The Cumulant Theory of Cyclostationary Time-Series, Part I: Foundation

William A. Gardner, *Fellow, IEEE,* and Chad M. Spooner, *Member, IEEE*

Abstract— **The problem of characterizing the sine-wave components in the output of a polynomial nonlinear system with a cyclostationary random time-series input is investigated. The concept of a pure nth-order sine wave is introduced, and it is shown that pure nth-order sine-wave strengths in the output time-series are given by scaled Fourier coefficients of the polyperiodic temporal cumulant of the input time-series. The higher order moments and cumulants of narrowband spectral components of time-series are defined and then idealized to the case of infinitesimal bandwidth. Such spectral moments and cumulants are shown to be characterized by the Fourier transforms of the temporal moments and cumulants of the time-series. It is established that the temporal and spectral cumulants have certain mathematical and practical advantages over their moment counterparts. To put the contributions of the paper in perspective, a uniquely comprehensive historical survey that traces the development of the ideas of temporal and spectral cumulants from their inception is provided.**

I. Introduction

THIS paper lays the foundation required for tackling problems in the general area of nonlinear processing of random signals with underlying periodicities, which are often called *cyclostationary signals*. Such signals commonly arise in communication, telemetry, radar, sonar, and control systems, and in various scientific disciplines that require analysis and processing of random measurement data obtained from systems subject to seasonal and other rhythmic variations, such as electrocardiograms and other physiological measurements, climatic, oceanic, meteorologic, and hydrologic data, and so on [33], [34].

The theory of linear and quadratic processing of cyclostationary signals is quite new and unfamiliar to many. It also is more technical than its more familiar counterpart for stationary signals. Consequently, this paper on more general nonlinear processing, which is more technical yet and deals with new concepts and results from mathematical statistics, must necessarily be tutorial and more lengthy than the typical research paper in the field of signal processing.

The central issue in this paper is the generation (and, in Part II, also in this issue, the utilization) of finite-strength additive sine-wave components from random data by nonlinear transformation, and we begin our development of the theoretical foundation for this in Section II. But first we motivate this area of study and place it in proper perspective relative to other areas of study with which the reader is assumed to be familiar.

The theory of generation of sine waves by using quadratic nonlinearities has been thoroughly laid out in [34]. Applications of this theory include weak-signal detection [36], [41], parameter estimation [43], spatial filtering [1], direction finding [91], system identification [38], and polyperiodic linear filtering [42] (see also [45]). Algorithms for these tasks that are designed using the theory of quadratically regenerated sine waves are particularly attractive when the signal of interest is heavily corrupted by noise and interfering signals because the regenerated sine-wave measurements are asymptotically independent of the noise and interference and are, therefore, tolerant to noise and interference in practice. Thus, the exploitation of regenerated sine waves allows the design of signal processing algorithms that exhibit a kind of signal selectivity. However, there are some signals from which sine waves cannot be generated by using a quadratic nonlinearity, but from which sine waves can be generated by using higher order nonlinearities. For example, no sine waves can be generated from a pulse-amplitude-modulated (PAM) signal with bandwidth equal to the Nyquist rate (e.g., as in partial-response signaling) by using a quadratic nonlinearity, but a sine wave with frequency equal to the symbol rate can be generated by using quartic nonlinearities. As another example, no sine waves with frequencies related to the carrier frequency can be generated from a balanced quadrature-phase-shift-keyed signal by using quadratic nonlinearities, but a sine wave with frequency equal to four times the carrier frequency can be generated by using a quartic nonlinearity. Thus, in order to take advantage of the desirable properties of sine-wave regenerative signal processing algorithms—namely, signal-selectivity and noise tolerance—we need to consider higher order nonlinearities.

Let us now consider the relationship of the study of higher order sine-wave generation to other parts of signal processing theory. The analysis of a signal into a set of spectral (sine-wave) components is a conceptually and mathematically useful concept. If the signal is periodic, such an analysis allows a representation in terms of a countable set of numerical values—the Fourier coefficients—as opposed to the uncountable set that is required for a pointwise representation of the signal over a single period. If the signal is erratic rather than periodic,

Manuscript received March 24, 1992; revised April 26, 1994. This work was supported by National Science Foundation under Grants MIP-88-12902 and MIP-91-12800, the United States Army Research Office, under Contract DAAL03-91-C-0018; and by E-Systems, Inc. (Greenville Division). The associate editor coordinating the review of this paper and approving it for publication was Prof. Jose A. R. Fonollosa.

W. A. Gardner is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA.

C. M. Spooner is with Mission Research Corporation, Monterey, CA 93940 USA.

IEEE Log Number 9406035.

but is persistent and has the property of stationarity, the power spectral density (PSD) is a useful spectral description in that it gives the average squared magnitudes of infinitesimal spectral components normalized by their infinitesimal bandwidths. Moreover, the effects of linear time-invariant operations on the signal can be easily expressed and understood through the use of the PSD. In the case of a periodic signal, a Fourier-series analysis gives the complex-valued strengths (magnitudes and phases) of the sine waves that comprise the signal, whereas in the case of a stationary erratic signal, the Wiener relation between the PSD and the autocorrelation function reveals that the PSD can be obtained by applying the Fourier transform to the autocorrelation function. Thus, we analyze either the signal itself or a probabilistic function of the signal—the autocorrelation—into spectral components that are finite and denumerable in the former case, and infinitesimal and nondenumerable in the latter.

The theory of cyclostationary signals combines these two different, but related, analysis concepts. A cyclostationary signal is itself erratic (except in uninteresting degenerate cases), but its autocorrelation function varies periodically or polyperiodically (with multiple incommensurate periods) with time. We use a Fourier-series analysis to determine the strengths of the sine-wave components that make up the (poly)periodically time-varying autocorrelation, and we use the Fourier transform of the resultant lag-dependent Fourier coefficients to obtain a generalization of the PSD. The theory of cyclostationary signals that is based on this time-varying autocorrelation function and its spectral characterization is called the theory of second-order cyclostationarity (SOCS), and is developed in [33], [34].

The theory of higher order statistics (HOS) characterizes the higher-than-second-order probabilistic functions (such as moments) of stationary signals, and measurements thereof [10]–[12], [66], [67], [72], [75] (and therefore can be called higher order stationarity). Similarly, the theory of higher order cyclostationarity (HOCS) characterizes the higher-than-second-order probabilistic functions of cyclostationary signals. These functions vary periodically or polyperiodically; hence they can be analyzed into Fourier series components. The individual Fourier coefficients of these time-varying functions can then be Fourier transformed in multiple dimensions to obtain the average higher order spectral behavior of the signal. These Fourier transforms also give the averages of products of multiple bandwidth-normalized infinitesimal spectral components of the erratic signal (which is a generalization of the Wiener relation to higher-than-second-order moments). Thus, HOCS deals with the time- and frequency-domain characterization of the strengths of the sine-wave components of (poly)periodically time-varying higher order probabilistic functions of cyclostationary signals. As such, it is more general than HOS in that HOS is subsumed by HOCS.

We have been purposely vague about our use of the phrase *erratic signal.* This is because there is another difference between the theory of HOCS and that of HOS (as it currently exists in the literature), namely, the difference in the mathematical framework used for the analyses described above. In the existing theory of HOS, the framework is that of stochastic processes (the ensemble-average framework), whereas in the theory of HOCS it is that of time-series (the time-average framework). Both HOS and HOCS can be studied within either framework, and therefore, *erratic signal* can mean either *stochastic process* or *time-series* in the preceding paragraphs; however, historically HOS has always been formulated in terms of stochastic processes, whereas the time-average framework of time-series has been preferred for practical reasons in the formulation of HOCS [39], [96], [99] (and in the autocorrelation theory of cyclostationarity (SOCS) [34]). If the process under consideration is stationary and ergodic, and the time-series under consideration can be viewed as a sample path of this process, then results from one HOS framework are generally true in the other (i.e., with probability equal to one). Similarly, if the process is cyclostationary and cycloergodic (terms that are defined subsequently), and if the time-series can be viewed as a sample path of this process, then results from one HOCS framework are again generally true in the other. Our reasons for adopting the time-average framework are as follows:

a) The time-average framework is conceptually and mathematically closer to the practice of signal processing.
b) The stochastic process framework entails a level of often unnecessary (although sometimes useful) abstraction.
c) The time-average framework avoids the pitfalls associated with using (unknowingly or otherwise) stochastic process models that are not ergodic or cycloergodic, cf. [45].

In summary, an *erratic signal* in HOS is a stochastic process, and the tool that defines the probabilistic functions is the mathematical operation of expectation, which requires a hypothetical ensemble of realizations chosen according to a probability rule (a probability measure that gives rise to probability density functions (PDF's)), whereas an *erratic signal* in HOCS is a single time-series with infinite duration, and the tool that defines the probabilistic functions is the infinite-time averaging operation, which leads to fraction-of-time probabilities.

We have also been purposely vague about the phrase *probabilistic function*. This is because we mean two different things by this one phrase. The first is the familiar *moment*, such as the mean or second moment of a random variable, which is the average (ensemble or time) of a product of quantities. The second thing that we mean by *probabilistic function* is the not-so-familiar *cumulant* (also called *semi-invariant, half-invariant, cumulative moment function, generalized correlation function*). The nth-order cumulant of a random variable is a nonlinear function of its moments for orders 1 through n, and the nth-order moment of a random variable is a nonlinear function of its cumulants for orders 1 through n.

Cumulants of stationary stochastic processes are the basic parameters in the theory of HOS primarily because the cumulants of Gaussian random variables and, therefore, stochastic processes for orders greater than two are identically zero, whereas the moments are not. A second reason for the central position of cumulants in HOS is that the cumulant for a sum of statistically independent random variables is the sum of the cumulants for each variable: cumulants are cumulative, whereas moments are not.

Thus, higher order cumulant measurements (rather than moment measurements) have a natural tolerance to Gaussian noise and interference that may be corrupting a signal of interest. It has also been recognized by several early researchers [10], [92] that cumulants can be used to construct well-behaved higher order spectral parameters (called *polyspectra*), whereas moments cannot, in general. Moments, however, are also important in that cumulants can be estimated by properly estimating and then combining moments.

Cumulants are fundamental in HOCS for the above reasons and for an additional important reason which is the reason attention in our study of cyclostationarity was originally drawn to them: we have derived the cumulant as the solution to a particular measurement problem associated with sine-wave generation, which is described in Section II.

Now that we have removed the intentional ambiguities in the terminology of the first part of this introduction, we can state precisely what the study of HOCS consists of: the frequency- and time-domain characterization of the strengths (magnitudes and phases) of the sine-wave components in (poly)periodically time-varying higher order moments and cumulants of time-series by using the time-average operation as the basic analysis tool. In the next section, the fundamental mathematical definitions and concepts of the time-average framework are briefly reviewed. More complete treatments are given in [34] and [40]. Following this is a brief but uniquely comprehensive historical survey tracing the development of the cumulant in statistics and engineering from its inception at the turn of the century to the present. This background material in Section I sets the stage for Section II, wherein the motivating problem of sine-wave generation by nonlinear transformation is introduced. The remainder of Section II and its companion Section III develop the time- and frequency-domain characterizations of nonlinearly generatable sine waves. The results of Sections II and III are discussed in Section IV, and conclusions are drawn in Section V.

A. Fraction-of-Time Probability

In direct analogy with the conventional probability distribution function that is defined and/or interpreted as an ensemble average, the nth-order fraction-of-time (FOT) probability distribution function for the real time-series $x(t)$, $t \in (-\infty, \infty)$ is defined by

$$F_{\boldsymbol{x}(t)}(\boldsymbol{y}) \triangleq \hat{E}^{\{\alpha\}}\left\{\prod_{j=1}^{n} U[y_j - x(t+t_j)]\right\} \tag{1}$$

where

$$\hat{E}^{\{\alpha\}}\{z(t)\} \triangleq \sum_{\alpha}\left\langle z(t+u)e^{-i2\pi\alpha u}\right\rangle = \sum_{\alpha}\left\langle z(u)e^{-i2\pi\alpha u}\right\rangle e^{i2\pi\alpha t} \tag{2}$$

is the multiple sine-wave extraction operation and

$$\langle w(u)\rangle \triangleq \lim_{Z\to\infty}\frac{1}{Z}\int_{Z/2}^{Z/2} w(u)\,du$$

is the usual time-averaging operation. The sum in (2) is over all values of α (which are assumed to be denumerable in number) that result in nonzero terms. In (1), $U[\cdot]$ is simply the event-indicator function

$$U[y_j - x(t+t_j)] \triangleq \begin{cases}1, & x(t+t_j) < y_j\\ 0, & \text{otherwise}\end{cases}$$

and

$$\boldsymbol{y} \triangleq [y_1 \cdots y_n]^{\dagger},$$
$$\boldsymbol{x}(t) \triangleq [x(t+t_1)\cdots x(t+t_n)]^{\dagger}$$

where $\dagger$ denotes matrix transposition. It is shown in [40] that $F_{\boldsymbol{x}(t)}(\boldsymbol{y})$ is nondecreasing in each of its arguments, only takes on values in $[0,1]$, and satisfies the boundary conditions $F_{\boldsymbol{x}(t)}(\boldsymbol{y}) \to 0$ for $\boldsymbol{y} \to -\infty$ and $F_{\boldsymbol{x}(t)}(\boldsymbol{y}) \to 1$ for $\boldsymbol{y} \to \infty$. Thus, (1) is a valid probability distribution function.

If the only nonzero term in the sum over the cycle frequency index α is that corresponding to $\alpha = 0$, then the time-series $x(t)$ is said to be *stationary of order* n. On the other hand, if some terms corresponding to $\alpha \neq 0$ are also nonzero, then $x(t)$ is said to *exhibit cyclostationarity*. If all such values of α are integer multiples of a single fundamental frequency, say $\alpha_0 = 1/T_0$, corresponding to a period T_0, then $x(t)$ is said to be *cyclostationary of order* n, otherwise $x(t)$ is said to be *polycyclostationary* (or *multiply-cyclostationary* or *almost cyclostationary*) *of order* n. In the latter case, the FOT distribution function can be expressed as

$$F_{\boldsymbol{x}(t)}(\boldsymbol{y}) = F_{\boldsymbol{x}}^{0}(\boldsymbol{y}) + \sum_{q}\left[F_{\boldsymbol{x}(t);T_q}(\boldsymbol{y}) - F_{\boldsymbol{x}}^{0}(\boldsymbol{y})\right] \tag{3}$$

where

$$F_{\boldsymbol{x}}^{0}(\boldsymbol{y}) \triangleq \left\langle \prod_{j=1}^{n} U[y_j - x(t+t_j)]\right\rangle$$

and where $F_{\boldsymbol{x}(t);T_q}(\boldsymbol{y})$ is the FOT distribution obtained by summing in (1)–(2) over only those values of α that are integer multiples of $1/T_q$. $F_{\boldsymbol{x}(t);T_q}(\boldsymbol{y})$ can also be obtained from the alternative but equivalent definition

$$F_{\boldsymbol{x}(t);T_q}(\boldsymbol{y}) \triangleq \lim_{M\to\infty}\frac{1}{2M+1} \times \sum_{m=-M}^{M}\left\{\prod_{j=1}^{n} U[y_j - x(t+t_j+mT_q)]\right\}. \tag{4}$$

The function $F_{\boldsymbol{x}}^{0}(\boldsymbol{y})$ must be subtracted from each $F_{\boldsymbol{x}(t);T_q}(\boldsymbol{y})$ in (3) because it appears in (4) as the $m = 0$ term for every value of T_q. That is, the $m = 0$ term in (4) is the same for each T_q, but should be included only once in the distribution function (1), as in (3).

The FOT probability density function (PDF) is given by the n-fold derivative of the distribution function

$$f_{\boldsymbol{x}(t)}(\boldsymbol{y}) \triangleq \frac{\partial^n}{\partial y_1 \cdots \partial y_n} F_{\boldsymbol{x}(t)}(\boldsymbol{y})$$

and has all the properties associated with the conventional probability density function.

The expectation operation associated with the FOT PDF is defined in the usual way for any PDF. Let $g[\boldsymbol{x}(t)]$ be a function of the vector of time-samples

$$\boldsymbol{x}(t) = [x(t+t_1) \cdots x(t+t_n)]^\dagger.$$

If we redefine the symbol $\hat{E}^{\{\alpha\}}\{\cdot\}$ (defined in (2)) to be the expected value with respect to the FOT PDF $f_{\boldsymbol{x}(t)}(\boldsymbol{y})$

$$\hat{E}^{\{\alpha\}}\{g[\boldsymbol{x}(t)]\} \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g[\boldsymbol{y}] f_{\boldsymbol{x}(t)}(\boldsymbol{y})\, d\boldsymbol{y} \tag{5}$$

then it can be shown that this expected value consists exactly of the finite-strength additive sine-wave components that are present in $g[\boldsymbol{x}(t)]$. Specifically

$$\hat{E}^{\{\alpha\}}\{g[\boldsymbol{x}(t)]\} = \sum_{\alpha} M_g^{\alpha} e^{i2\pi\alpha t} \tag{6}$$

where

$$M_g^{\alpha} \triangleq \langle g[\boldsymbol{x}(t)] e^{-i2\pi\alpha t} \rangle \tag{7}$$

which is consistent with the meaning given to $\hat{E}^{\{\alpha\}}\{\cdot\}$ in (2).

The FOT distributions and densities are not used explicitly (as in (5)) in this paper. Instead, the time-averages that characterize the FOT expectations are used directly, as in (6) and (7). Thus, the primary purpose of this discussion of FOT probability is to show that the multiple sine-wave extraction operation $\hat{E}^{\{\alpha\}}\{\cdot\}$ is completely analogous to the expectation operation $E\{\cdot\}$ and, therefore, that the subsequent use of terms such as moment, cumulant, and characteristic function (CF) is mathematically justified and appropriate. Examples of computation of the quantities (6) and (7) are given throughout the paper.

B. Statistical Independence in the FOT Framework

Since many of the properties of joint moments and cumulants of multiple time-series depend on their statistical dependence or lack thereof, it is important to understand the notion of statistical independence in the FOT framework.

The $(r+s)$th-order joint distribution function for the variables

$$\{\{y(t+u_j)\}_{j=1}^{r}\ \{z(t+v_k)\}_{k=1}^{s}\}$$

is denoted by $F_{\boldsymbol{y}(t)\boldsymbol{z}(t)}([\boldsymbol{u}\ \boldsymbol{v}])$. The time-series $y(t)$ and $z(t)$ are statistically independent if and only if the joint density factors

$$F_{\boldsymbol{y}(t)\boldsymbol{z}(t)}([\boldsymbol{u}\ \boldsymbol{v}]) = F_{\boldsymbol{y}(t)}(\boldsymbol{u}) F_{\boldsymbol{z}(t)}(\boldsymbol{v})$$

for all positive integers r and s, and all values of $\boldsymbol{u}$ and $\boldsymbol{v}$. It follows from this definition that any constant, periodic, or polyperiodic time-series is statistically independent of all time-series, including itself. This is analogous to the fact that a constant random variable is statistically independent of all random random variables, including itself.

It follows that the (temporal) expected value of a product of any number of time-samples of a (poly)periodic time-series and the time-samples of any other time-series is equal to the product of the samples of the (poly)periodic time-series multiplied by the expected value of the product of the samples of the other time-series. The previous statement is not true if the phrase expected value is replaced by the phrase *time-averaged value*, except in the special case where the (poly)periodic time-series is actually constant.

C. Cycloergodicity

We have drawn a distinction between the underlying concepts in the alternative frameworks of stochastic processes (used in HOS) and time-series (used in HOCS), namely the distinctions between stochastic expectation (ensemble average) and temporal expectation (time average or, more generally, sine-wave extraction), and between stationarity and cyclostationarity. In the light of the preceding paragraphs, these distinctions are clarified here.

A stochastic process is called kth-*order stationary* if all moments of order k or lower exist and are translation invariant. A time-series is called kth-*order stationary* if (i) its temporal moments up to order k exist, (ii) these moments are not identically zero, and (iii) there are no nonzero-frequency sine waves in any lag product with order k or lower (the temporal moment function contains no sine waves with frequencies $\alpha \neq 0$). If the third condition fails to hold, then the time-series is kth-*order cyclostationary* or kth-*order polycyclostationary.* A stochastic process *exhibits* kth-order cyclostationarity if the kth-order moment of the process exists and contains sine-wave components with nonzero frequencies α. The process *is* kth-order (almost or poly-) cyclostationary if its kth-order moments are (almost or poly-) periodic functions of the time-translation variable (i.e., if the moments are entirely composed of sums of sine waves). A stochastic process is called kth-*order cycloergodic* if the sine-wave components of every stochastic moment of order k or lower are equal to the sine-wave components of the corresponding temporal moment for almost every sample path; that is, stochastic moments are equal to temporal moments with probability one. In addition, we can define kth-*order ergodicity* to mean that the $\alpha = 0$ components of the stochastic moments are equal to the $\alpha = 0$ components of the corresponding temporal moments for almost every sample path of the process. It is evident that kth-order cycloergodicity implies kth-order ergodicity, but that the converse is not true.

Communication signals are often modeled as kth-order stationary stochastic processes even though the sample paths are not stationary time-series. This is done by introducing phase-randomizing variables. If in this case an ergodicity assumption is invoked, then measurements of statistics based on a single sample path of the process, such as moments and cumulants, can be in error in that they will not necessarily match the probabilistic functions of the process: the process is not necessarily kth-order cycloergodic even though it may be kth-order ergodic [35], [45]. This kind of modeling not only removes from consideration the potentially important sine waves, but also confuses the meaning and measurement of the $\alpha = 0$ components from sample paths of the process. This is illustrated with a numerical example in Part II, also in this issue.

It should be emphasized that all the results in this paper and its sequel can be obtained in the stochastic process framework provided that we restrict our attention to those cyclostationary processes that are cycloergodic. This duality is reflected in the notation we have chosen for the temporal expectation operation $\hat{E}^{\{\alpha\}}\{\cdot\}$, which is similar to the notation $E\{\cdot\}$ for the usual stochastic expectation used in the stochastic process framework.

D. Historical Survey of Cumulants and Cumulant Spectra

The history of the cumulant is traced in this section. A concise history, to begin with, is that the cumulant was born in mathematical statistics, developed in the probabilistic theory of stochastic processes, and after nearly 100 years found its way into electrical engineering through the field of higher order statistics as applied primarily to problems of time-series modeling and system identification.

In 1903, Danish astronomer Thorvald Nicolai Thiele published a book called *The Theory of Observations* [101] in which he tried to quantify the statistical nature of measurement errors. Thiele developed functions that he called *laws of presumptive errors* which are probability density functions. By expressing these functions in Maclaurin series form, he found that they could be characterized by moments or by cumulants, which he called *half-invariants*, because cumulants are invariant to additive constants in a random variable, but not to multiplicative constants. Thiele recognized that the half-invariants provided an easy way to test for the normal distribution: the higher order half-invariants are zero for Gaussian random variables. Although Thiele's introduction of cumulants was practically motivated, he did not arrive at them as the solution to a particular problem. Thiele gave no indication that he was aware of any other work on cumulants, and he did not use any term other than half-invariant to describe cumulants. Cramer [17], A. Fisher [29], and Graham [46] each claim that Thiele discovered cumulants.

Cumulants were introduced into the theory of sampling distributions, that is, the theory of the probability distribution of sample statistics, largely through the work of Fisher, Wishart, and Kendall [15], [30], [61], [107]. The basic problem here is to determine the distribution of statistics such as the sample moments or sample cumulants by, say, determining the moments or the cumulants of the sample statistic. Fisher knew that the mean of a statistic does not necessarily equal the corresponding population parameter. In the case of the sample variance $\hat{\sigma}^2$, the mean is

$$E\{\hat{\sigma}^2\} = \frac{N-1}{N}\sigma^2$$

where σ^2 is the population variance. Fisher proposed a new set of cumulant statistics, called *k-statistics*, for which the expected value is equal to the corresponding population cumulant. This set of statistics greatly facilitated work in sampling distribution theory. A detailed treatment of this topic is given in the excellent reference [61]. In 1937, Cornish [15] attributes the term *cumulant* to Laplace (without referencing a particular work of Laplace's) because Laplace called the logarithm of the characteristic function (LCF) the *cumulative function*, and cumulants are the coefficients of this function in Maclaurin series form. The LCF gets its alternative name of cumulative function from the property that the LCF for a sum of independent random variables is the sum of their LCF's: the LCF is cumulative. However, in 1928, Wishart [107] calls the LCF the *kappa generating function*, and calls the cumulants *cumulative moment functions*. The terms *cumulant* and *log-characteristic function* are used today.

Other early work, if taken further, could also have led to the cumulant. For example, in 1938, by generalizing the work of several others (Ursell, Darwin, and Fowler), Kahn and Uhlenbeck [102] presented a representation of the joint probability function in terms of a new function, and they represent this new function in terms of the probability function. We have observed that by using indicator functions to represent probability as an expectation, we can interpret the probability function as a joint moment of random indicator variables, and the new function can then be seen to be the cumulant of these indicator variables. Kahn and Uhlenbeck observed that their new function was zero when some subset of their events were independent of their remaining events, and they claimed that this was "the importance of the new function."

In the early 1950's, cumulants were used in an engineering context for the first time by Kuznetsov, Stratonovich, and Tikhonov in a study of the passage of stochastic processes through linear and nonlinear systems [66], [67]. Apparently without motivation, the authors decide to characterize the output stochastic process by using the logarithm of the joint characteristic function of samples of the output. They were therefore faced with a function which contained cumulants that were parameterized by the specific values of the sampling times. They called these cumulants *generalized correlation functions* because they were equal to the familiar correlation function for order two (the covariance). The main result in these two references is that the generalized correlation functions for the output process are related in a simple way to those for the input process. The authors also used the generalized correlation functions to characterize the degree to which the output process deviated from normality. A more general form of the relation between moments and cumulants than that given by Thiele in [101] is given in [67] apparently for the first time, but without proof, and is not central to the work therein. The authors of [66] and [67] used none of the terminology described in the previous paragraphs.

In the late 1950's and early 1960's the theory of cumulants, which are sometimes called *semi-invariants* (or *seminvariants*), was put on a firm theoretical foundation by Shiryaev and Leonov [69], [92]. In [69], the cumulants of the output of a polynomial nonlinearity are obtained in terms of the cumulants of the input process, thereby formalizing the earlier work in [66] and [67]. Also, a combinatorial proof of the relationships between moments and cumulants is given for arbitrary order. This appears to be the only published proof of these relations. The results of the work in [92] are essentially the probabilistic counterpart of the material presented in Sections II and III herein, restricted (for the most part) to stationary processes; [92] is a measure-theoretic approach to understanding higher order moments and spectra of stochastic processes. Shiryaev *defines* the polyspectrum as a cumulant with respect to the logarithm of a spectral characteristic function, which is the characteristic function of the spectral increments of the process, and also shows that the polyspectrum is equal to the Fourier transform of the time-domain cumulant function for generally nonstationary as well as stationary processes. Shiryaev does not specialize his results to the case of cyclostationary processes.

In the 1960's the properties of cumulants, both temporal and spectral, were investigated [10] and measurement techniques were developed [11], [12]. It is here that the term *polyspectrum* is introduced (which Brillinger attributes to Tukey) for the spectral cumulants or, equivalently, the Fourier transform of the temporal cumulants, and a case for the superiority of cumulants over moments for use in the theoretical development of HOS is made in [10]. The processes involved in [10] are assumed to be nth-order stationary, which means that all moments up to and including order n are translation invariant. In [10]–[12] polyspectra are *defined* to be the Fourier transforms of time-domain cumulants, but are also recognized to be spectral cumulants.

The 1970's saw minimal application of polyspectra and cumulants. The focus was on the third-order polyspectrum, called the *bispectrum*; the application was to the area of detection of phase-coupling in sinusoids [55], [62], [63], [75], [87], [88]. Three sinusoids with frequencies $\{f_i\}_{i=1}^3$ and random phases $\{\theta_i\}_{i=1}^3$ are said to be phase-coupled if $f_1 + f_2 = f_3$ and the sum $\theta_1 + \theta_2$ is statistically dependent on θ_3. This can be the case, for example, in the output of a linear-plus-quadratic system with the sum of sine waves with frequencies f_1 and f_2 at the input. Some progress in this area was made, and a corresponding interest in the statistical properties of estimates of the bispectrum was piqued.

In the 1980's a sector of the electrical engineering community became interested in HOS as a tool that could be used to perform system identification. Since the autocorrelation (and power spectral density) of a second-order stationary process does not contain phase information, it cannot be used to identify nonminimum-phase systems. Researchers were led to higher order statistics because higher order moments and cumulants do contain phase information. Many researchers claim that second-order statistics cannot be used to obtain phase information, but this is incorrect since the cyclic autocorrelation, which is a second-order statistic, is sensitive to phase for cyclostationary signals. Nevertheless, if the signal is stationary, or does not exhibit SOCS, then third or higher order statistics must be used to obtain phase information. System identification is still the dominant application area in HOS, as can be seen by noting that over half of the recent HOS tutorial paper by Mendel [72] is concerned with parametric system identification, that is, determining the coefficients of AR, MA and ARMA system models (see references in [72]). Other recent applications include synchronization [27], [6], random signal detection [47], [64], image reconstruction [81], tests for the Gaussian property and linearity of stochastic process [53], neural-network based estimation [104], radar signal processing [26], equalization [82] and direction-finding (source location) [13], [32], [54], [83], [110]. In most of these applications, the signals of interest are modeled as stationary stochastic processes (the exception is synchronization), and in many cases the highest order employed is three. There are good reasons for the latter restriction: for example, if the input signal-to-noise ratio (SNR) is below 0 dB, the SNR at the output of an nth-order homogeneous polynomial nonlinearity decreases as the order n increases. Also, computational complexity of algorithms that exploit HOS can grow rapidly as the order increases.

It is of some interest to note that the subject of cumulants has been largely neglected by the authors of the classic (or at least popular) texts in probability theory, mathematical statistics, stochastic processes, and time-series analysis. This is largely due to the long-standing emphasis on the correlation theory of processes and time-series wherein only the first and second moments are of interest. This theory is very powerful because it is sufficient for the explanation of the behavior of Gaussian processes, handles linear transformations of data easily and elegantly, and is computationally simpler (more tractable) than higher order theories. The treatment of cumulants and polyspectra in the most well-known texts is discussed next, followed by a brief description of three modern texts that do treat the topic of cumulants and polyspectra.

The texts considered in time-series analysis are [4], [9], [52], [57], [60], [65], [84]. The book [84] by Priestley contains the most material on cumulants. There is a short history that starts with Shiryaev's contribution [92], and then cumulants are defined through the LCF. Polyspectra are defined to be the Fourier transforms of these cumulants. Brillinger [9], on the other hand, defines cumulants in terms of their relation to moments and lists several of their elementary properties. The other books in the area of time-series analysis offer little (such as the moment/cumulant relations for $n = 1, 2, 3, 4$) or nothing on cumulants.

The texts considered in the area of stochastic processes are [7], [8], [24], [25], [51], [58], [59], [73], [77], [79], [86], [100], [103], [108], [109]. These texts pay very little attention to cumulants and polyspectra. The four texts [7], [86], [100], and [73] each define the cumulant through the series expansion of the LCF, but do little with them. The other references in the area of stochastic processes do not even mention cumulants.

References [16], [28], [29], [70], [71], [78], [80] are the texts considered in probability theory. Parzen [80] and Papoulis [78] both define the cumulant through the series expansion of the LCF, although Parzen does it in an exercise. Neither theory nor application of cumulants is developed in either of these books. A. Fisher essentially reproduces Thiele's work in his 1923 book [29]. The other references in the area of probability theory do not mention cumulants.

Finally, the texts considered in mathematical statistics are [17], [31], [61], [68], [89], [106]. Both Fisz [31] and Cramer

[17] define the cumulant through the series expansion of the LCF in the usual way, but go no further with the theory. The book by Kendall and Stuart [61] devotes a great deal of attention to cumulants, mostly in the context of sampling distribution theory (as mentioned previously). The authors were not concerned with stochastic processes or time-series theory, and do not define polyspectra. The other texts in mathematical statistics do not mention cumulants.

There are three modern texts that do treat the topics of cumulants and polyspectra for time-series and stochastic processes. The first is by Rosenblatt [90]. The material on cumulants and polyspectra in this text is essentially the same as in the two papers [11], [12], in which the cumulants and polyspectra of stationary stochastic processes are investigated, with emphasis on estimating these parameters from finite-length data records. The book by Priestley [85] contains a chapter devoted to estimation of the polyspectrum of a stationary stochastic process from a finite-length data record. The methods considered therein are the same as in [90]; however, Priestley's description of the frequency-domain method of estimating the cyclic polyspectrum is in error (see Section VI of Part II, the companion paper, also in this issue). The material in both of these texts is considered further in Part II, where the measurement of the parameters of HOCS is studied. The third modern text that treats cumulants is a collection of papers edited by Haykin [75]. The chapter by Nikias treats the topic of estimation of the polyspectrum; the emphasis of the chapter is on the use of such estimates in solving the problem of parametric (ARMA) system identification.

The nine research papers [37], [39], [44], and [93]–[97], all of which are from the same research group, address the topic of HOCS directly. In [37], the higher order temporal moments of cyclostationary time-series are introduced and shown to be related to the higher order moments of spectral components of the time-series by Fourier transformation, and it is established for the first time that the nth-order lag product of a time-series contains a finite-strength additive sine-wave component with frequency α if and only if the joint moment of n spectral components of the time-series is nonzero for some sets of n frequencies that sum to α. In [39], the concept of a *pure nth-order sine wave* is introduced and is shown to be characterized by a temporal cumulant function, whose Fourier transform is shown to be a spectral cumulant function, and the relationships between these functions and those that arise in the theory of HOS for stationary processes are discussed. higher order temporal and spectral moments of cyclostationary time-series are used in [44] to identify nonlinear systems. Techniques for measuring the higher order statistical functions (cyclic moments, cyclic cumulants, and cyclic polyspectra) for cyclostationary time-series are presented in [93], and some results on their performance are presented in [94]. A brief overview of the theory of higher order cyclostationary time-series is given in [95]. Finally, the application of the theory of HOCS to the problems of weak-signal detection and time-delay estimation is considered in [97].

There is one other research group that is studying higher order cyclostationarity. These researchers use the stochastic process framework to study the moments and cumulants of cyclostationary processes and their application to system identification, random-signal detection, modulation classification, and source location [18]–[23], [48]–[50]. The main differences between the approach taken by these researchers and the work herein are that we stress the sine-wave generation interpretation of cyclostationarity, whereas they focus on the probabilistic parameters (moments and cumulants) themselves; we have focused on developing the theory of higher order cyclostationarity, whereas they have focused on the application of system identification; we have focused on fourth-order moments and cumulants (because the third-order quantities are typically zero for manmade signals), whereas they have focused on third-order moments (which equal third-order cumulants for zero-mean processes); we have focused on continuous-time signals, whereas they have focused on discrete-time signals. These differences result in part from approaching the subject from the two distinct viewpoints of time-average-based second-order cyclostationarity on the one hand, and stationary-stochastic-process-based higher order statistics on the other.

Three research papers by other investigators also treat HOCS. One of these is [111], in which the stochastic process framework is used, and in which no connection is made to the sine-wave generation idea that is central to this paper, nor to cumulants, which are also central to this paper. The second is [5] in which a cyclic spectral analysis of the powers of a PAM signal is carried out, that is, the cyclic spectrum of the output of a nonlinear system with a PAM signal at the input is calculated. (The results herein are more general than those in [5].) The third is [2] in which the symmetry properties of nth-order polyspectra for $n \leq 6$ for cyclostationary stochastic processes are investigated.

The strong recent interest in HOS in the electrical engineering community is reflected in the recent workshops on higher order statistics, in the two IEEE PROCEEDINGS tutorial papers [72], [74], and in the special sections on HOS in the July 1990 IEEE TRANSACTIONS ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, and the January 1990 IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

II. THE TEMPORAL PARAMETERS OF HOCS

As one possible motivation for the definitions introduced in this section, let us consider all nonlinear signal processing operations that can be represented by a Volterra series. This includes (but not exclusively) all continuous, time-invariant, finite-memory, causal systems [76]. The output $y(t)$ of such an operation is expressed as[1]

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h_1(\tau_1)x(t+\tau_1)d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2) \\ &\quad \times x(t+\tau_1)x(t+\tau_2)d\tau_1 d\tau_2 + \cdots \\ &= \sum_n \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_n(\boldsymbol{\tau})L_x(t,\boldsymbol{\tau})_n d\boldsymbol{\tau} \end{aligned}$$

[1] Strictly speaking, the name Volterra is reserved for causal systems, for which $h_n(\boldsymbol{\tau}) = 0$ for any $\tau_j < 0$, but this restriction is lifted here. Also, infinite-memory systems, for which the $h_n(\cdot)$ have infinite support, are included here.

where $\boldsymbol{\tau} \triangleq [\tau_1 \cdots \tau_n]^\dagger$ and $L_x(t,\boldsymbol{\tau})_n$ is the nth-order lag product of the input $x(t)$

$$L_x(t,\boldsymbol{\tau})_n \triangleq \prod_{j=1}^{n} x(t+\tau_j). \quad (8)$$

We are interested in the finite-strength additive sine-wave components present in the output but absent in the input, that is, those sine waves that are generated by the action of the nonlinear operation on the input $x(t)$. For example, the strength (magnitude and phase) of the sine wave with frequency α in $y(t)$ is given by (assuming the order of $\hat{E}^{\{\alpha\}}$, $\sum$, and $\int$ can be interchanged)

$$\langle y(t)e^{-i2\pi\alpha t}\rangle = \sum_n \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_n(\boldsymbol{\tau}) \times \langle L_x(t,\boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle d\boldsymbol{\tau}.$$

Thus, we need only study the statistical quantities $\langle L_x(t,\boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle$, for arbitrary positive integers n, which are the strengths of the sine-wave components contained in the nth-order lag products of $x(t)$.

The nth-order lag product is an elementary nth-order homogeneous polynominal transformation of $x(t)$. This transformed time-series can be decomposed into a polyperiodic (or periodic) part and an aperiodic residual part

$$L_x(t,\boldsymbol{\tau})_n = p(t,\boldsymbol{\tau})_n + m(t,\boldsymbol{\tau})_n$$

where

$$\langle m(t,\boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle = 0 \quad (9)$$

for all real numbers α. The polyperiodic portion of $L_x(t,\boldsymbol{\tau})_n$ has associated with it the Fourier series

$$p(t,\boldsymbol{\tau})_n = \sum_\alpha R_x^\alpha(\boldsymbol{\tau})_n e^{i2\pi\alpha t} \quad (10)$$

where

$$R_x^\alpha(\boldsymbol{\tau})_n \triangleq \langle p(t,\boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle. \quad (11)$$

It is assumed herein that the partial sums in the Fourier series (10) converge uniformly in t for each $\boldsymbol{\tau}$ to $p(t,\boldsymbol{\tau})_n$. Then $p(\cdot,\boldsymbol{\tau})_n$ is a polyperiodic function, the limit (11) exists for each $\boldsymbol{\tau}$, and the set of values of the real variable α for which $R_x^\alpha(\boldsymbol{\tau})_n \not\equiv 0$ for each $\boldsymbol{\tau}$ is denumerable [14]. That is, there is at most a denumerable set of incommensurate periods in the polyperiodic (almost periodic) function $p(t,\boldsymbol{\tau})_n$ for each $\boldsymbol{\tau}$. It is further assumed that the union over all $\boldsymbol{\tau}$ of the sets of values of α for which $R_x^\alpha(\boldsymbol{\tau})_n \not\equiv 0$ is denumerable. For example, it is shown in [56] that this union is denumerable for $n=2$ if $p(t,\boldsymbol{\tau})_2$ is uniformly continuous in t and $\boldsymbol{\tau}$.

The lag-product time-series can therefore be expressed as

$$L_x(t,\boldsymbol{\tau})_n = \sum_\alpha R_x^\alpha(\boldsymbol{\tau})_n e^{i2\pi\alpha t} + m(t,\boldsymbol{\tau})_n \quad (12)$$

where the sum is over the denumerable set of real α for which $R_x^\alpha(\boldsymbol{\tau})_n \not\equiv 0$. From (9) and (12), we have

$$R_x^\alpha(\boldsymbol{\tau})_n = \langle L_x(t,\boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle. \quad (13)$$

Each value of α in the representation (12) is called an *impure nth-order cycle frequency*, and $R_x^\alpha(\boldsymbol{\tau})_n$ in (13) is called the *nth-order cyclic temporal moment function* (CTMF). From (13), it is evident that the CTMF arises quite naturally from a consideration of the finite-strength additive sine-wave components in the lag product (8). The sum of all such sine waves in $L_x(t,\boldsymbol{\tau})_n$ is given by the temporal expected value of the lag product (cf. Section 1.1)

$$\hat{E}^{\{\alpha\}}\{L_x(t,\boldsymbol{\tau})_n\} = \sum_\alpha R_x^\alpha(\boldsymbol{\tau})_n e^{i2\pi\alpha t}$$

which is called the *temporal moment function* (TMF), and is denoted by $R_x^\alpha(t,\boldsymbol{\tau})_n$

$$R_x(t,\boldsymbol{\tau})_n \triangleq \hat{E}^{\{\alpha\}}\{L_x(t,\boldsymbol{\tau})_n\} = \sum_\alpha R_x^\alpha(\boldsymbol{\tau})_n e^{i2\pi\alpha t}. \quad (14)$$

An individual component of the TMF, such as $R_x^\alpha(\boldsymbol{\tau})_n\, e^{i2\pi\alpha t}$, is called an *nth-order moment sine wave* (to distinguish it from a *cumulant sine wave* which is defined in Section II-C-4) or an *impure nth-order sine wave*. Time-series for which there exists at least one nth-order moment sine wave (with $\alpha \neq 0$) are called *nth-order cyclostationary* (CS_n) time-series. A potentially confusing property of CS_n time-series is that such time-series are in general CS_{2n}. (In contrast, for a stationary-of-order-n (S_n) time-series, we have S_n implies S_{n-1}.) A simple example illustrates this fact. Consider the time-series given by

$$x(t) = \cos(\omega t) + m(t)$$

where the zero-mean time-series $m(t)$ contains no sine-wave components. This time-series is CS_1. Any second-order lag product contains sine waves as well:

$$x(t+\tau_1)x(t+\tau_2) = \frac{1}{2}\cos(2\omega t + \omega[\tau_1+\tau_2]) + \frac{1}{2}\cos(\omega[\tau_1-\tau_2]) + \text{residue}.$$

Therefore, $x(t)$ is CS_2.

More interesting cases involve random time-series that do not themselves contain additive sine-wave components, because it is still true that, for example, CS_2 implies CS_4. For the purpose of illustrating the temporal moment functions, we next derive the second-order TMF and CTMF's for a time-series model of a common communication signal.

A. Example

Consider the time-series consisting of an amplitude-modulated carrier and an unmodulated carrier

$$x(t) = a(t)\cos(2\pi\nu_1 t + \theta_1) + \cos(2\pi\nu_2 t + \theta_2) \quad (15)$$

where $a(t)$ is a stationary time-series with $\hat{E}^{\{\alpha\}}\{a(t)\} = 0$ and $\nu_1 \neq \nu_2$. The second-order lag product is

$$\begin{aligned}
L_x(t, \boldsymbol{\tau})_2 &= [a(t+\tau_1)\cos(2\pi\nu_1[t+\tau_1]+\theta_1) \\
&\quad + \cos(2\pi\nu_2[t+\tau_1]+\theta_2)] \\
&\quad \times [a(t+\tau_2)\cos(2\pi\nu_1[t+\tau_2]+\theta_1) \\
&\quad + \cos(2\pi\nu_2[t+\tau_2]+\theta_2)] \\
&= \frac{a(t+\tau_1)a(t+\tau_2)}{2} \\
&\quad \times \left[\cos\left(4\pi\nu_1\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_1\right)\right. \\
&\quad \left. + \cos(2\pi\nu_1[\tau_1-\tau_2])\right] \\
&\quad + \frac{1}{2}\left[\cos\left(4\pi\nu_2\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_2\right)\right. \\
&\quad \left. + \cos(2\pi\nu_2[\tau_1-\tau_2])\right] \\
&\quad + \frac{a(t+\tau_1)}{2}[\cos(2\pi[(\nu_1+\nu_2)t \\
&\quad + \nu_1\tau_1+\nu_2\tau_2]+\theta_1+\theta_2) \\
&\quad + \cos(2\pi[(\nu_1-\nu_2)t+\nu_1\tau_1-\nu_2\tau_2]+\theta_1-\theta_2)] \\
&\quad + \frac{a(t+\tau_2)}{2}[\cos(2\pi[(\nu_1+\nu_2)t \\
&\quad + \nu_1\tau_2+\nu_2\tau_1]+\theta_1+\theta_2) \\
&\quad + \cos(2\pi[(\nu_1-\nu_2)t+\nu_1\tau_2-\nu_2\tau_1]+\theta_1-\theta_2)].
\end{aligned}$$

Since $a(t)$ is stationary and has zero mean, and $a(t)$ and $\cos(2\pi\nu t)$ are statistically independent time-series (cf. Section I-B), the TMF is given by

$$\begin{aligned}
&\hat{E}^{\{\alpha\}}\{L_x(t,\boldsymbol{\tau})_2\} \\
&= \frac{1}{2}R_a(\tau_1-\tau_2)\left[\cos(2\pi\nu_1[\tau_1-\tau_2])\right. \\
&\quad \left. + \hat{E}^{\{\alpha\}}\left\{\cos(4\pi\nu_1\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_1)\right\}\right] \\
&\quad + \frac{1}{2}\hat{E}^{\{\alpha\}}\left\{\cos\left(4\pi\nu_2\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_2\right)\right\} \\
&\quad + \frac{1}{2}\cos(2\pi\nu_2[\tau_1-\tau_2]).
\end{aligned} \tag{16}$$

Thus

$$R_x(t,\boldsymbol{\tau})_2 = \hat{E}^{\{\alpha\}}\{L_x(t,\boldsymbol{\tau})_2\} = \sum_{\alpha} R_x^{\alpha}(\boldsymbol{\tau})_2\, e^{i2\pi\alpha t} \tag{17}$$

where

$$R_x^{\alpha}(\boldsymbol{\tau})_2 = \tag{18}$$

$$\begin{cases}
\frac{1}{2}R_a(\tau_1-\tau_2)\cos(2\pi\nu_1[\tau_1-\tau_2]) \\
\quad +\frac{1}{2}\cos(2\pi\nu_2[\tau_1-\tau_2]), & \alpha = 0, \\
\frac{1}{4}R_a(\tau_1-\tau_2)\exp(\pm i(2\pi\nu_1[\tau_1+\tau_2]+2\theta_1)), & \alpha = \pm 2\nu_1, \\
\frac{1}{4}\exp(\pm i(2\pi\nu_2[\tau_1+\tau_2]+2\theta_2)), & \alpha = \pm 2\nu_2, \\
0, & \text{otherwise.}
\end{cases}$$

This example is continued in subsequent sections to illustrate the behavior of certain functions as they are introduced. Although the example is for the case $n = 2$ (primarily for brevity and clarity), it is important because it illustrates some unique features of the theory's parameters that are exhibited even for $n = 2$. A whole section is devoted to the specific modulation class of complex-valued pulse-amplitude-modulated signals in Part II, also in this issue. For now, a single example for $n > 2$ will suffice.

Let us consider an AM signal without the tonal interferer that is present in (15)

$$x(t) = a(t)\cos(2\pi f_c t + \theta) \tag{19}$$

and let us consider a fourth-order lag product

$$L_x(t,\boldsymbol{\tau})_4 = \prod_{j=1}^{4} x(t+\tau_j). \tag{20}$$

The sine wave with frequency $4f_c$ in the lag product (20) can be shown to be given by

$$\begin{aligned}
&R_x^{4f_c}(\boldsymbol{\tau})_4 e^{i2\pi 4 f_c t} = \\
&\frac{1}{16}\left\langle \prod_{j=1}^{4} a(t+\tau_j)\right\rangle \\
&\quad \times e^{i(2\pi f_c[\tau_1+\tau_2+\tau_3+\tau_4]+4\theta)} e^{i2\pi 4 f_c t}.
\end{aligned} \tag{21}$$

Since $R_a^0(\boldsymbol{\tau})_2 = \langle a(t+\tau_1)a(t+\tau_2)\rangle$, then the lag product for $a(t)$ can be represented by

$$a(t+\tau_1)a(t+\tau_2) = R_a^0(\tau_1,\tau_2)_2 + b(t,\tau_1,\tau_2) \tag{22}$$

for which $\langle b(t,\tau_1,\tau_2)\rangle \equiv 0$. By making use of (22) in the fourth-order lag product for $a(t)$, we can begin to see the pure and impure components of the fourth-order sine wave (21):

$$\begin{aligned}
R_x^{4f_c}(\boldsymbol{\tau})_4 &= \frac{1}{16}\langle [R_a^0(\tau_1,\tau_2)_2 + b(t,\tau_1,\tau_2)][R_a^0(\tau_3,\tau_4)_2 \\
&\quad + b(t,\tau_3,\tau_4)]\rangle e^{i(2\pi f_c[\tau_1+\tau_2+\tau_3+\tau_4]+4\theta)} \\
&= \frac{1}{16}[R_a^0(\tau_1,\tau_2)_2 R_a^0(\tau_3,\tau_4)_2 + \langle b(t,\tau_1,\tau_2) \\
&\quad \times b(t,\tau_3,\tau_4)\rangle] e^{i(2\pi f_c[\tau_1+\tau_2+\tau_3+\tau_4]+4\theta)}.
\end{aligned}$$

Therefore, there are components of the fourth-order moment sine wave (21) that consist of products of second-order moment sine waves, and there are—potentially—other components that do not consist of these products. We say *potentially* because there are other products of lower order sine-waves, namely those obtained by using a different factorization of the fourth-order lag product, and we are unsure at this point if these other impure sine waves are the only components of $\langle b(t,\tau_1,\tau_2)b(t,\tau_3,\tau_4)\rangle$.

In the case of the second-order lag products of (15), we can purify the second-order sine waves by operating directly on the data: we simply remove the sine-wave component $\cos(2\pi\nu_2 t + \theta_2)$ from $x(t)$. In the case of the fourth-order lag products of (19) (or (15)), we cannot purify the fourth-order sine waves by operating on the data because there are no sine waves in the data. Furthermore, we cannot simply subtract the second-order sine waves from the lag products

$x(t+\tau_1)x(t+\tau_2)$ and $x(t+\tau_3)x(t+\tau_4)$ because there are similar sine waves in other factorizations of the fourth-order lag product. In the next section we show how to properly purify the nth-order moment sine waves, thereby obtaining the *pure nth-order sine waves.*

B. Pure nth-Order Sine Waves

For low orders n, it is easy to mathematically characterize a pure nth-order sine wave in a way that matches our intuition. For $n = 1$, the moment sine waves are, by definition, pure first-order sine waves. For $n = 2$, all products of first-order moment sine waves can be subtracted from the second-order moment sine waves to obtain the pure second-order sine waves, which are denoted by $\sigma_x(t, \tau_1, \tau_2)_2$

$$\begin{aligned}\sigma_x(t,\tau_1,\tau_2)_2 &\triangleq \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\} \\ &\quad - \hat{E}^{\{\alpha\}}\{x(t+\tau_1)\}\hat{E}^{\{\alpha\}}\{x(t+\tau_2)\} \\ &= R_x(t,\boldsymbol{\tau})_2 - R_x(t,\tau_1)_1 R_x(t,\tau_2)_1.\end{aligned}$$

There are several interesting points to be made concerning pure second-order sine waves.

1) Since $R_x(t,\tau_1), R_x(t,\tau_2)$, and $R_x(t,\boldsymbol{\tau})_2$ are first- and second-order moments, then $\sigma_x(t,\tau_1,\tau_2)_2$ is a temporal covariance function.
2) If $R_x(t,\tau)_1 \equiv 0$, then there are no lower-than-second-order sine waves, and the second-order moment sine waves are equal to the pure second-order sine waves.
3) If the variables $x(t+\tau_1)$ and $x(t+\tau_2)$ are statistically independent (in the temporal sense [34], [40]), then $\hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\} = \hat{E}^{\{\alpha\}}\{x(t+\tau_1)\} \times \hat{E}^{\{\alpha\}}\{x(t+\tau_2)\}$ and therefore $\sigma_x(t,\tau_1,\tau_2)_2 = 0$, that is, there is no pure second-order sine wave for this particular pair of lags τ_1 and τ_2.

A recursion can be used to compute the pure third-order sine waves. Each distinct product of *pure* lower order sine waves must be subtracted from the third-order moment sine waves. Thus, products of pure second-order and pure first-order sine waves are subtracted from the third-order moment sine waves

$$\begin{aligned}\sigma_x(t,\boldsymbol{\tau})_3 &= \hat{E}^{\{\alpha\}}\left\{\prod_{j=1}^{3} x(t+\tau_j)\right\} \\ &\quad - \sigma_x(t,\tau_1,\tau_2)_2\sigma_x(t,\tau_3)_1 \\ &\quad - \sigma_x(t,\tau_1,\tau_3)_2\sigma_x(t,\tau_2)_1 \\ &\quad - \sigma_x(t,\tau_2,\tau_3)_2\sigma_x(t,\tau_1)_1 \\ &\quad - \sigma_x(t,\tau_1)_1\sigma_x(t,\tau_2)_1\sigma_x(t,\tau_3)_1. \end{aligned} \tag{23}$$

Note that all possible products of pure lower order sine waves appear in (23). The terms in the sum of products that are subtracted can be enumerated easily by considering the distinct *partitions* of the index set $\{1,2,3\}$. A partition of a set G is a collection of p subsets of G, $\{\nu_i\}_{i=1}^p$, with the following properties[2]:

$$G = \bigcup_{j=1}^{p} \nu_j, \qquad \nu_j \bigcap \nu_k = \emptyset \qquad j \neq k.$$

The set P_3 of distinct partitions of $\{1,2,3\}$ is

$$\begin{aligned}&p=1\colon \{1,2,3\}\\ &p=2\colon \{1,2\},\{3\} \quad \{1,3\},\{2\} \quad \{2,3\},\{1\}\\ &p=3\colon \{1\},\{2\},\{3\}.\end{aligned}$$

Thus, we can express the pure third-order sine waves $\sigma_x(t,\boldsymbol{\tau})_3$ as a sum over the elements of P_3

$$\sigma_x(t,\boldsymbol{\tau})_3 = R_x(t,\boldsymbol{\tau})_3 - \sum_{\substack{P_3\\ p\neq 1}}\left[\prod_{j=1}^{p}\sigma_x(t,\boldsymbol{\tau}_{\nu_j})_{n_j}\right]$$

where $\boldsymbol{\tau}_{\nu_j}$ is a vector of elements of $\{\tau_j\}_{j=1}^3$ that have indices in ν_j, and n_j is the number of elements in ν_j.

Note that, as in the case of $n = 2$, if the first-order moment sine waves are zero, $\hat{E}^{\{\alpha\}}\{x(t)\} = 0$, then the third-order moment sine waves are equal to the pure third-order sine waves. In this case, there are no products of lower order sine waves that can be subtracted from the moment.

Because there is a one-to-one correspondence between the set of distinct factorizations of a product of n factors and the set of distinct partitions of the set $\{1,2,\ldots,n\}$ (as illustrated for $n = 3$ above), the formula for the pure nth-order sine waves can be expressed in terms of these partitions

$$\sigma_x(t,\boldsymbol{\tau})_n = R_x(t,\boldsymbol{\tau})_n - \sum_{\substack{P_n\\ p\neq 1}}\left[\prod_{j=1}^{p}\sigma_x(t,\boldsymbol{\tau}_{\nu_j})_{n_j}\right] \tag{24}$$

where P_n is the set of distinct partitions of $\{1,2,\ldots,n\}$. The pure-sine-waves formula[3] (24) gives all the pure nth-order sine waves associated with the lag vector $\boldsymbol{\tau}$. A single pure nth-order sine wave with frequency β can be selected by computing the Fourier coefficient

$$\sigma_x^{\beta}(\boldsymbol{\tau})_n e^{i2\pi\beta t} = \langle \sigma_x(u,\boldsymbol{\tau})_n e^{i2\pi\beta(t-u)}\rangle \tag{25}$$

and can be expressed in terms of pure lower order sine waves by substituting the Fourier series for each σ_x

$$\sigma_x(t,\boldsymbol{w})_k = \sum_{\beta_k}\sigma_x^{\beta_k}(\boldsymbol{w})_k e^{i2\pi\beta_k t}, \tag{26}$$

where the sum is over all cycle frequencies β_k of order k, into (24). Thus

$$\sigma_x^{\beta}(\boldsymbol{\tau})_n = R_x^{\beta}(\boldsymbol{\tau})_n - \sum_{\substack{P_n\\ p\neq 1}}\left[\sum_{\boldsymbol{\beta}^{\dagger}\mathbf{1}=\beta}\prod_{j=1}^{p}\sigma_x^{\beta_j}(\boldsymbol{\tau}_{\nu_j})_{n_j}\right] \tag{27}$$

[2] The total number of distinct partitions of a set is called *Bell's number* [3], which must be computed by a recursion involving Stirling numbers of the second kind [46], [96]. This same recursion can be modified to yield the partitions themselves [96].

[3] This approach to obtaining pure nth-order sine waves can break down in some special anomalous cases involving degenerate time-series, which are described in Section B.

where $\boldsymbol{\beta}$ is the p-dimensional vector of cycle frequencies $[\beta_1 \cdots \beta_p]^\dagger$ and 1 is the p-dimensional vector of ones. Hence, the pure-sine-wave strength $\sigma_x^{\beta}(\boldsymbol{\tau})_n$ is given by the CTMF $R_x^{\beta}(\boldsymbol{\tau})_n$ with all products of pure lower order sine-wave strengths, for sine waves whose frequencies sum to β, subtracted out.

1) Example—Continued: Let us reconsider the example of Section II–A. The second-order TMF for the time-series consisting of the sum of an amplitude-modulated sine wave and an unmodulated sine wave is given by (17) and (18). The first-order TMF for this time-series is

$$\begin{aligned}R_x(t,\tau)_1 &= \hat{E}^{\{\alpha\}}\{x(t+\tau)\}\\ &= \frac{1}{2}\exp(i2\pi\nu_2(t+\tau)+i\theta_2)\\ &\quad+\frac{1}{2}\exp(-i2\pi\nu_2(t+\tau)-i\theta_2).\end{aligned}$$

To compute the pure second-order sine waves for this time-series, we subtract from the second-order TMF (17), (18) the products of all pure first-order sine waves, which are the first-order moment sine waves:

$$\begin{aligned}R_x(t,\tau_1)_1R_x(t,\tau_2)_1 = \frac{1}{2}\bigg[&\cos\left(4\pi\nu_2\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_2\right)\\ &+\cos(2\pi\nu_2[\tau_1-\tau_2])\bigg].\end{aligned}$$

Thus, the pure second-order sine waves are given by

$$\begin{aligned}\sigma_x(t,\tau_1,\tau_2)_2 &= \frac{1}{2}R_a(\tau_1-\tau_2)\cos(2\pi\nu_1[\tau_1-\tau_2])\\ &\quad+\frac{1}{2}R_a(\tau_1-\tau_2)\\ &\quad\times\cos\left(4\pi\nu_1\left[t+\frac{\tau_1+\tau_2}{2}\right]+2\theta_1\right)\end{aligned}$$

or, equivalently,

$$\sigma_x(t,\boldsymbol{\tau})_2 = \sum_{\beta}\sigma_x^{\beta}(\boldsymbol{\tau})_2 e^{i2\pi\beta t}$$

where

$$\sigma_x^{\beta}(\boldsymbol{\tau})_2 = \begin{cases}\frac{1}{2}R_a(\tau_1-\tau_2)\cos(2\pi\nu_1[\tau_1-\tau_2]), & \beta=0,\\ \frac{1}{4}R_a(\tau_1-\tau_2)\exp(\pm i(2\pi\nu_1 \\ \quad\times[\tau_1+\tau_2]+2\theta_1)), & \beta=\mp 2\nu_1,\\ 0, & \text{otherwise.}\end{cases} \tag{28}$$

It is evident that the set of pure second-order cycle frequencies $\{\beta\}$ is smaller than the set of impure second-order cycle frequencies $\{\alpha\}$ (there is no pure second-order sine wave with frequency $\pm 2\nu_2$). Note also that the function $R_x^0(\boldsymbol{\tau})_2$ is not integrable with respect to the variable $\tau = \tau_1 - \tau_2$ because of the presence of the sinusoidal term $\frac{1}{2}\cos(2\pi\nu_2[\tau_1-\tau_2])$, whereas the function $\sigma_x^0(\boldsymbol{\tau})_2$ is integrable. This problem with moments is compounded in the case of higher order moments, where there can be multiple sinusoidal factors in the CTMF, which result in products of impulse functions in its Fourier transform. This is studied in Section III.

In the next section, we show that the pure-nth-order-sine-waves function $\sigma_x(t,\boldsymbol{\tau})_n$ is, in fact, an nth-order cumulant function. Before doing this, a general introduction to cumulants is provided.

C. Cumulants

In the study of the statistical behavior of stochastic processes (and time-series), much attention is focused on the first- and second-order moment functions. There are good reasons for this, including the tractability of analysis, the primacy of the Gaussian distribution (which is completely characterized by these two moment functions), and the fact that the mean and variance of all linear transformations of a process are completely characterized by these two moment functions. For zero-mean random variables, stochastic processes, and time-series, the first and second moments are the first two elements of several distinct sets of statistical parameters that each completely describe the behavior of the quantity in question. These include moments, centralized moments, and cumulants. In fact, for zero-mean random variables, the moments, centralized moments, and cumulants are identical through order 3.

Thus, in this common case we are not faced with the prospect of choosing one set or the other for our work. However, when we have motivation to investigate statistics with orders larger than three, we must make a choice. To do this, we must understand the properties of each set of statistics.

The most commonly used property of cumulants is that for Gaussian random variables all cumulants of order 3 and higher are zero. However, cumulants possess other valuable properties that moments do not. In the remainder of this section, we present a brief tutorial treatment of cumulants of stochastic processes that includes their relationship to moments, and we define the analogous temporal cumulants of a time-series using the FOT probability framework. The connection between nth-order cumulants and pure nth-order sine waves will then become apparent.

1) Cumulants of a Single Random Variable[4]*:* Let the random variable X have probability density function (PDF) $f_X(u)$. The characteristic function (CF) is the conjugate Fourier transform of the PDF

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} f_X(u)e^{i\omega u}du = E\{e^{i\omega X}\}.$$

It is well known that the moments of X can be obtained from the CF by differentiation

$$E\{X^n\} = (-i)^n\frac{\partial^n}{\partial\omega^n}\Phi_X(\omega)\bigg|_{\omega=0} = m_n.$$

The nth-order moment of X is, therefore, the coefficient of the term corresponding to $(i\omega)^n/n!$ in the Maclaurin series expansion of the CF:

$$\Phi_X(\omega) = \sum_{n=0}^{\infty} m_n\frac{(i\omega)^n}{n!}. \tag{29}$$

The CF is a useful tool in the study of random variables, but it does have a drawback. Let Y be the sum of two independent

[4] The references cited in this and succeeding sections dealing with cumulants of random variables, stochastic processes, and nonstochastic time-series are believed to be the original sources where these quantities were first introduced.

random variables X_1 and X_2. Then the PDF for Y is the convolution of the PDFs for X_1 and X_2

$$f_Y(u) = \int_{-\infty}^{\infty} f_{X_1}(u-\lambda) f_{X_2}(\lambda) d\lambda$$

which implies that the CF for Y is the product of CF's for X_1 and X_2

$$\Phi_Y(\omega) = \Phi_{X_1}(\omega)\Phi_{X_2}(\omega). \tag{30}$$

By using (29) in (30), it can be shown that the nth-order moment of Y is related to the moments of X_1 and X_2 of *all* orders n and lower.

If we transform the multiplication in (30) to addition by applying the natural logarithm, we obtain the relation

$$\ln \Phi_Y(\omega) = \ln \Phi_{X_1}(\omega) + \ln \Phi_{X_2}(\omega). \tag{31}$$

These new functions are called *cumulative functions* [15], a term that is due to Laplace, and the coefficient of the term corresponding to $(i\omega)^n/n!$ in the Maclaurin series expansion of the left side of (31) (provided that it exists), is called the nth-order *cumulant* of the random variable Y [15]. The nth-order cumulant for Y is, therefore, the sum of the nth-order cumulants for X_1 and X_2. Note that if X is a Gaussian random variable, its cumulative function is a second-order polynomial in ω; hence

$$\frac{\partial^n}{\partial \omega^n} \ln \Phi_X(\omega) \equiv 0, \quad n \geq 3$$

and, therefore, all higher order cumulants of this random variable are zero.

2) Cumulants of a Set of Random Variables: The multivariate PDF for the set of r random variables $\{X_j\}_{j=1}^r$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^r}{\partial x_1 \cdots \partial x_r} F_{\boldsymbol{X}}(\boldsymbol{x}) \tag{32}$$

where $F_{\boldsymbol{X}}(\boldsymbol{x})$ is the multivariate distribution function

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \text{Prob}\left\{ \bigcap_{j=1}^{r} [X_j < x_j] \right\}.$$

The CF is the multidimensional conjugate Fourier transform of (32):

$$\Phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(\boldsymbol{x}) e^{i\boldsymbol{\omega}^\dagger \boldsymbol{x}} d\boldsymbol{x}.$$

The nth-order moment corresponding to the product $\prod_{j=1}^r X_j^{q_j}$, where $\Sigma_{j=1}^r q_j = n$ and q_j are positive integers $\forall j$, is given by the coefficient of the term corresponding to

$$\frac{i^n \prod_{j=1}^r \omega_j^{q_j}}{\prod_{j=1}^r q_j!}$$

in the multidimensional series expansion of the CF. We need only consider the case where $r = n$ and therefore $q_j = 1$ $\forall j$. This is so because if some of the q_j are greater than one, we can simply consider a larger set of variables $\{X'_j\}_{j=1}^n$, where some of the X'_j are identical according to the values of q_j.[5]

The cumulants are given by the coefficients in the series expansion of the cumulative function $\ln \Phi_{\boldsymbol{X}}(\boldsymbol{\omega})$. Since we consider only $r = n$, the resulting cumulants are called *simple cumulants* [69]. Thus, the nth-order simple cumulant for the variables $\{X_j\}_{j=1}^n$ is given by

$$C_{\boldsymbol{X}} \triangleq (-i)^n \frac{\partial^n}{\partial \omega_1 \cdots \partial \omega_n} \ln \Phi_{\boldsymbol{X}}(\boldsymbol{\omega}) \bigg|_{\boldsymbol{\omega}=\boldsymbol{0}}. \tag{33}$$

3) Multivariate Moment and Cumulant Relations: For the set of random variables $\{X_j\}_{j=1}^n$, the joint moment is

$$R_{\boldsymbol{X}} = E\left\{ \prod_{j=1}^n X_j \right\}. \tag{34}$$

Let ν_k be some nonempty subset of the set of indices $\{1, 2, \ldots, n\}$. Then the moment of order $n_k = |\nu_k|$ for those variables with subscripts in ν_k is

$$R_{\boldsymbol{X},\nu_k} = E\left\{ \prod_{j \in \nu_k} X_j \right\}.$$

The nth-order simple cumulant can be expressed in terms of the moments $R_{\boldsymbol{X},\nu_k}$ by using the distinct partitions of the index set $\{1, 2, \ldots, n\}$ denoted by $P_n = \{\nu_k\}_{k=1}^p$ [69]:

$$C_{\boldsymbol{X}} = \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^p R_{\boldsymbol{X},\nu_j} \right]. \tag{35}$$

Similarly, the nth-order moment $R_{\boldsymbol{X}}$ can be expressed in terms of lower order simple cumulants [69]

$$R_{\boldsymbol{X}} = \sum_{P_n} \left[\prod_{j=1}^p C_{\boldsymbol{X},\nu_j} \right] \tag{36}$$

where $C_{\boldsymbol{X},\nu_j}$ is the simple cumulant of the variables $\{X_k\}_{k \in \nu_j}$.

An important and useful property of multivariate cumulants is the *independence property*. Consider the set of variables

$$\{Z_m\}_{m=1}^n = \{X_j\colon j = 1, \ldots, r;\ Y_k\colon k = 1, \ldots, s\}, \quad n = r + s$$

where the X_j are independent of the Y_k. The nth-order joint PDF for these variables factors

$$f_{\boldsymbol{XY}}(\boldsymbol{z}) = f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{Y}}(\boldsymbol{y}), \quad \boldsymbol{z} = [x_1 \cdots x_r\ y_1 \cdots y_s]^\dagger$$

which implies that the CF is the product of CF's for $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$\Phi_{\boldsymbol{XY}}(\boldsymbol{\omega}) = \Phi_{\boldsymbol{X}}(\boldsymbol{\omega}_x)\Phi_{\boldsymbol{Y}}(\boldsymbol{\omega}_y).$$

Therefore

$$\ln \Phi_{\boldsymbol{XY}}(\boldsymbol{\omega}) = \ln \Phi_{\boldsymbol{X}}(\boldsymbol{\omega}_x) + \ln \Phi_{\boldsymbol{Y}}(\boldsymbol{\omega}_y)$$

[5] An exception to this that occurs in some special anomalous cases is described in Appendix B.

and the nth-order derivative in (33) is zero. Other elementary properties of cumulants are listed in [10], [72], and [99].

For a stochastic process $X(t)$, cumulants of order n are defined by simply selecting n time-samples, $\{X(\tau_j)\}_{j=1}^n$, identifying these as n random variables $\{X_j\}_{j=1}^n$, and applying the definition of the cumulant for n random variables.

4) Cumulants of a Time-Series: The relations (35) and (36) are more accessible than the CF expression (33) and are, therefore, used for the cumulants and moments of time-series. Comparing equations (24) and (36), it is apparent that the relationship for $\{x(t+\tau_j)\}_{j=1}^n$ between the pure-sine-waves function $\sigma_x(t, \boldsymbol{\tau})_n$ and the temporal moment function $R_x(t, \boldsymbol{\tau})_n$ is equivalent to that between the simple cumulant and moment for $\{X_j\}_{j=1}^n$. Moreover, by using the sine-wave-extraction operation, which is an expectation operation, we can reexpress the TMF in terms of lower order simple cumulants:

$$R_{x,n} = \sum_{P_n} \left[\prod_{j=1}^{p} C_{x,\nu_j}\right] = C_{x,n} + \sum_{\substack{P_n \\ p\neq 1}} \left[\prod_{j=1}^{p} C_{x,\nu_j}\right] \quad (37)$$

where

$$R_{x,n} \triangleq \hat{E}^{\{\alpha\}}\left\{\prod_{j=1}^{n} x(t+\tau_j)\right\} = R_x(t,\boldsymbol{\tau})_n \quad (38)$$

$$C_{x,n} = \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^{p} R_{x,\nu_j}\right] \quad (39)$$

$$C_{x,\nu_j} = \text{Cumulant } \{x(t+\tau_k)\}_{k\in\nu_j}.$$

Then, using the equivalence (24) and (37), we obtain

$$C_{x,n} = \sigma_x(t,\boldsymbol{\tau})_n$$

where (37), (38), and (39) are identical to (36), (34), and (35) respectively, except that $\hat{E}^{\{\alpha\}}\{\cdot\}$ is used in place of $E\{\cdot\}$. To build on this mathematical duality between the pure-sine-waves function and the cumulant function, the notation

$$C_x(t,\boldsymbol{\tau})_n \equiv \sigma_x(t,\boldsymbol{\tau})_n \quad (40)$$

is used, and this function is called the *temporal cumulant function* (TCF) [39]. The fundamental relation (39) then takes the form

$$C_x(t,\boldsymbol{\tau})_n = \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^{p} R_x(t,\boldsymbol{\tau}_{\nu_j})_{n_j}\right]. \quad (41)$$

A Fourier coefficient of this polyperiodic function of t is given by

$$C_x^{\beta}(\boldsymbol{\tau})_n \triangleq \left\langle C_x(t,\boldsymbol{\tau})_n e^{-i2\pi\beta t}\right\rangle$$
$$= \sum_{P_n} (-1)^{p-1}(p-1)! \left[\sum_{\boldsymbol{\alpha}^{\dagger}\mathbf{1}=\beta} \prod_{j=1}^{p} R_x^{\alpha_j}(\boldsymbol{\tau}_{\nu_j})_{n_j}\right] \quad (42)$$

and is called the *cyclic temporal cumulant function* (CTCF) [39]. An individual component of the TCF, such as $C_x^{\beta}(\boldsymbol{\tau})_n e^{i2\pi\beta t}$, is called an *nth-order cumulant sine wave* to distinguish it from an nth-order moment sine wave. It can be seen from (25) and (40) that the nth-order CTCF is identical to the (complex-valued) strength of the pure nth-order sine wave with frequency β that is contained in the nth-order lag product $L_x(t,\boldsymbol{\tau})_n$. The cyclic temporal moment function in (13), on the other hand, gives the strength of the entire sine wave with frequency α that is contained in $L_x(t,\boldsymbol{\tau})_n$, which can be called the *impure nth-order sine wave*.

It should be pointed out that this is the first instance (to the best of our knowledge) that cumulants have arisen as the solution to a practically motivated problem, namely the problem of pure nth-order sine-wave generation [39], rather than as a mathematical observation concerning the characteristic function [15], [101], [61], [67].

D. Properties of the Temporal Parameters of HOCS

We have seen that the TMF can be constructed from all of the CTMF's (cf. (14)), that the TCF can be constructed from all of the CTCF's (cf. (27)), that the TMF can be constructed from all of the lower order TCF's (cf. (37)), and that the TCF can be constructed from all of the lower order TMF's (cf. (41)). Thus, any CTCF can be obtained from all of the appropriate CTMF's, and vice versa. In other words, the sets of moment and cumulant functions for orders 1 through n contain the same information. How then should we determine which set of functions to work with in the study of sine-wave generation? To assist us in making the correct choice, we consider some important properties of these functions.

Signal Selectivity: Suppose our time-series $x(t)$ consists of the sum of M mutually independent time-series

$$x(t) = \sum_{m=1}^{M} y_m(t). \quad (43)$$

Then, the TCF for $x(t)$ is the sum of TCF's for $\{y_m(t)\}$ [96]

$$C_x(t,\boldsymbol{\tau})_n = \sum_{m=1}^{M} C_{y_m}(t,\boldsymbol{\tau})_n. \quad (44)$$

Thus, the pure nth-order sine waves in the lag products of each of $y_m(t)$ add to form the pure nth-order sine wave in the lag product of $x(t)$. The TMF does not obey this very useful cumulative relation.

To illustrate how (44) can be applied in practice, consider the situation where $\{y_m(t)\}_{m=1}^M$ represent M interfering signals that overlap in time and frequency, but which possess some distinct nth-order cycle frequencies, say $\{\alpha_m\}_{m=1}^M$. Then it follows from (44) that

$$C_x^{\alpha_m}(\boldsymbol{\tau})_n = C_{y_m}^{\alpha_m}(\boldsymbol{\tau})_n, \quad m = 1, 2, \ldots, M.$$

This indicates that the presence or absence of each of the signals $y_m(t)$ can be detected by measuring (estimating) the CTCF's of $x(t)$ for the cycle frequencies $\{\alpha_m\}$, and that parameters of each of the signals, on which these CTCF's depend, can be estimated. As illustrated in [1], [34], [36], [38], [41]–[43], and [91] for second order and in the companion Part II [98], in this issue, for higher order, this signal-selectivity property can be exploited in numerous ways to accomplish noise-and-interference-tolerant signal detection and estimation.

(This was first recognized in [39] but has since appeared elsewhere [18]–[23], [48]–[50].)

As another application, let $M = 2$, $y_1(t)$ be non-Gaussian, and $y_2(t)$ be Gaussian. Then $C_{y_2}(t, \boldsymbol{\tau})_n \equiv 0$ for $n \geq 3$ and, from (43), we have

$$C_x(t, \boldsymbol{\tau})_n = C_{y_1}(t, \boldsymbol{\tau})_n, \quad n \geq 3$$

which indicates the detectability of $y_1(t)$ with no knowledge about $y_2(t)$ (except that it is Gaussian).

Reduced-Dimension Integrability: By using (13) and (42), it can be shown that both the CTMF and the CTCF are sinusoidal jointly in the n variables $\boldsymbol{\tau}$:

$$C_x^{\beta}(\boldsymbol{\tau} + \mathbf{1}\Delta)_n = C_x^{\beta}(\boldsymbol{\tau})_n e^{i2\pi\beta\Delta}, \tag{45}$$

$$R_x^{\alpha}(\boldsymbol{\tau} + \mathbf{1}\Delta)_n = R_x^{\alpha}(\boldsymbol{\tau})_n e^{i2\pi\alpha\Delta}. \tag{46}$$

Hence, the CTMF and CTCF are not absolutely integrable with respect to $\boldsymbol{\tau}$. The periodicity suggests that we might reduce the dimension of the functions and retain all the information present in the functions. Reducing the dimension by one yields[6]

$$\bar{C}_x^{\beta}(\boldsymbol{u})_n \triangleq C_x^{\beta}([\boldsymbol{u}^{\dagger} \ 0]^{\dagger})_n,$$
$$\bar{R}_x^{\alpha}(\boldsymbol{u})_n \triangleq R_x^{\alpha}([\boldsymbol{u}^{\dagger} \ 0]^{\dagger})_n$$

(where $\boldsymbol{u} = [u_1 \cdots u_{n-1}]^{\dagger}$), which are not sinusoidal. The value of $C_x^{\beta}(\boldsymbol{\tau})_n$ ($R_x^{\alpha}(\boldsymbol{\tau})_n$) for any $\boldsymbol{\tau}$ can be obtained from the value of $\bar{C}_x^{\beta}(\boldsymbol{u})_n$ ($\bar{R}_x^{\alpha}(\boldsymbol{u})_n$) by using (45) ((46)). This leads us to ask if these reduced-dimension (RD) functions are integrable. We shall show that the function $\bar{R}_x^{\alpha}(\boldsymbol{u})_n$ (RD-CTMF) is not in general, whereas the function $\bar{C}_x^{\alpha}(\boldsymbol{u})_n$ (RD-CTCF) is in general for time-series possessing an asymptotic independence property, that is, consider the arbitrary two-set partition $\boldsymbol{\tau} = [\boldsymbol{\tau}_0 \quad \boldsymbol{\tau}_1]$ and assume that the FOT density for $\{x(t + \tau_j)\}_{j=1}^{n}$ factors asymptotically:

$$f_{\boldsymbol{x}(t)}(\boldsymbol{y}) \to f_{\boldsymbol{x}(t)}(\boldsymbol{y}_0) f_{\boldsymbol{x}(t)}(\boldsymbol{y}_1) \quad \text{as } \boldsymbol{\tau}_0 \to \infty$$

where $\boldsymbol{\tau}_0 \to \infty$ means that all of the elements of $\boldsymbol{\tau}_0$ are tending to infinity. This asymptotic factorization implies that the TMF, which is a moment corresponding to the PDF $f_{\boldsymbol{x}(t)}(\cdot)$, is asymptotically factorable as well

$$\begin{aligned}\hat{E}^{\{\alpha\}}\{L_x(t, \boldsymbol{\tau})_n\} &= \hat{E}^{\{\alpha\}}\{L_x(t, \boldsymbol{\tau}_0)_{n_0} L_x(t, \boldsymbol{\tau}_1)_{n_1}\} \\ &\to \hat{E}^{\{\alpha\}}\{L_x(t, \boldsymbol{\tau}_0)_{n_0}\} \\ &\quad \times \hat{E}^{\{\alpha\}}\{L_x(t, \boldsymbol{\tau}_1)_{n_1}\} \quad \text{as } \boldsymbol{\tau}_0 \to \infty.\end{aligned}$$

Thus, the TCF is asymptotically zero

$$C_x(t, \boldsymbol{\tau})_n \to 0 \quad \text{as } \boldsymbol{\tau}_0 \to \infty$$

because of the independence property of cumulants (cf. Section II-C-3), which implies that each CTCF is asymptotically zero. Generally, then, the CTCF is asymptotically zero as long as at least one of the n lag values—say τ_n—is fixed since the set of variables associated with $L_x(t, \boldsymbol{\tau}_0)_{n_0}$ is asymptotically independent of the set of variables associated with $L_x(t, \boldsymbol{\tau}_1)_{n_1}$, where $\boldsymbol{\tau}_0 = [\tau_1 \cdots \tau_{n-1}]^{\dagger}$ and $\boldsymbol{\tau}_1 = [\tau_n]^{\dagger}$. On the other hand, the cyclic temporal moment becomes

$$\begin{aligned}R_x^{\alpha}(\boldsymbol{\tau})_n &\to \langle R_x(t, \boldsymbol{\tau}_0)_{n_0} R_x(t, \boldsymbol{\tau}_1)_{n_1} e^{-i2\pi\alpha t}\rangle \quad \text{as } \boldsymbol{\tau}_0 \to \infty \\ &= \sum_{\gamma} R_x^{\gamma}(\boldsymbol{\tau}_0)_{n_0} R_x^{\alpha-\gamma}(\boldsymbol{\tau}_1)_{n_1}\end{aligned}$$

which is not necessarily zero. In fact, it is often nonzero.

If the rate of decay of the RD-CTCF is sufficiently large (e.g., $O(\|\boldsymbol{u}\|^{-2})$), then $\bar{C}_x^{\alpha}(\boldsymbol{u})_n$ is absolutely integrable and, therefore, Fourier transformable. The RD-CTMF is not, in general, Fourier transformable except in a generalized sense that accommodates Dirac delta functions, because it does not decay as its arguments grow without bound. We shall see in Section III that the Fourier transforms of the RD-CTCF and RD-CTMF can be very useful in characterizing a signal's higher order frequency-domain behavior. Before leaving the subject of temporal parameters, however, we continue with our AM example.

E. Example—Continued

We continue to develop the example of Sections II–A and II–B–1. Here we give the formulas for the RD-CTMF and RD-CTCF for the time-series

$$x(t) = a(t)\cos(2\pi\nu_1 t + \theta_1) + \cos(2\pi\nu_2 t + \theta_2).$$

The RD-CTMF is given by the CTMF with $\tau_2 = 0$, which is (cf. (18))

$$\bar{R}_x^{\alpha}(\tau_1)_2 = \begin{cases} \frac{1}{2}R_a(\tau_1)\cos(2\pi\nu_1\tau_1) + \frac{1}{2}\cos(2\pi\nu_2\tau_1), & \alpha = 0, \\ \frac{1}{4}R_a(\tau_1)\exp(\pm i(2\pi\nu_1\tau_1 + 2\theta_1)), & \alpha = \mp 2\nu_1, \\ \frac{1}{4}\exp(\pm i(2\pi\nu_2\tau_1 + 2\theta_2)), & \alpha = \mp 2\nu_2, \\ 0, & \text{otherwise} \end{cases}$$

and, from (28), the RD-CTCF is given by

$$\bar{C}_x^{\beta}(\tau_1)_2 = \begin{cases} \frac{1}{2}R_a(\tau_1)\cos(2\pi\nu_1\tau_1), & \beta = 0 \\ \frac{1}{4}R_a(\tau_1)\exp(\pm i(2\pi\nu_1\tau_1 + 2\theta_1)), & \beta = \mp 2\nu_1, \\ 0, & \text{otherwise.} \end{cases}$$

By comparing $\bar{R}_x^{\alpha}(\tau_1)_2$ with $\bar{C}_x^{\beta}(\tau_1)_2$, we see that the former is not integrable for some cycle frequencies (viz., $\alpha = 0, \mp 2\nu_2$), whereas the latter is integrable for all cycle frequencies, provided that $R_a(\tau_1)$ decays faster than $1/\tau_1$. In addition, it is clear that the RD-CTCF characterizes only those sine waves in the second-order lag product that are not the result of first-order sine-wave multiplications.[7]

[6] The reason for this particular choice of dimension reduction is made clear in Section III.

[7] An anomaly regarding the relationship between pure nth-order sine waves and the nth-order TCF is explained in the Appendix.

III. The Spectral Parameters of HOCS

The Fourier transform of $\bar{R}_x^0(\boldsymbol{u})_2$ is the power spectral density (PSD) of $x(t)$ and the Fourier transform of $\bar{C}_x^0(\boldsymbol{u})_2$ is the PSD of the centered version $x(t) - \hat{E}^{\{\alpha\}}\{x(t)\}$ of $x(t)$ (this is the Wiener relation [105], cf. Section IV–A and [34]). The Fourier transform of the symmetrized version $\langle x(t+u/2)x^*(t-u/2)e^{-i2\pi\alpha u}\rangle$ of $\bar{R}_x^\alpha(\boldsymbol{u})_2$ for $\alpha \neq 0$ is the spectral correlation function (cyclic spectral density function), and the Fourier transform of the corresponding symmetrized version of $\bar{C}_x^\alpha(\boldsymbol{u})_2$ is the spectral correlation function for the time-series $x(t)$ with its first-order sine waves removed (this is the cyclic Wiener relation [34], cf. Section IV–A). Therefore, we could define the spectral parameters of HOCS to be the Fourier transforms of $\bar{R}_x^\alpha(\boldsymbol{u})_n$ and $\bar{C}_x^\alpha(\boldsymbol{u})_n$, whenever such transforms exist. These transforms are indeed the central spectral parameters of the theory of HOCS, but it is more natural (for those interested in using the theory in practice) to derive them from a consideration of spectral moments and spectral cumulants; that is, from limiting versions (as bandwidth approaches zero) of moments and cumulants of narrowband spectral components of $x(t)$, and then to show that they can be characterized as Fourier transforms of temporal moments and cumulants.

It is assumed that $x(t)$ is absolutely integrable on finite intervals. We consider the complex envelope of the spectral component of a segment of $x(u)$ that is centered at t and has width T:

$$X_T(t,f) \triangleq \int_{t-T/2}^{t+T/2} x(v)e^{-i2\pi f v}\,dv. \tag{47}$$

The temporal moment of the set of n variables $\{X_T(t,f_j)\}_{j=1}^n$ is defined by[8]

$$\begin{aligned} S_{x_T}(\boldsymbol{f})_n &\triangleq \left\langle \prod_{j=1}^n X_T(t,f_j)\right\rangle, \quad \boldsymbol{f} \triangleq [f_1 \cdots f_n]^\dagger \\ &= \lim_{Z\to\infty} \frac{1}{Z}\int_{-Z/2}^{Z/2} \prod_{j=1}^n X_T(t,f_j)\,dt \end{aligned} \tag{48}$$

and is assumed for the time being to exist. If we now let the integration time T in (47) tend to infinity in (48), we obtain the *spectral moment function* (SMF)

$$\begin{aligned} S_x(\boldsymbol{f})_n &\triangleq \lim_{T\to\infty} S_{x_T}(\boldsymbol{f})_n \\ &= \lim_{T\to\infty}\lim_{Z\to\infty} \frac{1}{Z}\int_{-Z/2}^{Z/2} \prod_{j=1}^n X_T(t,f_j)\,dt. \end{aligned} \tag{49}$$

However, this limit exists only in a generalized sense that accommodates products of Dirac deltas (impulse functions). Nevertheless, we shall see that Dirac deltas can be avoided by working with the cumulant counterpart of this moment.

To see that the SMF (49) is composed of products of impulse functions, we proceed as follows. The function (48) can be expressed in terms of the CTMF as

$$\begin{aligned} S_{x_T}(\boldsymbol{f})_n &= \left\langle \int_{-T/2}^{T/2}\cdots\int_{-T/2}^{T/2} \prod_{j=1}^n x(t+v_j)e^{-i2\pi f_j(t+v_j)}\,d\boldsymbol{v}\right\rangle \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} w_T(\boldsymbol{v})_n R_x^{\alpha_0}(\boldsymbol{v})_n e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{v}}\,d\boldsymbol{v} \end{aligned} \tag{50}$$

where

$$\alpha_0 \triangleq \sum_{j=1}^n f_j = \boldsymbol{f}^\dagger \mathbf{1}, \quad w_T(\boldsymbol{v})_n \triangleq \prod_{j=1}^n \mathrm{rect}(v_j/T),$$

$$\mathrm{rect}(t) \triangleq \begin{cases} 1, & |t| \le 1/2, \\ 0, & |t| > 1/2. \end{cases}$$

Thus, $S_{x_T}(\boldsymbol{f})_n$ is nonzero only if the sum α_0 of the frequencies is equal to an nth-order cycle frequency α of $x(t)$. Now, assuming that $R_x^\alpha(\boldsymbol{v})_n$ is absolutely integrable on the hypercube of size T on a side, we see that the Fourier transform (50) exists and that (48) therefore exists. Assuming for the time being that $R_x^\alpha(\boldsymbol{v})_n$ is Fourier transformable on the entire space

$$S_x^\alpha(\boldsymbol{f})_n \triangleq \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} R_x^\alpha(\boldsymbol{\tau})_n e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{\tau}}\,d\boldsymbol{\tau} \tag{51}$$

we obtain from (50)

$$S_{x_T}(\boldsymbol{f})_n = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} S_x^{\alpha_0}(\boldsymbol{f}-\boldsymbol{g})_n \prod_{k=1}^n T\,\mathrm{sinc}(g_k T)\,d\boldsymbol{g}$$

where

$$\mathrm{sinc}(f) \triangleq \frac{\sin(\pi f)}{\pi f}.$$

Thus, the finite-time spectral moment $S_{x_T}(\boldsymbol{f})_n$ converges to

$$\begin{aligned} S_x(\boldsymbol{f})_n &\triangleq \lim_{T\to\infty} S_{x_T}(\boldsymbol{f})_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} S_x^{\alpha_0}(\boldsymbol{f}-\boldsymbol{g})_n \prod_{k=1}^n \delta(g_k)\,d\boldsymbol{g} \\ &= S_x^{\alpha_0}(\boldsymbol{f})_n \end{aligned} \tag{52}$$

where $\delta(\cdot)$ is the Dirac delta. Let us investigate this hypothetical Fourier transform (51). Using the fact that $R_x^\alpha(\boldsymbol{\tau})_n$ is sinusoidal in the translation variables (cf. (46)), we can formally show that

$$S_x^\alpha(\boldsymbol{f})_n = \bar{S}_x^\alpha(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger\mathbf{1} - \alpha), \quad \boldsymbol{f}' \triangleq [f_1 \cdots f_{n-1}]^\dagger \tag{53}$$

where $\bar{S}_x^\alpha(\boldsymbol{f}')_n$ is the Fourier transform of the RD-CTMF

$$\bar{S}_x^\alpha(\boldsymbol{f}')_n \triangleq \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \bar{R}_x^\alpha(\boldsymbol{u})_n e^{-i2\pi \boldsymbol{u}^\dagger \boldsymbol{f}'}\,d\boldsymbol{u}. \tag{54}$$

Thus, we have the formal result

$$S_x(\boldsymbol{f})_n = \begin{cases} \bar{S}_x^\alpha(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger\mathbf{1} - \alpha), & \boldsymbol{f}^\dagger\mathbf{1} = \alpha, \\ 0, & \boldsymbol{f}^\dagger\mathbf{1} \neq \alpha \end{cases} \tag{55}$$

for all cycle frequencies α of $x(t)$. The SMF can be reexpressed as

$$S_x(\boldsymbol{f})_n = \sum_\alpha \bar{S}_x^\alpha(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger\mathbf{1} - \alpha) \tag{56}$$

[8] It would be more consistent to use $\hat{E}^{\{\alpha\}}\{\cdot\}$ in place of $\langle\cdot\rangle$ in the definition of $S_{x_T}(\boldsymbol{f})_n$, but it can be shown that these two operations lead to the same function (49). Thus, we start with the time-average operation $\langle\cdot\rangle$.

which reveals that the SMF is a sum of components with impulsive factors. Moreover, we can show that $\bar{S}_x^\alpha(\boldsymbol{f}')_n$ can also be a sum of components with impulsive factors and even products of impulses (cf. Section II–E). Thus, neither the SMF nor the reduced-dimension SMF (RD-SMF) (54) are well-behaved functions.

The *spectral cumulant function* (SCF) is better behaved than the SMF. To establish this fact, we proceed in a manner analogous to that used for the SMF to obtain a characterization of the SCF in terms of the Fourier transform of the RD-CTCF. The simple cumulant of the variables $\{X_T(t, f_j)\}_{j=1}^n$ is given by

$$\begin{aligned} P_{x_T}(\boldsymbol{f})_n &\triangleq \text{Cumulant}\{X_T(t, f_j)\}_{j=1}^n \\ &= \sum_{P_n} \left[(-1)^{p-1}(p-1)! \prod_{j=1}^{p} S_{x_T}(\boldsymbol{f}_{\nu_j})_{n_j} \right]. \end{aligned} \tag{57}$$

This function is well-defined for finite T since each moment $S_{x_T}(\cdot)_{n_j}$ is finite. The spectral cumulant function is defined to be the limit

$$P_x(\boldsymbol{f})_n \triangleq \lim_{T\to\infty} P_{x_T}(\boldsymbol{f})_n. \tag{58}$$

We can use (50) to reexpress $P_{x_T}(\boldsymbol{f})_n$ in terms of lower order CTMF's:

$$P_{x_T}(\boldsymbol{f})_n = \sum_{P_n} k(p) \prod_{j=1}^{p} \times \left[\underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n_j} w_T(\boldsymbol{v}_{\nu_j})_{n_j} R_x^{\alpha_j}(\boldsymbol{v}_{\nu_j})_{n_j} e^{-i2\pi \boldsymbol{f}_{\nu_j}^\dagger \boldsymbol{v}_{\nu_j}} d\boldsymbol{v}_{\nu_j} \right]$$

where

$$k(p) \triangleq (-1)^{p-1}(p-1)!, \quad \alpha_j \triangleq \sum_{k\in\nu_j} f_k.$$

From this expression, we see that if for every partition in the set P_n (except that for $p = 1$), there is some α_j that is *not* a cycle frequency of order $n_j = |\nu_j|$, then the function $P_{x_T}(\boldsymbol{f})_n$ is equal to the function $S_{x_T}(\boldsymbol{f})_n$. If there is at least one partition such that all the α_j for that partition are cycle frequencies of order n_j, then the function $P_{x_T}(\boldsymbol{f})_n$ differs from $S_{x_T}(\boldsymbol{f})_n$. This is important when considering methods for measuring $P_{x_T}(\boldsymbol{f})_n$ and its components (cf. the companion Part II, in this issue).

We can reexpress $P_{x_T}(\boldsymbol{f})_n$ more compactly as

$$\begin{aligned} P_{x_T}(\boldsymbol{f})_n &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{P_n} k(p) \\ &\quad \times \left[\prod_{j=1}^{p} w_T(\boldsymbol{v}_{\nu_j})_{n_j} R_x^{\alpha_j}(\boldsymbol{v}_{\nu_j})_{n_j} \right] e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{v}} d\boldsymbol{v} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} w_T(\boldsymbol{v})_n \\ &\quad \times \left[\sum_{P_n} k(p) \prod_{j=1}^{p} R_x^{\alpha_j}(\boldsymbol{v}_{\nu_j})_{n_j} \right] e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{v}} d\boldsymbol{v}. \end{aligned} \tag{59}$$

By using (50), we can show that this last expression is equivalent to the Fourier transform of the CTCF on a hypercube of size T on a side:

$$P_{x_T}(\boldsymbol{f})_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} w_T(\boldsymbol{v})_n C_x^{\alpha_0}(\boldsymbol{v})_n e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{v}} d\boldsymbol{v}. \tag{60}$$

By analogy with the preceding argument for the SMF, we obtain

$$P_x(\boldsymbol{f})_n = \sum_{\beta} \bar{P}_x^\beta(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger \mathbf{1} - \beta) \tag{61}$$

where

$$\bar{P}_x^\beta(\boldsymbol{f}')_n \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{C}_x^\beta(\boldsymbol{u})_n e^{-i2\pi \boldsymbol{u}^\dagger \boldsymbol{f}'} d\boldsymbol{u} \tag{62}$$

is (by analogy with the accepted terminology *cyclic spectrum* for $\beta \neq 0$ and $n = 2$ and the accepted terminology *polyspectrum* for $\beta = 0$ and $n > 2$) defined to be the *cyclic polyspectrum* (CP). The transform (62) does exist (in the strict sense that excludes Dirac deltas) in general for time-series with asymptotically independent variables such that the RD-CTCF decays sufficiently rapidly in all directions so that $\bar{C}_x^\beta(\boldsymbol{u})_n$ is absolutely integrable, and hence Fourier transformable (cf. Section II–D–2).

Because of the characterization (61) of the SCF, we can see that the SCF is nonzero only on the hyperplanes specified by $\sum_{j=1}^n f_j = \beta$, where β is in the set of pure nth-order cycle frequencies of the time-series $x(t)$ (frequencies of pure nth-order sine waves). We can also see from (61) that the CP is the integrated SCF

$$\bar{P}_x^\beta(\boldsymbol{f}')_n = \int_{\beta^-}^{\beta^+} P_x(\boldsymbol{f})_n df_n$$

where (β^-, β^+) includes the value β but excludes all other pure cycle frequencies. Similarly, the RD-SMF is the integrated SMF.

In this section we have seen that the CP, which is the integrated SCF, is in general the only well-behaved spectral function in the theory of HOCS. The SMF and its reduced-dimension version $\bar{S}_x^\alpha(\boldsymbol{f}')_n$ in general contain products of impulses and are, therefore, not well-behaved functions. However, in the special case where the lowest order of cyclostationarity of $x(t)$ is n, the impure nth-order sine waves (with strengths given by the CTMF's) are identical to the pure nth-order sine waves (with strengths given by the CTCF's) and, as a result, the nth-order SCF is identical to the nth-order SMF, which results in equality between the CP and the RD-SMF (for $\alpha \neq 0$). In addition, there are certain values of the frequency vector $\boldsymbol{f}'$ for which the RD-SMF and the CP are equal even when $x(t)$ exhibits lower order cyclostationarity. For these $\boldsymbol{f}'$, the CP can be measured by measuring the RD-SMF, which is computationally simpler, since lower order moments do not have to be estimated and then combined. This is explained in Part II, in this issue.

A. Example—Continued

We finish the example of an AM time-series plus an unmodulated carrier (15). We can formally determine the SMF by Fourier transforming the CTMF

$$S_x(\boldsymbol{f})_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x^{\alpha_0}(\boldsymbol{\tau})_2 e^{-i2\pi \boldsymbol{f}^\dagger \boldsymbol{\tau}}\, d\boldsymbol{\tau}, \quad \alpha_0 = f_1 + f_2.$$

Thus, the SMF is given by

$$\begin{aligned} S_x(\boldsymbol{f})_2 = &\frac{1}{4}[S_a(f_1+\nu_1) + S_a(f_1-\nu_1) + \delta(f_1+\nu_2) \\ &+ \delta(f_1-\nu_2)]\delta(f_1+f_2) \\ &+ \frac{1}{4}e^{\pm i2\theta_1} S_a(f_1 \mp \nu_1)\delta(f_1+f_2\mp 2\nu_1) \\ &+ \frac{1}{4}e^{\pm i2\theta_2}\delta(f_1\mp\nu_2)\delta(f_2\mp\nu_2) \end{aligned} \quad (63)$$

and the SCF is given by the Fourier transform of the CTCF

$$\begin{aligned} P_x(\boldsymbol{f})_2 = &\frac{1}{4}[S_a(f_1+\nu_1) + S_a(f_1-\nu_1)]\delta(f_1+f_2) \\ &+ \frac{1}{4}e^{\pm i2\theta_1} S_a(f_1\mp\nu_1)\delta(f_1+f_2\mp 2\nu_1). \end{aligned} \quad (64)$$

The RD-SMF is given by the Fourier transform of the RD-CTMF

$$\bar{S}_x^{\alpha}(f_1)_2 = \begin{cases} \frac{1}{4}[S_a(f_1+\nu_1) + S_a(f_1-\nu_1) \\ \quad + \delta(f_1+\nu_2) + \delta(f_1-\nu_2)] & \alpha = 0, \\ \frac{1}{4}e^{\pm i2\theta_1} S_a(f_1\mp\nu_1), & \alpha = \mp 2\nu_1, \\ \frac{1}{4}e^{\pm i2\theta_2}\delta(f_1\mp\nu_2), & \alpha = \mp 2\nu_2, \\ 0, & \text{otherwise} \end{cases} \quad (65)$$

and the CP is the Fourier transform of the RD-CTCF

$$\bar{P}_x^{\beta}(f_1)_2 = \begin{cases} \frac{1}{4}[S_a(f_1+\nu_1) + S_a(f_1-\nu_1)], & \beta = 0, \\ \frac{1}{4}e^{\pm i2\theta_1} S_a(f_1\mp\nu_1), & \beta = \mp 2\nu_1, \\ 0, & \text{otherwise.} \end{cases} \quad (66)$$

It can be seen from (63)–(66) that the CP is the only well-behaved function for the time-series considered; that is, the CP does not contain any impulse functions. It can also be seen that the SCF and CP are signal selective in that they are completely determined by the properties of the AM signal $a(t)\cos(2\pi\nu_1 t+\theta_1)$ and do not reflect in any way the presence of the additive sine wave $\cos(2\pi\nu_2 t+\theta_2)$. This is a degenerate example of signal selectivity since the signal $\cos(2\pi\nu_2 t+\theta_2)$ is not erratic and does not necessarily spectrally overlap the other signal.

IV. Discussion

The relation (88) between the CP and the RD-CTCF is the generalization of the cyclic Wiener relation introduced in [34] from order 2 to order n. Within the stochastic process framework, the Fourier transform relation between temporal and spectral cumulant functions for generally nonstationary processes was first obtained in [92] and should be called the *Shiryaev-Kolmogorov (SK) relation*, which is the generalization of the Wiener-Khinchin relation for the PSD of a stationary process. Since the only type of nonstationarity for which temporal counterparts of stochastic moments and cumulants exist is cyclostationarity, we see that the relation (88) is the nonstochastic counterpart of the SK relation.

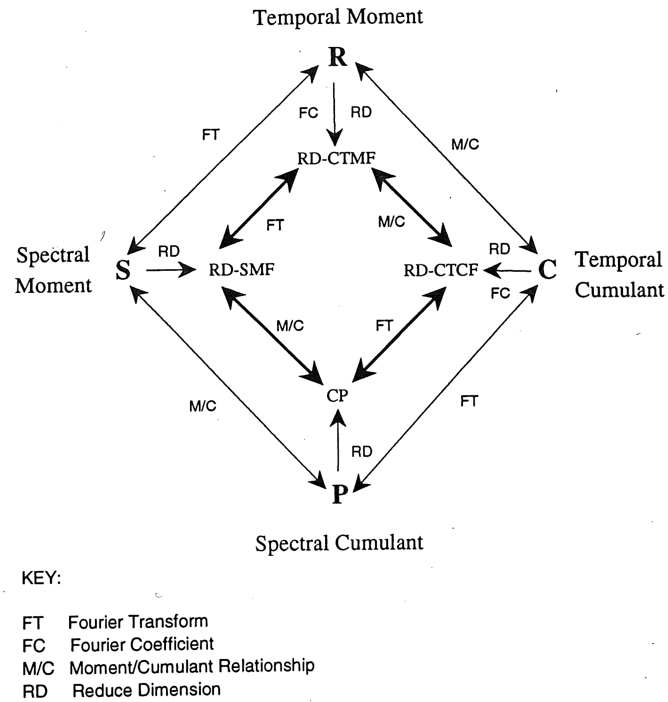


Fig. 1. A pictorial representation of the relationships between the parameters of higher order cyclostationarity. The parameters of higher order stationarity as they are typically defined correspond to the inner diamond (bold arrows) because in this case the time-invariance of the time-domain quantities suggests the dimension reduction. The quantities in the upper left half of the diagram are equivalent to those in the lower right for zero-mean signals and $n = 2$ and 3 because in this case moments and cumulants are equal. Thus, each of these halves contains the relation between the cyclic spectrum and the cyclic autocorrelation for zero-mean cyclostationary signals and, as a special case, contains the Wiener relation (bold arrow).

The relationships between all the parameters of higher order cyclostationarity are shown graphically in Fig. 1. In this figure, the lines represent functional relationships between the quantities at the arrowheads (see key).

A. HOCS and Second-Order Probabilistic Parameters

In this section we examine the relationship between the parameters of HOCS for $n = 2$ and the well-established parameters of second-order cyclostationarity (SOCS), which include the (nonstochastic) autocorrelation and PSD as special cases.

The autocorrelation function for a real time-series $x(t)$ is defined to be

$$R_x(\tau) \triangleq \langle x(t+\tau/2)x(t-\tau/2)\rangle \quad (67)$$

which is obtained by the time-averaging operation $\langle\cdot\rangle$. This function does not in general describe the second-order cyclostationarity (if any exists) of $x(t)$. To do that we need to use the sine-wave extraction operation to obtain the second-order TMF for $\tau_1 = \tau/2$ and $\tau_2 = -\tau/2$

$$R_x(t, \tau) \triangleq \hat{E}^{\{\alpha\}}\{x(t + \tau/2)x(t - \tau/2)\}. \tag{68}$$

For a stationary time-series, (67) and (68) are identical, but for a cyclostationary time-series we have

$$R_x(t, \tau) = \sum_{\alpha} R_x^{\alpha}(\tau) e^{i2\pi\alpha t} \tag{69}$$

where

$$R_x^{\alpha}(\tau) \triangleq \langle x(t + \tau/2)x(t - \tau/2)e^{-i2\pi\alpha t}\rangle. \tag{70}$$

Equations (69) and (70) define the central time-domain parameters of SOCS for real time-series [34]. The function (70) is called the *cyclic autocorrelation function*. We can relate the cyclic autocorrelation function to the RD-CTMF for $n = 2$ easily, since

$$\begin{aligned}\bar{R}_x^{\alpha}(u)_2 &= \langle x(t + u)x(t)e^{-i2\pi\alpha t}\rangle \\ &= \langle x(t + u/2)x(t - u/2)e^{-i2\pi\alpha t}\rangle e^{i\pi\alpha u}\end{aligned}$$

which implies that the RD-CTMF for $n = 2$ is related to the cyclic autocorrelation by a sinusoidal factor

$$\bar{R}_x^{\alpha}(u)_2 = R_x^{\alpha}(u)e^{i\pi\alpha u}. \tag{71}$$

The *spectral correlation function* or *cyclic spectrum* is the limit as the bandwidth tends to zero ($T \rightarrow \infty$) of the time-averaged product of spectral components with approximate bandwidth $1/T$ and frequency separation α

$$S_x^{\alpha}(f) = \lim_{T\rightarrow\infty} \left\langle \frac{1}{T} X_T(t, f + \alpha/2)X_T^*(t, f - \alpha/2)\right\rangle. \tag{72}$$

This function is the Fourier transform of the cyclic autocorrelation function (70)

$$S_x^{\alpha}(f) = \int_{-\infty}^{\infty} R_x^{\alpha}(\tau)e^{-i2\pi f\tau} d\tau. \tag{73}$$

The relation (73) is the *cyclic Wiener relation*, and it reduces to the *Wiener relation* between the PSD and the autocorrelation when $\alpha = 0$

$$S_x^{0}(f) = \int_{-\infty}^{\infty} R_x^{0}(\tau)e^{-i2\pi f\tau} d\tau. \tag{74}$$

Combining (54), (71), and (73) yields

$$\begin{aligned}\bar{S}_x^{\alpha}(f')_2 &= \int_{-\infty}^{\infty} \bar{R}_x^{\alpha}(u)_2 e^{-i2\pi f' u} du \\ &= \int_{-\infty}^{\infty} R_x^{\alpha}(u)e^{i\pi\alpha u}e^{-i2\pi f' u} du\end{aligned}$$

which implies that the Fourier transform of the RD-CTMF, $\bar{S}_x^{\alpha}(f')_2$, is related to the cyclic spectrum by a frequency shift

$$\bar{S}_x^{\alpha}(f')_2 = S_x^{\alpha}(f' - \alpha/2). \tag{75}$$

Note that the function $\bar{S}_x^{\alpha}(f')_2$ can contain impulses that are due to the first-order sine-wave components of the data $x(t)$. In the development of the theory of SOCS, it is most natural to assume that the data does not contain such finite-strength additive sine waves. In this case, $\bar{S}_x^{\alpha}(f')_2$ does not contain impulses and, as a result, the moments and cumulants are equal

$$\begin{aligned}\bar{R}_x^{\alpha}(u)_2 &\equiv \bar{C}_x^{\alpha}(u)_2, \\ \bar{S}_x^{\alpha}(f')_2 &\equiv \bar{P}_x^{\alpha}(f')_2.\end{aligned}$$

Thus, in this special case, the cyclic polyspectrum is equal to the shifted cyclic spectrum

$$\bar{P}_x^{\alpha}(f')_2 = S_x^{\alpha}(f' - \alpha/2)$$

and the CP for $\alpha = 0$ is equal to the PSD

$$\bar{P}_x^{0}(f')_2 = S_x^{0}(f').$$

We conclude that the parameters of HOCS that are defined in this paper are consistent (to within a frequency shift) with the previously developed second-order parameters for cyclostationary time-series, and are consistent with the notions of autocorrelation and power spectrum and are, therefore, properly referred to as generalizations of the second-order parameters. The same can be said of the parameters of HOCS and SOCS generalized from real-valued time-series to complex-valued time-series as done in the companion paper, Part II, for HOCS and in [34] for SOCS.

V. Concluding Remarks

In this paper, we have presented the appropriate parameters (statistical functions) for conveniently characterizing the sine waves that are generated by performing higher order nonlinear transformations on cyclostationary time-series. We have shown that the temporal cumulants of such time-series provide a mathematical characterization of the notion of a pure nth-order sine wave, which is that part of the sine wave present in an nth-order lag-product waveform that remains after removal of all parts that result from products of sine waves in lower order lag products obtained by factoring the nth-order product. We have also shown that the natural definitions of spectral moments and cumulants are characterized by Fourier transforms of temporal moments and cumulants. Most importantly, from a practical standpoint, we have shown that the temporal and spectral cumulants exhibit the property of signal selectivity, which means that they can be used to detect the presence of and/or estimate the parameters of a specific signal in a received waveform, even when that signal is corrupted by temporally and spectrally overlapping stationary and cyclostationary noise and interference, provided only that the signal has a unique cycle frequency (for instance, a unique keying rate or carrier frequency).

In the sequel to this Part I, the parameters of HOCS are generalized from real-valued time-series to complex-valued time-series; the effects of signal-processing operations on the HOCS parameters are determined; the parameters of HOCS for complex-valued pulse-amplitude-modulated time-series are calculated; measurement methods are described; and applications of the theory to signal processing problems, including weak-signal detection and time-delay estimation, are discussed.

APPENDIX A
SUMMARY OF THE PARAMETERS OF HOCS

Summary of the Temporal Parameters: In summary, the acronyms and parameter names for the temporal parameters of HOCS are as follows:

TMF	temporal moment function
CTMF	cyclic temporal moment function
RD-CTMF	reduced-dimension cyclic temporal moment function
TCF	temporal cumulant function
CTCF	cyclic temporal cumulant function
RD-CTCF	reduced-dimension cyclic temporal cumulant function

$$\text{TMF: } R_x(t, \boldsymbol{\tau})_n \triangleq \hat{E}^{\{\alpha\}}\left\{\prod_{j=1}^{n} x(t+\tau_j)\right\}, \tag{76}$$

$$\boldsymbol{\tau} = [\tau_1 \cdots \tau_n]^\dagger$$

$$\text{CTMF: } R_x^\alpha(\boldsymbol{\tau})_n \triangleq \langle R_x(t, \boldsymbol{\tau})_n e^{-i2\pi\alpha t}\rangle = \left\langle \prod_{j=1}^{n} x(t+\tau_j)e^{-i2\pi\alpha t}\right\rangle \tag{77}$$

$$\text{RD-CTMF: } \bar{R}_x^\alpha(\boldsymbol{u})_n \triangleq \left\langle x(t)\prod_{j=1}^{n-1} x(t+u_j)e^{-i2\pi\alpha t}\right\rangle, \tag{78}$$

$$\boldsymbol{u} = [u_1 \cdots u_{n-1}]^\dagger$$

$$\text{TCF: } C_x(t, \boldsymbol{\tau})_n = \sum_{P_n = \{\nu_k\}_{k=1}^p} \left[(-1)^{p-1}(p-1)! \times \prod_{j=1}^{p} R_x(t, \boldsymbol{\tau}_{\nu_j})_{n_j}\right] \tag{79}$$

$$\text{CTCF: } C_x^\beta(\boldsymbol{\tau})_n \triangleq \langle C_x(t, \boldsymbol{\tau})_n e^{-i2\pi\beta t}\rangle = \sum_{P_n = \{\nu_k\}_{k=1}^p} \left[(-1)^{p-1}(p-1)! \times \sum_{\boldsymbol{\alpha}^\dagger \mathbf{1} = \beta} \prod_{j=1}^{p} R_x^{\alpha_j}(\boldsymbol{\tau}_{\nu_j})_{n_j}\right] \tag{80}$$

$$\text{RD-CTCF: } \bar{C}_x^\beta(\boldsymbol{u})_n \triangleq C_x^\beta([\boldsymbol{u}\ 0]^\dagger)_n. \tag{81}$$

Summary of the Spectral Parameters: In summary, the acronyms and parameter names for the spectral parameters of HOCS are as follows:

SMF	spectral moment function
RD-SMF	reduced-dimension spectral moment function
SCF	spectral cumulant function
CP	cyclic polyspectrum

$$\text{SMF: } S_x(\boldsymbol{f})_n \triangleq \lim_{T\to\infty} S_{x_T}(\boldsymbol{f})_n \triangleq \lim_{T\to\infty} \left\langle \prod_{j=1}^{n} X_T(t, f_j)\right\rangle, \tag{82}$$

$$\boldsymbol{f} = [f_1 \cdots f_n]^\dagger$$

$$= \sum_\alpha \bar{S}_x^\alpha(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger \mathbf{1} - \alpha), \tag{83}$$

$$\boldsymbol{f}' = [f_1 \cdots f_{n-1}]^\dagger$$

$$\text{RD-SMF: } \bar{S}_x^\alpha(\boldsymbol{f}')_n \triangleq \mathcal{F}^{n-1}\{\bar{R}_x^\alpha(\boldsymbol{u})_n\} \tag{84}$$

$$\text{SCF: } P_x(\boldsymbol{f})_n \triangleq \lim_{T\to\infty} P_{x_T}(\boldsymbol{f})_n \tag{85}$$

$$\triangleq \lim_{T\to\infty} \sum_{P_n = \{\nu_k\}_{k=1}^p} \left[(-1)^{p-1}(p-1)! \times \prod_{j=1}^{p} S_{x_T}(\boldsymbol{f}_{\nu_j})_{n_j}\right] \tag{86}$$

$$= \sum_\beta \bar{P}_x^\beta(\boldsymbol{f}')_n \delta(\boldsymbol{f}^\dagger \mathbf{1} - \beta) \tag{87}$$

$$\text{CP: } \bar{P}_x^\beta(\boldsymbol{f}')_n \triangleq \mathcal{F}^{n-1}\{\bar{C}_x^\beta(\boldsymbol{u})_n\}. \tag{88}$$

APPENDIX B
PURE SINE WAVES AND THE TCF

In Section II–B, the temporal cumulant function is derived by considering the problem of pure nth-order sine-wave generation. That is, it is found that pure nth-order sine-wave strengths are characterized by the cyclic temporal cumulant function. However, we have found that some mathematically idealized models for time-series possess a degeneracy such that, for certain values of the lag vector τ, the CTCF is nonzero when there are no pure nth-order sine waves present. This degeneracy is illustrated and explained here for the case of $n = 4$.

Consider the fourth-order lag product

$$L_x(t, \boldsymbol{\tau})_4 = x(t+\tau_1)x(t+\tau_2)x(t+\tau_3)x(t+\tau_4)$$

for the binary (± 1) PAM signal

$$x(t) = \sum_{m=-\infty}^{\infty} a_m p(t + mT_0 + t_0)$$

with rectangular pulses

$$p(t) = \begin{cases} 1, & |t| \le T_0/2 \\ 0, & \text{otherwise.} \end{cases}$$

There can be sine waves associated with the factors

$$f_1(t) = x(t+\tau_1)x(t+\tau_2)$$

and

$$f_2(t) = x(t+\tau_3)x(t+\tau_4)$$

that make up $L_x(t, \boldsymbol{\tau})_4$, and if so, the product sine waves

$$\hat{E}^{\{\alpha\}}\{f_1(t)\}\hat{E}^{\{\alpha\}}\{f_2(t)\}$$

are subtracted, along with product sine waves corresponding to all other unique factorizations of $L_x(t, \boldsymbol{\tau})_4$, from the moment $\hat{E}^{\{\alpha\}}\{L_x(t, \boldsymbol{\tau})_4\}$ to obtain the cumulant sine waves (cf. (79)). Thus, a pure fourth-order sine wave contains no products of lower order sine waves, and cannot be equal to such a product; this is simply the intuitive notion of a pure fourth-order sine wave. To see the degeneracy that we have alluded to, $\tau_3 = \tau_4$ is chosen so that the fourth-order lag product under consideration is

$$L_x(t, \boldsymbol{\tau})_4 = x(t+\tau_1)x(t+\tau_2)x^2(t+\tau_3). \tag{89}$$

In this case, since $x^2(t+\tau_3) \equiv 1$, we have the equivalence

$$\begin{aligned} L_x(t, \boldsymbol{\tau})_4 &= [x(t+\tau_1)x(t+\tau_2)][1] \\ &= L_x(t, \tau_1, \tau_2)_2. \end{aligned}$$

Thus, $L_x(t, \boldsymbol{\tau})_4$ contains nothing other than products of second-order sine waves, namely the sine waves in $x(t+\tau_1)x(t+\tau_2)$ multiplied by the second-order sine wave with frequency zero, $x^2(t+\tau_3) \equiv 1$. Consequently, $L_x(t, \boldsymbol{\tau})_4$ can contain no pure fourth-order sine waves. However, from the companion paper Part II, Section V, we have the CTCF

$$\begin{aligned} C_x^\beta(\boldsymbol{\tau})_4 = \frac{C_{a,4}}{T_0} \int_{-\infty}^{\infty} & p(t+\tau_1)p(t+\tau_2)p^2(t+\tau_3) \\ & \times e^{-i2\pi\beta t} dt\, e^{i2\pi\beta t_0}, \quad \beta = k/T_0 \end{aligned}$$

which is not identically zero. Another way to see this is to use (79) to compute the TCF

$$\begin{aligned} C_x(t, \boldsymbol{\tau})_4 = \hat{E}^{\{\alpha\}}&\{L_x(t, \boldsymbol{\tau})_4\} \\ & - \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\} \\ & \times \hat{E}^{\{\alpha\}}\{x(t+\tau_3)x(t+\tau_3)\} \\ & - \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_3)\} \\ & \times \hat{E}^{\{\alpha\}}\{x(t+\tau_2)x(t+\tau_3)\} \\ & - \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_3)\} \\ & \times \hat{E}^{\{\alpha\}}\{x(t+\tau_2)x(t+\tau_3)\} \\ = -2&\hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_3)\} \\ & \times \hat{E}^{\{\alpha\}}\{x(t+\tau_2)x(t+\tau_3)\}. \end{aligned}$$

It is, in fact, reasonable to expect a nonzero result for the fourth-order TCF of the set $\{x(t+\tau_1), x(t+\tau_2), x(t+\tau_3), x(t+\tau_3)\}$ since there is no proper subset of this set that is statistically independent of the remaining subset (cf. Section II–C).

However, since there is a repeated factor in (89), we can treat the lag product as the product of three, rather than four, factors $\{x(t+\tau_1), x(t+\tau_2), x^2(t+\tau_3)\}$. In this case, there *is* an independent subset among the *three* variables and

$$\begin{aligned} &\hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)x^2(t+\tau_3)\} \\ &\quad \equiv \hat{E}^{\{\alpha\}}\{x(t+\tau_1)x(t+\tau_2)\}\hat{E}^{\{\alpha\}}\{x^2(t+\tau_3)\}. \end{aligned}$$

The third-order cumulant for these three variables is zero. However, the presence of a repeated factor in the lag product is not in itself the cause of the failure of the pure-sine-waves interpretation of the TCF to hold. It is the fact that the repeated factor is degenerate:

$$x^2(t+\tau_3) \equiv 1.$$

This fact leads us to conjecture that the only degenerate time-series (those for which the correspondence between pure nth-order sine waves and the TCF is not the simple cumulant) are those piece-wise constant signals with all values equal to the mth roots of any fixed real number. In such a case,

$$x^m(t) \equiv \text{ constant}$$

and the same fix applies: the $(n-m+1)$th-order cumulant for the variables

$$\{x(t+\tau_1), x(t+\tau_2), \ldots, x(t+\tau_{n-m}), x^m(t+\tau_{n-m+1})\}$$

is zero, which is in agreement with the absence of any pure sine waves.

References

[1] B. G. Agee, S. V. Schell, and W. A. Gardner, "Spectral self-coherence restoral: A new approach to blind adaptive signal extraction using antenna arrays," *Proc. IEEE*, vol. 78, pp. 753–767, 1990.

[2] V. G. Alekseev, "Symmetry properties of high-order spectral densities of stationary and periodic-nonstationary stochastic processes," *Problemy Peredachi Informatsii*, vol. 23, no. 3, pp. 48–53, 1987.

[3] G. E. Andrews, *The Theory of Partitions*. Reading, MA: Addison-Wesley, 1976.

[4] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis*. New York: Wiley, 1980.

[5] G. Bilardi and S. Pupolin, "Spectral analysis of the powers of a PAM digital signal," *Alta Frequenza*, vol. LIII, no. 2, pp. 70–76, Mar. 1984.

[6] N. M. Blachman and S. H. Mousavinezhad, "The spectrum of the square of a synchronous random pulse train," *IEEE Trans. Commun.*, vol. 38, no. 1, pp. 13–17, Jan. 1990.

[7] A. Blanc-Lapierre and R. Fortet. *Theory of Random Functions*, Vol. I. New York: Gordon and Breach, 1965.

[8] ——, *Theory of Random Functions*, Vol. II. New York: Gordon and Breach, 1968.

[9] D. R. Brillinger, *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart and Winston, 1975.

[10] ——, "An introduction to polyspectra," *Ann. Math. Stat.*, vol. 36, pp. 1351–1374, 1965.

[11] D. R. Brillinger and M. Rosenblatt, "Asymptotic theory of estimates of k-th order spectra," in *Spectral Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967.

[12] ——, "Computation and interpretation of k-th order spectra," in *Spectral Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967.

[13] H.-H. Chiang and C. L. Nikias, "A new method for adaptive time delay estimation for non-Gaussian signals," *IEEE Trans. Signal Processing*, vol. 38, no. 2, pp. 209–219, Feb. 1990.

[14] C. Corduneanu, *Almost Periodic Functions*, 2nd ed. New York: Chelsea, 1989.

[15] E. A. Cornish and R. A. Fisher, "Moments and cumulants in the specification of distributions," *Revue Int. Stat. Inst.*, vol. 5, pp. 307–320, 1937.

[16] H. Cramer, *Random Variables and Probability Distributions*. London: Cambridge Univ. Press, 1937.

[17] ——, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press, 1946.

[18] A. V. Dandawate and G. B. Giannakis, "Polyspectral analysis of non-stationary signals: System identification, classification and ambiguity functions," in *Proc. Int. Signal Processing Workshop Higher Order Stat.* (Chamrousse, France), July 10–12, 1991, pp. 147–150.

[19] ——, "Ergodicity results for non-stationary signals: Cumulants, ambiguity functions and wavelets," in *Proc. 25th Annual Conf. Inform. Sci., Syst.* (Johns Hopkins Univ., Baltimore) Mar. 20–22, 1991, pp. 976–983.

[20] ——, "Nonparametric polyspectral analysis of AM signals and processes with missing observations," in *Proc. 26th Conf. Inform. Sci., Syst.* (Princeton Univ., Princeton, NJ), Mar. 18–20, 1992.

[21] ——, "Cyclic-cumulant based identification of almost periodically time-varying systems: Parametric methods," in *Proc. Int. Conf. ASSP 1992* (San Francisco), Mar. 23–26, 1992, pp. 229–232.

[22] ——, "Detection and classification of cyclostationary signals using cyclic-HOS: A unified approach," in *Proc. Soc. Photo-Opt. Instr. Engr., Adv. Signal Processing Algorithms, Architectures, Implementations* (San Diego, CA), July 1992.

[23] ——, "Nonparametric identification of linear (almost) periodically time-varying systems using cyclic polyspectra," in *Sixth Workshop Stat., Signal, Array Processing* (Victoria, BC, Canada), Oct. 7–9, 1992.

[24] W. B. Davenport and W. L. Root, *Random Signals and Noise*. New York: McGraw-Hill, 1958.

[25] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.

[26] R. F. Dwyer, "The extraction of range and Doppler from fourth-order spectra," in *Proc. 1990 ICASSP*, pp. 2403–2406, 1990.

[27] T. T. Fang, "Analysis of self-noise in a fourth-power clock regenerator," *IEEE Trans. Commun.*, vol. 39, no. 1, pp. 133–140, Jan. 1991.

[28] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vols. I and II, 2nd ed. New York: Wiley, 1957.

[29] A. Fisher, *The Mathematical Theory of Probabilities*. New York: Macmillan, 1923.

[30] R. A. Fisher, "Moments and product moments of sampling distributions," in *Proc. London Math. Soc.* (series 2), vol. 30, pp. 199–238, 1930.

[31] M. Fisz, *Probability Theory and Mathematical Statistics*. New York: Wiley, 1963.

[32] P. Forster and C. L. Nikias, "Bearing estimation in the bispectrum domain," *IEEE Trans. Signal Processing*, vol. 39, no. 9, pp. 1994–2006, Sept. 1991.

[33] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. New York: Macmillan, 1985 (2nd ed.: New York: McGraw-Hill, 1990).

[34] ——, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[35] ——, "Common pitfalls in the application of stationary process theory to time-sampled and modulated signals," *IEEE Trans. Commun.*, vol. 35, no. 5, pp. 529–534, May 1987.

[36] ——, "Signal interception: A unifying theoretical framework for feature detection," *IEEE Trans. Commun.*, vol. 36, no. 8, pp. 897–906, Aug. 1988.

[37] ——, "Spectral characterization of Nth-order cyclostationarity," in *Proc. Fifth IEEE/ASSP Workshop Spectrum Estimat., Modeling* (Rochester, NY), Oct. 10–12, 1990, pp. 251–255.

[38] ——, "Identification of systems with cyclostationary input and correlated input/output measurement noise," *IEEE Trans. Automat. Contr.*, vol. 35, no. 4, pp. 449–452, Apr. 1990.

[39] W. A. Gardner and C. M. Spooner, "higher order cyclostationarity, cyclic cumulants, and cyclic polyspectra," in *Proc. Int. Symp. Inform. Theory, Applicat.* (Honolulu), Nov. 27–30, 1990, pp. 355–358.

[40] W. A. Gardner and W. A. Brown, "Fraction-of-time probability for time-series that exhibit cyclostationarity," *Signal Processing*, vol. 23, no. 3, pp. 273–292, June 1991.

[41] W. A. Gardner and C. M. Spooner, "Signal interception: Performance advantages of cyclic-feature detectors," *IEEE Trans. Commun.*, vol. 40, no. 1, pp. 149–159, Jan. 1992.

[42] W. A. Gardner, "Cyclic Wiener filtering: Theory and method," *IEEE Trans. Commun.*, vol. 41, no. 1, pp. 151–163, Jan. 1993.

[43] W. A. Gardner and C. K. Chen, "Signal-selective time-difference-of-arrival estimation for passive location of manmade signal sources in highly corruptive environments, Parts I and II," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1168–1197, May 1992.

[44] W. A. Gardner and T. L. Archer, "Exploitation of cyclostationarity for identifying the Volterra kernels of nonlinear systems," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 535–542, Mar. 1993.

[45] W. A. Gardner, "An introduction to cyclostationary signals," Chap. 1 in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, Ed. New York: IEEE Press, 1994.

[46] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*. Reading, MA: Addison-Wesley, 1989.

[47] G. B. Giannakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1284–1296, July 1990.

[48] G. B. Giannakis and A. V. Dandawate, "Polyspectral analysis of nonstationary signals: Bases, consistency and HOS-WV," in *Proc. Int. Signal Processing Workshop Higher Order Stats.* (Chamrousse, France), July 10–12, 1991, pp. 167–170.

[49] ——, "Polyspectral analysis of (almost) cyclostationary signals: LPTV system identification and related applications," in *Proc. 25th Asilomar Conf. Signals, Syst., Comput.* (Pacific Grove, CA), Nov. 4–6, 1991, pp. 377–382.

[50] G. B. Giannakis and S. Shamsunder, "Non-Gaussian source localization via exploitation of higher order cyclostationarity," in *6th Workshop Stat. Signal, Array Processing* (Victoria, BC, Canada), Oct. 7–9, 1992, pp. 193–196.

[51] I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*. Philadelphia: Saunders, 1969.

[52] E. J. Hannan, *Time-Series Analysis*. New York: Wiley, 1960.

[53] M. J. Hinich, "Testing for Gaussianity and linearity of a stationary time series," *J. Time Series Anal.*, vol. 3, no. 3, pp. 169–176, 1982.

[54] M. J. Hinich and G. R. Wilson, "Time delay estimation using the cross bispectrum," *IEEE Trans. Signal Processing*, vol. 40, no. 1, pp. 106–113, Jan. 1992.

[55] P. J. Huber *et al.*, "Statistical methods for investigating phase relations in stationary stochastic processes," *IEEE Trans. Audio, Electroacoust.*, vol. 19, no. 1, pp. 78–86, Mar. 1971.

[56] H. L. Hurd, "Correlation theory of almost periodically correlated processes," *J. Multivariate Anal.*, vol. 37, pp. 24–45, 1991.

[57] G. M. Jenkins and D. G. Watts, *Spectral Analysis and its Applications*. San Francisco: Holden-Day, 1968.

[58] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*. New York: Academic, 1975.

[59] ——, *A Second Course in Stochastic Processes*. New York: Academic, 1981.

[60] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[61] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics: Vol. 1, Distribution Theory*. London: Griffin, 1958.

[62] Y. C. Kim and E. J. Powers, "Digital bispectral analysis of self-excited fluctuation spectra," *Physics of Fluids*, vol. 21, no. 8, pp. 1452–1453, Aug. 1978.

[63] Y. C. Kim, J. M. Beall, and E. J. Powers, "Bispectrum and nonlinear wave coupling," *Physics of Fluids*, vol. 23, no. 2, pp. 258–263, Feb. 1980.

[64] D. Kletter and H. Messer, "Suboptimal detection of non-Gaussian signals by third-order spectral analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 6, pp. 901–909, June 1990.

[65] L. H. Koopmans, *The Spectral Theory of Time Series*. New York: Academic, 1974.

[66] P. I. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, "The transmission of certain random functions through linear systems," in *Non-Linear Transformations of Stochastic Processes*, P. I. Kuznetsov, Ed. Oxford, UK: Pergamon Press, 1965.

[67] ——, "The transmission of random functions through non-linear systems," in *Non-Linear Transformations of Stochastic Processes*, P. I. Kuznetsov, Ed. Oxford, UK: Pergamon Press, 1965.

[68] E. L. Lehmann, *Testing Statistical Hypothesis*. New York: Wiley, 1959.

[69] V. P. Leonov and A. N. Shiryaev, "On a method of calculation of semi-invariants," *Theory of Probability and Its Applicat.*, vol. 4, no. 3, pp. 319–328, 1959.

[70] M. Loeve, *Probability Theory*. New York: Van Nostrand, 1955.

[71] ——, *Probability Theory II*, 4th ed. New York: Springer-Verlag, 1963.

[72] J. M. Mendel, "Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results, and some applications," *Proc. IEEE*, vol. 79, no. 3, pp. 277–305, Mar. 1991.

[73] D. Middleton, *Statistical Communication Theory*. New York: McGraw-Hill, 1960.

[74] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE*, vol. 75, no. 7, pp. 869–891, July 1987.

[75] C. L. Nikias, "higher order spectral analysis," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice Hall, 1991.

[76] G. Palm and T. Poggio, "The Volterra representation and the Wiener expansion: Validity and pitfalls," *SIAM J. Appl. Math.*, vol. 33, no. 2, pp. 195–216, Sept. 1977.

[77] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.

[78] ——, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[79] E. Parzen, *Stochastic Processes*. San Francisco: Holden-Day, 1962.

[80] ——, *Modern Probability Theory and Its Applications*. New York: Wiley, 1960.

[81] A. P. Petropulu and C. L. Nikias, "Signal reconstruction from the phase of the bispectrum," *IEEE Trans. Signal Processing*, vol. 40, no. 3, pp. 601–610, Mar. 1992.

[82] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Trans. Signal Processing*, vol. 39, no. 2, pp. 522–526, Feb. 1991.

[83] ——, "Direction finding algorithms based on high-order statistics," *IEEE Trans. Signal Processing*, vol. 39, no. 9, pp. 2016–2024, Sept. 1991.

[84] M. B. Priestley, *Spectral Analysis and Time Series*. London: Academic, 1981.

[85] ——, *Non-Linear and Non-Stationary Time Series Analysis*. London: Academic, 1988.

[86] V. S. Pugachev, *Theory of Random Functions*. Oxford, UK: Pergamon Press, 1965.

[87] M. R. Raghuveer and C. L. Nikias, "Bispectrum estimation: A parametric approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 4, Oct. 1985, pp. 1213–1230.

[88] M. R. Raghuveer, "Time-domain approaches to quadratic phase coupling estimation," *IEEE Trans. Automat. Cont.*, vol. 35, no. 1, pp. 48–56, Jan. 1990.

[89] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*. Cambridge, MA: Harvard Business School, 1961.

[90] M. Rosenblatt, *Stationary Sequences and Random Fields*. Boston: Birkhauser, 1985.

[91] S. V. Schell and W. A. Gardner, "High-resolution direction finding," in *Handbook of Statistics*, Vol. 10, N. K. Bose and C. R. Rao, Eds. New York: Elsevier Science Publishers, 1993.

[92] A. N. Shiryaev, "Some problems in the spectral theory of higher order moments—I," *Theory of Probability and Its Applicat.*, vol. 5, no. 3, pp. 265–284, 1960.

[93] C. M. Spooner and W. A. Gardner, "Estimation of cyclic polyspectra," *Proc. 25th Ann. Asilomar Conf. Signals, Syst., Comput.* (Pacific Grove, CA), Nov. 4–6, 1991, pp. 370–376.

[94] ——, "Performance evaluation of cyclic polyspectrum estimators," *Proc. 26th Ann. Asilomar Conf. Signals, Syst., Comput.* (Pacific Grove, CA), Oct. 26–28, 1992.

[95] ——, "An overview of the theory of higher order cyclostationarity," in *Proc. Workshop Nonstationary Stochastic Processes and Their Applicat.*, A. G. Miamee, Ed. Singapore: World Scientific, 1992, pp. 110–125.

[96] C. M. Spooner, "Theory and application of higher order cyclostationarity," Ph.D. Dissertation, Dept. Elec., Comput. Eng., Univ. of CA, Davis, June 1992.

[97] C. M. Spooner and W. A. Gardner, "Exploitation of higher order cyclostationarity for weak-signal detection and time-delay estimation," in *Proc. 6th Workshop Stat. Signal, Array Processing* (Victoria, BC, Canada), Oct. 7-9, 1992, pp. 197–201.

[98] ——, "The cumulant theory of cyclostationary time-series, Part II: Development and applications," *IEEE Trans. Signal Processing*, this issue.

[99] C. M. Spooner, "Higher order statistics for nonlinear processing of cyclostationary signals," in *Cyclostationarity in Communications and Signal Processing* (W. A. Gardner, Ed.). New York: IEEE, 1993, Ch. 2.

[100] R. L. Stratonovich, *Topics in the Theory of Random Noise*, Vol. I. New York: Gordon and Breach, 1963.

[101] T. N. Thiele, *The Theory of Observations (1903)*, reprinted in *Annals of Math. Stat.*, vol. 2, pp. 165–308, 1931.

[102] B. Kahn and G. E. Uhlenbeck. "On the theory of condensation," *Physica*, vol. 5, pp. 399–416, May 1938.

[103] H. Urkowitz, *Signal Theory and Random Processes*. Norwood, MA: Artech House, 1983.

[104] L. Wang and J. M. Mendel, "Cumulant-based parameter estimation using neural networks," in *Proc. 1990 ICASSP*, 1990, pp. 2369–2372.

[105] N. Wiener, "Generalized harmonic analysis," *Acta Mathematika*, vol. 55, pp. 117–258, 1930.

[106] S. S. Wilks, *Mathematical Statistics*. Princeton, NJ: Princeton Univ. Press, 1950.

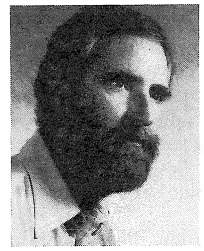
[107] J. Wishart, "A problem in combinatorial analysis giving the distribution of certain moment statistics," *Proc. London Math. Soc.*, (series 2), vol. 29, pp. 309–321, 1929.

[108] A. M. Yaglom, *Correlation Theory of Stationary and Related Random Functions II*. New York: Springer-Verlag, 1987.

[109] ——, *An Introduction to the Theory of Stationary Random Functions*. Englewood Cliffs, NJ: Prentice-Hall, 1962.

[110] W. Zhang and M. Raghuveer, "Nonparametric bispectrum-based time-delay estimators for multiple sensor data," *IEEE Trans. Signal Processing*, vol. 39, no. 3, pp. 770–774, Mar. 1991.

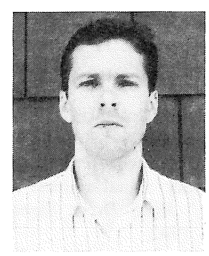
[111] G. D. Zivanovic, "Some aspects of the higher order cyclostationary theory," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 3497–3500, 1991.



William A. Gardner (S'64–M'67–SM'84–F'91) was born in Palo Alto, CA, in 1942. He received the M.S. degree from Stanford University in 1967, and the Ph.D. degree from the University of Massachusetts, Amherst, in 1972, both in electrical engineering.

He was a Member of the Technical Staff at Bell Laboratories in Massachusetts from 1967 to 1969. He has been a faculty member at the University of California at Davis since 1972, where he is Professor of Electrical and Computer Engineering. Since 1982, he has also been President of the engineering consulting firm Statistical Signal Processing, Inc., Yountville, CA. His research interests are in the general area of statistical signal processing, with primary emphasis on the theories of time-series analysis, stochastic processes, and signal detection and estimation and applications to communications and signals intelligence. Dr. Gardner is the author of: *Introduction to Random Processes with Applications to Signals and Systems,* Macmillan, New York, 1985 (2nd ed., McGraw-Hill, New York, 1990); *The Random Processes Tutor: A Comprehensive Solutions Manual for Independent Study*, McGraw-Hill, New York, 1990; and *Statistical Spectral Analysis: A Nonprobabilistic Theory,* Prentice-Hall, Englewood Cliffs, NJ, 1987. He is also Editor of *Cyclostationarity in Communications and Signal Processing,* IEEE Press, New York, 1994. He holds several patents and is the author of numerous research journal papers.

Dr. Gardner received the Best Paper of the Year Award from the European Association for Signal Processing in 1986 for the paper entitled "The Spectral Correlation Theory of Cyclostationary Signals," the 1987 Distinguished Engineering Alumnus Award from the University of Massachusetts, and the Stephen O. Rice Prize Paper Award in the Field of Communication Theory from the IEEE Communications Society in 1988 for the paper entitled "Signal Interception: A Unifying Theoretical Framework for Feature Detection." He organized and chaired the NSF/ONR/ARO/AFOSR-sponsored Workshop on Cyclostationary Signals in 1992. He is a member of the American Association for the Advancement of Science as well as the European Association for Signal Processing, and is a member of the honor societies Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and Alpha Gamma Sigma. He was elected Fellow of the IEEE in 1991 "for contributions to the development of time-series analysis and stochastic processes with applications to statistical signal processing and communication, and for contributions to engineering education."



Chad M. Spooner (S'88–M'92) was born in Des Moines, IA, in 1963. From 1981 to 1984, he attended Santa Rosa Junior College, Santa Rosa, CA, and was awarded the A.S. degree with high honors in 1984. He received the B.S. degree from the University of California at Berkeley in 1986, and the M.S. and Ph.D. degrees from the University of California at Davis in 1988 and 1992, respectively, all in electrical engineering.

He was a Teaching Assistant in the Department of Electrical and Computer Engineering at U.C. Davis from 1986 to 1988. From 1988 to 1992 he was a Research Assistant at U.C. Davis and a Consultant to Statistical Signal Processing, Inc. From 1992 to 1994 he was a Postdoctoral Researcher and Instructor at U.C. Davis, and is coauthor (with Dr. Stephan Schell) of a unique commercially available cyclic spectral analysis software package. He is currently with Mission Research Corporation in Monterey, CA. His research interests include statistical signal processing, weak-signal detection and parameter estimation, modulation classification, and the theory of the higher order statistics of cyclostationary signals.

Dr. Spooner is a member of Sigma Xi and The Mathematical Association of America. In 1992, he received the Award for Excellence in Ph.D. Research from the U.C. Davis Chapter of Sigma Xi.