# XIII. STATISTICAL COMMUNICATION THEORY


Prof. Y. W. Lee
Prof. A. G. Bose
Prof. I. M. Jacobs
Prof. J. B. Wiesner
V. R. Algazi
A. Bruce
D. F. De Long, Jr.
C. L. Feldman
F. J. Hernandez
H. Hemami
D. J. Sakrison
M. Schetzen
H. L. Van Trees, Jr.
S. Weinreb
C. E. Wernlein, Jr.
W. S. Widnall


## RESEARCH OBJECTIVES

This group is interested in a variety of problems in statistical communication theory. Our current research is concerned primarily with: measurement of correlation functions, location of noise sources in space by correlation methods, statistical behavior of coupled oscillators, nonlinear feedback systems, stochastic approximation methods in the analysis of nonlinear systems, measurement of the kernels of a nonlinear system, a problem in radio astronomy, and factors that influence the recording and reproduction of sound.

1. The measurement of the first-order and second-order correlation functions by means of orthogonal functions is being studied. Of primary concern are the measurement errors resulting from the truncation of the orthogonal set and from the use of finite time of integration.

2. Noise sources in space can be located by means of higher-order correlation functions. A study is being made of the errors, caused by finite observation time, in locating sources by this method.

3. Many physical processes may be phenomenologically described in terms of a large number of interacting oscillators. A study of these processes is producing some interesting results.

4. The design of a control system can be considered as a filtering problem with constraints imposed by fixed elements. By combining the functional power series and the differential equation methods of system characterization a formal solution to the problem can be found. Research is being conducted to determine the restrictions on the desired filtering operation and fixed elements that are necessary to achieve a practical system configuration.

5. Stochastic approximation methods have been considered for proving the convergence of certain iterative methods of adjusting the parameters of a system. The adjustment seeks to minimize the mean of some convex weighting function of the error. An investigation is being made of the types of systems and signals to which the methods are applicable.

6. A nonlinear system can be characterized by a set of kernels of all orders. The measurement of these kernels is a major problem in the theory of nonlinear systems. A method of measurement that depends upon crosscorrelation functions has been developed. Research on this problem is concerned primarily with the development of techniques that involve tape recording and digital computation, and the application of the method to various problems.

7. A project has been initiated that will have as its goal the measurement of the galactic deuterium-to-hydrogen abundance ratio. The approach to this problem will be based upon digital correlation techniques. The advantage of this method lies in the high degree of accuracy that can be obtained.

8. We are also studying the factors that influence the accurate recording and reproduction of sound. In this study the tools of statistical communication theory are applied to spectral analysis under different methods of recording, as well as to the computation and measurement of diffraction effects of the human head under various incident sound fields. In addition to the spectral studies, the transient behavior of the various links in

the reproduction process will be investigated. Associated with this project a filter of the Wiener-Lee type having controllable amplitude with a fixed phase over part of the audio spectrum will be constructed as a tool for studying the effects of magnitude and phase perturbations on sound signals.

Y. W. Lee

## A. MEASUREMENT OF THE KERNELS OF A NONLINEAR SYSTEM BY CROSSCORRELATION

In the Wiener theory of nonlinear systems (1) the input x(t) of a system A, as shown in Fig. XIII-1, is a white gaussian process. The output y(t) of the system is represented by the orthogonal expansion

$$y(t) = \sum_{n=1}^{\infty} G_n[h_n, x(t)] \tag{1}$$

in which $\{h_n\}$ is a set of kernels of the nonlinear system and $\{G_n\}$ is a complete set of orthogonal functionals. The orthogonal property of the functionals is expressed by the fact that the time average $\overline{G_n[h_n, x(t)]\, G_m[h_m, x(t)]} = 0$ for $m \neq n$. A nonlinear system

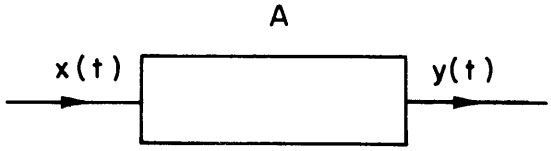



Fig. XIII-1. A nonlinear system with a white gaussian input.

is characterized by the set of kernels $\{h_n\}$. The first-order kernel $h_1(\tau_1)$, where $\tau_1$ is the time, is the linear kernel or the unit impulse response of a linear system. The second-order kernel, or the quadratic kernel, is $h_2(\tau_1, \tau_2)$. And the $n^{th}$-order kernel is $h_n(\tau_1, \ldots, \tau_n)$. The determination of the kernels is a major problem in the Wiener theory. Wiener expands the kernels in terms of a set of orthogonal functions such as the Laguerre functions. Thus if $\{\ell_m(\tau)\}$ is the set of Laguerre functions, then

$$\left.\begin{aligned} h_1(\tau_1) &= \sum_{m=0}^{\infty} c_m \ell_m(\tau_1) \\ h_2(\tau_1, \tau_2) &= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} c_{m_1 m_2} \ell_{m_1}(\tau_1)\, \ell_{m_2}(\tau_2) \\ &\vdots \\ h_n(\tau_1, \ldots, \tau_n) &= \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} c_{m_1 \ldots m_n} \ell_{m_1}(\tau_1) \cdots \ell_{m_n}(\tau_n) \end{aligned}\right\} \tag{2}$$

The determination of the coefficients of the Laguerre expansions, which leads to the determination of the G-functionals, is accomplished by a system of measurements. For reference, we list the first three terms of the G-functionals:

$$\left.\begin{aligned} G_1[h_1, x(t)] &= \int_{-\infty}^{\infty} h_1(\tau_1)\, x(t-\tau_1)\, d\tau_1 \\ G_2[h_2, x(t)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 - K\int_{-\infty}^{\infty} h_2(\tau_2, \tau_2)\, d\tau_2 \\ G_3[h_3, x(t)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_3)\, x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, d\tau_1 d\tau_2 d\tau_3 \\ &\quad - 3K\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_2)\, x(t-\tau_1)\, d\tau_1 d\tau_2 \end{aligned}\right\} \quad (3)$$

The leading term of the $n^{th}$-degree functional $G_n$ is a homogeneous functional of the $n^{th}$ degree, and the other terms of $G_n$ are each a homogeneous functional of degree lower than n. The $n^{th}$-degree homogeneous functional is

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n)\, x(t-\tau_1) \cdots x(t-\tau_n)\, d\tau_1 \cdots d\tau_n \quad (4)$$

The functional $G_n$ is constructed to be orthogonal to all functionals of degrees lower than n for a white gaussian input. The power density spectrum of this input is $\Phi_{xx}(\omega) = K/2\pi$ watts per radian per second so that the autocorrelation of the input is $\phi_{xx}(\tau) = Ku(\tau)$, where $u(\tau)$ is the unit impulse function.

We wish to introduce a method of determining the kernels of a nonlinear system that depends upon crosscorrelation techniques and avoids orthogonal expansions such as those of Eq. 2. This method is an extension of the crosscorrelation method that has been applied to linear systems (2).

## 1. Multidimensional-Delay White Gaussian Processes

First, we introduce a set of functionals that are formed by passing a white gaussian noise through a system of delay circuits as shown in Fig. XIII-2. In Fig. XIII-2(a) we have a delay circuit B with an adjustable delay time of $\sigma$ (seconds). The input x(t) is a white gaussian process whose power density spectrum is $K/2\pi$ watts per radian per second. The output $y_1(t)$ of the delay circuit is

$$y_1(t) = x(t-\sigma) \quad (5)$$

which can be written in the form of Eq. 4 as

$$y_1(t) = \int_{-\infty}^{\infty} u(\tau-\sigma)\, x(t-\tau)\, d\tau \tag{6}$$

The integral in Eq. 6 is a functional of the first degree. Let us call $y_1(t)$ a one-dimensional-delay white gaussian process.

In a similar manner we form a white gaussian process with a two-dimensional delay as shown in Fig. 2(b). Applying x(t) to the delay circuits $B_1$ and $B_2$ whose adjustable delay times are $\sigma_1$ and $\sigma_2$ and multiplying the outputs of $B_1$ and $B_2$ to form the output $y_2(t)$ of the system, we have

$$y_2(t) = x(t-\sigma_1)\, x(t-\sigma_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(\tau_1-\sigma_1)\, u(\tau_2-\sigma_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 \tag{7}$$

This expression is a homogeneous functional of the second degree. We shall refer to $y_2(t)$ as a two-dimensional-delay white gaussian process.

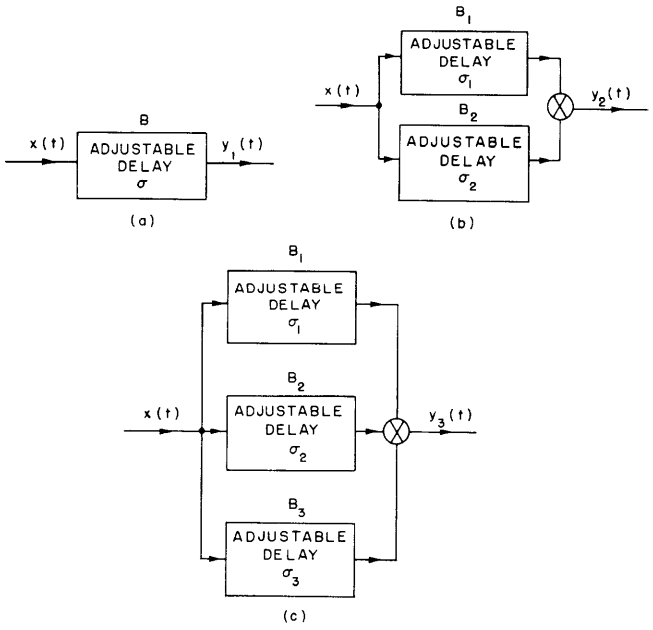



Fig. XIII-2. Delay circuits: (a) one-dimensional-delay circuit, (b) two-dimensional-delay circuit, (c) three-dimensional-delay circuit.

In Fig. XIII-2(c) we have x(t) applied to three delay circuits $B_1$, $B_2$, and $B_3$ whose adjustable delay times are $\sigma_1$, $\sigma_2$, and $\sigma_3$, and the outputs of the circuits are multiplied so that the product, which is the output of the whole system, is

$$y_3(t) = x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(\tau_1-\sigma_1)\, u(\tau_2-\sigma_2)\, u(\tau_3-\sigma_3)\, x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, d\tau_1 d\tau_2 d\tau_3 \qquad (8)$$

This is a three-dimensional-delay white gaussian process, and a homogeneous functional of the third degree. Obviously the n-dimensional-delay white gaussian process is

$$y_n(t) = (x-\sigma_1)\ldots(x-\sigma_n) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(\tau_1-\sigma_1)\ldots u(\tau_n-\sigma_n)\, x(t-\tau_1)\ldots x(t-\tau_n)\, d\tau_1\ldots d\tau_n \qquad (9)$$

The use of these functionals in the measurement of isolated kernels has been discussed by George (3). However, in the general case where a nonlinear system has more than one kernel he resorted to a Taylor series expansion. The method we present here does not depend upon expansions of the kernels in any form.

## 2. Determination of the First-Order Kernel

Now, consider that the nonlinear system A in Fig. XIII-3 is to be characterized; that is, the set of kernels $\{h_n\}$ of A are to be determined. By applying x(t) to A and

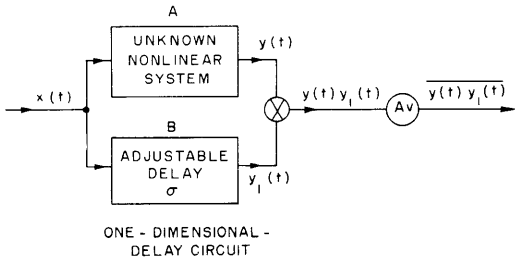



Fig. XIII-3. Measurement of the first-order kernel of a nonlinear system.

the delay circuit B of Fig. XIII-2(a), as indicated, then multiplying their outputs y(t) and $y_1(t)$, and finally averaging the product, we have

$$\overline{y(t)\, y_1(t)} = \overline{\left\{\sum_{n=1}^{\infty} G_n[h_n, x(t)]\right\} x(t-\sigma)} \qquad (10)$$

Since $x(t-\sigma)$ is a functional of the first degree, the functionals $G_n$, for $n > 1$, are orthogonal to $x(t-\sigma)$. Hence with $G_1$ as given in Eq. 3, we have

$$\overline{y(t)\, y_1(t)} = \overline{\left[\int_{-\infty}^{\infty} h_1(\tau_1)\, x(t-\tau_1)\, d\tau_1\right] x(t-\sigma)} = \int_{-\infty}^{\infty} h_1(\tau_1)\, \overline{x(t-\tau_1)\, x(t-\sigma)}\, d\tau_1$$

$$= \int_{-\infty}^{\infty} h_1(\tau_1)\, Ku(\sigma-\tau_1)\, d\tau_1 = Kh_1(\sigma) \tag{11}$$

Therefore, by applying a white gaussian process to the unknown nonlinear system A and to the one-dimensional-delay circuit B, and then crosscorrelating their outputs for various values of the delay time $\sigma$, we obtain the first-order kernel of the nonlinear system:

$$h_1(\sigma) = \frac{1}{K}\overline{y(t)\, y_1(t)} \tag{12}$$

3. Determination of the Second-Order Kernel

To measure the second-order kernel, we connect the system of Fig. XIII-2(b) to the unknown nonlinear system A in the manner shown in Fig. XIII-4. The average

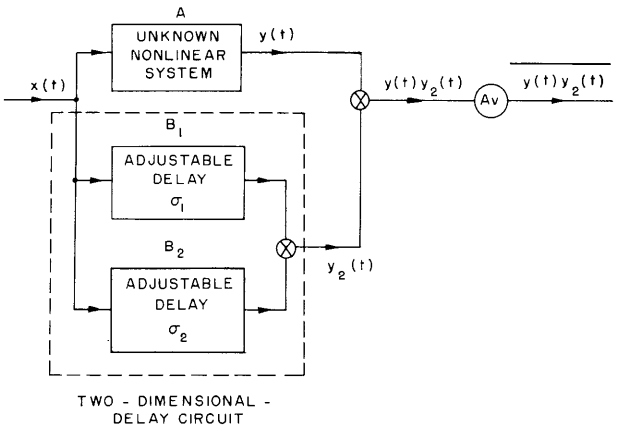



Fig. XIII-4. Measurement of the second-order kernel of a nonlinear system.

of the product of the outputs of the unknown nonlinear system and the two-dimensional-delay circuit is

$$\overline{y(t)\, y_2(t)} = \overline{\left\{\sum_{n=1}^{\infty} G_n[h_n, x(t)]\right\} x(t-\sigma_1)\, x(t-\sigma_2)} \tag{13}$$

We note that the $G_n$ for $n > 2$ are orthogonal to $x(t-\sigma_1)\, x(t-\sigma_2)$, which is a homogeneous functional of the second degree. Furthermore, for $n = 1$, we have

$$\overline{G_1[h_1, x(t)]\, x(t-\sigma_1)\, x(t-\sigma_2)} = \overline{\left[\int_{-\infty}^{\infty} h_1(\tau_1)\, x(t-\tau_1)\, d\tau_1\right] x(t-\sigma_1)\, x(t-\sigma_2)}$$

$$= \int_{-\infty}^{\infty} h_1(\tau_1)\, \overline{x(t-\tau_1)\, x(t-\sigma_1)\, x(t-\sigma_2)}\, d\tau_1 = 0 \tag{14}$$

(See Sec. XIII-C for the average of the product of gaussian variables.) Hence with $G_2$ as given in Eq. 3, Eq. 13 reduces to

$$\overline{y(t)\, x(t-\sigma_1)\, x(t-\sigma_2)} = \overline{G_2[h_2, x(t)]\, x(t-\sigma_1)\, x(t-\sigma_2)}$$

$$= \overline{\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 - K\int_{-\infty}^{\infty} h_2(\tau_2,\tau_2)\, d\tau_2\right] x(t-\sigma_1)\, x(t-\sigma_2)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)\, \overline{x(t-\tau_1)\, x(t-\tau_2)\, x(t-\sigma_1)\, x(t-\sigma_2)}\, d\tau_1 d\tau_2 - K^2 u(\sigma_1-\sigma_2)\int_{-\infty}^{\infty} h_2(\tau_2,\tau_2)\, d\tau_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)\, K^2[u(\tau_1-\tau_2)u(\sigma_1-\sigma_2)+u(\tau_1-\sigma_1)u(\tau_2-\sigma_2)+u(\tau_1-\sigma_2)u(\tau_2-\sigma_1)]\, d\tau_1 d\tau_2$$

$$- K^2 u(\sigma_1-\sigma_2)\int_{-\infty}^{\infty} h_2(\tau_2,\tau_2)\, d\tau_2$$

$$= K^2\left[u(\sigma_1-\sigma_2)\int_{-\infty}^{\infty} h_2(\tau_1,\tau_1)\, d\tau_1 + h_2(\sigma_1,\sigma_2) + h_2(\sigma_2,\sigma_1) - u(\sigma_1-\sigma_2)\int_{-\infty}^{\infty} h_2(\tau_2,\tau_2)\, d\tau_2\right]$$

$$= 2K^2 h_2(\sigma_1,\sigma_2) \tag{15}$$

Note that the kernels in Eq. 1 are symmetrical in the variables $\tau_1, \ldots, \tau_n$, so that for the second-order kernel we have $h_2(\tau_1, \tau_2) = h_2(\tau_2, \tau_1)$. The result in Eq. 15 means that if we apply x(t) to the unknown nonlinear system and to the two-dimensional-delay circuit and then crosscorrelate their outputs for various values of the delay times $\sigma_1$ and $\sigma_2$, we shall have the second-order kernel of the unknown nonlinear system given by

$$h_2(\sigma_1,\sigma_2) = \frac{1}{2K^2}\,\overline{y(t)\, y_2(t)} \tag{16}$$

4. Determination of the Third-Order Kernel

In a manner similar to the measurement of the first-order and second-order kernels we measure the third-order kernel of a nonlinear system as indicated in Fig. XIII-5. The crosscorrelation of the output of the unknown nonlinear system and the output of the three-dimensional-delay circuit as a function of the delay times $\sigma_1$, $\sigma_2$, and $\sigma_3$ is

$$\overline{y(t)\, y_3(t)} = \overline{\left\{\sum_{n=1}^{\infty} G_n[h_n, x(t)]\right\} x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)} \tag{17}$$

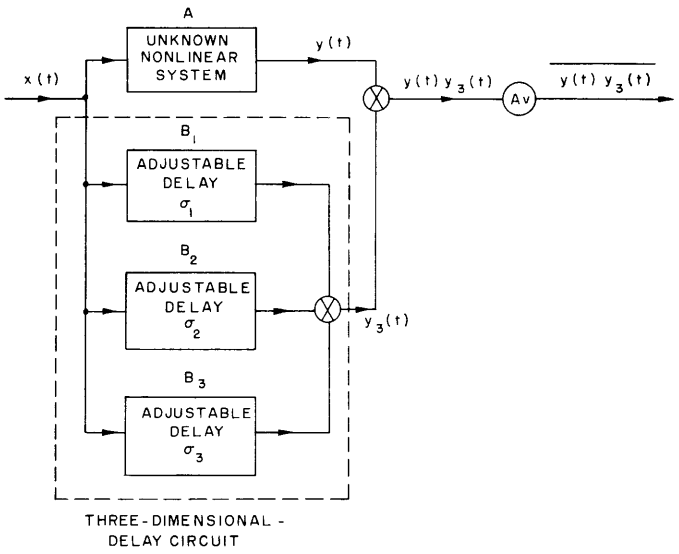



Fig. XIII-5. Measurement of the third-order kernel of a nonlinear system.

Since $x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)$ is a homogeneous functional of the third degree, it is orthogonal to $G_n$ for $n > 3$. When $n = 3$, we have, with $G_3$ as given in Eq. 3

$$\overline{G_3[h_3, x(t)]\, x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)}$$

$$= \overline{\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_3)\, x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, d\tau_1 d\tau_2 d\tau_3 \right.} \cdots$$

$$\cdots \overline{\left. - 3K \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_2)\, x(t-\tau_1)\, d\tau_1 d\tau_2 \right] x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_3)\, \overline{x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)}\, d\tau_1 d\tau_2 d\tau_3$$

$$- 3K \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1, \tau_2, \tau_2)\, \overline{x(t-\tau_1)\, x(t-\sigma_1)\, x(t-\sigma_2)\, x(t-\sigma_3)}\, d\tau_1 d\tau_2 \qquad (18)$$

The triple integral of Eq. 18 can be shown to be equal to

$$K^3 \left[ 6h_3(\sigma_1, \sigma_2, \sigma_3) + 3u(\sigma_2-\sigma_3) \int_{-\infty}^{\infty} h_3(\tau_1, \tau_1, \sigma_1)\, d\tau_1 + 3u(\sigma_1-\sigma_2) \int_{-\infty}^{\infty} h_3(\tau_3, \tau_3, \sigma_3)\, d\tau_3 \right.$$

$$\left. + 3u(\sigma_3-\sigma_1) \int_{-\infty}^{\infty} h_3(\tau_2, \tau_2, \sigma_2)\, d\tau_2 \right] \qquad (19)$$

and the last term of Eq. 18 can be shown to be equal to

$$-3K^3\left[u(\sigma_1-\sigma_2)\int_{-\infty}^{\infty}h_3(\tau_1,\tau_1,\sigma_3)\,d\tau_1+u(\sigma_3-\sigma_1)\int_{-\infty}^{\infty}h_3(\tau_1,\tau_1,\sigma_2)\,d\tau_1+u(\sigma_2-\sigma_3)\int_{-\infty}^{\infty}h_3(\tau_1,\tau_1,\sigma_1)\,d\tau_1\right] \tag{20}$$

Hence Eq. 18 reduces to

$$\overline{G_3[h_3, x(t)]\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)} = 6K^3h_3(\sigma_1, \sigma_2, \sigma_3) \tag{21}$$

To complete the evaluation of Eq. 17 we need to consider the crosscorrelation of the three-dimensional-delay white gaussian process with $G_1$ and $G_2$. The crosscorrelation involving $G_1$ is

$$\overline{G_1[h_1, x(t)]\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)} = \int_{-\infty}^{\infty}h_1(\tau_1)\,\overline{x(t-\tau_1)\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)}\,d\tau_1$$

$$= K^2\int_{-\infty}^{\infty}h_1(\tau_1)[u(\tau_1-\sigma_1)u(\sigma_2-\sigma_3)+u(\tau_1-\sigma_2)u(\sigma_1-\sigma_3)+u(\tau_1-\sigma_3)u(\sigma_1-\sigma_2)]\,d\tau_1$$

$$= K^2[u(\sigma_2-\sigma_3)h_1(\sigma_1)+u(\sigma_1-\sigma_3)h_1(\sigma_2)+u(\sigma_1-\sigma_2)h_1(\sigma_3)] \tag{22}$$

and the crosscorrelation involving $G_2$ is

$$\overline{G_2[h_2, x(t)]\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h_2(\tau_1, \tau_2)\,\overline{x(t-\tau_1)\,x(t-\tau_2)\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)}\,d\tau_1 d\tau_2$$

$$- K\int_{-\infty}^{\infty}h_2(\tau_2, \tau_2)\,\overline{x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)}\,d\tau_2 = 0 \tag{23}$$

since the mean of the product of an odd number of x's is zero.

Therefore our final result for Eq. 17 is

$$\overline{y(t)\,x(t-\sigma_1)\,x(t-\sigma_2)\,x(t-\sigma_3)}$$

$$= 6K^3h_3(\sigma_1, \sigma_2, \sigma_3) + K^2[u(\sigma_2-\sigma_3)h_1(\sigma_1)+u(\sigma_1-\sigma_3)h_1(\sigma_2)+u(\sigma_1-\sigma_2)h_1(\sigma_3)] \tag{24}$$

The first term on the right-hand side of this equation is the third-order kernel of the nonlinear system that we wish to determine. However, the second term on the same side of the equation gives rise to impulses when $\sigma_1 = \sigma_2$, $\sigma_1 = \sigma_3$, and $\sigma_2 = \sigma_3$. But when $\sigma_1 \neq \sigma_2$, $\sigma_1 \neq \sigma_3$, and $\sigma_2 \neq \sigma_3$, the term has zero value. Although theoretically the

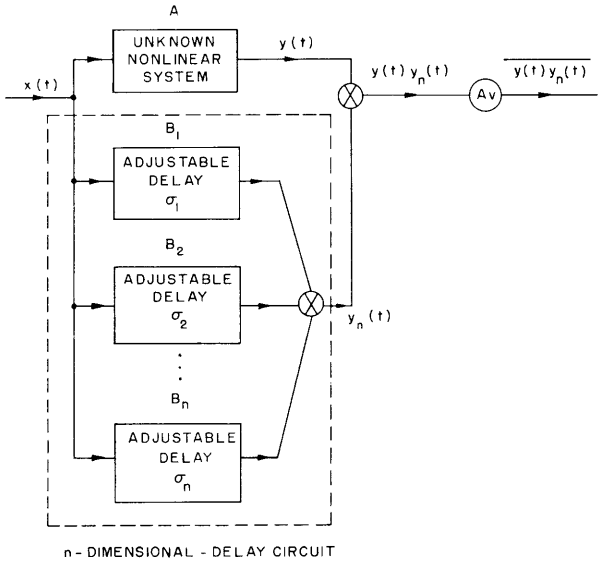



Fig. XIII-6. Measurement of the $n^{th}$-order kernel of a nonlinear system.

method does not yield the values of the third-order kernel at $\sigma_1 = \sigma_2$, $\sigma_1 = \sigma_3$, and $\sigma_2 = \sigma_3$, we should have no difficulty in the practical application of the method because we can come as close as we please to these points. Thus if we feed a white gaussian process to the unknown nonlinear system and to the three-dimensional-delay circuit and crosscorrelate their outputs for various values of the delays $\sigma_1$, $\sigma_2$, and $\sigma_3$, we can express the third-order kernel of the nonlinear system in terms of the crosscorrelation as

$$h_3(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{6K^3} \overline{y(t)\, y_3(t)} \qquad \text{for } \sigma_1 \neq \sigma_2,\ \sigma_2 \neq \sigma_3,\ \sigma_3 \neq \sigma_1 \tag{25}$$

## 5. Determination of the $n^{th}$-Order Kernel

To measure the $n^{th}$-order kernel in the manner shown in Fig. XIII-6 we have the crosscorrelation of the output of the unknown nonlinear system and the output of the n-dimensional-delay circuit given by

$$\overline{y(t)\, y_n(t)} = \overline{\left\{ \sum_{m=1}^{\infty} G_m[h_m, x(t)] \right\} x(t-\sigma_1)\, x(t-\sigma_2) \cdots x(t-\sigma_n)} \tag{26}$$

For $m > n$ the crosscorrelation is zero, and for $m = n$, we have

$$\overline{y(t)\, y_n(t)} = \overline{G_n[h_n, x(t)]\, x(t-\sigma_1)\, x(t-\sigma_2) \cdots x(t-\sigma_n)} \tag{27}$$

To evaluate this crosscorrelation, let us write the $n^{th}$-degree functional with $x(t-\sigma_1)\, x(t-\sigma_2) \ldots x(t-\sigma_n)$ as the leading term, in an orthogonal set $\{H_n[k_n, x(t)]\}$, as

$$H_n[k_n, x(t)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k_n(\tau_1, \ldots, \tau_n)\, x(t-\tau_1) \ldots x(t-\tau_n)\, d\tau_1 \ldots d\tau_n + F \tag{28}$$

where F is a sum of homogeneous functionals of degrees lower than n. It is clear from Eqs. 7 and 8 that $k_n(\tau_1, \ldots, \tau_n)$ in Eq. 28 is

$$k_n(\tau_1, \ldots, \tau_n) = u(\tau_1-\sigma_1) \ldots u(\tau_n-\sigma_n) \tag{29}$$

In terms of Eq. 28 the crosscorrelation of Eq. 27 is

$$\overline{y(t)\, y_n(t)} = \overline{G_n[h_n, x(t)]\{H_n[k_n, x(t)]-F\}} \tag{30}$$

Since $G_n$ is orthogonal to all functionals of degrees lower than n,

$$\overline{G_n[h_n, x(t)]\, F} = 0 \tag{31}$$

Hence

$$\overline{y(t)\, y_n(t)} = \overline{G_n[h_n, x(t)]\, H_n[k_n, x(t)]} \tag{32}$$

Formulas for the mean value of the product of functionals that are members of sets of orthogonal functionals are known (ref. 1, p. 41). In the present instance we can show that

$$\overline{G_n[h_n, x(t)]\, H_n[k_n, x(t)]} = n!\, K^n \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n)\, k_n(\tau_1, \ldots, \tau_n)\, d\tau_1 \ldots d\tau_n$$

$$= n!\, K^n \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n)\, u(\tau_1-\sigma_1) \ldots u(\tau_n-\sigma_n)\, d\tau_1 \ldots d\tau_n = n!\, K^n h_n(\sigma_1, \ldots, \sigma_n) \tag{33}$$

Note that $k_n$ is given by Eq. 29.

Combining Eqs. 31 and 33 in accordance with Eq. 30, which is the same as Eq. 27, we obtain

$$\overline{G_n[h_n, x(t)]\, x(t-\sigma_1) \ldots x(t-\sigma_n)} = n!\, K^n h_n(\sigma_1, \ldots, \sigma_n) \tag{34}$$

Our detailed work on $h_1$, $h_2$, and $h_3$ is in agreement with this general result. (See Eqs. 11, 15, and 21.)

We now return to Eq. 26 to consider the situation in which $m < n$. It is known that if m is even, then all of the terms in $G_m$ are functionals of even degrees; and if m is

odd, then all of the terms in $G_m$ are functionals of odd degrees. When n is even and m is even, the highest degree functional in $G_m$ involved in Eq. 26, for the case $m < n$, is of the degree n - 2. This condition means that the average

$$\overline{x(t-\tau_1) \cdots x(t-\tau_{n-2})\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \qquad \text{for } n > 2 \tag{35}$$

has to be taken in association with the highest degree functional in $G_m$. Since the mean of the product of gaussian variables can be reduced to a sum of products of the means of the products of pairs of the variables taken in all distinct ways (see Sec. XIII-C), and since in Eq. 35 there are two more $\sigma$'s than $\tau$'s, the result is that Eq. 35 is an impulse whenever two or more $\sigma$'s are equal and is zero otherwise. This fact is illustrated by Eq. 22 in the determination of $h_3$. Similarly, the average of the product of the n-dimensional-delay process and the other terms in $G_m$ for $m < n - 2$ is an impulse whenever two or more $\sigma$'s are equal and is zero otherwise. In other words, for n even and greater than 2, we have

$$\left.\begin{array}{r} \overline{x(t-\sigma_1) \cdots x(t-\sigma_n)} \\ \overline{x(t-\tau_1)\, x(t-\tau_2)\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \\ \vdots \\ \overline{x(t-\tau_1) \cdots x(t-\tau_{n-2})\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \end{array}\right\} = \begin{cases} 0 \text{ if no two } \sigma\text{'s are equal} \\ \text{an impulse if two or more } \sigma\text{'s are equal} \end{cases} \tag{36}$$

Furthermore, when n in $x(t-\sigma_1) \cdots x(t-\sigma_n)$ is even and m in $G_m$ is odd the crosscorrelation in Eq. 26 is zero because the mean of the product of an odd number of x's is zero.

When n in $x(t-\sigma_1) \cdots x(t-\sigma_n)$ is odd and m in $G_m$ is also odd, an argument similar to that just given will lead to the conclusion that for n odd and greater than 2

$$\left.\begin{array}{r} \overline{x(t-\tau_1)\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \\ \overline{x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \\ \vdots \\ \overline{x(t-\tau_1) \cdots x(t-\tau_{n-2})\, x(t-\sigma_1) \cdots x(t-\sigma_n)} \end{array}\right\} = \begin{cases} 0 \text{ if no two } \sigma\text{'s are equal} \\ \text{an impulse if two or more } \sigma\text{'s are equal} \end{cases} \tag{37}$$

This completes the discussion of Eq. 26 for $m > n$, $m = n$, and $m < n$. Combining the results that Eq. 26 is zero for $m > n$, that it is given by Eq. 34 for $m = n$, and that it has the properties of Eqs. 36 and 37 for $m < n$ we obtain the result that

$$h_n(\sigma_1, \ldots, \sigma_n) = \frac{1}{n!\, K^n} \overline{y(t)\, y_n(t)} \qquad \text{except when, for } n > 2, \text{ two or more } \sigma\text{'s are equal} \tag{38}$$

For the actual measurement of the kernels we can form the orthogonal functionals $H_n[k_n, x(t)]$ as given by Eqs. 28 and 29. By the use of these functionals we can determine the kernels $h_n(\sigma_1, \ldots, \sigma_n)$ for <u>all</u> values of the $\sigma$'s, without the restrictions stated in Eq. 38, by the method described in this report. We have not done so because without the additional complexity of forming the functionals $H_n[k_n, x(t)]$ we can come as close as we please to the set of points at which impulses occur.

## 6. Discussion

In comparison with the Wiener method of measurement of the kernels, the present method has the advantage of great simplicity. Digital computation and tape recording are particularly helpful in the application of the method. As we see from the theory of the method the only necessary data for the characterization of a nonlinear system – that is, the determination of its kernels of all orders – are the record of the white gaussian process that is fed into the nonlinear system and the corresponding output of the system. The record can be in the form of a twin-track recording on magnetic tape.

In the Wiener method of measurement the basis is the orthogonal expansion of the kernels and the representation of the orthogonal sets of functions by a system of linear networks and a system of nonlinear no-memory networks. Since in practical application the number of terms in an expansion must be finite, the Wiener method involves an error that is attributable to the truncation of the expansion. We know, however, that the error in the representation of a function by a finite orthogonal set of functions is the minimum integral square error. On the other hand, the method discussed in this report does not depend upon a series expansion of the kernels in any form. Hence another advantage of the method is that it involves no approximation error. In both methods, as we are aware, there is, among other errors, an error that is the result of using a finite time in taking the necessary average values.

We also note that the present method is a point-by-point method, whereas the Wiener method is, as pointed out before, a minimum-integral-square-error approximation method over the entire range of time. The determination of a set of coefficients determines the approximation over the entire range of time. We see that under certain circumstances these methods may complement each other. For instance, the Wiener method may indicate quickly the parts of the kernel curve that need greater details. These details may be more effectively obtained by the

present method. Again, in expanding the kernels by the Wiener method, we may wish to know whether the approximation is sufficiently good. A comparison of the approximation with the measurement by the present method should be a good check.

Y. W. Lee, M. Schetzen

## References

1. N. Wiener, Nonlinear Problems in Random Theory (Technology Press, M.I.T., and John Wiley and Sons, Inc., New York, 1958), pp. 28-38, 88-100.

2. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960), pp. 342-348.

3. D. A. George, Continuous Nonlinear Systems, Technical Report 355, Research Laboratory of Electronics, M.I.T., July 24, 1959, pp. 74-75.

## B. AN ITERATIVE PROCEDURE FOR SYSTEM OPTIMIZATION

[This report concludes the discussion of the filter optimization procedure that was introduced in Quarterly Progress Report No. 59, pages 98-105.]

We now consider condition iii. Condition iii(b) is always satisfied because the sequences v(m) and d(m) are uniformly bounded. In considering condition iii(a) it was assumed in the previous discussion that the term

$$F_n(\underline{x}_1) = E\{\| E\{\underline{Y}_n | \underline{x}_n, \underline{x}_1\} - E\{\underline{Y}_n | x_1\}\| \, | \underline{x}_1\} \tag{10}$$

approached zero at least as fast as $a_n/c_n$. This assumption is unrelated to the physical situation and is unduely presumptive in that it results in an estimate of the rate of convergence which is as rapid as that obtained when independent data are used for succeeding iterations. For this reason, we now impose the following restriction on the memory units of the filter:

$$|h_i(t)| \leq H_i e^{-\alpha t} \tag{11}$$

where $\alpha \geq a > 0$, $H_i \leq K < \infty$, and $i = 0, 1, \ldots, j$, and hypothesize a condition on the sequences v(m) and d(m). Let $g_1$ and $g_2$ be two continuous functions:

$$g_1 = g_1[s_1(t_1), \ldots, s_1(t_n), \ldots, s_j(t_1), \ldots, s_j(t_n), d(t_1), \ldots, d(t_n)]$$

$$g_2 = g_2[s_1(t_n+\tau_1), \ldots, s_1(t_n+\tau_m), \ldots, d(t_n+\tau_1), \ldots, d(t_n+\tau_m)]$$

where $t_j \geq t_i$; $\tau_j \geq \tau_i \geq 0$, with $j \geq i$; $|g_1| \leq G_1 < \infty$; and $|g_2| \leq G_2 < \infty$. We then make the following assumption concerning the rate at which the terms of the sequences v(m) and d(m) become independent:

$$\left|\overline{g_1 g_2} - \overline{g_1}\,\overline{g_2}\right| \leq K G_1 G_2 e^{-\beta \tau_1} \tag{12}$$

for all $\tau_1 \geq \tau_o$; $\tau_o < \infty$; $K < \infty$; and $\beta > 0$.

We shall now derive an estimate of $F_n(x_1)$ based only on the hypothesis expressed in inequality 12. To avoid notational difficulty, the discussion will be carried out in terms of the one-dimensional case shown in Fig. XIII-7; the methods used and results obtained carry over to the original k-dimensional case. The parameter x is now restricted to lie in a closed bounded interval X and the sequences v(m) and d(m) again assumed uniformly bounded. We will need to make the simplifying assumption that x, q(m), and d(m) are quantized. This is no practical restriction, since the iterative procedure is most likely to be carried out on a computer.

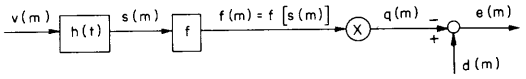



Fig. XIII-7. One-dimensional case of the filter of Fig. XI-5, Quarterly Progress Report No. 59, p. 100.

We first establish a simple moment theorem. Let x be a bounded random variable taking on the discrete values $x_i$, i = 0, 1, ..., N, with associated probabilities $p(x_i)$. Let Y be a bounded random variable. We wish to bound the quantity

$$E\{\left|E\{Y|x\} - E\{Y\}\right|\}$$

Let

$$f(x_i) = E\{Y|x_i\} - E\{Y\}$$

It is possible to find a polynomial

$$P(x_i) = \sum_{n=0}^{N} a_n (x_i)^n$$

with the property that $P(x_i) = f(x_i)$, i = 0, 1, ..., N. The coefficients $a_n$ can be uniformly bounded in terms of the maximum value of $|Y|$ and the quantization of x. Now

$$\overline{f^2} = \sum_{i=0}^{N} p(x_i)\, f^2(x_i) = \sum_{i=0}^{N} f(x_i)\, p(x_i) \sum_{n=0}^{N} a_n (x_i)^n = \sum_{n=0}^{N} a_n \overline{x^n f}$$

The assumption

$$\left|\overline{x^n Y} - \overline{x^n}\,\overline{Y}\right| = \left|\overline{x^n f}\right| \leq KAB\epsilon \tag{13}$$

where n = 0, 1, ..., N, i = 0, 1, ..., N, and

$$A = \sup\left|x^n\right| \qquad B = \sup\left|f\right|$$

thus implies

$$\overline{f^2} \leq K_1 \epsilon$$

since the $a_n$ are uniformly bounded in magnitude independently of $p(x_i)$. Thus by the Schwartz inequality,

$$|\bar{f}| = E\{|E\{Y|x\} - E\{Y\}|\} \leq (K_1)^{1/2} \epsilon^{1/2} \tag{14}$$

Inequality 13 still implies inequality 14 when x, instead of being a scalar, is a fixed m-tuple.

Now we consider $F_n(x_1)$. To clarify the expression for $F_n$, we shall denote by $e_n$ the data used to carry out the $n^{th}$ iteration (whether $e_n$ is a four-tuple, an eight-tuple, etc. is dependent upon how many samples are used for an iteration). Furthermore, let

$$\overline{\xi}^{e_n} \quad \text{and} \quad \overline{\xi}^{x_n}$$

denote averaging $\xi$ over $e_n$ and $x_n$, respectively. Then

$$F_n(x_1) = \int \left| \int Y_n(e_n, x_n)\, dP(e_n | x_n(x_1)) - \int Y_n(e_n, x_n)\, dP(e_n) \right| dP(x_n)$$

(where $x_n = x_n(x_1)$ is a family of random variables indexed by the parameter $x_1$), hence, using the moment theorem, we can show that

$$F_n(x_1) \leq (K_1)^{1/4} \epsilon_n^{1/4} \qquad \text{for all } x_1 \in X$$

if

$$\left| \overline{Y_n(e_n, x_n)\, x_n^m}^{x_n, e_n} - \overline{\overline{Y_n(x_n, e_n)}^{e_n}\, x_n^m}^{x_n} \right| \leq K_1^{1/2} \epsilon_n^{1/2} \tag{15}$$

for m = 0, 1, . . . , N and all $x_1 \in X$. But

$$\left| \overline{e_n^q x_n^m} - \overline{e_n^q}\,\overline{x_n^m} \right| \leq K_2 \epsilon_n \tag{16}$$

for q = 0, 1, . . . , M and for all $x_1 \in X$ ($x_n = x_n(x_1)$) again implies, by the moment theorem,

$$\int \left| \int x_n^m dP(x_n | e_n) - \int x_n^m dP(x_n) \right| dP(e_n) \leq K_2^{1/2} \epsilon_n^{1/2} \qquad \text{for all } x_1 \in X$$

Now let $B = \sup_{x \in X} |x|$ and $A = \sup |Y|$. Then, multiplying the integrand given above by $|Y_n|$, multiplying the right-hand side of the inequality by A, and using the relation $|\int a| \leq \int |a|$, we obtain inequality 15, with $K_1 = AK_2$. Thus inequality 16 implies

$$F_n(x_1) \leq (AK_2)^{1/4}\, \epsilon_n^{1/4} \qquad \text{for all } x_1 \in X$$

and we now need only estimate the moments in inequality 16. It should be noted that in inequality 16 $e_n^q$ is used symbolically. That is, $e_n$ might represent two samples of d(m), $d_a$ and $d_b$, and two samples of v(m), $v_a$ and $v_b$. Then $e_n^q$ is used to indicate all the products $d_a^p\, d_b^r\, v_a^s\, v_b^t$, where $p + r + s + t = q$.

Now

$$x_n = \sum_{j=1}^{n-1} \frac{a_j}{c_j}\, Y_j(e_j, e_{j-1}, \ldots, e_1, x_1)$$

hence

$$\left[\overline{e_n^q x_n^m} - \overline{e_n^q}\;\overline{x_n^m}\right] = \sum_{j_1=1}^{n-1} \cdots \sum_{j_m=1}^{n-1} \frac{a_{j_1}}{c_{j_1}} \cdots \frac{a_{j_m}}{c_{j_m}} \left[\overline{e_n^q Y_{j_1} \cdots Y_{j_m}} - \overline{e_n^q}\;\overline{Y_{j_1} \cdots Y_{j_m}}\right]$$

$$= \sum_{j_1=1}^{n-1} \sum_{j_2=1}^{j_1} \cdots \sum_{j_m=1}^{j_1} \frac{a_{j_1}}{c_{j_1}} \cdots \frac{a_{j_m}}{c_{j_m}} \left[\quad\right]$$

$$+ \sum_{j_2=1}^{n-1} \sum_{j_1=1}^{j_2-1} \sum_{j_3=1}^{j_2} \cdots \sum_{j_m=1}^{j_2} \frac{a_{j_1}}{c_{j_1}} \cdots \frac{a_{j_m}}{c_{j_m}} \left[\quad\right]$$

$$+ \ldots +$$

$$\sum_{j_m=1}^{n-1} \sum_{j_1=1}^{j_m-1} \cdots \sum_{j_{m-1}=1}^{j_m-1} \frac{a_{j_1}}{c_{j_1}} \cdots \frac{a_{j_m}}{c_{j_m}} \left[\quad\right]$$

$$= m \text{ terms} \left\{ \begin{array}{l} \displaystyle\sum_{j=1}^{n-1} \frac{a_j}{c_j} \left[\overline{e_n^q Y_j x_j^{m-1}} - \overline{e_n^q}\;\overline{Y_j x_j^{m-1}}\right] \\ \displaystyle + \sum_{j=1}^{n-1} \frac{a_j}{c_j} \left[\overline{e_n^q Y_j x_j^{m-2} x_{j-1}} - \overline{e_n^q}\;\overline{Y_j x_j^{m-2} x_{j-1}}\right] \\ + \ldots + \\ \displaystyle\sum_{j=1}^{n-1} \frac{a_j}{c_j} \left[\overline{e_n^q Y_j x_{j-1}^{m-1}} - \overline{e_n^q}\;\overline{Y_j x_{j-1}^{m-1}}\right] \end{array} \right. \tag{17}$$

Now, since

$$x_j = x_j(e_j, e_{j-1}, \ldots, e_1, x_1) \qquad Y_j = Y_j(e_j, e_{j-1}, \ldots, e_1, x_1)$$

we have, from assumption 12, for all $x_1 \in X$,

$$\left| \sum_{j=1}^{n-1} \frac{a_j}{c_j} \left[ \overline{e_n^q Y_j x_j^p x_{j-1}^r} - \overline{e_n^q} \, \overline{Y_j x_j^p x_{j-1}^r} \right] \right| \leq AB^{m-1} \left| e_n^q \right|_{max} K \sum_{j=1}^{n-1} \frac{a_j}{c_j} e^{-\beta s(n-j)}$$

where $p + r = m - 1$, and $s$ is the number of sample intervals allowed to elapse between the end of one iteration and the beginning of the next. Now, assuming that $a_j/c_j$ is monotonic, we have

$$\sum_{j=1}^{n-1} \frac{a_j}{c_j} e^{-\beta s(n-j)} \leq \frac{a_1}{c_1} e^{-\beta s(n-1)} + \frac{a_1}{c_1} \int_1^{n/2} e^{-\beta s(n-t)} \, dt + \frac{a_{n/2}}{c_{n/2}} \int_{n/2}^{n} e^{-\beta s(n-t)} \, dt$$

$$\leq \frac{a_1}{c_1} \left[ e^{-\beta s(n-1)} - \frac{1}{\beta s} \left( e^{-\beta s(n/2)} - e^{-\beta s(n-1)} \right) \right] + \frac{a_{n/2}}{c_{n/2}} \frac{1}{\beta s} \left[ 1 - e^{-\beta s(n/2)} \right]$$

The first term in brackets approaches zero at least as rapidly as $e^{-\beta s(n/2)}$, and the second term in brackets is bounded by 1. Thus for all $n$ greater than some $N_o$, $N_o < \infty$, we have

$$\left| \overline{e_n^q x_n^m} - \overline{e_n^q} \, \overline{x_n^m} \right| \leq mAB^{m-1} \left| e_n^q \right|_{max} K \frac{1}{\beta s} \frac{a_{n/2}}{c_{n/2}}$$

for all $x_1 \in X$, and hence

$$F_n(x_1) \leq [K_1]^{1/4} \left[ \frac{1}{\beta s} \right]^{1/4} \left[ \frac{a_{n/2}}{c_{n/2}} \right]^{1/4} \tag{18}$$

where $K_1$ is dependent only on A, B, $|e_n|_{max}$, and the quantization of $x_n$ and $e_n$. Statement 1 thus remains valid if assumption iii(a) is replaced by hypothesis 12 and we require

$$\sum_{j=1}^{\infty} \frac{a_j}{c_j} \left[ \frac{a_{j/2}}{c_{j/2}} \right]^{1/4} < \infty \tag{19}$$

(where $a_{j/2}/c_{j/2}$ is suitably interpolated for n odd). Now, if we set $a_n = An^{-\alpha}$ and $c_n = Cn^{-\gamma}$, then

$$\frac{a_{n/2}}{c_{n/2}} = 2^{\alpha-\gamma} \frac{a_n}{c_n} = 2^{\alpha-\gamma} n^{-(\alpha-\gamma)}$$

and statements 2 and 3 may be recast to read:

STATEMENT 2': Restrictions (a-e), assumption 12, and the choice $\alpha = 1$, $\gamma = 1/13$ imply

$$E\left\{\|\underline{x}_n - \underline{\theta}\|^2 \,|\, \underline{x}_1\right\} = O(n^{-2/13}) \qquad \text{for all } \underline{x}_1 \in X$$

STATEMENT 3': With the additional restriction that W(e) have a continuous third derivative, the choice $\alpha = 1$, $\gamma = 1/21\alpha$ implies

$$E\left\{\|\underline{x}_n - \underline{\theta}\|^2 \,|\, \underline{x}_1\right\} = O(n^{-4/21}) \qquad \text{for all } \underline{x}_1 \in X$$

The proofs of statements 2' and 3' follow exactly as do those of statements 2 and 3. We cannot, as in statements 2 and 3, state that the choices of $\alpha$ and $\gamma$ are optimum; rather they are the choices of $\alpha$ and $\gamma$ for which the estimates used guarantee the most rapid convergence.

REMARK 1: Note that restrictions (d) and (e) prohibit the use of the weighting function $W(e) = |e|$. We might remark that for practical purposes we could approximate $|e|$ arbitrarily closely by a function that satisfies restrictions (d) and (e). This answer is not entirely satisfactory, however. Restriction (d) is not troublesome. Indeed, any physical device that might be constructed (such as a rectifier) to obtain an approximation to $W(e) = |e|$ would almost certainly not behave as $|e|$ near the origin, but would possess continuous first and second derivatives. Therefore, when we set $W(e) = |e|$, we shall assume that restriction (d) is still satisfied.

Restriction (e), however, is more troublesome. Although it would be easy to construct a device that behaves as $|e|$ for large $|e|$, it would be difficult to build a device to approximate $|e|$ which is strictly convex.

We note that, other than assumption 12, we have not placed any restrictions on the signal d(m) except for uniform boundedness. We now add an additional restriction that permits the use of $W(e) = |e|$. Although the function $W(e) = |e|$ is not strictly convex, it is still convex. That is,

$$W[\alpha a+(1-\alpha)b] \leq \alpha W(a) + (1-\alpha)\, W(b) \qquad 0 \leq \alpha \leq 1 \tag{20}$$

Now if $S_{gn}(a) = -S_{gn}(b)$, then for any W(e) satisfying inequality 20 there exists an $E > 0$ with the property that for $\min [|a|, |b|] \geq \epsilon > 0$

$$W[\alpha a+(1-\alpha)b] \leq \alpha W(a) + (1-\alpha)\, W(b) - \alpha E|a-b| \qquad 0 \leq \alpha \leq 1/2 \tag{21}$$

Hence, if we replace assumption (e) by the milder condition (inequality 20), and assumption (c) by the stronger condition that there exist a $D > 0$ with the property that for all $\underline{x} \in X$

$$P\left\{S_{gn}\left(\sum_{i=1}^{k} x_i f_i(m) - d(m)\right) = -S_{gn}\left(\sum_{i=1}^{k} \theta_i f_i(m) - d(m)\right),\right.$$

$$\left.\min\left[\left|\sum_{i=1}^{k} x_i f_i(m) - d(m)\right|, \left|\sum_{i=1}^{k} \theta_i f_i(m) - d(m)\right|\right] > \epsilon\right\} > D\|\underline{x}-\underline{\theta}\| \qquad (22)$$

then we again obtain Eq. 7 and assumption (ii) is still satisfied.

The condition expressed by Eq. 22 is quite intractable; it would be extremely difficult in a practical situation to ascertain whether or not it is satisfied. Nevertheless, the condition is reasonable enough for carrying out the procedure with $W(e) = |e|$ with a fair amount of confidence that the procedure would converge.

REMARK 2: The discussion, thus far, has been in terms of discrete time-parameter sources. The adaptation of the method to continuous signals and systems is quite straightforward. We select some length of time T to be equivalent to one data sample. Then, assuming that the $n^{th}$ stage of the iterative procedure starts at time $t = \tau$, we make the 2k observations:

$$Y_n^1 = \frac{1}{T}\int_{\tau}^{\tau+T} W\left[d(t) - \sum_{i=1}^{k} x_i f_i(t)\right] dt$$

$$Y_n^2 = \frac{1}{T}\int_{\tau+T}^{\tau+2T} W\left[d(t) - \sum_{i=1}^{k} x_i f_i(t)\right] dt$$

$$\vdots$$

$$Y_n^{2k} = \frac{1}{T}\int_{\tau+(2k-1)T}^{\tau+2kT} W\left[d(t) - \sum_{i=1}^{k} x_i f_i(t)\right] dt$$

where

$$\underline{x} = \begin{cases} \underline{x}_n + c_n\underline{e}_1 & \tau < t < \tau + T \\ \underline{x}_n - c_n\underline{e}_1 & \tau + T < t < \tau + 2T \\ \quad\vdots & \\ \underline{x}_n - c_n\underline{e}_k & \tau + (2k-1)\,T < t < \tau + 2kT \end{cases}$$

and proceed exactly as in section 2. One iteration is thus performed in 2kT seconds.

In the continuous case,

$$M(\underline{x}) = \lim_{T\to\infty} \frac{1}{2T} \int_{-\infty}^{\infty} W[d(t)-q_{\underline{x}}(t)]\, dt$$

$$= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{m=-N}^{N} \int_{mT}^{(m+1)T} W[d(t)-q_{\underline{x}}(t)]\, dt$$

where $q_{\underline{x}}(t)$ is the output of the filter with the parameter set at $\underline{x}$. The derivations of the preceding discussion can thus be easily adapted to the continuous case.

[Note added in proof: See addenda in Sec. XIII-F for improved estimates of the rates of convergence given in statements 2' and 3'.]

D. J. Sakrison

## C. AVERAGE OF THE PRODUCT OF GAUSSIAN VARIABLES

The results given in this report are used extensively by Wiener (1) and by others. Since we have been unable to locate a detailed proof of these results, a proof is presented here for reference purposes.

The average of the product of N gaussian random variables is of basic importance in the statistical theory of nonlinear systems. If $\xi$ is a gaussian random variable, then

$$P_{\xi}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\overline{\xi})^2}{2\sigma^2}\right] \tag{1}$$

We may normalize $\xi$ by letting

$$\eta = \frac{\xi - \overline{\xi}}{\sigma} \tag{2}$$

Then

$$P_{\eta}(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) \tag{3}$$

We call $\eta$ a normalized random variable because $\overline{\eta} = 0$ and $\overline{\eta^2} = 1$. The result that we shall prove is that if $\eta_1, \eta_2, \ldots, \eta_{2N+1}$ (N=1, 2, ...) are normalized gaussian random variables, then

$$\overline{\eta_1\eta_2 \cdots \eta_{2N}} = \sum \prod \overline{\eta_i\eta_j} \tag{4}$$

and

$$\overline{\eta_1\eta_2 \cdots \eta_{2N+1}} = 0 \tag{5}$$

in which the notation $\sum \prod$ means the sum of all completely distinct ways of partitioning $\eta_1, \eta_2, \ldots, \eta_{2N}$ into pairs. The number of ways is $\frac{(2N)!}{N!\,2^N}$. For example, for N = 2,

$$\overline{\eta_1\eta_2\eta_3\eta_4} = \overline{\eta_1\eta_2}\,\overline{\eta_3\eta_4} + \overline{\eta_1\eta_3}\,\overline{\eta_2\eta_4} + \overline{\eta_1\eta_4}\,\overline{\eta_2\eta_3} \tag{6}$$

The number of terms in this expression is

$$\frac{(2N)!}{(N)!\,2^N} = \frac{4!}{2!\,2^2} = 3 \tag{7}$$

To prove this result, we begin by considering $P_\eta(y_1, y_2, \ldots, y_N)$ which is the joint probability density function of the N random variables, $\eta_1, \eta_2, \ldots, \eta_N$. The characteristic function of the joint probability density function is

$$M_\eta(a_1, a_2, \ldots, a_N) = \exp\left(\overline{j(a_1\eta_1+a_2\eta_2+\ldots+a_N\eta_N)}^{\,\eta}\right)$$

$$= \int_{-\infty}^{\infty} dy_1 \int_{-\infty}^{\infty} dy_2 \cdots \int_{-\infty}^{\infty} dy_N P_\eta(y_1, y_2, \ldots, y_N) \exp\left(j \sum_{i=1}^{N} a_i\eta_i\right) \tag{8}$$

The characteristic function can be expanded in a Taylor series:

$$M_\eta(a_1, a_2, \ldots, a_N) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} C_{k_1k_2\ldots k_N} a_1^{k_1} a_2^{k_2} \cdots a_N^{k_N} \tag{9}$$

in which

$$C_{k_1k_2\ldots k_N} = \frac{1}{k_1!\,k_2!\ldots k_N!} \frac{\partial^{k_1}}{\partial a_1^{k_1}} \frac{\partial^{k_2}}{\partial a_2^{k_2}} \cdots \frac{\partial^{k_N}}{\partial a_N^{k_N}} M_\eta(a_1, a_2, \ldots, a_N)\Big|_{a_1=a_2=\ldots=a_N=0} \tag{10}$$

However, from Eq. 8, we have

$$\frac{\partial^{k_1}}{\partial a_1^{k_1}} \frac{\partial^{k_2}}{\partial a_2^{k_2}} \cdots \frac{\partial^{k_N}}{\partial a_N^{k_N}} M_\eta(a_1, a_2, \ldots, a_N)\Big|_{a_1=a_2=\ldots=a_N=0} = \overline{\eta_1^{k_1}\eta_2^{k_2}\cdots\eta_N^{k_N}}\, j^{k_1} j^{k_2} \cdots j^{k_N} \tag{11}$$

so that

$$M_\eta(a_1, a_2, \ldots, a_N) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \overline{\eta_1^{k_1}\eta_2^{k_2}\cdots\eta_N^{k_N}}\, \frac{(ja_1)^{k_1}(ja_2)^{k_2}\cdots(ja_N)^{k_N}}{k_1!\,k_2!\ldots k_N!} \tag{12}$$

We note that the term for which $k_1 = k_2 = \ldots = k_N = 1$ is

$$\overline{\eta_1\eta_2\cdots\eta_N}(j\alpha_1)(j\alpha_2)\cdots(j\alpha_N) = \overline{\eta_1\eta_2\cdots\eta_N}(j)^N(\alpha_1\alpha_2\cdots\alpha_N) \tag{13}$$

This term contains the average that we want. Also, it is the only term in the expansion of Eq. 12 that contains this average. We shall now obtain another expansion of the characteristic function for the special case for which $\eta_1, \eta_2, \ldots, \eta_N$ are normalized gaussian random variables. We shall then obtain the desired result by equating the terms in the new expansion containing the product $(\alpha_1\alpha_2\cdots\alpha_N)$ with the term of Eq. 13.

If $\eta_1, \eta_2, \ldots, \eta_N$ are normalized gaussian random variables, their characteristic function can be shown (2) to be

$$M_\eta(\alpha_1, \alpha_2, \ldots, \alpha_N) = \exp\left(-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\overline{\eta_i\eta_j}\,\alpha_i\alpha_j\right) \tag{14}$$

Now, by the expansion

$$e^x = \sum_{p=0}^{\infty}\frac{x^p}{p!} \tag{15}$$

we can expand Eq. 14 as

$$M_\eta(\alpha_1, \alpha_2, \ldots, \alpha_N) = \sum_{p=0}^{\infty}\frac{1}{p!}\left(-\frac{1}{2}\right)^p\left[\sum_{i=1}^{N}\sum_{j=1}^{N}\overline{\eta_i\eta_j}\,\alpha_i\alpha_j\right]^p \tag{16}$$

The first few terms of Eq. 16 are

$$M_\eta(\alpha_1, \alpha_2, \ldots, \alpha_N) = 1 + \left(-\frac{1}{2}\right)\sum_{k_1=1}^{N}\sum_{k_2=1}^{N}\overline{\eta_{k_1}\eta_{k_2}}\,\alpha_{k_1}\alpha_{k_2}$$

$$+\frac{1}{2!}\left(-\frac{1}{2}\right)^2\sum_{k_1=1}^{N}\sum_{k_2=1}^{N}\sum_{k_3=1}^{N}\sum_{k_4=1}^{N}\overline{\eta_{k_1}\eta_{k_2}}\,\overline{\eta_{k_3}\eta_{k_4}}\,\alpha_{k_1}\alpha_{k_2}\alpha_{k_3}\alpha_{k_4} + \cdots \tag{17}$$

According to our previous discussion, we want only those terms that contain the product $\alpha_1\alpha_2\cdots\alpha_N$. We first note that the terms of the expansion, Eq. 16, contain only products of an even number of $\alpha$'s. Thus, if Eqs. 16 and 12 are to be equal, we require that the coefficient of the term of Eq. 13 shall be zero if N is odd. We thus have shown that

$$\overline{\eta_1\eta_2\cdots\eta_{2M+1}} = 0 \qquad M = 0, 1, 2, \ldots \tag{18}$$

This is Eq. 5. We now restrict our attention to the case for which N is even. Let N = 2M. We then note that the only terms of Eq. 16 that contain products of the

form $a_{k_1} a_{k_2} \cdots a_{k_{2M}}$ are those for which P = M. Those terms are

$$\frac{1}{M!}\left(-\frac{1}{2}\right)^M \sum_{k_1=1}^{2M} \sum_{k_2=1}^{2M} \cdots \sum_{k_{2M}=1}^{2M} \overline{\eta_{k_1}\eta_{k_2}} \cdots \overline{\eta_{k_{2M-1}}\eta_{k_{2M}}}\, a_{k_1} a_{k_2} \cdots a_{k_{2M}} \tag{19}$$

This sum contains many terms that we do not want. We want only those terms for which $k_1 \neq k_2 \neq \cdots \neq k_{2M}$, since this is the form of the term of Eq. 13. By eliminating all other terms from Eq. 19, we are left with the terms we desire, which may be written as

$$\frac{1}{M!}\left(-\frac{1}{2}\right)^M (a_1 a_2 \cdots a_{2M}) \sum \overline{\eta_{k_1}\eta_{k_2}} \cdots \overline{\eta_{k_{2M-1}}\eta_{k_{2M}}} \tag{20}$$

in which the sum is over all terms for which $k_1 \neq k_2 \neq \cdots \neq k_{2M}$. This sum contains many terms that have the same value. We can thus simplify Eq. 20 by summing together all such terms. To do this, we first note that $\overline{\eta_{k_i}\eta_{k_j}} = \overline{\eta_{k_j}\eta_{k_i}}$. There are $2^M$ terms of the sum in Eq. 20 that are identical in this manner, since each term is the product of M pairs. We also note that interchanging the order of the products of a term does not affect its value. This is so because

$$\overline{\eta_{k_1}\eta_{k_2}}\,\overline{\eta_{k_3}\eta_{k_4}} = \overline{\eta_{k_3}\eta_{k_4}}\,\overline{\eta_{k_1}\eta_{k_2}} \tag{21}$$

Since each term is the product of M pairs, there are M! permutations of this type. Thus there are M! terms of the sum of Eq. 20 that are identical in this manner. By summing all of these identical terms of Eq. 20, we can then write it in the form

$$(-1)^M (a_1 a_2 \cdots a_{2M}) \sum \prod \overline{\eta_{k_i}\eta_{k_j}} \tag{22}$$

in which the notation $\prod \overline{\eta_{k_i}\eta_{k_j}}$ means $\overline{\eta_{k_1}\eta_{k_2}}\,\overline{\eta_{k_3}\eta_{k_4}} \cdots \overline{\eta_{k_{2M-1}}\eta_{k_{2M}}}$, and the sum is over all completely distinct ways of forming the product. By equating Eqs. 22 and 13 we obtain

$$\overline{\eta_1\eta_2 \cdots \eta_{2M}} = \sum \prod \overline{\eta_i\eta_j} \tag{23}$$

This is Eq. 4, and our proof is complete.

By substituting Eq. 2 into Eq. 4, we obtain

$$\overline{\left(\frac{\xi_1 - \overline{\xi}_1}{\sigma_1}\right)\left(\frac{\xi_2 - \overline{\xi}_2}{\sigma_2}\right) \cdots \left(\frac{\xi_{2N} - \overline{\xi}_{2N}}{\sigma_{2N}}\right)} = \sum \prod \overline{\left(\frac{\xi_i - \overline{\xi}_i}{\sigma_i}\right)\left(\frac{\xi_j - \overline{\xi}_j}{\sigma_j}\right)} \tag{24}$$

By multiplying both sides of this equation by the product $\sigma_1\sigma_2 \cdots \sigma_{2N}$, we obtain

$$\overline{(\xi_1-\bar{\xi}_1)(\xi_2-\bar{\xi}_2)\cdots(\xi_{2N}-\bar{\xi}_{2N})} = \sum\prod \overline{(\xi_i-\bar{\xi}_i)(\xi_j-\bar{\xi}_j)} \tag{25}$$

Similarly, by substituting Eq. 2 into Eq. 5, we obtain

$$\overline{(\xi_1-\bar{\xi}_1)(\xi_2-\bar{\xi}_2)\cdots(\xi_{2N+1}-\bar{\xi}_{2N+1})} = 0 \tag{26}$$

To illustrate this last equation, consider the case for which N = 1. Then

$$\overline{(\xi_1-\bar{\xi}_1)(\xi_2-\bar{\xi}_2)(\xi_3-\bar{\xi}_3)} = 0 \tag{27}$$

By expanding the product, we thus obtain

$$\overline{\xi_1\xi_2\xi_3} = \bar{\xi}_1\overline{\xi_2\xi_3} + \bar{\xi}_2\overline{\xi_1\xi_3} + \bar{\xi}_3\overline{\xi_1\xi_2} - 2\bar{\xi}_1\bar{\xi}_2\bar{\xi}_3 \tag{28}$$

Note that we have not restricted the set of gaussian random variables, $\{\xi_i\}$, to be from the same ensemble. However, if they are from the same ensemble, we can write $\xi_i = x(t_i)$. Then, for the special case $\bar{\xi}_i = 0$, we have, from Eq. 25,

$$\overline{x(t_1)\,x(t_2)\cdots x(t_{2N})} = \sum\prod \overline{x(t_i)\,x(t_j)} \tag{29}$$

Thus, for a stationary ensemble, we have the result that

$$\begin{aligned}\phi_{xxxx}(\tau_1,\tau_2,\tau_3) &= \overline{x(t)\,x(t+\tau_1)\,x(t+\tau_2)\,x(t+\tau_3)} \\ &= \phi_{xx}(\tau_1)\,\phi_{xx}(\tau_3-\tau_2) + \phi_{xx}(\tau_2)\,\phi_{xx}(\tau_3-\tau_1) + \phi_{xx}(\tau_3)\,\phi_{xx}(\tau_2-\tau_1)\end{aligned} \tag{30}$$

in which $\phi_{xx}(\tau)$ is the autocorrelation function of x(t).

M. Schetzen

## References

1. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York, 1958).

2. H. Cramér, Methods of Statistics (Princeton University Press, 1946), Chap. 24.

## D. A METHOD FOR LOCATING SIGNAL SOURCES BY MEANS OF HIGHER-ORDER CORRELATION FUNCTIONS

The resolution of an antenna array is a basic problem of radio astronomy and of target location systems such as radar and sonar. The resolution of a receiving antenna may be taken as some fraction of its receiving beamwidth. Usually, by reciprocity, the receiving beamwidth is taken to be equal to that of the array when it is used as a

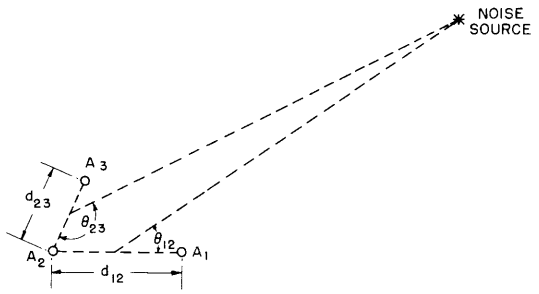



Fig. XIII-8. The geometry used for locating a noise source by an array of three antennas.

transmitting antenna (1). The physical limitations of the receiving antenna with regard to its beamwidth and bandwidth may then be determined (2). However, if the signals received by each element of the array are not processed linearly, then the assumption of reciprocity is no longer valid. The physical limitations of the array when used as a receiving antenna may then differ from those when it is used as a transmitting antenna. In this report, we shall present a method of locating noise sources in space by the use of higher-order correlation functions. We shall then obtain expressions for the ambiguity in locating a target when using this method. Some applications will then be presented.

A method for the location of a noise source in a plane by the use of second-order correlation functions has been discussed by Hayase (3). By this method, the noise source is located by crosscorrelating the signals received by three antennas as shown in Fig. XIII-8. If $f_1(t)$ is the signal received from the noise source by antenna no. 1, the signal received by antenna no. 2 is

$$f_2(t) = f_1(t-T_1) \tag{1}$$

and the signal received by antenna no. 3 is

$$f_3(t) = f_2(t-T_2) \tag{2}$$

in which

$$\begin{aligned} T_1 &= \frac{d_{12}}{c} \cos\theta_{12} \\ T_2 &= \frac{d_{23}}{c} \cos\theta_{23} \end{aligned} \tag{3}$$

where $d_{ij}$ is the distance between the $i^{th}$ and $j^{th}$ antennas and c is the velocity of the signal. The second-order crosscorrelation of the three received signals is then

$$\phi_{123}(\tau_1,\tau_2) = \overline{f_1(t)\, f_2(t+\tau_1)\, f_3(t+\tau_1+\tau_2)} = \overline{f_1(t)\, f_1(t-T_1+\tau_1)\, f_1(t-T_1-T_2+\tau_1+\tau_2)} \tag{4}$$

Since a second-order autocorrelation function has its maximum value at the origin, we note that $\phi_{123}(\tau_1, \tau_2)$ has its maximum value at $\tau_1 = T_1$ and $\tau_2 = T_2$ (ref. 4). Thus, by locating the peak of $\phi_{123}(\tau_1, \tau_2)$, the angles $\theta_{12}$ and $\theta_{23}$ can be determined from Eq. 3.

The exact location of the noise source is then given by the intersection of the direction lines as shown in Fig. XIII-8.

A limitation of this procedure is the difficulty of locating the peak of $\phi_{123}(\tau_1, \tau_2)$. The usual procedure for determining the second-order correlation function is to delay each of the time functions by means of delay lines, multiply the delayed time functions, and then average the product. In this manner, the correlation function is determined point-by-point in the $\tau_1$-$\tau_2$ plane. This is a time-consuming procedure and if the peak is to be accurately located, the points in the $\tau_1$-$\tau_2$ plane must be taken close together. However, to locate the position of the peak, we are really interested in the shape of the correlation function and not in its value at any one point in the $\tau_1$-$\tau_2$ plane. A method of determining the second-order correlation function which accomplishes this aim has been presented (5). By this method, the second-order correlation function is determined, with a minimum integral-square error, as the second-order impulse response of a network as shown in Fig. XIII-9. For this network, the impulse responses, $h_n(t)$, form a complete orthonormal set. That is,

$$\int_0^\infty h_i(t)\, h_j(t)\, dt = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{5}$$

The amplifier gains, $A_i$, are adjusted to be equal to certain averages of the signals received by the three antennas. Then, for a given delay, $\delta$, between the two impulses, the response of the network is the second-order correlation function along a line in the $\tau_1$-$\tau_2$ plane. If the averages are made over a finite time, the determination of the amplifier gains will be in error, which will cause an error in locating the peak of the correlation function. To determine how this latter error is related to the integration time, we consider an ensemble of measurements. In each measurement, the amplifier gains are determined by averaging for a time, T. Each amplifier gain can then be considered as a random variable. If we now write the experimentally determined gains, $A_i$, as

$$A_i = \overline{A_i} + B_i \tag{6}$$

in which $\overline{A_i}$ is the expectation of $A_i$, then the circuit of Fig. XIII-9 can be considered as two networks in parallel: one with the gains $\overline{A_i}$, and the other with the gains $B_i$. This is schematically depicted in Fig. XIII-10. Since $\overline{A_i}$ is the desired gain, the second-order impulse response of the circuit with the gains $\overline{A_i}$ is the desired correlation function. In this manner, we can consider the total second-order impulse response as being the desired response corrupted by noise; the noise being the response of the network with the random gains, $B_i$. Now, the error in locating the peak of the desired response along any line in the $\tau_1$-$\tau_2$ plane is proportional to the amplitude of the noise. Thus the mean-square error in locating the peak is proportional to the mean-square value of the noise.

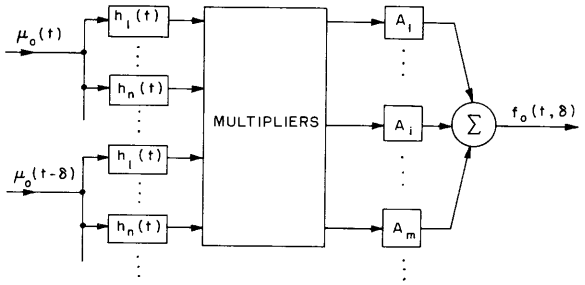



Fig. XIII-9. A network whose response, $f_o(t, \delta)$, is the second-order correlation function.

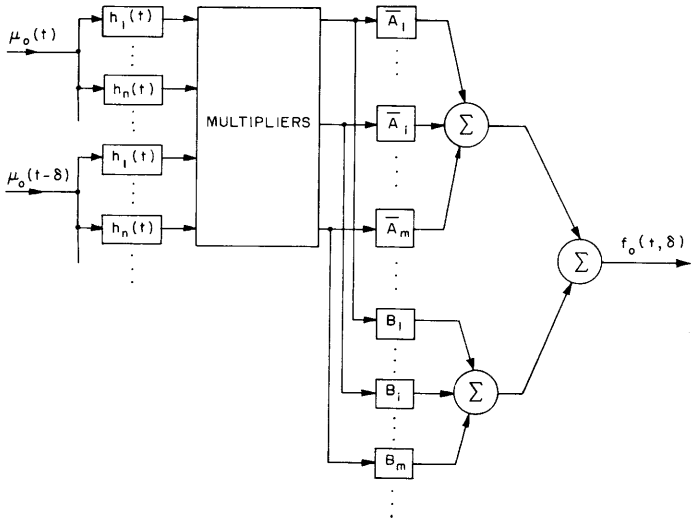



Fig. XIII-10. Pertaining to the calculation of the error in locating the peak of a second-order correlation function.

From Parseval's theorem, the expectation of the square of the noise integrated over the whole $\tau_1$-$\tau_2$ plane is $\sum_{i=1}^{\infty} \overline{B_i^2}$. However, it can be shown (6) that $\overline{B_i^2}$ is inversely proportional to the time of integration, T. Thus the experimental location of the peak of the correlation function in the $\tau_1$-$\tau_2$ plane may be said to lie within a circle of confusion whose radius, $R_2$, is inversely proportional to T (ref. 7). We shall define the ambiguity in locating a noise source as the area of this circle of confusion. The ambiguity in locating a noise source with an antenna array of three elements by the use of second-order correlation functions is thus inversely proportional to $T^2$.

By using more elements in the array, the ambiguity can be reduced. For example, consider the case in which there are four elements in the antenna array as shown in

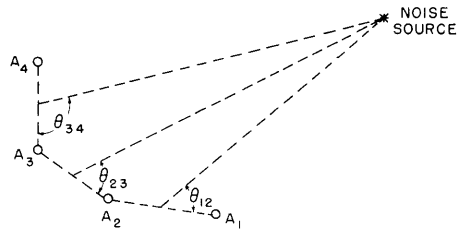



Fig. XIII-11. The geometry used for locating a noise source by an array of four antennas.

Fig. XIII-11. Then a third-order correlation function can be determined from the four received signals as the third-order impulse response of a network (5). As in our previous example, the angles $\theta_{12}$, $\theta_{23}$, and $\theta_{34}$ can be determined by locating the peak of the correlation function in the three-dimensional $\tau_1$-$\tau_2$-$\tau_3$ space. If the averages made to determine the amplifier gains of the network are over a finite time, T, then by the same method used in our previous example, the experimental location of the peak of the correlation function lies within a sphere of confusion whose radius, $R_3$, is inversely proportional to T. The ambiguity in locating a noise source with an array of four elements by the use of third-order correlation functions is thus inversely proportional to $T^3$. It is now clear that if the antenna array consists of N elements, then an (N-1) order correlation function can be determined from the N received signals as the (N-1) order impulse response of a network (5). For a finite time of observation, T, the location of the peak of the correlation function lies within an (N-1) dimensional sphere of confusion whose radius, $R_{N-1}$, is inversely proportional to T. The ambiguity in locating a noise source with an array of N elements is thus inversely proportional to $T^{N-1}$. For example, with a seven-element antenna array, the ambiguity in the location of a noise source can be reduced by a factor of two with only a ten per cent increase in the time of observation.

In order to attach physical interpretation to our definition of ambiguity, we must first briefly discuss the structure of the $\tau$-space. It should first be noted that the mapping of noise source positions to the $\tau$-space is not a one-one and onto mapping. An example of an array in which the mapping is not one-one is shown in Fig. XIII-12. In this example, the noise source above the array and the one symmetrically below it each has its peak at the same point in the $\tau$-space. Since such degeneracies arise from the symmetry properties of the array, the mapping can be made one-one by arranging the antennas of the array asymmetrically, but it will not be onto mapping. That is, every point in the $\tau$-space will not correspond to a noise source position. This is seen by noting from our previous discussion that each of the N coordinates of a point in the $\tau$-space is uniquely determined by one of the N direction lines from the antenna array. If the point corresponds

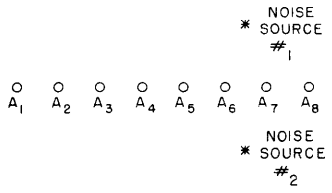



Fig. XIII-12. An example of an array for which the mapping is not one-one.

to a noise source position, then the corresponding direction lines intersect at a point which is the position of the noise source. If the angle of only one of the direction lines is changed, then the N direction lines no longer intersect at a point but at N points. In the N-dimensional $\tau$-space, this corresponds to moving parallel to one of the coordinate axes. From such considerations, it is seen that the locus of points in the $\tau$-space that correspond to noise source positions is a hypersurface. The exact shape of this hypersurface is a function of the relative positions of the array's antennas and can be determined by the simultaneous solution of the N equations for the N direction lines of the array. We now observe that an experimentally determined point in the $\tau$-space may not lie on this hypersurface. The center of the sphere of confusion in which it does lie, however, does correspond to the actual noise source position. Thus the hypersurface passes through the center of the sphere of confusion and the point of the hypersurface at the center of the sphere corresponds to the actual noise source position.

Before a measurement is made, we assume that all points on the hypersurface are equally likely. Consequently, after a measurement is made and an experimental point is obtained in the $\tau$-space off the hypersurface, the target position that one should choose is that point on the hypersurface which is closest to the experimental point. Thus the optimum choice of a target position is made by dropping a line from the experimental point perpendicular to the hypersurface. If the experimental point has the coordinates $(\tau_1, \tau_2, \ldots, \tau_N)$, and if we let $(\tau'_1, \tau'_2, \ldots, \tau'_N)$ be the coordinates of any point on the hypersurface, then by dropping a perpendicular to the hypersurface, we have chosen that point in space for which $\sum_{i=1}^{N} (\tau_i - \tau'_i)^2 = \sum_{i=1}^{N} (\Delta\tau_i)^2$ is a minimum.

We now wish to determine the probability, P, that the location in real space, to which this chosen point on the hypersurface corresponds, is within a given region about the true target position. We shall obtain an approximate expression for this probability. If the angle, $\alpha$, subtended by the region as seen from the antenna array is small, then from Eq. 3, the corresponding change in $\tau_1$, for example, is

$$\Delta\tau_1 \approx \frac{d_{12}}{c} \sin\theta_{12} \sin\alpha \approx \left[\frac{d_{12}}{c} \sin\theta_{12}\right] \alpha \tag{7}$$

Thus the change along any coordinate, $\tau_i$, can be approximated by a linear function of $\alpha$. The implication in the N-dimensional $\tau$-space is that the corresponding region of the hypersurface can be approximated by a hyperplane. For simplicity, let this region of interest on the hypersurface be a circle of radius $\epsilon_N$. The desired probability, P, is then the ratio of the partial volume of the sphere of confusion above and below the

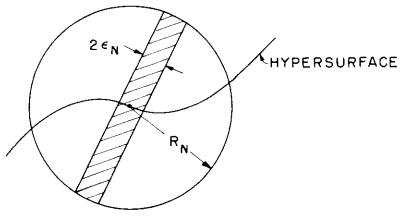



Fig. XIII-13. Pertaining to the calculation of the probability, P.

hyperplane of the circle to the total volume of the sphere. This partial volume is depicted by the shaded region of Fig. XIII-13. With the approximation that $\epsilon_N$ is small as compared with $R_N$, the ratio is given by

$$P = \frac{2\Omega(\beta)}{K_N R_N^{N-1}}$$

$$K_N = \frac{\pi^{N/2}}{\Gamma\left[\frac{N}{2}+1\right]} \tag{8}$$

in which $\Omega(\beta)$ is the area of the spherical cap of the partial volume. The equation for its area is (ref. 8)

$$\Omega(\beta) = \frac{(N-1)\,\pi^{(N-1)/2} R_N^{N-1}}{\Gamma[(N+1)/2]} \int_0^{\beta} \sin^{N-2} x\, dx \tag{9}$$

in which

$$\beta = \sin^{-1}\frac{\epsilon_N}{R_N} \approx \frac{\epsilon_N}{R_N} \tag{10}$$

Thus

$$\Omega(\beta) \approx \frac{\pi^{(N-1)/2} R_N^{N-1}}{\Gamma[(N+1)/2]} \left(\frac{\epsilon_N}{R_N}\right)^{N-1} \tag{11}$$

Substituting this last equation in Eq. 8, we find that the desired probability, P, is

$$P \approx \frac{2}{\sqrt{\pi}} \frac{\Gamma[(N/2)+1]}{\Gamma[(N+1)/2]} \left(\frac{\epsilon_N}{R_N}\right)^{N-1} \tag{12}$$

We have shown that $R_N$ is inversely proportional to T, the time of observation. Thus we observe from Eq. 12 that for a given array of N + 1 elements, the probability that the noise source is located within a given region about the true noise source position is proportional to $T^{N-1}$. Since $R_N$ is a function of the crosscorrelation function of the N + 1 received signals, both $R_N$ and $\epsilon_N$ are not only functions of the number of elements

in the array, but also of their relative positions in the array. These functions must be determined if we want to know the change in the probability, P, caused by a change in the array. From Eq. 12, we thus note that the optimum array is that one for which $(\epsilon_N/R_N)$ is a maximum.

The method of noise source location that we have just described is directly applicable to the design of receiving antenna arrays for use in radio astronomy. For target location systems such as radar and sonar, the target is not always an active source of noise. For such cases, the target may be made a passive source by illuminating it with some external noise source. For such cases, the $N^{th}$-order correlation function of the noise wave used for illumination can be tailored so that only a few terms of the orthonormal set, $h_n(t)$, of Fig. XIII-9 are required. In this manner, the additional error that results from truncation of the orthonormal set can be eliminated. A disadvantage of this method is that if several targets are present, they are no longer independent noise sources. As a result, false peaks will occur in the N-dimensional $\tau$-space. However, the location of these false peaks will be a function of the relative positions of the targets with respect to the illuminating noise source. To illustrate this, consider the simple case of two noise sources and an array of only two antennas. Let the signal received by the first antenna be

$$f_1(t) = N_1(t) + N_2(t) \tag{13}$$

in which $N_1(t)$ is the signal received from the first target and $N_2(t)$ is the signal received from the second target. The signal received by the second antenna will then be

$$f_2(t) = N_1(t-T_1) + N_2(t-T_2) \tag{14}$$

and the crosscorrelation of the two received signals is

$$\overline{f_1(t)\, f_2(t+\tau)} = \overline{[N_1(t)+N_2(t)][N_1(t-T_1+\tau)+N_2(t-T_2+\tau)]}$$

$$= \overline{N_1(t)\, N_1(t-T_1+\tau)} + \overline{N_2(t)\, N_2(t-T_2+\tau)} + \overline{N_1(t)\, N_2(t-T_2+\tau)} + \overline{N_2(t)\, N_1(t-T_1+\tau)} \tag{15}$$

The first term is the autocorrelation of the signal received from the first noise source and has a peak at $\tau = T_1$. Similarly, the second term has a peak at $\tau = T_2$. These are the two desired peaks. If the sources were independent, the third and fourth terms would be constants and the crosscorrelation of the received signals would contain only the two desired peaks. However, if the two targets are passive noise sources, then

$$N_2(t) = N_1(t-T_3) \tag{16}$$

in which $T_3$ is determined by the relative positions of the two targets with respect to

the illuminating noise source. For this case, the third and fourth terms become

$$\left.\begin{aligned} \overline{N_1(t)\, N_2(t-T_2+\tau)} &= \overline{N_1(t)\, N_1(t-T_3-T_2+\tau)} \\ \text{and} \\ \overline{N_2(t)\, N_1(t-T_1+\tau)} &= \overline{N_2(t)\, N_2(t+T_3-T_1+\tau)} \end{aligned}\right\} \tag{17}$$

Thus two false peaks at $\tau = T_2 + T_3$ and $\tau = T_1 - T_3$ arise from the crosscorrelation between the two targets. We should now observe that the maximum value of $T_1$ or $T_2$ that can occur is d/c, in which d is the distance between the two antennas and c is the velocity of the signal. Thus, for example, if $T_3$ is sufficiently large so that $|\tau_2+\tau_3|$ and $|\tau_1-\tau_3|$ are each greater than d/c, we know that they are false peaks, and there is no ambiguity. If this is not the case, then false peaks occur within the acceptable range of $\tau$. There are two possible methods of eliminating this ambiguity. First, since only the false peaks are a function of the position of the illuminating noise source, we can make a measurement for each of two different positions of the illuminator; the false peaks can then be determined by comparing the two measurements. The second method is to increase the number of antennas in the array. As this is done, not only can we increase the distance between targets in the N-dimensional $\tau$-space, but we also are imposing more constraints on the false peaks so that they can lie in the hypersurface corresponding to possible target positions. Thus, it should be possible to form an array by arranging a sufficient number of elements for which the false peaks that arise from dependent targets are separated from the hypersurface by distances greater than the radius of a sphere of confusion. This second method has an additional advantage. There is a second source of false peaks. They arise if the autocorrelation function of a noise source is not a monotonically decreasing function, but contains periodic components. By use of the second method it also should be possible to cause the location of such additional false peaks to be off the hypersurface.

It is interesting to note that this method of noise source location can be reversed to yield a method for navigation. Suppose that we want to locate the position of a receiver relative to several transmitting stations whose locations are known. If the signals transmitted by the several stations are coherent, then the receiver's position can be determined by crosscorrelating the several signals in the manner we have described and locating the peak of the correlation function.

M. Schetzen

(References and footnotes on following page)

References and Footnotes

1. S. Ramo and J. R. Whinnery, Fields and Waves in Modern Radio (John Wiley and Sons, New York, Second Edition, 1955), p. 558.

2. L. J. Chu, Physical limitations of omni-directional antennas, J. Appl. Phys. 19, No. 12, 1163-1175 (1948).

3. J. Y. Hayase, Properties of Second-Order Correlation Functions, Technical Report 330, Research Laboratory of Electronics, M.I.T. (1957).

4. We assume for the present that the $N^{th}$-order autocorrelation function of a noise source is a monotonically decreasing function. An example of such a random wave is the response of an RC filter for a gaussian white-noise input.

5. M. Schetzen, Measurement of correlation functions, Quarterly Progress Report No. 57, Research Laboratory of Electronics, M.I.T., April 15, 1960, pp. 93-98.

6. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, New York, 1960), Chap. 11.

7. Actually, since the expectation of the square of the noise is a function of $\phi_{123}(\tau_1, \tau_2)$, the region of confusion is an ellipse whose area, eccentricity, and orientation are a function of position in the $\tau_1$-$\tau_2$ plane. The major and minor axes of the ellipse, however, are inversely proportional to T.

8. C. Shannon, Probability of error for optimal codes in a gaussian channel, Bell Monograph 3259. Also, Bell System Tech. J. 38, 611-656 (1959).

## E. OPTIMUM COMPENSATION FOR NONLINEAR CONTROL SYSTEMS. II.

In Quarterly Progress Report No. 59, page 120, an algorithm was demonstrated which led from a solution of the associated optimum filter problem to the solution of the control problem. Before developing a similar procedure for the feedback compensator kernels, an example will be given in order to demonstrate the technique. This example is reasonably simple so that the technique will not be obscured by too much detail.

EXAMPLE 1. Series Compensation. Consider a system whose fixed elements are described by

$$y + ay + by = Cx + dx^3 \tag{18}$$

The desired operation of the over-all system is represented by a linear kernel $K_1$. The system is of the form shown in Fig. XI-14, Quarterly Progress Report No. 59, page 118.

We can write the linear term by inspection and then use the algorithm to find the higher-order kernels:

$$(S^2+aS+b)\, K_1(S) = CC_1(S_1)[1-K_1(S)] \tag{19}$$

This yields

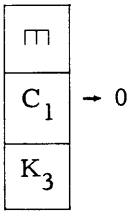
$$C_1(S) = \frac{1}{C} \frac{(S^2+aS+b)\, K_1(S)}{1 - K_1(S)} \tag{20}$$

The next nonzero equation is for $C_3$:

(a) $Q_3^1 = 0$, since $K_n = 0$ for $n \neq 1$

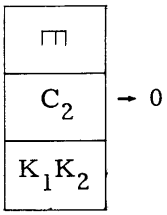
(b) $P_3^1 = P_3^1(1) + P_3^1(2) + P_3^1(3)$

For $P_3^1(1)$:



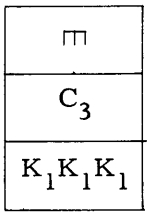

since $K_3 = 0$

For $P_3^1(2)$:




since $C_2 = 0$

For $P_3^1(3)$:




Therefore,

$$P_3^1 = P_3^1(3) = CC_3(S_1, S_2, S_3) \prod_{i=1}^{3} [1-K_1(S_i)] \tag{21}$$

(c) $P_3^3 = P_3^3(3) \rightarrow$

| | | |
|---|---|---|
| $C_1$ | $C_1$ | $C_1$ |
| $K_1$ | $K_1$ | $K_1$ |

$\rightarrow d \prod_{i=1}^{3} C_1(S_i)[1-K_1(S_i)]$

$$Q_3^1 = P_3^1 + P_3^3$$

$$0 = CC_3(S_1, S_2, S_3) \prod_{i=1}^{3} [1-K_1(S_i)] + d \prod_{i=1}^{3} C_1(S_i)[1-K_1(S_i)] \tag{22}$$

and we obtain

$$C_3(S_1, S_2, S_3) = -\frac{d}{C} C_1(S_1)\, C_1(S_2)\, C_1(S_3) \tag{23}$$

Similarly, for $C_5$

$$0 = +CC_5(S_1,S_2,S_3,S_4,S_5) \prod_{i=1}^{5} [1-K_1(S_i)] + 3dC_1(S_1)\, C_1(S_2)\, C_3(S_3,S_4,S_5) \prod_{i=1}^{5} [1-K_1(S_i)] \tag{24}$$

which yields

$$C_5(S_1, S_2, S_3, S_4, S_5) = -\frac{3d}{C} C_1(S_1)\, C_1(S_2)\, C_3(S_3, S_4, S_5) = +3\frac{d^2}{C^2} \prod_{i=1}^{5} C_1(S_i) \tag{25}$$

For $C_7$:

$$0 = +CC_7(S_1,S_2,S_3,\ldots S_7) \prod_{i=1}^{7} [1-K_1(S_i)] + 3dC_1(S_1)\, C_1(S_2)\, C_5(S_3,S_4,S_5,S_6,S_7) \tag{26}$$

and we obtain

$$C_7(S_1, S_2, S_3, \ldots S_7) = -\frac{9d^3}{C^3} \prod_{i=1}^{7} C_1(S_i) \tag{27}$$

In this case, we can see the form of the succeeding kernels. They consist of $C_1$ followed by a nonlinear no-memory element. The compensator could be synthesized as in Fig. XIII-14. But for $x^2 < \left|\frac{C}{3d}\right|$, the no-memory terms form a convergent power

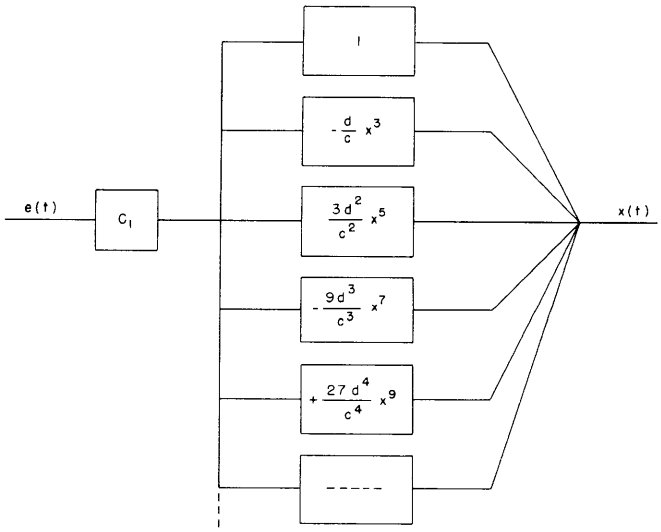



Fig. XIII-14. Series compensation – derived form.

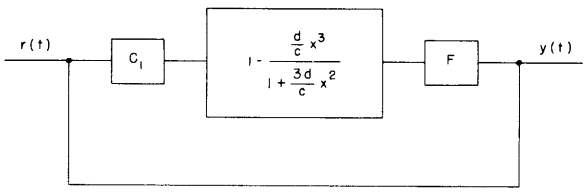



Fig. XIII-15. Series compensation – synthesis form.

series so that the system takes the form of Fig. XIII-15.

We see that the compensator kernels consist of the combination of a linear system and a nonlinear no-memory system. Consideration of other examples shows that for a large class of problems the optimum compensators are made up of combinations of three types of elements: linear filters, nonlinear no-memory devices, and the original filter kernels. The significance of this division is that it moves the filter-approximation problem outside the feedback loop.

(b) Algorithm for Determining Feedback Compensator Kernels

When $C_a = 1$, the system assumes the form shown in Fig. XI-14b in Quarterly Progress Report No. 59, p. 118. The problem is to determine a series of kernels for $C_b$ so that a desired filtering operation can be performed.

The equations describing the operation of the system are

$$C(t) = \int_0^t C_1(t-\tau)\, y(\tau)\, d\tau + \int_0^t \int_0^t C_2(t-\tau_1, t-\tau_2)\, y(\tau_1)\, y(\tau_2)\, d\tau_1 d\tau_2 + \dots \tag{28}$$

$$P_1(x, x', x'', \dots x^{(r)}) = P_2(y, y', y'', \dots y^{(s)}) \tag{29}$$

$$y(t) \int_0^t K_1(t-\tau)\, r(\tau)\, d\tau + \int_0^t \int_0^t K_2(t-\tau_1, t-\tau_2)\, r(\tau_1)\, r(\tau_2)\, d\tau_1 d\tau_2 + \dots \tag{30}$$

$$x(t) = r(t) - C(t) \tag{31}$$

Using the same approach as in the series compensator case, we can write an equation of the form

$$Q_m^s + Q_m^{s-1} + \dots Q_m^1 = P_m^r + P_m^{r-1} + \dots P_m^1 \tag{32}$$

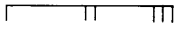
for each $m = 1, 2, \dots$ . Once again, the unknown kernel appears only in $P_m^1$. Therefore, each $C_m$ is determined successively.

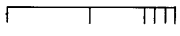
The method for determining $P_m^n$ and $Q_m^n$ will now be outlined. The techniques are the same as in the series case, but a complete discussion is included in order to maintain continuity.

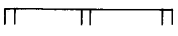
2.2 Construction of $P_m^n$; Contribution of an $n^{th}$-order nonlinearity of the input to an $m^{th}$-order feedback compensator kernel

As in the series case, $P_m^n$ consists of a sum of m-n+1 terms, $P_m^n(i)$. The structure of the component terms $P_m^n(i)$ is identical with the series compensation terms, but the actual functions are different. The construction of $P_m^n(i)$ involves three steps. First,

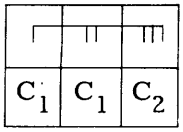
form all partitions of m objects into n cells. For example, the partitions for m = 6 and n = 3 are:

(I)  $p_{123}$

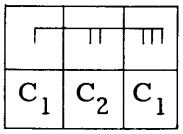
(II)  $p_{114}$

(III)  $p_{222}$

Second, consider all combinations of n compensator kernels $C_j$, where $\Sigma j = i$. Thus, for i = 4 and n = 3, the only set of kernels is $C_1$, $C_1$, $C_2$. For i = 5 and n = 3, the two possible sets are $C_1$, $C_2$, $C_2$ and $C_1$, $C_1$, $C_3$. Now, compare the various sets of kernels with the partitions from step 1 for compatibility. The number of objects in any cell of a partition represents the number of variables associated with that cell. The index on a compensator kernel $C_j$ represents the minimum number of variables necessary in its argument. Looking at partition I and the set of kernels $C_1$, $C_1$, $C_2$, we see that two compatible relations are:

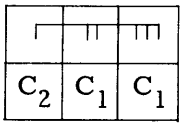
(Ia) 

| | | |
|---|---|---|
| $C_1$ | $C_1$ | $C_2$ |

and (Ib) 

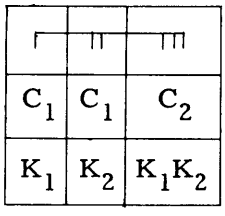
| | | |
|---|---|---|
| $C_1$ | $C_2$ | $C_1$ |

which could correspond to $C_1(S_1)\ C_1(S_2+S_3)\ C_2(S_4, S_5+S_6)$ for Ia and to $C_1(S_1)\ C_2(S_2, S_3)\ C_1(S_4+S_5+S_6)$ for Ib. An incorrect combination would be

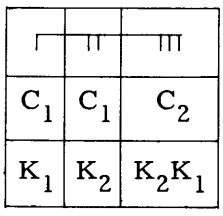


| | | |
|---|---|---|
| $C_2$ | $C_1$ | $C_1$ |

because there is only one variable in the first cell, and the argument of $C_2$ requires at least two variables. If we look at the original expression, we see that each compensator kernel $C_j$ has associated with it j filter kernels $K_\beta$. In the third step we look at the ways in which the filter kernels $K_\beta$ can combine with the compensator kernels $C_j$ with the restrictions that $\Sigma \beta = m$ and that the total number of kernels $K_\beta = i$. Looking at partition Ia, we see that the only possible arrangements are:

(Ia) 

| | | |
|---|---|---|
| $C_1$ | $C_1$ | $C_2$ |
| $K_1$ | $K_2$ | $K_1K_2$ |

or 

| | | |
|---|---|---|
| $C_1$ | $C_1$ | $C_2$ |
| $K_1$ | $K_2$ | $K_2K_1$ |

Since $C_2(S_1, S_2)$ may always be written in symmetrical form, the arrangements are

identical except for ordering. Thus, except for a numerical and characteristic coefficient, we can write the term that the partition represents.

$$[1-C_1(S_1)K_1(S_1)][C_1(S_2+S_3)K_2(S_2, S_3)][C_2(S_4, S_5+S_6)K_1(S_4)K_2(S_5, S_6)]$$

To complete $P_6^3(4)$, we repeat the same process for the remainder of $p_{123}$ and for the other major partitions $p_{114}$ and $p_{222}$. The procedure and results can be summarized in tabular form.

$P_6^3(4)$:

$p_{123}$ →

| | | | |
|---|---|---|---|
| $C_1$ | $C_1$ | $C_2$ | |
| $K_1$ | $K_2$ | $K_1K_2$ | $\rightarrow [1-C_1(S_1)K_1(S_1)][C_1(S_2+S_3)K_2(S_2, S_3)] \times [C_2(S_4, S_5+S_6)K_1(S_4)K_2(S_5, S_6)]$ |
| $C_1$ | $C_2$ | $C_1$ | |
| $K_1$ | $K_1K_1$ | $K_3$ | $\rightarrow [1-C_1(S_1)K_1(S_1)][C_2(S_2, S_3)K_1(S_2)K_1(S_3)] \times [C_1(S_4+S_5+S_6)K_3(S_4, S_5, S_6)]$ |

$p_{114}$ →

| | | | |
|---|---|---|---|
| $C_1$ | $C_1$ | $C_2$ | |
| $K_1$ | $K_1$ | $K_2K_2$ | $\rightarrow [1-C_1(S_1)K_1(S_1)][1-C_1(S_2)K_1(S_2)] \times [C_2(S_3+S_4, S_5+S_6)K_2(S_3, S_4)K_2(S_5, S_6)]$ |
| | | | |
| $K_1$ | $K_1$ | $K_1K_3$ | $\rightarrow [1-C_1(S_1)K_1(S_1)][1-C_1(S_2)K_1(S_2)] \times [C_2(S_3, S_4+S_5+S_6)K_1(S_3)K_3(S_4, S_5, S_6)]$ |

$p_{222}$ →

| | | | |
|---|---|---|---|
| $C_1$ | $C_1$ | $C_2$ | |
| $K_2$ | $K_2$ | $K_1K_1$ | $\rightarrow [C_1(S_1+S_2)K_2(S_1, S_2)][C_1(S_3+S_4)K_2(S_3, S_4)][C_2(S_5, S_6)K_1(S_5)K_1(S_6)]$ |

As in the series compensation case, three quantities remain to be specified: the sign of each p, the numerical coefficient of each p, and the characteristic coefficient associated with each p. The value of these quantities is exactly the same

as in the series case.

Construction of $Q_m^n$; Contribution of an $n^{th}$-order nonlinearity of the output to an $m^{th}$-order feedback compensator kernel

The basic equation for an $n^{th}$-order term is

$$[y(t)]^n = \left\{ \int_0^t K_1(t-\tau)\, r(\tau)\, d\tau + \int_0^t \int_0^t K_2(t-\tau_1, t-\tau_2)\, r(\tau_1)\, r(\tau_2)\, d\tau_1 d\tau_2 + \ldots \right\}^n \tag{33}$$

This is exactly the same basic equation as in the series compensator case. Therefore,

$$Q_m^n \text{ (series)} = Q_m^n \text{ (feedback)} \tag{34}$$

3. Summary

A general method for finding the optimum compensator for a control system has been shown. The range of input signal magnitudes for which the functional power series expansion is rapidly convergent will determine the practicality of the solution in a specific problem. Means of determining this radius of convergence will be shown in Quarterly Progress Report No. 61.

H. L. Van Trees, Jr.

## F. ADDENDA TO SECTION XIII-B

After the completion of Section XIII-B of this report it was found possible to simplify the derivations of that section so that it is necessary to apply only once the moment theorem used there. This results in improved estimates of the rates of convergence given in statements 2' and 3'. Equation 19 may be weakened to read:

$$\sum_{j=1}^{\infty} \frac{a_j}{c_j} \left[ \frac{a_{j/2}}{c_{j/2}} \right]^{1/2} < \infty$$

and statements 2' and 3' may be revised to read:

STATEMENT 2": Restrictions (a-c), assumption 12, and the choice $\alpha = 1$, $\gamma = 1/7$ imply

$$E\left\{ \left\| \underline{x}_n - \underline{\theta} \right\|^2 \middle| \underline{x}_1 \right\} = 0(n^{-2/7})$$

STATEMENT 3": With the additional restriction that W(e) shall possess a continuous third derivative, the choice $\alpha = 1$, $\gamma = 1/11$ implies

$$E\left\{\|\underline{x}_n - \underline{\theta}\|^2 \,|\, \underline{x}_1\right\} = 0(n^{-4/11})$$

Assumption 12 may also be weakened in the sense that the decay expressed in Eq. 12 need not be exponential but may fall off as slowly as $1/\tau_1^2$.

D. J. Sakrison