QUARTERLY OF APPLIED MATHEMATICS

VOLUME XXXVII July 1979 No. 2

NONLINEAR ANALYSIS WITH AN ARBITRARY STIMULUS ENSEMBLE*

BY

JONATHAN D. VICTOR AND BRUCE W. KNIGHT

The Rockefeller University

Abstract. A family of Wiener-type methods is discussed in a general context. These methods share the concept of expansion of an unknown transducer as an orthogonal series. The terms of the series are drawn from a hierarchy of subspaces of transducers that are orthogonal with respect to a particular stimulus ensemble. Choices of specific stochastic ensembles lead to previously described analytical methods, including the classical one of Wiener.

It is proposed that a sum of incommensurate or nearly incommensurate sinusoids forms a signal that leads to a useful orthogonal expansion. The family of orthogonal subspaces are presented explicitly. Projection of an unknown transducer into an orthogonal subspace amounts to isolation of Fourier components of the output of the unknown transducer at certain harmonics and combination frequencies of the input frequencies. Practical advantages of this technique include i) the ease of computation of the higher-order kernels, and ii) the opportunity for digital filtering of the response, which enhances the signal-to-noise ratio.

Finally, it is shown that the kernels obtained using a sum-of-sinusoids signal approach the Fourier transforms of the Wiener kernels as the number of sinusoids grows without bound. Thus, the sum-of-sinusoids technique retains a major theoretical advantage of the Wiener white-noise method: the kernels of simple model transducers have simple analytic forms.

Introduction. The practical application of functional analysis to the understanding of biological transducers is a rapidly expanding endeavor at the present time. This development stems naturally from the evident need to understand the function of biological organs in quantitative detail, from the great recent growth of laboratory techniques and of instrumentation which has arisen in response to that need, and from the recent availability of powerful computers which may be conveniently programmed to process the large arrays of numbers which the laboratory furnishes.

The current situation offers an unusual challenge to the theorist. Here we find not the familiar task of deducing the consequences of a well-defined set of dynamical laws, but rather the inverse challenge of organizing observed consequences in such a way that the underlying laws themselves may be deduced. The theorist's task is to propose experimental

* Received April 20, 1978. This research was supported, in part, by research grants from the National Eye Institute.

procedures which, in conjunction with powerful computational facilities, will lead to an enlightening description of the dynamical laws which govern the response of a given biological transducer.

The discipline which is sometimes called "nonlinear system identification in biology" and which is the subject of this paper, has a history which might be marked by four outstanding innovations:

1. Toward the end of the nineteenth century Volterra commenced his systematic exploration of functionals [56], or of functions which depend on a continuous set of values of other functions. Volterra observed that various familiar notions which apply to functions of several discrete variables survive a "passage from the discrete to the continuous". Thus a Taylor series in several variables passes naturally to a series of multiple integrals: a constant plus a single weighted integral over a given function plus a double integral over the products of all pairs of values of that given function, and so on. The whole series is the "Volterra series," and the weighting functions within the integrands—the analogues of Taylor coefficients—are the "Volterra kernels" of the various orders. It has long been common understanding that the present output of a transducer is a functional of its past input in Volterra's sense, and that a large class of nonlinear transducers may thus be characterized in terms of the Volterra series, which series takes the form of a sum of homogeneous functionals of the past input.

2. During and shortly after the Second World War, Wiener [58] explored the notion that a nonlinear transducer might be synthesized by using observations of its response to "Gaussian white noise", which noise eventually approximates, with arbitrary accuracy, any given input stimulus of fixed length. Gaussian white noise represents the "passage to the continuous" of a discrete set of variables which have a joint Gaussian probability distribution, and Wiener noted that just as such a distribution gives to a Taylor series a natural reorganization in terms of orthogonal Hermite polynomials, so likewise the passage to the continuous yields a reorganized description of a transducer in terms of component "Hermite polynomial" functionals, which are orthogonal transducers in the sense that they give zero cross-correlation in response to white noise. Just as series of orthogonal functions converge to a far larger set of functions than do Taylor series, so Wiener's series of orthogonal functionals encompassed not only transducers expressible in terms of Volterra series, but also included the full gamut of transducers which one might reasonably care to consider. The weighting functions in his orthognal expansion which characterize the individual transducer are now known as Wiener kernels, and Wiener advanced a clever analogue-electronic scheme whereby these kernels for an actual transducer might be evaluated.

3. In 1965, which was a time when computers with very large number-handling capacity were becoming available, an algorithm was advanced by Lee and Schetzen [29] for the direct evaluation of the Wiener kernels, by means of numerically cross-correlating the measured transducer response with products of Hermite polynomials of the white-noise input chosen at various time lags. Lee and Schetzen thus furnished the practical means for constructing the Wiener kernels of a given transducer.

4. In 1972 Marmarelis and Naka [32] (see also Møller [36]) combined the algorithm of Lee and Schetzen, and the use of a large modern computer, with very advanced experimental techniques of cellular electrophysiology; they thereby recorded the responses to a white-noise-modulated light stimulus of individual visual nerve-cells in the retina of the catfish. In this way they characterized the response of visual cells in terms of Wiener kernels. Their pioneering work has been followed by a great influx of scientific effort into this general area.

Aside from the procedures associated with the names of Volterra and Wiener, the use of theoretical systems analysis in conjunction with inputs of specially chosen functional forms to characterize biological transducers has a substantial recent history. The items cited below are only a sparse sample. Detection of sinusoidal flicker superimposed on steady light, by human subjects, was used by Ives [24] in an essentially system-analytic way before the general theory of linear transducer response to sinusoidal input was put in its modern form by Bode [6]. The classical sine-flicker threshold studies of deLange [12] bring together the work of Ives and Bode. An early application of the Bode methodology to electrophysiology data was that of Pringle and Wilson [42], who analyzed the modulated firing rate of a cockroach-leg mechanoreceptor cell in response to sinusoidal input derived from the vibrations of a weighted hacksaw-blade. The bulk electrical response of the wolf-spider's eye to sinusoidally modulated light was similarly analyzed by DeVoe [13]. Hughes and Maffei [23] analyzed the modulation in firing rate of gangion cells in the cat's retina in response to sine-flickering light. Spekreijse [48, 49] has used a battery of stimuli derived from both Gaussian noise and sine waves, in conjunction with very detailed theoretical system-analytic tools, to study both the linear and nonlinear response of the vertebrate visual system. An early system-analytic study of the crayfish stretch-receptor by Borsellino, Poppele, and Terzuolo [7] has led to a comparative study by Fohlmeister, Poppele and Purple [16, 17] of nerve-impulse generating transducers. The linear frequency response of the lateral eye of the horseshoe crab *Limulus* was given early investigation by Pinter [40], by Biederman-Thorson and Thorson [5, 52], and by Dodge, Knight and Toyoda [14, 26, 43]. Over the past decade a very detailed and predictively accurate linear system-analytic model has evolved for the visual neurophysiology of *Limulus* [8, 9, 44].

There now have been a substantial number of experimental forays unto applying methods of nonlinear system identification, in the general spirit of Wiener's proposal, to biological transducers [15, 19, 20, 28, 30, 32, 33, 35–37, 45, 46, 53, 55, 57]. Concurrently, the theoretical state of the subject has continued to advance. It was recognized before 1963 by Barrett [2] that Wiener's white-noise approach is just one member of a wide class of analytic procedures, each based on its own particular ensemble of input test signals. Recently this idea has been elaborated in some detail [25, 39, 60]. Specific procedures have also been advanced which adapt Wiener's orthogonal functional approach to an input ensemble which is a Poisson point process [27] and to an input ensemble which is a multilevel discrete noise [34]. Here we will discuss these theories as a class, hoping to gain insight into their common structure. The exposition below will begin with the presentation of a general theoretical framework which is easily specialized to recover any one of the various Wiener-like procedures which have been advanced to date.

All the Wiener-type identification procedures proposed to date manifest two common drawbacks when they are applied in the laboratory to biological transducers. The first drawback stems from the fact that in the laboratory the input test signal represents only a sample of the stochastic test ensemble from which it is drawn. To the degree to which that sample is statistically atypical of the whole input ensemble, the identification algorithm (which is based on statistical properties of the ensemble and not on the sample) may yield misleading results. The second drawback is that biological transducers typically generate substantial autonomous "noise" of their own, which noise the identification algorithm misinterprets as the response to input. The straightforward cure for both of these difficulties is to collect enough data to ensure adequate statistical accuracy; but here experimental biology does not cooperate, as it is often impractical to maintain the living transducer in a steady vital state for the time span which statistical accuracy demands.

The primary goal of the present paper is to introduce a family of test signals which avoids the first of the two drawbacks mentioned above and which much reduces the second. The new test signal is the sum of sinusoids, which in recent laboratory use has proven very effective in the analysis of linear [8], slightly nonlinear [59] and very nonlinear [46, 53, 55] neutral transducers. Although a sum of sinusoids has been used previously for the analysis of nonlinear systems [3, 31, 49, 50, 61], the use of this signal as the core of a Wiener-type procedure is novel.

The sum-of-sinusoids signal is deterministic rather than stochastic, which obviates the first drawback above which was the hazard of drawing an atypical input from a stochastic ensemble. Moreover, a nonlinear transducer will respond to a sum of sinusoids with only a discrete set of output frequencies at the harmonics and combination frequencies of the input. Thus the response may be numerically processed with sharply tuned digital filters set at these known frequencies, and the great majority of the noise which the transducer generates autonomously will be rejected because it appears at other frequencies. Thus the second drawback of the prior Wiener-type approaches is greatly reduced.

The sum-of-sinusoids procedure has further advantages: it proves simple to carry out preliminary online computer analysis, so that interpretations and judgements may be made in the laboratory, and acted upon while the experiment progresses.

Finally, the sum-of-sinusoids procedure has one very important theoretical advantage: for a wide class of nonlinear systems the kernels obtained in this way are closely related, in a clear way, to the kernels of Wiener's original procedure. In consequence, the kernels obtained from simple model systems with the sum-of-sinusoids technique have simple analytic forms (a virtue not shared by other implemented generalizations of Wiener's original procedure). This feature of the method allows the experimental data to suggest simple nonlinear transducers that may be appropriate for initial attempts at model-building.

In view of the length of the exposition to follow, we summarize its organization: in Secs. 1 and 2, we review the class of Wiener-like methods of nonlinear analysis. In Sec. 3, we specialize this general structure to the particular case in which the test signal is a sum of incommensurate sinusoids. The correspondence between the sum-of-sinusoids technique and the original method of Wiener is shown in Sec. 4. Sec. 5 shows how the results of Sec. 3 and 4 can be applied in a practical laboratory procedure in which the frequencies of the components of the sinusoidal sum are commensurate.

1. Wiener-type theories of nonlinear analysis. The Wiener-type theories of nonlinear analysis basically consist of a procedure for the expression of the input-output relation of an arbitrary transducer as a series of orthogonal functionals. The partial sums of this orthogonal series are themselves transducers. Each partial sum of the series is that transducer within a specified category that is the best approximation to the transducer under study. The specified category of transducers enlarges as the partial sums extend. Thus, the sequence of partial sums forms a sequence of improving approximations to the system under study.

The category of transducers from which the nth approximating transducer is drawn is the same in all theories considered here. It is the category of functions whose Volterra expansions [56] have terms of order no greater than n. However, the criterion of "best approximation" depends on the input signal used, and thus differs from theory to theory. The notion of "best approximation" relies on a notion of "distance" between transducers. The distance between two transducers is asserted to be the root-mean-squared discrepancy between their output, averaged over the entire stimulus ensemble. This distance induces an inner product in the space of transducers which is fundamental for the development of orthogonal expansions.

The distance, and thus the inner product, depend strongly on the choice of input signal. A Gaussian white noise leads to Wiener's original theory [58]; other choices result in the variations referred to above. The dependence of the geometry of the space of transducers on the choice of input signal is the primary source of the differences among the Wiener-type theories of nonlinear analysis. It reflects the fact that two nonlinear transducers may have different degrees of apparent dissimilarity when tested with different signals.

In the orthogonal expansion of a given transducer, each term of the series must be orthogonal to all transducers in the categories from which the previous terms were drawn. This condition is not satisfied by a series in which each term is a homogeneous functional of a different order, such as a Volterra series. Rather, the terms of the orthogonal series are inhomogeneous functionals that may be obtained by rearranging a Volterra series. A prescription for this rearrangement forms one principal element of a Wiener-type theory. In Sec. 2, we will show how this rearrangement may always be accomplished by a Gram-Schmidt procedure.

The second principal element of Wiener-type theories is a method for the determination of the approximating functionals from experimental data. Wiener proposed an analogue method [58] in his original theory. Lee and Schetzen [29] introduced a computational method based on the cross-correlations between the input signal and the transducer's response, which is suitable for Gaussian input signals. French [18] has suggested that those calculations may be performed more efficiently in the frequency domain. Krausz [27] and Marmarelis [34] propose procedures appropriate for their respective discrete non-Gaussian inputs. Most of these methods [27, 29, 34] are special cases of the methods of Klein and Yasui [25], which are applicable to Gaussian, white, and discrete input signals. For another kind of input signal, the sum-of-sinusoids signal, the determination of the approximating functionals is especially simple. This is shown in Sec. 3.

2. The orthogonal-series representation. In this section we review the general theory of the orthogonal decomposition of functionals [2] in order to provide a framework for the ensuing material.

The vector space of distinguishable transducers. The fundamental object of interest is the single-input, single-output transducer. Such transducers form a vector space in a natural way: addition in the vector space corresponds to parallel composition of the transducers. That is, if μ and ν are two transducers, then the response of the transducer $\mu + \nu$ to a given signal is just the pointwise sum of the responses of μ and ν to the signal s:

$$(\mu + \nu)(s)(t) = \mu(s)(t) + \nu(s)(t)\cdot$$

A natural scalar multiplication may also be defined in a pointwise fashion as a simple change in gain:

$$(\alpha\mu)(s)(t) = \mu(s)(t)\cdot\alpha.$$

We restrict consideration to "stationary" transducers, which satisfy

$$\mu(s_\tau)(t) = \mu(s)(t + \tau) \tag{1}$$

for

$$s_\tau(t) \equiv s(t + \tau)\ \cdot$$

The choice of input ensemble leads to a natural notion of "distance" between two transducers. Let Ω be a probability space from which the input signals s are drawn. We assume that Ω is also stationary: the weight of an input signal s and the weights of its time-shifts s_τ are identical. Then we define the squared distance between two transducers μ and ν in terms of their responses at time zero:

$$\|\mu - \nu\|^2 = <[(\mu - \nu)(s)(0)]^2>_\Omega$$

where $< >_\Omega$ indicates an average taken over the stimulus ensemble. This distance corresponds to the bilinear form

$$(\mu, \nu) = <\mu(s)(0) \cdot \nu(s)(0)>_\Omega \cdot \qquad (2)$$

Let $\hat{\mathfrak{M}}$ denote the vector space of stationary transducers for which the expression (μ, μ) (or $\|\mu\|^2$) exists and is finite. The expression (2) is not an inner product on $\hat{\mathfrak{M}}$ only because we might have $(\mu, \mu) = 0$ for some nonzero transducers μ in $\hat{\mathfrak{M}}$. But the transducers μ which satisfy $(\mu, \mu) = 0$ are those transducers whose response is zero almost everywhere in the stimulus ensemble Ω. Therefore, these transducers form a subspace $\hat{\mathfrak{M}}_0$ of $\hat{\mathfrak{M}}$. $\hat{\mathfrak{M}}_0$ consists of those transducers that cannot be distinguished from the zero transducer by an identification scheme based on the ensemble Ω.

This suggests that we direct our attention to the quotient space $\mathfrak{M} \equiv \hat{\mathfrak{M}}/\hat{\mathfrak{M}}_0$. An element of $\mathfrak{M}$ consists of a transducer μ of $\hat{\mathfrak{M}}$ together with all other transducers μ' of $\hat{\mathfrak{M}}$ that cannot be distinguished from μ with the testing ensemble Ω. That is, elements of $\mathfrak{M}$ are just those classes of transducers that can be distinguished from each other by the input ensemble Ω. We focus on the space $\mathfrak{M}$, and henceforth will use the symbols $\mu, \nu, \cdots$ to indicate the images in $\mathfrak{M}$ of the corresponding transducers in $\hat{\mathfrak{M}}$. Thus, $\mu, \nu, \cdots$ now represent classes of distinguishable transducers.

The form (μ, ν) of Eq. (2) is an inner product in $\mathfrak{M}$. It yields the fully positive-definite distance $\|\mu\| = (\mu, \mu)^{1/2}$ on $\mathfrak{M}$. The space $\mathfrak{M}$ may now be completed in the usual way to form a Hilbert space.

We remark that the three paragraphs above were needed only for reasons of tidiness and rigor. The "almost non-responding" transducers will manifest themselves only by perverse choice of transducer or of input ensemble. Consequently, we will shorten the term "transducer class" to "transducer" in what follows.

A sequence of orthogonal subspaces. Let M_j denote the subspace in $\mathfrak{M}$ of all distinguishable transducers that are functionals homogeneous of order j in the stimulus $s(t)$. An element μ_j of M_j is determined by a symmetric function $[\mu_j]$ of j prior times (u_j: $(\mathbb{R}^+)^j \to \mathbb{R}$). The function $[\mu_j]$ $(\tau_1, \tau_2, \cdots \tau_j)$ indicates how the input signal levels at j previous times interact to produce the current output. That is,

$$\mu_j(s)(t) = \iint \cdots \int [\mu_j](\tau_1, \tau_2, \cdots, \tau_j)\cdot s(t - \tau_1) \cdot s(t - \tau_2)\cdot \cdots \cdot s(t - \tau_j) d\tau_1 d\tau_2 \cdots d\tau_j, \qquad j \geq 1 \qquad (3)$$

and $\mu_0(s)(t) = [\mu_0]$, a constant.

The function $[\mu_j]$ may be a generalized function (cf. [38]). Thus, transductions such as

$$D_{j,\mathbf{T}}(s)(t) = \prod_{k=1}^{j} s(t - T_k) \qquad (4)$$

are included in (3), by choosing

$$[D_{j,\mathbf{T}}](\tau_1, \tau_2, \cdots, \tau_j) = c \prod_{k=1}^{j} \delta(\tau_k - T_k)$$

With the above notation, the Volterra expansion [56] of any transducer μ which has such an expansion may be written

$$\mu = \sum_{j=0}^{\infty} \mu_j\,, \qquad \mu_j \text{ in } M_j\,.$$

This representation exists only for the "analytic" elements in $\mathfrak{M}$ [21]. Non-analytic transductions such as rectifiers [22, 48, 49, 55] and fractional power laws [15, 47] are often contemplated as models in biology. The ability to include such transductions is a principal advantage of the Wiener approach.

The Wiener procedure extends the usefulness of the Volterra series through a term-by-term reorganization of that series into a new series of functionals. Each functional is related to a homogeneous counterpart in the Volterra series but is orthogonalized to all lower terms with respect to the inner product which is furnished by the input ensemble. This orthogonalization strengthens the convergence of the Wiener series in the same way that infinite series of orthogonal polynomials (Legendre, Hermite, Laguerre, for example) successfully converge to a set of functions far larger than the set expressed by convergent Taylor series [51].

In formal terms, orthogonal subspaces $\{K_j\}$ are constructed with the properties

$$\bigoplus_{j=0}^{n} K_j = \bigoplus_{j=0}^{n} M_j; \qquad K_n \perp K_j \quad \text{for} \quad j < n. \tag{5}$$

If the subspaces $\{K_j\}$ span the full space $\mathfrak{M}$, then every transduction μ will have a representation as an orthogonal series of "orthogonal transducers"

$$\mu = \sum_{j=0}^{\infty} \kappa_j\,, \qquad \kappa_j \quad \text{in} \quad K_j\,. \tag{6}$$

This program may be implemented explicitly by a Gram-Schmidt procedure. Suppose we have a Volterra functional μ_n which is homogeneous of nth order and that we wish to find that part of it, ν_n, which is orthogonal to all polynomial functionals of all lower orders. Suppose further that (by induction) we have established in each subspace K_j of lower order j a complete orthonormal basis $\{\kappa_{j,\beta}\}$ which are orthogonal in turn to all polynomial functionals of order lower than j. Then

$$\nu_n = \mu_n - \sum_{j=0}^{n-1} \sum_{\beta} (\mu_n\,, \kappa_{j,\beta})\kappa_{j,\beta}$$

manifestly is orthogonal to all the lower order subspaces. It also is invariant under orthogonal transformations among the $\{\kappa_{j,\beta}\}$, and therefore is independent of how the bases $\{\kappa_{j,\beta}\}$ were chosen. We proceed in the same way with a complete bases set $\{\mu_{n,\alpha}\}$ of homogeneous nth-order functionals. This yields a transformation $\phi_n: \mu_{n,\alpha} \to \nu_{n,\alpha}$, where the $\{\nu_{n,\alpha}\}$ satisfy

$$(\nu_{n,\alpha}\,, \kappa_{j,\beta}) = 0 \quad \text{for all} \quad j < n. \tag{7}$$

Extending the transformation $\phi_n: \mu_{n,\alpha} \to \nu_{n,\alpha}$ by linearity to the subspace M_n defines the counterpart subspace $K_n \equiv \phi_n(M_n)$ of orthogonal functionals. Each image function $\nu_{n,\alpha}$ has the homogeneous functional $\mu_{n,\alpha}$ as its leading term, which demonstrates the relation (5). The Gram-Schmidt reduction of the set $\{\nu_{n,\alpha}\}$ to an orthonormal basis $\{\kappa_{n,\beta}\}$ completes the induction. (The process may be started by assuming $\kappa_0 = 1$.)

Once the orthonormal bases have been constructed for the orthogonal subspaces, an explicit prescription is on hand for the presentation of an arbitrary transduction as an orthogonal series in the form of Eq. (5). The prescription is

$$\mu = \sum_{j=0}^{\infty} \left\{ \sum_{\beta} (\mu, \kappa_{j,\beta})\kappa_{j,\beta} \right\}$$

where the bracketed sum on β gives

$$\kappa_j = \sum_{\beta} (\mu, \kappa_{j,\beta})\kappa_{j,\beta} \tag{8}$$

for the vector κ_j of Eq. (6). Again, that bracketed term does not depend on the choice of basis. Eq. (8) expresses the jth element of the orthogonal expansion (6) in terms of the experimentally observable cross-correlations $(\mu, \kappa_{j,\beta})$ between the transducer and a standard set of transductions $\{\kappa_{j,\beta}\}$.

The orthogonal expansion (6) may be separated into two sums at the nth term:

$$\mu = \sum_{j=0}^{n} \kappa_j + \sum_{j=n+1}^{\infty} \kappa_j \, .$$

Clearly the two sums are mutually orthogonal. We may regard the first sum as an estimate of μ and the second sum as an error term not included in that estimate. The estimate vector and the error vector are at right angles, whence this estimate is the *best possible* estimate which can be chosen from within the nth-order subspace: its error vector has no component parallel to the estimate vector itself, and thus is smaller than the error vector for any other choice within the nth-order subspace, in the least-squares sense.

Convergence of the series (6) is guaranteed, since $\sum_{j=0}^{\infty} \|\kappa_j\|^2 \leq \|\mu\|^2$. We note that this is convergence as functionals in $\mathfrak{M}$, not pointwise in Ω. Whether the representation (6) exists for all elements μ of $\mathfrak{M}$ (that is, whether equality is attained in the previous relation) depends on whether the subspaces $\{K_j\}$ span $\mathfrak{M}$. When Ω is Gaussian white noise, conditions sufficient for spanning M are given in [10]. Conditions sufficient for the existence of the representation (6) are considered for other "white" noises in [39]; we consider this question for the sum-of-sinusoids signal in Sec. 3.

For particular ensembles Ω, the procedure indicated by Eq. (8) can be made entirely explicit. If Ω is Gaussian white noise as originally proposed by Wiener, a choice of $D_{j,\mathrm{T}}$ (Eq. (4)) for $\mu_{j,\alpha}$ leads to the Lee and Schetzen procedure [29] which explicitly evaluates the Wiener functional series. Choosing the Fourier transforms of $D_{j,\mathrm{T}}$ as basis elements results in French's modification [18]. Other choices for bases of M_j lead to Wiener's original analogue procedure [58] and the more general Cameron-Martin expansion [10]. For specific input ensembles that are white or discrete but not Gaussian, the procedure of Eq. (8) has been worked out by Marmarelis [34] and Krausz [27]. Klein and Yasui [25] and Palm and Poggio [39] have considered the general discrete non-Gaussian case. When Ω is Gaussian but not white, there is also a substantial simplification of (8), as is shown by Yasui [60] and Lee and Schetzen [29]. In Sec. 3, we specialize Eq. (8) for the sum-of-sinusoids signal.

3. The sum-of-sinusoids signal.

The considerations of Sec. 1 suggested the use of a sum of a large number of sinusoids as a test signal for the analysis of nonlinear transducers. In this approach, the transducer's response is characterized by frequencies that are equal to sums, differences, and higher harmonics of frequencies in the input signal. A test signal, which only probes the response at discrete frequencies, may overlook features which are "local in frequency space," and may be less than ideal for the "synthesis" problem. But this limitation from the "synthesis" point of view is balanced by an advantage from the "analysis" standpoint: because responses are measured at only a discrete mesh of frequencies, intrinsic noise of the transducer may be removed by digital filtering techniques. This advantage is especially important in neurophysiological applications [53, 54], where significant intrinsic noise is commonplace. Other practical advantages will be discussed below.

The inner product. Consider an input ensemble Ω whose members are each a sum of Q sinusoids,

$$s_{\psi}(t) = \sum_{r=1}^{Q} a_r \cos(\alpha_r t + \psi_r). \tag{9}$$

The coefficients a_r are fixed and the initial phases $\psi \equiv (\psi_1, \cdots, \psi_Q)$ are arbitrary. To begin with, we stipulate that the frequencies be a fixed sequence of increasing incommensurate positive numbers. For incommensurate frequencies α_r, a well-chosen time translation t_0 will bring one member of the ensemble $s_{\psi}(t + t_0)$ arbitrarily close to another, $s_{\psi'}(t)$, as all the $(\alpha_r t_0 + \psi_r) \bmod 2\pi$ simultaneously may be brought arbitrarily close to any corresponding ψ_r' by searching a long enough time span for the best t_0. However, no choice of t_0 leads to perfect registration in general. We complete the definition of our ensemble by postulating that s_{ψ} with distinct ψ appear with equal weight in the ensemble Ω. According to Eq. (9), the parameterization ψ of the members of the ensemble Ω forms a Q-dimensional torus $T^Q \equiv [0, 2\pi) \times [0, 2\pi) \times \cdots \times [0, 2\pi)$.

Because all test signals s_{ψ} are weighted equally in Ω, the inner product (2) becomes

$$(\mu, \nu) = \langle \mu(s_{\psi})(0) \cdot \nu(s_{\psi})(0) \rangle_{\psi \text{ in } T^Q} \equiv \frac{1}{(2\pi)^Q} \int \mu(s_{\psi})(0) \cdot \nu(s_{\psi})(0) d\psi. \tag{10}$$

For a large class of transducers, the average (10) over phases may be replaced by the time average in response to a single input.

The aim of the rest of this subsection is to give rigorous conditions sufficient for the ensemble average and the time average to be equal. That is, we give conditions sufficient for the relation

$$(\mu, \nu) = \langle \mu(s_0)(\tau) \cdot \nu(s_0)(\tau) \rangle_\tau \tag{11}$$

to be valid, where the subscript "0" stands for the vector of zero phases $\psi_1 = 0$, $\psi_2 = 0, \cdots$.

To prove Eq. (11), two technical conditions are required of the transducers μ and ν:

C1. $|\mu(s_{\psi})(0)|$ is less than some bound $B(\mu)$;

C2. μ is a continuous functional in the W^2 sense. That is, for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$D_{W^2}(s_{\psi} - s_{\psi'}) < \delta \text{ implies } D_{W^2}(\mu(s_{\psi}) - \mu(s_{\psi'})) < \epsilon.$$

The distance D_{W^2} is a mean-squared distance on the real line [4, pp. 71–72]:

$$D_{W^2}(f) \equiv \lim_{l \to \infty} \ \underset{t \text{ in}(-\infty,\infty)}{\text{u.b.}} \left\{ \frac{1}{l} \int_t^{t+l} (f(t'))^2 dt' \right\}^{1/2}. \tag{12}$$

The topology of the signals s_{ψ} in the D_{W^2}-norm is equivalent to the topology of the phase vectors ψ in the euclidean norm on the torus of phases T^Q, because

$$D_{W^2}(s_{\psi} - s_{\psi'}) = \left\{ \sum_{r=1}^{Q} 2a_r^2 \sin^2\left(\frac{\psi_r - \psi_r'}{2}\right) \right\}^{1/2}.$$

Therefore, the signals

$$s_\tau(t) \equiv \sum_{r=1}^{Q} a_r \cos(\alpha_r(t+\tau))$$

form a family that is dense in the ensemble of signals Ω, since the phase vectors $\psi = \tau\alpha$ are dense in T^Q.

By stationarity (1), the right-hand side of Eq. (11) is equal to $\langle \mu(s_\tau)(0)\cdot\nu(s_\tau)(0)\rangle_\tau$, an average over the dense subset $\{s_\tau\}$ of Ω. We now show that the smoothness conditions C1 and C2 allow us to replace the full ensemble average (10) by an average over the dense subset $\{s_\tau\}$. Define $s_{\psi,\tau}(t) \equiv s_\psi(t+\tau)$. Sweeping τ over a range of values is equivalent to sweeping ψ along a one-dimensional trajectory on the torus T^Q. Moreover, the entire torus may be regarded as the cluster of all such distinct trajectories. Formally, the phase vectors ψ in T^Q fall into equivalence classes based on whether the corresponding sum-of-sinusoids signals are identical, except for a translation in time. That is, ψ and ψ' are equivalent if there exists a time-translation τ such that $s_{\psi,\tau} = s_{\psi',0}$. Let T_0^Q be a subset of T^Q which contains exactly one member of each such equivalence class. Then

$$\langle \mu(s_\psi)(0)\cdot\nu(s_\psi)(0)\rangle_{\psi \text{ in } T^Q} = \langle \mu(s_{\psi,\tau})(0)\cdot\nu(s_{\psi,\tau})(0)\rangle_{\tau,\psi \text{ in } T_0^Q} =$$
$$\langle \mu(s_\psi)(\tau)\cdot\nu(s_\psi)(\tau)\rangle_{\tau,\psi \text{ in } T^Q},$$

where we have used the stationarity condition (1) in the last step.

It remains to be shown that conditions C1 and C2 permit us to replace the last average over T_0^Q by simple evaluation at $\psi = 0$. Define $\rho(s_\psi)(\tau) \equiv \mu(s_\psi)(\tau)\cdot\nu(s_\psi)(\tau)$. The Schwartz inequality and condition C1 imply

$$D_{W^2}(\rho(s_\psi) - \rho(s_0)) \leq B(\nu)\cdot D_{W^2}(\mu(s_\psi) - \mu(s_0)) + B(\mu)\cdot D_{W^2}(\nu(s_\psi) - (\nu(s_0)).$$

Since both ν and μ are hypothesized to be continuous in the W^2-sense, the above inequality shows that ρ is also continuous in the same sense. This in turn implies that the number $\langle\rho(s_\psi)(\tau)\rangle_\tau$ is a continuous function of s_ψ, in the W^2-sense. But since the frequencies $\{\alpha_r\}$ are incommensurable, by Eq. (12) the subset T_0^Q may always be chosen with all of its elements satisfying $D_{W^2}(s_\psi - s_0) < \delta$ for arbitrarily small $\delta > 0$. Thus, $\langle\rho(s_\psi)(\tau)\rangle_\tau = \langle\rho(s_0)(\tau)\rangle_\tau$, and the result (11) is obtained.

The smoothness conditions C1 and C2 are not strong conditions. The continuity is in a mean-squared sense over Ω, and is not point-wise. Thus, static transductions $\mu(s)(t) = f(s(t))$ conform to C2 even if f has a countable number of discontinuities which have finite total displacement. The condition C1 applies only to responses to signals in Ω, and thus is also fairly weak. In particular, these "reasonable smoothness" conditions do not exclude any reasonable biological transducers we might expect to encounter, unidentified, in the laboratory. However, these conditions allow us to construct a complete set of orthogonal functionals, as shown below.

The Fourier representation. We may anticipate that a wide variety of nonlinear transducers will respond to the sum of Q sinusoids, Eq. (9), with a response in the form of a Fourier sum

$$\mu(s_0)(t) = \sum_\beta \tilde{\mu}(\beta)\exp(i\beta t) \tag{13}$$

where the discrete set of real numbers β belong to the set of all sums, differences, and positive and negative integer multiplies of the Q incommensurate input frequencies, and where the real value of $\mu(t)$ evidently demands that $\tilde{\mu}(-\beta) = \overline{\tilde{\mu}(\beta)}$.

Among all the real values of β we may select the discrete set above by means of the Fourier integral

$$\tilde{\mu}(\beta) \equiv \lim_{l\to\infty} \frac{1}{l}\int_0^l \mu(s_0)(\tau) \exp(-i\beta\tau)\, d\tau = \langle \mu(s_0)(\tau) \exp(-i\beta\tau)\rangle_\tau . \tag{14}$$

Evidently the Fourier integral evaluates to zero unless β is drawn from the subset of values which appear in the Fourier sum of the previous equation. The discrete set of coefficients $\tilde{\mu}(\beta)$ of Eq. (14) forms the Fourier resolution of the transducer μ.

For a pair of transducers which permit a replacement of ensemble average by time average we have, by simple time average of the product of their Fourier sums,

$$(\mu, \nu) = \sum_\beta \overline{\tilde{\mu}(\beta)}\, \tilde{\nu}(\beta) \tag{15}$$

and in particular,

$$\|\mu\|^2 = \sum_\beta |\tilde{\mu}(\beta)|^2. \tag{16}$$

These Parseval relations affirm that inner product and quadratic distance take their natural forms in the vector space of the Fourier resolution.

The Fourier resolution has two other useful properties, which may be verified at once by use of Eq. (14). First, a linear transducer L with transfer function $\hat{L}(\omega)$ has a Fourier resolution

$$\begin{aligned} \tilde{L}(\beta) &= \tfrac{1}{2}a_r \hat{L}(\beta) \quad \text{for} \quad \beta = \pm\alpha_r\,, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Second, if μ is a transducer consisting of the arbitrary transducer ν followed by the linear transducer L, then

$$\tilde{\mu}(\beta) = \tilde{\nu}(\beta)\hat{L}(\beta). \tag{17}$$

Clearly the Fourier resolution falls short of a full characterization of the corresponding transducer, as the linear example above shows: a linear transducer chosen (perversely) with nulls at the Q frequencies α_r will map to the zero vector via the Fourier resolution. However, the laboratory challenge of an unclassified transducer is not so much unique identification, but rather an approximate description which is reasonably accurate and also concise. An input sum of Q sinusoids will, for example, uniquely specify a single linear transducer from a family of linear transducers whose transfer functions are rational functions of frequency and analytic in $2Q$ parameters or less. Nonlinear transducers yield Fourier coefficients at the much larger set of frequencies $\{|\beta_j|\}$, so the opportunity for identification of parameters is even greater than in the linear case.

We conclude the section by presenting rigorous and permissive conditions under which the Fourier resolution (14) will exist: the Fourier resolution exists for any transducer which satisfies the above smoothness conditions C1 and C2. First we note that such a transducer is completely characterized by its response to the input signal s_0. For if $\mu(s_0)(t) = \nu(s_0)(t)$ for almost all t, then the transducer classes μ and ν are equal:

$$\langle\{(\mu - \nu)(s_0)(t)\}^2\rangle_\tau = 0 \Rightarrow (\mu - \nu, \mu - \nu) = 0 \Rightarrow \mu = \nu$$

by Eq. (11).

We may now invoke the classical machinery which is used in the study of almost periodic functions. The condition C2 states that $\mu(s_0)(\tau)$ is an almost periodic function of τ, of class W^2 [4, p. 77]. In consequence the Fourier resolution (14) exists and is nonzero on a set of real values of β which is denumerable [4, p. 104].

For almost periodic functions, the Parseval relation

$$\{D_{W^2}(\mu(s_0))\}^2 = \sum_{\beta} |\tilde{\mu}(\beta)|^2$$

holds [4, p. 107] and is bounded by condition C1.

Hence the Fourier sum (13) converges in the W^2-sense. By (11), convergence in the W^2-sense implies convergence in $\mathfrak{M}$, for

$$\|\mu\|^2 \equiv (\mu, \mu) = \langle\{\mu(s_0)(\tau)\}^2\rangle_\tau = \lim_{l\to\infty} \frac{1}{l}\int_0^l \{\mu(s_0)(\tau)\}^2 d\tau$$

$$\equiv \{D_{W^2}(\mu(s_0))\}^2,$$

which establishes Eq. (16).

The orthogonal subspaces. For the sum-of-sinusoids ensemble Ω, we demonstrate that 1) the hierarchy of orthogonal subspaces $\{K_j\}$ of Sec. 2 may be determined without recourse to the Gram-Schmidt procedure, and 2) these subspaces are complete for transductions which satisfy conditions C1 and C2 above.

The subspace M_j of homogeneous jth-order transductions is certainly spanned by the "multiple time lag" transductions $D_{j,\mathbf{T}}$ of Eq. (4). The Fourier resolution $\tilde{D}_{j,\mathbf{T}}(\beta)$ is obtained by substituting Eq. (9) into Eq. (4), and the result into the definition (14) of the Fourier resolution:

$$\tilde{D}_{j,\mathbf{T}}(\beta) = \lim_{l\to\infty} \frac{1}{l}\int_0^l \exp(-i\beta\tau)\, D_{j,\mathbf{T}}(s_0)(\tau) d\tau$$

$$= \frac{1}{2^j}\lim_{l\to\infty} \frac{1}{l}\int_0^l \exp(-i\beta\tau) \prod_{m=1}^{j} \sum_{r=1}^{Q} (\exp(i\alpha_r(\tau - T_m)) + \exp(-i\alpha_r(\tau - T_m)))a_r\, d\tau.$$

The frequencies for which $\tilde{D}_{j,\mathbf{T}}(\beta) \neq 0$ are integer combinations of at most j of the input frequencies $\{\alpha_r\}$:

$$\beta = \sum_{r=1}^{Q} n_r\alpha_r, \quad \text{where} \quad \sum_{r=1}^{Q} |n_r| \leq j, \quad \text{and} \quad n_r \quad \text{in} \quad \{0, \pm 1, \pm 2, \cdots\}. \tag{18}$$

We will call the set of frequencies β satisfying (18) the lattice frequencies of order j. If β is a lattice frequency of order j but not of order $j - 1$, we will call β a lattice frequency primitive at order j. In other words, a lattice frequency primitive at order j is first encountered at order j, and equality holds in (18). In this case, we will call the vector on integers $\mathbf{n} \equiv (n_1, n_2, \cdots, n_Q)$ a *lattice point* of order j.

Thus, we have proven:

all the transducers in the jth-order polynomial subspace $\bigoplus_{k=0}^{j} M_k$ have a Fourier resolution (13) that is zero except on lattice frequencies of orders only up through order j.

Conversely, we now prove:

any transducer which yields a Fourier resolution which is nonzero only on lattice frequencies (or lattice points $\mathbf{n}$) of orders up through j is an element of $\bigoplus_{k=0}^{j} M_k$, the space of polynomial transducers of order j.

The proof follows from three simple observations: (1) A Fourier resolution which vanishes outside the lattice frequencies (or lattice points) of orders up through j may be expressed as a sum of Fourier resolutions each of which vanishes *on all lattice points except* on lattice points corresponding to a particular (unsigned) lattice frequency $|\beta|$ primitive at order $k \leq j$; (2) a homogeneous transduction of order $k \leqq j$ followed by a linear transduction yields a composed transduction still of order k; (3) a linear transduction may be chosen which nulls all lattice frequencies of order j except for the values $\pm\beta$, and this linear transducer may be preceded by a homogeneous transducer of order k whose Fourier resolution includes the unsigned frequency $|\beta|$ (see Eq. (17)). (For example, we may pick the homogeneous transducer $D_{k,\mathbf{T}}$ from above.) Thus the set of Fourier resolutions which vanish outside of lattice frequencies (or lattice points) of order j or less represent exactly the set of polynomial transducers of order j (in the quotient space $\mathfrak{M}$), the transducers which form the subspace $\bigoplus_{k=0}^{j} M_k$.

The orthogonal subspaces K_j may now be determined. K_j is that part of $\bigoplus_{k=0}^{j} M_k$ that is orthogonal to $\bigoplus_{k=0}^{j-1} M_k$, the space of all functional polynomials of order less than j. By the characterization in the above paragraph, K_j is exactly the subspace of transducers that are orthogonal to all transducers with nonzero Fourier resolutions on lattice frequencies of order up through $j - 1$. The expression (15) for the inner product in terms of the Fourier resolution characterizes this category simply: K_j is the subspace of transducer classes whose Fourier resolutions are zero except on the lattice frequencies primitive at order j.

The component of a transducer class μ that lies in K_j may easily be obtained from the Fourier resolution (14) of μ. One merely deletes all components at lattice frequencies primitive at order other than j. That is,

$$\tilde{\kappa}_j(\beta) = \tilde{\mu}(\beta), \beta \text{ a lattice frequency primitive at order } j$$
$$= 0 \quad \text{otherwise.}$$

This is precisely the result that would be obtained by the general procedure (8), by choosing transductions with Fourier resolutions nonzero on single primitive lattice point pairs as an orthogonal basis for K_j.

As a matter of rigor, to demonstrate completeness of the subspaces K_j, we need only show that the Fourier resolution (14) must be zero at any frequency that is not a lattice frequency. Suppose on the contrary that $\tilde{\mu}(\beta) \neq 0$, and that β is not of the form (18). Then, according to the theory of almost-periodic functions [4] either β is incommensurate with the input frequencies $\{\alpha_r\}$ or it is a submultiple of a lattice frequency. In either case, there exists an $h > 0$ such that for any $\delta > 0$, it is possible to choose a δ-W^2-almost period L_δ of s_0 such that $|\beta L_\delta| > h \pmod{2\pi}$. This implies that

$$D_{W^2}(\mu(s_0) - \mu(s_{L_\delta})) > \sqrt{2} \cdot \left|\sin \frac{h}{2}\right| \cdot |\tilde{\mu}(\beta)|,$$

which contradicts condition C2.

Thus the frequencies β at which a Fourier estimation $\tilde{\mu}(\beta)$ may be nonzero are of the form (18). This permits a change in notation that will facilitate later work. We index the components of the Fourier resolution by lattice points $\mathbf{n}$ rather than lattice frequencies:

$$\tilde{\mu}(\mathbf{n}) \equiv \tilde{\mu}(\mathbf{n}\cdot\alpha). \tag{19}$$

This indexing, combined with the relation (11), provides another useful form for the Fourier resolution:

$$\tilde{\mu}(\mathbf{n}) = \langle \mu(s_{\psi})(0) \cdot \exp(-i\mathbf{n} \cdot \psi) \rangle_{\psi} \tag{20}$$

4. Comparison with the Wiener technique. We have discussed the sum-of-sinusoids technique as a special case of the class of Wiener-type theories for the analysis of nonlinear systems. We have suggested that the technique has certain practical advantages owing to the deterministic nature of the input signal. In this section we wish to show that the sum-of-sinusoids technique retains a fundamental similarity to the original Wiener technique. We will also show that the Fourier resolution of a transducer on lattice points of order j approaches the Fourier transform of that transducer's jth-order Wiener kernel. The original Wiener technique [29, 58] relies on Gaussian noise as an input signal, and has been widely used in biology [19, 20, 30, 32, 33, 35–37, 45, 49, 57].

The correspondence between the Wiener kernels and the frequency resolution is important because it shows that the sum-of-sinusoids technique retains a key advantage of the Wiener theory, that the Wiener kernels of many nonlinear transducers have simple and distinctive functional forms. Thus, the qualitative features of the experimentally-determined frequency resolution of a transducer may suggest simple nonlinear models [55].

We prove the correspondence between the Wiener kernels and the frequency resolution by proving it for each element of a convenient basis set of transducers. These transducers consist of an arbitrary linear filter L_1, followed by a static nonlinearity whose operating curve is a Hermite polynomial, followed by a second linear filter, L_2. These basis transducers are chosen because their Wiener kernels are particularly simple (Eq. (34) below).

To see that these transducers are a basis for the space of all transducers with convergent Wiener expansions, we show that each of the multiple time-lag transducers $D_{j,\mathrm{T}}$ (Eq. (4)) is in the span of the linear-Hermite-linear transducers. The output of a transducer $D_{j,\mathrm{T}}$ is the product of its inputs at j previous times. It therefore can be expressed as a finite sum of transducers, each of which is a linear filter followed by jth-power law static nonlinearities, since

$$X_1 X_2 \cdots X_j = \frac{1}{j!} \sum_{k=1}^{j} (-1)^{j-k} \sum_{s_1 < \cdots < s_k} (X_{s_1} + X_{s_2} + \cdots X_{s_k})^j.$$

The jth-power law static nonlinearities can be expressed as a finite sum of Hermite polynomial static nonlinearities, so the proof is complete.

Even though the second linear filter L_2 is not necessary to generate a spanning set of transducers, we include it because it presents no additional complexities and because the linear-static nonlinear-linear sandwich model is frequently used as a model for biological transductions [15, 49, 55].

The nub of the calculation of the frequency resolution of the linear-static nonlinear-linear sandwich is the calculation of the frequency resolution of a static Hermite polynomial transducer. This we accomplish first; then we show how the calculation extends to the full sandwich.

Hermite polynomial nonlinearities. In this subsection, we calculate the Fourier resolution (20) of static nonlinear transducers whose operating curves are Hermite polynomials. The Hermite polynomials are constructed with respect to a Gaussian distribution whose variance is the variance P of the sum-of-sinusoids signal:

$$P = \langle s(t)^2 \rangle_\Omega = \tfrac{1}{2} \sum_{r=1}^{Q} a_r^2 . \tag{21}$$

The Hermite polynomials for a Gaussian of variance P may be defined by the generating function [51]

$$G(z, x) \equiv \sum_{j=0}^{\infty} \frac{z^j}{j!} H_j(x) = \exp\left(-P \frac{z^2}{2} + zx\right), \tag{22}$$

where $H_j(x)$ is the jth Hermite polynomial, parametric in P.

We treat the generating function (22) as a transducer parametric in z. The calculation of its frequency resolution $\tilde{G}(z, \mathbf{n})$ involves estimation of a product of independent terms. We substitute (22) into the expression (20) for the Fourier resolution, and use the expression (9) for s_{ψ} to obtain

$$\begin{aligned} \tilde{G}(z, \mathbf{n}) &\equiv \sum_{j=0}^{\infty} \frac{z^j}{j!} \tilde{H}_j(\mathbf{n}) \\ &= \left\langle \exp\left\{ z\left(\sum_{r=1}^{Q} a_r \cos\psi_r \right) - P\frac{z^2}{2} - i\mathbf{n}\cdot\psi \right\}\right\rangle_{\psi} \\ &= \prod_{r=1}^{Q} \left\langle \exp\left\{ za_r \cos\psi_r - \tfrac{1}{4}a_r^2 z^2 - in_r\psi_r \right\}\right\rangle_{\psi_r} = \prod_{r=1}^{Q} F_{n_r}(a_r), \end{aligned} \tag{23}$$

where

$$F_{n_r}(a_r) = \exp\left(-\tfrac{1}{4}a_r^2 z^2\right) \cdot \frac{1}{2\pi} \int_0^{2\pi} \exp\{-in_r\psi + za_r \cos\psi\}\, d\psi .$$

Thus, F_{n_r} is simply related to the Bessel function $J_{|n_r|}$, [1, Eq. (9.1.21)] and this allows us to expand F_{n_r} as a power series:

$$F_{n_r}(a_r) = \left(\frac{a_r z}{2}\right)^{|n_r|} \cdot \left(\sum_{l=0}^{\infty} \frac{1}{l!} \left(-\tfrac{1}{4}a_r^2 z^2\right)^l \right) \cdot \left(\sum_{k=0}^{\infty} \frac{\left(\tfrac{1}{4}a_r^2 z^2\right)^k}{k!(k + |n_r|)!} \right) .$$

The first few terms of the series for F_{n_r} are

$$\begin{aligned} F_{n_r}(a_r) = \frac{1}{|n_r|!} \cdot \left(\frac{a_r z}{2}\right)^{|n_r|} \Bigg\{ 1 &+ \frac{a_r^2 z^2}{4}\left(- \frac{|n_r|}{|n_r| + 1} \right) \\ &+ \frac{a_r^4 z^4}{16}\left(\frac{|n_r|^2 + |n_r| - 1}{2(|n_r| + 1)(|n_r| + 2)} \right) + \cdots \Bigg\} . \end{aligned} \tag{24}$$

We now consider the behavior of $F_{n_r}(a_r)$ as Q, the number of sinusoids in the input signal, grows. We assume that the power in the input signal, Eq. (21), remains constant. We also assume that no single sinusoid dominates the input signal. These conditions may be formulated

$$\max_r \frac{a_r^2}{2P} = \frac{1}{q} = \mathcal{O}\left(\frac{1}{Q}\right) . \tag{25}$$

Now let us fix the lattice point $\mathbf{n}$ of order $N = \sum_{r=1}^{Q} |n_r|$ and consider the behavior of the final product in Eq. (23) as Q approaches infinity under the condition (25). Nearly all of the terms of this product have $n_r = 0$. For these terms, an asymptotic estimate of $F_{n_r}(a_r)$

derived from (24) has an initial error term of $P^2z^4/16q^2$, which is $O(Q^{-2})$. At most N of the terms in the product of Eq. (23) have n_r nonzero. For each of these terms, the asymptotic estimate derived from (24) has an initial error term which is $O(Q^{-1})$. Thus, the product in Eq. (23) is estimated by

$$\left| \frac{\prod_{r=1}^{Q} F_{n_r}(a_r)}{\prod_{r=1}^{Q} \frac{1}{|n_r|!}\left(\frac{a_r z}{2}\right)^{|n_r|}} - 1 \right| \leq \frac{1}{q}\left\{\frac{Pz^2}{2}\left(\sum_{r=1}^{Q} \frac{|n_r|}{|n_r|+1}\right) + \frac{P^2z^4Q}{16q}\right\} + O(Q^{-2}),$$

where the last term contains only even powers of $z \geq 4$.

We now equate like powers of z on both sides of Eq. (23) to obtain formulae for $\tilde{H}_j(\mathbf{n})$:

$$\tilde{H}_N(\mathbf{n}) = U(N, \mathbf{n}, \mathbf{a}), \tag{26a}$$

$$\tilde{H}_{N+2}(\mathbf{n}) = U(N+2, \mathbf{n}, \mathbf{a}) \cdot \left(\frac{P}{2q} \sum_{r=1}^{Q} \frac{|n_r|}{|n_r|+1}\right), \tag{26b}$$

$$\tilde{H}_{N+4}(\mathbf{n}) = U(N+4, \mathbf{n}, \mathbf{a}) \cdot \left(\frac{P^2Q}{16q^2} + O(Q^{-2})\right), \tag{26c}$$

$$\tilde{H}_j(\mathbf{n}) = U(j, \mathbf{n}, \mathbf{a}) \cdot O(Q^{-2}), j > N + 4 \quad \text{and} \quad j \equiv N \pmod 2 \tag{26d}$$

$$\tilde{H}_j(\mathbf{n}) = 0, j < N \quad \text{or} \quad j \not\equiv N \pmod 2 \tag{26e}$$

where

$$U(j, \mathbf{n}, \mathbf{a}) = j! \prod_{r=1}^{Q} \frac{1}{|n_r|!} (\tfrac{1}{2}a_r)^{|n_r|}.$$

Thus, we have determined the Fourier resolution of a jth-order Hermite polynomial static nonlinearity H_j on an Nth-order lattice point $\mathbf{n}$. Unless the order j of the Hermite polynomial and the order N of the lattice point are equal, this value approaches zero as the number of frequencies increases. The rapidity of the approach is $O(Q^{-1})$ if $j = N + 2$ or $j = N + 4$, and is $O(Q^{-2})$ or faster if j is greater than $N + 4$, but j and N have the same parity. When j is less than N or if j and N have opposite parities, the Fourier resolution $\tilde{H}_j(\mathbf{n})$ is identically zero.

Finally, we obtain a rough estimate in dimensionless terms of the disparity between the orthogonal subspaces constructed with respect to the sum-of-sinusoids inner product and those constructed with respect to the Gaussian inner product. To do this, we consider the frequency resolutions of the transducers h_j, which are the static nonlinearities H_j normalized to unity in the Gaussian norm. Thus, the transducers h_j are vectors of unit length in the jth orthogonal subspace constructed with respect to the Gaussian inner product. These transducers are given by

$$h_j = (1/(j!P^j)^{1/2})H_j .$$

We assume that the input sinusoids have equal amplitude, so that $q = Q$ in the condition (25). Using Eqs. (26), we find

$$\tilde{h}_N(\mathbf{n}) = \left(\frac{N!}{P^N}\right)^{1/2} \prod_{r=1}^{Q} \frac{1}{|n_r|!} (\tfrac{1}{2}a_r)^{|n_r|}, \tag{27a}$$

$$\left|\frac{\tilde{h}_{N+2}(\mathbf{n})}{\tilde{h}_N(\mathbf{n})}\right| \leq \left(\frac{(N+1)(N+2)}{Q}\right)^{1/2} \cdot \tfrac{1}{2}\sum_{r=1}^{Q} \frac{|n_r|}{|n_r|+1} \leq N\left(\frac{(N+1)(N+2)}{4Q}\right)^{1/2}, \quad (27b)$$

$$\left|\frac{\tilde{h}_{N+4}(\mathbf{n})}{\tilde{h}_N(\mathbf{n})}\right| \leq \left(\frac{(N+1)(N+2)(N+3)(N+4)}{16Q}\right)^{1/2} + O(Q^{-2}), \quad (27c)$$

$$\left|\frac{\tilde{h}_j(\mathbf{n})}{\tilde{h}_N(\mathbf{n})}\right| = O(Q^{-2}), j > N+4 \quad \text{and} \quad j \equiv N \pmod 2, \quad (27d)$$

$$\tilde{h}_j(\mathbf{n}) = 0, j < N \quad \text{or} \quad j \not\equiv N \pmod 2. \quad (27e)$$

The second inequality in Eq. (27b) follows from

$$\sum_{r=1}^{Q} \frac{|n_r|}{|n_r|+1} \leq \frac{N}{2}, \qquad \text{for} \qquad \sum_{r=1}^{Q} |n_r| = N,$$

with the maximum occurring when exactly N of the n_r's are equal to ± 1.

Thus, we estimate that, up to $O(Q^{-1})$, the frequency resolutions at Nth-order lattice points are "contaminated" only by $(N+2)$nd-and $(N+4)$th-order Wiener kernels. The fraction of the value of the frequency resolution at an Nth-order lattice point that represents spillover from a unit-higher-order Wiener kernel is approximately $N^2/4Q$ for the $(N+2)$nd-order Wiener kernel, and $N^2/16Q$ for the $(N+4)$th-order Wiener kernel.

Nonstatic nonlinearities. In this subsection, we generalize the above results concerning the relationship of the Wiener kernels to the Fourier resolution obtained with a sum of sinusoids. We will show that the normalized frequency resolution of a linear-Hermite polynomial-linear sandwich approaches points on the Fourier transforms of the Wiener kernels of this transducer as the number of sinusoids in the input signal becomes large. As pointed out at the beginning of this section, these sandwich transducers are a basis for the space of transducers whose Wiener series are convergent, so the result has similar breadth.

We first define the normalized Fourier resolution. This is essentially the Fourier resolution (20), corrected for the amplitudes of the input sinusoids and the combinatorial coefficients of (27). The normalized Fourier resolution $\tilde{\mu}^0(\mathbf{n})$ of a transducer μ on a lattice point $\mathbf{n}$ of order N is defined by

$$\tilde{\mu}^0(\mathbf{n}) \equiv \tilde{\mu}(\mathbf{n}) \left(\frac{P^N}{N!}\right)^{1/2} \prod_{r=1}^{Q} (\tfrac{1}{2}a_r)^{-|n_r|} |n_r|!. \quad (28)$$

Thus, according to Eq. (27), the normalized Fourier resolution $\tilde{h}_N^0$ of the transducer h_N is 1 at all Nth-order lattice points, and at most $O(Q^{-1})$ at all other lattice points.

Now that we have shifted consideration to nonlinear systems that are not necessarily static, it is necessary to consider the limiting behavior of the frequencies $\{\alpha_r\}$ as well as that of the amplitudes $\{a_r\}$. As the number of sinusoids becomes large, we must require that the power spectrum of the sum-of-sinusoids signal approaches the power spectum $\mathcal{P}(\omega)$ of some particular Gaussian noise, in an appropriate integral sense. That is, we must choose the frequencies $\alpha_r(Q)$ and the amplitudes $a_r(Q)$ so that

$$\lim_{Q\to\infty} \sum_{\omega_1 \leq \alpha_r < \omega_2} \tfrac{1}{2}a_r^2 = \int_{\omega_1}^{\omega_2} \mathcal{P}(\omega)\,d\omega. \quad (29)$$

This is a strengthened version of the condition (25) that sufficed in the discussion of

Hermite polynomial static nonlinearities. The condition (29) can be satisfied for smooth $\mathcal{P}(\omega)$ with finite total power

$$P = \int_0^\infty \mathcal{P}(\omega)d\omega.$$

We now determine the limiting behavior of the frequency representation of a nonlinear transducer μ_j composed of a linear filter L_1 with transfer function $\hat{L}_1(\omega)$, followed by a Hermite polynomial static nonlinearity, followed by a second linear filter L_2 with transfer function $\hat{L}_2(\omega)$. We choose the static nonlinearity to be a Hermite polynomial of order j orthogonalized with respect to a Gaussian of variance equal to the power passed by the initial linear filter. This power, P_{L_1}, is given by

$$P_{L_1} = \int_0^\infty (L_1(\omega))^2 \mathcal{P}(\omega)d\omega = \lim_{Q\to\infty} \sum_{r=1}^{Q} \tfrac{1}{2}a_r^2(L_1(\omega))^2. \tag{30}$$

We write $\mu_j = L_1 \circ h_{j;P_{L_1}} \circ L_2$, where $h_{j;P_{L_1}}$ denotes the normalized Hermite polynomial of degree j constructed with respect to a Gaussian weighting of variance P_{L_1}.

It suffices to calculate the Fourier resolution of $\mu_j \equiv L_1 \circ h_{j;P_{L_1}}$ since, according to Eq. (17) and (19),

$$\tilde{\mu}_j(\mathbf{n}) = \tilde{\nu}_j(\mathbf{n}) \cdot \hat{L}_2(\alpha \cdot \mathbf{n}). \tag{31}$$

The initial linear filter L_1 transforms the input sum-of-sinusoids signal into a similar signal in which the amplitudes and phases of the component sinusoids are altered by the amplitudes and phases of $\hat{L}_1$. Thus, the response of the transducer ν_j to the signal s is identical to that of the static transducer $h_{j;P_{L_1}}$ which acts on an altered sum-of sinusoids signal s':

$$\nu_j(s_{\psi})(\tau) = h_{j;P_{L_1}}(s'_{\psi+\xi})(\tau),$$

where

$$s_{\psi}'(t) = \sum_{r=1}^{Q} |\hat{L}_1(\alpha_r)|\, a_r \cos(\alpha_r t + \psi_r),\ \xi \equiv (arg(\hat{L}_1(\alpha_1)), \cdots, arg(\hat{L}_1(\alpha_Q))).$$

According to Eq. (20),

$$\begin{aligned}\tilde{\nu}_j(\mathbf{n}) &= \langle h_{j;P_{L_1}}(s'_{\psi+\xi})(0)\cdot \exp(-i\mathbf{n}\cdot\psi)\rangle_{\psi\ in\ T^Q} \\ &= \exp(i\mathbf{n}\cdot\xi)\,\langle h_{j;P_{L_1}}(s'_{\psi})(0)\cdot\exp(-i\mathbf{n}\cdot\psi)\rangle_{\psi-\xi\ in\ T^Q} \\ &= \exp(i\mathbf{n}\cdot\xi)\tilde{h}'_{j;P_{L_1}}(\mathbf{n}).\end{aligned} \tag{32}$$

Here, $\tilde{h}'_{j;P_{L_1}}(\mathbf{n})$ denotes the Fourier resolution of the static transducer $h_{j;P_{L_1}}$ with respect to the input ensemble Ω' composed of the signals s_{ψ}'. The results of the previous section are now applicable, except for the fact that the power in each signal s_{ψ}' is not exactly P_{L_1}, but only approaches this value as Q grows (Eq. (29)). However, by re-expanding the polynomial $h_{j;P_{L_1}}$ in terms of Hermite polynomials orthogonal with respect to a Gaussian of variance equal to the power of the elements s_{ψ}' of Ω', it may be seen that this problem does not affect the outcome of the present calculation. We therefore apply Eq. (27) (using input amplitudes $a_r|\hat{L}_1(\alpha_r)|$) to obtain

$$\lim_{Q\to\infty} \tilde{h}'_{N;P_{L_1}}(\mathbf{n}) = (N!/(P_{L_1})^N)^{1/2} \prod_{r=1}^{Q} \frac{1}{|n_r|!}(\tfrac{1}{2}a_r|\hat{L}_1(\alpha_r)|)^{|n_r|},$$

$$\lim_{Q\to\infty} |\tilde{h}'_{j;P_{L_1}}(\mathbf{n})/\tilde{h}'_{N;P_{L_1}}(\mathbf{n})| = 0, \qquad j \neq N.$$

Next, we substitute the above result in Eqs. (31) and (32) to obtain

$$\lim_{Q\to\infty} \tilde{\mu}_N(\mathbf{n}) = \lim_{Q\to\infty} \exp(i\mathbf{n}\cdot\xi)\tilde{h}_{j;P_{L_1}}(\mathbf{n})\cdot\hat{L}_2(\alpha\cdot\mathbf{n})$$

$$= \left(\frac{N!}{(P_{L_1})^N}\right)^{1/2}\cdot\hat{L}_2(\alpha\cdot\mathbf{n})\prod_{r=1}^{Q}\frac{1}{|n_r|!}\{\tfrac{1}{2}a_r\hat{L}_1(\alpha_r sgn(n_r))\}^{|n_r|}$$

and

$$\lim_{Q\to\infty} \tilde{\mu}_j(\mathbf{n})/\tilde{\mu}_N(\mathbf{n}) = 0, \qquad j \neq N,$$

where we have recombined the amplitudes and phases of $\hat{L}_1(\alpha_r) \equiv |\hat{L}_1(\alpha_r)| \exp(i\xi_r)$.

We now obtain the desired expression for the limiting behavior of the normalized frequency resolution $\tilde{\mu}_j^0(\mathbf{n})$ by substitution of the above result in Eq. (28):

$$\lim_{Q\to\infty} \tilde{\mu}_j^0(\mathbf{n}) = \left(\frac{P}{P_{L_1}}\right)^{N/2}\cdot\hat{L}_2(\alpha\cdot\mathbf{n})\cdot\prod_{r=1}^{Q}\{\hat{L}_1(\alpha_r sgn(n_r))\}^{|n_r|},$$

$$\lim_{Q\to\infty} \tilde{\mu}_j^0(\mathbf{n}) = 0, \qquad j \neq N. \tag{33}$$

By the method of Lee and Schetzen [29] and Price's theorem [41], it may be shown (for example [49]) that the Fourier transform of the Nth Wiener kernel of μ_j measured with Gaussian noise of spectrum $\mathcal{P}(\omega)$ is

$$(P_{L_1})^{-N/2}\hat{L}_2(\omega_1 + \cdots + \omega_N)\prod_{k=1}^{N}\hat{L}_1(\omega_k), \qquad j = N$$

$$0, \qquad j \neq N. \tag{34}$$

Except for a factor of $P^{N/2}$, this agrees with Eq. (33), if we choose the N frequencies $\{\omega_k\}$ so that exactly $|n_r|$ of them are equal to $\alpha_r sgn(n_r)$.

Thus, we have found that the normalized Fourier resolution of the transducers $L_1 \circ h_{j;P_{L_1}} \circ L_2$ approaches values of the Fourier transforms of the Wiener kernels of this transduction. This correspondence, which extends to all transducers that have a Wiener expansion, is a crucial result. It allows one to calculate the frequency resolutions of simple classes of transducers, which is necessary for a rational approach to the development and evaluation of models for biological transductions [46, 55].

5. Commensurable frequencies. In Secs. 3 and 4, we assumed that the multiple sinusoids of the input had incommensurate frequencies. In a laboratory implementation, there are several practical advantages of using sinusoids whose frequencies are high multiples of a common repeat period. Such a signal is easy to generate on a digital computer, and the time average (14) may be evaluated for all lattice frequencies by a single application of a discrete Fourier transform. In this section, we discuss the transition from incommensurate to commensurable frequencies. The result of this generalization is a practical technique that has been applied with success in the analysis of biological transductions [8, 9, 46, 53, 54, 55, 59].

The transition to commensurable frequencies. In the preceding sections, the fact that the Q superimposed sinusoids of the input had incommensurate frequencies was a convenience but was not essential. The incommensurate nature of the frequencies was helpful in two ways. Every lattice point $\mathbf{n} = (n_1, n_2, \cdots, n_Q)$ leads to a unique frequency, $\beta = \mathbf{n} \cdot \alpha$. Thus the output frequencies themselves could be used to index the coefficients of the Fourier resolution. Once commensurate input frequencies are allowed, evidently some distinct choices of lattice points will lead to identical output frequencies, and that necessitates indexing the components of the Fourier resolution with the lattice points themselves. The hypothesis of incommensurate frequencies also allowed replacement of the phase ensemble average by the time average. This followed from the fact that the phase shifts resultant from the progress of time ultimately sample the torus of phases T^Q densely. Conversely, a choice of frequencies α which are commensurate must lead to an eventual common period of the sinusoids, at which period the trajectory in the torus of phases will close and no additional phase points will be sampled. In consequence, the theoretical discussion which involves ensemble averages, which are averages over phases, is not immediately and rigorously applicable to averages over time with a single input signal. As a practical laboratory procedure, the use of commensurate sinusoids permits precise digital filtering methods which rely upon the signal's repetition at the long but finite common period. On the other hand, there is the question whether the phase trajectory of a single input signal makes a sufficiently fine-grained sampling of the phase torus to allow the replacement of phase average by time average (in the same way that a Riemann integral may be replaced by an approximating sum). Otherwise, adequate approximation will require that data be taken in several passes, each starting with an inequivalent set of initial phases.

The treatment of incommensurate frequencies is applicable to the commensurate case, if we confine our attention to transducers μ which satisfy the further (and very permissive) condition:

C3: The transduction μ at time zero is a continuous function of the frequencies α. Transducer with finite memory satisfy this condition, and so will a much broader set of transducers whose memories fade arbitrarily close to zero after a long enough time.

Condition C3 guarantees that the Fourier resolution, Eq. (20), is a continuous function of the set of frequencies α. Thus, all the results that we have proven for an incommensurate set of frequencies carry over to a general set of frequencies, since there are incommensurate frequency sets arbitrarily close to any given frequency set.

In the case of commensurable frequencies, the time average (14) may be evaluated by an integration span which extends only over the finite common period of the sinusoids. However, that finitely-evaluated time average now represents only an incomplete sampling of the phase ensemble, and is no longer equal to the exact result (20) for the Fourier resolution. It is evident that by appropriate choice of inequivalent initial phases ψ the phase ensemble average (20) may be approximated to arbitrary accuracy by an expression of the form

$$\tilde{\mu}(\mathbf{n}) = \lim_{k \to \infty} \frac{1}{k} \sum \langle \mu(s_{\psi,k})(\tau) \exp(-i\alpha \cdot \mathbf{n}\tau) \rangle_\tau , \tag{35}$$

where the sum ranges over k inequivalent phase vectors ψ^k. Without proof, we quote a practical example. Let $Q = 8$, with $\alpha_j = (2^{j+2} - 1)\alpha_0$. Choose 8 inequivalent phase sets ψ^k, so that $M_{kj} \equiv \exp i\psi_j^k$ is the entry in the kth row and jth column of the Hadamard matrix:

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 & - & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix} \tag{36}$$

Then, in the approximation above, the evaluation of Fourier coefficients at lattice points of second order will not be contaminated by Fourier coefficients primitive at third, fourth, fifth, sixth, or seventh order [54].

Practical advantages of the sum-of-sinusoids method. The sum-of-sinusoids technique has been successful in analyzing linear [8, 9] and nonlinear [46, 53, 54, 55] neural transductions. Here we would like to suggest possible reasons for this success.

The sum-of-sinusoids signal is a deterministic one, in contrast to the Gaussian, Poisson, and other popular stochastic test signals. This means that the correlation properties of the actual test signal are known; for stochastic signals the correlation properties are known only for the abstract ensemble.

With stochastic signals, the problem often arises that the experimentally-determined characterization sequence becomes a poorer approximation to the transducer under study (in the mean-squared sense, as determined by the actual test signal) as higher-order terms are added [20, 28]. Presumably, this apparent paradox arises because the algorithm for extracting the terms of the orthogonal sequence, such as the Lee-Schetzen method [29], assumes that the ideal correlation properties of the ensemble are realized by the test signal. A failure of the mean-squared error to improve with successive terms thus indicates that the characterization procedure is focussing on the departure of the test signal from the ensemble average, rather than on properties of the transducer itself.

Here this undesirable outcome is avoided because of intrinsic features of the sum-of-sinusoids technique that stem from the use of a deterministic test signal. Experimental estimates of the approximating transducers κ_j (Eq. (6)) are essentially Fourier components $\tilde{\mu}(\beta)$ (Eq. (14)) or $\tilde{\mu}(\mathbf{n})$ (Eq. (35)) of the transducers output. Hence, inclusion of additional terms of the orthogonal expansion (8) can only increase the goodness of approximation to the transducer under study, since this procedure amounts to including additional terms of a Fourier expansion.

Additional advantages of the sum-of-sinusoids procedure are manifest if all of the input frequencies α are large integer multiples of a very low frequency, whose period is long in comparison to the time constants of interest in the system under study. In this instance, the ensemble Ω is composed of periodic signals. The average over τ in Eq. (35) may be calculated for all lattice points simultaneously by a single application of the Fast Fourier Transform [11]. This results not only in a great saving of computation, but in an enhancement of the signal-to-noise ratio. This is because all power in Fourier coefficients other than the lattice frequencies of interest may be disgarded as undriven responses. Such digital filtering is very helpful in testing transducers that are noisy or that have discontinuous outputs, or both (such as neural transducers).

Contributions from distinct lattice points $\mathbf{n}$ that share the same lattice frequency cannot be resolved without an average over inequivalent initial phases. In a finite length of time, a sampling of initial points may be explored. For certain sets α, an efficient procedure for such exploration (Eq. (36)) has been advanced [54, 55]. This procedure samples at a lattice of points of high density and symmetry. This sampling procedure has demonstrated the existence of significant fourth-order components in a biological transduction, that of the cat retinal ganglion Y-cell [54].

References

[1] M. Abramowitz and I. Stegun, eds., *Handbook of mathematical functions,* Dover, N.Y., 1959

[2] J. F. Barrett, *The use of functionals in the analysis of nonlinear physical systems,* J. Electronics and Control **15**, 567–615 (1963)

[3] E. Bedrosian and S. O. Rice, *The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs,* Proc. IEEE **59**, 1688–1707 (1971)

[4] A. S. Besicovitch, *Almost periodic functions,* Dover, N.Y., 1954

[5] M. Biederman-Thorson and J. Thorson, *Dynamics of excitation and inhibition in the light-adapted Limulus eye in situ,* J. Gen. Physiol. **58**, 1–19 (1971)

[6] H. W. Bode, *Network analysis and feedback amplifier design,* D. van Nostrand Company, N.Y., 1945

[7] A. Borsellino, R. E. Poppele, and C. A. Terzuolo, *Transfer functions of the slowly adapting stretch receptor organ of Crustacea,* Cold Spring Harbor Symp. Quant. Biol. **30**, 581–586 (1965)

[8] S. E. Brodie, B. Knight, and F. Ratliff, *The response of the Limulus retina to moving stimuli: a prediction by Fourier synthesis,* J. Gen. Physiol. **72**, 129–166 (1978)

[9] S. E. Brodie, B. Knight, and F. Ratliff, *The spatiotemporal transfer function of the Limulus lateral eye,* J. Gen. Physiol. **72**, 167–202 (1978)

[10] R. H. Cameron and W. T. Martin, *The orthogonal development of nonlinear functionals in a series of Fourier-Hermite functionals.* Ann. Math. **48**, 385–392 (1947)

[11] J. Cooley and J. Tukey, *An algorithm for the machine calculation of complex Fourier series,* Math. Computation **19**, 297–301 (1965)

[12] H. deLange, *Attenuation characteristics and phase-shift characteristics of the human fovea-cortex systems in relation to flicker-fusion phenomena,* Doctoral dissertation, Technische Hogeschool, Delft, 1957

[13] R. DeVoe, *Linear superposition of retinal action potentials to predict electrical flicker responses from the eye of the wolf spider, Lycosa baltimoriana (Keyserling),* J. Gen. Physiol. **46**, 75–96 (1962)

[14] F. Dodge, R. Shapley, and B. Knight, *Linear systems analysis of the Limulus retina,* Behav. Sci. **15**, 24–36 (1970)

[15] H. Duifhuis, *Cochlear nonlinearity and second filter: possible mechanism and implications,* J. Acoust. Soc. Am. **59**, 408–423 (1976)

[16] J. F. Fohlmeister, R. E. Poppele, and R. L. Purple, *Repetitive firing: a quantitative study of feedback in model encoders,* J. Gen. Physiol. **69**, 815–848 (1977)

[17] J. F. Fohlmeister, R. E. Poppele, and R. L. Purple, *Repetitive firing: quantitative analysis of encoder behavior of slowly adapting stretch receptor of crayfish and eccentric cell of Limulus,* J. Gen. Physiol. 69, 849–877 (1977)

[18] A. S. French, *Practical nonlinear system analysis by Wiener kernel estimation in the frequency domain,* Biol. Cybernetics **24**, 111–119 (1976)

[19] A. S. French and R. K. S. Wong, *Nonlinear analysis of sensory transduction in an insect mechanoreceptor,* Biol. Cybernetics **26**, 231–240 (1977)

[20] K. Fukurotani, K.-I. Hara, and Y. Oomura, *Dynamic characteristics of the receptive field of L-cells in the carp retina,* Vision Res. **15**, 1403–1405 (1975)

[21] E. Hille and R. S. Phillips, *Functional analysis and semigroups,* American Mathematical Society, Providence, R.I., 1957

[22] S. Hochstein and R. Shapley, *Linear and nonlinear spatial subunits in Y cat retinal ganglion cells,* J. Physiol. **262**, 265–284 (1976)

[23] G. W. Hughes and L. Maffei, *Retinal ganglion cell response to sinusoidal light stimulation,* J. Neurophysiol. **29**, 333–352 (1966)

[24] H. E. Ives, *A theory of intermittent vision,* J. Opt. Soc. Am. **6**, 343–361 (1922)

[25] S. Klein and S. Yasui, *Nonlinear systems analysis with non-Gaussian white stimuli: general basis functionals and kernels,* IEEE Trans. Inf. Theory, in press (1979)

[26] B. W. Knight, J.-I. Toyoda, and F. A. Dodge, *A quantitative description of the dynamics of excitation and inhibition in the eye of Limulus,* J. Gen. Physiol. **36**, 421–437 (1970)
[27] H. I. Krausz, *Identification of nonlinear systems using random impulse trains,* Biol. Cybernetics **19**, 217–230 (1975)
[28] H. I. Krausz and W. O. Friesen, *The analysis of nonlinear synaptic transmission,* J. Gen. Physiol. **70**, 243–265 (1977)
[29] Y. N. Lee and M. Schetzen, *Measurement of the kernels of a nonlinear system by cross-correlation,* Int. J. Control **2**, 237–254 (1965)
[30] E. D. Lipson, *White noise analysis of Phycomyces light growth response system I., II., III,* Biophys. J. **15**, 989–1045 (1975)
[31] G. Marchesini and G. Picci, *Sull' identificazione funzionale di sistemi nonlinearri in regime periodico,* Rendicotti dell' A.E.I. (1969)
[32] P. Z. Marmarelis and K.-I. Naka, *White noise analysis of a neuron chain: An application of the Wiener theory,* Science **175**, 1276–1278 (1972)
[33] P. Z. Marmarelis and K.-I. Naka, *Nonlinear analysis of receptive field responses in the catfish retina, II. One-input white noise analysis,* J. Neurophysiol. **36**, 619–633 (1973)
[34] V. Z. Marmarelis, *A family of quasi-white random signals and its optimum use in biological system identification I. Theory,* Biol. Cybernetics **27**, 49–56 (1977)
[35] G. D. McCann, *Nonlinear identification theory models for successive stages of visual nervous systems of flies,* J. Neurophysiol. **37**, 869–895 (1974)
[36] A. R. Møller, *Statistical evaluation of the dynamic properties of cochlear nucleus units using stimuli modulated with pseudorandom noise,* Brain Res. **57**, 443–456 (1973)
[37] K.-I. Naka, P. Z. Marmarelis, and R. Y. Chan, *Morphological and functional identification of catfish retinal neurons, III. Functional identification,* J. Neurophysiol. **38**, 92–131 (1975)
[38] G. Palm and T. Poggio, *The Volterra representation and the Wiener expansion: validity and pitfalls,* SIAM J. Appl. Math., **33**, 195–216 (1977)
[39] G. Palm and T. Poggio, *Stochastic identification methods for nonlinear systems: An extension of the Wiener theory,* personal communication (1977)
[40] R. B. Pinter, *Sinusoidal and delta function responses of visual cells of the Limulus eye,* J. Gen. Physiol. **49**, 565–593 (1966)
[41] R. A. Price, *A useful theorem for nonlinear devices having Gaussian inputs,* IRE Trans. Information Theory **IT-4**, 69–72 (1958)
[42] J. W. S. Pringle and V. J. Wilson, *The response of a sense organ to a harmonic stimulus,* J. Exp. Biol. **29**, 220–234 (1952)
[43] R. L. Purple and F. A. Dodge, *Interaction of excitation and inhibition in the eccentric cell in the eye of Limulus,* Cold Spring Harbor Symp. Quant. Biol. **30**, 529–537 (1965)
[44] F. Ratliff, B. Knight, F. Dodge, and H. K. Hartline, *Fourier analysis of dynamics of excitation and inhibition in the eye of Limulus: amplitude, phase, and distance,* Vision Res. **14**, 1155–1168 (1974)
[45] A. Sandberg and L. Stark, *Wiener G-functional analysis as an approach to nonlinear characteristics of human pupil light reflex,* Brain Res. **11**, 194–211 (1968)
[46] R. Shapley and J. Victor, *The effect of contrast on the transfer properties of cat retinal ganglion cells,* J. Physiol. **285**, 275–298 (1978)
[47] L. Sirovich and I. Abramov, *Photopigments and pseudopigments,* Vision Res. **17**, 5–16 (1977)
[48] H. Spekreijse, *Rectification in the goldfish retina: analysis by sinusoidal and auxiliary stimulation,* Vision Res. **9**, 1461–1472 (1969)
[49] H. Spekreijse, O. Estevez, and D. Reits, *Visual evoked potentials and the physiological analysis of visual processes in man,* in J. E. Desmedt (ed.), *Visual evoked potentials in man,* Clarendon Press, Oxford, 1977
[50] G. J. St.-Cyr and D. H. Fender, *Nonlinearities of the human oculomotor system: gain,* Vision Res. **9**, 1235–1246 (1969)
[51] G. Szego, *Orthogonal polynomials,* American Mathematical Society, Providence, R. I., 1939
[52] J. Thorson and M. Biederman-Thorson, *Distributed relaxation processes in sensory adaptation,* Science **183**, 161–172 (1974)
[53] J. Victor, *Nonlinear systems analysis: comparison of white noise and sum of sinusoids in a biological system,* Proc. Nat. Acad. Sci. USA **76**, 996–998 (1979)
[54] J. Victor and R. Shapley, *A method of nonlinear analysis in the frequency domain,* Biophys. J., submitted (1979)
[55] J. Victor, R. Shapley, and B. Knight, *Nonlinear analysis of cat retinal ganglion cells in the frequency domain,* Proc. Nat. Acad. Sci. USA **74**, 3068–3072 (1977)

[56] V. Volterra, *Theory of functionals and of integral and integro-differential equations*, London and Glasgow: Blackie and Sons, N.Y.: Dover, N.Y., 1959

[57] W. von Seelen and K. P. Hoffman, *Analysis of neuronal networks in the visual system of the cat using statistical signals*, Biol. Cybernetics **22**, 7–20 (1976)

[58] N. Wiener, *Nonlinear problems in random theory*, M.I.T. Press and Wiley, N.Y., 1965

[59] V. J. Wilson, B. Peterson, K. Fukushima, N. Hirai, and Y. Uchino, *Analysis of vestibulocollic reflexes by sinusoidal polarization of vestibular afferent fibers*, J. Neurophysiol. **42**, 331–346 (1979)

[60] S. Yasui, *Stochastic functional Fourier series, Volterra series, and nonlinear systems analysis*, IEEE Trans. Auto. Control, in press (1979)

[61] L. R. Young and L. Stark, *Variable feedback experiments testing a sampled data model for eye tracking movements*, IEEE Trans. **HFE-4**, 38–51 (1963)