Prof. Y. W. Lee
Prof. A. G. Bose
Prof. I. M. Jacobs
D. A. Chesler
D. A. George
A. D. Hause
K. L. Jordan, Jr.
W. Neugebauer
M. Schetzen
D. W. Tufts
C. E. Wernlein, Jr. (Absent)
G. D. Zames

A. CONTINUOUS NONLINEAR SYSTEMS

This study was completed and presented by D. A. George as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science, Department of Electrical Engineering, M.I.T., July 1959. The study will also be published as Technical Report 355.

Y. W. Lee

B. CONNECTIVITY IN PROBABILISTIC GRAPHS

This research was performed by I. M. Jacobs. The results, which were submitted to the Department of Electrical Engineering, M.I.T., August 1959, as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science, will also be published as Technical Report 356.

Y. W. Lee

C. A THERMAL CORRELATOR

This study was carried out by W. Neugebauer. The results were submitted to the Department of Electrical Engineering, M.I.T., 1959, as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

Y. W. Lee

D. NONLINEAR FILTER FOR THE REDUCTION OF RECORD NOISE

This study was completed by D. A. Shnidman. In June 1959 he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

A. G. Bose

E. NONLINEAR OPERATIONS ON GAUSSIAN PROCESSES

1. Orthogonal Functionals of Correlated Gaussian Processes

Wiener (1) and Barrett (2) have discussed an orthogonal hierarchy $\{G_N(t, K_N, x)\}$ of functionals of a zero-mean Gaussian process x(t). Wiener considered the case in which x(t) is white Gaussian noise; Barrett considered the case in which x(t) is non-white Gaussian noise. They have shown that these functionals of the same Gaussian process x(t) have the following type of orthogonality

$$\overline{G_N(t, K_N, x)\ G_M(t, L_M, x)} = 0 \qquad N \neq M \tag{1}$$

where the bar indicates ensemble average.

We shall prove that the same type of orthogonality that exists for these functionals of the same Gaussian process also exists for functionals of correlated Gaussian processes. That is, we shall prove that

$$\overline{G_N(t, K_N, x)\, G_M(t, L_M, y)} = 0 \qquad N \neq M \tag{2}$$

with x(t) and y(t) stationary, correlated, zero-mean Gaussian processes. The orthogonal G-functions are defined as

$$G_N(t, K_N, x) = \sum_{\nu=0}^{\left[\frac{N}{2}\right]} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_{N-2\nu}^{(N)} K_N(\tau_1, \ldots \tau_N)\, x(t-\tau_1) \ldots x(t-\tau_{N-2\nu})$$
$$\times R_{xx}(\tau_{N-2\nu+1}-\tau_{N-2\nu+2}) \ldots R_{xx}(\tau_{N-1}-\tau_N)\, d\tau_1 \ldots d\tau_N \tag{3}$$

and

$$G_M(t, L_M, y) = \sum_{\nu=0}^{\left[\frac{M}{2}\right]} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_{M-2\nu}^{(M)} L_M(\sigma_1, \ldots \sigma_M)\, y(t-\sigma_1) \ldots y(t-\sigma_{M-2\nu})$$
$$\times R_{yy}(\sigma_{M-2\nu+1}-\sigma_{M-2\nu+2}) \ldots R_{yy}(\sigma_{M-1}-\sigma_M)\, d\sigma_1 \ldots d\sigma_M \tag{4}$$

where $K_N(\tau_1, \ldots \tau_N)$ is an arbitrary symmetric function of the τ's, and $L_M(\sigma_1, \ldots \sigma_M)$ is an arbitrary symmetric function of the σ's. The set of constants $\left\{a_{N-2\nu}^{(N)}\right\}$ is the coefficient of $s^{N-2\nu}$ in the N^{th}-order Hermite polynomial $H_N(s)$, which is defined (3) as

$$H_N(s) = \sum_{\nu=0}^{\left[\frac{N}{2}\right]} a_{N-2\nu}^{(N)} s^{N-2\nu} \tag{5}$$

$$a_{N-2\nu}^{(N)} = \frac{(-1)^{\nu}\, N!}{2^{\nu}(N-2\nu)!\, \nu!} \tag{6}$$

The functions $R_{xx}(\tau)$ and $R_{yy}(\tau)$ are the autocorrelation functions of the x and y processes, respectively. The expression $\left[\frac{N}{2}\right]$ equals $\frac{N}{2}$ if N is even, and equals $\frac{N-1}{2}$ if N is odd.

We shall also prove that

$$\overline{G_N(t, K_N, x)\, G_N(t, L_N, y)} = N! \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_N(\tau_1, \ldots \tau_N)\, L_N(\sigma_1, \ldots \sigma_N)$$
$$\times R_{xy}(\tau_1-\sigma_1) \ldots R_{xy}(\tau_N-\sigma_N)\, d\tau_1 \ldots d\tau_N d\sigma_1 \ldots d\sigma_N \tag{7}$$

where $R_{xy}(\tau) = \overline{x(t)\, y(t+\tau)}$.

Since Eqs. 2 and 7 will be proved valid for an arbitrary symmetric kernel L_M, and because it can be shown that

$$G_M(t+a, L_M, y) = G_M(t, L'_M, y) \tag{8}$$

where

$$L'_M(\sigma_1, \dots \sigma_M) = L_M(\sigma_1+a, \dots \sigma_M+a) \tag{9}$$

then Eqs. 2 and 7 imply that the two following equations are true:

$$\overline{G_N(t, K_N, x)\, G_M(t+a, L_M, y)} = 0 \qquad N \neq M \tag{10}$$

$$\overline{G_N(t, K_N, x)\, G_N(t+a, L_N, y)} = N! \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_N(\tau_1, \dots \tau_N)\, L_N(\sigma_1+a, \dots \sigma_N+a)$$
$$\times R_{xy}(\tau_1-\sigma_1) \dots R_{xy}(\tau_N-\sigma_N)\, d\tau_1 \dots d\tau_N d\sigma_1 \dots d\sigma_N \tag{11}$$

In proving the orthogonality condition given by Eq. 2, we shall consider only the case of $N > M$, because the proof for $M > N$ is similar to that for $N > M$.

The expression for $G_M(t, L_M, y)$ in Eq. 4 contains only operations of degree M, or lower, in y. For example, the term

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} a^{(M)}_{M-2\nu} L_M(\sigma_1, \dots \sigma_M)\, y(t-\sigma_1) \dots y(t-\sigma_{M-2\nu})$$
$$\times R_{yy}(\sigma_{M-2\nu+1}-\sigma_{M-2\nu+2}) \dots R_{yy}(\sigma_{M-1}-\sigma_M)\, d\sigma_1 \dots d\sigma_M \tag{12}$$

contains the product of $M-2\nu$ y's and is thus of degree $M-2\nu$ in y. Therefore to prove Eq. 2 for $N > M$ it is sufficient to prove that

$$\overline{G_N(t, K_N, x) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} L_j(\sigma_1, \dots \sigma_j)\, y(t-\sigma_1) \dots y(t-\sigma_j)\, d\sigma_1 \dots d\sigma_j} = 0 \qquad j = 0, \dots N-1 \tag{13}$$

for all symmetric kernels G_N and L_j.

We shall now prove Eq. 13 valid. If we expand $G_N(t, K_N, x)$ as in Eq. 3 and interchange order of integration and averaging, then Eq. 13 becomes

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_j(\sigma_1, \ldots \sigma_j) \sum_{\nu=0}^{\left[\frac{N}{2}\right]} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_{N-2\nu}^{(N)} K_N(\tau_1, \ldots \tau_N)$$

$$\times \overline{x(t-\tau_1) \ldots x(t-\tau_{N-2\nu})\, y(\sigma_1) \ldots y(\sigma_j)}$$

$$\times R_{xx}(\tau_{N-2\nu+1}-\tau_{N-2\nu+2}) \ldots R_{xx}(\tau_{N-1}-\tau_N)\, d\tau_1 \ldots d\tau_N d\sigma_1 \ldots d\sigma_j = 0 \qquad j = 0, \ldots N-1 \tag{14}$$

The averaging operation in Eq. 14 is the average of a product of Gaussian variables. This average is the sum, over all ways of pairing, of products of averages of pairs of the Gaussian variables. We now divide the averaging operation in Eq. 14 into two parts: in part 1 each y is paired with an x for averaging; in part 2 at least two y's are paired together for averaging. If in part 2, for a particular j, we first average two y's that are paired together (because of the symmetry of L_j it does not matter which two) and integrate on their corresponding sigmas, then the rest of the expression is again in the form of the left-hand side of Eq. 14, except that j is smaller by 2. Therefore if we prove the set of Eqs. 14 in order of increasing j, then for each j it is only necessary to prove that the term from part 1 is zero, because the term from part 2 will have been proved equal to zero in the proof of the $(j-2)$ equation of Eqs. 14.

If we perform part 1 by pairing and averaging of each y with an x, and pair and average the remaining x's among themselves, then Eq. 14 can be replaced by

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L_j(\sigma_1, \ldots \sigma_j)\, K_N(\tau_1, \ldots \tau_N)\, R_{xy}(\tau_1-\sigma_1) \ldots R_{xy}(\tau_j-\sigma_j)$$

$$\times R_{xx}(\tau_{j+1}-\tau_{j+2}) \ldots R_{xx}(\tau_{N-1}-\tau_N)\, d\tau_1 \ldots d\tau_N d\sigma_1 \ldots d\sigma_j$$

$$\times \sum_{\nu=0}^{\frac{N-j}{2}} b_{N-2\nu, j} a_{N-2\nu}^{(N)} = 0 \qquad j = 0, \ldots N-1 \tag{15}$$

which is part 1 of the averaging, and then a proof of Eq. 15 will constitute a proof of Eq. 14. The constant $b_{n-2\nu, j}$ is the number of ways in which j y's can be paired with j of $N-2\nu$ x's and the remaining $N-2\nu-j$ x's paired among themselves. It should be noted that this pairing can exist if and only if j and N are either both even or odd, and $\nu \leq \frac{1}{2}(N-j)$.

To prove Eq. 15 valid we shall prove that

$$\sum_{\nu=0}^{\frac{N-j}{2}} b_{N-2\nu, j} a_{N-2\nu}^{(N)} = 0 \qquad j = 0, \ldots N-1 \tag{16}$$

The method of proof is to show that the orthogonality of the Hermite polynomials implies the validity of Eq. 16. The Hermite polynomials defined in Eqs. 5 and 6 are orthogonal with respect to the weighting function $e^{-s^2/2}$. That is,

$$\int_{-\infty}^{\infty} H_N(s)\, H_M(s)\, e^{-s^2/2}\, ds = \begin{cases} 0 & N \neq M \\ N! & N = M \end{cases} \tag{17}$$

Now let u be a Gaussian random variable of zero mean and unit variance. Thus, the probability density, p(u), is given by

$$p(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \tag{18}$$

The expected value of the product of two Hermite polynomials in u is then

$$\overline{H_N(u)\, H_M(u)} = \int_{-\infty}^{\infty} H_N(u)\, H_M(u)\, p(u)\, du$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_N(u)\, H_M(u)\, e^{-u^2/2}\, du \tag{19}$$

By substituting Eq. 17 in Eq. 19 it can be seen that

$$\overline{H_N(u)\, H_M(u)} = 0 \qquad N \neq M \tag{20}$$

Any positive integer power of u can be written as a weighted sum of Hermite polynomials in u of degree less than or equal to the power of u. That is, for any positive integer j there exists a set of constants $\{c_{j,M}\}$ which satisfies the equation

$$u^j = \sum_{M=0}^{j} c_{j,M} H_M(u) \tag{21}$$

It can be shown that

$$\overline{u^j H_N(u)} = 0 \qquad j = 0, \ldots N-1 \tag{22}$$

because from Eq. 21

$$\overline{u^j H_N(u)} = \sum_{M=0}^{j} c_{j,M} \overline{H_M(u)\, H_N(u)} \qquad j = 0, \ldots N-1 \tag{23}$$

and because from Eq. 20 the average values in Eq. 23 are zero.

If we use the expansion for the Hermite polynomial (Eq. 5), then Eq. 22 becomes

$$\overline{u_j \sum_{\nu=0}^{\left[\frac{N}{2}\right]} a_{N-2\nu}^{(N)} u^{N-2\nu}} = 0 \qquad j = 0, \ldots N-1 \tag{24}$$

Equation 24 contains the average of products of a Gaussian variable. In a manner identical to that used in deriving Eq. 15 from Eq. 14, it can be shown that in Eq. 24 it is only necessary to consider the pairing (for averaging purposes) of each u in u^j with a u in $u^{N-2\nu}$. With this pairing, Eq. 24 becomes

$$\sum_{\nu=0}^{\frac{N-j}{2}} b_{N-2\nu, j} a_{N-2\nu}^{(N)} \left(\overline{u^2}\right)^{\frac{N-2\nu+j}{2}} = 0 \qquad j = 0, \ldots N-1 \tag{25}$$

where the constant $b_{N-\nu, j}$ is the same as in Eq. 15. Since the variable u has unit variance, Eq. 25 becomes Eq. 16.

We have now proved Eq. 2, since Eq. 16 has been proved and it proves Eq. 15 which proves Eq. 13, which in turn proves Eq. 2.

We shall now prove Eq. 7. If we expand $G_N(t, L_N, y)$ as in Eq. 4 and apply the orthogonality expressed in Eq. 13, then the left-hand side of Eq. 7 becomes

$$\overline{G_N(t, K_N, x)\, G_N(t, L_N, y)}$$

$$= \overline{G_N(t, K_N, x) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_N^{(N)} L_N(\sigma_1, \ldots \sigma_N)\, y(t-\sigma_1) \ldots y(t-\sigma_N)\, d\sigma_1 \ldots d\sigma_N} \tag{26}$$

From Eq. 6 it can be seen that $a_N^{(N)} = 1$.

The averaging operation in the right-hand side of Eq. 26 involves the average of products of the Gaussian variables x and y. We now divide this averaging operation into two parts: in part 1 each y is paired with an x for averaging; in part 2 at least two y's are paired together for averaging. If in part 2 we first average two y's that are paired together and integrate on their respective sigmas, then that term has the form

$$\overline{G_N(t, K_N, x) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} L'_{N-2}(\sigma_1, \ldots \sigma_{N-2})\, y(t-\sigma_1) \ldots y(t-\sigma_{N-2})\, d\sigma_1 \ldots d\sigma_{N-2}} \tag{27}$$

where

$$L'_{N-2}(\sigma_1, \ldots \sigma_{N-2}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L_N(\sigma_1, \ldots \sigma_N)\, R_{yy}(\sigma_{N-1}-\sigma_N)\, d\sigma_{N-1} d\sigma_N$$

From Eq. 13 it is seen that expression 27 is equal to zero. Therefore in performing the averaging in the right-hand side of Eq. 26 it is necessary to consider only the terms of part 1, that is, those terms in which each of the N y's is paired with an x. However, in the expansion of Eq. 3 for $G_N(t, K_N, x)$, only the term in which $\nu = 0$ has at least N x's, and that term has exactly N x's. There are N! different ways in which N y's can be paired with N x's. If we use only the $\nu = 0$ term from Eq. 3, and use the pairing of each y with an x, and use the fact that $a_N^{(N)} = 1$, then Eq. 26 becomes Eq. 7, which was to be proved.

2. A Crosscorrelation Property of Certain Nonlinear Functionals of Correlated Gaussian Processes

We shall now show that the crosscorrelation between a G-function of one Gaussian process and a nonlinear no-memory function of a second correlated Gaussian process is, except for a scale factor, independent of the nonlinear no-memory operation. In particular, if x(t) and y(t) are stationary, correlated, zero-mean Gaussian processes, and if f is a nonlinear no-memory operation on y, then the crosscorrelation (a function of the delay α) between $G_N(t, K_N, x)$ and $f[y(t+\alpha)]$ is given by

$$\overline{G_N(t, K_N, x)\, f[y(t+\alpha)]} = C_f \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_N(\tau_1, \ldots \tau_N) \times R_{xy}(\tau_1+\alpha) \ldots R_{xy}(\tau_N+\alpha)\, d\tau_1 \ldots d\tau_N \tag{28}$$

It should be noted that in the right-hand side of Eq. 28, only the scale factor C_f depends on the nonlinear no-memory operation f.

Before proving Eq. 28 we shall relate this work to the results of others. Bussgang (4) proved that for stationary, correlated, zero-mean Gaussian processes the crosscorrelation between one process and a nonlinear no-memory function of the second process is identical, except for a scale factor, to the crosscorrelation between the two Gaussian processes. If we use the notation of the previous paragraph, Bussgang's result is given by

$$\overline{x(t)\, f[y(t+\alpha)]} = C_f' R_{xy}(\alpha) \tag{29}$$

where C_f' is a function of the nonlinear no-memory operation, f. We shall now show that Eq. 29 is a special case of Eq. 28 with the two parameters, N = 1 and $K_1(\tau_1) = \delta(\tau_1)$. Substituting these parameters for $G_N(t, K_N, x)$ in Eq. 3, we obtain

$$G_1(t, \delta(\tau_1), x) = \int_{-\infty}^{\infty} \delta(\tau_1)\, x(t-\tau_1)\, d\tau_1 = x(t) \tag{30}$$

If we substitute these parameters and Eq. 30 in Eq. 28, we obtain Bussgang's result:

$$\overline{x(t)\, f[y(t+a)]} = C_f \int_{-\infty}^{\infty} \delta(\tau_1)\, R_{xy}(\tau_1+a)\, d\tau_1 = C_f R_{xy}(a)$$

By setting $C_f = C_f'$, we have shown that Eq. 28 implies Eq. 29.

Another related paper is that of Nuttall (5) on separable processes. By using Nuttall's definition of "separable," it can be shown that $G_N(t, K_N, x)$ is separable with respect to $y(t)$. However, for $N \neq 1$, the processes $G_N(t, K_N, x)$ and $y(t)$ are uncorrelated, and therefore, for our purpose, some of Nuttall's theorems on properties of separable processes have to be modified slightly. For example, Nuttall proves that the fact that one process is separable with respect to a second process implies, and is implied by, the fact that the crosscovariance function between the first process and a nonlinear no-memory function of the second process is identical, except for a scale factor, to the crosscovariance function between the first process and the second process. This theorem could be modified to state: The fact that one process is separable with respect to a second process implies, and is implied by, the fact that the crosscovariance function between the first process and a nonlinear no-memory function of the second process is, except for a scale factor, independent of the nonlinear no-memory operation. With respect to our work, this modified theorem means that Eq. 28 implies that $G_N(t, K_N, x)$ is separable with respect to $y(t)$, and that the separability of $G_N(t, K_N, x)$ with respect to $y(t)$ implies the invariance of the crosscorrelation function to nonlinear no-memory operations.

Another of Nuttall's (5) results is that the square of a zero-mean Gaussian process $x(t)$ is separable with respect to itself, and hence has the invariance property

$$\overline{\left[x^2(t)-\overline{x^2(t)}\right] f_1\left[x^2(t+a)\right]} = C''_{f_1} \overline{\left[x^2(t)-\overline{x^2(t)}\right] x^2(t+a)} \tag{31}$$

By evaluating the average in the right-hand side of Eq. 31, we obtain

$$\overline{\left[x^2(t)-\overline{x^2(t)}\right] f_1\left[x^2(t+a)\right]} = 2C''_{f_1} R^2_{xx}(a) \tag{32}$$

We shall now show that Eq. 28 implies Eq. 32. From Eq. 3 we note that

$$G_2\left[t, \delta(\tau_1)\delta(\tau_2), x\right] = x^2(t) - \overline{x^2(t)} \tag{33}$$

The nonlinear no-memory operation f_1 on $x^2(t)$ can also be viewed as a different nonlinear no-memory operation f_2 on $x(t)$ as follows:

$$f_2(x) \equiv f_1(x^2) \tag{34}$$

If we substitute from Eqs. 33 and 34 in Eq. 32, then the left-hand side of Eq. 32 becomes

$$\overline{[x^2(t)-\overline{x^2(t)}]\, f_1[x^2(t+a)]} = \overline{G_2[(t, \delta(\tau_1)\delta(\tau_2), x]\, f_2[x(t+a)]} \tag{35}$$

Substituting Eq. 28 in the right-hand side of Eq. 35, we obtain

$$\overline{[x^2(t)-\overline{x^2(t)}]\, f_1[x^2(t+a)]} = C_{f_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(\tau_1)\, \delta(\tau_2)\, R_{xx}(\tau_1+a)\, R_{xx}(\tau_2+a)\, d\tau_1 d\tau_2 \tag{36}$$

If we carry out the integration, Eq. 36 becomes

$$\overline{[x^2(t)-\overline{x^2(t)}]\, f_1[x^2(t+a)]} = C_{f_2} R_{xx}^2(a) \tag{37}$$

By equating C_{f_2} with $2C''_{f_1}$ we see that we have derived Eq. 32 from Eq. 28.

With respect to Nuttall's (5) separable class, it can be shown that

$$G_N[t, \delta(\tau_1)\dots\delta(\tau_N), x]$$

is separable with respect to itself. To prove this separability, it is only necessary to prove that the invariance property exists, that is, that

$$\overline{G_N[t, \delta(\tau_1)\dots\delta(\tau_N), x]\, f\{G_N[t+a, \delta(\tau_1)\dots\delta(\tau_N), x]\}}$$

is, except for a scale factor, independent of the nonlinear no-memory operation, f. By viewing the nonlinear no-memory operation on

$$G_N[t+a, \delta(\tau_1)\dots\delta(\tau_N), x]$$

as a different nonlinear no-memory operation on $x(t+a)$, and by then using Eq. 28, we can see that the invariance property does exist.

We shall now prove Eq. 28 true. We expand $f[y(t+a)]$ in Hermite polynomials in $[R_{yy}(0)]^{-1/2}\, y(t+a)$. That is, if we define

$$u(t+a) = [R_{yy}(0)]^{-1/2}\, y(t+a) \tag{38}$$

then we can write

$$f[y(t+a)] = \sum_{j=0}^{\infty} c_j H_j[u(t+a)] \tag{39}$$

By use of the orthogonality (Eq. 17) of the Hermite polynomials, it can be seen that c_j is given by

$$c_j = \frac{1}{j!}\int_{-\infty}^{\infty} f(s)\, H_j(s)\, e^{-s^2/2}\, ds \tag{40}$$

We shall now show that each $H_j[u(t+\alpha)]$ can be expanded as a G-function of u. We first expand these Hermite polynomials as in Eq. 5. Thus

$$H_j[u(t+\alpha)] = \sum_{\nu=0}^{\left[\frac{j}{2}\right]} a_{j-2\nu}^{(j)} [u(t+\alpha)]^{j-2\nu} \tag{41}$$

The right-hand side of Eq. 41 can then be rewritten as

$$H_j[u(t+\alpha)] = \sum_{\nu=0}^{\left[\frac{j}{2}\right]} a_{j-2\nu}^{(j)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta(\sigma_1) \ldots \delta(\sigma_j)\, u(t+\alpha-\sigma_1) \ldots u(t+\alpha-\sigma_{j-2\nu})$$

$$\times R_{uu}(\sigma_{j-2\nu+1}-\sigma_{j-2\nu+2}) \ldots R_{uu}(\sigma_{j-1}-\sigma_j)\, d\sigma_1 \ldots d\sigma_j \tag{42}$$

The fact that the right-hand side of Eq. 42 is equal to the right-hand side of Eq. 41 is shown by performing the integration in Eq. 42 and by using the fact that

$$R_{uu}(0) = \overline{u(t)\, u(t)} = \overline{y(t)[R_{yy}(0)]^{-1/2}\, y(t)[R_{yy}(0)]^{-1/2}} = 1 \tag{43}$$

Comparing the right-hand side of Eq. 42 with the definition of a G-function given by Eq. 4, we find that

$$H_j[u(t+\alpha)] = G_j(t+\alpha, L_j, u) \tag{44}$$

where

$$L_j(\sigma_1, \ldots \sigma_j) = \delta(\sigma_1) \ldots \delta(\sigma_j) \tag{45}$$

We substitute Eq. 44 in Eq. 39 to obtain

$$f[y(t+\alpha)] = \sum_{j=0}^{\infty} c_j G_j(t+\alpha, L_j, u) \tag{46}$$

Using Eq. 46, we now evaluate the left-hand side of Eq. 28 and obtain

$$\overline{G_N(t, K_N, x)\, f[y(t+\alpha)]} = \sum_{j=0}^{\infty} c_j \overline{G_j(t+\alpha, L_j, u)\, G_N(t, K_N, x)} \tag{47}$$

Notice that u(t) and x(t) are correlated Gaussian variables with the crosscorrelation function given by

$$R_{xu}(\tau) = \overline{x(t)\ u(t+\tau)} = \overline{x(t)\ y(t+\tau)[R_{yy}(0)]^{-1/2}} = R_{xy}(\tau)[R_{yy}(0)]^{-1/2} \tag{48}$$

Changing orders of averaging and summation in Eq. 47, and using Eqs. 10 and 11 to evaluate the averages, we obtain

$$\overline{G_N(t, K_N, x)\ f[y(t+a)]} = c_N N! \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_N(\tau_1, \dots \tau_N)\ L_N(\sigma_1+a, \dots \sigma_N+a)$$
$$\times R_{xu}(\tau_1-\sigma_1) \dots R_{xu}(\tau_N-\sigma_N)\ d\tau_1 \dots d\tau_N d\sigma_1 \dots d\sigma_N \tag{49}$$

By substituting Eqs. 48 and 45 in Eq. 49, we obtain

$$\overline{G_N(t, K_N, x)\ f[y(t+a)]} = c_N N! [R_{yy}(0)]^{-N/2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_N(\tau_1, \dots \tau_N)\ \delta(\sigma_1+a) \dots \delta(\sigma_N+a)$$
$$\times R_{xy}(\tau_1-\sigma_1) \dots R_{xy}(\tau_N-\sigma_N)\ d\tau_1 \dots d\tau_N d\sigma_1 \dots d\sigma_N \tag{50}$$

After integration on the sigmas, Eq. 50 becomes

$$\overline{G_N(t, K_N, x)\ f[y(t+a)]} = c_N N! [R_{yy}(0)]^{-N/2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_N(\tau_1, \dots \tau_N)$$
$$\times R_{xy}(\tau_1+a) \dots R_{xy}(\tau_N+a)\ d\tau_1 \dots d\tau_N \tag{51}$$

Equation 51 proves Eq. 28 valid if C_f is defined as

$$C_f \equiv c_N N! [R_{yy}(0)]^{-N/2}$$

If we use the definition of c_N given in Eq. 40, C_f can be written as

$$C_f = [R_{yy}(0)]^{-N/2} \int_{-\infty}^{\infty} f(s)\ H_N(s)\ e^{-s^2/2}\ ds$$

D. A. Chesler

References

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

(References continued on following page)

3. H. Grad, Note on N-Dimensional Hermite Polynomials, Communications on Pure and Applied Mathematics, Vol. 2 (Interscience Publishers, Inc., New York, 1949), p. 330.

4. J. J. Bussgang, Crosscorrelation functions of amplitude-distorted Gaussian signals, Technical Report 216, Research Laboratory of Electronics, M.I.T., March 26, 1952.

5. A. H. Nuttall, Theory and application of the separable class of random processes, Technical Report 343, Research Laboratory of Electronics, M.I.T., May 26, 1958.

F. FM SPECTRA FOR GAUSSIAN MESSAGES

1. Introduction

Wiener has shown how to calculate the power density spectrum of an FM signal that is modulated by a Gaussian message (1). We shall consider here (a) the infinite series representation of slightly different FM signal spectra, (b) some bounds on the error, which are incurred by approximating these infinite series with finite partial sums, and (c) the asymptotic behavior of wideband FM spectra and bandwidth relations.

2. Spectrum Computation

Let z_t be a Gaussian random waveform from an ergodic ensemble, with the expected value

$$E\left[z_t^2\right] = 1 \tag{1}$$

If we call the derivative of z_t (which we shall assume to exist almost everywhere) a message, then the signal

$$g_t = e^{iaz_t} \tag{2}$$

can be regarded as a phasor representation of an FM signal with no carrier and a Gaussian modulating message. The quantity a, which is the rms phase deviation of the signal, is assumed to be positive. By representing g_t as a series of orthogonal homogeneous polynomial functionals, Wiener (2) obtained its autocorrelation function, $R_{gg}(\tau)$, in terms of the autocorrelation function of z_t, $R_{zz}(\tau)$, as

$$R_{gg}(\tau) = e^{-a^2} e^{a^2 R_{zz}(\tau)} \tag{3}$$

It then follows immediately that the power density spectrum of g_t, $S_{gg}(f)$, is given by

$$S_{gg}(f) = \int_{-\infty}^{\infty} R_{gg}(\tau)\, e^{i2\pi f\tau}\, d\tau$$

$$= e^{-a^2} \sum_{n=0}^{\infty} \frac{a^{2n}}{n!} H_n(f) \tag{4}$$

where

$$\left.\begin{aligned} H_o(f) &= \delta(f) \\ H_1(f) &= \int_{-\infty}^{\infty} R_{zz}(\tau)\, e^{i2\pi f\tau}\, d\tau = S_{zz}(f) \\ H_n(f) &= \int_{-\infty}^{\infty} H_{n-1}(x)\, H_1(f-x)\, dx \end{aligned}\right\} \tag{5}$$

We shall return to a discussion of this spectrum.

To review Wiener's theory and for purposes of comparison, we shall compute the autocorrelation function for

$$c_t = \cos(az_t) \tag{6}$$

without orthogonalizing the series of polynomial functionals, as Wiener did. We could just as easily do this for g_t of Eq. 2, but it is interesting to note the result for c_t. Employing a double angle formula, the cosine power series, and the binomial series, we obtain

$$R_{cc}(\tau) = E[\cos(az_{t+\tau}) \cos(az_t)]$$

$$= E\left[\frac{1}{2} \cos a(z_{t+\tau}-z_t) + \frac{1}{2} \cos a(z_{t+\tau}+z_t)\right]$$

$$= E\left[\frac{1}{2} \sum_{n=0}^{\infty} \frac{(-a^2)^n}{(2n)!} \left\{(z_{t+\tau}-z_t)^{2n}+(z_{t+\tau}+z_t)^{2n}\right\}\right]$$

$$= E\left[\sum_{n=0}^{\infty} \frac{(-a^2)^n}{(2n)!} \frac{1}{2} \sum_{m=0}^{2n} \binom{2n}{m} \left\{z_{t+\tau}^m(-z_t)^{2n-m}+z_{t+\tau}^m z_t^{2n-m}\right\}\right]$$

$$= E\left[\sum_{n=0}^{\infty} \frac{(-a^2)^n}{(2n)!} \sum_{p=0}^{n} \binom{2n}{2p} z_{t+\tau}^{2p} z_t^{2n-2p}\right] \tag{7}$$

in which 2p is substituted for m in the last step because odd values of m give only zero terms. Next, we interchange the order of summation and averaging. From the facts that the average of a product of Gaussian variables is equal to the sum over all ways of pairing variables of the products of averages of these pairs, and that we can use Eq. 1, we find

$$E\left[z_{t+\tau}^{2p} z_t^{2n-2p}\right] = \sum_{\substack{\text{even } j \\ \leq \min[2p,\, 2n-2p]}} \binom{2p}{j}\binom{2n-2p}{j} j!\, E^j[z_{t+\tau} z_t]$$

$$\cdot\ (2p-j-1)(2p-j-3)\ldots(1)\ E^{(2p-j)/2}\left[z_{t+\tau}^2\right]$$

$$\cdot\ (2n-2p-j-1)(2n-2p-j-3)\ldots(1)\ E^{(2n-2p-j)/2}\left[z_t^2\right]$$

$$= \sum_{k=0}^{\min[p,\, n-p]} \frac{(2p)!\,(2n-2p)!}{(2k)!\,(p-k)!\,(n-p-k)!\,2^{n-2k}} R_{zz}^{2k}(\tau) \tag{8}$$

in which j is replaced by $2k$ because only terms with even j contribute to the sum. Combining Eqs. 7 and 8, we obtain

$$R_{cc}(\tau) = \sum_{n=0}^{\infty} \sum_{p=0}^{n} \sum_{k=0}^{\min[p,\, n-p]} \frac{(-a^2)^n}{(2k)!\,(p-k)!\,(n-p-k)!\,2^{n-2k}} R_{zz}^{2k}(\tau) \tag{9}$$

If we let $\mu = p - k$, $\nu = n - 2k$, and sum in a different order, we obtain

$$R_{cc}(\tau) = \sum_{k=0}^{\infty} \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\nu} \frac{(-a^2)^{2k}\left(-\frac{a^2}{2}\right)^{\nu}}{(2k)!\,\mu!\,(\nu-\mu)!} R_{zz}^{2k}(\tau)$$

$$= \sum_{k=0}^{\infty} \frac{\left[a^2 R_{zz}(\tau)\right]^{2k}}{(2k)!} \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\nu} \frac{\left(-\frac{a^2}{2}\right)^{\nu}}{\mu!\,(\nu-\mu)!} \tag{10}$$

The first sum in Eq. 10 is a hyperbolic cosine. Substitution of λ for $\nu - \mu$ converts the double sum into two identical, independent exponential series, and we obtain

$$R_{cc}(\tau) = e^{-a^2} \cosh\left[a^2 R_{zz}(\tau)\right] \tag{11}$$

If we consider the signal,

$$s_t = \sin(a z_t) \tag{12}$$

a similar computation yields the autocorrelation function

$$R_{ss}(\tau) = e^{-a^2} \sinh\left[a^2 R_{zz}(\tau)\right] \tag{13}$$

We notice that the crosscorrelation function vanishes:

$$R_{cs}(\tau) = 0 \tag{14}$$

For each signal, g_t, c_t, or s_t, it follows from the preceding discussion that the corresponding power density spectrum can be expressed as a term-by-term transform of the series for the correlation function, in a form very much like that of Eqs. 4 and 5. Hereafter, we shall use the spectrum of Eq. 4 to represent the spectrum of an FM signal. Notice that for large arguments (values of x) both cosh x and sinh x approach $\frac{1}{2} e^x$. This fact ensures that the spectra in the three cases will be similar when a is sufficiently large. [We shall not be more precise about this matter.]

3. Spectrum Approximation and Errors

Observe that the FM signal power is equal to unity:

$$\int_{-\infty}^{\infty} S_{gg}(f)\, df = R_{gg}(0) = e^{-a^2} e^{a^2 R_{zz}(0)} = 1 \tag{15}$$

Also, since

$$\int_{-\infty}^{\infty} H_n(f)\, df = R_{zz}^n(0) = 1, \qquad \text{for all } n \tag{16}$$

the total power omitted by leaving out the terms in Eq. 4 for which n belongs to some index set E, is given by

$$\Delta = e^{-a^2} \sum_{n \in E} \frac{a^{2n}}{n!} \tag{17}$$

Because the total FM signal power is unity, Δ also represents the fractional power associated with the index set E. In the time domain, Δ can be interpreted as the actual and fractional mean-square error in the FM signal if the power density spectrum is shaped by ideal filtering so that its shape is that of the sum of the remaining terms. Hence, an interesting problem is to find the most important terms in the series of Eq. 4, and to bound the error, Δ, when all less important terms are left out.

If x is assumed to be positive, then all the terms of the series

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{18}$$

are positive. The difference between two adjacent terms is readily computed. We have

$$\frac{x^n}{n!} - \frac{x^{n-1}}{(n-1)!} = \left(\frac{x}{n} - 1\right) \frac{x^{n-1}}{(n-1)!} \tag{19}$$

Now we can see that the terms grow in magnitude with increasing n until a maximum is reached, and then decrease toward zero as n continues to increase indefinitely. The largest term is the one for which

$$n = N(x) \tag{20}$$

where N(x) is the integral part of x. (Actually, if x is an integer, we have a pair of largest terms, but it will not be necessary to distinguish this special case in the following discussion.) Hereafter, we shall assume that $x \geq 2$. Notice that this implies that $N(x) \geq 2$. [We shall also use $x < \frac{3}{2} N(x)$ and $x < 2N(x)$.] In any event, it should be clear that the important terms of the sum in Eq. 18 are grouped around N(x).

First, we calculate a loose upper bound on the tail of the series in Eq. 18 above the important terms. Choose a number u that satisfies the inequality

$$u > \frac{1}{x-1} \tag{21}$$

It then follows that

$$N[(1+u)N(x)] - N(x) + 1 \geq (1+u)\, N(x) - 1 - N(x) + 1 = uN(x) > u\,\frac{2x}{3} \tag{22}$$

Since terms decrease with increasing n when $n > N(x)$, all terms in a partial sum beyond N(x) will exceed the last term. In particular, if r is the value of the $N[(1+u)N(x)]$ term, then

$$e^x > \sum_{n=N(x)}^{N[(1+u)N(x)]} \frac{x^n}{n!} \geq ru\,\frac{2x}{3} \tag{23}$$

where we have underbounded the number of terms in the partial sum by using inequality 22. Notice that the $N[(1+u)N(x)] + 1$ term is $r\,\frac{x}{N[(1+u)N(x)]+1}$, which is bounded by $r\,\frac{x}{(1+u)\,N(x)}$, the next term is bounded by $r\left[\frac{x}{(1+u)\,N(x)}\right]^2$, and so on. We can bound the sum over the tail beyond $N[(1+u)N(x)]$ by

$$\sum_{n=N[(1+u)N(x)]+1}^{\infty} \frac{x^n}{n!} < \sum_{m=1}^{\infty} r\left[\frac{x}{(1+u)N(x)}\right]^m = \frac{r\,\frac{x}{(1+u)\,N(x)}}{1 - \frac{x}{(1+u)\,N(x)}}$$

$$= \frac{r\,\frac{x}{N(x)}}{1 + u - \frac{x}{N(x)}} < \frac{2r}{u - \left[\frac{x}{N(x)} - 1\right]} \tag{24}$$

We have

$$\frac{x}{N(x)} - 1 < \frac{x}{x-1} - 1 = \frac{1}{x-1} \tag{25}$$

This explains why u was chosen to obey inequality 21. By substituting inequality 25 and r from inequality 23 in inequality 24, we can rewrite the bound as

$$\sum_{n>(1+u)x} \frac{x^n}{n!} < \frac{3e^x}{u\left(u-\frac{1}{x-1}\right)x} \tag{26}$$

Similar manipulations yield a bound on the tail of the series in Eq. 18 below the important terms. Choose a number v that satisfies the inequality

$$0 < v \leqslant 1 \tag{27}$$

This time, we underbound the number of important terms below $N(x)$ as follows:

$$N(x) - N[(1-v)N(x)] + 1 \geqslant N(x) - (1-v)\,N(x) + 1 = vN(x) + 1 > vx - v + 1 \geqslant vx \tag{28}$$

If s is the value of the $N[(1-v)N(x)]$ term, then,

$$e^x > \sum_{n=N[(1-v)N(x)]}^{N(x)} \frac{x^n}{n!} \geqslant svx \tag{29}$$

The $N[(1-v)N(x)] - 1$ term is $s\,\frac{N[(1-v)N(x)]}{x}$, which is bounded above by $s(1-v)$, the preceding term is bounded by $s(1-v)^2$, and so on. Thus,

$$\sum_{n=0}^{N[(1-v)N(x)]} \frac{x^n}{n!} < \sum_{m=0}^{\infty} s(1-v)^m = \frac{s}{1-(1-v)} = \frac{s}{v} \tag{30}$$

Combining Eqs. 28, 29, and 30, we obtain

$$\sum_{0\leqslant n\leqslant(1-v)(x-1)} \frac{x^n}{n!} < \frac{e^x}{v^2x} \tag{31}$$

What do Eqs. 26 and 31 mean? They imply that the power density spectrum of an FM signal (Eq. 4) with $a^2 \geqslant 2$, can be approximated by the partial sum

$$S_{gg}(f) \approx e^{-a^2} \sum_{(1-v)(a^2-1)<n\leqslant(1+u)a^2} \frac{a^{2n}}{n!} H_n(f) \tag{32}$$

with an error (from Eq. 17),

$$\Delta < \frac{3}{u\left(u-\frac{1}{a^2-1}\right)a^2} + \frac{1}{v^2a^2} \tag{33}$$

Equation 33 shows, among other things, that the terms that contribute significantly to the sum tend to cluster around $n \approx a^2$ in a narrowing band (expressed as a fraction of a^2, not in absolute terms) as a increases. In fact, if $H_n(f)$ does not change too rapidly with n when n is large, we shall have a good approximation for sufficiently large a:

$$S_{gg}(f) \approx H_{N(a^2)}(f) \tag{34}$$

This interesting behavior suggests that we take a look at the general behavior of $H_n(f)$ for large n to determine what the limiting spectrum is like.

4. Limiting Spectrum and Bandwidth for Wideband FM

Recall the definition of $H_n(f)$ from Eq. 5. Remember that power density spectra are non-negative, so that $H_n(f) \geq 0$, for all n, f. From Eq. 16, the integral over any $H_n(f)$ is unity. This suggests interpreting $H_1(f)$ as the probability density function of a random variable with zero mean (symmetry of power density spectra about the origin). Assume that the variance, σ^2, is finite, with $\sigma^2 = \int_{-\infty}^{\infty} f^2 H_1(f)\, df$. $H_n(f)$ looks like the density function of the distribution of a sum of n independent random variables, each variable being distributed according to $H_1(f)$. Under these conditions, the Central Limit Theorem is valid (3). Therefore, the cumulative distribution of the sum approaches that of a Gaussian random variable with zero mean and variance $n\sigma^2$, as n increases. Under somewhat more restrictive conditions on $H_1(f)$, we can say that, for increasing n,

$$H_n(f) \to (2\pi n\sigma^2)^{-1/2}\, e^{-f^2/(2n\sigma^2)} \qquad \text{almost everywhere} \tag{35}$$

In general, the convergence will not be uniform. On the other hand, if we are mainly interested in the location of the major part of the power, the integrated power density spectrum is a natural expression with which to work, and we shall have convergence under less restrictive conditions, as we have already pointed out, for

$$\int_{-\infty}^{f} H_n(x)\, dx \to \int_{-\infty}^{f} (2\pi n\sigma^2)^{-1/2}\, e^{-x^2/(2n\sigma^2)}\, dx \tag{36}$$

In either case, the wideband FM spectrum tends to a Gaussian shape, a fact that is intuitively justified by the Gaussian amplitude density of the modulating message.

We return for a moment to the finite variance assumption, $\sigma^2 < \infty$. Remember that $H_1(f) = S_{zz}(f)$ is the power density spectrum of the normalized phase signal, z_t. Hence, we conclude that the finite variance assumption implies an essentially bandlimited phase signal spectrum, with tails that can be estimated by a Tchebychev bound in general, or by a less conservative bound in more restricted cases. Moreover, the power density spectrum of the message itself must be limited. The power density spectrum of the message, $\frac{1}{2\pi} z_t'$ (converting from radian to cyclic measure) is

$$S_{\frac{z'}{2\pi}\frac{z'}{2\pi}}(f) = f^2 H_1(f) \tag{37}$$

The finite variance assumption implies that, given any $\epsilon > 0$, there exists a finite f_ϵ with the property that

$$\int_{-\infty}^{-f_\epsilon} S_{\frac{z'}{2\pi}\frac{z'}{2\pi}}(f)\, df + \int_{f_\epsilon}^{\infty} S_{\frac{z'}{2\pi}\frac{z'}{2\pi}}(f)\, df = 2\int_{f_\epsilon}^{\infty} f^2 H_1(f)\, df \leq \epsilon \tag{38}$$

Thus, by choosing ϵ as small as desired and finding the smallest f_ϵ that satisfies inequality 38, we can say that the message bandwidth is essentially f_ϵ. Therefore, the finite variance assumption can be interpreted as a limitation on message bandwidth.

We note one final bandwidth relationship. The power of the unnormalized (cyclic measure) message, $\frac{a}{2\pi} z_t'$, is $a^2\sigma^2$. Since Gaussian waves seldom exceed approximately three times their rms value, we can say that the peaks of the message waveform almost always fall below $3a\sigma$. On the other hand, Eqs. 34 and 36 imply that, for large values of a, the FM signal power density spectrum will fall almost entirely within a band of frequencies extending from 0 up to approximately $3a\sigma$. Hence, in the limit for wide-band FM, the peak frequency deviation of the message does, indeed, control the signal bandwidth.

A. D. Hause

References

1. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press, Cambridge, Massachusetts, and John Wiley and Sons, Inc., New York, 1958), pp. 49-55.

2. Ibid., p. 54.

3. W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 1 (John Wiley and Sons, Inc., New York, 2d ed., 1957), p. 229.

G. AN OPTIMUM RECEIVED PULSE SHAPE FOR PULSE CODE MODULATION

In discussions of pulse code modulation (1) it is usually assumed that the receiver consists of a threshold detector (or level selector) that samples a received waveform every T_o seconds and decides in which of J prescribed voltage intervals the sample lies. The received waveform is assumed to consist in part of an infinite train of identical pulses with unit peak value and period T_o. These pulses have been modulated so that the peak voltage of each pulse can take on any one of J possible values. Statistically independent Gaussian noise is added to this modulated pulse train to form the received waveform. There exists a one-to-one correspondence between the J voltage intervals and the J possible pulse peak voltages.

Since the sampling is assumed to be periodic, it is possible (2) to choose a band-limited pulse in such a way that only the amplitude of one pulse (plus the Gaussian noise) is measured at any sampling instant. Each of the other pulses makes no contribution to the voltage at that instant. We thus have Gaussian noise interference, but no interpulse interference. Pulse code modulation is known (3) to be quite efficient for combating additive Gaussian noise.

In this report we shall relax the assumption that the sampling occurs strictly periodically and allow a jitter in the sampling times. However, the sampling times will be mutually independent.

We may contend with additive Gaussian noise in the standard manner (1). However, we must now cope with the fact that spurious contributions from many pulses will be superimposed upon the contribution of the desired information-bearing pulse. These spurious contributions will be called "interpulse interference."

We shall now define a criterion for deciding which of two received pulse shapes gives rise to more or less interpulse interference. Our next step will be to discuss how we can obtain an optimum received pulse shape.

1. Criterion for an Optimum Received Pulse Shape

If we are given two unmodulated individual output pulses, $r_1(t)$ and $r_2(t)$, whose peak values lie in the time interval $\left(-\frac{T_o}{2}, \frac{T_o}{2}\right)$, we shall say that $r_1(t)$ gives rise to less interpulse interference if $I[r_1(t)] < I[r_2(t)]$, where

$$I[r(t)] = \int_{-\infty}^{\infty} q(t)[r(t)-d(t)]^2 \, dt \tag{1}$$

with

$$d(t) = \begin{cases} 1 & \text{for all } t \text{ satisfying } p(t) > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$q(t) = \sum_{n=-\infty}^{\infty} p(t-nT_o)$$

and $p(t)$ is the probability density function for the sampling instant during the time interval $\left(-\frac{T_o}{2}, \frac{T_o}{2}\right)$.

We assume that the detector samples the received waveform once and only once during each time interval $\left(\left(m-\frac{1}{2}\right)T_o, \left(m+\frac{1}{2}\right)T_o\right)$, with $m = 0, \pm 1, \pm 2, \ldots$, and that the probability density function for the sampling instant within each interval is the same for all intervals.

We can justify formula 1 as a criterion for comparing our pulses because (a) the desired unmodulated pulse, d(t), has unit value at all instants when a sample might be taken [in order to retrieve any information voltage associated with the pulse whose peak value lies in the time interval $\left(-\frac{T_o}{2}, \frac{T_o}{2}\right)$]. The function d(t) has the value zero at all other possible sampling instants and thus cannot interfere with any other pulses. (b) The squared difference between r(t) and d(t) in the integrand of formula 1 is weighted by the factor q(t). This restricts the integration to possible sampling instants only.

Thus, if $I[r(t)]$ is small, then r(t) must be close to d(t) at all possible sampling instants. Hence, the receiver should be able to extract the voltage information from each pulse separately, since it must cope with only a small amount of interpulse interference.

2. The Minimization of $I[r(t)]$

By using the elementary calculus of variations and Fourier methods the writer has found that a necessary condition for $I[r(t)]$ to be minimum by varying r(t) is

$$\frac{1}{T_o} \sum_{n=-\infty}^{\infty} P\left(\frac{n}{T_o}\right) R\left(f - \frac{n}{T_o}\right) = P(f) \tag{2}$$

where T_o is the signaling period, and

$$R(f) = \int_{-\infty}^{\infty} r(t)\, e^{-j2\pi ft}\, dt; \quad P(f) = \int_{-\infty}^{\infty} p(t)\, e^{-j2\pi ft}\, dt$$

Equation 2 has the solution:

$$r(t) \text{ is } \begin{cases} 1 & \text{whenever } p(t) > 0 \\ \text{undefined} & \text{whenever } q(t) = 0 \\ 0 & \text{elsewhere} \end{cases}$$

Thus, at possible sampling times, r(t) = d(t).

If we now add the constraint that R(f) be zero outside the frequency interval (-W, W), then we replace Eq. 2 with

$$\frac{1}{T_o} \sum_{n=-\infty}^{\infty} P\left(\frac{n}{T_o}\right) R\left(f - \frac{n}{T_o}\right) = P(f) \qquad \text{for } f \text{ in } (-W, W) \tag{3}$$

The writer has obtained a proof of the existence and uniqueness of the solution of Eq. 3. This solution is

$$R(x_j) = \begin{cases} \dfrac{P(x_1)\, D_M^{(1,j)} + P(x_2)\, D_M^{(2,j)} + \ldots + P(x_M)\, D_M^{(M,j)}}{D_M} & \text{for } x_j \text{ in } (-W, W) \\ 0 & \text{elsewhere} \end{cases} \tag{4}$$

where D_M is the determinant

$$D_M = \left(\frac{1}{T_o}\right)^M \begin{vmatrix} P(0) & P\left(\frac{1}{T_o}\right) & P\left(\frac{2}{T_o}\right) & \cdots & P\left(\frac{M-1}{T_o}\right) \\ P\left(\frac{-1}{T_o}\right) & P(0) & P\left(\frac{1}{T_o}\right) & \cdots & P\left(\frac{M-2}{T_o}\right) \\ \vdots & \vdots & \vdots & & \vdots \\ P\left(\frac{-M+1}{T_o}\right) & P\left(\frac{-M+2}{T_o}\right) & P\left(\frac{-M+3}{T_o}\right) & \cdots & P(0) \end{vmatrix}$$

$D_M^{(i,j)}$ is the cofactor of the (i,j) element of D_M, and x_j is any prescribed value of f, say f_o, in the interval $(-W, W)$.

The set $\{x_i\}$, $i = 1, 2, \ldots, M$ is the set of ordered $1/T_o$ translates of x_j that lie within the band $(-W, W)$. That is,

(a) x_i is contained in $(-W, W)$ for $i = 1, 2, \ldots, M$;

(b) $x_{i+1} = x_i + 1/T_o$ for $i = 1, 2, \ldots, M - 1$;

(c) $x_j = f_o$ for some j, $j = 1, 2, \ldots, M$;

and we compute M as follows: Let d be the largest positive integer (or zero if there is none) that is such that $f_o + \frac{d}{T_o}$ is contained in the interval $(-W, W)$. Let g be the largest positive integer (or zero if there is none) that is such that $f_o - \frac{g}{T_o}$ is contained in the interval $(-W, W)$. Then $M = d + g + 1$.

3. Examples

Having defined the symbols in formula 4, let us now apply the formula to some specific examples.

EXAMPLE 1. Suppose that we are signaling at a rate that is at least as fast as the Nyquist rate (i.e., $1/T_o \geq 2W$), and that $(-W, W)$ is an open interval. Then, when we are computing the integer M, we find that $d = g = 0$. Hence $M = 1$. Thus formula 4 reduces to

$$R(x_1) = \begin{cases} T_o\, P(x_1) & -W < x_1 < W \\ 0 & \text{elsewhere} \end{cases}$$

Therefore, when we are signaling at a rate faster than the Nyquist rate, the best that we can do is to shape the received pulse spectrum so that it matches the spectrum of the probability density function, $p(t)$, as far out as we allow our pulse bandwidth to extend.

When we signal at a rate that is slower than the Nyquist rate, then $M > 1$, and more terms enter into formula 4. Thus there is a marked difference in the solution when $2WT_o$ is less than or greater than unity.

EXAMPLE 2. This example is intended to demonstrate what computations must be made when we apply formula 4. Let $2WT_o = 2$, and

$$p(t) = \begin{cases} \frac{1}{\delta}\left[1 + \cos 2\pi \frac{t}{\delta}\right] & -\frac{\delta}{2} \leq t \leq \frac{\delta}{2} \\ 0 & \text{elsewhere} \end{cases}$$

where $0 < \delta \leq T_o$. Hence,

$$P(f) = \frac{\sin \pi\delta f}{\pi\delta f} + \frac{1}{2} \frac{\sin \pi\delta\left(f-\frac{1}{\delta}\right)}{\pi\delta\left(f-\frac{1}{\delta}\right)} + \frac{1}{2} \frac{\sin \pi\delta\left(f+\frac{1}{\delta}\right)}{\pi\delta\left(f+\frac{1}{\delta}\right)}$$

If $R(f)$ is to be bandlimited to the open interval $(-W, W)$, then we compute M for different values of f_o as follows:

For $-W < f_o < 0$, $d = 1$, $g = 0$; hence $M = 2$.

For $f_o = 0$, $d = 0$, $g = 0$; hence $M = 1$.

For $0 < f_o < W$, $d = 0$, $g = 1$; hence $M = 2$.

Thus

$$D_M = \begin{cases} D_2 = \frac{1}{T_o^2} \begin{vmatrix} P(0) & P\left(\frac{1}{T_o}\right) \\ P\left(\frac{-1}{T_o}\right) & P(0) \end{vmatrix} = \frac{1}{T_o^2}\left[1 - P\left(\frac{-1}{T_o}\right) P\left(\frac{1}{T_o}\right)\right] & \text{if } M = 2 \\ D_1 = \frac{P(0)}{T_o} = \frac{1}{T_o} & \text{if } M = 1 \end{cases}$$

Direct application of formula 4 now yields

$$R(f) = \begin{cases} T_o \dfrac{P(f) - P\left(f + \frac{1}{T_o}\right) P\left(\frac{1}{T_o}\right)}{1 - P\left(\frac{-1}{T_o}\right) P\left(\frac{1}{T_o}\right)} & -W < f < 0 \\ T_o & f = 0 \\ T_o \dfrac{P(f) - P\left(f - \frac{1}{T_o}\right) P\left(\frac{-1}{T_o}\right)}{1 - P\left(\frac{-1}{T_o}\right) P\left(\frac{1}{T_o}\right)} & 0 < f < W \end{cases}$$

D. W. Tufts

References

1. B. M. Oliver, J. R. Pierce, and C. E. Shannon, The philosophy of PCM, Proc. IRE 36, 1324-1331 (1948).

2. Ibid., p. 1330 (Appendix I).

3. Ibid., pp. 1323-1325 (Section II).

H. CONSERVATION OF BANDWIDTH IN NONLINEAR OPERATIONS

When a bandlimited signal is filtered nonlinearly, the width of the spectrum of the resulting signal, in general, has no bounds. Expansion of the spectrum is characteristic of nonlinear operations, and hinders their use in communication channels whose bandwidth must usually be constrained.

However, filtering does not, loosely speaking, add any new degrees of freedom to a signal. For example, when a bandlimited signal is filtered by a nonlinear operator without memory that has an inverse, a sampling of the output signal at the Nyquist rate obviously specifies the situation completely even though the output is not bandlimited; the inverse operation can be performed on the samples instead of on the signal; this yields, in effect, a Nyquist sampling of the original signal, which is completely specified by the samples. It seems plausible, therefore, that some of the spectrum added by invertible nonlinear filtering is redundant, and hence that it could be discarded without affecting the possibility of recovering the original signal.

We shall discuss a situation (Fig. VIII-1) in which a function of time, x(t), whose bandwidth is $2\omega_0$ centered about zero, and whose energy is finite, is filtered by any nonlinear operator without memory, N, that has an inverse and has both a finite maximum slope and a minimum slope that is greater than zero. The spectrum of y(t) is then

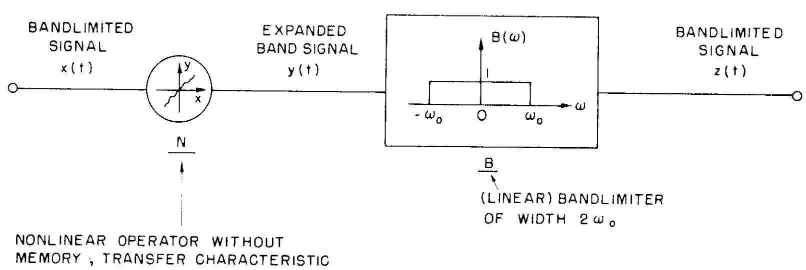

Fig. VIII-1. Bandlimiter following a nonlinear operator without memory.

narrowed down to the original passband by means of a bandlimiter, $\underline{B}$. It will be shown that the cascade combination of $\underline{B}$ following $\underline{N}$ has an inverse. This is equivalent to our stated hypothesis concerning recoverability.

a. Notation

Lower-case letters (for example, x) are used to denote functions of time. The value assumed by x at time t is x(t). An operational relation $\underline{N}$ (underlined, capital letter) between two functions of time, say x and y, is indicated by

$$y = \underline{N}(x); \qquad y(t) = F(x(\tau)) \tag{1}$$

in which the second equation relates the specific value of y at time t to that of x at time τ. Spectra are denoted by capital letters; for example, $X(\omega)$ is the spectrum of x(t).

$\underline{I}$ denotes the identity operator that maps each function into itself. We shall have occasion to use the following symbols:

$\underline{A} + \underline{B}$, denoting the sum of operators $\underline{A}$ and $\underline{B}$.

$\underline{A} * \underline{B}$, denoting the cascade combination of operator $\underline{A}$ following $\underline{B}$.

$\underline{A}^{-1}$, denoting the inverse of the operator $\underline{A}$, with the property that $\underline{A} * \underline{A}^{-1} = \underline{I}$.

b. Outline of the Method of Inversion

It is required to show that $\underline{B} * \underline{N}$ has an inverse and to find it. The essential difficulty is that the bandlimiter $\underline{B}$ has no inverse by itself, since its spectrum is identically zero outside the passband. It would have an inverse (simply the identity operator) if the

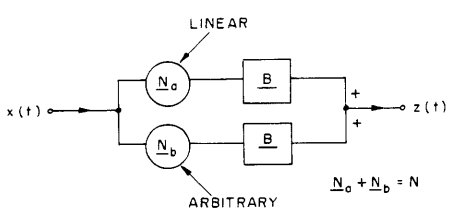

Fig. VIII-2. An equivalent split form for $\underline{B} * \underline{N}$.

input to it lay entirely in the passband, but y is not such a signal. To circumvent this difficulty, we resort to the following device: $\underline{N}$ is split into the sum of two parts (Fig. VIII-2)

$$\underline{N} = \underline{N}_a + \underline{N}_b \tag{2}$$

of which $\underline{N}_a$ is linear (hence $\underline{N}_a$ is a pure gain). Thus

$$\underline{B} * \underline{N} = \underline{B} * (\underline{N}_a + \underline{N}_b) \tag{3}$$

and, since $\underline{B}$ is linear and hence distributive,

$$\underline{B} * \underline{N} = \underline{B} * \underline{N}_a + \underline{B} * \underline{N}_b \tag{4}$$

We can now find the inverse of the part $\underline{B} * \underline{N}_a$ because the spectrum of the output of $\underline{N}_a$ lies within the passband of $\underline{B}$. However, $\underline{B} * \underline{N}_b$ suffers from the original difficulty. Nevertheless, we can find the inverse of $\underline{B} * \underline{N}$ by virtue of the fact that the inverse of a sum of two operators, one of which is linear, can be expressed as an iteration involving the inverse of only the linear operator (1). For, consider the sum operator

$$\underline{H} = \underline{H}_a + \underline{H}_b \tag{5}$$

in which $\underline{H}_a$ is linear and has a known inverse, $\underline{H}_a^{-1}$, whereas $\underline{H}_b$ is arbitrary. Consider also the following iteration formula, which involves no inverse other than that of $\underline{H}_a$:

$$\begin{aligned} \underline{K}_1 &= \underline{H}_a^{-1} \\ \underline{K}_n &= \underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-1} \qquad n = 2, 3, \ldots \end{aligned} \tag{6}$$

If $\underline{K}_n$ converges to $\underline{K}$, which itself satisfies the iteration formula, Eq. 7, then $\underline{K}$ must be the inverse of $\underline{H}$, because

$$\underline{K} = \underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K} \tag{7}$$

Multiplying by $\underline{H}_a$ and rearranging, we obtain

$$\underline{H}_a * \underline{K} + \underline{H}_b * \underline{K} = \underline{I}$$

or, using Eq. 5, we have

$$\underline{H} * \underline{K} = I \tag{8}$$

which shows that $\underline{K}$ is indeed the inverse of $\underline{H}$.

It will be shown that the iteration converges as required whenever the operator $\underline{H}_a^{-1} * \underline{H}_b$ is a contraction, a condition that is fulfilled when $\underline{H}$, whose inverse is sought, is not too different from $\underline{H}_a$, whose inverse is known.

These concepts will now be made precise in a series of lemmas. First, the convergence of the iteration will be established for arbitrary operators satisfying certain conditions. Then we shall identify $\underline{B} * \underline{N}$ with $\underline{H}$ and, after studying properties of $\underline{B}$ and $\underline{N}$, show that $\underline{B} * \underline{N}$ may be split in such a way as to satisfy these conditions.

c. The Space L_2

Consider (2) a collection, L_2, of all functions x(t) that are defined on the infinite interval, $-\infty < t < \infty$, with the property that the following integral, denoted $\|x\|$, exists and is finite:

$$\|x\| = \int_{-\infty}^{\infty} |x(t)|^2 \, dt < \infty \tag{9}$$

With each pair x, y, in L_2 a distance, $\|x-y\|$, is associated. By definition, x = y whenever $\|x-y\| = 0$.

We shall be concerned with operators such as $\underline{H}$ that map L_2 into itself. An operator $\underline{H}$ is called (3) a contraction in L_2 if it satisfies the condition

$$\|\underline{H}(x) - \underline{H}(y)\| \leq a\|x-y\| \qquad 0 \leq a < 1 \tag{10}$$

for all pairs x, y, of members of L_2.

d. Lemma 1: Convergence of the Iteration for the Inverse

Any operator $\underline{H}$ that maps the space L_2 into itself has an inverse that may be computed by an iteration formula, provided that the following conditions are satisfied:

(i) $\underline{H}$ may be split into a sum of two operators. Thus

$$\underline{H} = \underline{H}_a + \underline{H}_b \tag{11}$$

in which $\underline{H}_a$ is linear and has an inverse, $\underline{H}_a^{-1}$, and $\underline{H}(0) = \underline{H}_a(0) = \underline{H}_b(0) = 0$. (The condition $\underline{H}(0) = 0$ is not at all necessary, but it simplifies the iteration.)

(ii) The cascade operator $\underline{H}_a^{-1} * \underline{H}_b$ is a contraction.

The iteration formula has the form

$$\begin{aligned} &\underline{K}_o = 0 \\ &\underline{K}_1 = \underline{H}_a^{-1} \\ &\underline{K}_n = \underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-1} \qquad n = 2, 3, \ldots \end{aligned} \tag{12}$$

Proof: We shall show that the sequence $\{\underline{K}_n\}$ is a Cauchy sequence.

Using formula 12, we obtain

$$\begin{aligned} \underline{K}_n - \underline{K}_{n-1} &= \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-1}\right) - \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-2}\right) \\ &= \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-2} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-1} \qquad n = 2, 3, \ldots \end{aligned} \tag{13}$$

Since $\underline{H}_a^{-1} * \underline{H}_b$ is a contraction, we can apply Eq. 10 to Eq. 13 to obtain, for all z in L_2,

$$\begin{aligned} \|(\underline{K}_n - \underline{K}_{n-1})(z)\| &\leq a\|(\underline{K}_{n-1} - \underline{K}_{n-2})(z)\| \\ &\leq a^{n-1}\|(\underline{K}_1 - \underline{K}_o)(z)\| \\ &= a^{n-1}\|\underline{H}_a^{-1}(z)\| \qquad 0 \leq a \leq 1, \; n = 2, 3, \ldots \end{aligned} \tag{14}$$

in which the second inequality is obtained by applying the first inequality (n-1) times, and making use of the fact that $\underline{H}_a^{-1} * \underline{H}_b(0) = 0$ by hypothesis.

Suppose, next, that m > n, with n = 1, 2, ... Then, by the triangle inequality,

$$\|(\underline{K}_m - \underline{K}_n)(z)\| = \left\| \left([\underline{K}_m - \underline{K}_{m-1}] + [\underline{K}_{m-1} - \underline{K}_{m-2}] + \dots [\underline{K}_{n+1} - \underline{K}_n]\right)(z)\right\|$$

$$\leq \|(\underline{K}_m - \underline{K}_{m-1})(z)\| + \dots + \|(\underline{K}_{n+1} - \underline{K}_n)(z)\| \tag{15}$$

Then, by applying Eq. 14 to Eq. 15, we obtain

$$\|(\underline{K}_m - \underline{K}_n)(z)\| \leq (\alpha^{m-1} + \dots \alpha^n) \|\underline{H}_a^{-1}(z)\|$$

$$= \alpha^n(\alpha^{m-n-1} + \dots + 1) \|\underline{H}_a^{-1}(z)\|$$

$$= \frac{\alpha^n(1 - \alpha^{m-n})}{1 - \alpha} \|\underline{H}_a^{-1}(z)\|$$

$$\leq \alpha^n \frac{\|\underline{H}_a^{-1}(z)\|}{(1-\alpha)} \qquad m > n, \ n = 1, 2, \dots \tag{16}$$

where we have summed the geometric progression. Since $\underline{H}_a^{-1}(z)$ belongs to L_2, $\|\underline{H}_a^{-1}(z)\|$ is finite, and the right-hand side can be made arbitrarily small by choosing n to be sufficiently large. Hence $\underline{K}_n$ is a Cauchy sequence. The Riesz-Fisher theorem therefore ensures the existence of a unique x with the property that

$$\underline{K}_n(z) \to x \qquad \text{as } n \to \infty \tag{17}$$

The operator $\underline{K}$ is now defined by

$$\underline{K}(z) = x \tag{18}$$

$\underline{K}$ satisfies the iteration formula, Eq. 7, because

$$\left\|\underline{K}(z) - \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}\right)(z)\right\|$$

$$= \left\|\underline{K}(z) - \underline{K}_n(z) + \underline{K}_n(z) - \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}\right)(z)\right\|$$

$$\leq \|\underline{K}(z) - \underline{K}_n(z)\| + \left\|\underline{K}_n(z) - \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}\right)(z)\right\| \tag{19}$$

$$= \|x - \underline{K}_n(z)\| + \left\|\left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}_{n-1}\right)(z) - \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}\right)(z)\right\| \tag{20}$$

$$\leq \|x - \underline{K}_n(z)\| + \|\underline{K}_{n-1}(z) - x\| \tag{21}$$

The triangle inequality was used to get Eq. 19, and Eqs. 18 and 12 were used to get Eq. 20 from Eq. 19. The contraction condition was used to get Eq. 21.

Since $\underline{K}_n(z) \to x$, by Eq. 17, the right-hand side of Eq. 21 can be made arbitrarily small, so that

$$\underline{K}(z) = \left(\underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K}\right)(z) \tag{22}$$

Since this is true for all z in L_2, $\underline{K}$ satisfies the iteration formula

$$\underline{K} = \underline{H}_a^{-1} - \underline{H}_a^{-1} * \underline{H}_b * \underline{K} \tag{23}$$

Rearranging terms and multiplying by $\underline{H}_a^{-1} *$, we have

$$(\underline{H}_a + \underline{H}_b) * \underline{K} = \underline{H} * \underline{K} = \underline{I} \tag{24}$$

so that $\underline{K}$ is the inverse of $\underline{H}$.

e. Bandlimiting

The bandlimiting operator in L_2, $\underline{B}$, is defined by

$$\underline{B}(y) = z; \qquad z(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} y(t-\tau) \frac{\sin \tau}{\tau} \, d\tau \tag{25}$$

Since both $x(t-\tau)$ and $(\sin \tau)/\tau$ are in L_2, their product is in L_1, and the integral in Eq. 25 exists in the Lebesgue sense (4). The Fourier transforms also exist in L_2 and are given by Wiener (5) as

$$Y(\omega) = \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} y(t) \, e^{-j\omega t} \, dt \tag{26}$$

$$B(\omega) = \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} \frac{\sin t}{t} \, e^{-j\omega t} \, dt$$

$$= 0, \ |\omega| > \omega_o; \qquad 1, \ |\omega| < \omega_o \tag{27}$$

$$Z(\omega) = \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} z(t) \, e^{-j\omega t} \, dt \tag{28}$$

Moreover, $Y(\omega)\, B(\omega)$ must be in L_2 so that

$$Z(\omega) = Y(\omega)\, B(\omega) \tag{29}$$

Lemma 2: If $\underline{B}$ denotes the bandlimiting operator in L_2 (Eq. 25), then for all y in L_2

$$\|\underline{B}(y)\| \leq \|y\| \tag{30}$$

Proof: Parseval's theorem is applicable. By definition (Eq. 9),

$$\|\underline{B}(y)\| = \int_{-\infty}^{\infty} |z(t)|^2 \, dt = \int_{-\infty}^{\infty} |Z(\omega)|^2 \, d\omega = \int_{-\omega_o}^{\omega_o} |Y(\omega)|^2 \, d\omega$$

$$\leq \int_{-\infty}^{\infty} |Y(\omega)|^2 \, d\omega = \|y\| \tag{31}$$

f. Lemma 3: Splitting the Nonlinear Operator

A nonlinear operator without memory $\underline{N}$ defined on L_2 by

$$\underline{N}(x) = y; \qquad \underline{N}(x(t)) = y(t) \tag{32}$$

which satisfies the dual Lipschitz conditions

$$\beta\|x_2 - x_1\| \leq \|\underline{N}(x_2) - \underline{N}(x_1)\| \leq \gamma\|x_2 - x_1\| \qquad 0 < \beta \leq \gamma < \infty \tag{33}$$

for all x_1 and x_2 in L_2, and which is monotonic, can be split into the sum of two operators:

$$\underline{N} = \underline{N}_a + \underline{N}_b \tag{34}$$

of which $\underline{N}_a$ is linear, and has the property that $\underline{N}_a^{-1} * \underline{N}_b$ is a contraction. For simplicity, we assume that $\underline{N}(0) = \underline{N}_a(0) = \underline{N}_b(0) = 0$.

Proof: Let $\underline{N}_a$ be the linear operator defined by

$$\underline{N}_a(x) = y; \qquad y(t) = \frac{1}{2} (\gamma+\beta) \, x(t) \tag{35}$$

and let $\underline{N}_b$ be defined by

$$\underline{N}_b(x) = y; \qquad y(t) = \underline{N}(x(t)) - \frac{1}{2} (\gamma+\beta) \, x(t) \tag{36}$$

so that Eq. 34 is satisfied by construction. From Eqs. 35 and 36 it follows that

$$\underline{N}_a^{-1} * \underline{N}_b(x) = y; \qquad \underline{N}_a^{-1} * \underline{N}_b(x(t)) = \frac{2}{\gamma + \beta} \left[\underline{N}(x(t)) - \frac{1}{2} (\gamma+\beta) \, x(t)\right]$$

$$= \frac{2}{\gamma + \beta} \underline{N}(x(t)) - x(t) \tag{37}$$

The operator $\underline{N}_a^{-1} * \underline{N}_b$ will be shown to be a contraction. Let $x_1(t)$ and $x_2(t)$ be any two real numbers, and suppose that $x_2(t) > x_1(t)$. Using Eq. 37, we obtain

$$\underline{N}_a^{-1} * \underline{N}_b(x_2(t)) - \underline{N}_a^{-1} * \underline{N}_b(x_1(t)) = \frac{2}{\gamma + \beta} \left[\underline{N}(x_2(t)) - \underline{N}(x_1(t))\right] - \left[x_2(t) - x_1(t)\right] \tag{38}$$

The right-hand side of Eq. 38 is the difference of two expressions, each of which is positive because $x_2(t) > x_1(t)$, and $\underline{N}$ is assumed to be monotonic. Suppose, first, that the difference is positive or zero. Then, by using the upper Lipschitz condition of Eq. 33, we obtain

$$\underline{N}_a^{-1} * \underline{N}_b(x_2(t)) - \underline{N}_a^{-1} * \underline{N}_b(x_1(t)) \leq \frac{2}{\gamma+\beta}\,\gamma[x_2(t) - x_1(t)] - [x_2(t) - x_1(t)]$$

$$= \frac{\gamma-\beta}{\gamma+\beta}\,[x_2(t) - x_1(t)] \qquad (39)$$

When the difference in Eq. 38 is negative the lower Lipschitz condition may be used instead to give Eq. 39 with the inequality and sign reversed. Hence

$$\left|\underline{N}_a^{-1} * \underline{N}_b(x_2(t)) - \underline{N}_a^{-1} * \underline{N}_b(x_1(t))\right| \leq a\left|x_2(t) - x_1(t)\right| \qquad 0 \leq a = \frac{\gamma-\beta}{\gamma+\beta} < 1 \qquad (40)$$

which implies that

$$\left\|\underline{N}_a^{-1} * \underline{N}_b(x_2) - \underline{N}_a^{-1} * \underline{N}_b(x_1)\right\| \leq a\left\|x_2 - x_1\right\| \qquad 0 \leq a < 1 \qquad (41)$$

and therefore $\underline{N}_a^{-1} * \underline{N}_b$ is a contraction.

We can now prove our principal hypothesis.

g. Theorem on Bandwidth Redundancy

The operator $\underline{B} * \underline{N}$, consisting of a bandlimiter, $\underline{B}$, following a monotonic nonlinear operator without memory, $\underline{N}$, restricted to inputs that are in L_2 (have finite energy), and whose spectrum is nonzero only in the passband of $\underline{B}$, has an inverse in L_2 that can be found as the limit in the mean of an iteration, provided that the following conditions are satisfied:

(i) $\underline{N}$ satisfies the dual Lipschitz conditions

$$\beta\left\|x_2(t) - x_1(t)\right\| \leq \left\|N(x_2(t)) - \underline{N}(x_1(t))\right\| \leq \gamma\left\|x_2(t) - x_1(t)\right\| \qquad 0 < \beta \leq \gamma < \infty \qquad (42)$$

for all real $x_1(t)$, $x_2(t)$.

(ii) For simplicity only, the passband of $\underline{B}$ is confined to $-\omega_o \leq \omega \leq \omega_o$, and $\underline{N}(0) = 0$.

The iteration formula is

$$(\underline{B}*\underline{N})^{-1} = \underset{n\to\infty}{\text{l.i.m.}}\ \underline{K}_n$$

$$\underline{K}_1 = \underline{N}_a^{-1} \qquad (43)$$

$$\underline{K}_n = \underline{N}_a^{-1} - \underline{B} * \underline{N}_b * \underline{K}_{n-1} \qquad n = 2, 3, \ldots$$

in which

$$\underline{N}_a(x) = y; \qquad y(t) = \frac{1}{2}\,(\beta+\gamma)\,x(t) \tag{44}$$

and

$$\underline{N}_b = \underline{N} - \underline{N}_a \tag{45}$$

Proof: Since $\underline{B}$ is linear, we have

$$\underline{B} * \underline{N} = B * (\underline{N}_a + \underline{N}_b) = \underline{B} * \underline{N}_a + \underline{B} * \underline{N}_b \tag{46}$$

But because $\underline{N}_a$ is linear and $\underline{B} * \underline{N}_a$ is restricted to inputs within the passband,

$$\underline{B} * \underline{N}_a = \underline{N}_a \tag{47}$$

so that

$$(\underline{B} * \underline{N}_a)^{-1} = \underline{N}_a^{-1} \tag{48}$$

Moreover, $(\underline{B} * \underline{N}_a)^{-1} * (\underline{B} * \underline{N}_b)$ is a contraction, because

$$(\underline{B} * \underline{N}_a)^{-1} * (\underline{B} * \underline{N}_b) = \underline{N}_a^{-1} * (\underline{B} * \underline{N}_b) \tag{49}$$

$$= \underline{B} * \underline{N}_a^{-1} * \underline{N}_b \tag{50}$$

where the sequence of $\underline{B}$ and $\underline{N}_a$ can be interchanged, since both are linear. Now, by using Lemma 2, first, and then Lemma 3, we obtain

$$\left\| \underline{B} * \underline{N}_a^{-1} * \underline{N}_b(x_2) - \underline{B} * \underline{N}_a^{-1} * \underline{N}_b(x_1) \right\|$$

$$\leq \left\| \underline{N}_a^{-1} * \underline{N}_b(x_2) - \underline{N}_a^{-1} * \underline{N}_b(x_1) \right\| \leq \alpha \left\| x_2 - x_1 \right\| \qquad 0 \leq \alpha = \frac{\gamma - \beta}{\gamma + \beta} < 1 \tag{51}$$

Hence $(\underline{B} * \underline{N}_a)^{-1} * (\underline{B} * \underline{N}_b)$ is a contraction. Since $(B * \underline{N}_a)^{-1}$ exists and is linear, Lemma 1 is applicable and the theorem is proved. (Note that $\underline{B} * \underline{N}_a^{-1}$ and $\underline{B} * \underline{N}_b^{-1}$ both map bandlimited signals into bandlimited signals.)

h. Extensions of the Theorem

With slight modifications the theorem is valid when $\underline{N}(0) \neq 0$ and for arbitrary passbands.

The no-memory condition is not essential – any monotonic operator that meets the slope conditions will do. In fact, the theorem is valid for any operator that is close (in the specified sense) to a linear operator.

The theorem can be shown to be valid in the more general case of a continuous

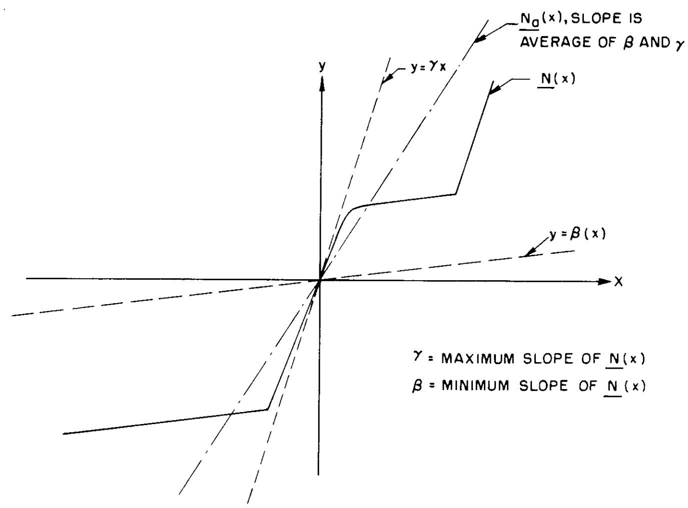

Fig. VIII-3. Graphs of the operator $\underline{N}$, and its linear part, $\underline{N}_a$.

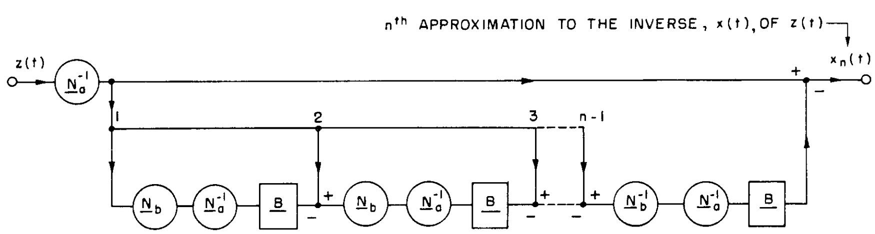

Fig. VIII-4. Schematic form for the n^{th} approximation to the inverse of $\underline{B} * \underline{N}$. (Note that $\underline{N}_a^{-1}$ is a pure gain.)

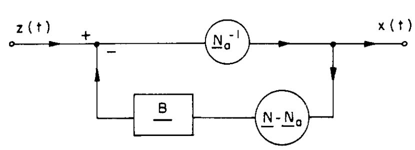

Fig. VIII-5. Realization of the inverse of $\underline{B} * \underline{N}$ by means of a feedback system.

operator that is monotonically increasing, and that satisfies only the upper Lipschitz condition in some arbitrarily small neighborhood of the origin and outside some other arbitrarily large neighborhood of the origin. Such an operator can be approximated in L_2 as the limit of a sequence of bounded slope operators.

i. Interpretation of the Theorem

When $\underline{N}$ has a derivative the theorem is applicable, provided that this derivative has an upper bound, γ, and a lower bound, β, that is greater than zero. The graph of $\underline{N}$ can then be bounded by two straight lines, as in Fig. VIII-3. The graph of $\underline{N}_a(x)$ is a straight line, whose slope is the average of the slopes of the bounding lines.

The iteration has the schematic form shown in Fig. VIII-4, and may be realized, at least formally, as the feedback system shown in Fig. VIII-5.

The rapidity of convergence of the iteration is greater than that of the geometric series, $1 + \alpha + \alpha^2 + \dots$, in which $\alpha = (\gamma-\beta)/(\gamma+\beta)$. For rapid convergence it is therefore desirable that the difference between maximum and minimum slopes be small.

j. Imperfect Bandlimiting

An ideal bandlimiter is not realizable. However, it can be approximated, for example, by a Butterworth filter. We shall compute a bound on the error in the inverse signal, x(t), that results from this approximation.

We have shown that x(t) satisfies without error the following relation:

$$x = (\underline{B}*\underline{N})^{-1}(z) = \underline{N}_a^{-1}(z) - \underline{B} * \underline{N}_a^{-1} * \underline{N}_b(x) \tag{52}$$

In place of $\underline{B}$ we use an approximate bandlimiter, $\underline{B}'$, which satisfies the restriction

$$|B'(\omega)| \leq 1 \qquad -\infty < \omega < \infty$$

whence, just as with B (Eq. 30), we have, for y in L_2

$$\|B'(y)\| \leq \|y\| \tag{53}$$

This ensures that the iteration converges when $\underline{B}'$ is used in place of $\underline{B}$. That is, there exists some x' with the property that

$$x' = \underline{N}_a^{-1}(z) - \underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x') \tag{54}$$

Subtracting Eq. 53 from Eq. 54, we get for the error in x,

$$\|x' - x\| = \|\underline{B} * \underline{N}_a^{-1} * \underline{N}_b(x) - \underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x')\|$$

$$= \left\|\left(\underline{B} * \underline{N}_a^{-1} * \underline{N}_b(x) - \underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x)\right) + \left(\underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x) - \underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x')\right)\right\| \tag{55}$$

Applying the triangle inequality, we get

$$\|x' - x\| \leqslant \|(\underline{B}-\underline{B}') * \underline{N}_a^{-1} * \underline{N}_b(x)\| + \|\underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x) - \underline{B}' * \underline{N}_a^{-1} * \underline{N}_b(x')\| \tag{56}$$

Since the operator $\underline{B}' * \underline{N}_a^{-1} * \underline{N}_b$ in the second term on the right-hand side is a contraction, we may write

$$\|x' - x\| \leqslant \|(\underline{B}-\underline{B}') * \underline{N}_a^{-1} * \underline{N}_b(x)\| + \alpha\|x' - x\| \tag{57}$$

whence we get

$$\|x' - x\| \leqslant \frac{1}{1-\alpha} \|(\underline{B}-\underline{B}') * \underline{N}_a^{-1} * \underline{N}_b(x)\|$$

$$= c\|(\underline{B}-\underline{B}') * \underline{N}_b(x)\| \tag{58}$$

in which we have combined the constants $\frac{1}{1-\alpha}$ and $\underline{N}_a^{-1}$ into c.

It is clear, then, that the error in determining the original signal, x, is proportional to the error with which $\underline{B}'$ operates on $\underline{N}_b(x)$, and becomes small as $\underline{B}'$ approaches $\underline{B}$.

If the approximate bandlimiter, $\underline{B}'$, is realizable, there is an irreducible error. However, if a delay in the inversion is tolerable, x can be recovered with any desired accuracy by combining the bandlimiter with a delay, and delaying the signal z by a corresponding length of time in each iteration cycle.

G. D. Zames

References

1. G. D. Zames, Nonlinear operators – cascading, inversion, and feedback, Quarterly Progress Report No. 53, Research Laboratory of Electronics, M.I.T., April 15, 1959, p. 93.

2. N. Wiener, The Fourier Integral and Certain of Its Applications (Dover Publications, Inc., New York, 1933), p. 18.

3. A. N. Kolmogoroff and S. V. Fomin, Elements of the Theory of Functionals, Vol. 1 (Graylock Press, Rochester, New York, 1957), p. 43.

4. N. Wiener, op. cit., p. 19.

5. N. Wiener, op. cit., pp. 69-71.