Statistica Sinica 1(1991), 465-476

SOME HISTORY OF THE STUDY OF HIGHER-ORDER MOMENTS AND SPECTRA

David R. Brillinger

University of California, Berkeley

Abstract: Some of the history of the development of higher-moments and spectra for random processes, particularly time series and point processes, is presented for the years preceding 1980. Time-side and frequency-side ideas are contrasted. Some uses of the concepts and associated techniques are mentioned. So too are some of the computational procedures that have been employed.

Key words and phrases: Bispectrum, cumulant, cumulant spectrum, higher-order spectrum, history, product-moment, point process, polyspectrum, product density, random process, system, time series.

1. Introduction and Disclaimer

It is not easy to present the history of concepts used in diverse scientific fields. The material is inevitably limited by the writer's experience. Still it seems worth attempting, even if for no other reason than to induce others to provide their views on the matter. Since the work of the present paper is meant to be historical rather than review, consideration will be restricted to the pre-1980 period. Also, focus will be on the general case rather than the particular cases of the power spectrum, cross spectrum or bispectrum.

There are a variety of "sides" from which one can discuss the matter: theoretical-empirical, time-frequency, ordinary series-generalized process, discrete time-continuous time, computational-distributional, univariate-vector amongst others. There is insufficient space and time to cover many of these aspects in any detail, but a variety of comments will be made in attempting this.

One of the purposes of this work is to make available to young researchers a listing of some of the original sources. In many cases these can be read much more productively than works that have appeared later.

2. Second-Order Moments and Spectra

There will be minimal discussion of the second-order, that is of the power spectral or the cross spectral, cases. (The term second-order refers to the fact that the basic quantities involved are quadratic.) The principal concern, instead, is with the higher-order situation. There also exists an extensive commentary concerning the second-order case in Yaglom (1987b). The bibliography in Wold (1965) contains a listing of many pre-1960 works.

On the time-side, two early references to empirical work with the autocovariance function that will be mentioned as being of more than ordinary interest are Hooker (1901) and Taylor (1921).

By analysis on the frequency-side is meant making essential use of sinusoids, bands of (angular) frequency and Fourier transforms in the study of time series and related processes. On the frequency-side a remarkable reference, turned up by A. M. Yaglom, is Einstein (1914). This paper was greeted with surprise and excitement. In it Einstein constructed the first specifically consistent estimate of the power spectrum, although in 1891 Michaelson (see e.g. Michaelson (1907)) had already proposed a sensible estimate. Commentaries on this paper are given in Masani (1986) and Yaglom (1987a). Another early reference to frequency-side analysis, pointed out in Rice (1945), is Kenrick (1929).

3. Higher-Order Moments

The moments employed in the analysis of random processes and time series are direct extensions of those of ordinary statistics, the main properties of which are given, for example, in Kendall and Stuart (1969). The moment approach in statistics is usually identified with the name of Karl Pearson. In anticipation of later development of time series and random process techniques, we remark that Pearson's method of moments has largely been replaced by R. A. Fisher's likelihood approach as the years have passed, and this seems likely to continue.

The joint product-moment of the k-variate random variable $\mathbf{X} = (X_1, \ldots, X_k)$ is

$$E\{X_1 \cdots X_k\}.$$

The joint cumulant of the variate $\mathbf{X}$ is that elementary combination of the joint product moments of subsets of the components of $\mathbf{X}$, which vanishes if any subset of the components is statistically independent of the remainder (see Brillinger (1965)). More simply, it is the coefficient of $\theta_1 \cdots \theta_k$ in the Taylor expansion of the log moment generating function of $\mathbf{X}$. In what follows the joint cumulant will be written

$$\mathrm{cum}\{X_1, \ldots, X_k\}.$$

It has the property of vanishing for $k > 2$ in the case of jointly Gaussian variates.

In essence what has been done in the time series case is simple. The product moment function of order k of the process $X(\cdot)$ is defined by

$$m_k(t_1,\dots,t_k) = E\{X(t_1)\cdots X(t_k)\} \tag{1}$$

for $-\infty < t_j < \infty$ while the cumulant function is defined by

$$c_k(t_1,\dots,t_k) = \mathrm{cum}\{X(t_1),\dots,X(t_k)\}.$$

In case the process is stationary one has the simplification

$$m_k(t+u_1,\dots,t+u_{k-1},t) = m_k(u_1,\dots,u_{k-1})$$

and

$$c_k(t+u_1,\dots,t+u_{k-1},t) = c_k(u_1,\dots,u_{k-1}) \tag{2}$$

for $-\infty < t < \infty$. That is, one deals with functions of one fewer argument.

In fact it seems that the earliest substantial development of higher-order moments took place for point (and particle) processes. The point process parameters may be viewed as resulting from a correspondence $X(t) \equiv dN(t)$. Here points are scattered along a line and the increment $dN(t)$ takes on the values 0 or 1 depending on whether or not there is a point in the small interval $(t, t+dt]$ of the line.

For point processes the moment measures are defined via the product-moment and cumulant expressions

$$E\{dN(t_1)\cdots dN(t_k)\}, \quad \mathrm{cum}\{dN(t_1),\dots,dN(t_k)\},$$

respectively. The point process case is notable in that when the points are isolated, the first quantity here has a naive interpretation as

$$\mathrm{Prob}\{\text{point in } (t_1, t_1+dt_1) \text{ and} \cdots \text{and point in } (t_k, t_k+dt_k)\}.$$

This expression often leads to a density function, being equal to

$$p_k(t_1,\dots,t_k)dt_1\cdots dt_k \tag{3}$$

in case all the t's are distinct. The functions $p_k(\cdot)$ have appeared in various guises in the physics literature for many years. In particular the references Ursell (1927), Yvon (1935), Bogoliubov (1946), Bhabha (1950), Ramakrishnan (1950), are to be noted. In Yvon (1935) and Born and Green (1949) concern was with the molecular theory of fluids. Continuing, Rice (1945) set down such density functions in a study of the crossings of random processes, Ramakrishnan (1950) set down formal definitions, Kuznetsov and Stratonovich (1956) following the lead of Bogoliubov developed a set of correlation functions and in particular suggested the consideration of cumulant functions. Macchi (1969, 1975) makes further formal development of product densities. Daley and Milne (1973) provide a bibliography of the point process literature.

Consideration now turns to the time series case. A central idea here is that of Kolmogorov (see Shiryaev (1960)) to base analyses on the cumulant functions rather than the product moment functions. The use of the cumulant functions in the time series case may be motivated in several ways. Cumulants "remove" the lower order information in a sense because they vanish if some proper subset of the $X(t)$'s is independent of the remainder. In many cases of interest as functions of t they tend to 0 for large arguments and thus can have convenient analytic properties, e.g. integrability. They also turn up in investigations of ergodicity, e.g. Leonov (1960), Brillinger (1965). Cumulant functions provide a means by which to introduce mixing, leading to the later development of central limit results useful in suggesting statistical approaches to problems.

In connection with these developments references include Leonov (1960, 1964), Shiryaev (1960, 1963), Sinai (1963a,b) and Brillinger (1965). Shiryaev (1989) includes the following: "In the late 1950's and early 1960's Kolmogorov suggested to his pupils V. P. Leonov and A. N. Shiryaev a series of problems related to the issues of nonlinear analysis of random processes (in particular, in radio technology) which brought about the techniques of calculating cumulants under nonlinear transformations, and the development of the theory of spectral analysis of the high-order moments of stationary random processes." Other Eastern European work includes Zhurbenko (1970) introducing an alternate form of mixing condition and Zuev (1973), Statulevicus (1977) developing bounds and then using them in developing various large sample expressions.

4. Higher-Order Spectra

It often provides greater insight to note that, in the stationary case a process $X(\cdot)$ has a spectral representation

$$X(t) = \int_{-\infty}^{\infty} e^{it\lambda} dZ(\lambda) \tag{4}$$

$-\infty < t < \infty$ involving the random function $Z(\cdot)$. This representation leads directly to the definition of higher-order spectra, and specifically, when they exist, the cumulant spectra, $f_k(\cdot)$, are given by

$$\text{cum}\{dZ(\lambda_1), \ldots, dZ(\lambda_k)\} = \delta(\lambda_1 + \cdots + \lambda_k) f_k(\lambda_1, \ldots, \lambda_{k-1}) d\lambda_1 \cdots d\lambda_k \tag{5}$$

with $\delta(\cdot)$ the Dirac delta function. (The concentration of the mass on the subspace $\lambda_1 + \cdots + \lambda_k = 0$ results from the stationarity of the process.) The indirect definition of the cumulant spectrum f_k is as the Fourier transform of the cumulant function of the right hand side of expression (2).

Moments of order k of the dZ are considered in the seminal work of Blanc-Lapierre and Fortet (1953). It is interesting that in counterdistinction to the second-order case, the cumulant functions do not necessarily have a representation as the Fourier transform of a measure (see Kolmogorov (1960) and Sinai (1963a)), that is f_k of (5) may have to be treated as a generalized function of some type. In many cases of interest, however, the higher-order spectra are proper functions.

Turning to an empirical aspect, Brillinger and Rosenblatt (1967a,b) develop the result

$$\mathrm{cum}\{d^T(\lambda_1),\dots,d^T(\lambda_k)\} \approx (2\pi)^{k-1}\Delta^T(\lambda_1+\cdots+\lambda_k)f_k(\lambda_1,\dots,\lambda_{k-1})$$

for the case of discrete time, where

$$\Delta^T(\lambda)=\sum_{t=0}^{T-1} e^{-i\lambda t} \quad \text{and} \quad d^T(\lambda)=\sum_{t=0}^{T-1} e^{-i\lambda t}X(t).$$

This result suggests that even had cumulant spectra not been defined in their own right, researchers would have been led to them as they developed the statistical properties of empirical Fourier transforms. Brillinger (1965) and Brillinger and Rosenblatt (1967a,b) present estimates of cumulant spectra of general order and develop some properties of those estimates including their asymptotic independence and normality.

The case of $k = 3$ has been studied in some detail. In particular one may mention the works of Tukey (1953, 1959), MacDonald (1963), Hasselman et al. (1963), Godfrey (1965), Rosenblatt and Van Ness (1965), Van Ness (1966a), Shaman (1966), Hinich and Clay (1968), Kleiner (1971) and Subba Rao and Gabr (1980). For example Tukey (1953) constructs a process with a general function as bispectrum. A bibliography for the case $k = 3$ has been prepared by Tryon (1981).

In the case of a stationary point process there is a spectral representation analogous to (4) above, namely

$$N(t)=\int \frac{e^{it\lambda}-1}{i\lambda}\,dZ(\lambda). \tag{6}$$

The cumulant spectra are again given by (5) and the spectral representation is seen to provide a unifying treatment of the time series and point process cases.

5. Terminology

J. W. Tukey has introduced many of the terms of spectral analysis. In particular he called the spectra of a single series for $k = 3,4$ the bi- and tri-spectra respectively. He introduced the general term–polyspectrum. Other terms commonly employed for the concept are cumulant spectrum and higher- or kth order spectrum. Tukey seems also to have introduced the terms bifrequency and bicoherence in the case $k = 3$ for (λ_1, λ_2) and $|f_3(\lambda_1, \lambda_2)|^2/[f_2(\lambda_1)f_2(\lambda_2)f_2(\lambda_1 + \lambda_2)]$ respectively.

There has been some difference of opinion over whether, for example, the bispectral case ($k = 3$) should be referred to as second-order because of the essential dependence on just two arguments. This shows itself, for example in the use of the prefix "bi".

The moment functions (1) and (3) are often called correlation functions in the physical sciences literature, even in the case of general k (see e.g. Glauber (1963)).

6. Some Uses

Single series

Brillinger (1965) points out several uses of higher-order cumulant spectra. For example, they may be used to examine a process for Gaussianity and they may be used to examine a process for linearity. The latter use is investigated in some detail for the case $k = 3$ by Subba Rao and Gabr (1980). Brillinger (1965) also indicates how higher-order spectra might be employed in looking back at the genesis of an observed series from more elementary series. Peaks in the second-order spectrum at frequencies in elementary relation are suggestive of the operation of a nonlinearity at some earlier stage as follows, for example, from the result that if $X(\cdot)$ is Gaussian with mean 0 and second-order spectrum $f(\cdot)$ then the process $X(t)^2$ has power spectrum

$$2\int_{-\infty}^{\infty} f(\lambda - \alpha)f(\alpha)d\alpha$$

which will show a peak at $\beta + \gamma$ if $f(\cdot)$ has peaks at β and γ.

Van Ness (1966b) mentions how polynomial functional expansions may be employed for prediction. Lii et al. (1976) and Rosenblatt (1978) show how bispectra occur in connection with energy transfer between distinct frequencies – a phenomenon not possible with linear systems. Rosenblatt (1979) studies how the bispectrum may be employed to estimate the phase function of a nonGaussian linear process. Lumley and Takeuchi (1976) investigate the higher-order spectra of turbulent flows. Higher-order spectra of Gaussian series vanish. Hence by working with higher-order spectra, additive Gaussian noise is "removed". Their use in studying nonGaussian series is likewise apparent. Higher-order spectra are useful for detecting nonlinearities.

As a further use, one can note that higher-order spectra appear in the variances of estimates of lower-order moment and cumulant functions and so must be estimated to provide indications of the latter's uncertainty.

System identification

In his book, Wiener (1958) sets down a "polynomial" representation for a system with Gaussian white noise input and discusses the analysis and synthesis of the system. Such representations are often called Volterra expansions.

Tick (1961) considers the particular case of the identification of a quadratic system with Gaussian process input. Specifically he considers a system

$$Y(t) = a_0 + \sum_u a_1(u)X(t-u) + \sum_u \sum_v a_2(u,v)X(t-u)X(t-v) + \varepsilon(t)$$

with $X(\cdot)$ stationary Gaussian and with $\varepsilon(\cdot)$ a noise series independent of X. For such a system it may be shown that the second- and third-order cross-spectra are given by

$$f_{YX}(\lambda) = A_1(\lambda) f_{XX}(\lambda) \tag{7}$$

and

$$f_{XXY}(\lambda,\mu) = 2A_2(-\lambda,-\mu) f_{XX}(\lambda) f_{XX}(\mu) \tag{8}$$

respectively. Here A_1 and A_2 are the Fourier transforms of a_1 and a_2. In doctoral theses, Feuerverger (1972) and Gasser (1972) develop further aspects of the identification of quadratic systems. In particular, Feuerverger determines some statistical properties of estimates of A_1, A_2 developed from (7) and (8).

Lee and Schetzen (1965) set down a way to estimate the kernels of Wiener's polynomial expansion by cross-correlating Gaussian white noise input with the output. An early computation of a second-order kernel is given in Stark (1963). The book by Marmarelis and Marmarelis (1978) presents details and many examples of the use of the Lee-Schetzen method. Priestley (1978) gives expressions for the transfer functions which arise in the Volterra expansions of a bilinear model.

Point processes

In the case of a point process, Davies (1977) makes use of the product densities to examine a point process for Poissonness. Fourth-order product densities are estimated systematically for seismic and astronomical data in Kagan (1981) and Fry and Peebles (1978) respectively. Ogura (1972) introduces a Volterra expansion based on the Poisson process.

7. Computational Procedures

Higher-order spectra may be estimated in a variety of fashions and higher-order cumulant function estimates may be computed in at least two.

The product moments may be estimated directly by the method of moments i.e. by equating sample and population moments. Then one way to estimate the cumulant functions is to simply substitute into the formula giving cumulants in terms of product-moments. An indirect fashion is to inverse-Fourier transform the periodogram of order k avoiding the submanifolds in which proper subsets of frequencies sum to 0 or in the discrete time case sum to a multiple of 2π.

Brillinger (1965) suggests estimating higher-order spectra by complex demodulation, by narrow-band filtering, or by computing windowed Fourier transforms of empirical cumulants. Brillinger and Rosenblatt (1967b) suggest estimating cumulant spectra by smoothing higher-order periodograms avoiding the submanifolds on which the lower order spectra were concentrated. Lii et al. (1976) in the case of $k = 3$ suggest estimating the bispectrum by averaging third-order periodograms based on separate time stretches of the series. This technique has the further advantage of allowing one to estimate the variability of the estimate directly.

It may be remarked that for the last form of estimate, tapering the data before computing the Fourier transform can be quite crucial. It goes almost without remarking that a fast Fourier transform (FFT) can be central to the computations required. As this is a historical paper it would be remiss not to point out the reference Heideman et al. (1984) which makes the case that Gauss knew about an FFT in 1805.

Another issue arising in the computations is how the fundamental domain of computation is restricted via the periodicities and symmetries present. A related idea is that of aliasing. Meaningful research seems to have been done on this topic, so far, only for the cases of $k = 2, 3, 4$.

8. Extensions

There are near-immediate extensions of the concepts of higher-order moments and spectra to spatial, particle, generalized and stationary increment processes. Streater and Wightman (1964) consider higher-order moments in a quantum mechanical situation. There are further generalizations to hybrid processes of the type $X(\tau_j)$ where $X(\cdot)$ is an ordinary process and τ_j a point process. Other types of nonlinear systems, e.g. those containing an instantaneous nonlinearity or bilinear systems, may be studied by higher-order spectra. The second-order procedure that Whittle (1953) introduces for estimating finite-dimensional parameters may be extended to the higher-order case.

9. Discussion

Availability of large samples has allowed higher-order spectra to enter practice. One sees engineers putting the ideas into practice. At the same time theoreticians have continued to develop the formalism.

Key ideas that may be recognized on the time-side are the suggestion of Kuznetsov-Stratonovich-Kolmogorov that cumulant functions be the basic entities employed (many had considered product moment functions) and that of Wiener-Lee-Schetzen that polynomial systems may be identified via Gaussian white noise input and cross-correlation. A key idea on the frequency-side is that of Blanc-Lapierre and Fortet of considering moments of the $dZ(\cdot)$ variates.

It is interesting to note the parallel development of the ideas for the time series and point process cases. It is also interesting to see some of the ideas coming out of practical engineering problems while others result from strictly mathematical development. It is further interesting to note the corresponding Eastern European and North American work on the topic.

Given the history, one can speculate on future developments. These would seem to include discovery and study of special stochastic processes, concern for efficient procedures, robust-resistant techniques, and more subtle estimation of the uncertainty of estimates. Undoubtedly there will be syntheses of concepts that at this point seem distinct.

Acknowledgement

This research was supported by the National Science Foundation Grants MCS-8316634 and DMS-8900613. Murray Rosenblatt and John Tukey made helpful comments.

References

Bhabha, H. J. (1950). On the stochastic theory of continuous parametric systems and its application to electron cascades. *Proc. Roy. Soc. London Ser.A* **202**, 301-332.

Blanc-Lapierre, A. and Fortet, R. (1953). *Théorie des Fonctions Aléatoires.* Masson, Paris.

Bogoliubov, N. N. (1946). *Problems of a Dynamical Theory in Statistical Physics.* Gostekhizdat, Moscow.

Born, M. and Green, H. S. (1949). *A General Kinetic theory of Liquids.* Cambridge University Press, Cambridge.

Brillinger, D. R. (1965). An introduction to polyspectra. *Ann. Math. Statist.* **36**, 1351-1374.

Brillinger, D. R. and Rosenblatt, M. (1967a). Asymptotic theory of estimates of k-th order spectra. In *Spectral Analysis of Time Series* (Edited by B. Harris), 153-188. John Wiley, New York.

Brillinger, D. R. and Rosenblatt, M. (1967b). Computation and interpretation of k-th order spectra. In *Spectral Analysis of Time Series* (Edited by B. Harris), 189-232. John Wiley, New York.

Daley, D. J. and Milne, R. K. (1973). The theory of point processes : A bibliography. *Internat. Statist. Rev.* **41**, 183-201.

Davies, R. B. (1977). Testing the hypothesis that a point process is Poisson. *Adv. in Appl. Probab.* **9**, 724-746.

Einstein, A. (1914). Méthode pour la determination de valeurs statistiques d'observations concernant grandeurs soumises à des fluctuations irrégulières. *Arch. Sci. Phys. Natur.* **37**, 254-255. This paper is reproduced in Masani (1986).

Feuerverger, A. (1972). On the Cumulant Spectra Approach to Polynomial Regression of Stationary Time Series. Ph.D. Thesis, University of California, Berkeley.

Fry, J. N. and Peebles, P. J. E. (1978). Statistical analysis of catalogs of extragalactic objects. IX. The four-point galaxy correlation function. *Astrophys. J.* **221**, 19-33.

Gasser, T. (1972). System Identification, Polyspectra and Related Functions. Ph.D. Thesis, Swiss Federal Institute of Technology, Zurich.

Glauber, R. J. (1963). The quantum theory of optical coherence. *Phys. Rev.* **130**, 2529-2539.

Godfrey, M. D. (1965). An exploratory study of the bi-spectrum of economic time series. *Appl. Statist.* **14**, 48-69.

Hasselman, K., Munk, W. and MacDonald, G. J. F. (1963). Bispectra of ocean waves. In *Time Series Analysis* (Edited by M. Rosenblatt), 125-139. John Wiley, New York.

Heideman, M. T., Johnson, D. H. and Burrus, C. S. (1984). Gauss and the history of the fast Fourier transform. *IEEE ASSP Mag.* 14-21.

Hinich, M. J. and Clay, C. S. (1968). The application of the discrete Fourier transform in the estimation of power spectra, coherence and bispectra of geophysical data. *Rev. Geophys.* **6**, 347-363.

Hooker, R. H. (1901). Correlation of marriage-rate with trade. *J. Roy. Statist. Soc. Ser.A* **64**, 485-492.

Kagan, Y. Y. (1981). Spatial distribution of earthquakes: the four-point moment function. *Geophys. J. Roy. Astr. Soc.* **67**, 719-733.

Kendall, M. G. and Stuart, A. (1969). *The Advanced Theory of Statistics*, Vol. 1, 3rd edition. Griffin, London.

Kenrick, G. (1929). The analysis of irregular motions with applications to the energy frequency spectrum of static and of telegraph signals. *Philos. Mag. Ser.7* **7**, 176-196.

Kleiner, B. (1971). Die Berechnung von Bispektren. Ph.D. Thesis, Swiss Federal Institute of Technology, Zurich.

Kolmogorov, A. N. (1960). On the $\Phi^{(n)}$ classes of Fortet and Blanc-Lapierre. *Theory Probab. Appl.* **5**, 337.

Kuznetsov, P. I. and Stratonovich, R. L. (1956). A note on the mathematical theory of correlated random points. In *Non-Linear Transformations of Stochastic Processes* (Edited by P. I. Kuznetsov, R. L. Stratonovich and V. I. Tikhonov, 1965), 101-115. Pergammon, Oxford.

Lee, Y. W. and Schetzen, M. (1965). Measurement of the Wiener kernels of a nonlinear system by cross-correlation. *Internat. J. Control* **2**, 237-254.

Leonov, V. P. (1960). The use of the characteristic functional and semi-invariants in the ergodic theory of stationary processes. *Soviet Math.* **1**, 878-881.

Leonov, V. P. (1964). *Some Applications of Higher Semi-Invariants to the Theory of Stationary Random Processes.* Izdatistvo Nauka, Moscow.

Lii, K. S., Rosenblatt, M. and Van Atta, C. (1976). Bispectral measurements in turbulence. *J. Fluid Mech.* **77**, 45-62.

Lumley, J. L. and Takeuchi, K. (1976). Application of central-limit theorems to turbulence and higher-order spectra. *J. Fluid Mech.* **74**, 433-468.

Macchi, O. (1969). Processus ponctuels et multicoincidences. *Comptes rendus Acad. Sci. Paris Ser.A* **268**, 1616-1619.

Macchi, O. (1975). The coincidence approach to stochastic point processes. *Adv. in Appl. Probab.* **7**, 83-122.

MacDonald, G. J. F. (1963). The bispectra of atmospheric pressure records. In *Proceedings IBM Scientific Computing Symposium on Statistics*, 247-264. IBM, White Plains.

Marmarelis, P. Z. and Marmarelis, V. Z. (1978). *Analysis of Physiological Systems.* Plenum, New York.

Masani, P. R. (1986). Einstein's contribution to Generalized Harmonic Analysis and his intellectual kinship with Norbert Wiener. *Jahrbuch Uberlicke Mathematik* **109**, 190-209.

Michaelson, A. A. (1907). *Light Waves and Their Uses.* Chicago University Press, Chicago.

Ogura, H. (1972). Orthogonal functionals of the Poisson process. *IEEE Trans. Inform. Theory* **18**, 473-481.

Priestley, M. B. (1978). Non-linear models in time series analysis. *The Statistician* **27**, 159-176.

Ramakrishnan, A. (1950). Stochastic processes relating to particles distributed in a continuous infinity of states. *Proc. Cambridge Philos. Soc.* **46**, 595-602.

Rice, S. O. (1945). Mathematical analysis of random noise. *Bell System Technical J.* **24**, 46-156.

Rosenblatt, M. (1978). Energy transfer for the Burgers' equation. *Phys. Fluids* **21**, 1694-1697.

Rosenblatt, M. (1979). Linearity and nonlinearity in time series: prediction. *Proc. 42nd Session Internat. Statist. Inst.* **1**, 423-434.

Rosenblatt, M. and Van Ness, J. W. (1965). Estimation of the bispectrum. *Ann. Math. Statist.* **36**, 1120-1136.

Shaman, P. (1966). Large-sample approximations to the first- and second-order moments of bispectral estimates. Technical Report, School of Engineering, New York University.

Shiryaev, A. N. (1960). Some problems in the spectral theory of higher-order moments, I. *Theory Probab. Appl.* **5**, 265-284.

Shiryaev, A. N. (1963). On conditions for ergodicity of stationary processes in terms of higher order moments. *Theory Probab. Appl.* **8**, 436-439.

Shiryaev, A. N. (1989). Kolmogorov: life and creative activities. *Ann. Probab.* **17**, 866-944.

Sinai, Ya. G. (1963a). Properties of spectra of ergodic dynamical systems. *Soviet Math.* **4**, 875-877.

Sinai, Ya. G. (1963b). On higher order spectral measures of ergodic stationary processes. *Theory Probab. Appl.* **8**, 429-436.

Stark, L. (1963). Black-box description and physical element identification in the pupil system. *Quarterly Progress Report*, Research Laboratory for Electronics, Massachusetts Institute of Technology, **67**, 247-250.

Statulevicus, V. A. (1977). Application of semi-invariants to asymptotic analysis of distributions of random processes. In *Multivariate Analysis*, IV (Edited by P. R. Krishnaiah), 325-337. North-Holland, Amsterdam.

Streater, R. F. and Wightman, A. S. (1964). *PCT, Spin & Statistics, and All That.* Benjamin, New York.

Subba Rao, T. and Gabr, M. M. (1980). *An Introduction to Bispectral Analysis and Bilinear Time Series Models.* Lecture Notes in Statistics **24**, Springer-Verlag, New York.

Taylor, G. I. (1921). Diffusion by continuous movements. *Proc. London Math. Soc. Ser.2* **20**, 196-212.

Tick, L. J. (1961). The estimation of "transfer functions" of quadratic systems. *Technometrics* **3**, 563-567.

Tryon, P. V. (1981). The bispectrum and higher-order spectra: A bibliography. Technical Note 1036, National Bureau of Standards, USA.

Tukey, J. W. (1953). The spectral representation and transformation properties of the higher moments of stationary time series. Reprinted in *The Collected Works of John W. Tukey*, Vol. 1 (Edited by Brillinger, D. R.), 165-184. Wadsworth, Belmont.

Tukey, J. W. (1959). An introduction to the measurement of spectra. In *Probability and Statistics: The Harald Cramer Volume* (Edited by U. Grenander), 300-330. Almqvist and Wiksell, Stockholm.

Ursell, H. D. (1927). The evaluation of Gibbs' phase-integral for imperfect gases. *Proc. Cambridge Philos. Soc.* **23**, 685-697.

Van Ness, J. W. (1966a). Asymptotic normality of bispectral estimates. *Ann. Math. Statist.* **37**, 1257-1272.

Van Ness, J. W. (1966b). Empirical nonlinear prediction and polyspectra. Technical Report No. 18, Statistics Department, Stanford University.

Whittle, P. (1953). The analysis of multiple stationary time series. *J. Roy. Statist. Soc. Ser.B* **15**, 125-139.

Wiener, N. (1958). *Nonlinear Problems in Random Theory.* MIT Press, Cambridge.

Wold, H. (1965). *Bibliography on Time Series and Stochastic Processes.* Oliver and Boyd, Edinburgh.

Yaglom, A. M. (1987a). Einstein's 1914 paper on the theory of irregularly fluctuating series of observations. *IEEE ASSP Mag.* 7-11.

Yaglom, A. M. (1987b). *Correlation Theory of Stationary and Related Random Functions*, Two Volumes. Springer-Verlag, New York.

Yvon, J. (1935). *La Théorie Statistique des Fluids et l'Equation d'Etat.* Herman, Paris.

Zhurbenko, I. K. (1970). On the estimation of mixed semi-invariants for a certain class of random processes. *Theory Probab. Appl.* **15**, 525-528.

Zuev, N. M. (1973). Bounds of mixed semi-invariants of random processes. *Math. Notes Acad. Sci. USSR* **13**, 351-354.

Department of Statistics, University of California, Berkeley, CA 94720, U.S.A.

(Received January 1990; accepted January 1991)