VERIFICATION OF A METHOD TO ESTIMATE THE WIENER KERNEL TRANSFORMS OF A NONLINEAR SYSTEM

Peter D. Burns

COPYRIGHT NOTICE

P. D. Burns, 'Verification of a Method to Estimate the Wiener Kernels of a Nonlinear System,' M. Eng. Thesis, Electrical and Computer Eng., Clarkson University, 1977.

Published by the author

CLARKSON COLLEGE OF TECHNOLOGY*

DEPARTMENT OF ELECTRICAL ENGINEERING

Verification of a Method to Estimate the Wiener Kernel Transforms of a Nonlinear System

A Thesis

by

Peter D. Burns

Submitted in partial fulfillment of the requirements

for the degree of

Master of Engineering

(Electrical Engineering)

April 13, 1977

*Clarkson University

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont'd)

LIST OF TABLES

Verification of a Method to Estimate the Wiener Kernel Transforms of a Nonlinear System

Peter D. Burns

ABSTRACT

A method for estimating the Fourier transforms of the Wiener kernels of a nonlinear system is presented. In order to verify the method, it is applied to various systems in several experiments. The estimates for the Fourier transforms of the first, second and third kernels are obtained and found to converge to the theoretical kernel transforms. The number of calculations necessary for a desired mean square error of the experimental estimates is considered in comparison with a time domain method for Wiener kernel estimation.

ACKNOWLEDGMENTS

The author wishes to thank the following faculty members for their invaluable help in this endeavor; J. Koplowitz, R. Mukundan and R. Meyer.

1.0 INTRODUCTION

In general, a system is specified by an input — output relationship. Knowledge of the output which due to a known input is available if one has a mathematical description of the system. Such a description is a useful tool for both analysis and design.

Any system can be classified as either linear or nonlinear. A nonlinear system is defined as one where superposition does not hold. The linear time invariant system has been extensively studied and is usually described by its impulse response or, in the frequency domain, the system function. The mathematical system description used here is a generalization of the impulse response, the set of Wiener kernels. Only single-valued, shift (time)-invariant, zero-memory systems will be addressed. Any interconnection of such linear and nonlinear systems can be described by the set of Wiener kernels.

The kernel description represents a general method for nonlinear system analysis. Previous methods have generally had more restricted applications [1]. Transient descriptions of systems are available through solution of nonlinear differential equations. Solutions, however, are not readily available for third- or higher-order equations. The graphical phase-plane method can be used to solve nonlinear differential equations of no higher than the second order. The "describing function" method can be used for steady-state nonlinear system analysis. This method, however, neglects the effects of harmonics in the system and therefore does not have general application.

Nonlinear systems have also been described in terms of linear

approximations and total harmonic distortion. Such methods depend on the system's response being essentially linear over a given operating range.

Since nonlinear systems occur frequently in many practical applications it is appropriate to discuss examples of where the Wiener kernel description is of value. Consider a transistor whose d.c. transfer characteristics are parabolic rather than linear. Instead of concerning oneself with nonlinear system analysis one may choose to merely specify the percent total harmonic distortion. Harmonic distortion will, however, vary with input signal amplitude and therefore does not completely describe the system response. The output signal may be thought of as the system's output owing to the given input, rather than a distorted linear response. Furthermore, if a mathematical model is available to describe the system, its output can be predicted from a knowledge of the input.

A similar situation occurs in the large-area (low-frequency) characteristics of a photographic process [2]. Instead of input and output being voltage and current values (as with the previous case), one deals with optical reflectance, transmittance and density values. A typical transfer characteristic curve of one such process is shown in Fig. 1. If such a system is considered linear (eg. through a small signal assumption), the terms optical transfer function (O.T.F.) and point spread function are directly analogous to the system function and impulse response of the electronic system.

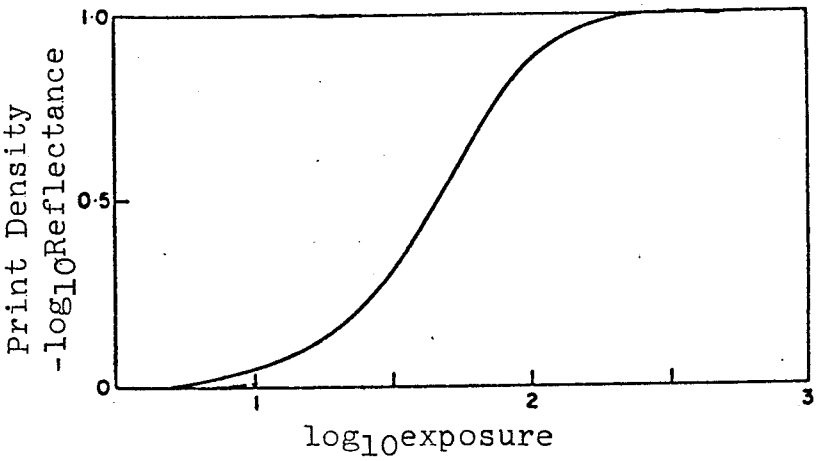


FIGURE 1. Typical characteristic curve for photographic film

The objective of this work is to verify a method for estimating the Wiener kernel transforms of a nonlinear system. In Section 2.0 the Volterra and Wiener kernel system descriptions are presented. Lee and Schetzen [3] developed estimates for the Wiener kernels based on a systems output which is due to a white Gaussian input and their work is described in Section 3.0. Also in Section 3.0, estimates for the Fourier transforms of the Wiener kernels are presented. French and Butz [4] first noted the form of these estimates and they were later more clearly developed by Koplowitz.

Experiments to verify the frequency domain Wiener kernel estimates are described in Section 4.0 where results are also given. Consideration of the kernel transform estimate variance, and comparison of the computation necessary for time and frequency domain estimation are given in Section 5.0.

2.0 VOLTERRA AND WIENER SERIES EXPANSION

A linear shift invariant system is described by its impulse response and its Fourier transform, the system function. The system input — output relationship is given by the convolution integral and its Fourier transform. The nonlinear descriptions used here are functional series involving generalizations of the linear system descriptors.

2.1 VOLTERRA SERIES EXPANSION

In the same way that a function operates on a variable x a functional operates on a function f(x). A functional series studied by Volterra [5] described the output of a nonlinear system. The terms of the series are n-fold convolution integrals, functionals, which operate on a set of multidimensional impulse responses, the Volterra kernels.

The output of a nonlinear system expressed by the Volterra series is

$$y(t) = h_o + \int h_1(\tau)x(t-\tau)d\tau + \iint h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2 + \iiint h_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)x(t-\tau_3)d\tau_1,d\tau_2 d\tau_3 + \ldots \tag{2.1}$$

where y(t) is the system output, x(t) the input and h_n the set of Volterra kernels. The limits of integration are to be taken from $-\infty$ to $+\infty$ unless otherwise indicated. The system output, y(t), is seen as the sum of terms of the form

$$\int \cdots \int h_n(\tau_1, \ldots, \tau_n) x(t-\tau_1) \ldots x(t-\tau_n) d\tau_1 \ldots d\tau_n \tag{2.2}$$

which is the n-fold convolution integral.

The successful use of the Volterra functional system representation depends on the ability to represent the nonlinearity by a power series or polynomial. For example, if the system nonlinearity could be described by the power series

$$y(x) = a_o + a_1 x + a_2 x^2 + \ldots a_n x^n \tag{2.3}$$

then the system can be described by the Volterra kernels of order up to n. The more violent the system nonlinearity, the further the series of (2.1) must be summed for close convergence to y(t). The d.c. component of y(t) not due to x(t) is h_o. This is not the total d.c. bias of the output, since a zero mean input to a square law device results in a nonzero mean output. The d.c. output component which is due to this element appears as a result of the second-order Volterra functional.

A first-order system output would be given by

$$y(t) = h_o + \int h_1(\tau) x(t-\tau) d\tau. \tag{2.4}$$

The second term on the right-hand side is the convolution integral of a linear system and h_1 is the impulse response. For a linear system, since the convolution integral specifies the output, the only Volterra kernel of (2.1) that is nonzero is h_1. This is because, for such a system, the only constant of (2.3) not equal to zero is h_1. If the output of a system is given by

$$y(t) = \int\int h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2 \tag{2.5}$$

this results from a quadratic (squaring) nonlinearity and is seen as a two-dimensional convolution integral.

A linear system can be described in both the time and frequency domains and so, too, can a nonlinear system. Just as the linear system function is of interest, so is the Volterra kernel transform

$$H_n(w_1,\ldots,w_n) = \int\ldots\int h_n(\tau_1,\ldots,\tau_n)e^{-j(w_1\tau_1+\ldots+w_n\tau_n)}d\tau_1\ldots d\tau_n \tag{2.6}$$

where H_n is the n-dimensional Fourier transform of h_n.

Just as the kernels are used to determine system response to an input, so can kernel transforms. Express the second Volterra kernel transform in polar notation,

$$H_2(w_1,w_2) = |H_2(w_1,w_2)|e^{j\phi(w_1,w_2)} \tag{2.7}$$

where $|H|$ and ϕ are the magnitude and phase angle, respectively. Thomas [6] has shown that the output which is due to the second-order functional for an input

$$x(t) = \cos w_1 t \tag{2.8}$$

neglecting dc terms is

$$y(t) = \left|\frac{H_2(w_1,w_1)}{8\pi^2}\right| \cos\left(2w_1 t + \phi(w_1,w_1)\right). \tag{2.9}$$

If the input is the sum of two cosines,

$$x(t) = \cos w_1 t + \cos w_2 t \tag{2.10}$$

the corresponding output is the sum of four terms,

$$\begin{aligned} y(t) = &\left|\frac{H_2(w_1,w_1)}{8\pi^2}\right| \cos\left(2w_1 t + \phi(w_1,w_1)\right) \\ &+ \left|\frac{H_2(w_2,w_1)}{4\pi^2}\right| \cos\left((w_1 + w_2)t + \phi(w_2,w_1)\right) \\ &+ \left|\frac{H_2(-w_2,w_1)}{4\pi^2}\right| \cos\left((w_1 - w_2)t + \phi(-w_2,w_1)\right) \\ &+ \left|\frac{H_2(w_2,w_2)}{8\pi^2}\right| \cos\left((2w_2 t + \phi(w_2,w_2)\right) \quad . \end{aligned} \tag{2.11}$$

Consider the system of Fig. 2 consisting of a linear subsystem, A, followed by a nonlinear operator, x^n. The impulse response of A is a(t). The output of a is

$$y(t) = \int a(\tau)\ x\ (t-\tau)d\tau. \tag{2.12}$$

The output of the nonlinear element is

$$y^n(t) = \left[\int a(\tau)\ x\ (t-\tau)d\tau\right]^n \tag{2.13}$$

or

$$g_1(t) = \int \cdots \int a(\tau) \cdots a(\tau_n) x\ (t-\tau_1) \cdots x(t-\tau_n) d\tau_1 \cdots d\tau_n \ .$$

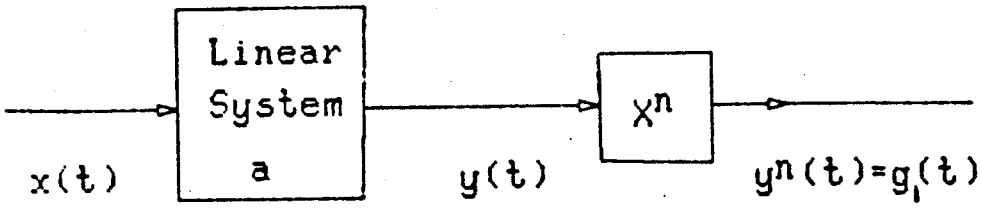


FIGURE 2. Nonlinear system of nth order

However $g_1(t)$ can also be expressed by the Volterra series expansion. Equating (2.1) and (2.13) gives the Volterra kernels,

$$h_m = 0, \text{ for } m \neq n \tag{2.14}$$

$$h_n(\tau_1,\tau_2\ldots\tau_n) = a(\tau_1)\ a(\tau_2)\ldots a(\tau_n)$$

The corresponding kernel transform using (2.6) is given by

$$H_n(w_1,\ldots,w_n) = A(w_1)\ldots A(w_n) \tag{2.15}$$

where A(w) is the Fourier transform of a(t).

If the linear system A has an impulse response

$$a(\tau) = e^{-a\tau}, \tag{2.16}$$

which is the case for a capacitor, the nth Volterra kernel is given by

$$h_n(\tau_1,\ldots,\tau_n) = e^{-a(\tau_1 + \ldots + \tau_n)}. \tag{2.17}$$

The kernel transform is

$$H(w_1,\ldots,w_n) = \frac{a^n}{(a-jw_1)\ldots(a-jw_n)}. \tag{2.18}$$

Now consider the system of Fig. 3 which consists of the previous system cascaded with a linear system B. The impulse responses of A and B are a(t) and b(t) respectively. The following notation is used for the Volterra kernel and kernel transforms of the system A followed by the nonlinear element, the nth Volterra kernel is given by

$$a_n(\tau_1,\ldots,\tau_n) = a(\tau_1)\ldots a(\tau_n)$$

$$A_n(w_1,\ldots,w_2) = A(w_1)\ldots A(w_2) \tag{2.19}$$

The system output is

$$g_2(t) = \int b(\tau) g_1(t-\tau)d\tau, \tag{2.20}$$

using (2.13) and (2.19)

$$g_1(t-\tau) = \int\ldots\int a_n(\tau_1-\tau,\ldots,\tau_n-\tau) \times (t-\tau_1-\tau)\ldots$$

$$x(t-\tau_n-\tau)d\tau_1\ldots d\tau_n. \tag{2.21}$$

substituting (2.21) into (2.20) yields

$$g_2(t) = \int\ldots\int b(\tau)a_n(\tau_1-\tau,\ldots,\tau_n-\tau) \times (t-\tau_1-\tau)\ldots$$

$$x(t-\tau_n-\tau)d\tau_1\ldots d\tau_n d\tau \tag{2.22}$$

The system output, $g_2(t)$ can be expressed by its Volterra series expansion. The output which is due to the nth kernel is

$$\int\ldots\int h_n(\tau_1,\ldots,\tau_n) \times (t-\tau_1)\ldots \times (t-\tau_n)d\tau_1\ldots d\tau_n \ . \tag{2.23}$$

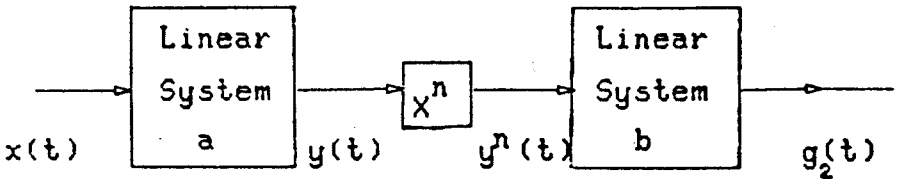


FIGURE 3. Nonlinear system of Fig. 2 cascaded with a linear subsystem

After equating (2.22) and (2.23), h_n is found to be similar in form to the right-hand side of (2.20),

$$h_n(\tau_1,\ldots,\tau_n) = \int b(\tau)a_n(\tau_1-\tau,\ldots,\tau_n-\tau)d\tau. \tag{2.24}$$

This is the nth Volterra kernel of the system of Fig. 3. From (2.1) the system output is seen as the sum of terms which are integrals involving the set of Volterra kernels. Equation (2.22) shows, however, that the nth-order term (functional) is the complete output, $g_2(t)$. From this one concludes

$$h_m = 0 \tag{2.25}$$

for $m \neq n$.

The Volterra kernel transform corresponding to (2.24) can be found as follows:

$$H_n(w_1,\ldots,w_n) = \int\ldots\int h_n(\tau_1,\ldots,\tau_n)e^{-j(w_1\tau_1+\ldots+w_n\tau_n)}d\tau_1\ldots d\tau_n \tag{2.26}$$

$$= \int\ldots\int b(\tau)a_n(\tau_1-\tau,\ldots,\tau_n-\tau)e^{-j(w_1\tau_1+\ldots+w_n\tau_n)}$$

$$d\tau_1\ldots d\tau_n d\tau. \tag{2.27}$$

Integrating the right-hand side of (2.26) with respect to $\tau_1, \ldots \tau_n$ gives

$$H_n(w_1,\ldots,w_n) = \int b(\tau) A_n(w_1,\ldots,w_n) e^{-j(w_1+\ldots+w_n)\tau} d\tau \qquad (2.28)$$

where $A_n(w_1,\ldots,w_n)$ is the n-dimensional Fourier transform of $a_n(\tau_1, \ldots,\tau_n)$. Integrating (2.27) with respect to τ gives

$$H_n(w_1,\ldots w_n) = B(w_1+\ldots+w_n)A_n(w_1,\ldots,w_n). \qquad (2.29)$$

Substituting (2.19) into (2.24) and (2.29) shows the Volterra kernel and kernel transform in terms of the impulse responses and system functions of A and B,

$$h_n(\tau_1,\ldots,\tau_n) = \int b(\tau) a(\tau_1-\tau)\ldots a(\tau_n-\tau) d\tau_1 \ldots d\tau_n d\tau \qquad (2.30)$$

$$H_n(w_1,\ldots,w_n) = B(w_1+\ldots+ w_n)A(w_1)\ldots A(w_n). \qquad (2.31)$$

As an example of the above system configuration, consider the case when

$$\begin{aligned} a(\tau) &= e^{-a\tau} \\ b(\tau) &= e^{-b\tau} \end{aligned} \qquad (2.32)$$

The nth-order Volterra kernel is

$$h_n(\tau_1,\ldots,\tau_n) = \int e^{-b\tau} e^{-a(\tau_1+\ldots+\tau_n)} e^{an\tau} d\tau \qquad (2.33)$$

and the corresponding kernel transform is

$$H_n(w_1,\ldots,w_n) = \frac{a^n}{(a-jw_1)\ldots(a-jw_n)} \cdot \frac{b}{b - j(w_1+\ldots+w_n)} . \qquad (2.34)$$

2.2 WIENER SERIES EXPANSION

Wiener [7] developed a functional series based on Volterra's work for the analysis of a nonlinear system response to a white Gaussian input. The functionals of Wiener's expansion are an orthogonal set of Volterra functionals which operate on the system Wiener kernels.

The system output expressed by the Wiener series expansion is

$$\begin{aligned} y(t) = {} & k_o + \int k_1(\tau)x(t-\tau)d\tau \\ & + \int\int k_2(\tau_1,\tau_2)x(t-\tau_2)d\tau_1 d\tau_2 \\ & - C\int k_2(\tau,\tau)d\tau \\ & + \int\int\int k_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 \\ & - 3C \int\int k_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau, + \ldots \end{aligned} \tag{2.35}$$

where the input power density spectrum is C watts/Hz, and $\{k_n\}$ is the set of Wiener kernels. Equation (2.35) can also be expressed as

$$y(t) = \sum_{n=o}^{\infty} G_n\left[k_n, x(t)\right] \tag{2.36}$$

where $\{G_n\}$ is the set of orthogonal Wiener functionals. The corresponding Volterra and Wiener functionals are listed in Table 1. The Wiener functionals are orthogonal and this is expressed as

$$E\left[G_n \cdot G_m\right] = 0 \tag{2.37}$$

TABLE 1

VOLTERRA AND WIENER FUNCTIONALS

Order	Functional
0	$\int h_o(\tau)d\tau$
1	$\int h_1(\tau)x(t-\tau)d\tau$
2	$\iint h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2$
3	$\iiint h_3(\tau_1,\tau_2,\tau_3)x(t-\tau)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3$

Wiener Expansion

Order	Functional
0	$\int k_o(\tau)d\tau$
1	$\int k_1(\tau)x(t-\tau)d\tau$
2	$\iint k_2(\tau_1,\tau_2)x(t-\tau)x(t-\tau_2)d\tau_1 d\tau_2$
	$- C\int k_2(\tau_2,\tau_2)d\tau_2$
3	$\iiint k_3(\tau_1,\tau_2,\tau_3)x(t-\tau)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3$
	$-3C\iint k_3(\tau_1,\tau_2\tau_3)x(t-\tau)d\tau_1 d\tau_2$

The power density spectrum of the input process is C watts/Hz.

To illustrate the differences between the Volterra and Wiener kernels which result from the orthogonality property, consider the system of Fig. 4. From (2.14) the output is only due to the third Volterra functional,

$$y(t) = \iiint h_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 \quad . \tag{2.38}$$

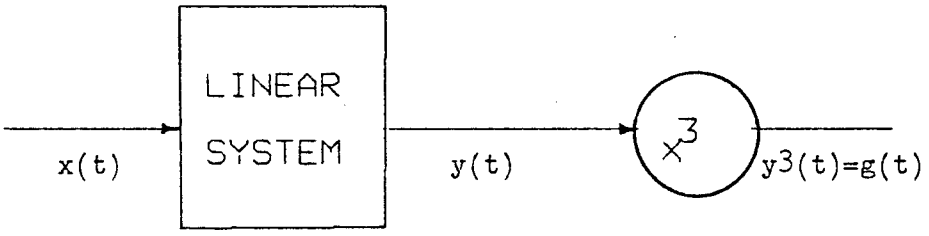


FIGURE 4. Third-order nonlinear system.

Consider the system output expressed by the Wiener functional expansion operating on the Volterra kernels

$$\begin{aligned} y(t) \stackrel{?}{=} & \iiint h_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_1)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 \\ & - 3C\iint h_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau_1 \quad . \end{aligned} \tag{2.39}$$

To satisfy (2.38), (2.39) becomes

$$\begin{aligned} y(t) = & \iiint h_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 \\ & - 3C\iint h_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau_1 + 3C\iint h_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau_1 \end{aligned} \tag{2.40}$$

Comparison of Eqs (2.35) and (2.40) shows that the last term added has the form of the first Wiener functional. In fact, if

$$\left.\begin{aligned} k_3(\tau_1,\tau_2,\tau_3) &= h_3(\tau_1,\tau_2,\tau_3) \\ k_1(\tau) &= 3C\int h_3(\tau,\tau_1,\tau_1)d\tau_1 \\ k_2(\tau_1,\tau_2) &= k_o = 0 \end{aligned}\right\} \tag{2.41}$$

Eq. (2.40) becomes

$$\begin{aligned} y(t) = &\int k_1(\tau)x(t-\tau)d\tau \\ &+ \iiint k_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 \\ &- 3C\int k_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau_1 \end{aligned} \tag{2.42}$$

Equations (2.41) satisfy the Volterra (2.38) and Wiener (2.42) expansions of the system output. For the above system, the first Wiener functional

$$\int k_1(\tau)x(t-\tau)d\tau \tag{2.43}$$

can be thought of as the linear component present at the output owing to the cubic nonlinearity.

Similarly, for a system whose output is due only to the second Volterra functional,

$$\begin{aligned} k_2(\tau_1,\tau_2) &= h_2(\tau_1,\tau_2) \\ k_1(\tau) &= 0 \\ k_o &= \int h_2(\tau_1,\tau)d\tau \end{aligned} \tag{2.44}$$

Another property results from the orthogonality of the Wiener functional expansion. The output of a system can be represented, from (2.36), as

$$y(t) = \lim_{n\to\infty}\left[G_o\left[k_o,x\right] + G_1\left[k_1,x\right] + \ldots + G_n\left[k_n x\right]\right]. \tag{2.45}$$

Since $\{G_n\}$ is a set of orthogonal functions, previously determined terms are not changed by taking more terms in the approximation to $y(t)$ [7]. It should also be noted that each G_n operates only on the kernel of order n. Therefore if one were to obtain successive Wiener kernels, for every one known the corresponding functional of (2.52) can be found.

Wiener also shows that the Wiener kernels are, or can be made symmetrical with respect to their arguments. Suppose $k_2(\tau_1,\tau_2)$ of Eq. (2.35) is not symmetrical with respect to τ_1 and τ_2. The Wiener kernel can be made symmetrical by interchanging τ_1 and τ_2, adding this to the original and dividing by 2. This has not changed $y(t)$. The same argument can be applied to Volterra kernels.

3.0 WIENER KERNEL AND KERNEL TRANSFORM ESTIMATES

3.1 WIENER KERNEL ESTIMATES

Lee and Schetzen [3] present a method for estimating the Wiener kernels by cross correlation. They introduce a set of functionals formed from delay circuits with a white Gaussian input. The input process is ergodic and the nonlinear system is time invariant. Figure 5 shows the circuit used for estimating $k_2(\tau_1,\tau_2)$, the second Wiener kernel.

The output of the delay circuit is given by

$$y_2(t) = x(t-\sigma_1)x(t-\sigma_2). \tag{3.1}$$

The nonlinear system output is, by the Wiener expansion

$$\begin{aligned} y(t) = k_o &+ \int k(\tau)x(t-\tau)d\tau \\ &+ \iint k_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau)d\tau_1 d\tau_2 \\ &- C\int k_2(\tau,\tau)d\tau \end{aligned} \tag{3.2}$$

The outputs of the delay circuits are multiplied to form $y_2(t)$. The output of the nonlinear system is then multiplied by $y_2(t)$ and averaged. The signal $y(t)y_2(t)$ (using (3.1) and (3.2)) is

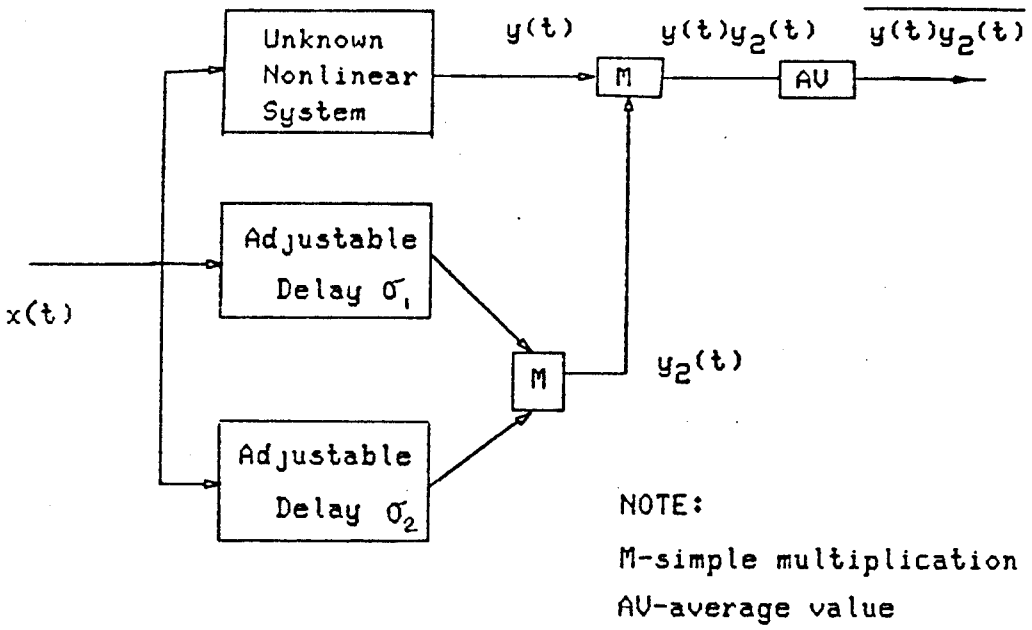


FIGURE 5. Measurement circuit for second Wiener kernel estimate

$$y(t)y_2(t) = \left[k_o + \int k(\tau)x(t-\tau)d\tau - \right.$$

$$+\iint k_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2$$

$$\left. - C\int k_2(\tau,\tau)d\tau\right] x(t-\sigma_1)x(t-\sigma_2) \qquad (3.3)$$

The average of $y(t)y_2(t)$ is

$$\overline{y(t)y_2(t)} = k_o\overline{x(t-\sigma_1)x(t-\sigma_2)} + \int k_1(\ )x(\overline{t-\tau_1})x(t-\sigma_1)x(t-\sigma_2)d\tau$$

$$+\iint k_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)x(t-\sigma_1)x(t-\sigma_2)d\tau_1 d\tau_2$$

$$- C\int k_2(\tau,\tau)x\overline{(t-\sigma_1)x(t-\sigma_2)}d\tau \qquad (3.4)$$

where $\bar{x}$ indicates the average value. Several terms of (3.4) are zero because the average of the product of an odd number of zero mean Gaussian random variables is zero. Using the properties of the white input and symmetric Wiener kernels (3.4) becomes, after several steps,

$$\overline{y(t)y_2(t)} = 2C^2 k_2(\sigma_1,\sigma_2) + C k_o \delta(\sigma_1-\sigma_2) \tag{3.5}$$

where $\delta(t)$ is the unit impulse. Since the impulse function is zero except where σ_1 equals σ_2, (2.5) becomes

$$\overline{y(t)y_2(t)} = 2C^2 k_2(\sigma_1,\sigma_2) \tag{3.6}$$

for $\sigma_1 \neq \sigma_2$. Noting that the input process is ergodic, the nonlinear system time invariant (2.6) can be rewritten

$$k_2(\sigma_1,\sigma_2) = \frac{1}{2C^2} E\left[y(t)x(t-\sigma_1)x(t-\sigma_2)\right] \tag{3.7}$$

where E indicates the expected value.

In the same manner, the nth kernel can be estimated using n delays. The first four Wiener kernel estimates (which are due to Lee and Schetzen) are

$$k_o = E[y(t)]$$

$$k_1(\tau) = \frac{1}{C} E[y(t)x(t-\tau)]$$

$$k_2(\tau_1,\tau_2) = \frac{1}{2C^2} E[y(t)x(t-\tau_1)x(t-\tau_2)] \tag{3.8}$$

for $\tau_1 \neq \tau_2$

$$k_3(\tau_1,\tau_2,\tau_3) = \frac{1}{6C^3} E\left[y(t)\, x\,(t-\tau_1)\, x\,(t-\tau_2)\, x\,(t-\tau_3)\right]$$

for $\tau_1 \neq \tau_2$, $\tau_2 \neq \tau_3$, $\tau_1 \neq \tau_3$.

3.2 WIENER KERNEL TRANSFORM ESTIMATES

From the above Wiener kernel estimates a heuristic argument leading to kernel transform estimates is presented. Consider a nonlinear system described by the Wiener kernel expansion. The output is given by Equation (3.2). Lee and Schetzen show that for a white Gaussian input

$$k_2(\tau_1, \tau_2) = \frac{1}{2C^2} E\left[y(t)\, x\,(t-\tau_1)\, x\,(t-\tau_2)\right]$$

for $\tau_1 \neq \tau_2$. If the process is ergodic, the time and ensemble averages are equal. Taking the average over 2T and defining

$$x_T(t) = \begin{cases} x(t) & -T<t<T \\ o & \text{elsewhere} \end{cases}$$

and (3.9)

$$y_T(t) = \begin{cases} y(t) & -T<+<T \\ o & \text{elsewhere} \end{cases}$$

gives

$$k_2(\tau_1, \tau_2) = \frac{1}{2TC^2} \int_{-T}^{T} y_T(t)\; x_T(t-\tau_1) x_T(t-\tau_2)\, dt \qquad (3.10)$$

$$\triangleq \hat{k}_2(\tau_1,\tau_2)$$

where the superscript $\hat{k}_2$ indicates the estimate of k_2 from input and output over the interval (-T,T).

The aim, however, is to estimate the Fourier transform of $k_2(\tau_1,\tau_2)$. Consider the Fourier transform of $\hat{k}_2\ (\tau_1,\tau_2)$,

$$F\left[\hat{k}_2(\tau_1,\tau_2)\right] \triangleq \hat{K}_2(W_1,W_2) \tag{3.11}$$

$$\hat{K}_2(W_1,W_2) = \int \hat{k}_2(\tau_1,\tau_2)e^{-j(W_1\tau_1 + W_2\tau_2)}d\tau_1 d\tau_2 \tag{3.12}$$

$$= \frac{1}{2TC^2}\int_{-T}^{T}\iint y_T(t)X_T(t-\tau_1)X_T(t-\tau_2)e^{-j(W_1\tau_1 + W_2\tau_2)}d\tau_1 d\tau_2 dt \quad . \tag{3.13}$$

Equation (3.13) can be simplified by noting that

$$\int X(t-\tau)e^{-jw}\,d\tau = e^{-jwt}\int X(t-\tau)e^{-jw(t-\tau)}d\tau$$

$$= e^{-jwt}X^*(W) \tag{3.14}$$

where $X^*(W)$ denotes the complex conjugate of the Fourier transform of $x(t)$. Using (3.14), (3.13) becomes

$$\hat{K}_2(w_1,w_2) = X^*(w_1)X^*(w_2)\frac{1}{2C^2T}\int_{-T}^{T}y_T(t)e^{-j(w_1t_1 + w_2t_2)}dt_1 dt_2 \quad . \tag{3.15}$$

As T approaches ∞, the integral of (3.15) approaches $Y_T(w_1 + w_2)$. Hence as T approaches ∞ one might expect that

$$\hat{K}_2(w_1,w_2) \to \frac{X^*(w)X^*(w)Y(w_1 + w_2)}{2C^2} \quad . \tag{3.16}$$

However, this is not true.

Letting the limits of (3.15) approach $-\infty$ and ∞, and integrating, leads to a false conclusion, (3.16). A similar situation occurs in power spectrum estimation. This is discussed briefly by Papoulis [8], and more thoroughly by Davenport and Root [9]. It is shown that the autocorrelation function, $R(\tau)$ can be approximated by

$$\hat{R}(\tau) = \frac{1}{2T}\int_{-T}^{T}X(t)\ x\ (t-\tau)d\tau \tag{3.17}$$

and that as T approaches ∞, $\hat{R}(\tau)$ approaches $R(\tau)$. This is not the case, however, in the frequency domain when estimating the power spectrum. The Fourier transform of $\hat{R}(\tau)$ is

$$F\left[\hat{R}(\tau)\right] = \frac{|X_T(w)|^2}{2T} \tag{3.18}$$

where $|X(w)|$ indicates the magnitude of $X(w)$. Davenport and Root prove that

$$\lim_{T\to\infty} \frac{|X_T(w)|^2}{2T} \tag{3.19}$$

does not approach the power spectrum. However, the power spectrum is equal to the expected value

$$\lim_{T\to\infty} \frac{E\left[|X_T(w)|\right]}{2T} \quad . \tag{3.20}$$

Similarly in a proof by Koplowitz it is shown in Appendix I that

$$K_2(w_1,w_2) = \lim_{T\to\infty} \frac{E\left[X^*(w_1)X^*(w_2)Y(w_1+w_2)\right]}{2C^2} - \frac{k_o\delta^2(w_1+w_2)}{2C} \tag{3.21}$$

The expression (3.15) is now seen as a random variable whose expected value approaches the true kernel transform. The first three Wiener kernel transform estimates are

$$\begin{aligned} K_1(w) &= E\left[X^*(w)\ Y\ (w)\right]/C \\ K_2(w_1,w_2) &= \frac{E\left[X^*(w_1)\ X^*(w_2)\ Y\ (w_1+w_2)\right]}{2C^2} \end{aligned} \tag{3.22}$$

for $w_1 \neq w_2$

$$K_3(w_1,w_2,w_3) = \frac{E\left[X^*(w_1)X^*(w_2)X^*(w_3)Y(w_1+w_2+w_3)\right]}{6C^3}$$

for $w_1 \neq -w_2, w_2 \neq -w_3, w_1 \neq -w_3$.

For the second and third-order estimates, impulse functions occur when any two of the arguments sum to zero. This is shown in Appendix I for $\hat{K}_2$ and Appendix II for $\hat{K}_3$.

4.0 EXPERIMENTS TO ESTIMATE WIENER KERNEL TRANSFORMS

In order to verify the Wiener kernel transform estimates presented in Section 3.0, several experimental computer simulations were performed. The first-, second- and third-order Wiener kernel transforms were estimated for several systems. The diagram in Fig. 6 outlines the experimental procedure which involved simulating a nonlinear system response to a discrete, white, stationary, Gaussian input process. A random-number generator was used to generate the inpute sequence. To avoid any possible effects associated with the initial output of the number generator it was allowed to reach "steady state." The initial points, three time constants of the linear system, were not used for each single estimate.

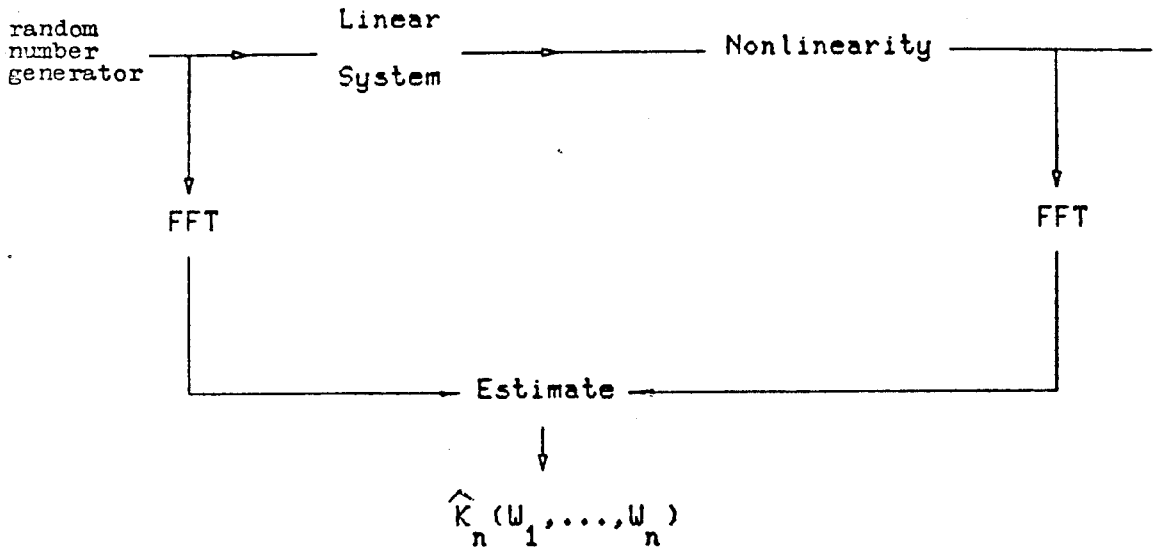


FIGURE 6. Outline for kernel transform estimation experiment.

The estimates necessitate the discrete Fourier transform of both input and output signals (sequences) and these are calculated via the fast

Fourier transform algorithm. The experiments were repeated many times, since the estimate expected values should approach the true Wiener kernel transforms. A flow chart and listing of the FORTRAN encoded programs are given in Appendix V. A 64-point data length was used.

4.1 FIRST-ORDER TRANSFORM ESTIMATE EXPERIMENTS

The first-order kernel transofrm is merely the linear system function (Section 2.0). The systems chosen for these experiments were two recursive (infinite impulse response) digital filters of first and second order, respectively.

The first-order difference equation for the low-pass filter used in the experiments is

$$y(n) = 0.3\ x\ (n) + 0.7y\ (n-1)\quad n = 0,1,... \tag{4.1}$$

with the initial condition

$$y(-1) = 0\ .$$

The frequency magnitude response is [10],

$$\left|T_1(w)\right| = \frac{0.3}{(1.49 - 1.4\ \cos\ w)^{1/2}} \tag{4.2}$$

for $0 \leq w \leq 2\pi$, where $w = \pi$ is the Nyquist frequency. The second-order difference equation for the digital resonator used is

$$y(n) = 0.198\ x\ (n) + 1.20y(n-1) - 0.723\ y(n-2) \tag{4.3}$$

with the initial conditions

$$y(-1) = y(-1) = 0\ .$$

The corresponding frequency magnitude response [10] is

$$\left|T_2(w)\right| = \frac{0.198}{\left[(1-1.2\ \cos\ w - 0.723\ \cos\ 2\ w)^2 + (1.2\ \sin\ w + 0.723\ \sin\ 2\ w)^2\right]^{1/2}} \tag{4.4}$$

for $0 \leq w \leq 2\pi$.

The magnitude and phase angle for the calculated and estimated kernel transforms after 1000 averages are given in Figs 7 — 10. For the estimate magnitude, mean squared errors of 1.12×10^{-3} and 1.3×10^{-3} were achieved for the first- and second-order digital filters, respectively. In all plots the frequency is normalized, so that w = 1 is the Nyquist frequency.

4.2 SECOND-ORDER TRANSFORM ESTIMATE EXPERIMENTS

Three experiments to verify the second Wiener kernel transform estimate were performed. The first two nonlinear systems consisted of the previous filters of Equations (4.1) and (4.3), respectively, whose output was squared as shown in Fig. 11. The second-order kernel transform for a system of the above form is given by Equation (2.15),

$$K_2(w_1,w_2) = T(w_1)\ T(w_2) \tag{4.5}$$

where T(w) is the system function of the appropriate digital filter. Perspective plots of the magnitude of the kernel transforms and estimate after 2500 averages, for each system, are given in Figs 12 — 16.

The second-order estimate approached $K_2(w_1,w_2)$ where $w_1=-w_2$. As shown in Fig. 14 where the two arguments sum to zero, an impulse function is present as in Equation (3.21). For this reason, points close to this region were excluded when calculating the mean square error which was 1.54×10^{-3} for the low-pass system and 1.73×10^{-3} for the resonator system after 2500 averages. The magnitude and phase angle of two cross sections for each estimate are shown in Figs 17 — 20.

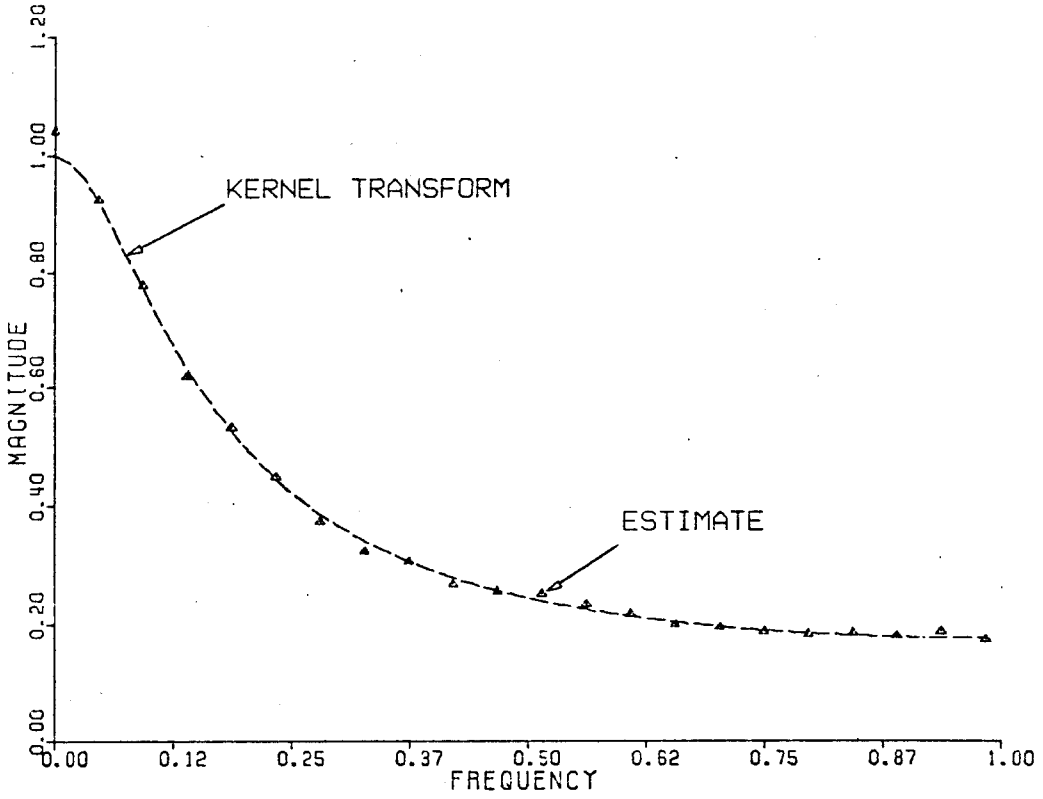


FIGURE 7. First-order estimate magnitude after 1000 averages — low-pass

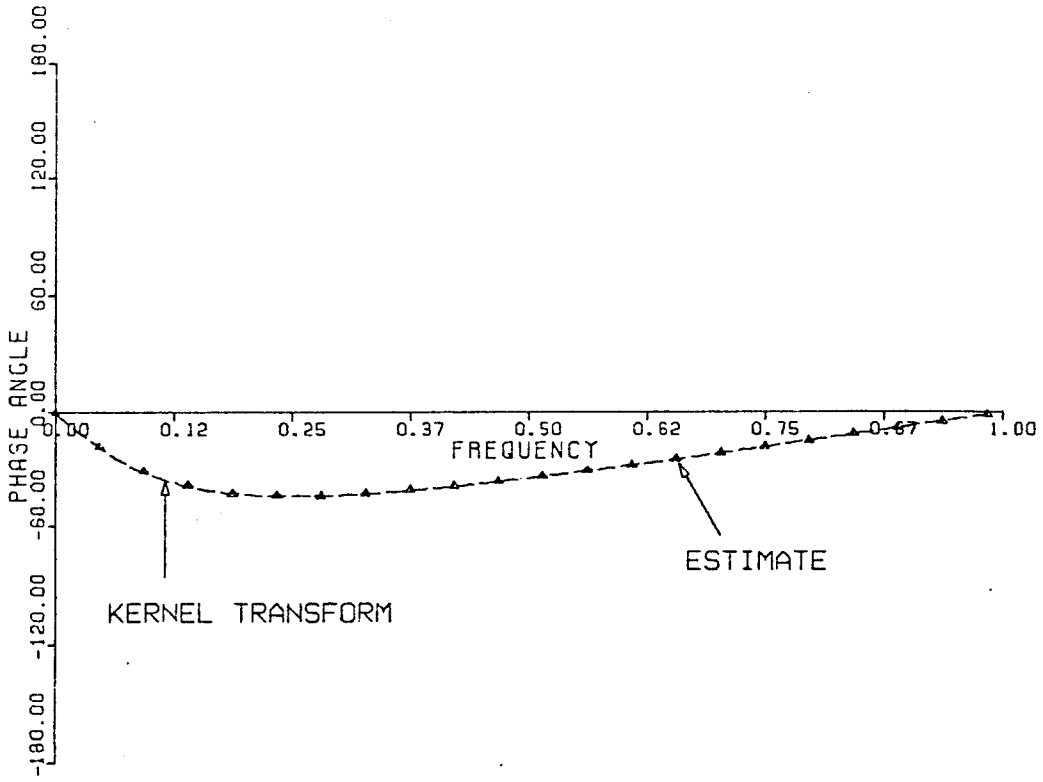


FIGURE 8. First-order estimate phase angle after 1000 averages — low-pass system.

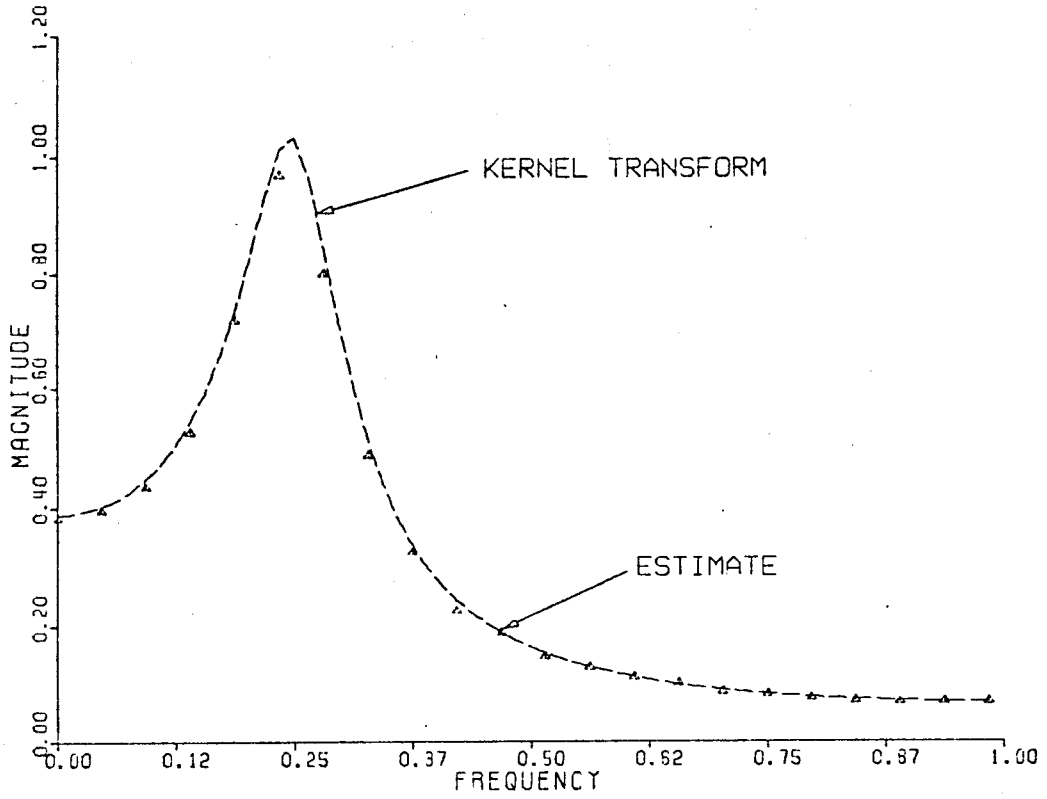


FIGURE 9. First-order estimate magnitude after 1000 averages — resonator system.

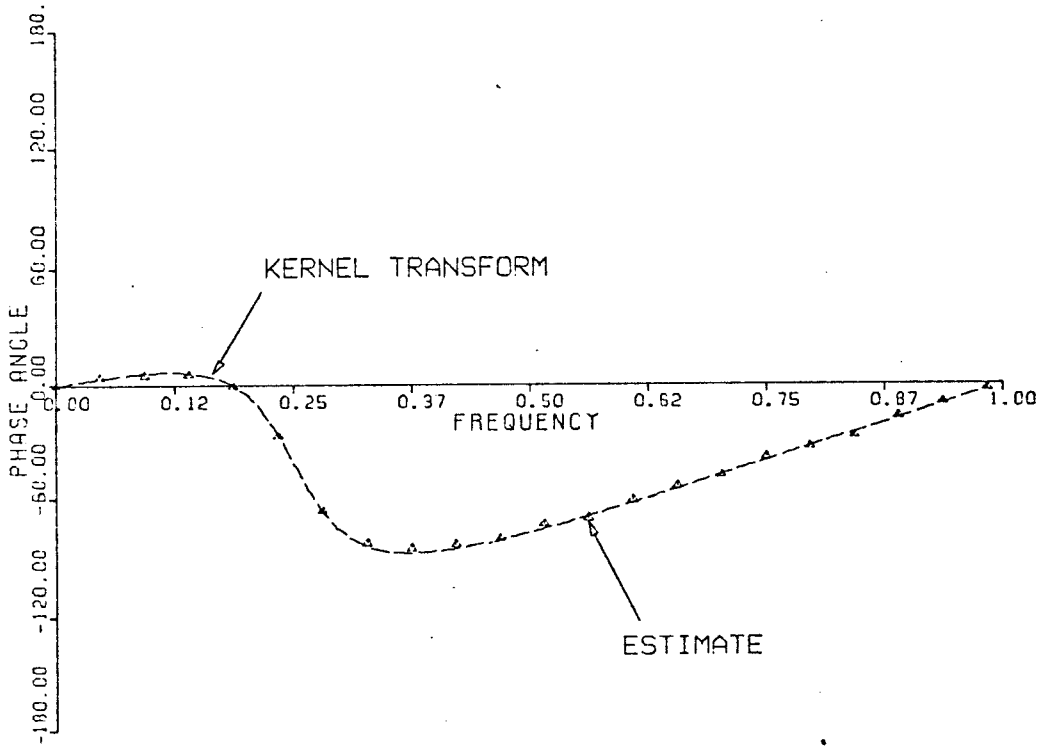


FIGURE 10. First-order estimate phase angle after 1000 averages — resonator system.

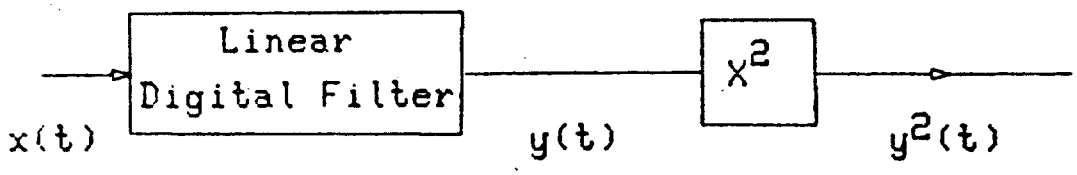


FIGURE 11. System configuration used in second-order kernel transform experiments.

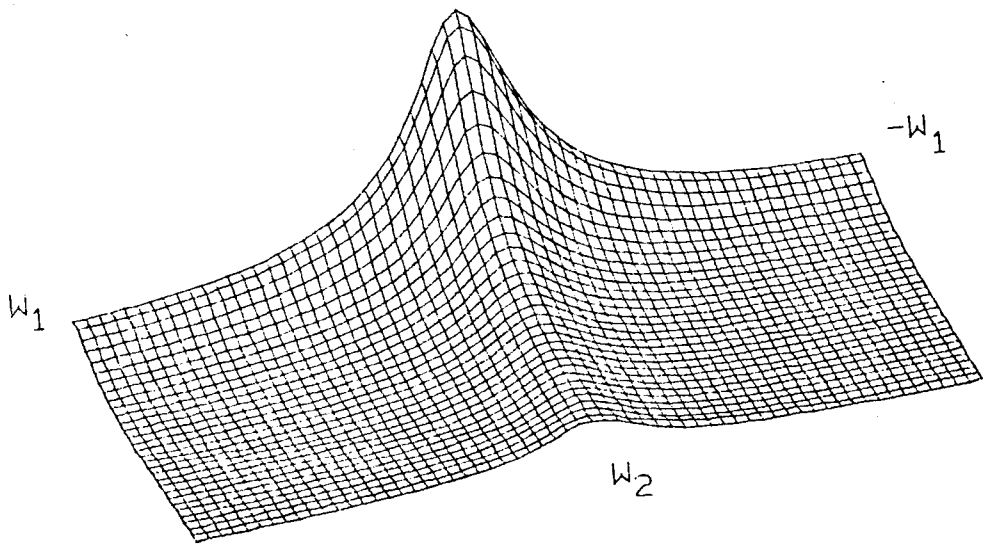


FIGURE 12. Second-order kernel transform magnitude — low-pass nonlinear system.

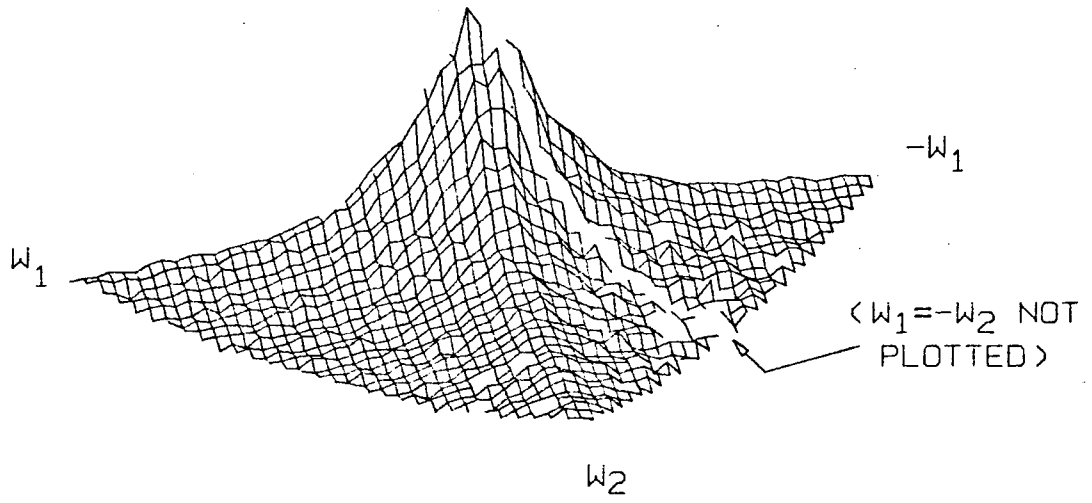


FIGURE 13. Second-order estimate magnitude after 2500 averages — low-pass nonlinear system.

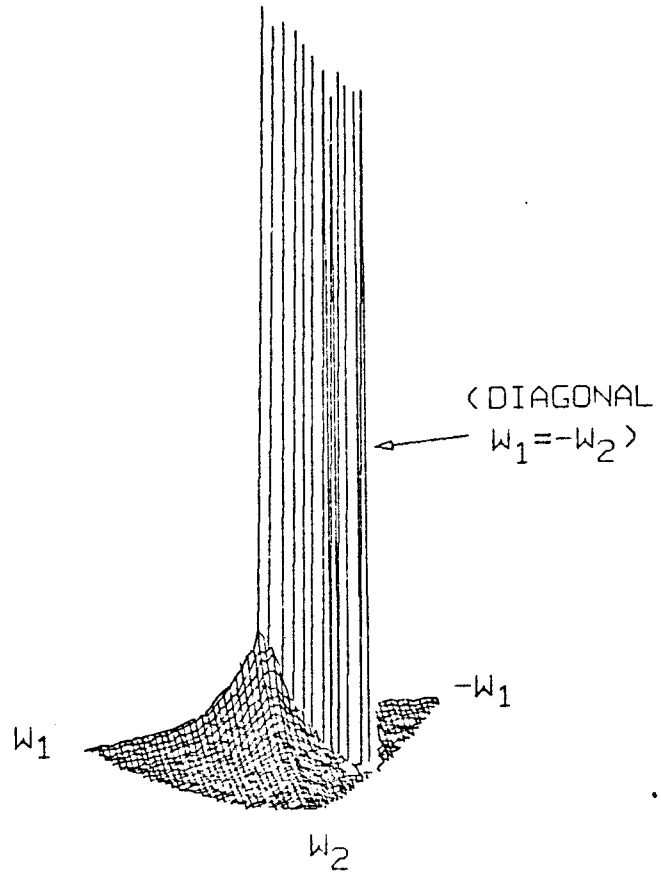


FIGURE 14. Second-order estimate magnitude after 2500 averages with diagonal — low-pass nonlinear system.

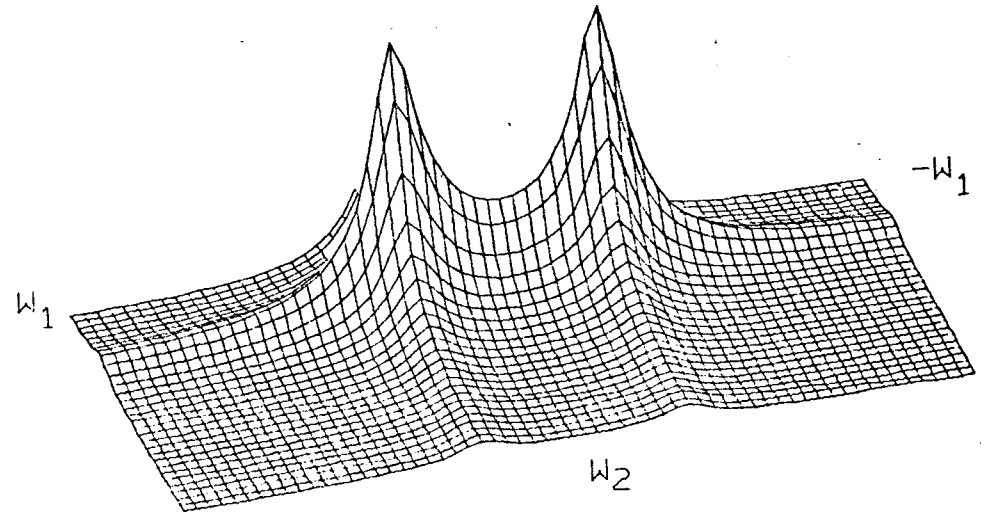


FIGURE 15. Second-order kernel transform magnitude — resonator nonlinear system.

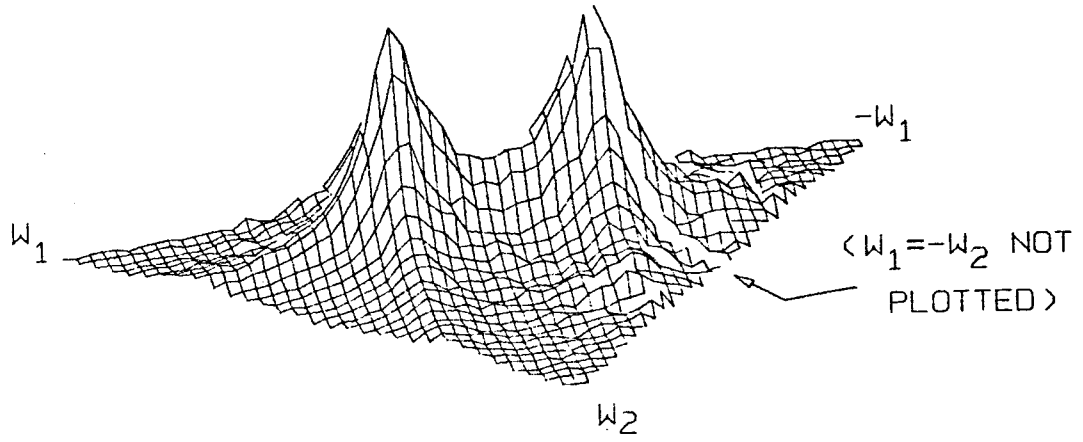


FIGURE 16. Second-order estimate magnitude after 2500 averages — resonator nonlinear system.

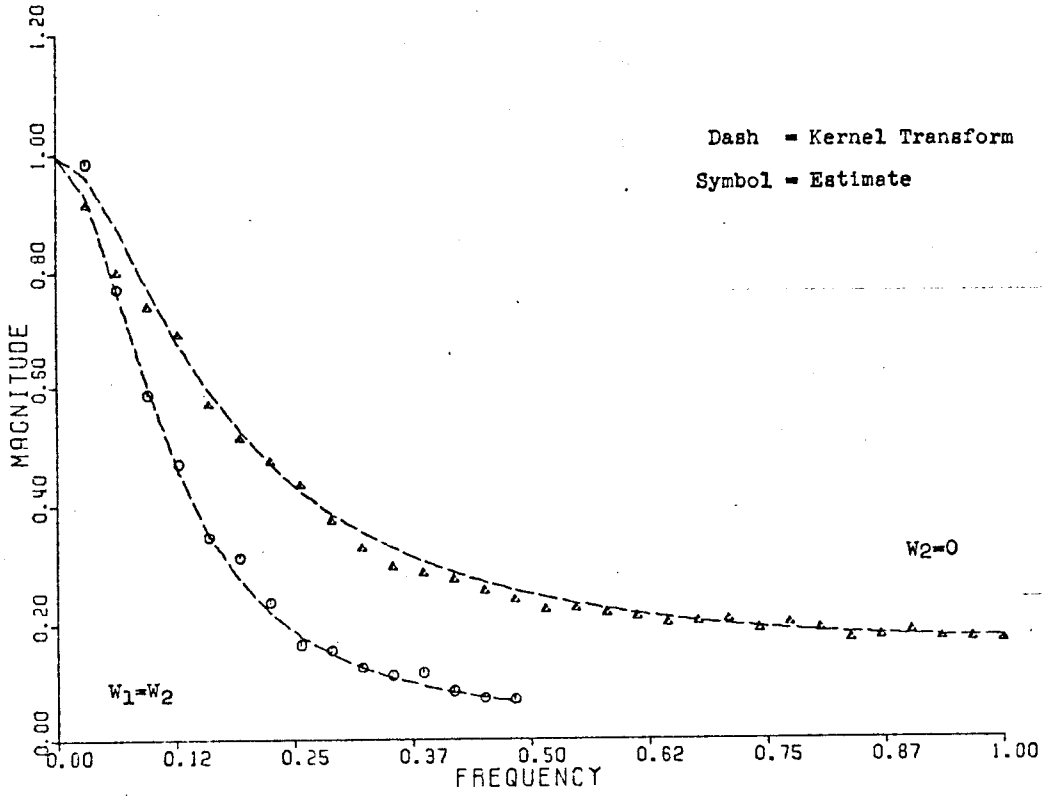


FIGURE 17. Second-order estimate-magnitude cross sections after 2500 averages — low-pass nonlinear system.

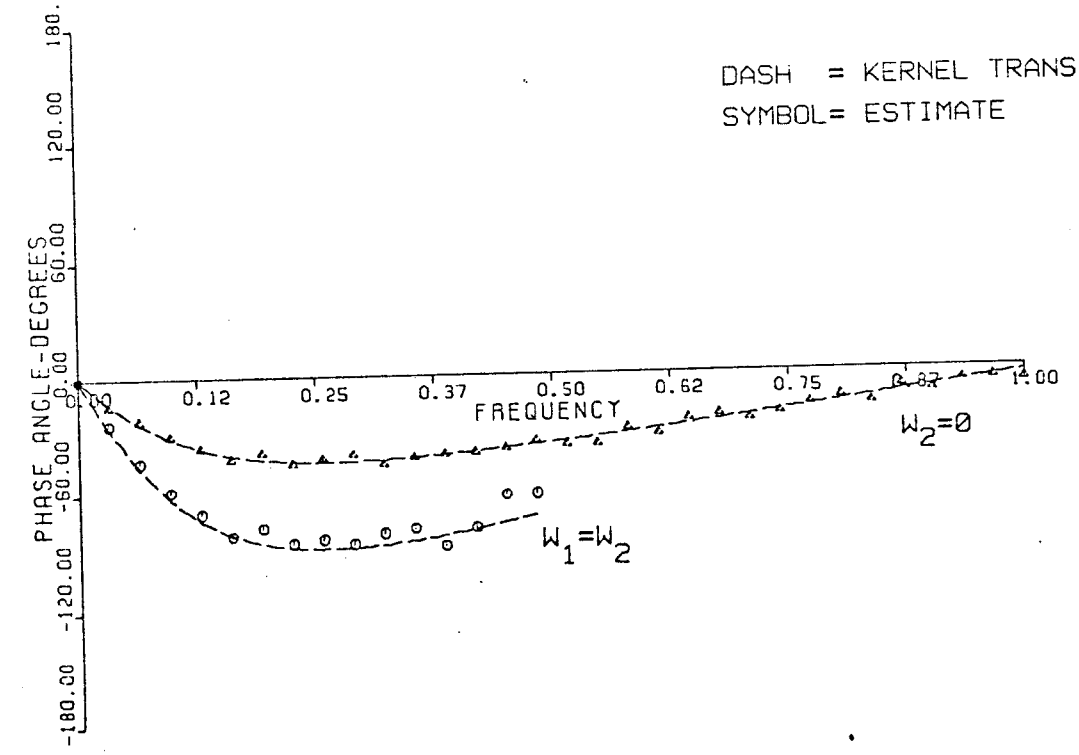


FIGURE 18. Second-order estimate phase angle cross sections after 2500 averages — low-pass nonlinear system.

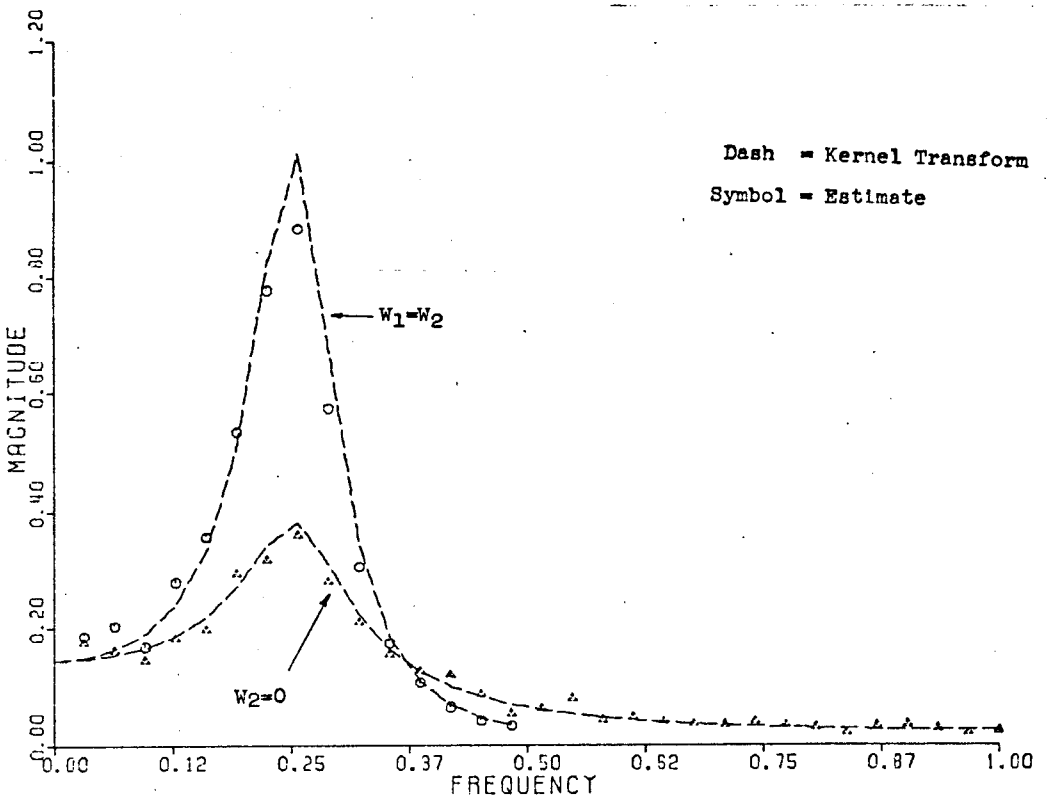


FIGURE 19. Second-order estimate magnitude cross sections after 2500 averages — resonator nonlinear system.

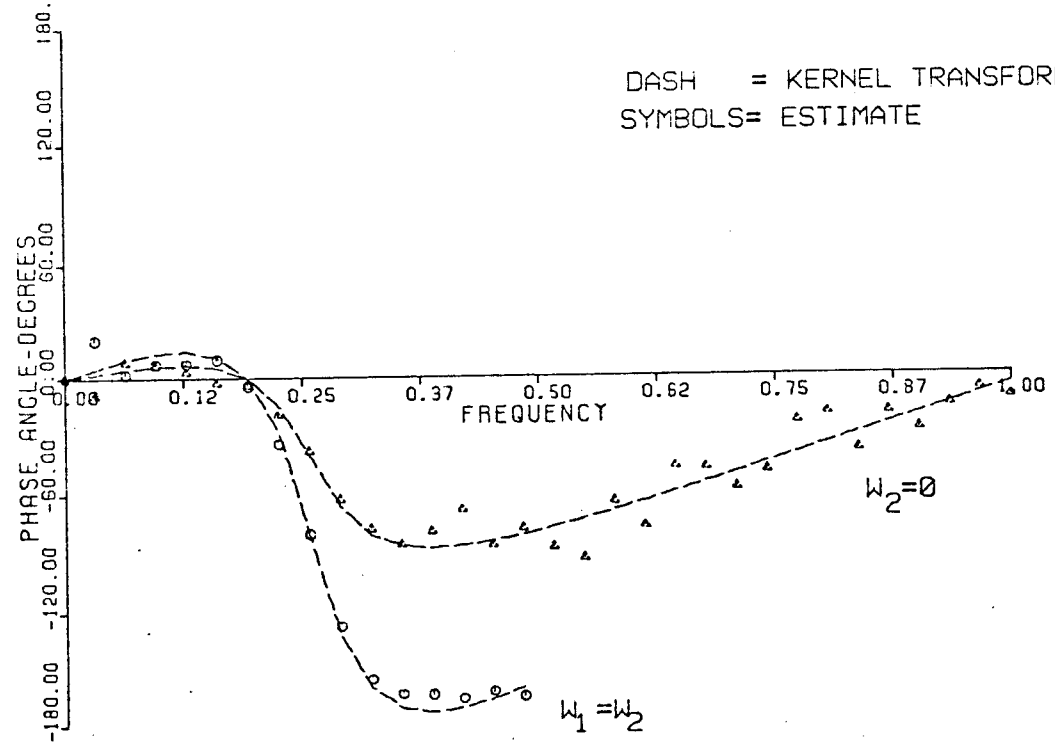


FIGURE 20. Second-order estimate phase angle cross sections after 2500 averages — resonator nonlinear system.

The third nonlinear system whose second-order kernel transform was estimated is shown in Fig. 21. This system is similar to the previous one with the addition of an output filter. The second-order Wiener kernel transform for the system is given by Equation (2.31),

$$K_2(w_1,w_2) = T_2(w_1)T_2(w_1)T_1(w_1+w_2) \tag{4.6}$$

where $T_1(w)$ and $T_2(w)$ are the system functions of the low-pass filter and resonator of Equations (4.1) and (4.3). The kernel transform magnitude and its estimate after 5000 averages are shown in Figs 22 and 23. The mean squared error achieved was 9.03×10^{-4}.

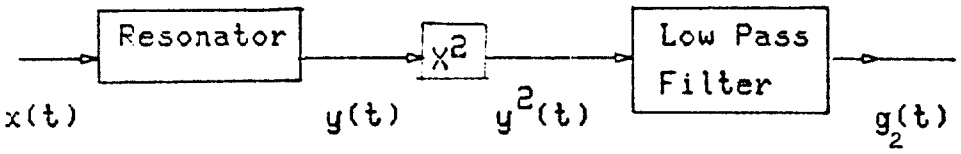


FIGURE 21. System configuration used in second-order kernel transform estimate experiment.

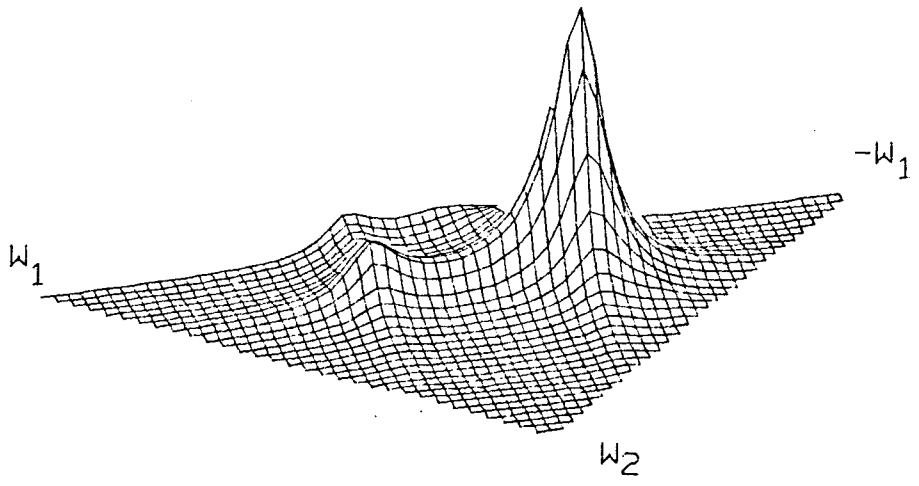


FIGURE 22. Second-order kernel transform magnitude for resonator – low-pass nonlinear system.

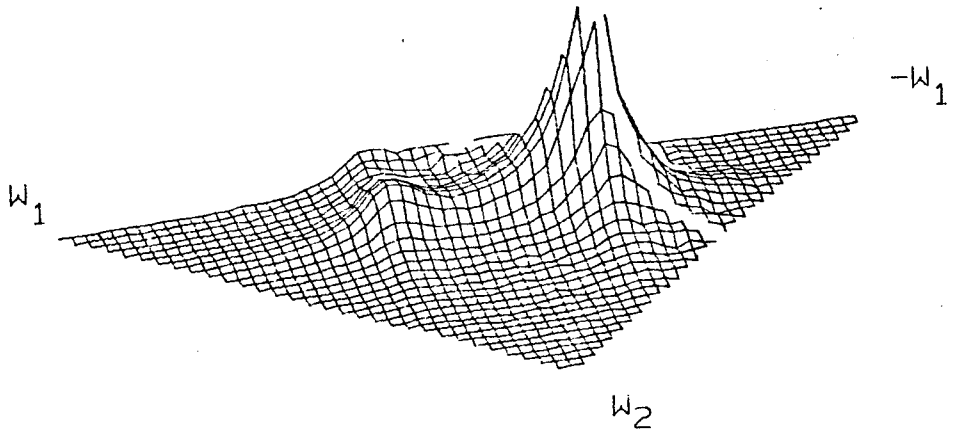


FIGURE 23. Second-order estimate magnitude after 5000 averages for resonator — low-pass nonlinear system.

4.3 THIRD-ORDER TRANSFORM ESTIMATE EXPERIMENTS

Two experiments were performed to verify the third-order Wiener kernel transform estimate. The two nonlinear systems consisted of the linear filters of Equations (4.1) and (4.3) whose output was cubed as shown in Fig. 4.

The Wiener kernel transform for the two systems of this form is given by Equation (2.15),

$$K_3(w_1,w_2,w_3) = T(w_1)T(w_2)T(w_3) \tag{4.7}$$

where $T(w)$ is the system function of the filter used. The cross-section plots, for $w_3 = 0.125$, for the kernel transform and estimate after 5000 averages are shown in Figs 24 — 27. A mean squared error of 2.74×10^{-3}

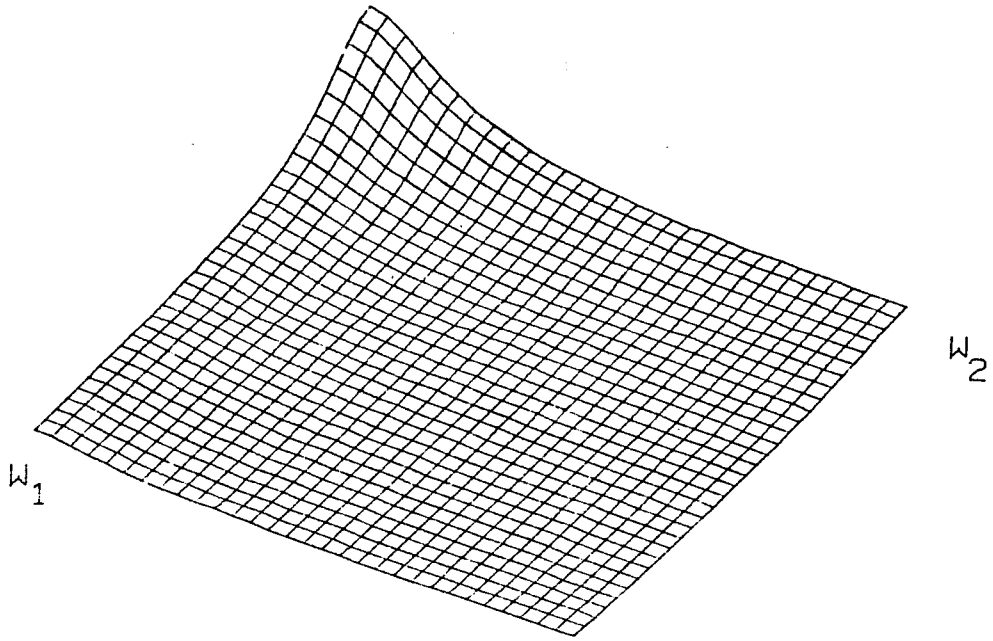


FIGURE 24. Third-order kernel transform magnitude cross-section ($w_3 = 0.125$) — low-pass nonlinear.

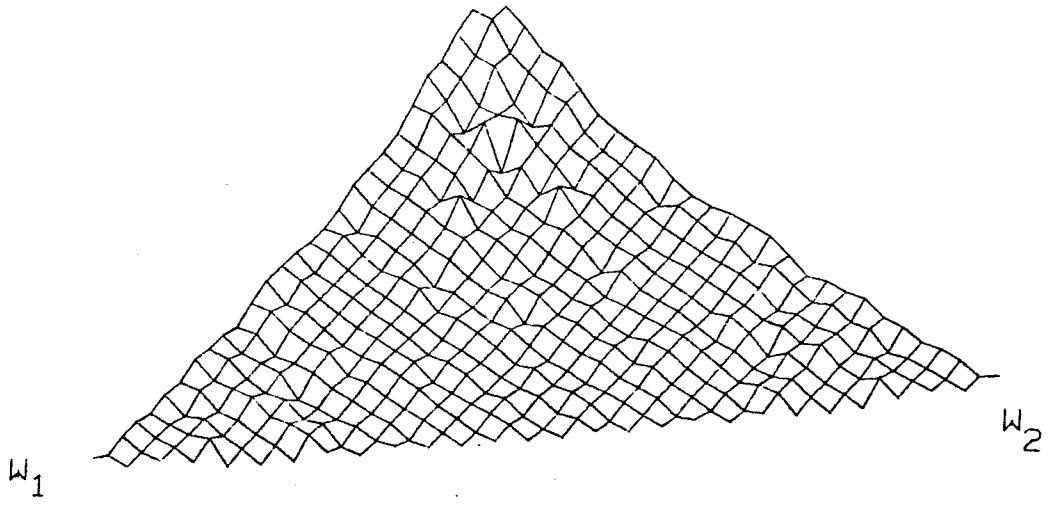


FIGURE 25. Third-order estimate magnitude cross-section ($w_3 = 0.125$) after 5000 averages — low-pass nonlinear system.

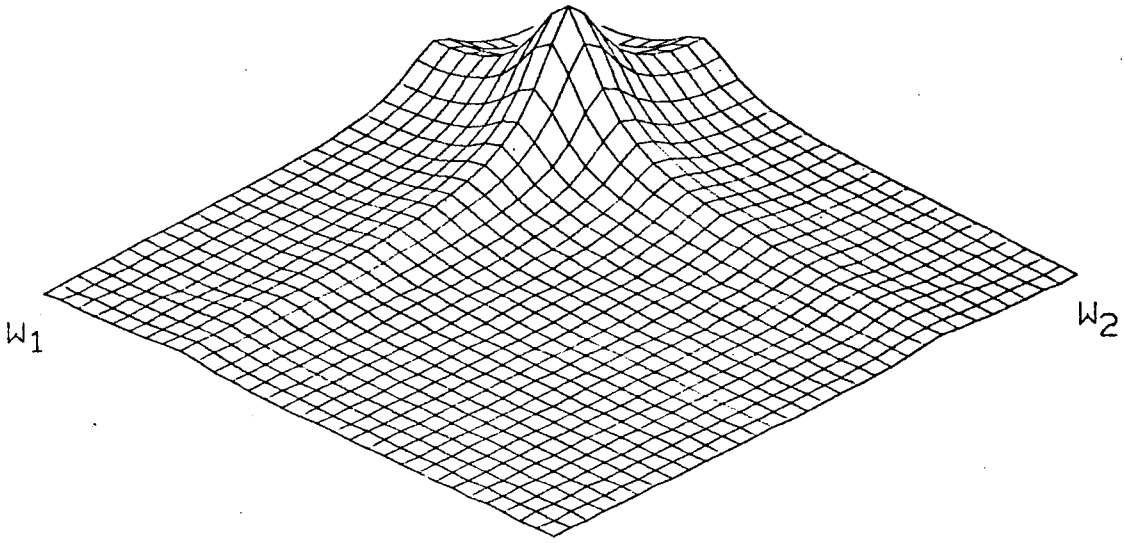


FIGURE 26. Third-order kernel transform magnitude cross-section ($w_3 = 0.125$) — resonator nonlinear system.

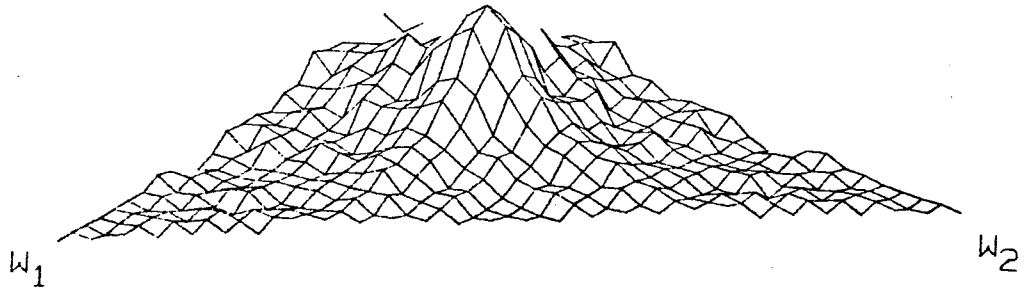


FIGURE 27. Third-order estimate magnitude cross section ($w_3 = 0.125$) after 5000 averages — resonator nonlinear system.

was found for the low-pass system and 2.86×10^{-3} for the resonator system.

As with the second-order case, the estimate does not converge to the kernel where

$$\begin{aligned} w_1 &= -w_2 \\ w_2 &= -w_3 \\ w_3 &= -w_1 \end{aligned} \tag{4.8}$$

In these regions impulse functions are present as shown in Appendix II.

In addition to estimating the third kernel transform of the above system, the first kernel transofrm was estimated for the low-pass nonlinear system. The first Wiener kernel of the system of Fig. 4 is given by Equations (2.14) and (2.41)

$$\begin{aligned} k_1(\tau) &= 3C \int a(\tau)\ a(\tau_1)\ a(\tau_1)\ d\tau_1 \\ &= 3C\ \ a(\tau) \int a^2(\tau)\ d\tau \end{aligned} \tag{4.9}$$

where $a(t)$ is the impulse response of the linear system. The corresponding kernel transform is

$$K_1(w) = \left[3C \int a^2(\tau_1)\ d\tau_1\right] A(w) \tag{4.10}$$

where $A(w)$ is the Fourier transform of $a(t)$. For this experiment the low-pass cubic nonlinear system first kernel transform was estimated.

Equation (4.11) was calculated for the continuous nonlinear system of Fig. 4. The system in question, however, is its discrete equivalent. Instead of $a(t)$ we have

$$a(n) \qquad n = 0,1,2\ldots$$

for the digital filter impulse response. The discrete equivalent of the integration of (4.10) is summation, so that (4.10) becomes

$$K_1(w) = \left[3C \sum_{n=-\infty}^{\infty} a^2(n)\right] T_1(w) \quad . \tag{4.11}$$

Since the filter is causal [11],

$$a(n) = 0$$

for $n < 0$ and Equation (4.11) becomes

$$K_1(w) = \left[3C \sum_{n=0}^{\infty} a^2(n)\right] T_1(w) \tag{4.12}$$

The low-pass filter (Equation 4.1) output to the digital impulse input

$$x(n) = \begin{cases} 1 & \text{for } n=0 \\ 0 & \text{for } n\neq 0 \end{cases} \tag{4.13}$$

is seen as

$$a(n) = 0.3(0.7)^n \qquad n=0,1,\ldots \tag{4.14}$$

and

$$a^2(n) = 0.09(0.7)^{2n} \quad n=0,1,\ldots \tag{4.15}$$

The summation of (4.16) was numerically found equal to 0.174. Equation (4.12) is now

$$K_1(w) = 0.522\ T_1(w) \tag{4.17}$$

for an input process variance of unity. The first kernel magnitude transform and its estimate are given in Fig. 28. A mean squared error of 1.12×10^{-3} was found after 1000 averages.

Similarly, the first kernel transform of the resonator cubic nonlinear system was estimated. The estimate after 1000 averages is shown in Figs 30 and 31.

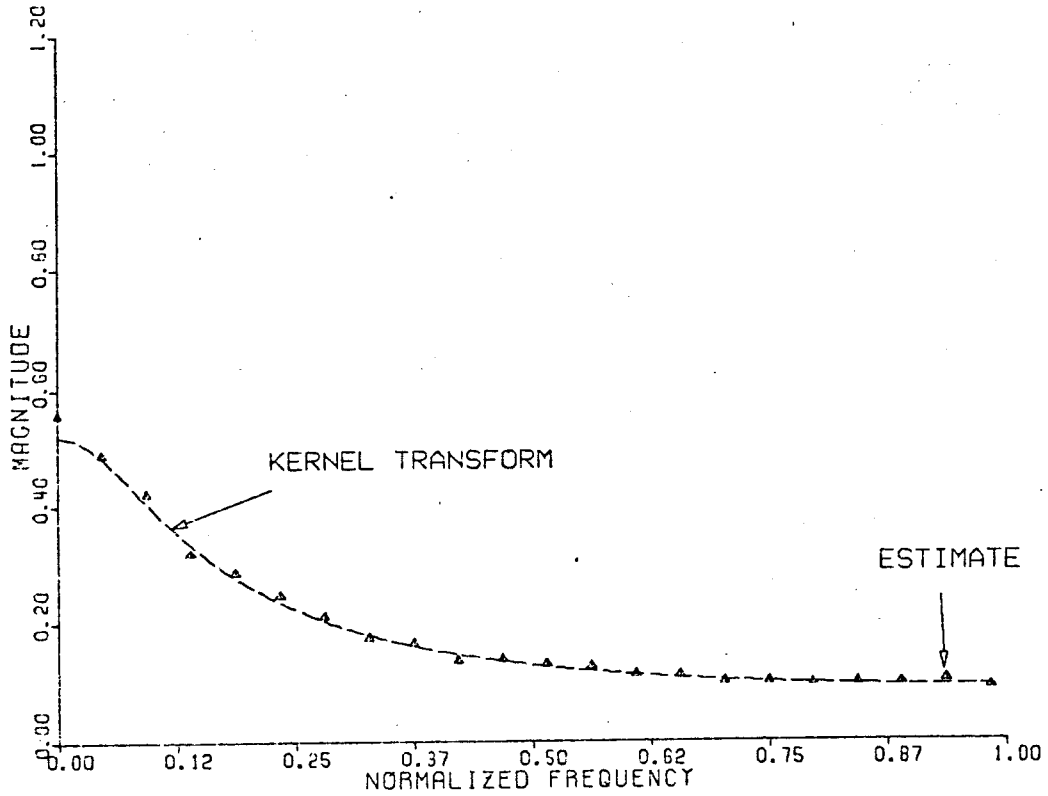


FIGURE 28. First-order kernel transform estimate magnitude of the low-pass cubic nonlinear system after 1000 averages.

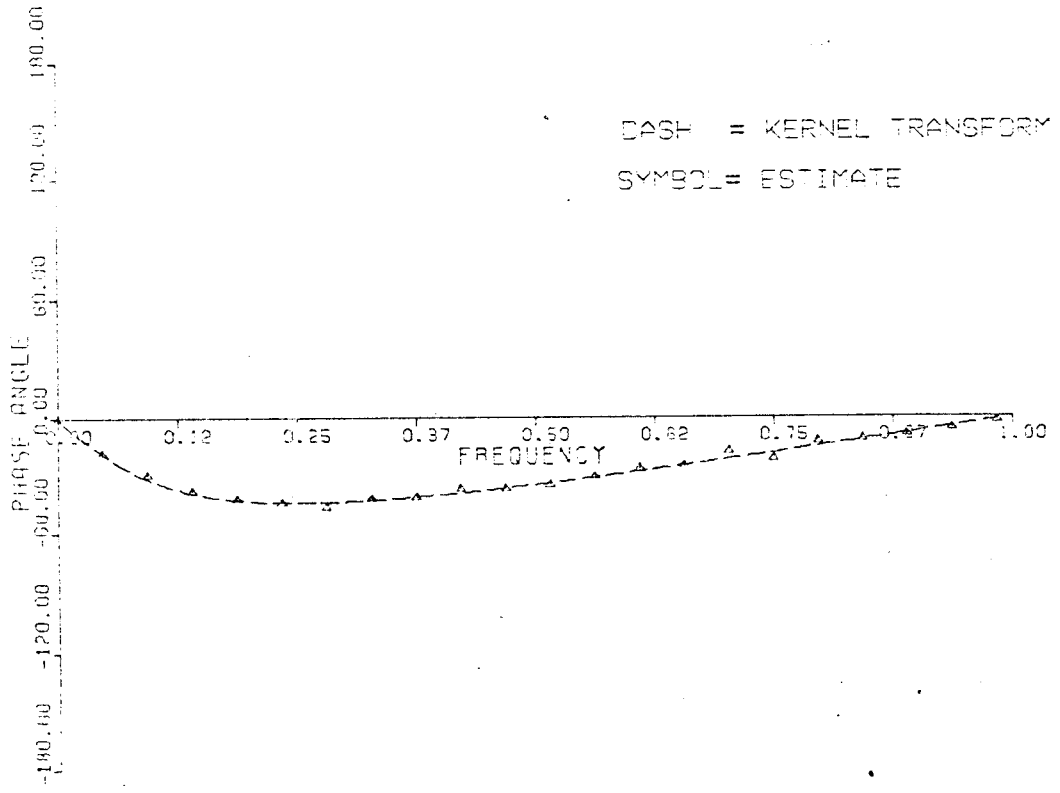


FIGURE 29. First-order kernel transform phase angle of the low-pass cubic nonlinear system after 1000 averages.

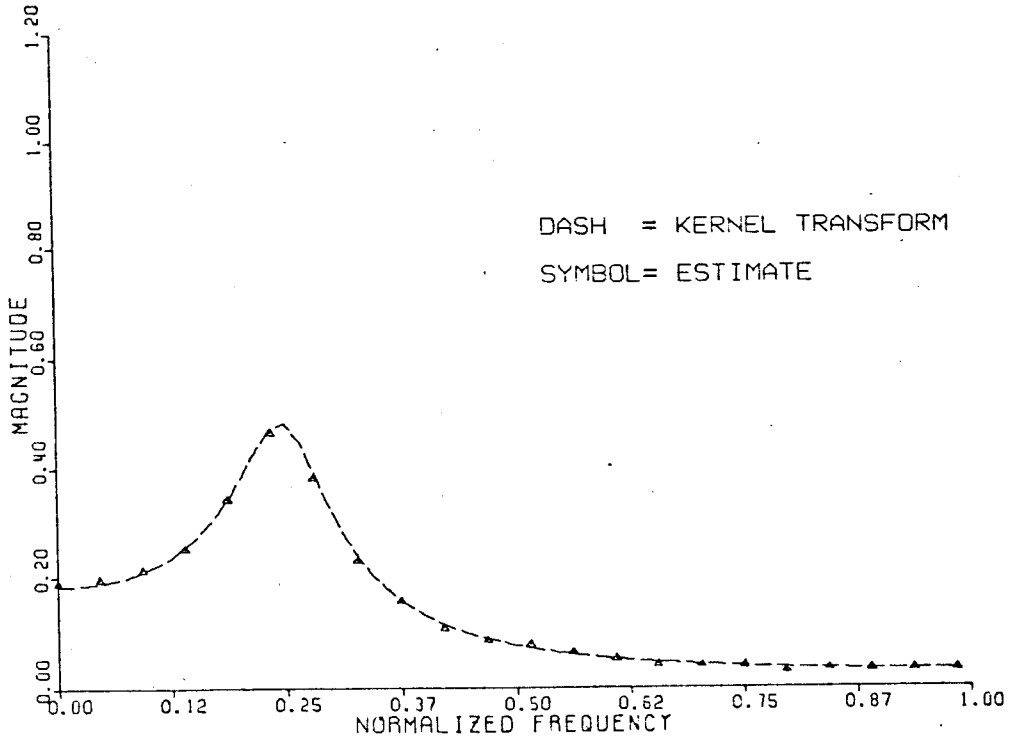


FIGURE 30. First-order kernel transform estimate magnitude of the resonator cubic nonlinear system after 1000 averages.

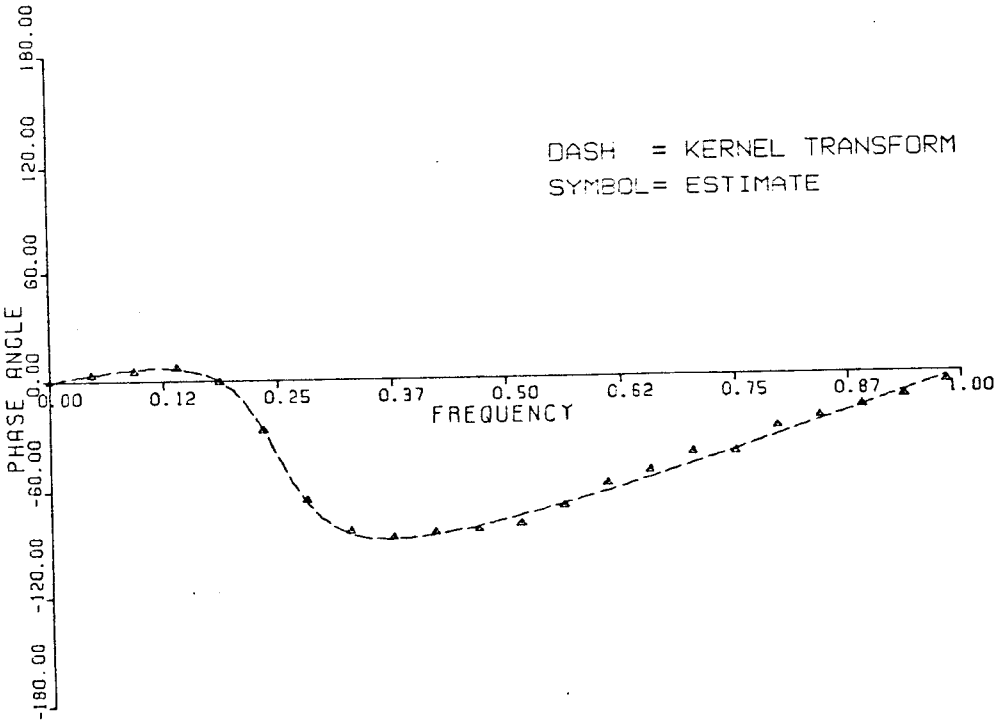


FIGURE 31. First-order kernel transform estimate phase angle of the resonator cubic nonlinear system after 1000 averages.

5.0 ESTIMATE VARIANCE AND THE COMPUTATION NECESSARY FOR WIENER KERNEL ESTIMATION

In Section 4, experiments were described estimating the Wiener kernel transforms of several systems. The results show that the estimates approach the kernel transforms. In order to assess the usefulness of the estimates, the variances of the estimates are now considered.

5.1 ESTIMATE VARIANCE AND STANDARD DEVIATION

The power spectrum estimate is being widely used and is discussed here only for purposes of comparison with the Wiener kernel transform estimates. Consider the power spectrum estimate, the periodogram, of a random process $x(t)$, given by

$$\hat{P}(w) = \frac{|X(w)|^2}{2T} \tag{5.1}$$

where $X(w)$ is the Fourier transform of $x(t)$ over the interval $[-T,T]$.

The expected value of the estimate of (5.1) approaches the process power spectrum. For the discrete case, a good approximation [11] for the periodogram variance is

$$\sigma^2 = P^2(w) \left[1 + \left(\frac{\sin w M}{M \sin w}\right)^2\right] \tag{5.2}$$

where $P(w)$ is the power spectrum of the process and M is the estimate data length. As M increases, the variance approaches the square of the spectrum. The estimate variance decreases [11] as $1/M$, where M is the number of independent estimates averaged when taking the expected value. Knowing the variance of a single estimate, and the effect of averaging, the number of averages is now known for a given average estimate variance. The above discussion points out that knowing the estimate variance is of

practical value when making the power spectrum estimate. The same information is useful for the Wiener kernel estimates.

During the experiments described in Section 4, an estimate of the kernel transform estimate magnitude variance was obtained. For this discussion, the kernel transform estimate will be referred to as the 'estimator.'

An estimate of the estimator magnitude variance [12] is

$$\hat{\sigma}_n^2(w_1,\ldots,w_n) = \frac{\sum_{i=1}^{N}\left[\left|\hat{K}_n^i(w_1,\ldots,w_n)\right| - \left|(\bar{K}_n(w_1,\ldots,w_n)\right|\right]^2}{N-1} \qquad (5.3)$$

where $|\hat{K}_n^i|$ is the magnitude of the ith estimator of order n and $|\bar{K}_n|$ is the magnitude of the estimator magnitude after M (complex) averages. The value of M was sufficiently large that $|\bar{K}_n|$ had approached its (complex) mean value, which took up to 7000 averages for the third-order case.

Table 2 lists the first-order kernel transform magnitude, estimator magnitude and $\hat{\sigma}$, for the two linear filters of Equations (4.1) and (4.3). Data are also given for the first-order estimator of the cubic nonlinear system of Fig. 4. The average ratio of $\hat{\sigma}$ divided by the estimator magnitude is close to 1.0. It was expected that the estimator magnitude variance of the nonlinear system would differ further from that of the linear filter systems. For the nonlinear systems, $\hat{\sigma}$ was somewhat greater, however investigation with other nonlinear systems is needed to verify these results.

The same information is listed in Tables 3 and 4 for several points of the second kernel transform and estimator. Data are given for the low-pass and resonator second-order nonlinear systems and also the resonator — low-pass system. The results

TABLE 2

First Wiener Kernel Transform Estimate after 1000 averages (Nyquist frequency = 64)

system	Frequency	$\lvert K_1 \rvert$	$\lvert \hat{K}_1 \rvert$	$\hat{\sigma}$	$\hat{\sigma} / \lvert \hat{K}_1 \rvert$
low-pass	5	.877	.892	.816	.915
	15	.470	.481	.462	.960
	25	.307	.305	.305	1.00
	35	.235	.243	.245	1.01
	50	.189	.189	.192	1.02
	Average				.981
resonator	5	.414	.416	.400	.963
	15	.927	.912	.893	.979
	25	.338	.327	.331	1.01
	35	.145	.144	.147	1.02
	50	.082	.081	.083	1.03
	Average				1.00
low-pass cubic	5	.462	.476	.554	1.16
	15	.254	.261	.287	1.10
	25	.161	.167	.125	1.05
	35	.126	.123	.151	1.23
	50	.099	.104	.123	1.24
	Average				1.16
resonator cubic	5	.195	.205	.241	1.17
	15	.438	.441	.499	1.13
	25	.159	.157	.201	1.28
	35	.068	.073	.102	1.39
	50	.039	.039	.053	1.36
	Average				1.27

TABLE 3

SECOND WIENER KERNEL TRANSFORM ESTIMATE — LOW-PASS SYSTEM AFTER 2500 AVERAGES

(Nyquist frequency = 32)

Frequency (W_1, W_2)	$\lvert K_1 \rvert$	$\lvert \hat{K}_1 \rvert$	$\hat{\sigma}$	$\hat{\sigma} / \lvert \hat{K}_1 \rvert$
3,22	0.154	0.163	0.389	2.99
4,2	0.746	0.765	2.23	2.91
5,2	0.596	0.627	1.84	2.93
7,4	0.356	0.370	1.11	3.01
9,5	0.252	0.284	0.841	2.96
9,8	0.130	0.113	0.374	3.32
9,10	0.151	0.146	0.448	3.07
9,15	0.109	0.131	0.412	3.14
9,19	0.092	0.091	0.294	3.23
10,7	0.181	0.159	0.494	3.11
15,12	0.084	0.085	0.259	3.05
-2,27	0.160	0.136	0.406	2.98
-4,1	0.789	0.827	2.35	2.84
-8,1	0.469	0.433	1.33	3.07
-14,1	0.289	0.326	0.925	2.84
-24,1	0.196	0.254	0.801	3.15
-28,2	0.159	0.160	0.509	3.18
-29,1	0.180	0.202	0.733	3.63
-30,0	0.177	0.176	0.545	3.10
Average				3.08

TABLE 4

SECOND WIENER KERNEL TRANSFORM ESTIMATE — RESONATOR SYSTEM AFTER 2500 AVERAGES

(Nyquist frequency = 32)

Frequency w_1, w_2	$\lvert K_2 \rvert$	$\lvert \hat{K}_2 \rvert$	$\hat{\sigma}$	$\frac{\hat{\sigma}}{\lvert K_1 \rvert}$
0,-1	0.192	0.253	1.53	6.1*
1,-5	0.361	0.409	1.11	2.71
2,-10	0.132	0.143	0.446	3.12
3,-11	0.131	0.127	0.391	3.07
6,-14	0.155	0.172	0.557	3.23
2,1	0.163	0.107	1.50	14.0*
3,3	0.205	0.212	1.03	4.86
5,5	0.345	0.356	0.935	2.63
5,9	0.550	0.498	1.34	2.69
7,7	0.632	0.655	1.67	2.55
8,8	0.696	0.708	1.98	2.79
10,9	0.501	0.511	1.29	2.52
15,9	0.178	0.206	0.484	2.35
20,9	0.105	0.081	0.270	3.33
23,9	0.301	0.323	0.832	2.58
Average				2.96
Resonator — Low-Pass System				
0,2	0.135	0.236	8.04	34.1*
2,4	0.104	0.109	0.409	3.75
3,12	0.056	0.054	0.214	3.96
4,13	0.050	0.054	0.184	3.41
6,15	0.071	0.070	0.246	3.51
7,9	0.184	0.174	0.496	2.85
9,10	0.134	0.122	0.314	2.57
9,12	0.070	0.059	0.187	3.17
Average				3.31

*These points are close to the diagonal $w_1 = -w_2$ and are excluded from the average.

suggest that the estimator magnitude standard deviation is three times its (complex) average value. For frequencies close to and in the region where $w_1 = -w_2$, $\hat{\sigma}_2$ increases substantially as does the bias of the estimate.

The corresponding data for the low-pass and resonator cubic nonlinear systems' third-order estimators are given in Tables 5 and 6. The results suggest that $\hat{\sigma}_3$ is approximately nine times the value of the kernel transform magnitude. Again, for points at or near where the arguments sum to zero, increased variance and bias are evident.

The above results suggest that the single kernel transform estimate standard deviation is approximately one, three and nine times the value of the kernel transform for the 1st, 2nd and 3rd kernels. Assuming this to be true, the number of averages (used in taking the expectation) for a given estimate variance is known. This follows if the variance decreases as 1/M where M is the number of independent estimates averaged, as for the power spectrum estimate.

5.2 COMPUTATION NECESSARY FOR VARIOUS SINGLE ESTIMATES

In order to compare the amount of computation required for Wiener kernel transform estimation the approximate number of real multiplications needed are considered. Again, the power spectrum estimate is included for comparison. In this section, the regions where the arguments sum to zero, for second- and higher-order estimates, are included. Their inclusion has a small effect compared to the other computation necessary.

TABLE 5

THIRD WIENER KERNEL TRANSFORM ESTIMATE — LOW-PASS SYSTEM AFTER 5000 AVERAGES

(Nyquist frequency = 32)

Frequency W_1, W_2, W_3	$\lvert K_3 \rvert$	$\lvert \hat{K}_3 \rvert$	$\hat{\sigma}$	$\frac{\hat{\sigma}}{\lvert \hat{K}_3 \rvert}$
0,7,10	0.182	0.107	1.020	9.53
0,12,10	0.119	0.121	1.115	9.22
0,17,10	0.091	1.10	0.922	8.95
1,2,5	0.658	0.660	5.02	7.90
1,7,5	0.317	0.293	3.01	10.3
1,12,5	0.208	0.184	1.81	9.84
1,17,5	0.159	0.142	1.26	8.87
3,1,5	0.505	0.497	4.70	9.46
4,0,5	0.524	0.496	4.32	8.71
4,2,5	0.402	0.375	3.72	9.92
5,5,10	0.278	0.287	2.81	9.79
6,4,5	0.241	0.246	2.30	9.34
6,6,10	0.265	0.259	2.14	8.26
7,0,5	0.356	0.359	3.48	9.69
7,10,1	0.086	0.075	0.682	9.09
9,0,5	0.287	0.301	2.89	9.60
9,7,5	0.123	0.140	1.26	9.00
12,0,5	0.223	0.216	1.99	9.21
14,12,5	0.057	0.071	0.671	9.45
17,0,5	0.166	0.143	1.31	9.16
Average				9.27

5.2.1 Power Spectrum Estimate

The estimate of the power spectrum of a random process $x(t)$ is given by

$$\hat{P}(w) = \frac{|X(w)|^2}{2T} \tag{5.4}$$

where $X(w)$ is the Fourier transform of $x(t)$ over the interval $[-T,T]$. In practice, the discrete Fourier transform (DFT) is calculated via the fast Fourier transform (FFT) algorithm. The number of real multiplications necessary for the FFT of $x(t)$ of length N, from Appendix III, is

$$N \log_2 \frac{N}{2} + 2N \quad . \tag{5.5}$$

TABLE 6

THIRD WIENER KERNEL TRANSFORM ESTIMATE AFTER 5000 AVERAGES — RESONATOR SYSTEM

(Nyquist Frequency = 32)

Frequency w_1, w_2, w_3	$\lvert K_3 \rvert$	$\lvert \hat{K}_3 \rvert$	$\hat{\sigma}$	$\frac{\hat{\sigma}}{\lvert \hat{K}_3 \rvert}$
0,2,5	0.076	0.277	4.31	15.6*
0,2,7	0.116	0.142	2.29	16.1*
2,4,5	0.123	0.185	1.65	8.90
2,4,7	0.142	0.141	1.32	9.34
2,7,5	0.180	0.171	1.36	8.01
3,12,5	0.071	0.078	0.740	9.49
3,12,7	0.103	0.100	0.962	9.62
4,6,5	0.173	0.129	1.21	9.38
6,7,4	0.252	0.240	2.19	9.14
6,8,5	0.354	0.340	2.96	8.72
7,8,6	0.516	0.481	4.28	8.91
7,9,5	0.367	0.327	2.75	8.41
9,5,6	0.417	0.387	3.32	8.58
9,10,-5	0.291	0.313	2.74	8.75
9,10,7	0.424	0.377	3.29	8.73
9,12,5	0.143	0.107	0.954	8.92
9,15,7	0.134	0.146	1.31	8.97
Average				8.92

*These points where close to the region where two of w_1, w_2, w_3 sum to zero and are excluded from the average.

Since X(w) in general is complex, Equation (5.4) can be expressed as

$$\hat{P}(w) = \frac{1}{2T}\left[\mathrm{Re}(X(w))^2 + \mathrm{Im}(X(w))^2\right] \qquad (5.6)$$

where Re(X) and Im(X) are the real and imaginary parts of X, respectively. Equation (5.6) involves two real multiplications and, for real x, X(w) is conjugate symmetric, so that only N/2 estimate points are needed. The number of real multiplications for the power spectrum estimate of Equation (5.6) is therefore

$$\frac{N}{2} \cdot 2$$

plus those needed for the FFT, or, in total,

$$N \log_2 \frac{N}{2} + 3N \quad . \qquad (5.7)$$

5.2.2 First Wiener Kernel Transform Estimate

The first Wiener kernel transform estimate is the expected value

$$K_1(w) = E\frac{[X^*(w)Y(w)]}{C} \tag{5.8}$$

where X and Y are the Fourier transforms of the input and output signal. The asterisk indicates the complex conjugate. For this estimate both the input and output sequence are transformed, leading to twice the number of expression (5.5), or

$$2N \log_2 \frac{N}{2} + 4N \tag{5.9}$$

real multiplications. As with the previous estimate, $K_1(w)$ is symmetric, so that expression (5.8) need only be computed for N/2 points. Expression (5.8) involves one complex multiplication, but each one requires four real multiplications. The number of real multiplications needed, excluding the FFT, is

$$\frac{N}{2} \text{ points} \cdot \frac{4 \text{ real mult.}}{\text{point}} = 2N \text{ real mult.}$$

In total, adding (5.9),

$$2N \log_2 \frac{N}{2} + 6N \tag{5.10}$$

real multiplications are needed.

5.2.3 Second Wiener Kernel Transform Estimate

The second Wiener kernel transform estimate is the expected value,

$$\hat{K}_2(w_1,w_2) = \frac{E[X^*(w_1)\ X^*(w_2)\ Y(w_1+w_2)]}{2C^2} \tag{5.11}$$

As with the previous estimate, both the input and output are transformed, leading to

$$2N \log_2 \frac{N}{2} + 4N \tag{5.12}$$

real multiplications.

The original sampling of X(t) and Y(t) specifies the Nyquist frequency, or highest frequency where we have knowledge of Y(w). Thus we can only estimate $K_2(w_1,w_2)$ when

$$\left.\begin{aligned} |w_1| &\leq w_n \\ |w_2| &\leq w_n \\ |w_1 + w_2| &\leq w_n \end{aligned}\right\} \tag{5.13}$$

where w_n is the Nyquist frequency, which is shown in Fig. 32. When computing $\hat{K}_2$, use can be made of the kernel transform symmetry. The Fourier transforms X(w) and Y(w) are conjugate symmetric so that

$$\hat{K}_2(w_1,w_2) = \hat{K}_2^*(-w_1,-w_2) \quad . \tag{5.14}$$

the kernel transform is symmetric with respect to its arguments so that

$$\hat{K}_2(w_1,w_2) = \hat{K}_2(w_2,w_1) \quad . \tag{5.15}$$

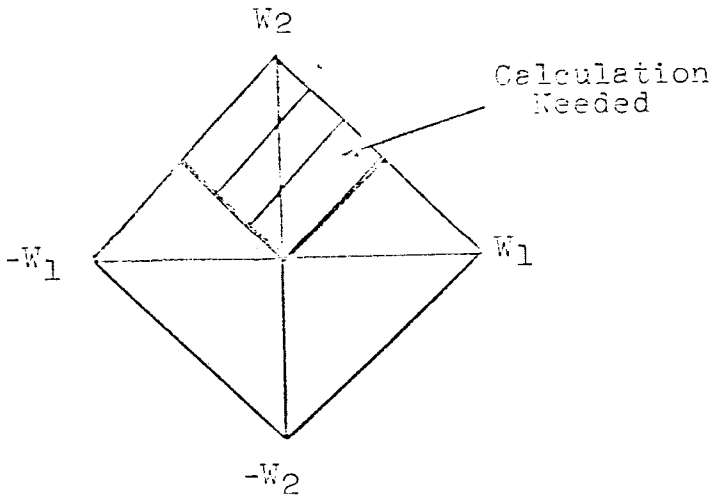


FIGURE 32. Regions of second kernel transform estimate.

Consequently, one-fourth of the points enclosed by the equations of (5.13) need be calculated, as shown by the shaded area of Fig. 32. The number of points in the shaded area is $N^2/8$. Expression (5.11) involves two complex multiplications. The number of real multiplications for the second-order estimator, excluding the FFT, is

$$\frac{N^2}{8} \text{ points} \cdot \frac{2 \text{ complex mult.}}{\text{point}} \cdot \frac{4 \text{ real mult.}}{\text{complex mult.}} = N^2 \text{ mult.} \qquad (5.16)$$

In total, from (5.12) and (5.16),

$$2 N \log_2 \frac{N}{2} + 4N + N^2 \qquad (5.17)$$

real multiplications complete the estimate.

5.2.4 Third Wiener Kernel Transform Estimate

The third-order Wiener kernel transform estimate is the expected value

$$\hat{K}_3(w_1,w_2,w_3) = \frac{E[X^*(w_1)X^*(w_2)X^*(w_3)Y(w_1+w_2+w_3)]}{6C^3} . \qquad (5.18)$$

As with the previous estimate, both the input and output are transformed, leading to

$$2N \log_2 \frac{N}{2} + 4 N \qquad (5.19)$$

real multiplications.

The symmetry noted for the second-order case (Equations 5.14) is again present because of the conjugate symmetric nature of X(w) and Y(w). This can be expressed as

$$K_3(w_1,w_2,w_3) = K_3^*(-w_1,-w_2,-w_3) \qquad (5.20)$$

which results in a reduction by a factor of 2 in the number of points to be calculated. The kernel symmetry is expressed by

$$\begin{aligned}\hat{K}_3(w_1,w_2,w_3) &= \hat{K}_3(w_1,w_3,w_2) \\ &= \hat{K}_3(w_3,w_1,w_2)\end{aligned} \tag{5.21}$$

etc.

This property results in further computation savings by a factor of 6. So far, the number of points to be calculated is

$$\frac{N^3}{12} \quad . \tag{5.22}$$

When estimating K_3, as with K_2, the original sampling specifies the Nyquist frequency. Therefore $K_3(w_1,w_2,w_3)$ can only be estimated when

$$\left.\begin{aligned} |w_1| &\le w_N \\ |w_2| &\le w_N \\ |w_3| &\le w_N \\ |w_1 + w_2 + w_3| &\le w_N \end{aligned}\right\} \quad . \tag{5.23}$$

the resultant reduction in the number of points of the estimate is not as obvious as for the second-order case. French and Butz [4] state that the reduction can be shown to be by a factor of n! where n in this case is 3. This is shown to be true in Appendix IV, and the reduction is by a factor of 6. The number of points to be calculated is one-sixth of (5.22) or

$$\frac{N^3}{(12)(6)} \quad . \tag{5.23}$$

The expression (5.18) involves three complex multiplications. The number of real multiplications required, excluding the FFT, is

$$\frac{N^3}{(12)(6)} \text{ points} \cdot \frac{3 \text{ complex mult.}}{\text{point}} \cdot \frac{4 \text{ real mult.}}{\text{complex mult.}} = \frac{N^3}{6} \text{ real mult.} \tag{5.24}$$

In total, from (5.19) and (5.24),

$$2N \log_2 \frac{N}{2} + 4N + \frac{N^3}{6} \tag{5.25}$$

real multiplications are required for a single estimate.

From expressions (5.10), (5.17) and (5.25) a general expression for the approximate number of real multiplications for an estimate of the above form for the nth order is

$$2N \log \frac{N}{2} + 4N + \frac{2nN^n}{(n!)2} \tag{5.26}$$

or

$$2\left[N \log \frac{N}{2} + 2N + \frac{nN^n}{(n!)^2}\right] . \tag{5.27}$$

The terms of (5.27) resulting from the fast Fourier transforms, dependent on N, are independent of kernel order, n. As N and/or n become large, the term

$$\frac{nN^n}{(n!)^2}$$

accounts for most of the multiplications. If the kernel were only to be estimated for a limited range of frequencies, substantially less computation is required.

5.3 EXAMPLE OF THE NUMBER OF MULTIPLICATIONS FOR FREQUENCY DOMAIN ESTIMATES

In Section 5.1, results were presented suggesting that the single kernel transform estimate standard deviation is one, three and nine times the value of the kernel transform, for the first, second and third kernel, respectively. If this is true, the number of averages (used in taking the expectation) for a given estimate variance is known. This

follows if the variance decreases as 1/M where M is the number of independent estimates.

The number of averages necessary to achieve 1% variance for several frequency domain estimates is given in Table 7. The number of real multiplications necessary (from Equations 5.7, 5.10, 5.17 and 5.25) for a 1% variance with a data length of 64 is also given in Table 7. The fifth column of Table 7, "mult./point," is the total number of multiplications for the entire estimate, divided by the number of points of the estimate. This value is approximately proportional to the single estimate variance inferred from the experimental results.

The number of real multiplications necessary if one only needed to estimate the kernel transform of one frequency point is shown in Table 8. Ratios similar to those obtained for the previous case were obtained for the number of multiplications per point.

5.4 COMPARISON OF WIENER KERNEL AND KERNEL TRANSFORM ESTIMATE

A final comparison is made of the relative efficiency of the time- and frequency-domain Wiener kernel estimates. The estimate of the Wiener kernel discussed in Section 3 is of the form

$$k_n(\tau_1,\ldots,\tau_n) = \frac{1}{n!C^n} E\left[y(t)x(t-\tau_1)\ldots x(t-\tau_n)\right] \quad . \tag{5.28}$$

As before, n is the kernel order, and X(t) and y(t) are the input and output, respectively.

Consider the number of real multiplications needed for the above estimate of the nth kernel. The estimate, assuming an ergodic input process, is found from the time average

$$k_n(\tau_1,\ldots,\tau_n) = \frac{1}{n!C^n} \int y(t)x(t-\tau_1)\ldots x(t-\tau_n)dt \quad . \tag{5.29}$$

TABLE 7

COMPARISON OF MULTIPLICATION NECESSARY FOR VARIOUS ESTIMATES (DATA LENGTH 64, 1% VARIANCE).

Estimate	$\hat{\sigma}^2/\|K\|$	No. of Averages	No. of real mult.	No. points / Estimate	mult. / point	Ratio
Power Spectrum	1	100	5.1×10^4	64	936	1
1st Kernel Transform	1	100	1.0×10^5	64	1.9×10^3	2
2nd Kernel Transform	9	900	4.5×10^6	2048	2.2×10^3	2.3
3rd Kernel Transform	81	8100	3.6×10^8	4.4×10^4	8.3×10^3	8.8

TABLE 8

Estimate	No. mult. Single Estimate	No. of Averages	No. of mult. 1% Variance	Ratio
Power Spectrum	450	100	4.50×10^4	1
1st kernel transform	900	100	9.00×10^4	2
2nd kernel transform	904	900	8.13×10^5	18
3rd kernel transform	908	8100	7.35×10^6	163

Multiplication required for a single point (frequency) of various estimates (N=64, 1% variance).

If the estimate is calculated from discrete data of length M, Equation (5.29) becomes, for the first kernel estimate

$$\hat{k}_1(i) = \frac{1}{C\,M} \sum_{p=0}^{M-1} y(p)X(p-i) \tag{5.30}$$

for $i << M$. Implementing (5.30) requires M real multiplications. For an N point estimate where

$$N << M \quad ,$$

NM real multiplications are required. Similarly, the second-order estimate is calculated from discrete data by

$$\hat{k}_2(i,j) = \frac{1}{C^2 M} \sum_{p=0}^{M-1} y(p)X(p-i)X(p-j) \tag{5.31}$$

which requires 2M real multiplications. The second-order estimate is calculated for N^2 points where again $N << M$. In total, therefore,

$$2MN^2$$

real multiplications complete this second Wiener kernel estimate. In general,

$$nMN^n \tag{5.32}$$

multiplications are needed for the nth-order estimate.

The number of points at which the estimate need be calculated can be reduced if use is made of the kernel symmetry. Recall from Section 2.2 that

$$\begin{aligned} k_2(\tau_1,\tau_2) &= k_2(\tau_2,\tau_1) \\ k_3(\tau_1,\tau_2,\tau_3) &= k_3(\tau_1,\tau_3,\tau_2) \\ &= k_3(\tau_3,\tau_1,\tau_3) \\ &\text{etc.} \end{aligned} \tag{5.33}$$

This leads to a reduction in the number of points of the estimate to be calculated by n! The number of real multiplications required for the nth-order kernel estimate is from (5.32)

$$\frac{nMN^n}{n!} = \frac{MN^n}{(n-1)!} \quad . \tag{5.34}$$

The cross-correlation estimate of Equation (5.28) can also be calculated via the FFT algorithm. Consider the first kernel estimate

$$\hat{k}_1(\tau) = \int y(t)x(t-\tau)d\tau \quad . \tag{5.35}$$

To see how the FFT can be used, consider the Fourier transform of $\hat{k}_1(\tau)$, $\hat{K}_1(w)$, given by

$$\hat{K}_1(w) = \int \hat{k}_1(\tau)e^{-jw\tau}d\tau \quad . \tag{5.36}$$

Substituting (5.35) into (5.36) gives

$$\hat{K}_1(w) = \iint y(t)x(t-\tau)e^{-jw\tau}dtd\tau \quad . \tag{5.37}$$

Equation (5.37) can be simplified by noting, from (3.14),

$$\int x(t-\tau)e^{-jw\tau}d\tau = e^{-jwt}X^*(w) \quad , \tag{5.38}$$

Where X(w) is the Fourier transform of x(t) and the asterisk indicates the complex conjugate. Substituting (5.38) into (5.37) gives

$$\begin{aligned}\hat{K}_1(w) &= X^*(w)\int y(t)e^{-jwt}dt \\ &= X^*(w)Y(w) \quad .\end{aligned} \tag{5.39}$$

This is the form of the first kernel transform estimate without the expectation of Equation (3.22).

The above suggests a method for calculating the Wiener kernel estimate via the FFT. Instead of computing the cross correlation over M data points, compute it over N points at a time and average M/N complete estimates. This is accomplished by transforming the first N points of x(t)

and y(t) and calculating $\hat{K}_1(w)$ as in Equation (5.39). The inverse transform of $\hat{K}_1(w)$, which is $\hat{k}_1(\tau)$, is calculated.

The estimate is now found for all N points. This procedure is equivalent to calculating the estimate for a data length N as

$$\hat{k}_1(i) = \frac{1}{C\ (N-i)} \sum_{k=1}^{N} y(k)x(k-i) \tag{5.40}$$

for i = 0,...,N-1. It should be noted that a higher estimate variance will be evident as i increases to N. This is because each estimate point results from the average of N-i calculations.

Consider the second-order Wiener kernel estimate from (5.29),

$$\hat{k}_2(\tau_1,\tau_2) = \frac{1}{2C^2}\int y(t)x(t-\tau_1)x(t-\tau_2)dt \quad . \tag{5.41}$$

Let us evaluate (5.41) for a fixed τ_1 value. First calculate

$$z(t) = y(t)\ x(t-\tau_1) \tag{5.42}$$

for all t. Substituting (5.42) into (5.41) gives

$$\hat{k}_2(\tau_1,\tau_2) = \frac{1}{2C^2}\int z(t)x(t-\tau_2)dt \tag{5.43}$$

which is of the same form as Equation (5.35). As before, Equation (5.43) can be implemented with short data sequences via the FFT. The Fourier transform of $\hat{k}_2(\tau_1,\tau_2)$ is calculated, as in (5.39), as,

$$\hat{K}_2^{\tau_1}(w) = X^*(w)\ Z(w) \tag{5.44}$$

where the superscript τ_1 indicates that τ_1 is fixed. The inverse transform of $\hat{K}_2^{\tau_1}(w)$ gives

$$\hat{k}_2(\tau_1,\tau_2)$$

for τ_1 a fixed value, and

$$\tau_2 = 1,\ldots,N \quad .$$

This must be repeated N times for different τ_1. This procedure is equivalent to calculating the estimate of Equation (5.31) for M=N. Again, as τ_1 and τ_2 increase, the variance of the estimate increases. This is because each estimate point results from an average of N minus the greater of τ_1 and τ_2 calculations. From the data of length M, M/N are calculated and averaged.

As with the previous estimates, consideration is now given to the number of real multiplications required to implement the kernel estimate via the FFT. To implement Equation (5.39), X(t) and y(t) must be transformed which requires

$$2N \log_2 \frac{N}{2} + 4N \tag{5.45}$$

real multiplications as shown in Appendix III. In addition, Equation (5.39) requires N complex multiplications, or 4N real multiplications. The inverse FFT of $\hat{K}_1(w)$ requires

$$2N \log N$$

real multiplications, as shown in Appendix III, since $\hat{K}_1(w)$ is complex. In total,

$$2N \log_2 \frac{N}{2} + 4N + 4N + 2N \log_2 N$$

or

$$2N \left[4 + \log_2 \frac{N}{2} + \log N\right] \tag{5.46}$$

real multiplications are needed for the single estimate of length N. The original data are of length M, so that the above estimate is calculated M/N times and averaged. This leads to

$$2M \left[4 + \log_2 \frac{N}{2} + \log_2 N\right] \tag{5.47}$$

real multiplications.

To find the approximate number of multiplications for the second-order estimate consider Equation (5.42). This requires M real multiplications. Note, however that is this must be performed for all τ_1, leading to

$$N + N - 1 + N - 2 + \ldots + 0$$

real multiplications. The sum of this series is

$$\frac{N(N+1)}{2} \tag{5.48}$$

Next, both x(t) and z(t) must be transformed, leading to

$$N \log_2 \frac{N}{2} + 2N$$

real multiplications each. This must be performed N times for z(t) but only once for x(t), requiring

$$(N+1)(N \log_2 \frac{N}{2} + 2N) \tag{5.49}$$

real multiplications. Equation (5.44) requires N complex multiplications or 4 real multiplications which must also be performed N times, resulting in

$$4N^2 \tag{5.50}$$

real multiplications. The inverse transform of $\hat{K}_2^{\tau_1}(w)$ is calculated N times, requiring

$$2N^2 \log_2 N$$

real multiplications since $\hat{K}_2^{\tau_1}(w)$ is complex. In total, from Equations (5.48, (5.49) and (5.50), the total number of real multiplications is

$$N^2[\log_2 \frac{N}{2} + \log_2 N + 6\frac{1}{2}] + N[\log_2 \frac{N}{2} + \frac{5}{2}] \tag{5.51}$$

As with the previous estimate, the second-order estimate is calculated M/N

times and averaged. This requires, from Equation (5.51),

$$M[N\ (\log_2 \frac{N}{2} + \log_2 N + 6\frac{1}{2}) + \log_2 \frac{N}{2} + \frac{5}{2}] \tag{5.52}$$

real multiplications.

Lee and Schetzen [3] estimated the second Wiener kernel of a nonlinear system. The kernel estimate was from data, M, equal to 30,000. That is, each point of the estimate was the result of 30,000 averages, and achieved an RMS error of 0.6% of the maximum value. A similar experiment was performed with the second kernel transform estimate. The mean squared error was found inversely proportional to the number of averages. After 100 averages the MS error was found to be 3.55%. Extrapolating, it would take 98,600 averages to reach an RMS error of 0.6%.

Consider the number of real multiplications required for the second kernel estimate for M equal to 30,000 and the second-order kernel transform estimate averaged 98,600 times for N equal to 128.

From Equation (5.34) the number of real multiplications required for the cross-correlation estimate is

$$\frac{MN^n}{(n-1)!} = (30,000)(128)^2$$

$$= 4.92 \times 10^8 \quad . \tag{5.53}$$

As noted earlier, the variance of the cross-correlation estimate via the FFT increases as τ_1 and τ_2. This is because each estimate point results from the average of N minus the greater of τ_1 and τ_2, times M/N calculations. Consider the case when the data length M is equal to 45,000. The number of real multiplications, from Eq. (5.52) is (for N equal to 128)

$$1.13 \times 10^8 \tag{5.54}$$

The number of real multiplications needed for each kernel transform estimate, from Equation (5.17), is

$$2N \log_2 \frac{N}{2} + 4N + N^2 = 2 \cdot 128 \cdot 6 + 4 \cdot 128 + (128)^2$$
$$= 1.84 \times 10^4 \quad . \tag{5.55}$$

this must be calculated 98,600 times, which requires

$$98{,}600 \cdot 1.84 \times 10^4 = 1.81 \times 10^9 \tag{5.56}$$

real multiplications. Dividing (5.53), (5.54) and (5.56) by 4.92×10^8 gives the following ratios of the number of real multiplications required.

Cross correlation	1.0
Kernel transform	3.6
Cross correlation via FFT	0.23

(5.57)

From (5.56) it is seen that the Wiener kernel transform estimate requires 3.6 times as many multiplications as the cross-correlation estimate. It should also be noted that it requires 3.3 times as many averages, which accounts for a large part of the difference.

Consider the number of real multiplications needed per estimate point for the above estimates. The cross-correlation estimate yields N^2 estimate points. The kernel transform estimate results in $N^2/2$ estimate points, from Fig. 32. The cross-correlation estimate gives N^2 points; however, as τ_1 and τ_2 increase, so does the estimate variance. The estimate points result from

$$(N-\tau) \frac{M}{N}$$

average where τ is the greater of τ_1 and τ_3. In this case, M is 45,000 if the points where τ_1 and τ_2 are less than 0.7N, each will be the result of between 45,000 and 15,000 averages. These points will be considered useful for this comparison. These points number $N^2/2$. For the above estimates, then, the number of real multiplications per point is

Cross correlation	3.0×10^4	—	1	
Kernel Transform	2.2×10^5	—	7	(5.58)
Cross correlation via FFT	1.4×10^4	—	0.46	

Clearly, using the FFT in the computation of the cross-correlation estimate results in a saving in required multiplication.

An alternate method of estimating the Wiener kernel would be to first estimate the kernel transform and then, after averaging, compute the inverse DFT. The two-dimensional inverse FFT requires

$$2N[2N \log_2 N] \tag{5.59}$$

real multiplications, or in this case 4.58×10^5. Since this method of estimating the Wiener kernel requires more real multiplications than either direct kernel estimate, it may be of limited value.

One can also use the cross-correlation estimate to estimate the kernel transform by performing the two-dimensional FFT of $\hat{k}_2(\tau_1,\tau_2)$.

One can also use the cross-correlation method to estimate the kernel transform. As above, after the Wiener kernel is estimated after averaging, the FFT is computed. The two-dimensional FFT requires, as in (5.59), 4.58×10^5 real multiplications. If the kernel transform is estimated in this manner the number of real multiplications required are

Cross-correlation estimate	4.92×10^8
Cross-correlation via FFT	1.13×10^8

(5.60)

Either way of calculating the cross-correlation estimate and then transforming requires fewer real multiplications than directly estimating the kernel transform.

Before ending this discussion of the number of multiplications necessary for various estimates, consider the situation where an estimate of $K_2(w_1,w_2)$ over a particular region (bandwidth) is required. If only one point of $\hat{K}_2(w_1,w_2)$ is to be estimated from Equations (5.16)

and (5.17),

$$2N \log_2 \frac{N}{2} + 4N + 8 \tag{5.61}$$

real multiplications are required. For the above example, where N is 128, Equation (5.61) becomes

$$2.06 \times 10^3$$

real multiplications. This is repeated 98,600 times requiring 2.03×10^8 real multiplications.

The above estimate could also be found through cross correlation and transforming. To estimate the kernel by either cross correlation estimate calculation, refer to Equations (5.53) and (5.54). The two-dimensional FFT is calculated by first transforming that data column by column (N times) and then doing the same to the rows. If only one point of the kernel transform is to be estimated the kernel estimate is transformed over each row and then over one column (N+1) times. This requires

$$(N+1)\ 2N \log_2 N = 2.3 \times 10^5 \tag{5.62}$$

real multiplications. Adding (5.62) to Equations (5.53) and (5.54) gives the number of real multiplications for a single-point kernel transform estimate below:

Cross correlation	4.92×10^8	
Kernel transform	2.03×10^8	(5.63)
Cross correlation FFT	1.13×10^8	

For applications where the cross correlation via FFT estimate is unacceptable because of high estimate variance, the Wiener kernel estimate offers multiplication savings by a factor of 2.4.

6.0 CONCLUSIONS

Any interconnection of single-valued time-invariant, zero-memory nonlinear, and linear systems can be described by its Volterra series. The Volterra series expansion for a nonlinear system is a series of generalized convolution integrals, the Volterra functionals. These functionals are integrals over generalized impulse responses, the Volterra kernels.

The output of a nonlinear system of the above type which is due to a white Gaussian noise input can be described by the Wiener series expansion. The terms of the Wiener series are orthogonal functions of the Volterra convolution integrals. Since a Wiener functional is orthogonal to all others, previously determined terms of the series are not changed by taking more terms in the approximation. The Wiener functionals are functions of the system input and the Wiener kernels.

The use of the Wiener kernel description of nonlinear systems is dependent on knowledge of the Wiener kernels. Two methods for estimating the Wiener kernels from knowledge of the system input and output signals have been discussed. The Wiener kernels can be estimated by a cross-correlation technique which is due to Lee and Schetzen [3]. Estimates of the Fourier transforms of the Wiener kernels have been presented in this work.

In order to verify the Wiener kernel transform estimate method it was applied to several nonlinear systems in a series of computer simulation experiments. The systems consisted of feed-through interconnections of linear and nonlinear subsystems. The estimates of the Fourier

transforms of the first, second and third kernels were obtained, and found to approach the theoretical kernel transforms.

During the experiments the variance of the first-, second- and third-order estimate magnitude was estimated. Experimental results suggest that the single kernel transform estimate standard deviation is one, three and nine times the value of the kernel transform for the first, second and third kernels, respectively.

As a measure of the computation, the approximate number of real multiplications necessary for the kernel transform estimates has been considered. For a data length of 64 and 1% variance the first, second, and third kernel transform estimates need 2, 2.3 and 8.8 times as many real multiplications per estimate point, as did the power spectrum estimate. The number of real multiplications needed substantially increased with increases order n and data length. The amount of multiplication required for the cross-correlation estimate has also been considered.

For comparison of the time- and frequency-domain second Wiener kernel estimates, the number of real multiplications necessary for each to achieve a desired RMS error was calculated. The large variance of the kernel transform estimate resulted in approximately three times more averaging being required. This resulted in 3.6 times more real multiplications. Calculation of the cross-correlation estimate using the fast Fourier transform (FFT) results in considerable computation savings but increased estimate variance. When the Wiener kernel transform is to be estimated over few points (narrow bandwidth), the second kernel transform estimate showed a saving of real multiplications by a factor of 2.4 over the cross-correlation method.

REFERENCES

1. George, D., Continuous Nonlinear Systems, Technical Report 355, Research Laboratory of Electronics, M.I.T., Cambridge Mass. 24 July 1959, pp. 2-3.

2. Dainty, C., and Shaw, R., Image Science, Academic Press, London, 1974.

3. Lee, Y.W. and Schetzen, M., "Measurement of the Wiener kernels of a Nonlinear system by Cross-correlation," Int.J. Control Vol. 2 pp. 237-254, 1965.

4. French, A.S. and Butz, E.G., "Measuring the Wiener Kernels of a Nonlinear System using the Fast Fourier Transform Algorithm," Int.J. Control vol. 17, pp. 529-539, 1973.

5. Volterra, N., Theory of Functionals and of Integral and Integro-Differential Equations, Blackie and Sons, London, 1930.

6. Thomas, E.J., "Some Considerations on the Application of Volterra Representation of Nonlinear Networks to Adaptive Echo Cancellors," Bell Telephone Labs Mem. 31, March 1969.

7. Wiener, N., Nonlinear Problems in Random Theory, M.I.T. Press, Cambridge Mass. 1958. pp. 28-38.

8. Papoulis, A., Probability, Random Variables and Stochastic Processes, McGraw-Hill Book Co., New York, 1965, pp. 343-344.

9. Davenport, W., and Root, W., An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Co., New York, 1958, pp. 106-107.

10. Rabiner L.R., and Gold, B., Theory and Application of Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N.J. 1975, pp. 22-24.

11. Oppenheim, A.V. and Schafer, R.W., Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N.J. 1975, pp. 15.

12. Freund, J.E., Mathematical Statistics, Prentice-Hall Englewood Cliffs, N.J. 1962, pp. 210.

13. Cooley, J., Lewis, P., and Welch, P., "The Fast Fourier Transform: Programming Considerations in the Calculation of Sine, Cosine and Laplace Transforms," J. Sound Vib., vol 12, pp. 315-337, July 1970.

14. Marmarelis, P.Z. and Naka, K., "White-Noise Analysis of a Neuron Chain: An Application of Wiener Theory," Science, vol. 175, pp. 1276-1278, March 1972.

15. Marmarelis, P.Z., "Nonlinear Identification of Bioneural Stems Through White Noise Stimulation," Proc. 13th Joint Automatic Control Conf. (Stanford, Calif.) 1972, pp. 117-126.

16. Bansal, V.S., "Volterra Series Analysis of a Diode-Ring Multiplier," Indian J. Pure and Applied Physics, vol. 2, 1973 pp. 529-532.

17. Selby, M. ed. Standard Mathematical Tables, The Chemical Rubber Co., Cleveland Ohio, 1970 pp. 19.

APPENDIX I

Second-Order Wiener Kernel Transform Estimate

The estimate of $K_2(w_1,w_2)$ is given by $E\left[\hat{K}_2(w_1,w_2)\right]$ where

$$\hat{K}(w_1,w_2) = X^*(w_1)\ X^*(w_2)\ Y(w_1+w_2) \quad . \tag{1}$$

The superscript $\hat{}$ denotes the estimate from information over the interval [-T,T]. The complex conjugate of X is X* and Y is the output. Equation (1) may be written as

$$\hat{K}_2(w_1,w_2) = \frac{1}{2T}\int_{-T}^{T} X(t_1)e^{jw_1t_1}dt \int_{-T}^{T} x(t_2)e^{jw_2t_2}dt_2 \int_{-T}^{T} y(t)e^{-j(w_1+w_2)t_3}dt_3 \tag{2}$$

$$= \frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} y(t)X(t_1)X(t_2)e^{j(w_1t_1+w_2t_2)}e^{-j(w_1+w_2)t_3}dt_1dt_2dt_3 \ . \tag{3}$$

The system output may be expressed in terms of the Wiener functional expansion

$$y(t) = k_o + \int k_1(\tau)\ X(t-\tau)d\tau$$

$$+ \iint k_2(\tau_1,\tau_2)X(t+\tau_1)X(t-\tau_2)d\tau_1d\tau_2 - C\int k_2(\tau_1\tau)d\tau \ ^1, \tag{4}$$

where the input process variance is σ^2. The limits of integration are $[-\infty,\infty]$ unless otherwise indicated. The integral notation will follow convention with the inner integral sign associated with the first variable of integration.

Taking the expected value of $\hat{K}_2(w_1,w_2)$ and letting T approach ∞, substituting (4) into (3) gives

[1]The input process power density spectrum is C watts/Hz.

$$\lim_{T\to\infty}\bar{E}[\hat{K}_2] = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\overline{k_o x(t_1)X(t_2)}\,e^{j(w_1t_1+w_2t_2)}\,e^{-j(w_1+w_2)}\,dt_1dt_2dt$$

$$+\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_1(\tau)\overline{X(t-\tau)X(t_1)X(t_2)}[\cdot]dt_1dt_2dt\,d\tau$$

$$+\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_2(\tau_1,\tau_2)\overline{X(t-\tau_1)X(t-\tau_2X(t_1)X(t_2)}[\cdot] \qquad (5)$$

$$\cdot dt_1dt_2dtd\tau_1d\tau_2$$

$$-C\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_2(\tau,\tau)\overline{X(t_1X(t_2)}[\cdot]dt_1dt_2dtd\tau \quad .$$

The first term of (5) is

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_o\overline{X(t_1)X(t_2)}[\cdot]dt_1dt_2dt \qquad (6)$$

Since the input process, X, is white,

$$\overline{X(t_1)X(t_2)} = \sigma^2\delta(t_1-t_2) \qquad (7)$$

Substituting (7) into (6) gives

$$\lim_{T\to\infty}\frac{1}{2T}\sigma^2\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_o\,\delta(t_1-t_2)e^{j(w_1t_1+w_2t_2)}\,e^{-j(w_1+w_2)t}\,dt_1dt_2dt \qquad (8)$$

Integrating over t_1, t_2 and t as $T \to \infty$ gives

$$\sigma^2 k_o \delta^2 (w_1 + w_2) \quad . \tag{9}$$

The second term of (5) is

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_1(\tau)\ \overline{X(t-\tau)X(t_1)X(t_2)}\ [\cdot]dt_1dt_2dt\ d\tau \tag{10}$$

This expression is zero since, Wiener [7] has shown that for a Gaussian process

$$\overline{X(t_1)X(t_2)\ldots X(t_n)} = 0 \tag{11}$$

for n odd.

The third term of (5) is

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_2(\tau_1,\tau_2)X(t-\tau_1)X(t-\tau_2)X(t_1)X(t_2)[\cdot]$$

$$\cdot dt_1dt_2dtd\tau_1d\tau_2 \quad . \tag{12}$$

To simplify (11) consider that for Gaussian random variables [7]

$$\overline{X_1X_2X_3X_4} = \overline{(X_1X_2)}\ \overline{(X_3X_4)} + \overline{(X_1X_3)}\ \overline{(X_2X_4)} + \overline{(X_1X_4)}\ \overline{(X_2X_3)} \quad . \tag{13}$$

Substituting (13) into (12) using (7) gives

$$\lim_{T\to\infty} \frac{\sigma^4}{2T} \int\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_2(\tau_1,\tau_2) e^{j(w_1t+w_2t_2)} e^{-j(w_1+w_2)t}$$

$$\cdot \left[\delta(t_1-t_2)\delta(\tau_2-\tau_1) + \delta(t_1-t+\tau_1)\delta(t_2-t+\tau_2)\right. \tag{14}$$

$$\left. + \delta(t_1-t+\tau_2)\delta(t_2-t+\tau_1)\right] dt_1dt_2dtd\tau_1d\tau_2 \quad .$$

The above expression (14) is the sum of three terms. The first term of (14) is

$$\lim_{T\to\infty} \frac{\sigma^4}{2T}\int\int\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_2(\tau_1,\tau_2)e^{j(w_1t_1+w_2t_2)}e^{-j(w_1+w_2)t} \tag{15}$$

$$\cdot\ \delta(t_1-t_2)\delta(\tau_1,\tau_2)dt_1dt_2dtd\tau_1d\tau_2 \quad .$$

Integrating with respect to τ_1 over $[-\infty,\infty]$ and t_1 over $[-T,T]$ as T approaches ∞ we get

$$\lim_{T\to\infty} \frac{\sigma^4}{2T}\int\int_{-T}^{T}\int_{-T}^{T} k_2(\tau_2,\tau_2)e^{jt_2(w_1+w_2)}e^{-jt(w_1+w_2)}dt_2dtd\tau_2 \tag{16}$$

Expression (16) can be rewritten as

$$\lim_{T\to\infty} \frac{\sigma^4}{2T}\int k_2(\tau_2,\tau_2)d\tau_2 \int_{-T}^{T} e^{jt_2(w_1+w_2)}dt_2\int_{-T}^{T} e^{-j(w_1+w_2)t}dt \quad . \tag{17}$$

Since

$$\lim_{T\to\infty} \int_{-T}^{T} e^{-jt_2(w_1+w_2)}dt_2 = \delta(w_1+w_2). \tag{18}$$

and

$$\lim_{T\to\infty} \int_{-T}^{T} e^{j(w_1+w_2)t}dt = 2T\ \delta(w_1+w_2) \tag{19}$$

expression (17) becomes

$$\sigma^4\delta^2(w_1+w_2)\int k_2(\tau_2,\tau_2)d\tau \quad . \tag{20}$$

The second term of (14) is

$$\lim_{T\to\infty} \frac{\sigma^4}{2T} \int\int \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} k_2(\tau_1,\tau_2) e^{j(t_1w_1+t_2w_2)} e^{-jt(w_1+w_2)}$$

$$\cdot \delta(t_1-t_2+\tau)\delta(t_2-t+\tau_2)dt_1dt_2dtd\tau_1d\tau_2 \quad . \tag{21}$$

Integrating over t_1 and t_2 and letting $T \to \infty$, gives

$$\lim_{T\to\infty} \int\int \int_{-T}^{T} k_2(\tau_1,\tau_2) e^{-j(w_1\tau_1+w_2\tau_2)} e^{jt(w_1+w_2)} e^{-jt(w_1+w_2)} \tag{22}$$

$$\cdot dtd\tau_1d\tau_2 \quad .$$

The terms involving t inside the integral cancel. Noting that

$$\int_{-T}^{T} dt = 2T \quad , \tag{23}$$

expression (22) becomes

$$\sigma^4 \int\int k_2(\tau_1,\tau_2) e^{-j(w_1\tau_1+w_2\tau_2)} d\tau_1d\tau_2 \quad . \tag{24}$$

Integrating (24) with respect to τ_1 and τ_2 yields

$$\sigma^4 K_2(w_1,w_2) \quad . \tag{25}$$

The third term of (14) is

$$\lim_{T\to\infty} \frac{\sigma^4}{2T} \int\int \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} k_2(\tau_1,\tau_2) e^{j(w_1t_1+w_2t_2)} e^{-jt(w_1+w_2)} \tag{26}$$

$$\cdot \delta(t_1-t+\tau_2)\delta(t_2-t+\tau_1)dt_1dt_2dtd\tau_1d\tau_2 \quad .$$

Integrating with respect to t_1 and t_2, as $T \to \infty$, expression (26) becomes

$$\lim_{T\to\infty} \frac{\sigma^4}{2T}\int\int\int_{-T}^{T} k_2(\tau_1,\tau_2)e^{-j(w_2\tau_1+w_1\tau_2)} e^{j(w_1+w_2)t} e^{-j(w_1+w_2)t} \cdot dtd\tau_1 d\tau_2 \quad . \tag{27}$$

As in (23), after integrating with respect to t as $T \to \infty$, expression (27) becomes

$$\sigma^4\int\int k_2(\tau_1,\tau_2)e^{-j(w_2\tau_1+w_1\tau_2)} d\tau_1 d\tau_2 \quad . \tag{28}$$

Integrating with respect to τ_1 and τ_2 yields

$$\sigma^4 H_2(w_2,w_1) \quad . \tag{29}$$

After substituting (20), (25) and (29), expression (13) is now

$$\sigma^4\left[\delta(w_1+w_2)\int k_2(\tau,\tau)d\tau + H_2(w_1,w_2) + H_2(w_2,w_1)\right] \quad . \tag{30}$$

Wiener (1958) shows that the Wiener kernels are, or can be made, symmetrical without any loss of generality. That is,

$$k_2(\tau_1,\tau_2) = k(\tau_2,\tau_1) \quad . \tag{31}$$

Consider the Fourier transform of $k_2(\tau_1,\tau_2)$:

$$K_2(w_1,w_2) = \int\int k_2(\tau_1,\tau_2)e^{-j(w_1\tau_1+w_2\tau_2)} d\tau_1 d\tau_2 \quad . \tag{32}$$

Substituting (31) into (32) gives

$$K_2(w_1,w_2) = \int\int k_2(\tau_1,\tau_2)e^{-j(w_1\tau_1+w_2\tau_2)} d\tau_1 d\tau_2 \quad . \tag{33}$$

Integrating with respect to τ_1 and τ_2 yields

$$K_2(w_1,w_2) = K_2(w_2,w_1) \quad . \tag{34}$$

Therefore expression (30) becomes

$$\sigma^4\left[2K_2(w_1,w_2) + \delta^2(w_1+w_2)\int k_2(\tau,\tau)d\tau\right] \quad . \tag{35}$$

The last term of (5) is

$$-\lim_{T\to\infty}\frac{-\sigma^4}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_2(\tau,\tau)e^{-j(w_1t_1+w_2t_2)} \cdot e^{-t(w_1+w_2)}\delta(t_1-t_2)dt_1dt_2dtd\tau \quad . \tag{36}$$

Integrating with respect to t_1, t_2 and t and letting $T \to \infty$ gives

$$-\sigma^4\,\delta^2(w_1+w_2)\int k_2(\tau,\tau)d\tau \quad . \tag{37}$$

The sum of expressions (9), (35) and (37) is the expected value of $\hat{K}_2(w_1,w_2)$ as T approaches ∞, from (5),

$$\lim_{T\to\infty} E[\hat{K}_2] = \sigma^2 k_o \delta^2(w_1+w_2) + 2\sigma^4\, H_2(w_1,w_2) \quad . \tag{38}$$

Rewriting (38) using (1)

$$K_2(w_1,w_2) = \lim_{T\to\infty}\frac{1}{2\sigma^4}E\,[X^*(w_1)X^*(w_2)Y(w_1+w_2)] - \frac{k_o\sigma^2(w_1+w_2)}{2\sigma^2} \tag{39}$$

or

$$K_2(w_1,w_2) = \lim_{T\to\infty}\frac{1}{2\sigma^4}E[X^*(w_1)X^*(w_2)Y(w_1+w_2)] \tag{40}$$

for $w_1 \neq -w_2$.

APPENDIX II

Third-Order Wiener Kernel Transform Estimate

The estimate of $K_3(w_1,w_2,w_3)$ is given by $E\left[\hat{K}(w_1,w_2,w_3)\right]$ where

$$\hat{K}_3(w_1,w_2,w_3) = X^*(w_1)X^*(w_2)X^*(w_3)Y(w_1+w_2+w_3). \tag{1}$$

The superscript $\hat{}$ denotes the estimate from information over the interval [-T,T]. The complex conjugate of X is X* and Y is the output. Equation (1) may be written as

$$\hat{K}_3(w_1,w_2,w_3) = \frac{1}{2T}\left[\int_{-T}^{T} x(t_1)e^{jw_1t_1}dt \int_{-T}^{T} x(t_2)e^{jw_2t_2} \int_{-T}^{T} x(t_3)e^{jw_3t_3} dt_3 \bullet \int_{-T}^{T} y(t)e^{-j(w_1+w_2+w_3)} dt\right] \tag{2}$$

Equation (2) can be rewritten as

$$\hat{K}_3(w_1,w_2,w_3) = \frac{1}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} y(t)x(t_1)x(t_2)x(t_3)e^{j(w_1t_1+w_2t_2+w_3t_3)} e^{-j(w_1+w_2+w_3)t} dt_1dt_2dt_3dt \quad . \tag{3}$$

The system output may be expressed in terms of the Wiener functional expansion

$$y(t) = k_o + \left[\int k_1(\tau)x(t-\tau)dt + \iint k_2(\tau_1,\tau_2)x(t-\tau_1) \bullet x(t-\tau_2)d\tau_1d\tau_2 - C\int k_2(\tau,\tau)d\tau + \iiint k_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1d\tau_2d\tau_3 - 3C\iint k_3(\tau,\tau_1,\tau_1)x(t-\tau)d\tau d\tau_1\right] \tag{4}$$

where the power density spectrum of the input is C watts/HZ. The limits of integration are $[-\infty,\infty]$ unless otherwise indicated. The integral notation will follow convention with the inner integral sign associated with the first variable of integration.

Taking the expected value of $\hat{K}_3(w_1,w_2,w_3)$ and letting T approach ∞, substituting (4) into (3) gives, assuming ergodicity,

$$\lim_{T\to\infty} E[\hat{K}_3] = \lim_{T\to\infty} \frac{1}{2T}\Bigg[\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_o\, \overline{x(t_1x(t_2)x(t_3)}\, e^{j(w_1t_1+w_2t_2+w_3t_3)}$$

$$\bullet\, e^{-j(w_1+w_2+w_3)t}\, dt_1dt_2dt_3dt + \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_1(\tau)\ .$$

$$\overline{x(t-\tau)x(t_1)x(t_2)x(t_3)}\,[\bullet]dt_1dt_2dt_3dtd\tau$$

$$+\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_2(\tau_1,\tau_2)\ \overline{x(t-\tau_1)x(t-\tau_2)x(t_1)x(t_2)x(t_3)}\,[\bullet]$$

$$\bullet\, dt_1dt_2dt_3dtd\tau_1d\tau - C\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_2(\tau,\tau)\ \overline{x(t_1x(t_2)x(t_3)}\,[\bullet]$$

$$dt_1dt_2dt_3dtd\tau + \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_3(\tau_1,\tau_2,\tau_3)$$

$$\overline{x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)x(t_1x(t_2)}\ \bullet\ X(t_3)\,[\bullet]dt_1dt_2dt_3dtd\tau_1d\tau_2d\tau_3$$

$$-3C\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_3(\tau,\tau_1,\tau_1)\ \overline{x(t-\tau)x(t_1x(t_2)x(t_3)}\,[\bullet]$$

$$\bullet\, dt_1dt_2dt_3dtd\tau d\tau_1 \qquad . \tag{5}$$

The bar in the equation indicates the time average.

The right-hand side of (5) is seen as the sum of six terms. The first term of Equation (5) is

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_o \overline{x(t_1)x(t_2)x(t_3)[\bullet]}\, dt_1 dt_2 dt_3 dt \quad . \tag{6}$$

Equation (6) can be simplified since Wiener [7] has shown that for a Gaussian white process

$$\begin{aligned} \overline{x(t_1)x(t_2)\ldots x(t_n)} &= 0 \qquad \text{for } n \text{ odd} \\ &= \sigma^{2n} \sum_{ij} \prod \delta(t_i - t_j) \quad \text{for } n \text{ even} \end{aligned} \tag{7}$$

where σ^2 is the input process variance.

The sum of (7) is over all ways of dividing n objects into distinct pairs and the product is over all pairs formed in this manner. The expression (6) is equal to zero.

The second term of Equation (5) is

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_1(\tau)\, \overline{x(t-\tau_1)x(t_1)x(t_2)x(t_3)} \bullet e^{j(w_1t_1+w_2t_2+w_3t_3)} e^{-j(w_1+w_2+w_3)t} dt_1 dt_2 dt_3 dt d\tau_1 \quad . \tag{8}$$

Substituting (7) into (8) leads to four terms of the form

$$\lim_{T\to\infty} \frac{\sigma^2 1}{2T} \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_1(\tau)\delta(t_1-t_2) e^{j(w_1t_1+w_2t_2+w_3t_3)} e^{-j(w_1+w_2+w_3)t} dt_1 dt_2 dt d\tau \quad . \tag{9}$$

Integrating over t_1 and t_2 yields

$$\sigma^2\delta(w_1+w_2)\int_{-T}^{T}\int_{-T}^{T}k_1(\tau)[\bullet]d\tau \tag{10}$$

where $\delta(t)$ is the unit impulse function. The expression (10) is zero except where $w_1 = -w_2$. The other terms of (8) lead to expressions which are zero except where two arguments of w_1, w_2, w_3 sum to zero.

The third term of (5) is

$$\lim_{T\to\infty}\frac{1}{2T}\int\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_2(\tau_1,\tau_2)\ x(t-\tau_1)x(t-\tau_2)x(t_1)x(t_2)x(t_3)$$

$$[\bullet]dt_1dt_2dt_3dtd\tau_1d\tau_2 \quad . \tag{11}$$

The expression (11) is zero because of (7).

The fourth term of (5) is

$$\lim_{T\to\infty}\frac{C}{2T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_2(\tau,\tau)[\bullet]\ \overline{x(t_1)x(t_2)x(t_3)}\ dt_1dt_2dt_3dtd\tau \tag{12}$$

which is also zero because of (7).

The fifth term of (5) is

$$\lim_{T\to\infty}\frac{1}{2T}\int\int\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}k_3(\tau_1,\tau_2,\tau_3)[\bullet]\ \overline{x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)}$$

$$\overline{x(t_1x(t_2)x(t_3)}\ dt_1dt_2dt_3dtd\tau_1d\tau_2d\tau_3 \quad . \tag{13}$$

Using (7) to simplify the terms under the bar in (13) leads to the sum of fifteen terms. Six are of the form

$$\lim_{T\to\infty} \frac{\sigma^6}{2T} \int\int\int \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} k_3(\tau_1,\tau_2,\tau_3)\,[\bullet]\,\delta(t_3+\tau_1-t)\,\delta(t_1+\tau_2-t)$$

$$\delta(t_2+\tau_3-t)\,dt_1\ldots d\tau_3 \quad . \tag{14}$$

Integrating (14) with respect to t_1,t_2,t_3,t_4 yields

$$\sigma^6 \int\int\int k_3(\tau_1,\tau_2,\tau_3)\, e^{-j(w_3\tau_1+w_2\tau_2+w_2\tau_3)}\, d\tau_1 d\tau_2 d\tau_3 \quad . \tag{15}$$

Integrating (15) with respect to τ_1,τ_2,τ_3 gives

$$\sigma^6 K_3(w_3,w_1,w_2) \quad . \tag{16}$$

Since the Wiener kernels and Wiener kernel transforms are symmetric, (16) becomes

$$\sigma^6 K_3(w_1,w_2,w_3) \quad . \tag{17}$$

The remaining nine terms, after applying (7) to (13), are of the form

$$\lim_{T\to\infty} \frac{1}{2T} \int\int\int \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} k_3(\tau_1,\tau_2,\tau_3)\,[\bullet]\,\delta(\tau_3-\tau_2)\,\delta(t_3-t_2)$$

$$\bullet\,\delta(t_1+\tau_1-t)\,dt_1\ldots d\tau_3 \quad . \tag{18}$$

Integrating (17) with respect to τ_3 and t_3 gives

$$\lim_{T\to\infty} \frac{1}{2T} \int\int \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} k_3(\ _1,\ _2,\ _3)\, e^{jt_2(w_2+w_3)}\, e^{jt_1w_1}$$

$$\cdot e^{-jt_4(w_1+w_2+w_3)} \delta(t_1+\tau_1-t)\,dt_1dt_2dtd\tau_1d\tau_2 \tag{19}$$

Integrating (19) over t_2 gives

$$\int e^{jt_2(w_2+w_3)}\,dt_2 = \delta(w_2+w_3) \tag{20}$$

and after integrating with respect to t_2 and t_4 becomes

$$\delta(w_2+w_3)\iint k_3(\tau_1,\tau_2,\tau_2)e^{-j\tau_1 w_1}d\tau_1d\tau_2 \tag{21}$$

which is zero except where $w_2 = -w_3$. The nine terms of the form of (18) lead to expressions as in (21) with arguments of the impulse functions (w_1+w_2), (w_2+w_3) etc. After simplification via (7) as in (14) and (18), equation (13) becomes

$$6\sigma^6 K_3(w_1,w_2,w_3) \tag{22}$$

for $w_1 \neq -w_2, w_2 \neq -w_3, w_1 \neq -w_3$.

The sixth term of (5) is

$$\lim_{T\to\infty} \frac{3C}{2T}\iiint_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} k_3(\tau,\tau_1,\tau_1)[\bullet]$$

$$\overline{x(t-\tau)x(t_1)x(t_2)x(t_3)}\;dt_1dt_2dt_3dtd\tau d\tau_1 \quad . \tag{23}$$

This expression leads to terms which are zero except where two of w_1,w_2,w_3 sum to zero as in Equations (8) — (10).

The expected value of $\hat{K}_3$, in the limit as T approaches is

$$6\sigma^6(K_3(w_1,w_2,w_3)$$

for $w_1 = -w_2, w_2 = -w_3, w_1 = -w_3$, or

$$K_3(w_1,w_2,w_3) = \lim_{T\to\infty} E \frac{\left[X^*(w_1)X^*(w_2)X^*(w_3)Y(w_1+w_2+w_3)\right]}{6\sigma^6} \qquad (24)$$

for $w_1 = -w_2, w_2 = -w_3, w_1 = -w_3$.

APPENDIX III

Discrete Fourier Transform of Real Data Sequence via FFT (radix 2)

For N complex data, radix 2, the FFT requires approximately

$$\frac{N}{2} \log_2 N$$

complex multiplications (Oppenheim & Schafer 1975).

The transform of a real sequence of length N can be implemented by an FFT of length $\frac{N}{2}$ complex points as follows (Cooley, Lewis & Welch 1970). Given the input sequence;

$$x(n) \quad , \quad n = 0,1,\ldots,N-1.$$

1. Form
$$\left.\begin{aligned} x_1(n) &= x(2n) \\ x_2(n) &= x(2_n + 1) \end{aligned}\right\} \qquad (1)$$

which are the even and odd points of x(n).

2. Form,
$$x_3(n) = x_1(n) + j\, x_2(n) \qquad n=0,..,\frac{N-1}{2} \qquad (2)$$

and transform via FFT of length $\frac{N}{2}$

$$x_3(n) \longleftrightarrow X_3(k) \qquad k=0,\ldots,\frac{N-1}{2} \qquad (3)$$

3. Separate $X_1(k)$ and $X_2(k)$, the transforms of $x_1(n)$ and $x_2(n)$, from $X_3(k)$,

$$\begin{aligned} X_1(k) &= \frac{1}{2}\left[X_3^*\left(\frac{N}{2} - k\right) + X(k)\right] \\ X_2(k) &= \frac{j}{2}\left[X_3^*\left(\frac{N}{2} - k\right) - X(k)\right] \end{aligned} \qquad (4)$$

Where the superscript * indicates complex conjugate and $j = \sqrt{-1}$.

4. The Transform of the original sequence x(n) is X(k) and is found as follows;

$$X(k) = \frac{1}{2}\left\{X_1(k) + X_2(k)\, W_N^{\,k}\right\} \qquad (5)$$

where $W_N k = \exp\left(\frac{2\pi jR}{N}\right)$.

Since x(n) is real, X(k) is conjugate symmetric so that (5) need only be computed for k = 0,1...,N/2, which involves N/2 complex multiplications.

The DFT of the N point real sequence involves $\frac{N}{4} \log_2 \frac{N}{2}$ complex multiplications for the length $\frac{N}{2}$ complex FFT. In addition Equation (5) involves $\frac{N}{2}$ more complex multiplications. One complex multiplication involves four real multiplications. In total,

$$4\left(\frac{N}{4} \log_2 \frac{N}{2} + \frac{N}{2}\right) = N \log_2 \frac{N}{2} + 2N \qquad (6)$$

real multiplications are required.

APPENDIX IV

The Argument Range of the Third Wiener Kernel Estimate

The third-order Wiener kernel transform estimate is given by

$$\hat{K}_3[w_1,w_2,w_3] = \frac{E\left[X^*(w_1)X^*(w_2)X^*(w_3)Y(w_1+w_2+w_3)\right]}{6\sigma^2} \tag{1}$$

where E indicates the expectation and σ^2 is the input process variance. The Fourier transform of the input, x, is X and of the output, y, is Y and the asterisk indicates the complex conjugate.

When estimating K_3, the original sampling specifies the Nyquist frequency, or highest frequency where one has knowledge of X(w) and Y(w). Consequently K_3 can only be estimated where

$$\left.\begin{aligned} |w_1| &\leq w_N \\ |w_2| &\leq w_N \\ |w_3| &\leq w_N \end{aligned}\right\} \tag{2}$$

and

$$|w_1 + w_2 + w_3| \leq w_N \tag{3}$$

where w_N is the Nyquest frequency.

The Equations (2) limit the absolute value of each of the arguments to less than or equal to w_N. The arguments can be thought of as three orthogonal coordinates. The Equations (2) restrict the arguments to a cube of side length $2w_N$ centered at the origin. Let a quadrant be the volume in which each argument is either positive or negative. There are eight quadrants per cube.

The volume restricted by Equations (2) and (3) will be considered as a fraction of the total cube volume. The cube and the volume enclosed by Equations (2) and (3) are shown in Fig. A1. The volume of the cube is

$$V_C = (\text{side length})^3 = w_N^3 \tag{4}$$

The volume enclosed between the origin and the shaded plane of Fig. A1 can be replotted as shown in Fig. A2. This solid is seen as a prismatoid whose upper base has zero area [17]. The volume is given by [17].

$$V = \frac{h}{6}\left[A_B + 4A_M\right] \tag{5}$$

where h is the altitude, A_B is the area of the base and A_M is the modsection area. The midsection is a cross section parallel to the base at a distance of h/2.

The base is an equilateral triangle with side length $\sqrt{2}w_N$, as shown in Fig. A3. The length p of Fig. A3 is, by the Pythagorean relation, equal to $\frac{\sqrt{3}}{\sqrt{2}} w_N$. The area of the base is

$$A_B = \text{side} \times p$$

where p is the triangle altitude, or

$$A_B = \frac{\sqrt{3}}{2} w_N^2 \quad . \tag{6}$$

The midsection is an equilateral triangle similar to the base, as shown in Fig. A3, with a side half that of the base. The midsection area is, therefore

$$A_M = \frac{\sqrt{3}}{4} w_N^2 \quad . \tag{7}$$

To find the altitude of the prismatoid, h, consider the diagram of A4. The line between the center of the base and a vertex bisects the vertex angle, since the base is an equilateral trangle. The angle α is therefore 30°. The length S of Fig. A4 is found since

$$\cos \alpha = \frac{\sqrt{2}\, w_N}{2P} \tag{8}$$

or

$$p = \frac{\sqrt{2}\, w_N}{2 \cos 30°}$$

$$= \frac{1.414}{2(0.866)} w_N$$

$$= 0.8164\, w_N \quad . \tag{9}$$

The altitude, h, is found by the Pythagorean relation,

$$1 = h^2 + (0.814)^2 w_N^2 \tag{10}$$

or

$$h = 0.578\, w_N \quad . \tag{11}$$

Substituting (6), (7) and (11) into (5) gives

$$V = \frac{0.568}{6} w_N \left[\frac{\sqrt{3}}{2} w_N^2 + \frac{\sqrt{3}}{2} w_N^2\right]$$

$$= \frac{(0.578)(1.732)}{6} w_N^3$$

$$= \frac{w_N^3}{6} \tag{12}$$

Equations (2) and (3) limit the estimate to one-sixth of those points possible if Equation (2) was the only restriction. The above analysis holds for all eight quadrants where $\hat{K}_3$ is calculated.

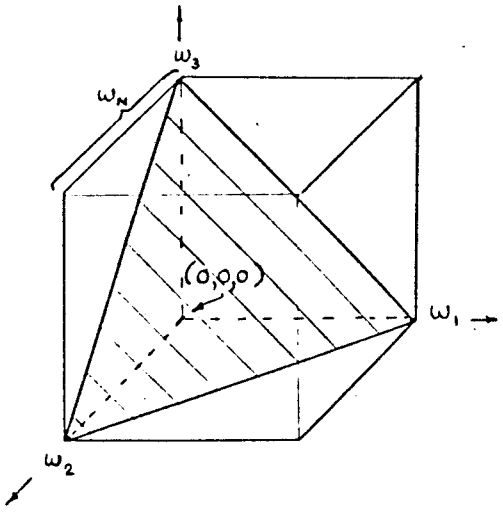


FIGURE A1. One quadrant of the third Wiener kernel transform.

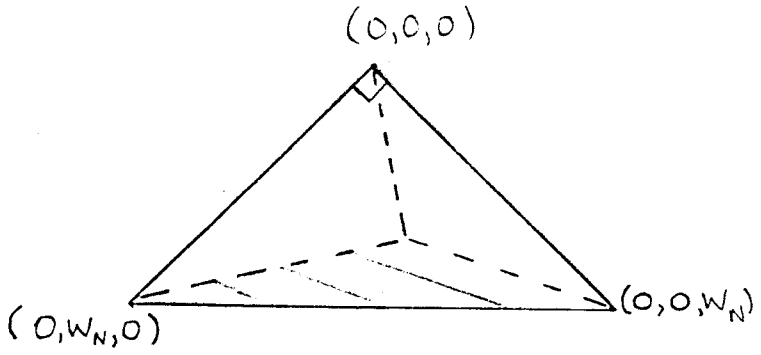


FIGURE A2. Prismatoid.

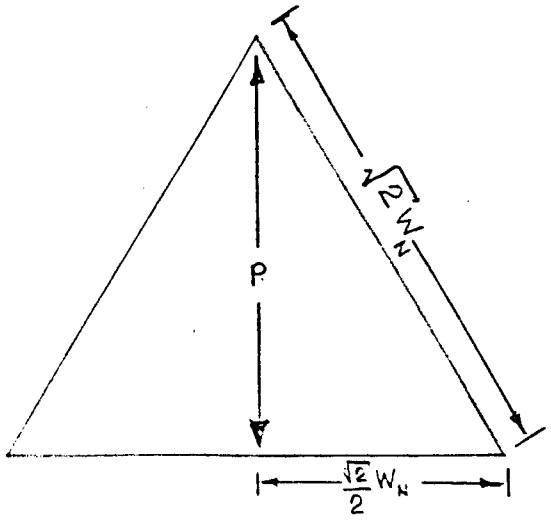


FIGURE A3. Prismatoid base.

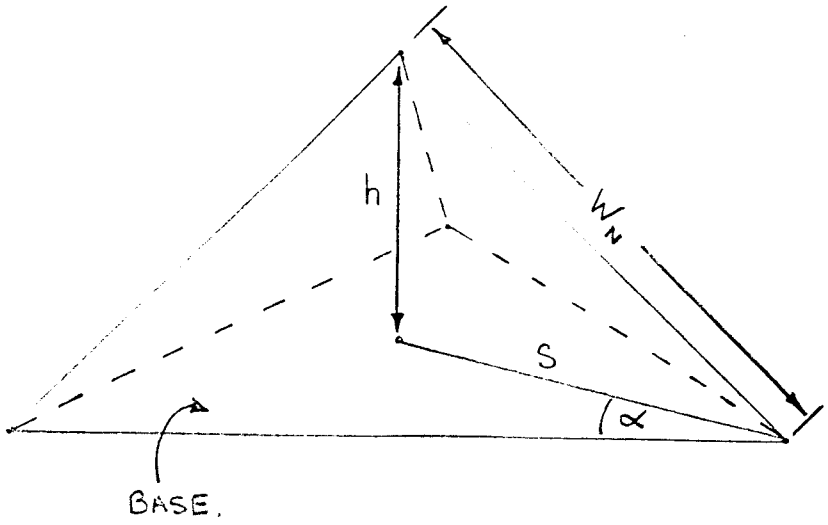
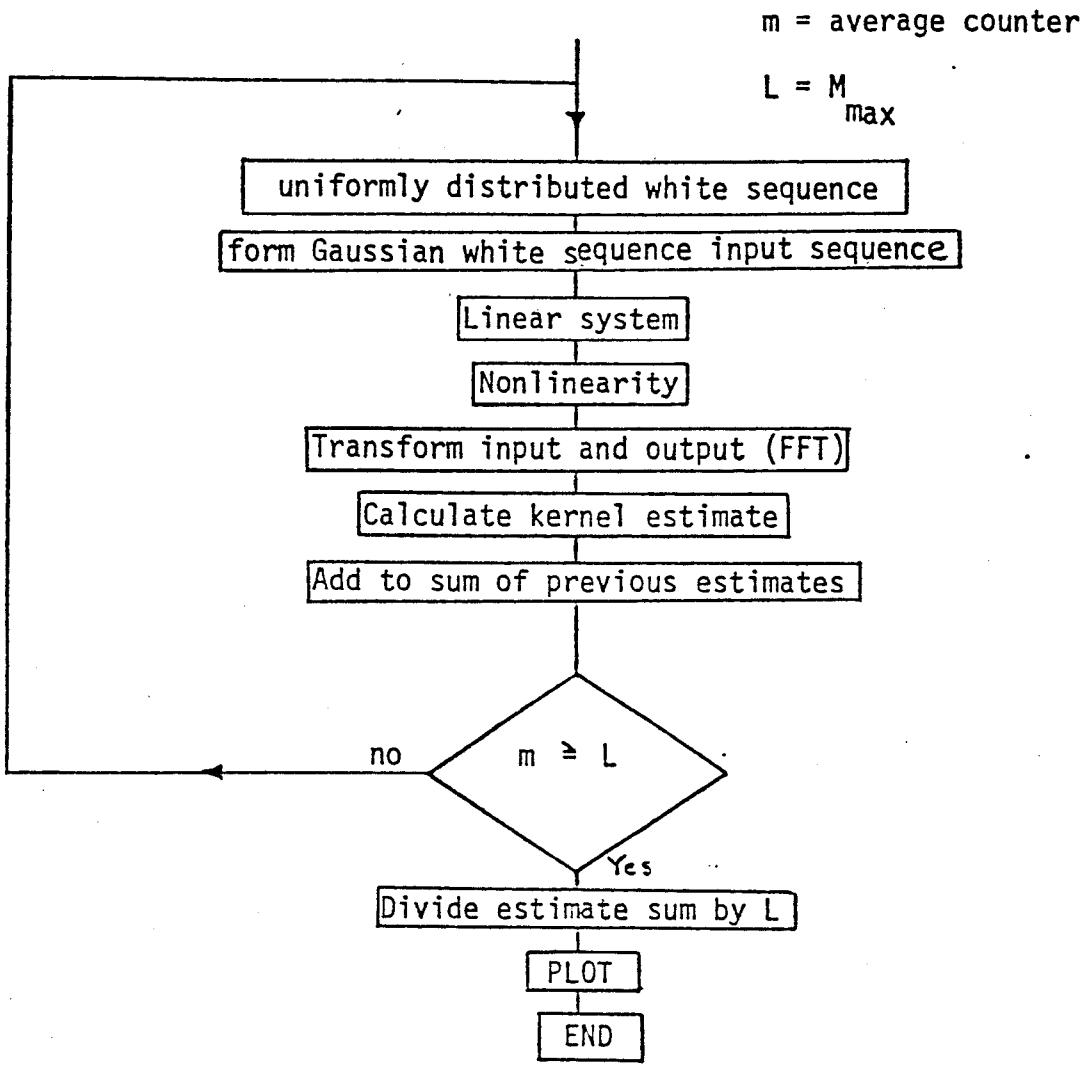


FIGURE A4. Prismatoid showing altitude h.

APPENDIX V

Computer Programs



Estimate experiment program flow chart.