Prof. Y. W. Lee
Prof. A. G. Bose
Prof. A. H. Nuttall
D. A. Chesler
D. A. George
A. D. Hause
I. M. Jacobs
K. L. Jordan, Jr.
W. Neugebauer
J. W. Pan
M. Schetzen
D. W. Tufts
C. E. Wernlein, Jr.
G. D. Zames

A. SECOND-ORDER NONLINEAR FILTERS

1. White Noise Input

The problem considered in this report is that of optimizing the filter of Fig. X-1. We are given a white Gaussian-noise input x(t) of unity power per cycle per second, and a desired output z(t) which is some nonlinear function of the past of x(t); we must choose the parameters of the filter of Fig. X-1 so that its output u(t) is as close as possible, in the mean-square sense, to z(t). If E stands for expectation and $\mathcal{E}$ for the mean-square error, then we want to minimize

$$\mathcal{E} = E\{[u(t) - z(t)]^2\} \tag{1}$$

The filter of Fig. X-1 is composed of a dc voltage c_o, a linear filter with impulse response $L_1(t)$, and a set of N normalized linear filters with impulse response $\{h_n(t)\}$ each of which is followed by a no-memory squaring device and a pure gain $\{a_n\}$. The normalization of $\{h_n(t)\}$ is defined by

$$\int_0^\infty h_n^2(t)\, dt = 1 \tag{2}$$

The quadratic part, $y_2(t)$, of the filter output is a little more general than it might

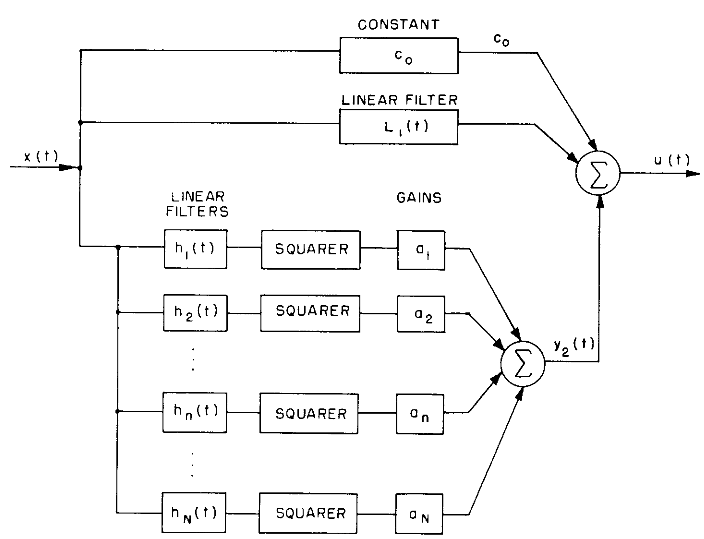

Fig. X-1. Second-order nonlinear filter.

at first appear. That is, given any output $y_2'(t)$ defined by

$$y_2'(t) \equiv \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} \int_0^\infty k_i(\tau_1)\, x(t-\tau_1)\, d\tau_1 \int_0^\infty k_j(\tau_2)\, x(t-\tau_2)\, d\tau_2 \tag{3}$$

then there exists a normalized set of N linear filters with impulse response $\{h_n(t)\}$ and a set of N gains $\{a_n\}$ with the property that

$$y_2'(t) = y_2(t) \equiv \sum_{n=1}^{N} a_n \left[\int_0^\infty h_n(\tau)\, x(t-\tau)\, d\tau\right]^2 \tag{4}$$

The $\{h_n(t)\}$ are linear combinations of $\{k_n(t)\}$.

We shall now make use of Wiener's hierarchic expansion (ref. 1) to separate the error $\mathscr{E}$, defined in Eq. 1, into a sum of parts from each member of the hierarchy. The desired output z(t) can be written as an infinite sum of members of Wiener's hierarchy $\{G_n(t)\}$.

$$z(t) = \sum_{n=0}^{\infty} G_n(t) \tag{5}$$

The filter output contains no terms higher than second order in the input. Therefore the filter output can be written as

$$u(t) = \sum_{n=0}^{2} Y_n(t) \tag{6}$$

in which the members of the set $\{Y_n(t)\}$ are members of Wiener's hierarchy and will be defined explicitly later.

The members of the sets $\{G_n(t)\}$ and $\{Y_n(t)\}$ have the following useful properties:

$$E\{G_n(t)\, G_m(t)\} = 0 \qquad n \neq m \tag{7}$$

$$E\{Y_n(t)\, Y_m(t)\} = 0 \qquad n \neq m \tag{8}$$

$$E\{G_n(t)\, Y_m(t)\} = 0 \qquad n \neq m \tag{9}$$

Substituting Eqs. 5 and 6 in Eq. 1, we obtain

$$\mathscr{E} = E\left\{\left[\sum_{n=0}^{\infty} G_n(t) - \sum_{m=0}^{2} Y_m(t)\right]^2\right\} \tag{10}$$

If we expand Eq. 10 and make use of the linear independence given in Eqs. 7, 8, and 9, we obtain the desired separation of error.

$$\mathscr{E} = E\left\{[G_o(t) - Y_o(t)]^2\right\} + E\left\{[G_1(t) - Y_1(t)]^2\right\} + E\left\{[G_2(t) - Y_2(t)]^2\right\} + \sum_{n=3}^{\infty} E\left\{G_n^2(t)\right\} \tag{11}$$

Note that only the first three terms of Eq. 11 contain the set $\{Y_n(t)\}$.

We now express the first three terms of the set $\{G_n(t)\}$ in terms of a set of kernels $\{K_n\}$. We also express the set $\{Y_n(t)\}$ in terms of the parameters of the filter of Fig. X-1.

$$G_o(t) = g_o \tag{12}$$

$$G_1(t) = \int_0^{\infty} K_1(\tau)\, x(t-\tau)\, d\tau \tag{13}$$

$$G_2(t) = \int_0^{\infty}\int_0^{\infty} K_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 - \int_0^{\infty} K_2(\tau, \tau)\, d\tau \tag{14}$$

$$Y_o(t) = c_o + \sum_{n=1}^{N} a_n \tag{15}$$

$$Y_1(t) = \int_0^{\infty} L_1(t)\, x(t-\tau)\, d\tau \tag{16}$$

$$Y_2(t) = \int_0^{\infty}\int_0^{\infty} \sum_{n=1}^{N} a_n h_n(\tau_1) h_n(\tau_2)\, x(\tau_1)\, x(\tau_2)\, d\tau_1 d\tau_2 - \sum_{n=1}^{N} a_n \tag{17}$$

The kernel $K_2(\tau_1, \tau_2)$ is symmetric in τ_1 and τ_2. The first term of the error in Eq. 11 becomes

$$E\left\{[G_o(t) - Y_o(t)]^2\right\} = \left[g_o - \left(c_o + \sum_{n=1}^{N} a_n\right)\right]^2 \tag{18}$$

The error given by expression 18 can be made equal to zero by choosing

$$c_o = g_o - \sum_{n=1}^{N} a_n \tag{19}$$

By substituting Eqs. 13 and 16 in the second term of Eq. 11, and by performing the averaging, we obtain

$$E\left\{[G_1(t) - Y_1(t)]^2\right\} = \int_0^{\infty} [K_1(\tau) - L_1(\tau)]^2\, d\tau \tag{20}$$

The error given by expression 20 can be made equal to zero by choosing

$$L_1(\tau) = K_1(\tau) \tag{21}$$

Similarly, the third term of Eq. 11 becomes

$$E\left\{\left[G_2(t)-Y_2(t)\right]^2\right\} = 2\int_0^\infty\int_0^\infty \left[K_2(\tau_1,\tau_2)-\sum_{n=1}^{N} a_n h_n(\tau_1)h_n(\tau_2)\right]^2 d\tau_1 d\tau_2 \tag{22}$$

It follows directly from Hilbert-Schmidt theory (2) that the error given in expression 22 is minimized if the members of the set $\{a_n\}$ are chosen as the N eigenvalues $\{\lambda_n\}$ of largest magnitude of the symmetric kernel $K_2(\tau_1,\tau_2)$ and if the members of the set $\{h_n(t)\}$ are chosen as the corresponding normalized eigenfunctions $\{\phi_n(t)\}$. That is, if

$$\lambda_n\,\phi_n(\tau_2) = \int_0^\infty K_2(\tau_1,\tau_2)\,\phi_n(\tau_1)\,d\tau_1 \tag{23}$$

and if

$$\int_0^\infty \phi_n^2(\tau)\,d\tau = 1 \tag{24}$$

and if

$$|\lambda_n| \geqslant |\lambda_{n+1}| \qquad \text{for all } n \tag{25}$$

then we should choose

$$a_n = \lambda_n \qquad n = 1,\ldots N \tag{26}$$

$$h_n(t) = \phi_n(t) \qquad n = 1,\ldots N \tag{27}$$

2. Non-White Noise Input

We now consider the case in which the input to the filter of Fig. X-1 is Gaussian noise x(t) that is non-white. This problem can be immediately reduced to the previous case by passing the input noise through a whitening filter with a realizable inverse and using the output of the whitening filter as the input to a filter of the form of Fig. X-1. This cascade arrangement still has the form of the filter of Fig. X-1, since the whitening filter can be incorporated into each of the linear filters of this figure. The whitening process does not increase the minimum attainable error, since each of the linear filters can invert the effects of the whitening filter.

D. A. Chesler

References

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

2. F. Riesz and B. Sz-Nagy, Functional Analysis (F. Ungar Publishing Company, New York, 1955), p. 243.

B. CONTINUOUS FEEDBACK SYSTEMS

In a previous report (1) an algebraic representation for continuous nonlinear systems was introduced. The present report considers the representation of nonlinear feedback systems. In the algebra of continuous systems

$$f(t) = \int_0^t \dots \int_0^t h_n(t-\tau_1, \dots, t-\tau_n)\, x(\tau_1) \dots x(\tau_n)\, d\tau_1 \dots d\tau_n \qquad (1)$$

is represented by

$$f = \underline{H}_n[x] \qquad (2)$$

or

$$f = \underline{H}_n[x^n] \qquad (3)$$

where these forms are equivalent. All elements in the feedback systems considered will be assumed to be describable in terms of this algebra.

1. Additive Feedback

The equation

$$f = x + \underline{H}[f] \qquad (4)$$

describes the feedback system of Fig. X-2. If

$$f = \underline{F}[x] \qquad (5)$$

then

$$\underline{F} = \underline{I} + \underline{H} * \underline{F} \qquad (6)$$

where $\underline{I}$ is the identity system and $*$ denotes the cascade operation. Equation 6 is called a system equation, and contains $\underline{F}$ implicitly. The problem is to find an explicit representation for $\underline{F}$ in terms of the algebraic operations and $\underline{H}$.

A necessary first step is to investigate the uniqueness of output f. If f is not unique, then the physical system represented by Eq. 4 will exhibit some erratic behavior. A technique for investigating uniqueness has been developed. One result is:

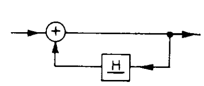

Fig. X-2. Additive feedback system.

"If $\underline{H}$ is a bounded system then f is unique." A bounded system, $\underline{H}$, is one for which all kernels $h_n(t_1, t_2, \dots, t_n)$ are bounded functions.

Other writers (2, 3) have obtained a

representation for $\underline{F}$ by assuming that

$$\underline{F} = \underline{F}_1 + \underline{F}_2 + \ldots + \underline{F}_n + \ldots \tag{7}$$

and substituting in Eq. 4. Thus

$$\underline{F}_1 + \underline{F}_2 + \ldots = \underline{I} + (\underline{H}_1 + \underline{H}_2 + \ldots) * (\underline{F}_1 + \underline{F}_2 + \ldots) \tag{8}$$

Equating equal orders across the equality sign gives

$$\underline{F}_1 = \underline{I} + \underline{H}_1 * \underline{F}_1 \tag{9}$$

$$\underline{F}_2 = \underline{H}_2 \circ (\underline{F}_1 \cdot \underline{F}_1) + \underline{H}_1 \circ \underline{F}_2 \tag{10}$$

and so on. The meaning of the operation "$\circ$" can be obtained from the following example.

$$\{\underline{H}_3 \circ (\underline{A}_1 \cdot \underline{A}_2 \cdot \underline{A}_5)\}[x^8] = \underline{H}_3\left[\underline{A}_1[x] \cdot \underline{A}_2[x^2] \cdot \underline{A}_5[x^5]\right] \tag{11}$$

in the notation of the previous report (1). After rearrangement we have

$$\underline{F}_1 = (\underline{I} - \underline{H}_1)^{-1} \tag{12}$$

$$\underline{F}_2 = (\underline{I} - \underline{H}_1)^{-1} * \underline{H}_2 * (\underline{I} - \underline{H}_1)^{-1} \tag{13}$$

and so on. Thus, an infinite series representation for $\underline{F}$ has been developed. However, the reliable use of any series depends upon its convergence.

Brilliant (2) developed a method for testing convergence, but this method is conservative and tends to show only analyticity, that is, convergence for sufficiently small inputs. Furthermore, the series (Eq. 7) developed above can be shown to diverge in cases in which the system is well behaved.

The nature of this series is more apparent when it is noted that it is a generalized Taylor series. Consider

$$f(t) = ax(t) + \underline{H}[f(t)] \tag{14}$$

where a is some real number. Since f(t) depends on $\underline{a}$, it can be written as f(t, a). Then

$$f(t, a) = ax(t) + \underline{H}[f(t, a)] \tag{15}$$

or

$$f(a) = ax + \underline{H}[f(a)] \tag{16}$$

because f(a, t) is considered as a function of $\underline{a}$ only for a fixed time t. Then, under the usual constraints,

$$f(a) = f(0) + f'(0)\, a + \frac{1}{2!} f''(0)\, a^2 + \dots + \frac{1}{n!} f^{(n)}(0)\, a^n + \dots \tag{17}$$

where

$$f^{(n)}(0) = \left. \frac{d^n f(a)}{da^n} \right|_{a=0} \tag{18}$$

If these derivations are obtained from the feedback equation (Eq. 16) and f(a) is expanded as in Eq. 17 for a = 1, the series

$$\underline{F}_1 + \underline{F}_2 + \dots + \underline{F}_n + \dots \tag{19}$$

that has been developed is the same as the series of Eq. 7. Therefore, the series of Eq. 7 is a generalized Taylor series for the system, expanded about zero input. Because of the high degree of continuity required for a Taylor series, it is not surprising that convergence difficulties should arise.

In an attempt to overcome some of this difficulty, an iteration technique has been developed to obtain an explicit representation for the feedback system $\underline{F}$. Now

$$\underline{F} = \underline{I} + \underline{H} * \underline{F} \tag{20}$$

and the following sequence of approximations for $\underline{F}$ is formed:

$$\underline{F}_{(1)} = \underline{I} \tag{21}$$

$$\underline{F}_{(2)} = \underline{I} + \underline{H}[\underline{F}_{(1)}] = \underline{I} + \underline{H}[\underline{I}] = \underline{I} + \underline{H} \tag{22}$$

$$\underline{F}_{(3)} = \underline{I} + \underline{H}[\underline{F}_{(2)}] = \underline{I} + \underline{H}[\underline{I}+\underline{H}] \tag{23}$$

and so on. In the limit

$$\lim_{n\to\infty} \underline{F}_{(n)} \tag{24}$$

satisfies the equation, and so

$$\underline{F} = \lim_{n\to\infty} \underline{F}_{(n)}$$

$$= \underline{I} + \underline{H}[\underline{I} + \underline{H}[\underline{I} + \underline{H}[\ \dots\ ]\dots] \tag{25}$$

The series may be truncated to form an approximation to $\underline{F}$.

Some techniques for investigating convergences have been developed. A particular case is of interest. If

$$|\underline{H}[x]| \leq |\underline{K}_1[x]| \tag{26}$$

where $\underline{K}_1$ is a suitably chosen linear system, and $\underline{H}$ is a bounded system, then the series can be shown to converge over any finite time interval for any bounded input. It is also possible to estimate truncation error. It would be expected that such a limit could be imposed on most physical systems because of the saturation phenomena.

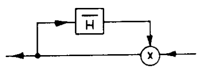

Fig. X-3. Multiplicative feedback system.

The convergence of this "iteration series" seems to be easier to study than that of the Taylor series. Furthermore, the convergence of the iteration series has certain connections with system stability, and this is being studied. The relationships between the two series are also being studied.

2. Multiplicative Feedback

A multiplicative feedback system is shown in Fig. X-3. The equations corresponding to this system are

$$f = x \cdot \underline{H}[f] \tag{27}$$

or

$$\underline{F} = \underline{I} \cdot (\underline{H} * \underline{F}) \tag{28}$$

where

$$f = \underline{F}[x] \tag{29}$$

Examples of such systems are found in AGC systems and in some FM detector circuits.

It should be noted that $\underline{H}$ must contain a zero-order term ($\underline{H}_0$), or the output will be identically zero. Hence Eq. 27 becomes

$$f = ax + x \cdot \underline{H}[f] \tag{30}$$

or

$$\underline{F} = a\underline{I} + \underline{I} \cdot (\underline{H} * \underline{F}) \tag{31}$$

As before, the uniqueness of f can be investigated and in the particular case of bounded $\underline{H}$, f is unique.

The Taylor series form can be developed as before. This gives

$$\underline{F} = a\underline{I} + 2\underline{I}(a\underline{H}_1 + a^2\underline{H}_2) + \dots \tag{32}$$

The iteration series gives

$$\underline{F}_{(1)} = a\underline{I} \tag{33}$$

$$\underline{F}_{(2)} = a\underline{I} + \underline{I} \cdot \underline{H}[a\underline{I}] \tag{34}$$

$$\underline{F}_{(3)} = a\underline{I} + \underline{I} \cdot \underline{H}[a\underline{I} + \underline{I} \cdot \underline{H}[a\underline{I}]] \tag{35}$$

and so on.

Similar studies of convergence can be made. If

$$|\underline{H}[x]| \leqslant |\underline{K}_1[x]| \tag{36}$$

and $\underline{H}$ is bounded, and $\underline{K}_1$ is some linear system, then the iteration series converges.

D. A. George

References

1. D. A. George, An algebra of continuous systems, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1958, p. 95.

2. M. Brilliant, Theory of the analysis of nonlinear systems, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., February 1958, and Technical Report 345, Research Laboratory of Electronics, M.I.T., March 3, 1958.

3. J. F. Barrett, The use of functionals in the analysis of non-linear physical systems, Statistical Advisory Unit Report 1/57, Ministry of Supply, Great Britain, 1957.

C. PATH AND TREE PROBABILITIES FOR A COMPLETE GRAPH

A probabilistic graph is an ensemble of line graphs and an associated probability law. The ensemble is generated by first constructing a base graph (set of nodes and connecting links) and then randomly erasing links in such a manner that in the resulting ensemble every link is present with a probability, p, and absent with a probability, q, independent of the presence or absence of all other links.

We are concerned with deriving the probabilities that a path or a tree is present in the ensemble generated by a complete base graph. (A graph is "complete" if there is one link between every pair of nodes. A "path" is present in a graph if there exists an unbroken chain of links connecting two prechosen terminal nodes. A "tree" is present if there is a path between every pair of nodes.) The choice of terminal nodes does not affect the path probability in a complete probabilistic graph because of the symmetry of the graph.

The complete graph was chosen for this analysis for two reasons. First, the complete graph was an integral part of the base graphs used to prove sufficiency in the limit theorems of the previous report (1). Second, the complete probabilistic graph is equivalent to an ensemble in which the base graph is chosen at random. In this case, the probability of a link being present in the ensemble can be thought of as the product of the probability that the link was chosen for use in the base graph and the probability that the link was not erased.

This analysis demonstrates the power of the factoring theorem and the intimate connection between the study of probabilistic graphs and combinatorial analysis. The following derivation could, in principle, be accomplished by the use of the inclusion-exclusion technique for calculating the probability of a union of nondisjoint events. That is, the probability of a tree could be found by listing all sets of links that form trees, adding the probabilities of these sets, subtracting the probability of all possible pairs of sets, adding the probabilities of all possible sets taken three at a time, and so on. The difficulty of this approach is evident when we note that the number of trees in an n-node complete graph is n^{n-2}, and the bookkeeping quickly becomes impossible. The advantage of the factoring technique is that it takes full advantage of the symmetries of the complete graph.

The factoring technique is based upon the factoring theorem which asserts that the probability of a path (or tree), P, can be written in the form

$$P = pP_1 + qP_2 \tag{1}$$

where P_1 is the probability of path (or tree) when a selected link is shorted in the base graph, and P_2 is the probability of path (or tree) when the same link is opened in the base graph. The factoring technique used in this derivation involves factoring not one link but many links.

Consider an (n+2)-node complete graph arranged as in Fig. X-4, with the terminal nodes for which the path probability is to be calculated placed in the extreme right and left positions, and the remainder of the nodes placed in a column in the center. There are n nodes in the center column connected by $\binom{n}{2} = n(n-1)/2$ links. The derivation of path and tree probabilities is accomplished by first factoring the graph on all $\binom{n}{2}$ central

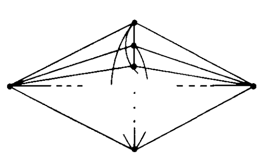

Fig. X-4. Desired node arrangement of the (n+2)-node complete graph.

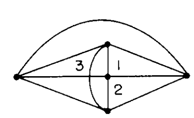

Fig. X-5. Five-node complete graph.

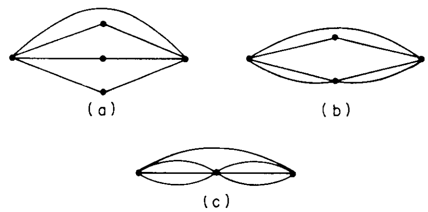

Fig. X-6. (a) Links $\bar{1}\,\bar{2}\,\bar{3}$ factored with probability q^3.
(b) Links $1\,\bar{2}\,\bar{3}$, $\bar{1}\,2\,\bar{3}$ or $\bar{1}\,\bar{2}\,3$ factored with probability $3pq^2$.
(c) Links $1\,2\,\bar{3}$, $1\,\bar{2}\,3$, $\bar{1}\,2\,3$ or $1\,2\,3$ factored with probability $3p^2q + p^3$.

links, and then calculating path and tree probabilities in the resulting structure. For example, factoring the $\binom{3}{2} = 3$ central links in the 5-node complete graph shown in Fig. X-5 results in the factored structures shown in Fig. X-6a, b, and c. The probability that each of these structures will occur is obtained by adding the probabilities of the various sets of shorted links and opened links that can produce them. These probabilities and sets are also given in Fig. X-6. (The notation, $a\bar{b}\bar{c}$, indicates that link a is shorted and links b and c are opened.)

The general appearance of a factored graph is shown in Fig. X-7. Note that this structure is completely defined by a partition of n; that is, by an n-component vector $\ell = (\ell_1, \dots, \ell_n)$, where $1\ell_1 + 2\ell_2 + \dots + n\ell_n = n$. For convenience, we shall use L_n to designate the set of all possible partitions, ℓ, of n. Also, we shall denote a set of i paths joined together at the center node as an i-path. There are three 1-paths, two 2-paths, and one 4-path depicted in Fig. X-7. The correspondence between partitions of n and factored structures is obtained by letting the i^{th} component of the partition vector, ℓ_i, represent the number of i-paths in the structure. Thus, the partitions (3, 0, 0), (1, 1, 0), and (0, 0, 1) correspond to the structures in Fig. X-6a, b, and c. It is important to note that when factoring an (n+2)-node complete graph, a structure is obtained for each and every partition of n; that is, for every member of the set L_n.

Since the partition completely specifies the factored structure, the probability of a path and the probability of a tree in the structure is a function only of ℓ. Call these

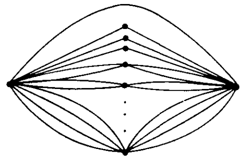

Fig. X-7. General form of complete graph after factoring.

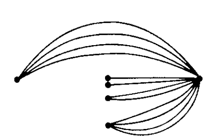

Fig. X-8. ℓ-structure with subset of center nodes connected to left node.

probabilities $A(\ell)$ and $B(\ell)$, respectively. Furthermore, denote by $M(\ell)$ the probability of obtaining the partition ℓ. Then, the probability of a path, P_P, and the probability of a tree, P_T, can be written in the form

$$P_P = \sum_{\ell \in L_n} A(\ell)\, M(\ell) \tag{2a}$$

$$P_T = \sum_{\ell \in L_n} B(\ell)\, M(\ell) \tag{2b}$$

Let us first calculate $A(\ell)$. The probability that no path exists between the terminal (right and left) nodes, $1 - A(\ell)$, is equal to the probability that the direct link and all of the i-paths fail. The probability that an i-path does not fail is equal to the probability that at least one link does not fail between the left node and the center node and between the center node and the right node, out of the i links that are available for each connection. Thus,

$$\begin{aligned} \text{Probability of i-path failing} &= 1 - (1 - q^i)(1 - q^i) \\ &= q^i(2 - q^i) \end{aligned} \tag{3}$$

Taking into account all i-paths and the direct link, the probability of a path existing in an ℓ-structure, $A(\ell)$, is

$$\begin{aligned} A(\ell) &= 1 - q \prod_{i=1}^{n} [q^i(2 - q^i)]^{\ell_i} \\ &= 1 - q^{n+1} \prod_{i=1}^{n} [2 - q^i]^{\ell_i} \end{aligned} \tag{4}$$

The calculation of the probability of a tree in an ℓ-structure, $B(\ell)$, is a bit more difficult. The attack proceeds as follows. Some of the center nodes of the various i-paths are assumed to have intact at least one link that connects them to the left node. The remaining center nodes are not directly connected to the left node. For each i-path center node connected to the left, there are now i possible links that connect left to right. This situation is shown in Fig. X-8. A tree is then present if at least one of the single links that connect left to right is intact and if the center nodes that are not connected to the left are connected to the right.

The probability of these events will now be calculated. The probability $P(m_i)$ that m_i out of the ℓ_i i-path center nodes have at least one connection intact to the left, and that the remaining $(\ell_i - m_i)$ i-path center nodes have no connections intact to the left, is

$$P(m_i) = \binom{\ell_i}{m_i} (1 - q^i)^{m_i} (q^i)^{\ell_i - m_i} \tag{5}$$

Each of the i-path center nodes connected to the left provide i links connecting left to right. Thus, including the direct link between terminals, there exist m links between left and right, where

$$m = 1 + \sum_{i=1}^{n} i\, m_i \tag{6}$$

The probability that at least one of these m links is intact is

$$1 - q^m \tag{7}$$

Finally, the probability that each of the remaining $(\ell_i - m_i)$ i-path center nodes have at least one connection intact with the right terminal is

$$(1 - q^i)^{\ell_i - m_i} \tag{8}$$

The probability of a tree, $B(\ell)$, can now be written. It is only necessary to sum the product of these probabilities over all possible combinations of the m_i. (Although it has not been explicitly stated, this derivation corresponds to an application of the factoring technique by which all links passing between the center nodes and the left node have been factored.) Thus,

$$B(\ell) = \sum_{m_1=0}^{\ell_1} \dots \sum_{m_n=0}^{\ell_n} (1 - q^m) \prod_{i=1}^{n} P(m_i)(1 - q^i)^{\ell_i - m_i} \tag{9}$$

Substituting from Eqs. 5 and 6 in Eq. 9, and performing the indicated sums, we obtain

$$B(\ell) = \prod_{i=1}^{n} (1 - q^{2i})^{\ell_i} - q^{n+1} \prod_{i=1}^{n} 2^{\ell_i}(1 - q^i)^{\ell_i} \tag{10}$$

Then it is only necessary to calculate the probability of obtaining an ℓ-structure. Because of the symmetry of the complete graph this calculation can be broken into two parts. First, we count the number of distinct ways of assigning n available center nodes to form the various i-paths of the partition ℓ. If we call this number $N(\ell)$, then

$$N(\ell) = \frac{n!}{\ell_1! \dots \ell_n! (1!)^{\ell_1} \dots (n!)^{\ell_n}} \tag{11}$$

Equation 11 follows from noticing that there are n! ways of ordering n numbers, but that every ordering does not result in a distinct partition. In particular, the ℓ_i sets of nodes forming i-paths can be ordered in $\ell_i!$ ways, and the i-nodes in any i-path can be ordered in i! ways.

Now, since every node in a complete graph looks like every other node, it is possible to calculate the probability of a partition for one particular node assignment and then multiply by the total number of node assignments possible for that partition, $N(\ell)$. Since $N(\ell)$ is the number of distinct partitions, this multiplication is valid because each assignment represents a disjoint event.

Now, what combinations of shorted links and open links can result in a specified partition, ℓ? Well, there must be a sufficient number of links to connect each set of i-nodes forming an i-path, and all links connecting this set of i-nodes with other center nodes must be open. Let us assume the following convention. All $\binom{n}{2}$ center nodes are initially considered open. This event occurs with probability $q^{\binom{n}{2}}$. Whenever a shorted link is used to connect nodes, a factor p/q is introduced into the probability expression. The numerator, p, accounts for the probability of the link being shorted, and the denominator, q, reduces the number of links that are open by subtracting one from the exponent of $q^{\binom{n}{2}}$. Thus, it is not necessary to count the links that must be open to separate i-paths from one another, since this is done automatically. That is, all nodes are assumed to be open unless specifically used to join the i nodes of an i-path together.

The number of ways of joining i labeled nodes together with j links is a familiar problem in combinatorial analysis (2, 3). Let us introduce the generating function of the number of connected i-node graphs, $G_i(x)$, where the enumeration is by the number of links. That is,

$$G_i(x) = \sum_{j=0}^{\infty} g_{ij} x^j \tag{12}$$

where g_{ij} is the number of connected graphs with i nodes and j links, or, in other words, the number of ways of joining i nodes with j links. The following generating functions are easily obtained:

$$\left.\begin{aligned} G_1(x) &= 1 \qquad \text{(by convention)} \\ G_2(x) &= x \\ G_3(x) &= 3x^2 + x^3 \\ G_4(x) &= 16x^3 + 15x^4 + 6x^5 + x^6 \end{aligned}\right\} \tag{13}$$

An explicit formula for $G_i(x)$ will be given after the discussion of Bell polynomials.

Now, since g_{ij} is the number of different ways of connecting i nodes with j links, and since each way represents a disjoint event, the probability of connecting a particular i-path with j links is $g_{ij}(p/q)^j$. The probability of connecting a particular i-path with an unrestricted number of links is $G_i(p/q)$. Since the probabilities are independent for different i-paths, the probability of obtaining a set of shorted and opened links that results in a specified partition is

$$q^{\binom{n}{2}}\left[G_1\left(\frac{p}{q}\right)\right]^{\ell_1}\left[G_2\left(\frac{p}{q}\right)\right]^{\ell_2}\cdots\left[G_n\left(\frac{p}{q}\right)\right]^{\ell_n} \tag{14}$$

Thus, by using Eqs. 11 and 14, $M(\ell)$ is found to be

$$M(\ell) = \frac{n!\,q^{\binom{n}{2}}}{\ell_1!\dots\ell_n!}\prod_{i=1}^{n}\left[\frac{G_i(p/q)}{i!}\right]^{\ell_i} \tag{15}$$

Combining Eqs. 2, 4, 10, and 15, we obtain for the probability of a path, P_P, and the probability of a tree, P_T, in the (n+2)-node complete graph

$$P_P = 1 - q^{\binom{n}{2}+n+1}\sum_{\ell\in L_n}\frac{n!}{\ell_1!\dots\ell_n!}\prod_{i=1}^{n}\left[\frac{(2-q^i)\,G_i(p/q)}{i!}\right]^{\ell_i} \tag{16a}$$

$$P_T = q^{\binom{n}{2}}\sum_{\ell\in L_n}\frac{n!}{\ell_1!\dots\ell_n!}\left\{\prod_{i=1}^{n}\left[\frac{(1-q^{2i})\,G_i(p/q)}{i!}\right]^{\ell_i} - q^{n+1}\prod_{i=1}^{n}\left[\frac{2(1-q^i)\,G_i(p/q)}{i!}\right]^{\ell_i}\right\} \tag{16b}$$

Equations 16a and 16b can be put into a more satisfying form by introducing the Bell polynomials (3, 4). Let f and y be n component vectors. The Bell polynomial, $A_n(f;y)$, is defined as

$$A_n(f;y) = \sum_{\ell\in L_n}\frac{n!\,f_s}{\ell_1!\dots\ell_n!}\left(\frac{y_1}{1!}\right)^{\ell_1}\cdots\left(\frac{y_n}{n!}\right)^{\ell_n} \tag{17}$$

where $s = \sum_{i=1}^{n}\ell_i$. For example,

$$
\begin{aligned}
A_1(f;y) &= f_1y_1 \\
A_2(f;y) &= f_1y_2 + f_2y_1^2 \\
A_3(f;y) &= f_1y_3 + f_2(3y_2y_1) + f_3y_1^3 \\
A_4(f;y) &= f_1y_4 + f_2\left(4y_3y_1 + 3y_2^2\right) + f_3\left(6y_2y_1^2\right) + f_4y_1^4
\end{aligned}
\tag{18}
$$

One of the basic applications of the Bell polynomials is in differentiating composite functions. It can be shown that

$$D_z^n f(y(z)) = A_n(f;y) \tag{19a}$$

where D_z^n denotes the n^{th} derivative with respect to z, and the components of the vectors f and y are

$$
\begin{aligned}
f_s &= D_y^s f(y)\Big|_{y=y(z)} \\
y_i &= D_z^i y(z)
\end{aligned}
\tag{19b}
$$

Making use of results presented by Riordan (3), we can express the generating function for complete graphs, $G_i(x)$, in terms of Bell polynomials, as follows:

$$G_i(x) = A_i\left(\{(-1)^{k-1}\ (k-1)!\}\ ;\ \left\{(1+x)^{\binom{k}{2}}\right\}\right) \tag{20}$$

Equations 16a and 16b can also be conveniently rewritten as

$$P_P = 1 - q^{\binom{n}{2}+n+1} A_n\left(\{1\}\ ;\ \left\{(2-q)^i\, G_i\left(\frac{p}{q}\right)\right\}\right) \tag{21}$$

and

$$P_T = q^{\binom{n}{2}} A_n\left(\{1\}\ ;\ \left\{(1-q^{2i})\, G_i\left(\frac{p}{q}\right)\right\}\right) - q^{\binom{n}{2}+n+1} A_n\left(\{1\}\ ;\ \left\{2(1-q^i)\, G_i\left(\frac{p}{q}\right)\right\}\right) \tag{22}$$

where $\{a_i\}$ represents an n-component vector whose i^{th} component is a_i.

Exact expressions for the probability of a path and of a tree in the (n+2)-node complete graph have now been derived. The expressions have a rather simple form in terms of Bell polynomials. Unfortunately, these expressions are not simple to work with analytically, and limiting behavior for large n is not known. However, knowledge of the exact expressions simplifies the search for realistic bounds.

I. M. Jacobs

References

1. I. M. Jacobs, Connectivity in random graphs – two limit theorems, Quarterly Progress Report No. 52, Research Laboratory of Electronics, M.I.T., Jan. 15, 1959, p. 64.

2. E. N. Gilbert, Enumeration of labelled graphs, Can. J. Math. 8, 405-411 (1956).

3. J. Riordan, Introduction to Combinatorial Analysis (John Wiley and Sons, Inc., New York, 1958).

4. E. T. Bell, Exponential polynomials, Ann. Math. 35, 258-277 (1934).

D. MINIMIZATION OF TRUNCATION ERROR IN SERIES EXPANSIONS OF RANDOM PROCESSES

In Quarterly Progress Report No. 52 it was shown (1) that a certain set of functions minimized the truncation error of the series expansion of a random process. A better proof has been found by using a theorem from functional analysis, and it will be given in the following discussion.

The theorem (2) reads as follows: If $K(t,s) = K_1(t,s) + K_2(t,s)$, where $K(t,s)$, $K_1(t,s)$, and $K_2(t,s)$ are positive definite kernels and $K_2(t,s)$ has only a finite number, N, of eigenvalues, then

$$\mu_{i+N} \leqslant \mu_{1i} \qquad i = 1, 2, \ldots \tag{1}$$

where μ_i and μ_{1i} are the eigenvalues of $K(t,s)$ and $K_1(t,s)$, respectively, and are so arranged that

$$\mu_1 \geqslant \mu_2 \geqslant \mu_3 \geqslant \ldots$$

$$\mu_{11} \geqslant \mu_{12} \geqslant \mu_{13} \geqslant \ldots$$

Now we consider the problem of the minimization of the truncation error. This error is defined as

$$\xi_N[\{\psi_n(t)\}] = E\left[\int_a^b \left[x(t) - \sum_{n=1}^{N} a_n \psi_n(t)\right]^2 dt\right] \tag{2}$$

where $x(t)$ is a random process with correlation function $R(t,s)$ and $\{\psi_n(t)\}$ is any set of N orthonormal functions. The a_n's are the Fourier coefficients. We shall show that the complete orthonormal set of eigenfunctions $\{\phi_n(t)\}$ of the integral equation

$$\int_a^b R(t,s)\, \phi_n(t)\, dt = \beta_n \phi_n(s) \qquad a \leqslant s \leqslant b \tag{3}$$

where

$$\beta_1 \geqslant \beta_2 \geqslant \beta_3 \geqslant \dots \tag{4}$$

minimizes this error, so that

$$\min_{\{\psi_n(t)\}} \xi_N[\{\psi_n(t)\}] = \xi_N[\{\phi_n(t)\}] = \sum_{n=N+1}^{\infty} \beta_n \tag{5}$$

It was shown in the previous report (3) that

$$\xi_N[\{\psi_n(t)\}] = \int_a^b R(t,t)\, dt - \sum_{n=1}^{N} \lambda_{nn} \tag{6}$$

where

$$\lambda_{nn} = E\left[a_n^2\right] \tag{7}$$

For the set $\{\phi_n(t)\}$ (see ref. 4) $\lambda_{nn} = \beta_n$, so that

$$\xi_N[\{\phi_n(t)\}] = \int_a^b R(t,t)\, dt - \sum_{n=1}^{N} \beta_n \tag{8}$$

Let us form the kernel

$$K_2(t,s) = \sum_{n=1}^{N} \lambda_{nn}\, \psi_n(t)\, \psi_n(s) \tag{9}$$

and consider R(t, s) as the sum of two kernels, as follows:

$$\begin{aligned} R(t,s) &= [R(t,s) - K_2(t,s)] + K_2(t,s) \\ &= K_1(t,s) + K_2(t,s) \end{aligned} \tag{10}$$

We know from Mercer's theorem (5) that any positive definite kernel K(t, s) can be expanded in a uniformly convergent series of eigenfunctions,

$$K(t,s) = \sum_{i=1}^{\infty} \mu_i\, \eta_i(t)\, \eta_i(s) \tag{11}$$

where μ_i and $\eta_i(t)$ are the eigenvalues and eigenfunctions, respectively, of K(t, s). Integrating term by term, we have

$$\int_a^b K(t,t)\, dt = \sum_{i=1}^{\infty} \mu_i \tag{12}$$

Applying this to Eq. 8, we have

$$\xi_N[\{\phi_n(t)\}] = \sum_{n=1}^{\infty} \beta_n - \sum_{n=1}^{N} \beta_n = \sum_{n=N+1}^{\infty} \beta_n \tag{13}$$

We do the same thing for $K_1(t, s)$ and obtain

$$\int_a^b K_1(t, t)\, dt = \int_a^b R(t, t)\, dt - \int_a^b K_2(t, t)\, dt \tag{14}$$

and we have, from Eqs. 9 and 6,

$$\int_a^b K_1(t, t)\, dt = \int_a^b R(t, t)\, dt - \sum_{n=1}^{N} \lambda_{nn} = \xi_N[\{\psi_n(t)\}] \tag{15}$$

After applying relation 12, it follows that

$$\int_a^b K_1(t, t)\, dt = \sum_{i=1}^{\infty} \chi_i \tag{16}$$

where the χ_i are the eigenvalues of $K_1(t, s)$. Now we apply Eq. 1 of our theorem and Eq. 13, to obtain

$$\xi_N[\{\psi_n(t)\}] = \int_a^b K_1(t, t)\, dt = \sum_{n=1}^{\infty} \chi_n \geq \sum_{n=1}^{\infty} \beta_{n+N} = \sum_{n=N+1}^{\infty} \beta_n = \xi_N[\{\phi_n(t)\}] \tag{17}$$

and hence

$$\xi_N[\{\psi_n(t)\}] \geq \xi_N[\{\phi_n(t)\}] = \sum_{n=N+1}^{\infty} \beta_n \tag{18}$$

which was to be proved.

K. L. Jordan, Jr.

References

1. K. L. Jordan, Jr., Minimization of truncation error in series expansions of random processes, Quarterly Progress Report No. 52, Research Laboratory of Electronics, M.I.T., Jan. 15, 1959, pp. 70-78.

2. F. Riesz and B. Sze.-Nagy, Functional Analysis, translated into English by L. F. Boron (Ungar Publishing Company, New York, 1955), p. 238.

3. K. L. Jordan, Jr., op. cit., p. 76, cf. the unnumbered equation preceding Eq. 12.

4. Ibid, p. 72, Eq. 3.

5. F. Riesz and B. Sze.-Nagy, op. cit., p. 245.

E. A SAMPLING THEOREM FOR STATIONARY RANDOM PROCESSES

Let [x(t)] be a stationary process with correlation function $\phi_x(\tau)$, and let [f(t)] be a process (in general, nonstationary) generated as follows.

If x(t) is a particular member of the ensemble [x(t)], then f(t), the corresponding member of the ensemble [f(t)], is given by

$$f(t) = \sum_{n=-N}^{N} x(nT_o)\, s(t - nT_o) \tag{1}$$

where the sequence $\{x(nT_o)\}$ represents samples of x(t) taken uniformly over a finite time interval $(-NT_o, NT_o)$, and s(t) is an interpolatory function to be determined.

We wish to discover what interpolatory function, s(t), will give us best agreement between f(t) and x(t) during the time interval $(-NT_o, NT_o)$ for all members of the ensemble [x(t)]. By this "best agreement" we mean that we wish to minimize

$$I = \frac{1}{2NT_o} \int_{-NT_o}^{NT_o} E \cdot \left[[x(t) - f(t)]^2 \right] dt \tag{2}$$

where E · [ ] stands for ensemble average of [ ].

Letting $g_N(t)$ be a function that is one in the time interval $(-NT_o, NT_o)$ and zero elsewhere, we find from our previous work (1) that the expression for I may be rewritten as follows:

$$I = \frac{1}{2NT_o} \sum_{m=-N}^{N} \int_{-\infty}^{\infty} g_N(t)[-2\phi_x(t - mT_o)\, s(t - mT_o)]\, dt$$

$$+ \sum_{k=-2N}^{0} \sum_{m=-N}^{k+N} \int_{-\infty}^{\infty} \phi_x(kT_o) \cdot s(t - mT_o) \cdot s(t - mT_o + kT_o)\, g_N(t)\, dt$$

$$+ \sum_{k=1}^{2N} \sum_{m=k-N}^{N} \int_{-\infty}^{\infty} \phi_x(kT_o)\, s(t - mT_o)\, s(t - mT_o + kT_o)\, g_N(t)\, dt + C \tag{3}$$

in which C is $E \cdot [(x(t))^2]$. (Note that Eq. 3 is corrected from Eq. 3 in ref. 1.)

We now make a change of variable, $u = t - mT_o$, and interchange the orders of summation and integration. Then we have

$$I = \frac{1}{2NT_o} \int_{-\infty}^{\infty} \left\{ \left[-2\phi_x(u) \cdot s(u)\right] \cdot \sum_{m=-N}^{N} g_N(u+mT_o) \right.$$

$$+ \left[\sum_{k=-2N}^{0} \phi_x(kT_o)\ s(u+kT_o) \right] \cdot s(u) \cdot \sum_{m=-N}^{k+N} g_N(u+mT_o)$$

$$\left. + \left[\sum_{k=1}^{2N} \phi_x(kT_o)\ s(u+kT_o) \right] \cdot s(u) \cdot \sum_{m=k-N}^{N} g_N(u+mT_o) \right\} du + C \qquad (4)$$

The expression

$$f(u) = \sum_{m=-N}^{N} g_N(u+mT_o)$$

can be described as follows.

$$f(u) \begin{cases} = 2N & \text{for } |u| \leq T_o \\ = (2N-1) & \text{for } T_o < |u| \leq 2T_o \\ \cdot & \\ \cdot & \\ \cdot & \\ = 1 & \text{for } (2N-1)\ T_o < |u| \leq 2NT_o \\ = 0 & \text{elsewhere} \end{cases}$$

Let

$$f_k(u) \begin{cases} = \sum_{m=-N}^{k+N} g_N(u+mT_o) & \text{for } -2N \leq k \leq 0 \\ = \sum_{m=k-N}^{N} g_N(u+mT_o) & \text{for } 1 \leq k \leq 2N \\ = 0 & \text{for all other } k \end{cases}$$

Note that $f_o(u) = f(u)$, and $f_k(u-kT_o) = f_{-k}(u)$.

We can rewrite Eq. 4 as

$$I = \frac{1}{T_o} \int_{-\infty}^{\infty} \left\{ \left[-2\phi_x(u)\ s(u)\right] \cdot \frac{1}{2N} f(u) \right.$$

$$\left. + s(u) \sum_{k=-2N}^{2N} \phi_x(kT_o) \cdot s(u+kT_o) \cdot \frac{1}{2N} \cdot f_k(u) \right\} du + C \qquad (5)$$

If we now substitute $s(u) + \epsilon\eta(u)$ for $s(u)$ in Eq. 5 and assume that the function $s(u)$ minimizes I, then a necessary condition for I to be minimum is that

$$\left.\frac{\partial I}{\partial \epsilon}\right|_{\epsilon=0} = 0$$

In our case,

$$\left.\frac{\partial I}{\partial \epsilon}\right|_{\epsilon=0} = \frac{1}{T_o}\int_{-\infty}^{\infty} \eta(u)\left[-2\phi_x(u) \cdot \frac{f(u)}{2N}\right] du$$

$$+ \frac{1}{T_o}\int_{-\infty}^{\infty} \eta(u)\left[\sum_{k=-2N}^{2N} \phi_x(kT_o)\ s(u+kT_o)\ \frac{f_k(u)}{2N}\right] du$$

$$+ \frac{1}{T_o}\int_{-\infty}^{\infty} s(u)\left[\sum_{k=-2N}^{2N} \phi_x(kT_o)\ \eta(u+kT_o)\ \frac{f_k(u)}{2N}\right] du$$

Making the change of variable, $w = u + kT_o$, in the last integral, and noting that $\phi_x(kT_o) = \phi_x(-kT_o)$, and $f_k(w-kT_o) = f_{-k}(w)$, we have

$$\left.\frac{\partial I}{\partial \epsilon}\right|_{\epsilon=0} = \frac{1}{T_o}\int_{-\infty}^{\infty} \eta(u) \cdot \left[-2\phi_x(u) \cdot \frac{f(u)}{2N} + 2\sum_{k=-2N}^{2N} \phi_x(kT_o) \cdot s(u+kT_o) \cdot \frac{f_k(u)}{2N}\right] du$$

Since $\eta(u)$ is arbitrary, our necessary condition for I to be minimum becomes

$$\phi_x(u) \cdot \frac{f(u)}{2N} = \sum_{k=-2N}^{2N} \phi_x(kT_o) \cdot s(u+kT_o) \cdot \frac{f_k(u)}{2N} \tag{6}$$

Let us now consider some properties of the solution of Eq. 6. Since $f(u)$, $f_k(u)$, and $f_{-k}(u)$ are all zero for $|u| > 2NT_o$, $s(u)$ is arbitrary for $|u| > 2NT_o$. For simplicity, we shall let $s(u)$ be zero for $|u| > 2NT_o$.

For the special case in which $\phi(kT_o) = 0$ for $k = \pm 1, \pm 2, \ldots, \pm 2N$, Eq. 6 becomes

$$\phi_x(u) \cdot \frac{f(u)}{2N} = \phi_x(0) \cdot s(u) \cdot \frac{f(u)}{2N} \qquad [\text{since } f(u) = f_o(u)]$$

The solution is

$$s(u) = \frac{\phi_x(u)}{\phi_x(0)} \qquad \text{for } |u| \leq 2NT_o$$

$$s(u) = 0 \qquad \text{for } |u| > 2NT_o$$

Another example is given by the following case:

$$\phi_x(\tau) = 2 - \frac{|\tau|}{T_o} \qquad \text{for } 0 \leq |\tau| \leq 2T_o$$

$$\phi_x(\tau) = 0 \qquad \text{elsewhere}$$

In this case, Eq. 6 may be written as follows for N = 1:

$$\left(2 - \frac{u}{T_o}\right) \cdot 1 = 1 \cdot s(u-T_o) \cdot 1 + 2 \cdot s(u) \cdot 1 + 1 \cdot s(u+T_o) \cdot \frac{1}{2}$$

with $0 \leq u \leq T_o$, and as

$$\left(2 - \frac{w}{T_o}\right) \cdot \frac{1}{2} = 1 \cdot s(w-T_o) \cdot \frac{1}{2} + 2 \cdot s(w) \cdot \frac{1}{2}$$

for $T_o < w \leq 2T_o$.

The form of Eq. 6 indicates that we should look for a symmetric solution, that is, we assume that $s(u-T_o) = s(T_o-u)$. Making the change of variable, $w = u + T_o$, gives us the following system of equations which is valid for $0 < u < T_o$:

$$2 - \frac{u}{T_o} = s(T_o-u) + 2s(u) + \frac{1}{2} s(u+T_o)$$

$$\frac{1}{2}\left(1 - \frac{u}{T_o}\right) = \frac{1}{2} s(u) + s(u+T_o)$$

A solution to these equations (one that satisfies the conditions at $u = 0$, $u = \pm T_o$, and $u = \pm 2T_o$) is

$$s(u) \begin{cases} = 1 - \frac{|u|}{T_o} & \text{for } 0 \leq |u| \leq T_o \\ = 0 & \text{elsewhere} \end{cases}$$

Let us now substitute Eq. 6 in Eq. 5 to obtain an expression for the minimum value of I (when 2N + 1 samples have been taken). We shall call this minimum value I_{2N+1}. It is given by the formula

$$I_{2N+1} = C - \frac{1}{T_o} \int_{-\infty}^{\infty} \phi(u) \cdot s(u) \cdot \frac{f(u)}{2N}\, du \qquad (7)$$

For the special case that has just been discussed, C is 2. From the definition of f(u) we recall that f(u)/2N is 1 in the interval $(-T_o, T_o)$, and s(u) and $\phi(u)$ have been defined. Thus Eq. 7 becomes

$$I_{2N+1} = 2 - \frac{1}{T_o}\int_{-T_o}^{T_o} \left(2 - \frac{|\tau|}{T_o}\right)\left(1 - \frac{|\tau|}{T_o}\right) d\tau$$

$$= 2 - \frac{2}{T_o}\int_{0}^{T_o} \left(2 - \frac{\tau}{T_o}\right)\left(1 - \frac{\tau}{T_o}\right) d\tau$$

$$= 2 - \frac{2}{T_o}\left[2\tau - \frac{3\tau^2}{2T_o} + \frac{\tau^3}{3T_o^2}\right]_0^{T_o}$$

$$= \frac{1}{3}$$

We might interpret this in terms of a ratio of average signal power to average "noise" (or error) signal power into a unit resistor load. For our special case,

$$\frac{S}{N} = \frac{2}{\frac{1}{3}} = \frac{6}{1}$$

We have been discussing the reconstruction, over a finite interval, of a member of a stationary ensemble from a finite number of uniformly spaced samples. We have constrained the problem by requiring that each sample be treated the same; that is, the same interpolatory function is used for each sample value.

We now turn to the case of reconstruction from an infinite number of samples. (As before, we do not restrict $[x(t)]$ to be bandlimited.) In this situation, f(t) is given by

$$f(t) = \sum_{n=-\infty}^{\infty} x(nT_o)\, s(t-nT_o)$$

As Bennett (2) has shown, $[f(t)]$ is a cyclostationary process, that is, its statistics vary periodically with period T_o. Thus it is reasonable to minimize I, as follows:

$$I = \frac{1}{T_o}\int_0^{T_o} E \cdot [x(t) - f(t)]^2\, dt$$

$$= \frac{1}{T_o}\int_0^{T_o} \left[C - 2\sum_{n=-\infty}^{\infty} \phi_x(t-nT_o)\, s(t-nT_o) + \sum_{m,\,n-\infty}^{\infty}\sum \phi_x(nT_o - mT_o)\, s(t-mT_o)\, s(t-nT_o)\right] dt$$

Letting $u = t - nT_o$, we have

$$I = C - \frac{1}{T_o} \sum_{n=-\infty}^{\infty} \int_{-nT_o}^{(-n+1)T_o} \left[2\phi_x(u)\, s(u) - s(u) \sum_{m=-\infty}^{\infty} \phi_x(nT_o - mT_o)\, s(u + nT_o - mT_o) \right] du$$

Making the change of summation index, k = n - m, and then carrying out the summation on n, yields

$$I = C - \frac{1}{T_o} \int_{-\infty}^{\infty} \left[2\phi_x(u)\, s(u) - s(u) \sum_{k=-\infty}^{\infty} \phi_x(kT_o)\, s(u + kT_o) \right] du \tag{8}$$

A necessary condition for I to be minimum (by varying s) is that

$$\phi_x(u) = \sum_{k=-\infty}^{\infty} \phi_x(kT_o)\, s(u + kT_o) \tag{9}$$

If $\Phi_1(f)$ is the Fourier transform of $\phi_x(u)$, then Eq. 9 may be written in the frequency domain as

$$S_1(f) = \frac{\Phi_1(f)}{\sum_{k=-\infty}^{\infty} \phi_x(kT_o)\, e^{+j2k\pi T_o f}} = \frac{\Phi_1(f)}{\sum_{k=-\infty}^{\infty} \phi_x(kT_o) \cos 2\pi k T_o f} \tag{10}$$

where $S_1(f)$ is the Fourier transform of s(t), provided that the series in the denominator has no zeros, where $\Phi_1(f)$ is nonzero.

The analysis given by Bennett (3) shows that we can rewrite Eq. 10 as

$$S_1(f) = \frac{\Phi_1(f)}{\frac{1}{T_o} \sum_{n=-\infty}^{\infty} \Phi_1(f - nf_r)} \tag{11}$$

where $f_r = 1/T_o$.

In terms of $w = 2\pi f$ (and thus $w_r = 2\pi/T_o$), we have

$$2\pi \cdot S(w) = \frac{2\pi \cdot \Phi(w)}{\frac{2\pi}{T_o} \sum_{n=-\infty}^{\infty} \Phi(w - nw_r)}\ ; \qquad \text{or } S(w) = \frac{T_o \cdot \Phi(w)}{2\pi \sum_{n=-\infty}^{\infty} \Phi(w - nw_r)} \tag{12}$$

with the definition $F(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t)\, e^{-jwt}\, dt$ used.

We can obtain an expression for the minimum value of I, which we shall call I_∞, by substituting Eq. 9 in Eq. 8. Thus

$$I_{\infty} = C - \frac{1}{T_o} \int_{-\infty}^{\infty} \phi_x(u)\, s(u)\, du \qquad (13)$$

Use of Eq. 12 or Eq. 11, and Parseval's relation, yields

$$I_{\infty} = C - \int_{-\infty}^{\infty} \frac{[\Phi(w)]^2}{\sum_{n=-\infty}^{\infty} \Phi(w - nw_r)}\, dw \qquad (14)$$

or

$$I_{\infty} = C - \int_{-\infty}^{\infty} \frac{[\Phi_1(f)]^2}{\sum_{n=-\infty}^{\infty} \Phi_1(f - nf_r)}\, df \qquad (15)$$

It is clear from Eq. 11 that $S_1(f)$ [and hence $S(w)$] is well defined for all values of f, because of the fact that $\Phi_1(f)$ must be everywhere non-negative. That is, the zeros of the denominator are, at most, only the zeros of $\Phi_1(f)$. No extra zeros can be added.

For the special case in which $\Phi_1(f)$ is zero outside the range $(-W, W)$, and $f_r \geq 2W$ [i.e., $T_o \leq 1/(2W)$], we have as one solution

$$S_1(f) \begin{cases} = T_o & \text{for } |f| \leq W \\ = 0 & \text{elsewhere} \end{cases}$$

$$s(t) = \frac{2W}{f_r} \cdot \frac{\sin 2Wt}{2Wt}$$

and

$$I_{\infty} = C - \int_{-\infty}^{\infty} \Phi_1(f)\, df = 0$$

D. W. Tufts

References

1. D. W. Tufts, A sampling theorem for stationary random processes, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1958, pp. 111-116.

2. W. R. Bennett, Statistics of regenerative digital transmission, Bell System Tech. J. 37, pp. 1501-1542 (1958).

3. Ibid, pp. 1505-1506.

F. NONLINEAR OPERATORS – CASCADING, INVERSION, AND FEEDBACK

This report is concerned with some of the properties of cascades of nonlinear operators and with methods of inversion. The formulas of Brilliant (1) and Barrett (2) are related to trees and partitions and are tabulated for convenient reference in sections 2 and 3. An inversion algorithm is presented and is used for determining the inverse of a polynomial, as well as for studying the effects of feedback on the nonlinear distortion of an amplifier.

The problem of inversion is crucial in nonlinear theory, for the solutions of nonlinear operator equations can be expressed as inverses. Thus, with the use of George's (3) notation, the solution of

$$z(t) = \underline{H}[y(t)] \tag{1}$$

is

$$y(t) = \underline{H}^{-1}[z(t)] \tag{2}$$

where y(t) and z(t) are a pair of functions of the independent variable t, and $\underline{H}$ is the operator whose inverse is being sought.

The inverse $\underline{H}^{-1}$, itself, is a solution of the operator equation

$$\underline{H} * \underline{K} = \underline{I} \tag{3}$$

in which $\underline{K}$ and $\underline{I}$ are the unknown and the identity operator, respectively, and the symbol * denotes "cascading," that is, application in sequence.

As a consequence of Eq. 3, inversion and cascading are intimately related, and cascading merits study.

1. Cascading of Polynomial Operators

We deal with those nonlinear operators that can be represented by polynomials in integrals, in this manner:

$$z(t) = \int_{-\infty}^{\infty} k_1(t-\tau)\, y(\tau)\, d\tau + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_2(t-\tau_1, t-\tau_2)\, y(\tau_1)\, y(\tau_2)\, d\tau_1 d\tau_2 + \dots \tag{4}$$

The kernels k_1, k_2, ... completely specify each operator.

When two operators are cascaded, as in Fig. X-9, a new operator results. As Brilliant (1) has shown, its kernels can be determined from those of the original operators by means of standard transformations. These are derived by substituting a polynomial with undetermined kernels for y(t) wherever y(t) occurs in Eq. 4. Terms of equal order in y(t) are collected after the orders of integration have been interchanged (2). The resulting expressions are sums of multiple convolutions of the original kernels. For example, the kernel of order 2, denoted

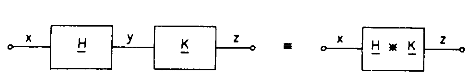

Fig. X-9. Cascaded operators.

by $(k*h)_2$ is given by

$$[k*h]_2(\tau_1, \tau_2) = \int_{-\infty}^{\infty} k_1(\sigma)\, h_2(\tau_1 - \sigma, \tau_2 - \sigma)\, d\sigma$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_2(\sigma_1, \sigma_2)\, h_1(\tau_1 - \sigma_1)\, h_2(\tau_2 - \sigma_2)\, d\sigma_1 d\sigma_2 \tag{5}$$

Integrations can be eliminated by taking a Fourier transformation (of two variables) of Eq. 5.

$$[K* H]_2(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [k*h]_2(\tau_1, \tau_2) \exp[j(\omega_1\tau_1 + \omega_2\tau_2)]\, d\tau_1 d\tau_2$$

$$= K_1(\omega_1 + \omega_2)\, H_1(\omega_1, \omega_2) + K_2(\omega_1, \omega_2)\, H_1(\omega_1)\, H_1(\omega_2) \tag{6}$$

The spectrum of the cascaded kernel of order 2 (and of every order greater than 2) is therefore a sum of products of spectra, instead of the product that is encountered in linear theory. Fourier transforms of several variables are discussed by Van der Pol and Bremmer (4). Other cascading equations are listed in Table X-1.

2. Inversion of Polynomial Operators

The inverse $\underline{H}^{-1}$ of the operator $\underline{H}$ is given by a solution for $\underline{K}$ of the equation

$$\underline{H} * \underline{K} = \underline{I} \tag{7}$$

in which $\underline{I}$ denotes the identity operator.

If the inverse has a representation in polynomial form, it can be determined by substituting a polynomial with unspecified kernels for $\underline{K}$ in Eq. 7, collecting terms of equal order on both sides of Eq. 7, and solving the resulting infinite set of equations. If these equations are arranged in their natural order, they can be solved in that order, for then each contains only one unknown. Thus the equation of order 1 is

$$H_1(\omega)\, K_1(\omega) = 1 \tag{8}$$

whence

$$K_1(\omega) = \frac{1}{H_1(\omega)} \tag{9}$$

just as in the linear case.

Next, we have

$$K_2(\omega_1,\omega_2)\, H_1(\omega_1)\, H_1(\omega_2) + K_1(\omega_1+\omega_2)\, H_2(\omega_1,\omega_2) = 0 \tag{10}$$

whence, with the use of Eq. 9,

$$K_2(\omega_1,\omega_2) = -\frac{H_2(\omega_1,\omega_2)}{H_1(\omega_1)\, H_1(\omega_2)\, H_1(\omega_1+\omega_2)} \tag{11}$$

This procedure is, of course, valid only as long as the inverse has the specified form, and then only for sufficiently small inputs that are within the circle of convergence (1). Such a solution exists whenever the original operator has an invertible linear term. However, when it does not, the method fails, even though an inverse may exist. For example, although a square has an inverse – the square root – that inverse exhibits infinite slope at the origin, and cannot be represented by a power series in that vicinity.

These results were obtained by Barrett (1), and Brilliant (2), and are entirely analogous to those for ordinary polynomials (5). Cascading and inversion polynomials are listed in Tables X-1 and X-2 for reference. They are related to the Bell polynomials, which are discussed by Riordan (6). They are derived in an asymmetrical form for operators that are themselves symmetrical.

Table X-1.

Cascading Polynomials (Orders 1-4).

$$[K*H]_1(\omega) = K_1(\omega)\, H_1(\omega)$$

$$[K*H]_2(\omega_1,\omega_2) = K_1(\omega_1+\omega_2)\, H_2(\omega_1,\omega_2) + K_2(\omega_1,\omega_2)\, H_1(\omega_1)\, H_1(\omega_2)$$

$$\begin{aligned}[K*H]_3(\omega_1,\omega_2,\omega_3) = {} & K_1(\omega_1+\omega_2+\omega_3)\, H_3(\omega_1,\omega_2,\omega_3) + 2K_2(\omega_1,\omega_2+\omega_3)\, H_1(\omega_1)\, H_2(\omega_2+\omega_3) \\ & + K_3(\omega_1,\omega_2,\omega_3)\, H_1(\omega_1)\, H_1(\omega_2)\, H_1(\omega_3)\end{aligned}$$

$$\begin{aligned}[K*H]_4(\omega_1,\omega_2,\omega_3,\omega_4) = {} & K_1(\omega_1+\omega_2+\omega_3+\omega_4)\, H_4(\omega_1,\omega_2,\omega_3,\omega_4) \\ & + 2K_2(\omega_1,\omega_2+\omega_3+\omega_4)\, H_1(\omega_1)\, H_3(\omega_2,\omega_3,\omega_4) \\ & + K_2(\omega_1+\omega_2,\omega_3+\omega_4)\, H_2(\omega_1,\omega_2)\, H_2(\omega_3,\omega_4) \\ & + 3K_3(\omega_1,\omega_2,\omega_3+\omega_4)\, H_1(\omega_1)\, H_1(\omega_2)\, H_2(\omega_3,\omega_4) \\ & + K_4(\omega_1+\omega_2+\omega_3+\omega_4)\, H_1(\omega_1)\, H_1(\omega_2)\, H_1(\omega_3)\, H_1(\omega_4)\end{aligned}$$

Table X-2.

Inversion Polynomials (Orders 1-4).

$$[H^{-1}]_1(\omega) = \frac{1}{H_1(\omega)}$$

$$[H^{-1}]_2(\omega_1,\omega_2) = \frac{H_2(\omega_1,\omega_2)}{H_1(\omega_1)H_1(\omega_2)H_1(\omega_1+\omega_2)}$$

$$[H^{-1}]_3(\omega_1,\omega_2,\omega_3) = \frac{1}{H_1(\omega_1+\omega_2+\omega_3)\prod_{i=1}^{3}H_i(\omega_i)}\left\{-H_3(\omega_1,\omega_2,\omega_3)+\frac{2H_2(\omega_1,\omega_2+\omega_3)H_2(\omega_2,\omega_3)}{H_1(\omega_2+\omega_3)}\right\}$$

$$\begin{aligned}[H^{-1}]_4(\omega_1,\omega_2,\omega_3,\omega_4) = \frac{1}{H_1(\omega_1+\omega_2,\omega_3+\omega_4)\prod_{i=1}^{4}H_i(\omega_i)}\Bigg\{&-H_4(\omega_1,\omega_2,\omega_3,\omega_4)\\ &+\frac{2H_2(\omega_1,\omega_2+\omega_3+\omega_4)H_3(\omega_2,\omega_3,\omega_4)}{H_1(\omega_2+\omega_3+\omega_4)}\\ &-\frac{H_2(\omega_1+\omega_2,\omega_3+\omega_4)H_2(\omega_1,\omega_2)H_2(\omega_3,\omega_4)}{H_1(\omega_1+\omega_2)H_1(\omega_3+\omega_4)}\\ &+\frac{3H_3(\omega_1,\omega_2,\omega_3+\omega_4)H_2(\omega_3,\omega_4)}{H_1(\omega_3+\omega_4)}\\ &-\frac{4H_2(\omega_1,\omega_2+\omega_3+\omega_4)H_2(\omega_2,\omega_3+\omega_4)H_2(\omega_3,\omega_4)}{H_1(\omega_2+\omega_3+\omega_4)H_1(\omega_3+\omega_4)}\Bigg\}\end{aligned}$$

Table X-3.

Feedback Polynomials.

$$G_1(\omega) = \frac{H_1(\omega)}{1+H_1(\omega)}$$

$G_n(\omega_1, \dots \omega_n)$, $n > 2$ is derived from $[H^{-1}]_n(\omega_1, \dots \omega_n)$ in Table X-2 by replacing H_1 with $1 + H_1$ and changing signs.

3. Feedback

When an operator $\underline{H}$ is placed in a feedback loop (Fig. X-10), the resulting operator, $\underline{G}$, is given by any one of the following equations:

$$\underline{G} = \underline{I} - (\underline{I}+\underline{H})^{-1} \tag{12}$$

$$\underline{G} = \underline{H} * (\underline{I}+\underline{H})^{-1} \tag{13}$$

$$\underline{G} = (\underline{I} + \underline{H}^{-1})^{-1} \tag{14}$$

$$\underline{G} = \underline{H} * (\underline{I}-\underline{G}) = \underline{I} - \underline{H}^{-1} * \underline{G} \tag{15}$$

These equations are valid only when the indicated inverses exist. (When $\underline{H}$ is a polynomial operator, (I+H) has an inverse in some vicinity of the origin whenever its linear part is invertible.)

Equation 12 indicates that the feedback polynomials can be derived from the inversion polynomials by replacing H_1 with $1 + H_1$ wherever $\underline{H}_1$ occurs, and changing signs, except in the case of G_1 which is given by

$$G_1(\omega) = 1 - \frac{1}{1 + H_1(\omega)} = \frac{H_1(\omega)}{1 + H_1(\omega)} \tag{16}$$

4. Partitions and Trees

The structures of the cascading and inversion polynomials can be related to the structures of trees and partitions, respectively. This provides a relatively easy means for writing them, and offers some insight into the behavior that they describe.

A partition of order n is a division of n identical objects into s cells. The representation (10:3, 3, 2, 1, 1), or that shown in Fig. X-11, denotes a partition of 10 objects into 5 cells containing 3, 3, 2, 1, and 1 objects, respectively. This partition has two "repetitions" of 3, and two of 1. We associate a coefficient, a, with each partition, which is the number of its rearrangements and is given by

$$a = \frac{s!}{r_1!\, r_2! \ldots} \tag{17}$$

where r_1, r_2, ... are the numbers of repetitions. In our example,

$$a = \frac{5!}{2!\,2!} = 30$$

A tree of order n is a sequence of partitions, the first of which has order n,

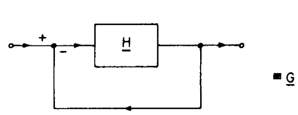

Fig. X-10. Operator in a feedback loop.

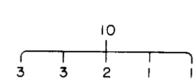

Fig. X-11. Representation of a partition.

terminating in cells containing only ones. It is denoted by the graph shown in Fig. X-12.

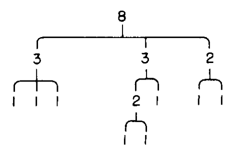

Fig. X-12. Representation of a tree.

The "multiplicity" m of a tree is the number of its partitions or nodes (m = 5 in Fig. X-12). The number of rearrangements of a tree, b, is the product of the coefficients a for all the partitions. It is

$$b = \frac{3!}{2!} \times \frac{3!}{3!} \times 2! \times \frac{2!}{2!} \times \frac{2!}{2!} = 6$$

The cascade spectrum of order n, $[k*h]_n$, is a sum of terms of positive sign. There is one term corresponding to every possible partition without regard to arrangement of order n, multiplied by the associated coefficient a. The correspondence is most easily illustrated by an example.

The spectrum $[k*h]_{10}$ has a term corresponding to the partition (10:3, 3, 2, 1, 1) and that term is

$$30K_5(\omega_1 + \omega_2 + \omega_3, \omega_4 + \omega_5 + \omega_6, \omega_7 + \omega_8, \omega_9, \omega_{10})\, H_3(\omega_1, \omega_2, \omega_3)\, H_3(\omega_4, \omega_5, \omega_6)$$
$$H_2(\omega_7, \omega_8)\, H_1(\omega_9)\, H_1(\omega_{10})$$

There is a single K-factor of order s (5, in this example) whose independent variables are sums of frequencies occurring in cells of 3, 3, 2, 1, and 1. There are s H-factors, each with the order and variables of one of these cells. The arrangement of the cells and the assignment of the subscripts is immaterial, as long as the indicated pairing is maintained.

There is a similar correspondence between inverse spectra and trees of the same order. An inverse spectrum of order n is a sum of terms, one for each partition of order n, with a coefficient $(-1)^m b$. Thus $[H^{-1}]_8$ has a term corresponding to the tree in Fig. X-12 and that term is

$$\frac{(-1)^5 \times 6 \times H_3(\omega_1+\omega_2+\omega_3, \omega_4+\omega_5+\omega_6, \omega_7+\omega_8)\, H_3(\omega_1, \omega_2, \omega_3)\, H_2(\omega_4+\omega_5, \omega_6)\, H_2(\omega_7, \omega_8)\, H_2(\omega_4, \omega_5)}{\prod_{i=1}^{8} H_i(\omega_i)\, H_1(\omega_1+\omega_2+\omega_3+\omega_4+\omega_5+\omega_6+\omega_7+\omega_8)\, H_1(\omega_1+\omega_2+\omega_3)\, H_1(\omega_4+\omega_5+\omega_6)\, H_1(\omega_7+\omega_8)\, H_1(\omega_4+\omega_5)}$$

The numerator is constructed by inspection of the tree, one partition at a time. The denominator has an H_1 factor for every H factor in the numerator, with a sum of the variables that occur. The subscripts correspond to downward paths in the tree.

The analogy between inverses and trees and the fact that a tree is a sequence of partitions, which are themselves analogous to cascades, suggests that an inverse might be constructed out of a sequence of cascades. This reasoning led to an algorithm, for which a simpler derivation, which is given in section 5a, was found later. It has also been found that this algorithm is a generalization of one of Newton's approximation methods for polynomials.

5. An Algorithm for the Inverse

a. Theorem

The inverse of any polynomial operator that has an analytic inverse can be represented by an algorithm; the n^{th} cycle of the algorithm differs from the true inverse, at most, with respect to terms whose order is greater than n. Sufficient (though not necessary) conditions for the convergence of the algorithm are stated in section 5c.

Consider a polynomial operator $\underline{H}$ that has an analytic inverse, $\underline{H}^{-1}$. Denote the linear part of $\underline{H}$ by $\underline{H}_1$, and the strictly nonlinear (7) remainder by $\underline{H}_\xi$, so that

$$\underline{H} = \underline{H}_1 + \underline{H}_\xi \tag{18}$$

Let $\underline{A}^n$ be the n^{th} cycle of the algorithm

$$\underline{A}^n = (\underline{H}_1)^{-1} - (\underline{H}_1)^{-1} * \underline{H}_\xi * \left\{(\underline{H}_1)^{-1} - (\underline{H}_1)^{-1} * \underline{H}_\xi * \left\{\dots \left\{(\underline{H}_1)^{-1}\right\} \dots\right\}\right\} \tag{19}$$

which is shown schematically in Fig. X-13. Let $\underline{E}^n$ denote the resulting error operator,

$$\underline{E}^n = \underline{H}^{-1} - \underline{A}^n \tag{20}$$

Then $\underline{E}^n$ has no terms whose order is less than (n+1). We denote this fact by the equation

$$N(\underline{E}^n) = n + 1 \tag{21}$$

in which $N(\underline{H})$ denotes a lower bound to the order of the lowest-order term of any operator $\underline{H}$.

This representation is significant in that it expresses the inverse of a nonlinear

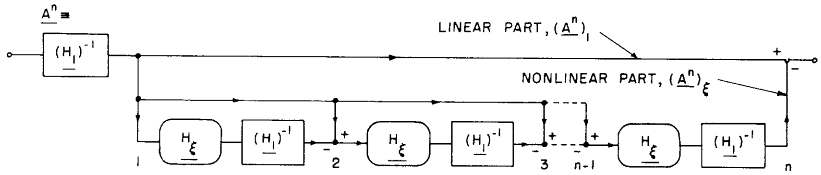

Fig. X-13. N^{th} cycle, $\underline{A}^n$, of an inversion algorithm.

operator explicitly in terms of its nonlinear part (which is specified) and the inverse of its linear part (which can be determined by the usual inversion methods of linear theory).

b. Proof of "Low-Order Convergence"

We develop an implicit formula for the inverse. Let $\underline{K}$ denote the inverse of $\underline{H}$. That is,

$$\underline{H} * \underline{K} = \underline{I} \tag{22}$$

Decomposing Eq. 22 into linear and strictly nonlinear parts (7), we have

$$(\underline{H}_1 + \underline{H}_\xi) * (\underline{K}_1 + \underline{K}_\xi) = \underline{I} \tag{23}$$

Since $\underline{H}_1$ is linear and therefore distributive,

$$\underline{H}_1 * \underline{K}_1 + \underline{H}_1 * \underline{K}_\xi + \underline{H}_\xi * (\underline{K}_1 + \underline{K}_\xi) = \underline{I} \tag{24}$$

Since $\underline{H}_1 * \underline{K}_1$ is the only term on the left-hand side that is linear [for, by construction, $N(\underline{K}_\xi) = N(\underline{H}_\xi) = 2$, which ensures that the remaining terms are strictly nonlinear (7)], and $\underline{I}$ is linear, we can split Eq. 20 into two parts:

$$\begin{aligned} &\underline{H}_1 * \underline{K}_1 = \underline{I} \\ &\underline{H}_1 * \underline{K}_\xi + \underline{H}_\xi * (\underline{K}_1 + \underline{K}_\xi) = \underline{O} \end{aligned} \tag{25}$$

Solving the linear equation and substituting the result, $(H_1)^{-1}$, in the nonlinear equation, we obtain an implicit formula in which $\underline{K}_\xi$ is the only unknown:

$$\underline{K}_\xi = -(\underline{H}_1)^{-1} * \underline{H}_\xi * \left\{(\underline{H}_1)^{-1} + \underline{K}_\xi\right\} \tag{26}$$

This is the basis for our algorithm (Eq. 19).

Next, we develop an iteration formula for E^n, which will allow us to prove our contention. By our definition, Eq. 20,

$$\underline{E}^n = \underline{K} - \underline{A}^n \tag{27}$$

whence

$$\begin{aligned} \underline{E}^n - \underline{K} &= -\underline{A}^n \\ \underline{E}^{n-1} - \underline{K} &= -\underline{A}^{n-1} \end{aligned} \tag{28}$$

but, by inspection of Eq. 19,

$$\underline{A}^n = (\underline{H}_1)^{-1} - (\underline{H}_1)^{-1} * \underline{H}_\xi * \underline{A}^{n-1} \tag{29}$$

therefore, using Eqs. 28 and 29, we have

$$\underline{E}^n - \underline{K} = -(\underline{H}_1)^{-1} + (\underline{H}_1)^{-1} * \underline{H}_\xi * \{\underline{K} - \underline{E}^{n-1}\} \tag{30}$$

or, using an equation similar to Eq. 18 for $\underline{K}$ on the left, we have

$$\underline{E}^n = (\underline{H}_1)^{-1} * (\underline{H}_\xi) * \{\underline{K} - \underline{E}^{n-1}\} + \underline{K}_\xi \tag{31}$$

Substitution of Eq. 26 for $\underline{K}_\xi$ yields our iteration formula

$$\underline{E}^n = (\underline{H}_1)^{-1} * \underline{H}_\xi * \{\underline{K} - \underline{E}^{n-1}\} - (\underline{H}_1)^{-1} * \underline{H}_\xi * \underline{K} \tag{32}$$

in which we have gathered $\underline{K}_\xi$ and $(\underline{H}_1)^{-1}$ into $\underline{K}$.

Now, for any equation of the type

$$\underline{A} = \underline{B} * (\underline{C}+\underline{D}) - \underline{B} * \underline{C} \tag{33}$$

it is true that

$$N(\underline{A}) \geqslant \{N(\underline{B}) - 1\} \min \{N(\underline{C}), N(\underline{D})\} + \max \{N(\underline{C}), N(\underline{D})\} \tag{34}$$

(This is a fairly obvious property of the order of the lowest-order terms of polynomials.) Since $N\left\{(\underline{H}_1)^{-1} * \underline{H}_\xi\right\} \geqslant 2$, while $N(\underline{K}) = 1$, applying Eq. 34 to Eq. 32, we obtain

$$N(\underline{E}^n) \geqslant 1 + N(\underline{E}^{n-1}) \geqslant n + 1 \tag{35}$$

where the second inequality in Eq. 35 follows by induction because $N(\underline{E}^1) = 2$. This completes the proof of the property that we shall refer to as "low-order convergence."

We wish to emphasize that this property is the result of the absence of a linear term in the operator $(\underline{H}_1)^{-1} * \underline{H}_\xi$ which operates once in every cycle. The feedback algorithm generated by the formula

$$\underline{G} = \underline{H} * (\underline{I}-\underline{G}) \tag{36}$$

does not have this property.

If the operator $\underline{H}_\xi$ lacks higher-order terms as well, then low-order convergence proceeds at a rate of more than one per cycle. We shall call the operator $(\underline{H}_1)^{-1} * \underline{H}_\xi$ the "constrictor" of the algorithm.

c. Metric Convergence

Although the algorithm $\underline{A}^n$ converges in low order, it does not always converge in a metric (8) sense. (But the former ensures that the latter implies convergence to $\underline{H}^{-1}$.)

The necessary conditions for metric convergence are not understood. However, sufficient conditions are easy enough to find. Thus it is sufficient for the constrictor $(\underline{H}_1)^{-1} * \underline{H}_\xi$ to be a contraction (9). [Contractions are discussed by Kolmogoroff and Fomin (10).]

It follows from the definition (9) of a contraction that if $\underline{X}$ is such an operator, while $\underline{P}$ and $\underline{Q}$ are arbitrary, then

$$\|\underline{X} * \underline{P} - \underline{X} * \underline{Q}\| \leq \alpha \|\underline{P} - \underline{Q}\|, \qquad \alpha < 1 \tag{37}$$

in which the vertical bars denote norm (8) in the metric.

Since $(\underline{H}_1)^{-1} * \underline{H}_\xi$ is a contraction and, from Eq. 30, we have

$$\underline{E}^n = (\underline{H}_1)^{-1} * \underline{H}_\xi * \{\underline{K} - \underline{E}^{n-1}\} - (\underline{H}_1)^{-1} * \underline{H}_\xi * \underline{K} \tag{38}$$

we can apply inequality 37, which yields

$$\begin{aligned} \|\underline{E}^n\| &\leq \alpha \|\underline{E}^{n-1}\| \\ &\leq \alpha^{n-1} \|\underline{E}^1\| \\ &= \alpha^{n-1} \|\underline{K}_\xi\|, \qquad \alpha < 1 \end{aligned} \tag{39}$$

Hence the error term is reduced by a factor less than α in each cycle, and convergence is ensured.

Alternative methods of producing contractions, for example, by splitting into other than linear and nonlinear parts, are being studied.

d. Bounds on the Lipschitz Constant, α

The existence of a contraction is investigated by a study of the constant α.

For any ordinary polynomial in one variable α is simply the maximum slope.

A linear operator whose kernel is $k(\tau)$ has:

$$\alpha \leq \int_{-\infty}^{\infty} |k(\tau)| \, d\tau \tag{40}$$

For a cascaded pair and for a product of operators the product of the individual coefficients is the pair coefficient. For a sum operator it is the sum of the individual coefficients.

A bound on α for the general polynomial operator is obtained as follows. The absolute difference between two outputs $y_1(t), y_2(t)$ corresponding to any two inputs $x_1(t), x_2(t)$ is given by

$$|y_1(t) - y_2(t)| = \left|\sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} k_n(t-\tau_1, \dots, t-\tau_n) \{[x_1(\tau_1) \dots x_1(\tau_n)] - [x_2(\tau_1) \dots x_2(\tau_n)]\} d\tau_1 \dots d\tau_n\right| \tag{41}$$

$$\leq \sum_{n=1}^{\infty} \| [x_1(\tau_1) \dots x_1(\tau_n)] - [x_2(\tau_1) \dots x_2(\tau_n)] \| \times \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |k_n(\tau_1, \dots, \tau_n)| \, d\tau_1 \dots d\tau_n \tag{42}$$

$$\leq \|x_1(\tau) - x_2(\tau)\| \sum_{n=1}^{\infty} \{\|x_1(\tau)\| + \|x_2(\tau)\|\}^{n-1} \times \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |k_n(\tau_1, \dots, \tau_n)| \, d\tau_1 \dots d\tau_n \tag{43}$$

where the last inequality has been derived as follows: Let

$$\begin{aligned} (x_1(\tau_1) \dots x_1(\tau_n) - x_2(\tau_1) \dots x_2(\tau_n)) &= D^n \\ (x_1(\tau_1) \dots x_1(\tau_n) + x_2(\tau_1) \dots x_2(\tau_n)) &= S^n \end{aligned} \tag{44}$$

Therefore

$$\begin{aligned} D^{n+1} &= \frac{1}{2} [x_1(\tau_{n+1}) - x_2(\tau_{n+1})] \, S^n + \frac{1}{2} [x_1(\tau_{n+1}) + x_2(\tau_{n+1})] \, D^n \\ S^{n+1} &= \frac{1}{2} [x_1(\tau_{n+1}) + x_2(\tau_{n+1})] \, S^n + \frac{1}{2} [x_1(\tau_{n+1}) - x_2(\tau_{n+1})] \, D^n \end{aligned} \tag{45}$$

Denoting $\|x_1(\tau) - x_2(\tau)\|$ by δ and $\|x_1(\tau)\| + \|x_2(\tau)\|$ by σ, and applying the triangle inequality to the absolute values of Eqs. 45, we obtain

$$\begin{aligned} |D^{n+1}| &\leq \frac{1}{2} \delta |S^n| + \frac{1}{2} \sigma |D^n| \\ |S^{n+1}| &\leq \frac{1}{2} \delta |D^n| + \frac{1}{2} \sigma |S^n| \end{aligned} \tag{46}$$

(Equations 45 and 46 were suggested by I. M. Jacobs.)

Since $\delta \leq \sigma$ by definition, it follows from Eq. 46 that if the following equations are true for n, then they are true for $n + 1$.

$$|D^n| \leq \delta \sigma^{n-1}$$
$$|S^n| \leq \sigma^n \tag{47}$$

Moreover, Eqs. 47 are true for n = 1, so that they are true for all n, by induction. Hence Eqs. 47 can be written:

$$\|D^n\| = \|[x_1(\tau_1) \ldots x_1(\tau_n)] - [x_2(\tau_1) \ldots x_2(\tau_n)]\| \leq \|x_1(\tau) - x_2(\tau)\| \left\{\|x_1(\tau)\| + \|x_2(\tau)\|^{n-1}\right\} \tag{48}$$

This is the inequality that was used in Eq. 43. Equation 43 implies (by the definition of a) that

$$a = \max_{x_1(\tau), x_2(\tau)} \sum_{n=1}^{\infty} \{\|x_1(\tau)\| + \|x_2(\tau)\|\}^{n-1} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} |k_n(\tau_1, \ldots, \tau_n)| \, d\tau_1 \ldots d\tau_n \tag{49}$$

If all inputs are bounded by

$$\|x_i(\tau)\| \leq B, \qquad i = 1, 2 \tag{50}$$

then we can apply expression 50 to Eq. 49 and express the result in the form of a derivative, which gives

$$a \leq \max_{\theta < 2B} \frac{d}{d\theta} \sum_{n=1}^{\infty} n^{-1} a_n \theta^n, \qquad a_n = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} |k_n(\tau_1, \ldots, \tau_n)| \, d\tau_1 \ldots d\tau_n \tag{51}$$

Thus a is bounded by the maximum slope of a polynomial in θ on the domain $\theta \leq 2B$.

This is a very general bound, and consequently will usually give pessimistic results. Better bounds can be fitted to individual circumstances.

e. An Example

A nonlinear operator without memory has the sufficient requirements of section 5c if the maximum absolute value of the slope of its nonlinear part is less than that of its linear part. Such a nonlinear operator without memory has been drawn at random in Fig. X-14, and the successive cycles of its algorithm are shown in Fig. X-15.

6. Feedback Algorithms

The inversion algorithm can be applied directly to the solution of feedback problems by means of the feedback equations, Eqs. 12, 13, and 14.

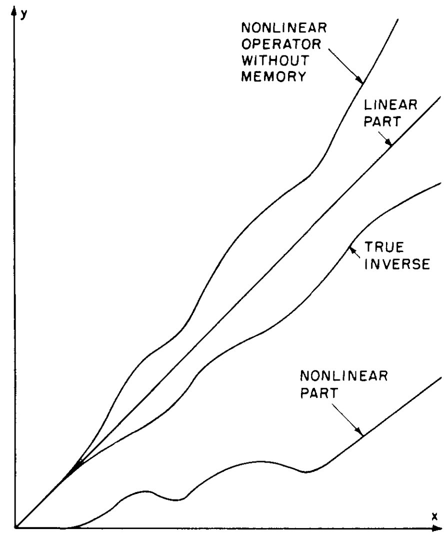

Fig. X-14. Nonlinear operator without memory and its inverse.

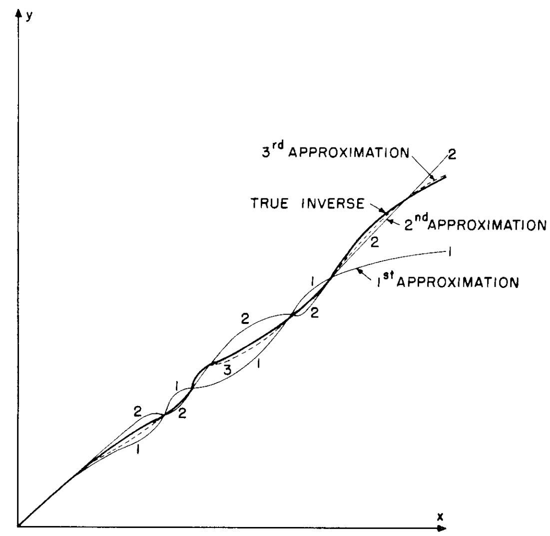

Fig. X-15. Successive approximations to the inverse.

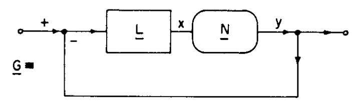

Fig. X-16. Linear-nonlinear-operator-without memory pair in a feedback loop.

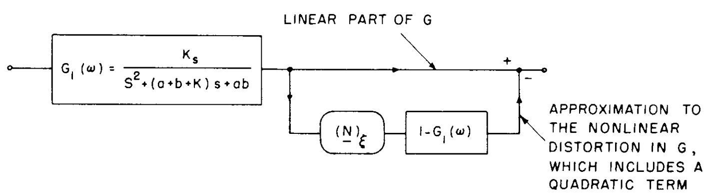

Fig. X-17. Second approximation to a high-gain amplifier with feedback.

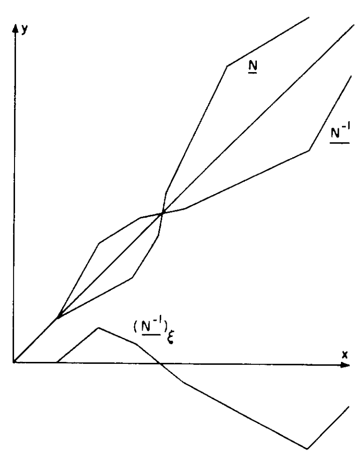

Fig. X-18. Nonlinear operator without memory, $\underline{N}$, its inverse, $\underline{N}^{-1}$, and the nonlinear part of its inverse, $(\underline{N}^{-1})_\xi$.

For example, we wish to study a linear operator $\underline{L}$, cascaded with a nonlinear operator without memory, $\underline{N}$, in a feedback loop, as in Fig. X-16.

We pick the feedback equation that gives the best rate of convergence. Thus if L is a high-gain bandpass amplifier with spectrum

$$L(\omega) = \frac{Ks}{(s+a)(s+b)} \tag{52}$$

and the input spectrum is negligible outside the high-gain parts of the passband (so that frequency distortion occurs at the edges of the passband), then rate of convergence will be high for high gain if the norm of the constrictor varies inversely as the gain. Equation 14 accomplishes this, for, applying the algorithm to it and simplifying, we obtain

$$\underline{G} = \underline{H} * (\underline{I}+\underline{L})^{-1} - (\underline{I}+\underline{L})^{-1} * (\underline{N}^{-1})_{\xi} * \left\{\underline{H} * (\underline{I}+\underline{L})^{-1} - (\underline{I}+\underline{L})^{-1} * (\underline{N}^{-1})_{\xi} * \left\{\dots\right.\right. \tag{53}$$

and if we restrict inputs to the vicinity of the passband, $\|(\underline{I}+\underline{L})^{-1}\|$ varies approximately inversely as the gain.

The norm of $(\underline{I}+\underline{L})^{-1} * (\underline{N}^{-1})_{\xi}$ is

$$\|(\underline{I}+\underline{L})^{-1} * (\underline{N}^{-1})_{\xi}\| = \alpha s$$

in which α is the norm of $(\underline{I}+\underline{L})^{-1}$, and s is the maximum slope of $(\underline{N}^{-1})_{\xi}$ (normalized with respect to $(\underline{N}^{-1})_1$).

If the gain is 100 and the half-power points are 20 and 20,000 cps, α is less than 3, so that s less than 0.33 ensures convergence, whatever the input is. However, if we exclude very low and very high frequencies, α assumes a much smaller value.

The second approximation to G is given by the system shown in Fig. X-17. The error in the linear and quadratic terms is zero, while that in the nonlinear part is less than 100 αs per cent. $(\underline{N}^{-1})_{\xi}$ is obtained graphically, as shown in Fig. X-18, by reflecting the graph of $\underline{N}$ about the 45° line and subtracting it.

G. D. Zames

References and Footnotes

1. M. B. Brilliant, Theory of the analysis of nonlinear systems, Sc.D. Thesis, Department of Electrical Engineering, M.I.T., Jan. 15, 1958.

2. J. F. Barrett, The use of functionals in the analysis of non-linear physical systems, Statistical Advisory Unit Report 1/57, Ministry of Supply, Great Britain, 1957.

3. D. A. George, An algebra of continuous systems, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1958, p. 95.

4. B. Van der Pol and H. Bremmer, Operational Calculus Based on the Two-Sided Laplace Integral (Cambridge University Press, London, 1950).

5. K. Knopp, Infinite Series and Sequences (Dover Publications, Inc., New York, 1956).

6. J. Riordan, An Introduction to Combinatorial Analysis (John Wiley and Sons, Inc., New York, 1958).

7. "Strictly nonlinear" or "N = 2" means a sum of terms that are homogeneous of order 2 and higher; "linear" means distributive with respect to linear combinations, and hence homogeneous of order 1.

8. A metric assigns a distance $\|F-G\|$ to every pair of elements F and G. We employ a metric in which $\|F\|$ (norm of F) denotes maximum absolute value on the infinite interval when F is a function, and maximum norm of the range (output) function when F is an operator.

9. A contraction is an operator that reduces the distances between all pairs of domain (input) functions by factors γ with the property that $\gamma \leqslant a < 1$.

10. A. N. Kolmogorov and S. V. Fomin, Elements of the Theory of Functionals, Vol. 1 (Graylock Press, Rochester, New York, 1957).