By the continuity of G_ϵ^{-1},

$$\lim_{\epsilon \to 0} x'(F^{-1}-G_\epsilon^{-1})x = x'(F^{-1}-G^{-1})x \geqslant 0$$

and $F^{-1}-G^{-1}$ is nonnegative definite.

References

[1] D. C. Fraser and J. E. Potter, "The optimum linear smoother as a combination of two optimum linear filters," *IEEE Trans. Automatic Control*, vol. AC-14, pp. 387–390, Aug. 1969.

[2] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: M.I.T. Press, 1974.

[3] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

[4] P. B. Liebelt, *An Introduction to Optimal Estimation*. Reading, MA: Addison-Wesley, 1967.

[5] K. S. Miller, *Complex Stochastic Processes*. Reading, MA: Addison-Wesley, 1974.

[6] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA J.*, vol. 3, no. 8, pp. 1445–1450, Aug. 1965.

Nonlinear Systems Analysis with Non-Gaussian White Stimuli: General Basis Functionals and Kernels

STANLEY KLEIN AND SYOZO YASUI, MEMBER, IEEE

***Abstract*—The Wiener–Lee–Schetzen scheme of using Gaussian white noise to test a nonlinear dynamical system is extended in two ways. 1) An arbitrary non-Gaussian white noise stationary signal can be used as the test stimulus. 2) An arbitrary function of this stimulus can then be used as the analyzing function for cross correlating with the response to obtain the kernels characterizing the system. Closed form expressions are given for the generalized orthogonal basis functions. The generalized kernels are expanded in terms of Volterra kernels and Wiener kernels. The expansion coefficients are closely related to the cumulants of the stimulus probability distribution. These results are applied to the special case of a Gaussian stimulus and a three-level analysis function. For this case a detailed analysis is made of the magnitude of the deviation of the kernels obtained with the ternary truncation as compared to the Wiener kernels obtained by cross correlating with the same Gaussian as was used for the stimulus. The deviations are found to be quite small.**

Introduction

A nonlinear analytic system can be described through a Volterra functional expansion [1]. Wiener [2] facilitated the practical usefulness of the functional expansion by introducing a set of orthogonal functions which completely characterize the system. Wiener's functionals and their associated kernels are constructed with respect to a Gaussian white noise input. Lee and Schetzen [3] showed how the various Wiener kernels could be measured by cross correlating the system's response with moments of the Gaussian noise input. The Wiener–Lee–Schetzen white noise method has been extensively applied to biological systems [4].

The purpose of this paper is twofold. First a formalism will be developed to handle the most general white noise test stimulus. There have already been several efforts to extend Wiener's scheme to non-Gaussian stimuli [4]–[13]. However, these authors do not clarify how the kernels obtained with non-Gaussian stimuli are related to the basic Volterra and Wiener kernels. Several of these authors [8], [11] are interested in the Cameron–Martin expansion rather than the Wiener expansion. Our study is unique in its focus upon the relationship between Wiener-like expansions. The simplicity of the leading terms in the expansion relating non-Gaussian kernels to the Gaussian (Wiener) kernels (see (20)) may help remove the stigma against using non-Gaussian stimuli.

A second purpose of this paper is to consider the case in which the output is cross correlated not with the stimulus, but with a nonlinear function of the stimulus. This case commonly occurs in practice since no stimulator is perfectly linear. The *intended* stimulus (used for cross correlation) may be a true Gaussian, for example, but the *actual* stimulus will be a truncated Gaussian due to physical limits on the upper and lower stimulus levels. It is shown here how the measured kernels depend upon the stimulus function (the actual input to the system) and upon the analysis function (used for cross correlating with the system's output). The use of an analysis function which differs from the stimulus function may furthermore be useful when rapid real-time calculations of higher order kernels are desired. For example, consider the real-time evaluation of a 50×50 element second-order kernel when the sampling time is 5 ms. This calculation requires at least one multiplication every 4 μs. Present computers are too slow. Replacing the Gaussian stimulus with its binary or ternary quantization allows all multiplications to be replaced by much more rapid additions and subtractions.

Since the signal used for cross correlation to obtain the kernel estimates may differ from the stimulus signal, there is the danger of losing the orthogonality of the expansion. We shall develop a new set of "dual-space" kernels and "dual-space" functionals which preserve orthogonality. These dual-space kernels will be expanded in terms of Volterra kernels and then related to Wiener kernels. The last section will consider the case where a Gaussian stimulus is used for testing and the ternary function is used for rapid computation. Factors contributing to the differences between the dual-space kernels and the Wiener kernels will be explored. It will be shown that the first- and second-order ternary kernels differ minimally from Wiener kernels.

Orthonormal Dual-Basis Functions

The output of an analytic time-invariant stable system, $y(t)$, can be related to its input $x(t)$ through the Volterra functional expansion [1]:

$$y(t)=\sum_{n,\tau_j}\delta^n g_n(\tau_1\cdots\tau_n)\prod_{i=1}^{n}x(t-\tau_i). \tag{1}$$

Several studies have examined the validity and convergence of the Volterra expansion for both deterministic and stochastic inputs [13]–[15]. In our formulation the values of the input and output are sampled every δ s, so the time variables τ_i and t are integers rather than continuous variables. Quantized rather than continuous time intervals are used, because they avoid the singularities associated with stochastic integrals [14] and they allow attention to be focused on diagonal kernel elements which are important for the results of this correspondence. Further-

Manuscript received January 28, 1977; revised November 28, 1978. This work was supported in part by Grant No. NS 03627 from the National Institutes of Health and the Alfred P. Sloan Foundation.

S. Klein is with the Division of Biology, California Institute of Technology, Pasadena, CA 91125 and with the Joint Science Department, Claremont Colleges, Claremont, CA.

S. Yasui was with the Division of Biology, California Institute of Technology, Pasadena, CA 91125. He is now with the National Institute for Basic Biology, 38 Nishigonaka, Myodijicho Okazaki, 444 Japan.

0018-9448/79/0700-0495$00.75

more, this is the formulation which is implemented on a digital computer.

If n_i of the time intervals τ_i in (1) are equal, then the power function $x^{n_i}(t-\tau_i)$ appears in the expansion. Instead of the Volterra expansion with $x^{n_i}(t-\tau_i)$ as the basis functions, it is possible to introduce an *orthogonal* basis $X_{n_i}(t-\tau_i)$:

$$y(t)=\sum_n y_n(t)$$

$$y_n(t)=\sum_{\tau_j}\delta^n h_n(\tau_1\cdots\tau_n)\prod_{i=1}^{k}X_{n_i}(t-\tau_i) \quad (2)$$

where n_i is the number of repetitions of τ_i in h_n, $n=\sum_{i=1}^{k}n_i$, and where k is the number of time intervals τ_i which are different. The summation is over all nonnegative values of τ_j.

The basis functions $X_n(t)$ can be constructed to be orthonormal to a dual set of basis functions $V_n(t)$ by a Gram–Schmidt orthogonalization procedure. The basis functions can be written as

$$X_n(t)=\det\left[M_n(u,w)\right]/\det\left[\overline{w_{n-1}M_{n-1}(u,w)}\right] \quad (3)$$

$$V_n(t)=\det\left[M_n(w,u)\right]/\det\left[\overline{u_nM_n(w,u)}\right] \quad (4)$$

where a bar over a quantity is the expectation operator or the time average of that quantity assuming ergodicity. The functions $u_i(t)$ and $w_i(t)$ are linear or nonlinear zero-memory time-invariant functions of $x(t)$ which should satisfy $\overline{u_iw_i}\neq 0$ and $|\overline{u_i^m w_i^{m'}}|<\infty$. The matrix $M_n(u,w)$ which assures orthogonality is given by

$$M_n(u,w)=\begin{bmatrix} 1 & \overline{u_1} & \cdots & \overline{u_n}\\ \overline{w_1} & \overline{w_1u_1} & \cdots & \overline{w_1u_n}\\ \cdots & \cdots & \cdots & \cdots\\ \overline{w_{n-1}} & \overline{w_{n-1}u_1} & \cdots & \overline{w_{n-1}u_n}\\ 1 & u_1(t) & \cdots & u_n(t)\end{bmatrix}. \quad (5)$$

The matrix $M(w,u)$ is the same as $M(u,w)$ but with u and w exchanged throughout.

The basis function given by (3) and (5) were originally proposed for the special case $u_i=w_i$ by Barrett [6]. We shall now demonstrate that also for the general case $u_i\neq w_i$ the basis functions (3)–(5) satisfy the orthonormality condition

$$\overline{X_n(t)V_{n'}(t+\Delta)}=\delta_{nn'}\delta_{\Delta 0} \quad (6)$$

where $\delta_{ij}=1$ for $i=j$, and $\delta_{ij}=0$ for $i\neq j$.

The expectation vanishes for $\Delta\neq 0$ because of the "whiteness" of the input stimulus. The whiteness condition means

$$\overline{(x(t)-\bar{x})(x(t+\Delta)-\bar{x})}=0, \qquad \text{for } \Delta\neq 0.$$

This condition implies

$$\overline{\left(X_n(t)-\overline{X}_n\right)\left(V_{n'}(t+\Delta)-\overline{V}_{n'}\right)}=0, \qquad \text{for } \Delta\neq 0.$$

The factors $\overline{X}_n$ and $\overline{V}_{n'}$ vanish (except for the trivial case $n=n'=0$) since the first and last row of the matrix $\overline{M}_n$ are equal, thereby causing det $\overline{M}_n$ to vanish in (3) and (4).

We now examine the case $\Delta=0$. In order to prove orthogonality for $n<n'$, $X_n(t)$ can be expanded:

$$\overline{X_nV_{n'}}=\sum_{i=0}^{n}a_i\overline{u_iV_{n'}}.$$

For $n'<n$, $V_n(t)$ can be expanded:

$$\overline{X_nV_{n'}}=\sum_{i'=0}^{n'}a'_{i'}\overline{X_nw_{i'}}.$$

However, $\overline{u_iV_{n'}}=0$ for $i<n'$ (and $\overline{X_nw_{i'}}=0$ for $i'<n$) since the expectation operator makes the bottom row of the determinant (5) equal to a preceding row, causing the determinant to vanish. Therefore, $\overline{X_nV_{n'}}=0$ for $n\neq n'$. For the case $n=n'$ we have $\overline{X_nV_n}=\overline{u_nV_n}$ because of the normalization factor in (3), and $\overline{u_nV_n}=1$ because of the normalization factor in (4). The determinant formalism is thus seen to be a natural method for enforcing orthonormality using general basis functions.

Throughout this paper the functions u_n and w_n will be chosen as follows: $u_n(t)=x^n(t)$ and $w_n(t)=v^n(t)$ where $v(t)$ is either equal to $x(t)$ or is a nonlinear zero-memory time-invariant function of $x(t)$ which satisfies $\overline{xv}\neq\bar{x}\bar{v}$. The analysis function $v(t)$ which is used for cross correlating with the response in order to obtain the system kernels need not be identical to the stimulus function $x(t)$. In this paper we examine the behavior of the system kernels for general stimulus distributions and general analysis functions. The basis functions $X_n(t)$ and $V_n(t)$ for $n\leqslant 4$ are tabulated in Table I.

In order for the orthonormal bases to be complete it is necessary that each function $x^n(t)$ and $v^n(t)$ be linearly independent of lower order functions. An example of an *incomplete* analysis basis is given by the multilevel function $v(t)$, where during each time interval, $v(t)$ is equal to one of n possible fixed scalar values v_i. For such a distribution $v^n(t)$ is linearly dependent on lower powers of $v(t)$ as shown by

$$\prod_{i=1}^{n}(v(t)-v_i)=0.$$

In the final section the case in which $v(t)$ has a ternary (3-level) distribution is considered. In this case, $V_n(t)$ for $n\geqslant 3$ becomes indeterminate (the normalization factor in the denominator vanishes as well as the numerator), so kernels with three or more repeated time indices are not calculable. An indeterminate kernel estimate means that the *variance* of the kernel estimate is infinite. The inability to calculate diagonal elements of high-order kernels is not a severe limitation, since most analyses focus on kernels with less than three time indices. The limitation would be severe, however, for nonlinear zero-memory systems where the diagonal elements contain all the information.

Relationship of Dual Kernels to Volterra and Wiener Kernels

The object of this section is to relate our dual kernels to the Volterra and the Wiener kernels. The functional expansion (2) can be inverted by cross correlating the response with the orthonormal analysis basis functions:

$$\delta^nH_n(\tau_1^{n_1}\cdots\tau_k^{n_k})\equiv\overline{y(t)\prod_{i=1}^{k}V_{n_i}(t-\tau_i)}. \quad (7)$$

The relationship between H_n and h_n can be found using (2) and (6):

$$H_n(\tau_1^{n_1}\cdots\tau_k^{n_k})=h_n(\tau_1\cdots\tau_n)n!/\prod_{i=1}^{k}n_i! \quad (8)$$

with $n\equiv\sum_{i=1}^{k}n_i$. The factorials give the number of ways a particular set of time intervals occurs in (2). Our use of H_n rather than h_n will eliminate most of the combinatorial factors in the forthcoming equations. The functional expansion (2) can be rewritten by introducing a time-ordered (TO) summation (time ordering is introduced to avoid further factorials):

$$y(t)=\sum_{k,n_j}\sum_{\tau_j}^{\text{TO}}H_n(\tau_1^{n_1}\cdots\tau_k^{n_k})\delta^n\prod_{i=1}^{k}X_{n_i}(t-\tau_i) \quad (9)$$

where $\sum_{\tau_j}^{\text{TO}}$ means that $\tau_j<\tau_{j+1}$. Similarly we can introduce a time-ordered Volterra expansion to rewrite (1) as

$$y(t)=\sum_{p,m_j}\sum_{\tau_j}^{\text{TO}}G_m(\tau_1^{m_1}\cdots\tau_p^{m_p})\delta^m\prod_{i=1}^{p}x^{m_i}(t-\tau_i) \quad (10)$$

TABLE I
ORTHOGONAL DUAL BASIS FUNCTIONS*

	Symmetric Case ($\overline{x^i v^j}$ = 0 for i + j odd)	General Case
X_0	1	1
$X_1(t)$	$x(t)$	$x(t)-\overline{x}$
$X_2(t)$	$x^2(t)-\overline{x^2}$	$x^2(t) - \frac{\overline{vx^2} - \overline{v}\,\overline{x^2}}{\overline{vx} - \overline{v}\,\overline{x}}\,x(t) - \frac{\overline{x^2}\,\overline{vx} - \overline{x}\,\overline{vx^2}}{\overline{vx} - \overline{v}\,\overline{x}}$
$X_3(t)$	$x^3(t) - \frac{\overline{vx^3}}{\overline{vx}}\,x(t)$	
$X_4(t)$	$x^4(t) - \frac{\overline{v^2x^4} - \overline{v^2}\,\overline{x^4}}{\overline{v^2x^2} - \overline{v^2}\,\overline{x^2}}\,x^2(t) + \frac{\overline{v^2x^4}\,\overline{x^2} - \overline{v^2x^2}\,\overline{x^4}}{\overline{v^2x^2} - \overline{v^2}\,\overline{x^2}}$	
V_0	1	1
$V_1(t)$	$v(t)/\,\overline{vx}$	$(v(t) - \overline{v})/(\overline{vx} - \overline{v}\,\overline{x})$
$V_2(t)$	$(v^2(t) - \overline{v^2})/(\overline{v^2x^2} - \overline{v^2}\,\overline{x^2})$	$\frac{(\overline{vx} - \overline{v}\,\overline{x})\,v^2(t) - (\overline{v^2x} - \overline{x}\,\overline{v^2})v(t) - (\overline{v^2}\,\overline{vx} - \overline{v}\,\overline{v^2x})}{\overline{v^2x^2}\,(\overline{vx} - \overline{v}\,\overline{x}) - (\overline{v^2x} - \overline{x}\,\overline{v^2})\,\overline{vx^2} - (\overline{v^2}\,\overline{vx} - \overline{v}\,\overline{v^2x})\,\overline{x^2}}$
$V_3(t)$	$(v^3(t) - \frac{\overline{v^3x}}{\overline{vx}}\,v(t))/(\overline{v^3x^3} - \frac{\overline{v^3x}\,\overline{vx^3}}{\overline{vx}})$	

*The basis functions X_n and V_n are presented in the upper and lower halves of the table, respectively. The basis functions for a general stimulus function $x(t)$ and general analysis function $v(t)$ are presented in the right column for $n \leq 2$. Basis functions for the special case of symmetric $x(t)$ and $v(t)$ (odd moments of x and v vanish) are presented in the left column.

where $m=\sum_{i=1}^{p} m_i$ and

$$G_m(\cdots)=g_m(\cdots)m!/\prod_{i=1}^{p} m_i!. \tag{11}$$

From (7) and (10) the dual-space kernels can be expanded in terms of Volterra kernels:

$$H_n(\tau_1^{n_1}\cdots\tau_k^{n_k})=\sum_{p,m_j}\delta^{m-n}\prod_{i=1}^{p}\overline{x^{m_i}V_{n_i}}\sum_{\tau_j,j>k}^{\mathrm{TO}}G_m(\tau_1^{m_1}\cdots\tau_p^{m_p}) \tag{12}$$

where $n_i=0$ for $i>k$. The time-ordered summation in which diagonal elements are treated differently from off-diagonal elements allows the expansion coefficients $\overline{x^m V_n}$ to be easily calculated using (4) and (5). For example:

$$\overline{x^m V_2}=\det\begin{vmatrix}1 & \overline{v} & \overline{v^2}\\ \overline{x} & \overline{vx} & \overline{v^2x}\\ \overline{x^m} & \overline{vx^m} & \overline{v^2x^m}\end{vmatrix}\Bigg/\det\begin{vmatrix}1 & \overline{v} & \overline{v^2}\\ \overline{x} & \overline{vx} & \overline{v^2x}\\ \overline{x^2} & \overline{vx^2} & \overline{v^2x^2}\end{vmatrix}. \tag{13}$$

The time-ordered summation (12), however, has the problem that the relative contribution of diagonal and off-diagonal terms is affected by the size of the sampling interval δ. In order to eliminate the dependence upon sampling interval, the time-ordered condition must be removed. The resulting summations would be continuous, and the remaining diagonal contributions would then be the same as the diagonal contributions resulting from an arbitrarily small sampling interval. As a first step towards a continuous summation, the time-ordered summation will be replaced by an exclusive summation (EX) in which each time delay τ_i can take on any value, except that it cannot equal any other time delay:

$$H_n(\tau_1^{n_1}\cdots\tau_k^{n_k})=\sum_r\frac{1}{r!}\sum_{m_j}\delta^{m-n}\prod_{i=1}^{k+r}\overline{x^{m_i}V_{n_i}}\sum_{\tau_j\neq\tau_{j'}}^{\mathrm{EX}}G_m(\tau_1^{m_1}\cdots\tau_{k+r}^{m_{k+r}}) \tag{14}$$

where $r=p-k$ is the number of distinct time intervals which are summed over. The factor $1/r!$ is needed to compensate for the multiple counting which occurs when the time-ordered restriction is eliminated.

The restriction $\tau_j\neq\tau_{j'}$ in (14) can be removed by introducing expansion coefficients Q_{mn}, and inclusive (IN) summations over τ_j, allowing the summations to be continuous:

$$H_n(\tau_1^{n_1}\cdots\tau_k^{n_k})=\sum_r\frac{1}{r!}\sum_{m_j}\prod_{i=1}^{k+r}Q_{n_im_i}\sum_{\substack{\tau_j\\ \text{for } j>k}}^{\mathrm{IN}}\delta^r G_m(\tau_1^{m_1}\cdots\tau_{k+r}^{m_{k+r}}). \tag{15}$$

The expansion coefficients $Q_{n_im_i}$ can be determined by comparing (14) and (15). In order to compare diagonal elements, it is first necessary to compensate for the normalization chosen in (11). This is most easily done by using (11) to reexpress both (14) and (15) in terms of g_n. Equating the coefficients of $g_m(\tau^m)$ in (14) and (15) for the case $k=1$ gives:

$$\delta^{m-n}\overline{x^m V_n}=\sum_r\frac{\delta^r}{r!}\sum_{m_j}m!\,\frac{Q_{nm_0}}{m_0!}\,\frac{Q_{0m_1}}{m_1!}\cdots\frac{Q_{0m_r}}{m_r!}$$

TABLE II
EXPANSION COEFFICIENTS FOR INCLUSIVE SUMMATION*

Q_{11}	$= 1$
Q_{01}	$= \overline{x} = 0$
Q_{02}	$= \overline{x^2\delta} \equiv P$
Q_{03}	$= \overline{x^3\delta^2}$
Q_{04}	$= (\overline{x^4} - 3\overline{x^2}^2)\delta^3$
Q_{05}	$= (\overline{x^5} - 10\,\overline{x^3}\,\overline{x^2})\delta^4$
Q_{06}	$= (\overline{x^6} - 15\,\overline{x^4}\,\overline{x^2} - 10\,\overline{x^3}^2 + 30\,\overline{x^2}^3)\delta^5$
Q_{07}	$= (\overline{x^7} - 21\,\overline{x^5}\,\overline{x^2} - 35\,\overline{x^3}\,\overline{x^4} + 140\,\overline{x^3}\,\overline{x^2}^2)\delta^6$
Q_{12}	$= \delta\overline{vx^2} / \overline{vx}$
Q_{13}	$= \delta^2(\overline{vx^3} - 3\,\overline{vx}\,\overline{x^2})/\overline{vx}$
Q_{14}	$= \delta^3(\overline{vx^4} - 4\,\overline{vx}\,\overline{x^3} - 6\,\overline{vx^2}\,\overline{x^2})/\overline{vx}$
Q_{15}	$= \delta^4(\overline{vx^5} - 5\,\overline{vx}\,\overline{x^4} - 10\,\overline{vx^2}\,\overline{x^3} - 10\,\overline{vx^3}\,\overline{x^2}$ $+ 30\,\overline{vx}\,\overline{x^2}^2)/\overline{vx}$
Q_{23}	$= (\overline{v^2x^3} - \overline{v^2}\,\overline{x^3})/(\overline{v^2x^2} - \overline{v^2}\,\overline{x^2})$

*The functions $x(t)$ and $v(t)$ need not be symmetric, but they have been normalized to have a zero mean.

where $m=\sum_{i=0}^r m_i$. An alternate expression is obtained by grouping together factors with equal m_i:

$$\delta^{m-n}\overline{x^m V_n} = \sum_{m_0} m!\frac{Q_{nm_0}}{m_0!}\sum_{r_j}\delta^r\prod_j\left((Q_{0j}/j!)^{r_j}/r_j!\right) \quad (16)$$

where $r=\sum_j r_j$ and $m=m_0+\sum_j j^{r_j}$.

The Q_{nm} can be determined iteratively by expressing Q_{nm} in terms of $Q_{n'm'}$ with $n' \leqslant n$ and $m' \leqslant m$. In order to simplify these calculations, the stimulus mean will be normalized to zero ($\bar{x}=0$) causing Q_{01} to vanish. For example, consider the case $m=5$, $n=1$:

$$\delta^4\overline{x^5V_1} = Q_{15} + \frac{\delta 5!}{4!}Q_{11}Q_{04} + \frac{\delta 5!}{3!2!}Q_{13}Q_{02} + \frac{\delta 5!}{2!3!}Q_{12}Q_{03} + \frac{\delta^2 5!}{2!^3}Q_{11}Q_{02}^2.$$

So

$$Q_{15} = \frac{\delta^4}{\overline{vx}}\left[\overline{vx^5} - 5\,\overline{vx}\left(\overline{x^4} - 3\,\overline{x^2}^2\right) - 10\left(\overline{vx^3} - 3\,\overline{vx}\,\overline{x^2}\right)\overline{x^2} - 10\,\overline{vx^2}\,\overline{x^3} - 15\,\overline{vx}\,\overline{x^2}^2\right].$$

The first few Q_{nm} have been calculated and are presented in Table II.

The expansion (16) is similar to the expansion of the characteristic function of a probability distribution in terms of the distribution cumulants. The expansion coefficients Q_{0m} (and Q_{1m} if $v=x$) are in fact the cumulants of x.

It is convenient to group together the lowest order terms (for each value of r) in the Volterra expansion (15) which are specified by $m_i=n_i$ for $i \leqslant k$:

$$H_n^W(\tau_1^{n_1}\cdots\tau_k^{n_k}) = \sum_r \frac{P^r}{r!}\sum_{\substack{\tau_i\\ i>k}}^{\mathrm{IN}}\delta^r G_m(\tau_1^{n_1}\cdots\tau_k^{n_k},\tau_{k+1}^2\cdots\tau_{k+r}^2). \quad (17)$$

The superscript W is chosen since the expansion (17) is precisely the expansion of a Wiener kernel in which the stimulus function and analysis function are Gaussian. The expansion (17) is obtained for Gaussian stimuli, since $Q_{nm}=0$ if $m>n>0$, $Q_{0m}=0$ for $m>2$, and $Q_{02}=P$, the power density. The usefulness of the normalizations (8) and (11) receives further support from the simplicity of (17). In terms of H^W, the Volterra expansion of a dual space kernel (15) becomes

$$H_n(\tau_1^{n_1}\cdots\tau_k^{n_k}) = \sum_r\frac{1}{r!}\sum_{\substack{m_j>3\\ \text{for } j>k}}\prod_{i=1}^{k+r}Q_{n_im_i}\sum_{\substack{\tau_j\\ \text{for } j>k}}\delta^r H_m^W(\tau_1^{m_1}\cdots\tau_{k+r}^{m_{k+r}}). \quad (18)$$

The difference between the Volterra-kernel expansion (15) and the Wiener-kernel expansion (18) is that in the latter m_j is restricted to be greater than 2 for $j>k$. The terms for $m_j=2$ are included within the Wiener kernel. For example:

$$\begin{aligned} H_1(\tau_1) = {} & \sum_{m_1>1}H_{m_1}^W(\tau_1^{m_1})Q_{1m} + \sum_{m_1>1}\sum_{m_2>3}Q_{1m_1}Q_{0m_2} \cdot \sum_{\tau_2}^{\mathrm{IN}}H_m^W(\tau_1^{m_1},\tau_2^{m_2})\delta \\ & + 1/2\sum_{m_1>1}\sum_{m_2>3}\sum_{m_3>3}Q_{1m_1}Q_{0m_2}Q_{0m_3} \cdot \sum_{\tau_2\tau_3}^{\mathrm{IN}}\delta^2 H_m^W(\tau_1^{m_1},\tau_2^{m_2},\tau_3^{m_3}) + \cdots \end{aligned} \quad (19)$$

$$= H_1^W(\tau_1) + \frac{\overline{vx^2}\,\delta}{\overline{vx}}H_2^W(\tau_1^2) + \left(\frac{\overline{vx^3}}{\overline{vx}} - 3\overline{x^2}\right)\delta^2 H_3^W(\tau_1^3) + \text{4th order terms.} \quad (20)$$

The Wiener kernels which appear in this expansion are for a Gaussian stimulus with the same power density $P=\overline{x^2\delta}$ as the original stimulus. When x is Gaussian and $v=x$, the dual kernels are identical to Wiener kernels.

The terms in (19) other than the first term will cause the dual-space kernel for general stimulus and analysis to differ from the Wiener kernels. In order to assess the magnitude of the deviation of the dual kernel from the Wiener kernel, it is useful to compare the size of the nonleading terms of (20) to the size of the leading term. The ratios of the sum of squares of the terms explicitly shown in (20) are

$$\sum_\tau\delta\left[\frac{\overline{vx^2}}{\overline{vx}}\delta H_2^W(\tau^2)\right]^2 \Big/ \sum_\tau\left[H_1^W(\tau)\right]^2\delta = \left(\frac{\overline{vx^2}^2}{\overline{vx}^2\,\overline{x^2}}\right)\left(\frac{\delta}{T_2}\right)\frac{\overline{y_2^2}}{\overline{y_1^2}} \quad (21)$$

$$\sum_\tau\delta\left[\left(\frac{\overline{vx^3}}{\overline{vx}} - 3\,\overline{x^2}\right)\delta^2 H_3^W(\tau^3)\right]^2 \Big/ \sum_\tau\left[H_1^W(\tau)\right]^2\delta = \left(\frac{\overline{vx^3}}{\overline{vx}\;\overline{x^2}} - 3\right)^2\left(\frac{\delta}{T_3}\right)^2\frac{\overline{y_3^2}}{\overline{y_1^2}} \quad (22)$$

where the response $y_n(t)$ from the nth order kernel was defined in (2). The mean square nth order response is given by

$$\overline{y_n^2} = P^n \sum_{\tau_i}^{\text{TO}} \delta^n \left[H_n^W(\tau_1 \cdots \tau_n) \right]^2.$$

The average integration times T_n are defined by

$$T_n^{n-1} \equiv \frac{\sum_{\tau_i}^{\text{TO}} \delta^n \left[H_n^W(\tau_1 \cdots \tau_n) \right]^2}{\sum_{\tau} \delta \left[H_n^W(\tau^n) \right]^2}.$$

The integration time T_n is a measure of the extent of significant off-diagonal terms in H_n. If, for example, H_n is a constant when $|\tau_i - \tau_{i'}| < \Delta$ for all $i \neq i'$ and vanishes further from the main diagonal ($|\tau_i - \tau_{i'}| > \Delta$), then $T_n = \Delta/2$.

The right sides of (21) and (22) have three factors which limit the discrepancy between Wiener kernels and dual kernels.

1) The factors $\overline{vx^2}/\overline{vx}(\overline{x^2})^{1/2}$ and $(\overline{vx^3}/\overline{vx}\,\overline{x^2}) - 3$ become vanishingly small as the stimulus and analysis distributions approach the Gaussian case.

2) The factor $\overline{y_n^2}/\overline{y_1^2}$ gives the ratio between the contribution of the nth order kernel and the first-order kernel to the mean square response. It is often found that the first-order contribution to the mean square response is greater than the *sum of all* higher order contributions.

3) The factor $(\delta/T_n)^{n-1}$ has special significance for our present considerations. The integration time T_n provides a measure of the temporal extent of significant contributions away from the main diagonal of the nth order kernel. By choosing the sampling time δ to be small, the terms shown in (21) and (22) can be made small, and the kernels obtained with a non-Gaussian stimulus or analysis function are very close to the kernels for the Gaussian case.

Change of Basis

How are the kernels for stimulus and analysis functions x and v_1 related to the kernels obtained with the *same stimulus* x, but a different analysis function v_2? The kernels H_n^1 for the first pair (x, v_1) can be expressed in terms of the kernels H_n^2 for the second pair (x, v_2) by expanding the first orthonormal basis in terms of the second:

$$V_n^{(1)}(t) = \sum_{m \geq n} V_m^{(2)}(t)\, \overline{X_m^{(2)} V_n^{(1)}}. \tag{23}$$

The terms with $m < n$ are not present since $\overline{x^m V_n} = 0$ for $m < n$. Inserting (23) into (7) leads to

$$H_n^1(\tau_1^{n_1} \cdots \tau_k^{n_k}) = \sum_{m_i} \delta^{m-n} \cdot \prod_{m_i > n_i} \overline{X_{m_i}^{(2)} V_{n_i}^{(1)}}\, H_m^2(\tau_1^{m_1} \cdots \tau_k^{m_k}). \tag{24}$$

This transformation is particularly simple since there is no summation over time.

For *symmetric* stimulus and analysis functions the leading terms of the expansion are

$$H_0^1 = H_0^2$$

$$H_1^1(\tau) = H_1^2(\tau) + H_3^2(\tau^3)\delta^2\, \overline{X_3^{(2)} V_1^{(1)}} + \cdots$$

$$H_2^1(\tau_1, \tau_2) = H_2^2(\tau_1, \tau_2) + \left[H_4^2(\tau_1, \tau_2^3) + H_4^2(\tau_1^3, \tau_2) \right] \delta^2\, \overline{X_3^{(2)} V_1^{(1)}} + \cdots$$

$$H_2^1(\tau^2) = H_2^2(\tau^2) + H_4^2(\tau^4)\delta^2\, \overline{X_4^{(2)} V_2^{(1)}} + \cdots.$$

TABLE III
Expectation Values for Ternary Analysis and Gaussian Stimulus*

$\overline{v^m x^n}$	$= \overline{vx^n}\, v_o^{(m-1)}$ for m odd
	$= \overline{v^2 x^n}\, v_o^{(m-2)}$ for m even
	$= \sigma^2 \dfrac{d(\sigma \overline{v^m x^{n-2}})}{d\sigma}$ for all m,n
	$= 0$ for m + n odd
$\overline{x^{2n}}$	$= \overline{x^2}^{\,n}\, 1 \cdot 3 \cdots (2n-1)$
$\overline{x^2}$	$= \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \int_0^\infty x^2 e^{-x^2/2\sigma^2} dx = \sigma^2$
$\overline{xv}$	$= \sqrt{\frac{2}{\pi}} \frac{v_o}{\sigma} \int_a^\infty x e^{-x^2/2\sigma^2} dx = \sqrt{\frac{2}{\pi}}\, v_o \sigma\, e^{-a^2/2\sigma^2}$
$\overline{v^2}$	$= \sqrt{\frac{2}{\pi}} \frac{v_o^2}{\sigma} \int_a^\infty e^{-x^2/2\sigma^2} dx = v_o^2\, \mathrm{erf}(a/\sigma)$
$\overline{vx^3}$	$= (a^2 + 2\sigma^2)\, \overline{vx}$
$\overline{v^2 x^2} - \overline{v^2}\, \overline{x^2}$	$= a\, v_o \overline{vx}$
$\overline{v^2 x^4} - \overline{v^2}\, \overline{x^4}$	$= a\, v_o (a^2 + 3\sigma^2) \overline{vx}$

*The ternary function $v(t)$ is given by (26). The zero-mean Gaussian $x(t)$ has a standard deviation, σ.

From Table I the leading coefficients are

$$\overline{X_3^{(2)} V_1^{(1)}} = \frac{\overline{x^3 v_1}}{\overline{x v_1}} - \frac{\overline{x^3 v_2}}{\overline{x v_2}}$$

$$\overline{X_4^{(2)} V_2^{(1)}} = \frac{\overline{v_1^2 x^4} - \overline{v_1^2}\, \overline{x^4}}{\overline{v_1^2 x^2} - \overline{v_1^2}\, \overline{x^2}} - \frac{\overline{v_2^2 x^4} - \overline{v_2^2}\, \overline{x^4}}{\overline{v_2^2 x^2} - \overline{v_2^2}\, \overline{x^2}}. \tag{25}$$

Gaussian-Ternary Dual Kernels

A simple example which illustrates the preceding formalism is the case in which $x(t)$ is zero-mean Gaussian and $v(t)$ is the ternary function of $x(t)$ given by

$$v(t) = \pm v_0, \quad \text{for } x(t) \gtrless \pm a$$
$$v(t) = 0, \quad \text{for } |x(t)| \leq a. \tag{26}$$

The value of v_0 is irrelevant (one can take $v_0 = 1$), since the normalizations of V_n and X_n eliminate all occurrences of v_0 from the dual kernels. The expectation values for ternary analysis and Gaussian stimulation are given in Table III.

The dual ternary kernels h^T are related to Wiener kernels by either (19) or by (24), (25):

$$h_1^T(\tau) = h_1^W(\tau) + \delta P\left(\frac{a^2}{\sigma^2} - 1\right) h_3^W(\tau^3) + \cdots$$

$$h_2^T(\tau_1, \tau_2) = h_2^W(\tau_1, \tau_2) + 2\delta P\left(\frac{a^2}{\sigma^2} - 1\right) \cdot \left[h_4^W(\tau_1, \tau_2^3) + h_4^W(\tau_1^3, \tau_2) \right] + \cdots$$

$$h_2^T(\tau^2) = h_2^W(\tau^2) + \delta P\left(\frac{a^2}{\sigma^2} - 3\right) h_4^W(\tau^4) + \cdots \tag{27}$$

where $\sigma^2 \equiv \overline{x^2}$, and $P = \sigma^2 \delta$ is the power density of the stimulus at low frequency.

Because of the following property of a ternary function:

$$v^n(t) = v(t)v_0^{n-1}, \qquad \text{for } n \text{ odd},$$

$$v^n(t) = v^2(t)v_0^{n-2}, \qquad \text{for } n \text{ even},$$

the kernels with three or more repeated time indices are indeterminate. This is because both the numerator and denominator of $V_n(t)$ vanish for $n \geqslant 3$.

The use of ternary instead of Gaussian analysis causes the dual kernels to deviate from the Wiener kernels in two respects. First, there are contributions from diagonal elements of high-order kernels (18)–(22) and (27). Second, because the analysis basis is incomplete, certain diagonal elements of the ternary kernels are indeterminate—having infinite variance. Both deviations involve diagonal elements with at least three repeated time indices.

The symmetric ternary truncation requires a choice of a/σ, the cutoff parameter, which relates the Gaussian distribution to the ternary threshold levels. There are several alternative choices for a/σ which can be justified on three different grounds.

1) Eliminate the contribution of $g_3(\tau\tau\tau)$ to $h_1(\tau)$. The condition $\overline{vx^3} - 3\overline{vx}\,\overline{x^2} = 0$ leads to $a/\sigma = 1$, since $\overline{vx^3}/\overline{vx} = a^2 + 2\sigma^2$ and $\overline{x^2} = \sigma^2$.

2) Normalize the second-order kernel diagonal elements to equal the normalization for off-diagonal elements. This item may be the most important, since a *filtered* Gaussian stimulus produces kernels without a well defined diagonal. The condition is $\overline{x^2v^2} - \overline{x^2}\,\overline{v^2} = 2\overline{vx}^2$ which leads to $(a/\sigma)^2 = 8/\pi e^{-a^2/\sigma^2}$. The solution of this transcendental equation is $a/\sigma \simeq 0.98$. This condition is essentially the same as the previous condition $a/\sigma = 1$.

3) Minimize the variance of the first-order kernel (16). The variance is proportional to the factor $f = \overline{x^2v^2}/\overline{vx}^2 = \pi/2e^{a^2/\sigma^2} \cdot \mathrm{erf}\,(a/\sigma)$. The correlation between x and v is given by $f^{-1/2}$. For $a = \sigma$ the factor f equals 1.37, which is 10 percent greater than the minimum value of f which occurs at $a = 0.7\sigma$. The factor $f = 1.37$ implies that an experiment must run 37 percent longer if the ternary rather than the original Gaussian function is used for cross correlation. This may be a small penalty to pay for the gain in computation speed.

The truncation value $a/\sigma = 1$ meets all three criteria satisfactorily, so this value can be used with the assurance that the kernels obtained have minimal systematic and statistical deviations from Wiener kernels.

References

[1] V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*. London and Glasgow: Blackie and Sons, 1930, Ch. 1.

[2] N. Wiener, *Nonlinear Problems in Random Theory*. New York: Wiley, 1958.

[3] Y. W. Lee and M. Schetzen, "Measurement of the Wiener kernels of a nonlinear system by cross correlation," *Int. J. Contr.*, vol. 2, pp. 237–254, 1965.

[4] G. D. McCann and P. Z. Marmarelis, Eds., *Proc. 1st Symp. Test. Ident. Nonlinear Syst.*, Calif. Inst. of Technology, Pasadena, March 1975.

[5] K. Ito, "Multiple Wiener integral," *J. Math. Soc. Japan*, vol. 13, pp. 157–169, 1951.

[6] J. F. Barrett, "The use of functionals in the analysis of nonlinear physical systems," *J. Electron Contr.*, vol. 15, pp. 567–615, 1963.

[7] T. Hida and N. Ikeda, "Analysis on Hilbert space with reproducing kernel arising from multiple Wiener integrals," in *Proc. 5th Berkeley Symp. Math. Stat. Prob.*, Univ. of Calif. Press, vol. 2, Part 1, pp. 117–143, 1965.

[8] D. B. Brick, "Wiener's non-linear expansion procedure applied to cybernetic problems," *IEEE Trans. Syst. Sci. Cybern.*, vol. 1, pp. 67–74, 1965.

[9] H. Ogura, "Orthogonal functions of the Poisson process," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 473–481, 1972.

[10] H. I. Krausz, "Identification of nonlinear systems using random impulse train inputs," *Biolog. Cybern.*, vol. 19, pp. 217–230, 1975.

[11] A. Segall and T. Kailath, "Orthogonal functionals of independent increment processes," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 287–298, 1976.

[12] V. Z. Marmarelis and G. D. McCann, "A family of quasi-white random signals and its optimal use in biological systems identification, Part I: Theory; Part II: Application of the photoreceptor of Calliphora erythrocephala," *Biolog. Cybern.*, vol. 27, pp. 41–82, 1977.

[13] Y. H. Ku and A. A. Wolf, "Volterra–Wiener functionals for the analysis of nonlinear systems," *J. Franklin Inst.*, vol. 3, pp. 9–26, 1966.

[14] G. Palm and T. Poggio, "The Volterra representation and the Wiener expansion: Validity and pitfalls," *SIAM J. Appl. Math.*, vol. 33, pp. 195–216, 1977.

[15] E. Gilbert, "Functional expansions for the response of nonlinear differential systems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 909–921, 1977.

[16] S. Yasui and S. Klein, "Clipping technique for measurement of Wiener-like Fourier kernels," Calif. Inst. of Technology preprint, 1979.

Comments on "Binary Single-Sideband Phase-Modulated Communication Systems"

S. K. MANOCHA

In the above paper,[1] equations (8) and (9) should have read

$$\exp\left[\hat{f}_0(t)\right]\sin f_0(t) = -\exp\left[\hat{f}_1(t)\right]\sin f_1(t), \tag{1}$$

$$\exp\left[\hat{f}_0(t)\right]\cos f_0(t) = -\exp\left[\hat{f}_1(t)\right]\cos f_1(t). \tag{2}$$

Squaring and adding (1) and (2) gives

$$\exp\left[2\hat{f}_0(t)\right] = \exp\left[2\hat{f}_1(t)\right],$$

or

$$\hat{f}_0(t) = \hat{f}_1(t). \tag{3}$$

This implies that

$$f_0(t) = f_1(t) + K, \tag{4}$$

where K is an arbitrary constant. Equation (4) holds since the Hilbert transform of a constant is zero. Also, by dividing (1) and (2),

$$\tan f_0(t) = \tan f_1(t)$$

$$f_0(t) = f_1(t) + (2n+1)\pi, \tag{5}$$

which verifies that $f_0(t)$ and $f_1(t)$ are equal to within an additive constant $(2n+1)\pi$. Thus equation (6)[1] can be satisfied by such functions $f_0(t)$ and $f_1(t)$, and it is possible to obtain antipodal signals using single-sideband phase modulated formulation.

Author's Reply[2, 3]

I believe that Mr. Manocha is correct in his conclusion that it is possible to obtain antipodal signals with the single-sideband phase modulated (SSBPM) formulation. The basic conclusion of the paper concerning the suboptimum performance of an SSBPM system with square modulation remains, however. If Mr. Manocha is pursuing this subject, it would be of interest to see if he could discover a class of modulating waveforms which would minimize both the bandwidth and the intersymbol interference in a practical digital SSBPM system.

Manuscript received June 12, 1978.

The author is with the Research and Development Department, Indian Telephone Industries, Ltd., Naini, Allahabad 211010, India.

[1]H. D. Chadwick, *IEEE Trans. Info. Theory*, vol. IT-18, pp. 214–215, January 1972.

[2]Manuscript received August 1, 1978

[3]The author is with Satellite Business Systems, 8003 Westpark Dr., McLean, VA 22102.