infty$ in the continuous case.

(ii) If the series involved are real

$$(3.6) \qquad C^*_{1 \cdots k}(w_1, \cdots, w_k) = C_{1 \cdots k}(-w_1, \cdots, -w_k).$$

(iii) If the series are identical, $C_{1 \cdots k}(w_1, \cdots, w_k)$ is symmetric in its arguments.

(iv) If a process $\{X_1(t), \cdots, X_k(t)\}$ satisfies (a) and (b) above and for ϕ equal both c and C,

$$\int \cdots \int |\phi_{h_1 \cdots h_j}(t_1, \cdots, t_{j-1}, 0)| \, dt_1 \cdots dt_{j-1} < \infty$$

in the continuous case, or

$$\sum \cdots \sum |\phi_{h_1 \cdots h_j}(t_1, \cdots, t_{j-1}, 0)| < \infty$$

in the discrete case, then the process belongs to $\Psi^{(k)}$. In this case the Fourier relation (3.5) may in fact be inverted;

(v) $C_{1 \cdots k}(w_1, \cdots, w_k)$ being a complex number, for some purposes it may be useful to express it in terms of an amplitude and phase;

(vi) If $\{X_1(t), \cdots, X_k(t)\}$ is in fact $\{X(t), \cdots, X(t)\}$, i.e. all of the components are identical, and if $X(t)$ is real then $\Psi^{(k)}$ reduces to the class $\Delta^{(k)}$ introduced by Kolmogorov (see [24].)

This section will be concluded with a number of examples of polyspectra.

Example 1. Suppose $\{X(t), X(t)\}$ denotes a two dimensional real process with identical components, then the first order polyspectrum $C_{11}(w_1, w_2)$ reduces to the power spectrum of $X(t)$.

Example 2. Suppose $\{X_1(t), X_2(t)\}$ denotes a two dimensional real process, then the first order polyspectrum $C_{12}(w_1, w_2)$ reduces to the cross-spectrum of the two series $X_1(t)$ and $X_2(t)$.

Example 3. Suppose $X(t) = \int g(t - u) \, dY(u)$ where $\int |g(u)| \, du < \infty$ and $Y(u)$ is a process with stationary and independent increments. Denote the jth cumulant of $Y(1) - Y(0)$ by K_j (it being assumed to exist). The $(j - 1)$th order polyspectrum of $X(t)$ (really of $\{X(t), \cdots, X(t)\}$) is given by

$$K_j G(w_1) \cdots G(w_j) \tag{3.7}$$

where $G(w) = \int e^{-iwu} g(u)\, du$ and $w_1 + \cdots + w_j = 0$.

This result may be demonstrated by making use of the characteristic functional of the process as derived in [1], p. 148 for example.

EXAMPLE 4. Suppose

$$X(t) = \int a(t-u)\, dW(u) + \iint b(t-u, t-v)\, dW(u)\, dW(v) \tag{3.8}$$

where $W(t)$ is a Wiener process, $a(u)$ and $b(u, v)$ have Fourier transforms $A(w)$, $B(w_1, w_2)$ respectively and $b(u, v)$ is assumed symmetric in u and v for convenience. In this case the second order polyspectrum or bispectrum of $X(t)$ is given by,

$$\begin{aligned} 2[A(w_1)A(w_2)B(-w_1, -w_2) &+ A(w_2)A(w_3)B(-w_2, -w_3) \\ &+ A(w_3)A(w_1)B(-w_3, -w_1)] \\ + 8 \int \overline{B(w, w_1 - w)B(w_2 + w, -w)B(w - w_1, -w - w_2)}\, dw \end{aligned} \tag{3.9}$$

where the bar denotes the mean of all permutations of (w_1, w_2, w_3).

This result may be demonstrated by making use of the formula for the kth order product moment of a Wiener process (see [32]).

EXAMPLE 5. Suppose

$$\begin{aligned} X_1(t) &= \int a(t-u)\, dW_1(u), \\ X_2(t) &= \int b(t-u)\, dW_2(u), \\ X_3(t) &= \iint c(t-u, t-v)\, dW_1(u)\, dW_2(v), \end{aligned} \tag{3.10}$$

where $W_1(t)$ and $W_2(t)$ are independent Wiener processes and where $a(u)$, $b(u)$ and $c(u, v)$ have Fourier transforms $A(w)$, $B(w)$ and $C(w_1, w_2)$ respectively. In this case the second order polyspectrum of $\{X_1(t), X_2(t), X_3(t)\}$ is given by,

$$A(w_1)B(w_2)C(-w_1, -w_2). \tag{3.11}$$

If in fact $W_1(t)$ and $W_2(t)$ are not independent, but are completely dependent, $W_1(t) = W_2(t)$, then the polyspectrum is given by,

$$A(w_1)B(w_2)\{C(-w_1, -w_2) + C(-w_2, -w_1)\} \tag{3.12}$$

or if $c(u, v)$ is symmetric in u and v by,

$$2A(w_1)B(w_2)C(-w_1, -w_2). \tag{3.13}$$

4. Estimation. In this section it will be supposed that an observed stretch $\{X_1(t), \cdots, X_k(t); 0 \leqq t \leqq T\}$ of a real discrete time series belonging to $\Psi^{(k)}$, $k \geqq 2$, is available. (The corresponding procedures for a stretch of a continuous time series are immediately apparent.) Three distinct techniques for the estimation of polyspectra will be proposed.

The first technique follows directly from the definition (3.5) which indicates that the polyspectrum $C_{1\cdots k}(w_1, \cdots, w_k)$, $\sum w_j = 0$, is the $(k-1)$-dimensional Fourier transform of $c_{1\cdots k}(t_1, \cdots, t_{k-1}, 0)$. The technique is the following; estimate the product moment $m_{1\cdots k}(t_1, \cdots, t_{k-1}, 0)$, and all necessary lower order product moments by formulae of the form,

$$(4.1) \quad \hat{m}_{1\cdots k}(t_1, \cdots, t_{k-1}, 0) = T^{-1} \sum_{t=0}^{s} X_1(t_1 + t) \cdots X_{k-1}(t_{k-1} + t) X_k(t)$$

where $s = T - \max\{t_j\}$, and for $(t_1, \cdots, t_{k-1})$ in a set I to be specified later. The joint cumulant $c_{1\cdots k}(t_1, \cdots, t_{k-1}, 0)$ may now be estimated by substituting in (3.4), i.e. by forming

$$(4.2) \quad \hat{c}_{1\cdots k}(t_1, \cdots, t_{k-1}, 0) = \sum (-1)^{p-1}(p-1)!\, \hat{m}_{v_1}(t_{v_1}) \cdots \hat{m}_{v_p}(t_{v_p})$$

where the summation extends over all grouping of the integers $1, \cdots, k$ and $t_k = 0$. (Some workers may wish to divide by $T - s + 1$ rather than T in the expression (4.1). For a discussion of this point in the first order case see [29]. Also some workers may perhaps wish to substitute into the formulas for Fisher's k-statistics. For a definition of these latter see [16].) The estimate (4.2) has one undesirable property, namely it is not invariant under changes $X_j(t) \to X_j(t) + h_j$, whereas the corresponding population cumulant is. This defect may be remedied by first subtracting the sample means from the series before calculating the estimate. In this case the summation in (3.4) extends only over groupings containing no first-order elements.

From (3.5) we see that $\hat{c}_{1\cdots k}(t_1, \cdots, t_{k-1}, 0)$ is estimating

$$(4.3) \quad \int \cdots \int \exp[i(w_1 t_1 + \cdots + w_{k-1} t_{k-1})] C_{1\cdots k}(w_1, \cdots, w_k)\, dw_1 \cdots dw_{k-1}$$

where $\sum w_j = 0$. That is it is estimating the coefficient of a term in the Fourier series expansion of $C_{1\cdots k}(w_1, \cdots, w_k)$, $\sum w_j = 0$. The sum of a number of such terms may be used to approximate the function itself; however classical Fourier analysis indicates that the use of a summability technique may well improve the approximation (see [13] for example). This leads one to consider estimates of the form,

$$(4.4) \quad (1/2\pi)^{k-1} \sum_I \lambda_{t_1 \cdots t_{k-1}}^{(n)} \exp[-i(w_1 t_1 + \cdots + w_{k-1} t_{k-1})] \cdot \hat{c}_{1\cdots k}(t_1, \cdots, t_{k-1}, 0)$$

where the $\lambda_{t_1 \cdots t_{k-1}}^{(n)}$ are the convergence factors of a summability method.

Convergence factors that seem appropriate for this situation include;

(a) a product of one dimensional convergence factors, i.e.

$$(4.5) \quad \lambda_{t_1 \cdots t_{k-1}}^{(n)} = \lambda_{t_1}^{(n)} \cdots \lambda_{t_{k-1}}^{(n)}$$

where for example (Féjer summability)

$$(4.6) \quad \begin{aligned} \lambda_t^{(n)} &= 1 - |t|/n, \quad 0 \leqq |t| \leqq n \\ &= 0 \quad \text{otherwise} \end{aligned}$$

or (Tukey summability)

$$\lambda_t^{(n)} = .54 + .46 \cos \pi t/n, \qquad 0 \leqq |t| \leqq n \tag{4.7}$$
$$= 0 \quad \text{otherwise;}$$

(b) a genuine multidimensional factor such as (Bochner-Riesz summability, see [4])

$$\lambda_{t_1 \cdots t_{k-1}}^{(n)} = (1 - |t|^2/n^2)^m, \qquad 0 \leqq |t| \leqq n \tag{4.8}$$
$$= 0 \quad \text{otherwise}$$

where $|t|^2 = t_1^2 + \cdots + t_{k-1}^2$.

This last factor has the advantage that the convergence of the approximating series at a specified point depends only on the behaviour of the function in the neighborhood of the point.

We see that the set I mentioned earlier is in fact determined by the non-zero values of the convergence factors.

Before describing the second estimation technique, let us introduce another means of looking at the polyspectrum. Because $\Psi^{(k)} \subseteq \Psi^{(2)}$, the series $\{X_j(t)\}$ has a Cramér representation

$$\{\textstyle\int e^{iwt}\, dZ_j(w)\} \tag{4.9}$$

where $Z_j(w)$ is a stochastic function.

In terms of this representation the cumulant $c_{1\cdots k}(t_1, \cdots, t_k)$ may be written

$$\int \cdots \int \exp\,[i(w_1t_1 + \cdots + w_kt_k)]\mathcal{C}(dZ_1(w_1), \cdots, dZ_k(w_k)) \tag{4.10}$$

where $\mathcal{C}(x_1, \cdots, x_k)$ denotes the joint cumulant of $x_1, \cdots, x_k$. (This results from the fact that the joint cumulant of $y_1, \cdots, y_m$ where $y_k = \sum a_{i_k k} x_{i_k k}$ is given by

$$\sum \cdots \sum a_{i_1 1} \cdots a_{i_m m}\mathcal{C}(x_{i_1 1}, \cdots, x_{i_m m}). \tag{4.11}$$

In fact (4.11) would appear to be one of the main reasons why cumulants prove so useful. It states that one can write down immediately the joint cumulant of a number of linear combinations of independent or dependent random variables, in terms of their joint cumulants.)

Comparing (3.5) and (4.10) and assuming that the Fourier transform is unique almost everywhere,

$$\delta(w_1 + \cdots + w_k)C_{1\cdots k}(w_1, \cdots, w_k)\, dw_1 \cdots dw_k \tag{4.12}$$
$$= \mathcal{C}(dZ_1(w_1), \cdots, dZ_k(w_k)) \quad \text{(almost everywhere).}$$

This indicates that with realizations of the spectral functions $dZ_j(w_j)$ and a proper normalization one can estimate the polyspectrum.

After this introduction, it can be stated that the second proposed technique of estimating polyspectra is based upon obtaining realizations of the spectral functions by means of the procedure of complex demodulation ([8], [29]). Given an observed stretch of series $\{X_j(t), 0 \leqq t \leqq T\}$ the steps are as follows:

(i) form $X_j(t) \cos w_0 t$ and $X_j(t) \sin w_0 t$, $0 \leqq t \leqq T$,

(ii) form the series,

$$U_j(t, w_0) = (2k+1)^{-1} \textstyle\sum_{s=-k}^{k} \lambda_{t-s}^{(k)} X_j(s) \cos w_0 s, \tag{4.13}$$

$$U_j^H(t, w_0) = (2k+1)^{-1} \textstyle\sum_{s=-k}^{k} \lambda_{t-s}^{(k)} X_j(s) \sin w_0 s, \tag{4.14}$$

$k \leqq t \leqq T - k$, where for example $\lambda_t^{(k)}$ is given by (4.6) or (4.7). $\Delta Z_j(w_j)$ may now be approximated by $U_j(t, w_j) + iU_j^H(t, w_j)$.

(iii) $C_{1\cdots k}(w_1, \cdots, w_k)$, $\sum w_j = 0$, is now estimated by forming

$$\textstyle\sum (-1)^{p-1}(p-1)!\, \hat{m}_{v_1} \cdots \hat{m}_{v_p}, \tag{4.15}$$

the summation extending over all groupings of $1, 2, \cdots, k$ and $\hat{m}_v$ is given by,

$$T^{-1} \textstyle\sum_{t=k}^{T-k} \{U_{h_1}(t, w_{h_1}) + iU_{h_1}^H(t, w_{h_1})\} \cdots \{U_{h_p}(t, w_{h_p}) + iU_{h_p}^H(t, w_{h_p})\} \tag{4.16}$$

where $v = (h_1, \cdots, h_p)$.

The final technique proposed for the estimation of a polyspectrum is based upon the fact that the expression (4.12) is also equal to

$$\mathcal{C}(e^{iw_1 t}\, dZ_1(w_1), \cdots, e^{iw_k t}\, dZ_k(w_k)). \tag{4.17}$$

The polyspectrum can consequently be estimated by obtaining realizations of the frequency components $e^{iw_j t}\Delta Z_j(w_j)$. These realizations may be obtained by deriving estimates of $X_j(t, w_0)$, the component of frequency w_0 in the series $X_j(t)$ and $X_j^H(t, w_0)$ the corresponding Hilbert transform (see [8]). A useful technique for obtaining $X_j(t, w_0)$ and $X_j^H(t, w_0)$ is described below:

$$e^{iw_j t}\Delta Z_j(w_j) \quad \text{may be estimated by} \quad X_j(t, w_j) + iX_j^H(t, w_j).$$

$C_{1\cdots k}(w_1, \cdots, w_k)$, $\sum w_j = 0$, may be estimated by forming the expression (4.15) where U and U^H in (4.16) are replaced by X and X^H respectively.

The promised technique for obtaining $X_j(t, w_0)$ and $X_j^H(t, w_0)$ evolves from a procedure suggested in [11], pp. 77–78. Define

$$a_m(t) = N^{-1} \textstyle\sum_{s=-N}^{N} X(t+s) \cos \pi m s/N, \qquad m = 0, 1, \cdots, N, \tag{4.18}$$

$$b_m(t) = N^{-1} \textstyle\sum_{s=-N}^{N} X(t+s) \sin \pi m s/N, \qquad m = 1, \cdots, N-1, \tag{4.19}$$

where $w_0 = \pi m/N$. The advantage of this definition is that the $a_m(t)$, $b_m(t)$ may be generated by recursion,

$$\begin{aligned} a_m(t+1) = {} & a_m(t) \cos \pi m/N + b_m(t) \sin \pi m/N \\ & + [(-1)^m/N][X(N+1+t) - X(-N+t)], \end{aligned} \tag{4.20}$$

$$b_m(t+1) = -a_m(t) \sin \pi m/N + b_m(t) \cos \pi m/N. \tag{4.21}$$

Use of these recursion relations greatly decreases the number of arithmetical operations involved.

The proposed estimates of $X(t, w_0)$ and $X^H(t, w_0)$ are now,

$$.23a_{m-1}(t) + .54a_m(t) + .23a_{m+1}(t), \tag{4.22}$$

$$.23b_{m-1}(t) + .54b_m(t) + .23b_{m+1}(t), \tag{4.23}$$

respectively. (The coefficients .23 and .54 used here are derived from Tukey's weights.)

For later reference it is noted here that in terms of the spectral representation $a_m(t, w_0) + ib_m(t, w_0)$ may be written,

$$\int \exp\,(iwt + i\tfrac{1}{2}\theta)\,[\sin N\theta/N \sin \tfrac{1}{2}\theta]\, dZ(w) \tag{4.24}$$

where $\theta = w + w_0$. This means that $X(t, w_0) + iX^H(t, w_0)$ equals

$$\int e^{iwt} Q(w + w_0)\, dZ(w) \tag{4.25}$$

with an elementary function $Q(w)$.

The final two estimation techniques proposed above have several advantages over the first. They are easily adapted to obtain running estimates of the polyspectrum and so the presence of nonstationarities may be investigated. Also once an initial effort has been made to obtain the series $X(t, w_0) + iX^H(t, w_0)$ or $U(t, w_0) + iU^H(t, w_0)$, they may be put to a variety of uses with few additional calculations; for example polyspectra of various orders, involving various series may be calculated. These series should have to be calculated only once in the history of a series, provided enough foresight is shown in the bandwidths of the filters employed. The series $U + iU^H$ has a further advantage; typically it is fairly smooth so not every value need necessarily be retained.

5. Some statistical properties of the proposed estimates. The discussion in this section will be restricted to the discrete case; however the continuous case follows in an identical manner, sums in the time domain being replaced by integrals, and integrals in the frequency domain having their range increased from $-\pi, \pi$ to $-\infty, \infty$.

Suppose a stretch of a time series $\{X_1(t), \cdots X_k(t);\ -T' \leqq t \leqq T'\}$ is available. When the second and third estimates of Section 4 are examined in detail for this case, it is seen that they have the form,

$$\hat{C}_{1\cdots k} = \sum (-1)^{p-1}(p-1)!\, \hat{n}_{v_1} \cdots \hat{n}_{v_p} \tag{5.1}$$

where when $v = (j_1, \cdots, j_m)$,

$$\hat{n}_v = (2T+1)^{-1} \textstyle\sum_{t=-T}^{T} Y_{j_1}(t) \cdots Y_{j_m}(t) \tag{5.2}$$

with

$$Y_j(t) = \sum g_j(t-u) X_j(u) \tag{5.3}$$

for (complex valued) functions $g_j(t)$ related to the filters employed and where $T' > T > 0$.

We will restrict consideration to estimates of this form throughout this section.

The $g_j(t)$ appearing in (5.3) will be said to be *absolutely summable* if they are such that

$$\sum_{t=-\infty}^{\infty} |g_j(t)| < \infty. \tag{5.4}$$

Also if the time series $\{X_1(t), \cdots, X_k(t)\}$ is such that for joint cumulants of all orders and all sets of subscripts $(i_1, \cdots, i_m)$,

$$\sum_{t_2} \cdots \sum_{t_m} |c_{i_1 \cdots i_m}(0, t_2, \cdots, t_m)| < \infty. \tag{5.5}$$

Then the series is said to satisfy *Condition A*.

Expressions will be required for the joint cumulants of a number of non-elementary random variables. Before presenting these expressions however, let us introduce some terminology of [17].

Consider the two-way table

$$\begin{matrix} (1, 1) & \cdots & (1, k_1) \\ \vdots & & \vdots \\ (j, 1) & \cdots & (j, k_j) \end{matrix} \tag{5.6}$$

and a partition $P_1 \cup P_2 \cup \cdots \cup P_m$ of its elements. We shall say that the sets P_{i^1} and P_{i_2} of the partition *hook* if there exist $(i_1, j_1) \varepsilon P_{i_1}$ and $(i_2, j_2) \varepsilon P_{i_2}$ such that $i_1 = i_2$. We shall say that the sets $P_{i'}$ and $P_{i''}$ *communicate* if there exists a sequence of sets $P_{i_1} = P_{i'}, P_{i_2}, \cdots, P_{i_r} = P_{i''}$ such that P_{i_j} and $P_{i_{j+1}}$ hook for each j. A partition is said to be *indecomposable* if and only if all its sets communicate.

If the rows of table (5.6) are denoted by $R_1, \cdots, R_j$ then $\{P_1, \cdots, P_m\}$ is indecomposable if and only if there exist no sets $P_{i_1}, \cdots, P_{i_r}(r < m)$ and rows $R_{j_1}, \cdots, R_{j_s}(s < j)$ with

$$P_{i_1} \cup \cdots \cup P_{i_r} = R_{j_1} \cup \cdots \cup R_{j_s}. \tag{5.7}$$

The indecomposable partitions correspond to the arrays of [16], Rule 3, p. 283, when the rule is extended to the higher dimensional case.

LEMMA 5.1. *Consider a (not necessarily rectangular) array* $\|x_{mn}\|$ *of random variables* x_{mn}. *Consider the* j *random variables*

$$y_m = \prod_{n=1}^{k_m} x_{mn}. \tag{5.8}$$

The joint jth *order cumulant* $\mathcal{C}(y_1, \cdots, y_j)$ *is given by*

$$\sum_v \mathcal{C}_{v_1} \cdots \mathcal{C}_{v_p} \tag{5.9}$$

where $\mathcal{C}_v = \mathcal{C}(x_{a_1}, \cdots, x_{a_m})$ *when* $v = (a_1, \cdots, a_m)$, *(the a's are pairs of integers selected from table* (5.6)), *and the summation in* (5.9) *extends over all the indecomposable partitions of* (5.6).

PROOF. This result follows immediately from a theorem of [17].

LEMMA 5.2. *Consider series* $Z_1(t), \cdots, Z_k(t)$ *of the form*

$$Z_j(t) = \sum h_j(t - u)X_j(u) \tag{5.10}$$

with $h_j(t)$ *complex-valued, bounded*[2] *and absolutely summable. Suppose that the*

[2] The boundness follows from the absolute summability in fact.

series $\{X_1(t), \cdots, X_k(t)\}$ *satisfies Condition A and let* $(v_1, \cdots, v_p)$ *denote an indecomposable partition of table* (5.6) *and* $f_{v_i}(s_{v_i})$, $(s_1 = 0)$, *the joint cumulant of elements selected from the table*

$$
\begin{matrix} Z_1(0) & \cdots & Z_k(0) \\ Z_1(s_2) & \cdots & Z_k(s_2) \\ \vdots & \cdots & \vdots \\ Z_1(s_j) & \cdots & Z_k(s_j) \end{matrix} \tag{5.11}
$$

in accordance with v_i. *Under these conditions*

$$\sum_{s_2=-\infty}^{\infty} \cdots \sum_{s_j=-\infty}^{\infty} |f_{v_1}(s_{v_1}) \cdots f_{v_p}(s_{v_p})| < \infty.$$

PROOF. If $v = \{(i_1, j_1), \cdots, (i_m, j_m)\}$, define $\tilde{h}_v(t_v)$ as

$$h_{j_1}(t_{i_1, j_1}) \cdots h_{j_m}(t_{i_m, j_m}). \tag{5.12}$$

Now $\sum_{s_2} \cdots \sum_{s_j} f_{v_1}(s_{v_1}) \cdots f_{v_p}(s_{v_p})$

$$= \sum_{t_{11}} \cdots \sum_{t_{jk}} \sum_{s_2} \cdots \sum_{s_j} \tilde{h}_{v_1}(s_{v_1} - t_{v_1}) \cdots \tilde{h}_{v_p}(s_{v_p} - t_{v_p})\, c_{v_1}(t_{v_1}) \cdots c_{v_p}(t_{v_p}) \tag{5.13}$$

$$= \sum_{u_{v_1}} \cdots \sum_{u_{v_p}} \sum_{t_1} \cdots \sum_{t_p} \sum_{s_2} \cdots \sum_{s_j} \tilde{h}_{v_1}(s_{v_1} - u_{v_1} - t_1) \cdots \tilde{h}_{v_p}(s_{v_p} - u_{v_p} - t_p) c_{v_1}(u_{v_1}) \cdots c_{v_p}(u_{v_p}), \tag{5.14}$$

where t_i is one of the arguments of t_{v_i} and $u_{v_i} = t_{v_i} - t_i$. (We are here taking advantage of the stationarity of the process.) Now in the arguments of the h's there occur a variety of $s_i - t_m$. Since the partition is indecomposable, there exist $p + j - 1$ of these such that the relationship

$$s_{i_n} - t_{m_n} = a_n, \tag{5.15}$$

$n = 1, \cdots, p + j - 1$ is non-singular.

Let us substitute the a's into (5.14) retaining $p + j - 1$ h's with arguments of the form $a_m - u_n$ and note that the remaining h's are bounded. Thus the absolute value of (5.13) is

$$\leqq M \sum_{u_{v_1}} \cdots \sum_{u_{v_p}} |c_{v_1}(u_{v_1}) \cdots c_{v_p}(u_{v_p})| \cdot \sum_{a_1} \cdots \sum_{a_{p+j-1}} |h(a_1 - u_{m_1}) \cdots h(a_{p+j-1} - u_{m_{p+j-1}})|. \tag{5.16}$$

That (5.16) is bounded now follows from the discrete analog of Theorem 33 of [5]. The interchanges of the various summations in this lemma may be justified by Fubini's theorem.

Define a_T to be

$$(2T + 1)^{-1} \sum_{-T}^{T} Y_1(t) \cdots Y_k(t) - \sum_{p>1} \prod_{j=1}^{p} \sum \cdots \sum \tilde{g}_{v_j}(-t_{v_j}) c_{v_j}(t_{v_j}). \tag{5.17}$$

THEOREM 5.1. *Consider the real-valued random variable* $\alpha a_T + \alpha^* a_T^*$ *where the* $g_j(t)$ *are bounded*[3] *and absolutely summable and the series* $\{X_1(t), \cdots, X_k(t)\}$

[3] The boundness follows from the absolute summability in fact.

satisfies Condition A. Under these conditions the jth *cumulant of* $(\alpha a_T + \alpha^* a_T^*) \cdot (2T+1)^{1-j^{-1}}$ *approaches*

$$\sum_v \sum_\epsilon \sum_{s_2} \cdots \sum_{s_j} \alpha^{\epsilon_1} \cdots \alpha^{\epsilon_j} d_{v_1}(s_{v_1}) \cdots d_{v_p}(s_{v_p}) \tag{5.18}$$

where $\epsilon = (\epsilon_1, \cdots, \epsilon_j)$ *ranges over* $\epsilon_i =$ *"blank"* or "*", v *ranges over the indecomposable partitions* $(v_1, \cdots, v_p)$ *of the table*

$$\begin{matrix} (1,1) & \cdots & (1,k) \\ \vdots & \cdots & \vdots \\ (j,1) & \cdots & (j,k) \end{matrix} \tag{5.19}$$

and $d_{v_i}(s_{v_i})$ *denotes the joint cumulant of elements selected from the table*

$$\begin{matrix} Y_1^{\epsilon_1}(0) & \cdots & Y_k^{\epsilon_1}(0) \\ Y_1^{\epsilon_2}(s_2) & \cdots & Y_k^{\epsilon_2}(s_2) \\ \vdots & \cdots & \vdots \\ Y_1^{\epsilon_j}(s_j) & \cdots & Y_k^{\epsilon_j}(s_j) \end{matrix} \tag{5.20}$$

in accordance with v_i.

PROOF. Consider the case $j = 1$ first.

$$\begin{aligned} \mathcal{C}_1(\alpha a_T + \alpha^* a_T^*) &= E(\alpha a_T + \alpha^* a_T^*) \\ &= \alpha E a_T + \alpha^* (E a_T)^* \end{aligned} \tag{5.21}$$

where

$$\mathrm{E} a_T = \sum_{t_1} \cdots \sum_{t_k} g_1(-t_1) \cdots g_k(-t_k) c_{1\cdots k}(t_1, \cdots, t_k), \tag{5.22}$$

giving the stated result. Next, if $j > 1$, using the result of Lemma 5.1 $\mathcal{C}_j(\alpha a_T + \alpha^* a_T^*)$ equals

$$(2T+1)^{-j} \sum_{-T}^{T} \cdots \sum_{-T}^{T} \sum_v \sum_\epsilon \alpha^{\epsilon_1} \cdots \alpha^{\epsilon_j} d_{v_1}(t_{v_1}) \cdots d_{v_p}(t_{v_p}). \tag{5.23}$$

Taking advantage of the stationarity of the series involved, (5.23) may be written

$$\begin{aligned} (2T+1)^{-j} \sum_v \sum_\epsilon \sum_{s_2=-\infty}^{\infty} \cdots \sum_{s_j=-\infty}^{\infty} \sum_{t_1=-\infty}^{\infty} \alpha^{\epsilon_1} \cdots \alpha^{\epsilon_j} d_{v_1}(s_{v_1}) \\ \cdots d_{v_p}(s_{v_p}) \phi(t_1/T) \phi[(t_1+s_2)/T] \cdots \phi[(t_1+s_j)/T] \end{aligned} \tag{5.24}$$

where $\phi(x) = 1$ for $|x| \leqq 1$ and $= 0$ otherwise, and where $s_j = t_j - t_1$. In turn (5.24) equals

$$\begin{aligned} (2T+1)^{-j+1} \sum_v \sum_\epsilon \sum_{s_2} \cdots \sum_{s_j} \alpha^{\epsilon_1} \cdots \alpha^{\epsilon_j} d_{v_1}(s_{v_1}) \\ \cdots d_{v_p}(s_{v_p}) \Phi_T(s_2/T, \cdots, s_j/T) \end{aligned} \tag{5.25}$$

where $\Phi_T(s_2/T, \cdots, s_j/T)$ is given by

$$(2T+1)^{-1} \sum_t \phi(t/T) \phi[(t+s_2)/T] \cdots \phi[(t+s_j)/T]. \tag{5.26}$$

(5.26) may be seen to be measurable, uniformly bounded in T and convergent to 1. Taking advantage of the absolute summability result of Lemma 5.2, Lebesgue's bounded convergence theorem may be applied and the stated result seen to be true.

COROLLARY 5.1.1. *Under the assumptions of the theorem, a_T is asymptotically complex Gaussian with mean (5.22) and variance covariance matrix*

$$(2T+1)^{-1} \begin{Vmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{Vmatrix} \tag{5.27}$$

where

$$r_{11} = \tfrac{1}{4}\textstyle\sum_v \sum_\epsilon \sum_s d_{v_1}(s_{v_1}) \cdots d_{v_p}(s_{v_p}), \tag{5.28}$$

$$r_{12} = (1/4i)\textstyle\sum_v \sum_s [d_{v_1}(s_{v_1}) \cdots d_{v_p}(s_{v_p}) - d^*_{v_1}(s_{v_1}) \cdots d^*_{v_p}(s_{v_p})], \tag{5.29}$$

$$r_{22} = \tfrac{1}{4}\textstyle\sum_v \sum_\epsilon \sum_s d_{v_1}(s_{v_1}) \cdots d_{v_p}(s_{v_p})\alpha^{\epsilon_1}\alpha^{\epsilon_2}, \tag{5.30}$$

with $\alpha = 1/i$ in (5.30). In (5.28), (5.30) $\epsilon = (\epsilon_1, \epsilon_2)$ extends over ϵ_i = "blank" or "". In (5.28), (5.29), (5.30), v extends over all indecomposable partitions selected from the table*

$$\begin{matrix} (1, 1) & \cdots & (1, k) \\ (2, 1) & \cdots & (2, k). \end{matrix} \tag{5.31}$$

In (5.28), (5.30) $d_{v_i}(s_{v_i})$ denotes the joint cumulant of elements selected from the table

$$\begin{matrix} Y_1^{\epsilon_1}(0) & \cdots & Y_k^{\epsilon_1}(0) \\ Y_1^{\epsilon_2}(s) & \cdots & Y_k^{\epsilon_2}(s) \end{matrix} \tag{5.32}$$

in accordance with v_i. In (5.29) $d_{v_i}(s_{v_i})$ denotes the joint cumulant of elements selected from the table

$$\begin{matrix} Y_1(0) & \cdots & Y_k(0) \\ Y_1(s) & \cdots & Y_k(s) \end{matrix} \tag{5.33}$$

in accordance with v_i.

(As in the case of this corollary table (5.31) has but two rows, v extends over partitions such that at least one set of the partition has an element from both row 1 and row 2.)

This corollary results from the fact that the cumulants of order > 2 of $(2T+1)^{\frac{1}{2}}a_T$ tend to 0.

Let

$$b_T = (2T+1)^{-1}\textstyle\sum_{-T}^{T} Y_{i_1}(t) \cdots Y_{i_r}(t), \tag{5.34}$$

$$c_T = (2T+1)^{-1}\textstyle\sum_{-T}^{T} Y_{j_1}(t) \cdots Y_{j_s}(t). \tag{5.35}$$

COROLLARY 5.1.2. *Under the assumptions of the theorem, b_T and c_T are asymptotically joint complex Gaussian.*

This result may be demonstrated by considering the joint cumulants of b_T

and c_T. The result obviously extends to the case of more than two estimates as well.

COROLLARY 5.1.3. *Under the conditions of the theorem*

$$(2T + 1)^{-1}\sum_{-T}^{T} Y_j(t) \tag{5.36}$$

is asymptotically complex Gaussian for all j.

LEMMA 5.3. *The estimate* $\hat{C}_{1\cdots k}$ *given in* (5.1) *may be written in the form*

$$(2T + 1)^{-1}\sum_{-T}^{T} Y_1(t) \cdots Y_k(t) - \sum_{p>1} \hat{C}_{v_1} \cdots \hat{C}_{v_p}. \tag{5.37}$$

PROOF. $\hat{C}_{1\cdots k}$ and $\hat{m}$ (as defined in (5.2)) are actually the cumulant and moments of the random variable taking on the values

$$\{Y_1(t), \cdots, Y_k(t)\}, \qquad -T \leqq t \leqq T, \tag{5.38}$$

with probability $(2T + 1)^{-1}$. Applying the moment-cumulant relation inverse to (3.4) to the variable (5.38) yields

$$\hat{m}_{1\cdots k} = \sum_{p\geqq 1} \hat{C}_{v_1} \cdots \hat{C}_{v_p} \tag{5.39}$$

where the summation extends over all partitions of the integers 1, 2, $\cdots$, k. The stated result is now evident.

THEOREM 5.2. *Consider the estimate* $\hat{C}_{1\cdots k}$ *given at* (5.1) *where the* $g_j(t)$ *are bounded*[4] *and absolutely summable and where the series* $\{X_1(t), \cdots, X_k(t)\}$ *satisfies Condition A.* $\hat{C}_{1\cdots k}$ *is asymptotically complex Gaussian with mean* (5.22) *and variance-covariance matrix* (5.27).

PROOF. Corollary 5.1.3 indicates that the stated result is true for $k = 1$. Lemma 5.4 yields the representation (5.37). We will use this representation to prove the stated result by means of induction. Suppose that the result is true for $K \leqq k - 1$; therefore the $\hat{C}_v$ appearing on the right hand side of (5.37) are asymptotically normal. Now on consideration of $(2T + 1)^{\frac{1}{2}}\hat{C}_{v_1} \cdots \hat{C}_{v_p}$, and the rate at which the $\hat{C}_v$ are tending to asymptotic normality, one sees that

$$\hat{C}_{v_1} \cdots \hat{C}_{v_p} = \prod_{j=1}^{p} \sum \cdots \sum \tilde{g}_{v_j}(-t_{v_j})c_{v_j}(t_{v_j}) + o_p(2T + 1)^{-\frac{1}{2}}. \tag{5.40}$$

Thus the asymptotic distribution of $\hat{C}_{1\cdots k}$ is the same as that of

$$(2T + 1)^{-1} \sum_{-T}^{T} Y_1(t) \cdots Y_k(t) - \sum_{p>1} \prod_{j=1}^{p} \cdot\sum \cdots \sum \tilde{g}_{v_j}(-t_{v_j})c_{v_j}(t_{v_j}). \tag{5.41}$$

The distribution of (5.41) was derived in Theorem 5.1.

COROLLARY 5.2.1. *Consider a pair of estimates* $\hat{C}_{v_1}$, $\hat{C}_{v_2}$ *of lower order polyspectra. Under the conditions of the theorem these estimates have asymptotically a joint complex Gaussian distribution.*

THEOREM 5.3. *In the frequency domain expressions* (5.22), (5.28), (5.29), (5.30) *take the form*

[4] The boundness follows from the absolute summability in fact.

$$(5.42)\quad \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \delta(w_1 + \cdots + w_k) G_1(w_1) \cdots G_k(w_k) C_{1\cdots k}(w_1, \cdots, w_k)\, dw_1 \cdots dw_k,$$

$$(5.43)\quad \tfrac{1}{4} \int_{w_1} \int_{w_2} G_1(w_{11}) \cdots G_k(w_{1k}) G_1(w_{21}) \cdots G_k(w_{2k}) \sum_{\epsilon} \sum_{v} D_{v_1}(w_{v_1}) \cdots D_{v_p}(w_{v_p}),$$

$$(5.44)\quad (1/4i) \int_{w_1} \int_{w_2} [G_1(w_{11}) \cdots G_k(w_{1k}) G_1(w_{21}) \cdots G_k(w_{2k})] \cdot [\textstyle\sum_{v} \delta(\tilde{w}_{v_1}) C_{v_1}(w_{v_1}) \cdots \delta(\tilde{w}_{v_p}) C_{v_p}(w_{v_p}) - \delta(-\tilde{w}_{v_1}) C_{v_1}(-w_{v_1}) \cdots \delta(-\tilde{w}_{v_p}) C_{v_p}(-w_{v_p})]\, dw_{11} \cdots dw_{2k},$$

$$(5.45)\quad \tfrac{1}{4} \int_{w_1} \int_{w_2} G_1(w_{11}) \cdots G_k(w_{1k}) G_1(w_{21}) \cdots G_k(w_{2k}) \sum_{\epsilon} \sum_{v} D_{v_1}(w_{v_1}) \cdots D_{v_p}(w_{v_p}) \alpha^{\epsilon_1} \alpha^{\epsilon_2}, \quad \alpha = 1/i.$$

In (5.43), (5.45) $\epsilon = (\epsilon_1, \epsilon_2)$ *extends over* $\epsilon_i =$ "+1" or "−1". *In* (5.43), (5.44), (5.45), v *extends over all indecomposable partitions selected from table* (5.31). *In* (5.43) *and* (5.45) $D_{v_i}(w_{v_i})$ *denotes the joint cumulant of elements selected from the table*

$$(5.46)\qquad \begin{array}{c} dZ_1(\epsilon_1 w_{11}) \cdots dZ_k(\epsilon_1 w_{1k}) \\ dZ_1(\epsilon_2 w_{21}) \cdots dZ_k(\epsilon_2 w_{2k}) \end{array}$$

in accordance with v_i. $Z_i(w)$ *comes from the spectral representation of* $X_i(t)$, *and* $g_j(t)$ *is given by*

$$(5.47)\qquad (2\pi)^{-1} \int_{-\pi}^{\pi} e^{iwt} G_j(w)\, dw$$

with $G_j(w)$ *real.*

PROOF. This result may be proved by noting that

$$(5.48)\qquad Y_j(t) = \int_{-\pi}^{\pi} e^{iwt} G_j(w)\, dZ_j(w).$$

The reader will have noted that throughout this section a limiting process leading to estimates of averaged polyspectra (as in (5.42)) was used rather than one actually leading to $C_{1\cdots k}(w_1, \cdots, w_k)$. One could obtain $C_{1\cdots k}$ in the limit by letting the g_j's employed depend on T; however it is felt that the limiting procedure employed yields results more representative of the finite T case.

It is perhaps of interest to mention the paper [1] where it is shown that the moment estimates derived from a stationary Gaussian process are asymptotically Gaussian.

6. Applications of the theory. The intention of this section is to present a number of situations in which the estimation of polyspectra or associated polyspectral coefficients may be of use.

Suppose that we are interested in a real-valued time series $X(t)$. Are we wise to carry out a harmonic analysis of $X(t)$ or does some function of $X(t)$, say $\log X(t)$, have a simpler harmonic analysis? This question may be answered to a limited extent by evaluating certain polyspectra. To begin, note that many functional relationships may be approximated by relationships of the form

$$x = y + \alpha y^{k-1}, \tag{6.1}$$

where α is small. Consider the time series relationship

$$X(t) = Y(t) + \alpha[Y(t)]^{k-1}, \tag{6.2}$$

where α is small and $Y(t)$ is a simpler series than $X(t)$; simpler in the sense that cumulants of order j, $2 < j \leqq k$ are negligible. Evaluating the $(k-1)$th order polyspectrum of $X(t)$ in terms of the polyspectra of $Y(t)$, using (6.2) and retaining terms of first and lower order gives

$$C(w_1, \cdots, w_k) = \alpha(k-1)\sum f(w_{j_1}) \cdots f(w_{j_{k-1}}) \tag{6.3}$$

where $C(w_1, \cdots, w_k)$ denotes the $(k-1)$th order polyspectrum of $X(t)$, $f(w)$ denotes the power spectrum of $X(t)$, and in (6.3) the summation extends over the indices $1, \cdots, k$ taken $(k-1)$ at a time. (Remember that in the case of a single series $X(t)$, when we are considering the $(k-1)$th order polyspectrum we are really thinking of the series as $\{X(t), \cdots, X(t)\}$.) Thus we see that if a relationship of the form (6.2) holds, α is given approximately by

$$C(w_1, \cdots, w_k)/(k-1)\sum f(w_{j_1}) \cdots f(w_{j_{k-1}}). \tag{6.4}$$

This coefficient may be estimated by substituting estimates of the $(k-1)$th order polyspectrum and the power spectrum of $X(t)$ into (6.4).

In this connection we have,

THEOREM 6.1. *Let $X(t)$ denote a time series satisfying the conditions of Theorems 5.2, 5.3. Let $\hat{C}(w_1, \cdots, w_k)$, $\hat{f}_j(w_j)$ denote estimates of $C(w_1, \cdots, w_k)$, $f(w_j)$ respectively of the form of the estimates of Theorem 5.2. The random variable*

$$\hat{C}(w_1, \cdots, w_k)/(k-1)\sum \hat{f}_{j_1}(w_{j_1}) \cdots \hat{f}_{j_{k-1}}(w_{j_{k-1}}) \tag{6.5}$$

tends to

$$\frac{\int \cdots \int \delta(w_1 + \cdots + w_k)G_1(w_1) \cdots G_k(w_k)C(w_1, \cdots, w_k)\, dw_1 \cdots dw_k}{(k-1)\sum h_{j_1}(w_{j_1}) \cdots h_{j_{k-1}}(w_{j_{k-1}})} \tag{6.6}$$

in probability, where

$$h_j(w_j) = \int_{-\pi}^{\pi} |G_j(w)|^2 f(w)\, dw. \tag{6.7}$$

(6.5) *is also asymptotically complex Gaussian.*

PROOF. This theorem results from Theorem 5 and Corollary 3 of [19] and Theorems 5.2 and 5.3 of this paper.

Turning to another application of the theory, consider the following heuristic model of a frequency component being produced by the beating or multiplication together of a number of individual frequency components. Suppose we are considering real-valued time series $X_1(t), \cdots, X_k(t)$ with spectral representations

$$X_j(t) = \int e^{iwt}\, dZ_j(w). \tag{6.8}$$

Consider the following question, does the component at frequency w_k in the series $X_k(t)$ come about as the product of the components at frequencies $-w_j$ in the series $X_j(t)$, $j = 1, \cdots, k-1$, where $\sum w_j = 0$?

In terms of the spectral functions $Z_j(w)$, we are wondering if $\Delta Z_k(w_k)$ is of the approximate form

$$\beta \Delta Z_1(-w_1) \cdots \Delta Z_{k-1}(-w_{k-1}) \tag{6.9}$$

for some constant β where

$$\Delta Z_j(w_j) = Z_j(w_j + \Delta w_j) - Z_j(w_j). \tag{6.10}$$

Since the series involved are real, (6.9) may be written

$$\beta \Delta Z_1^*(w_1) \cdots \Delta Z_{k-1}^*(w_{k-1}). \tag{6.11}$$

The linear regression coefficient of $\Delta Z_k(w_k)$ on (6.10) is therefore

$$E\Delta Z_1(w_1) \cdots \Delta Z_k(w_k)/E|\Delta Z_1(w_1) \cdots \Delta Z_{k-1}(w_{k-1})|^2 \tag{6.12}$$

and the coefficient of determination is

$$\frac{|E\, \Delta Z_1(w_1) \cdots \Delta Z_k(w_k)|^2}{E\, |\Delta Z_1(w_1) \cdots \Delta Z_{k-1}(w_{k-1})|^2\, E\, |\Delta Z_k(w_k)|^2}. \tag{6.13}$$

If the w_j satisfy no relation of the form

$$\textstyle\sum_{i=1}^{s} w_{j_i} = 0 \qquad (s < k) \tag{6.14}$$

and the Δw_j are small then (6.12) and (6.13) are given by

$$C_{1\cdots k}(w_1, \cdots, w_k)/f_1(w_1) \cdots f_{k-1}(w_{k-1}) \tag{6.15}$$

and

$$[|C_{1\cdots k}(w_1, \cdots, w_k)|^2/f_1(w_1) \cdots f_k(w_k)]\cdot[|\Delta w_1 \cdots \Delta w_{k-1}|/|\Delta w_k|], \tag{6.16}$$

respectively.

Thus when one is considering the question of frequency components beating together, one is led to consider the coefficients (6.15) and

$$|C_{1\cdots k}(w_1, \cdots, w_k)|^2/f_1(w_1) \cdots f_k(w_k). \tag{6.17}$$

This latter represents the relative appropriateness at various polyfrequencies of the beating together of frequency components model. (Relative because of the additional factor in (6.16).)

These coefficients may be estimated by substituting estimates of the required polyspectra, and we may prove,

THEOREM 6.2. *Let* $\{X_1(t), \cdots, X_k(t)\}$ *denote a time series satisfying the conditions of Theorems* 5.2, 5.3. *Let* $\hat{C}_{1\cdots k}(w_1, \cdots, w_k)$, $\hat{f}_j(w_j)$ *denote estimates of*

$C_{1\cdots k}(w_1, \cdots, w_k)$, $f_j(w_j)$ *respectively of the form of the estimates considered in Theorem* 5.2. *The random variables*

$$\hat{C}_{1\cdots k}(w_1, \cdots, w_j)/\hat{f}_1(w_1) \cdots \hat{f}_{k-1}(w_{k-1}) \tag{6.18}$$

and

$$|\hat{C}_{1\cdots k}(w_1, \cdots, w_k)|^2/\hat{f}_1(w_1) \cdots \hat{f}_k(w_k) \tag{6.19}$$

tend in probability to

$$\frac{\int \cdots \int \delta(w_1 + \cdots + w_k) G_1(w_1) \cdots G_k(w_k) C_{1\cdots k}(w_1, \cdots, w_k)\, dw_1 \cdots dw_k}{h_1(w_1) \cdots h_{k-1}(w_{k-1})} \tag{6.20}$$

and

$$\frac{\left|\int \cdots \int \delta(w_1 + \cdots + w_k) G_1(w_1) \cdots G_k(w_k) C_{1\cdots k}(w_1, \cdots w_k)\, dw_1 \cdots dw_k\right|^2}{h_1(w_1) \cdots h_k(w_k)}, \tag{6.21}$$

respectively where

$$h_j(w_j) = \int |G_j(w)|^2 f_j(w)\, dw. \tag{6.22}$$

Moreover, asymptotically the estimates are joint complex Gaussian.

PROOF. The proof proceeds on the same lines as the proof of Theorem 6.1. It is perhaps of interest to point out the values of (6.15) and (6.17) in the case of one of the examples considered earlier. Suppose $X(t)$ is the process of Example 3, Section 3. In this case (6.15) and (6.17) are given by

$$[K_k/(K_2)^{k-1}] \cdot G_k(w_k)/|G_1(w_1) \cdots G_{k-1}(w_{k-1})| \tag{6.23}$$

and

$$K_k^2/K_2^k \tag{6.24}$$

and we see that an examination of the coefficients (6.19) for constancy provides a test for the model of this example.

The reader will have noted that in the derivation of the coefficients (6.15) and (6.17) it was assumed that the w_j satisfy no relation of the form (6.14). This assumption is reasonable in view of the fact that if the process satisfies an ergodicity requirement to be presented in the next section, then components whose frequencies are such that (6.14) is true, are uncorrelated with the remaining components and a relation of the form (6.9) is then inconsistent.

7. Moments or cumulants? At this point the reader is no doubt wondering why the polyspectrum was defined as the Fourier transform of the cumulant rather than of the product moment or of the central product moment. In this section a justification of this definition will be provided for a class of processes. The essential property that these processes have is a form of ergodicity.

Let us begin by noting that the Fourier transform of at most one of the product moment, central product moment or cumulant can be "nice" in the sense of being a proper function. Suppose for example that the polyspectra are proper functions and consider the Fourier transform of $m_{1\cdots k}(t_1, \cdots, t_k)$. As derived from the relation inverse to (3.4) it is,

$$(7.1) \quad \delta(w_1 + \cdots + w_k)M(w_1, \cdots, w_k) = \sum \delta(\tilde{w}_{v_1}) \cdots \delta(\tilde{w}_{v_p})C_{v_1}(w_{v_1}) \cdots C_{v_p}(w_{v_p})$$

where $\tilde{w}_v = w_{i_1} + \cdots + w_{i_j}$ if v denotes the grouping $(i_1, \cdots, i_j)$. $M(w_1, \cdots, w_k)$ is seen to contain many delta functions if the lower order polyspectra do not vanish (as the ordinary power spectrum must not). Thus we see that if the polyspectra are proper functions, then the Fourier transforms of the product moments are not. The converse of this statement may be seen to be true by considering the expansion (3.4). By considering similar expansions involving central moments, we are led to the conclusion that at most one of the definitions may lead to proper functions.

It will now be shown that for processes satisfying a form of ergodicity requirement, the property of having a proper function as a polyspectrum is not evidently inconsistent, whereas the corresponding property for moments and central moments is inconsistent. The class $\Psi^{(k)}$ introduced earlier is thus perhaps a reasonable one so far as ergodic type processes are concerned.

The following notation will be adhered to in the remainder of this section:

(a) if v denotes a group of distinct integers $(i_1, \cdots, i_j)$ selected from $(1, \cdots, k)$, then $\tilde{X}_v(t_v)$ denotes the product $X_{i_1}(t_{i_1}) \cdots X_{i_j}(t_{i_j})$,

(b) if u, v denote distinct groupings, then the refinement grouping obtained by inserting the subdivisions of u into v will be denoted by $u \otimes v$,

(c) if u is the grouping (u_1, u_2) and $t = (t_1, \cdots, t_k)$, then $\bar{t}$ will denote $(\bar{t}_1, \cdots, \bar{t}_k)$ where $\bar{t}_i = t_i + \tau$ if $i \varepsilon u_1$ and $\bar{t}_i = t_i$ if $i \varepsilon u_2$.

The process $X(t) = \{X_1(t), \cdots, X_k(t)\}$ is said to satisfy *Condition* $I(k)$ if the joint moments of order $\leq k$ exist, and for all groupings u and v (u consisting of two subgroups), and the X's corresponding to the different subgroups v_j of v being from independent realizations of $X(t)$,

$$(7.2) \qquad (2T)^{-1} \int_{-T}^{T} U(\tau)\, d\tau$$

approaches

$$(7.3) \qquad m_{r_1}(t_{r_1}) \cdots m_{r_p}(t_{r_p})$$

in probability where $r = u_1 \otimes v$ and $U(\tau)$ denotes the product of the individual X terms in $\tilde{X}_{v_1}(\bar{t}_{v_1}) \cdots \tilde{X}_{v_p}(\bar{t}_{v_p})$ involving τ.

Condition I(k) is seen to be a form of ergodicity requirement. In fact if we are concerned with a univariate weakly mixing process $X(t)$ belonging to $\Phi^{(\infty)} \cap S^{(\infty)}$, then $X(t)$ satisfies Condition I(k) for every k (see [7]).

The process $X(t) = \{X_1(t), \cdots, X_k(t)\}$ is said to satisfy *Condition* $II(k)$ if,

(i) there exists $\delta > 0$ such that for $j \leq k$ and distinct indices $i_1, \cdots, i_k$

selected from $1, \cdots, k$,

$$E|X_{i_1}(t_1) \cdots X_{i_j}(t_j)|^{1+\delta} < \infty, \tag{7.4}$$

(ii) there exists T_0 and $M > 0$ such that for $T > T_0$,

$$(2T)^{-1} \int_{-T}^{T} E_1|\tilde{X}_{v_1}(\bar{t}_{v_1})|^{1+\delta} \cdots E_p|\tilde{X}_{v_p}(\bar{t}_{v_p})|^{1+\delta} \, d\tau < M \tag{7.5}$$

for all groupings v, where the subscripts on the expected value operators denote independent realizations of the process.

The following lemma will be required.

LEMMA 7.1. *Let* $\{U_n\}$ *be a sequence of random variables tending to* μ *in probability. Let* V *be a random variable such that* (i) *for some* $\delta > 0$, $E|V|^{1+\delta}$ *exists,* (ii) *there exist* N *and* $M > 0$ *such that for* $n > N$, $E|U_nV|^{1+\delta} < M$, *then* $EU_nV \to \mu EV$.

PROOF.

$$\begin{aligned} |EU_nV - \mu EV| &= |E(U_n - \mu)V| \leqq E|(U_n - \mu)V| \\ &= \int |U_n - \mu| \cdot |V| \, dP_n(U, V) \\ &= \int_{|U-\mu| \leqq \epsilon} |U - \mu| \cdot |V| \, dP_n(U, V) \\ &\quad + \int_{|U-\mu| > \epsilon} |U - \mu| \cdot |V| \, dP_n(U, V) \end{aligned} \tag{7.6}$$

where $P_n(U, V)$ denotes the joint cdf of U_n and V. The first term in (7.6) is $\leqq \epsilon E|V|$ and consequently may be made arbitrarily small by a choice of ϵ. The second term is less than or equal to

$$\{\int_{|U-\mu|>\epsilon} dP_n(U, V)\}^{\delta/(1+\delta)} \{\int |U - \mu|^{1+\delta} |V|^{1+\delta} \, dP_n(U, V)\}^{1/(1+\delta)}. \tag{7.7}$$

The first term in (7.7) may be made arbitrarily small as a result of the convergence in probability of $\{U_n\}$ to μ, while the second term remains bounded. Consequently (7.6) may be made arbitrarily small and the lemma follows.

THEOREM 7.1. *Consider the process* $X(t) = \{X_1(t), \cdots, X_k(t)\}$ *that satisfies Conditions* $I(k)$ *and* $II(k)$. *For any groupings* $(v_1, \cdots, v_p)$ *and* (u_1, u_2) *of* $(1, \cdots, k)$,

$$\begin{aligned} \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} m_{v_1}(\bar{t}_{v_1}) \cdots m_{v_p}(\bar{t}_{v_p}) \, d\tau \\ = m_{r_1}(t_{r_1}) \cdots m_{r_p}(t_{r_p}) m_{s_1}(t_{s_1}) \cdots m_{s_p}(t_{s_p}) \end{aligned} \tag{7.8}$$

where $r = u_1 \otimes v$ *and* $s = u_2 \otimes v$.

PROOF.

$$\begin{aligned} (2T)^{-1} \int_{-T}^{T} m_{v_1}(\bar{t}_{v_1}) \cdots m_{v_p}(\bar{t}_{v_p}) \, d\tau \\ = (2T)^{-1} \int_{-T}^{T} E_1\tilde{X}_{v_1}(\bar{t}_{v_1}) \cdots E_p\tilde{X}_{v_p}(\bar{t}_{v_p}) \, d\tau \end{aligned} \tag{7.9}$$

$$= E_1 \cdots E_p (2T)^{-1} \int_{-T}^{T} \tilde{X}_{v_1}(\bar{t}_{v_1}) \cdots \tilde{X}_{v_p}(\bar{t}_{v_p}) \, d\tau, \tag{7.10}$$

since under the stated conditions Tonelli's theorem applies.

The result now follows from the lemma taking U_T to be (7.2) and $U(\tau)$ to be the product of the individual X terms in $\tilde{X}_{v_1}(\bar{t}_{v_1}) \cdots \tilde{X}_{v_p}(\bar{t}_{v_p})$ involving τ.

Let us next prove,

THEOREM 7.2. *Consider the process* $X(t) = \{X_1(t), \cdots, X_k(t)\}$ *satisfying Conditions* $I(k)$ *and* $II(k)$.

$$\lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} c_{1\cdots k}(\bar{t})\, d\tau = 0 \tag{7.11}$$

for any grouping (u_1, u_2).

PROOF.

$$c_{1\cdots k}(\bar{t}) = \sum (-1)^{p-1}(p-1)!\, m_{v_1}(\bar{t}_{v_1}) \cdots m_{v_p}(\bar{t}_{v_p}). \tag{7.12}$$

Thus,

$$(2T)^{-1} \int_{-T}^{T} c_{1\cdots k}(\bar{t})\, d\tau = \sum (-1)^{p-1}(p-1)!\, (2T)^{-1} \int_{-T}^{T} m_{v_1}(\bar{t}_{v_1}) \cdots m_{v_p}(\bar{t}_{v_p})\, d\tau. \tag{7.13}$$

From the preceding theorem this tends to

$$\sum (-1)^{p-1}(p-1)!\, m_{r_1}(t_{r_1}) \cdots m_{r_p}(t_{r_p}) m_{s_1}(t_{s_1}) \cdots m_{s_p}(t_{s_p}) \tag{7.14}$$

where $r = u_1 \otimes v$ and $s = u_2 \otimes v$.

We note that (7.14) is the joint kth order cumulant of the process $X(t) = \{X_1(t), \cdots, X_k(t)\}$ wherein the components with subscripts in u_1 are statistically independent of those with subscripts in u_2. The expression must consequently be 0 as this cumulant is 0.

Before proceeding to the next theorem, let us make one last definition:

$\Phi_k^{(k)}$ denotes the class of k-dimensional processes $X(t) = \{X_1(t), \cdots, X_k(t)\}$ with finite kth order absolute moments and such that for $v = (i_1, \cdots, i_j)$ any group of j distinct integers from $1, \cdots, k$ there exist complex totally finite measures $M_v(\Omega)$ such that

$$EX_{i_1}(t_1) \cdots X_{i_j}(t_j) = \int \cdots \int \exp\,[i(w_1 t_1 + \cdots + w_j t_j)] M_v(dw_1, \cdots, dw_j). \tag{7.15}$$

As in [24] it is possible to introduce in an obvious manner a polyspectral measure

$$\sum (-1)^{p-1}(p-1)!\, M_{v_1} \times \cdots \times M_{v_p}(\Omega) \tag{7.16}$$

for this class where Ω is a Borel set of R^k.

THEOREM 7.3. *Consider the process* $X(t) = \{X_1(t), \cdots, X_k(t)\}$ *belonging to* $\Phi_k^{(k)}$ *and satisfying Conditions* $I(k)$ *and* $II(k)$. *Given the grouping* (u_1, u_2), *let* Ω_1 *be the flat* $\tilde{w}_{u_1} = 0$, Ω_2 *the flat* $\tilde{w}_{u_2} = 0$, *and* Ω *a measurable subset of* $\Omega_1 \times \Omega_2$, *then*

$$\sum (-1)^{p-1}(p-1)!\, M_{v_1} \times \cdots \times M_{v_p}(\Omega) = 0. \tag{7.17}$$

PROOF.

$$c_{1\cdots k}(t_1, \cdots, t_k) = \int \cdots \int \exp\,[i(t_1 w_1 + \cdots + t_k w_k)] \cdot \sum (-1)^{p-1}(p-1)!\, M_{v_1} \times \cdots \times M_{v_p}(dw), \tag{7.18}$$

$$(7.19)\quad (2T)^{-1}\int_{-T}^{T} c_{1\cdots k}(\bar t_1, \cdots, \bar t_k)\,d\tau = \int\cdots\int \exp[i(t_1w_1 + \cdots + t_kw_k)] \cdot \sum (-1)^{p-1}(p-1)!\, M_{v_1}\times\cdots\times M_{v_p}(dw)(2T)^{-1}\int_{-T}^{T}\exp[i\tilde w_{v_1}\tau]d\tau.$$

But

$$(7.20)\quad \lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp[i\tilde w_{v_1}\tau]d\tau = \epsilon(\tilde w_{v_1}) = 1 \quad \tilde w_{v_1} = 0$$
$$= 0 \quad \tilde w_{v_1} \neq 0.$$

Thus we have

$$(7.21)\quad \int\cdots\int \exp[i(t_1w_1 + \cdots + t_kw_k)]\epsilon(\tilde w_{v_1}) \cdot \sum (-1)^{p-1}(p-1)!\, M_{v_1}\times\cdots\times M_{v_p}(dw) = 0 \quad \text{for all } t$$

and

$$(7.22)\quad \int\cdots\int \chi(\Omega)\epsilon(\tilde w_{v_1})\sum(-1)^{p-1}(p-1)!\,M_{v_1}\times\cdots\times M_{v_p}(dw) = 0$$

for any measurable set Ω of the flat $\sum w_i = 0$, where $\chi(\Omega)$ is the characteristic function of Ω, and we see that

$$\sum(-1)^{p-1}(p-1)!\,M_{v_1}\times\cdots\times M_{v_p}(\Omega) = 0$$

if in fact $\Omega \subseteq \Omega_1 \times \Omega_2$.

This argument parallels an argument in [25].

Now the Lebesgue measure of the set Ω of this theorem is 0, consequently the measure $\sum(-1)^{p-1}(p-1)!\,M_{v_1}\times\cdots\times M_{v_p}(dw)$ satisfies a necessary condition for it to be absolutely continuous with respect to $(k-1)$-dimensional Lebesgue measure. We conclude that the polyspectrum, which is an attempt to provide a density of this measure with respect to Lebesgue measure, is not evidently inconsistent.

The ergodicity of stationary processes is also investigated in [26].

A different type of justification of the use of cumulants is the following; in the Gaussian case all the information is contained in the first two moments. Consequently a kth order product moment $k > 2$, has no new information to provide, nor does its Fourier transform. The kth order cumulant is a function of the product moments of orders k and less which is zero in the Gaussian case. The consideration of the cumulant in this case is not liable to deceive one into believing that he has gained some information. In the non-Gaussian case the cumulant provides an indication of the non-Gaussianity. The cumulants appear to provide a form of harmonic analysis of the distribution in fact.

It seems appropriate to end the paper on a note of pessimism. Experience with real random variables indicates that higher order moments are typically not efficient estimates of scientifically relevant parameters; consequently as the specifications of stochastic processes become tighter, polyspectra are likely to prove less pertinent in a similar manner.

8. Acknowledgment. The author would like to acknowledge many conversations with Professor John W. Tukey and Dr. Michael Godfrey that were most helpful in the preparation of this paper.

REFERENCES

[1] BARTLETT, M. S. (1961). *An Introduction to Stochastic Processes*. Cambridge Univ. Press.

[2] BLACKMAN, R. B. and TUKEY, J. W. (1959). *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover, New York.

[3] BLANC-LAPIERRE, A. and FORTET, R. (1953). *Theorie des Fonctions Aleatoires*. Masson, Paris.

[4] BOCHNER, S. (1936). Summation of multiple Fourier series by spherical means. *Trans. Amer. Math. Soc.* **40** 175–207.

[5] BOCHNER, S. and CHANDRASEKHARAN, K. (1949). *Fourier Transforms*. Princeton Univ. Press.

[6] BRILLINGER, D. R. (1964). A technique for measuring the spectral density matrix of two signals. *Proc. I.E.E.E.* **52** 103–104.

[7] BRILLINGER, D. R. (1965). The ergodicity of stationary processes. Submitted to *Theor. Prob. Appl.*

[8] BRILLINGER, D. R. and HATANAKA, M. (1965). An harmonic analysis of non-stationary time series with special reference to the Permanent Income Hypothesis. To be submitted to *Econometrica*.

[9] CRAMÉR, H. (1942). On harmonic analysis in certain functional spaces. *Ark. Mat. Astr. Fys.* **28B**.

10] 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, New York University.

[11] HAMMING, R. W. (1962). *Numerical Methods for Scientists and Engineers*. McGraw-Hill, New York.

[12] HASSELMAN, K., MUNK, W., and MACDONALD, G. (1963). Bispectrum of ocean waves. *Time Series Analysis*, (M. Rosenblatt ed.) 125–139. Wiley, New York.

[13] JACKSON, D. (1930). *The Theory of Approximation*. Amer. Math. Soc., Providence.

[14] JENKINS, G. M. (1961). General considerations in the analysis of spectra. *Technometrics* **3** 133–166.

[15] JENKINS, G. M. (1963). Cross-spectral analysis and the estimation of linear open loop transfer functions. *Time Series Analysis*, (M. Rosenblatt ed.) 267–278. Wiley, New York.

[16] KENDALL, M. G. and STUART, A. (1958). *The Advanced Theory of Statistics*, **1**. Griffin, London.

[17] LEONOV, V. P. and SHIRYAEV, A. N. (1959). On a method of calculation of semi-invariants. *Theor. Prob. Appl.* **4** 319–329.

[18] MAGNESS, T. A. (1954). Spectral response of a quadratic device to non-Gaussian noise. *J. Appl. Phys.* **25** 1357–1365.

[19] MANN, H. B. and WALD, A. (1943). On stochastic limit and order relationships. *Ann. Math. Statist.* **14** 217–226.

[20] MAZUR, S. and ORLICZ, W. (1934). Grundlegende Eigenschaften der polynomischen Operationen. *Studia Math.* **5** 50–68 and 170–189.

[21] PAPOULIS, P. (1962). *The Fourier Integral and Its Applications*. McGraw-Hill, New York.

[22] PARZEN, E. (1961). Mathematical considerations in the estimation of spectra. *Technometrics* **3** 167–190.

[23] ROSENBLATT, M. and VAN NESS, J. W. (1964). Estimates of the bispectrum of stationary random processes. Brown University Technical Report. Contract NONR562(29).

[24] SHIRYAEV, A. N. (1960). Some problems in the spectral theory of higher-order moments, I. *Theor. Prob. Appl.* **5** 265–284.

[25] SHIRYAEV, A. N. (1963). On conditions for ergodicity of stationary processes in terms of higher order moments. *Theor. Prob. Appl.* **8** 436–439.

[26] SINAI, YA. G. (1963). On higher order spectral measures of ergodic stationary processes. *Theor. Prob. Appl.* **8** 429–436.

[27] SUN, T. C. (1963). A central limit theorem for non-linear functions of a normal stationary processes. *J. Math. Mech.* **12** 945–978.

[28] TUKEY, J. W. (1959). An introduction to the measurement of spectra. *Probability and Statistics*, (U. Grenander ed.) 300–330. Wiley, New York.

[29] TUKEY, J. W. (1961). Discussion, emphasizing the connection between analysis of variance and spectrum analysis. *Technometrics* **3** 191–219.

[30] TUKEY, J. W. (1963). *An Introduction to the Frequency Analysis of Time Series*. Notes on Mathematics 596, course offered at Princeton University.

[31] VAN DER TWEEL, L. H. (1961). Some problems in vision regarded with respect to linearity and frequency response. *Ann. New York Acad. Sci.* **89** 829–856.

[32] WIENER, N. (1958). *Nonlinear Problems in Random Theory*. M.I.T. Press, Cambridge.