

X. STATISTICAL COMMUNICATION THEORY

Prof. Y. W. Lee	D. A. George	R. E. Wernikoff
Prof. A. G. Bose (Absent)	J. Y. Hayase	C. E. Wernlein, Jr.
M. B. Brilliant	I. M. Jacobs	H. E. White
	A. H. Nuttall	

RESEARCH OBJECTIVES

This group is interested in a variety of problems in statistical communication theory. Current research is primarily concerned with: analytic nonlinear systems, crosscorrelation functions under nonlinear transformation, output probability densities of linear systems, signal theory, second-order correlation functions, and probability distribution analyzers.

1. A generalization of time-invariant linear systems to some nonlinear systems is being developed. The generalization is based upon replacing the condition of linearity by that of analyticity (expressibility in a power series). An analytic system is characterized by an infinite sequence of system functions. One advantage of this development is that two analytic systems in cascade are equivalent to one analytic system, the Fourier transforms of whose system functions are series of products of the transforms of the system functions of the components.

2. The relationship between the crosscorrelation functions of two time functions before and after nonlinear transformation is of considerable theoretical and practical importance. For a certain class of time functions, any nonlinear no-memory transformation leaves the crosscorrelation function unchanged, except for a scale factor. This interesting property is being studied.

3. The determination of probability density functions at the output of a linear system, in the general case, is an exceedingly difficult problem. A method is being developed to find the output density functions in terms of a series expansion.

4. Under development is a theory of signals. The work consists of three parts: (a) analysis of the nature of signals, taking into account the finite resolution of the physical detecting devices; (b) an abstract formulation of signal properties and their algebra, retaining only operationally meaningful concepts; and (c) application to communication theory, notably to the time-domain synthesis of both active and passive networks. The last part includes a study of the representation of certain linear operators as linear combinations of chosen sets of operators.

5. The project on second-order correlation functions continues with particular emphasis on the properties of these functions.

6. The development of the analog probability density analyzer continues.

Y. W. Lee

A. ANALYTIC NONLINEAR SYSTEMS

In view of the evident power of the theory of time-invariant linear systems, in contrast to the theory of more general systems, it seems desirable to generalize this theory to some nonlinear systems. Rather than attempting to generalize immediately to the class of all time-invariant systems, we shall replace the linearity condition by the weaker condition of analyticity. (This amounts to replacing linear functions by power series; and in studying systems that are built up from smaller component systems, we hope to use advantageously the computational device of collecting terms of like degree.)

We shall, in fact, place four restrictions on the class of systems to be treated, of

which three are essentially taken from the theory of linear systems: single input, single output; time-invariance; analyticity; continuity.

The first condition means that the system can be represented mathematically as an operator. To every function of time $f(t)$ (within certain restrictions), called the input, there corresponds another function of time $g(t)$, called the output; the correspondence is indicated by $g = \mathcal{H}f$. Specifying $f(t)$ for every t defines the function f ; the operator \mathcal{H} defines $g = \mathcal{H}f$, and the output at any time t has the value $g(t) = \mathcal{H}f(t)$.

The condition of time-invariance allows us to represent the system by a functional. (A functional assigns to every given function a corresponding number; thus it is easier to specify than an operator.) Given an input function f , and fixing a "present time" t , we can define

$$u(\tau) = f(t-\tau) \quad (1)$$

which expresses the input on a time scale measured (backward) relative to the present time t . (We use a backward relative time so that ultimately we get a form comparable to the convolution integral.) Then, as a consequence of time-invariance, the present output $g(t)$ depends only on the function u , and we can write it as the functional

$$g(t) = \mathcal{H}^*u \quad (2)$$

Given a functional \mathcal{H}^* , Eqs. 1 and 2 show how it can be used to define a time-invariant system. (Note that such properties as boundedness, measurability, periodicity, and so forth, hold for u if they hold for f .)

If the system were linear, and

$$f(t) = \sum_k c_k f_k(t) \quad (3a)$$

implying that

$$u(\tau) = \sum_k c_k u_k(\tau) \quad (3b)$$

we would have $\mathcal{H}^*u = \sum_k c_k \mathcal{H}^*u_k$. Replacing linearity by analyticity, that is, replacing the linear function of the c_k by a power series, we write

$$\begin{aligned} \mathcal{H}^*u &= \sum_{n=0}^{\infty} \sum_{k_1} \dots \sum_{k_n} A_{u_{k_1} \dots u_{k_n}} \prod_{i=1}^n c_{k_i} \\ &= A + \sum_k A_{u_k} c_k + \sum_{k_1} \sum_{k_2} A_{u_{k_1} u_{k_2}} c_{k_1} c_{k_2} + \dots \end{aligned} \quad (4)$$

(X. STATISTICAL COMMUNICATION THEORY)

where each $A_{u_{k_1} \dots u_{k_n}}$ is a number that is specified when the functions u_{k_1}, \dots, u_{k_n} are specified, independently of the order of these functions, and independently of whatever other u_k may be involved in other terms. We shall say that a system is analytic if, whenever the sum in Eq. 3a is an infinite series that converges everywhere absolutely, Eq. 4 holds and the sums on the k 's in Eq. 4 converge absolutely. (This implies that Eq. 4 must hold if the sum is finite.)

We shall restrict ourselves to bounded inputs, for both mathematical and physical reasons. The physical reason is that the characteristics of practical nonlinear systems can be measured only for a finite range of input values. We shall, however, admit input functions that may be nonzero over the whole infinite time range $(-\infty, \infty)$.

The continuity condition appropriate to this class of inputs is that we can make $\left| A_{u_{k_1} \dots u_{k_n}} - A_{u'_{k_1} \dots u'_{k_n}} \right|$ arbitrarily small by making $|u_{k_i}(\tau) - u'_{k_i}(\tau)| < \epsilon$ for all i and all τ , with a sufficiently small positive number ϵ .

Under these conditions we can derive a generalization of the convolution integral to analytic systems. A rigorous derivation is obtained by considering those functions u which are characteristic functions of sets of values of τ ; that is, defined so that $u(\tau) = 1$, for a certain set of values of τ , and zero for all other values. Then $A_{u_{k_1} \dots u_{k_n}}$ defines a completely additive set function in n -dimensional space, and, by the continuity condition, the sum on k_1, \dots, k_n becomes an integral with respect to this set function. We are well on the way to a Lebesgue-Stieltjes integral. We can avoid it by the convenient mathematical fiction that if we use impulses we can differentiate any function; alternatively, if we strengthen the continuity condition by replacing "all τ " by "all τ except on a set of sufficiently small measure," the set function becomes the integral of an absolutely integrable function of n variables. We then have

$$\mathcal{H}^*(u) = \sum_{n=0}^{\infty} \int \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n u(\tau_i) d\tau_i \quad (5)$$

the integrals being taken on $(-\infty, \infty)$.

Another less rigorous derivation is obtained by considering those functions u which are impulses. Then $u_{k_i}(\tau) = \delta(\tau - \tau_i)$. The sums on k_i become integrals on τ_i ; Eq. 3b becomes the absolutely convergent integral

$$u(\tau) = \int u(\tau') \delta(\tau - \tau') d\tau' \quad (6)$$

so that c_k has become $u(\tau')$. Then $A_{u_{k_1} \dots u_{k_n}}$ becomes the function of n variables

$h_n(\tau_1, \dots, \tau_n)$ and Eq. 4 becomes Eq. 5. It is evident that the functions h_n are absolutely integrable.

Alternatively, we can expand $u(\tau)$ in a normal orthogonal sequence of functions u_k , so that if

$$c_k = \int u(\tau) u_k(\tau) d\tau \quad (7)$$

then Eq. 3b holds, at least formally. Then if we use Eq. 7 to substitute for c_{k_i} in Eq. 4, and define

$$h_n(\tau_1, \dots, \tau_n) = \sum_{k_1} \dots \sum_{k_n} A_{u_{k_1} \dots u_{k_n}} \prod_{i=1}^n u_{k_i}(\tau_i) \quad (8)$$

we again obtain Eq. 5, at least formally. But with this method it is not easy to apply the conditions of absolute convergence and boundedness to obtain the absolute integrability of the functions h_n .

Having obtained Eq. 5, we make the substitutions indicated by Eqs. 1 and 2 and obtain

$$\begin{aligned} g(t) &= \sum_{n=0}^{\infty} \int \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n f(t - \tau_i) d\tau_i \\ &= h_0 + \int h_1(\tau) f(t - \tau) d\tau + \int \int h_2(\tau_1, \tau_2) f(t - \tau_1) f(t - \tau_2) d\tau_1 d\tau_2 + \dots \end{aligned} \quad (9)$$

This is the desired generalization of the convolution integral.

The functions h_n thus obtained are symmetric functions. They are uniquely determined if the system is specified; and it is evident that if they are specified they determine the system. They are, in fact, determined if the response of the system is specified only for all inputs $f(t)$ for which $|f(t)| < M$ for any arbitrarily small positive M ; so that systems, like limiters, whose large-signal response is not predictable from their small-signal response, are not analytic.

Equation 9 evidently defines the response to any input $f(t)$ for which $|f(t)|$ is always less than the radius of convergence of the power series

$$\sum_{n=0}^{\infty} x^n \int \dots \int |h_n(\tau_1, \dots, \tau_n)| d\tau_1 \dots d\tau_n \quad (10)$$

because this power series dominates Eq. 9. (It may also define a response for some

(X. STATISTICAL COMMUNICATION THEORY)

other input.) If the radius of convergence of Eq. 10 is infinite, we conclude that, for every bounded input, Eq. 9 defines a bounded output. We can then consider a chain of a number of such systems in cascade with no mathematical doubts about achieving a well-defined output.

Equation 9 is a power series in a certain sense; the n^{th} term is a term of n^{th} degree, in that multiplying the input by a constant A results in multiplying the n^{th} term by A^n . With the degree of each term thus defined, we can apply the device of "collecting terms of like degree" to the study of the cascade chains mentioned above.

We note one restriction which was not imposed on the class of systems treated. This is the physical realizability condition that the present output be independent of future input. It does not simplify the mathematics. Wherever it is necessary, it can be imposed by taking all integrals on $(0, \infty)$ instead of on $(-\infty, \infty)$ or by requiring the h_n functions to be zero whenever any of their arguments are negative.

Many of the formulas and concepts mentioned above can be found in the works of Volterra, Wiener, Ikehara, and Deutsch (see refs. 1-6).

1. Fourier Transforms and Analytic Systems

Having generalized the convolution integral, we would like to find a corresponding generalization of the Fourier transform relations of linear system theory. Accordingly, we consider the n^{th} term of Eq. 9

$$g_n = \int \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n f(t - \tau_i) d\tau_i \quad (11)$$

and define the transforms

$$H_n(\omega_1, \dots, \omega_n) = \int \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n e^{-j\omega_i \tau_i} d\tau_i \quad (12)$$

$$F(\omega) = \frac{1}{2\pi} \int f(t) e^{-j\omega t} dt \quad (13)$$

$$G(\omega) = \frac{1}{2\pi} \int g(t) e^{-j\omega t} dt \quad (14)$$

We then obtain formally, by assuming the validity of the inversion formulas,

$$g_n(t) = \int \dots \int \exp\left(jt \sum_{i=1}^n \omega_i\right) H_n(\omega_1, \dots, \omega_n) \prod_{i=1}^n F(\omega_i) d\omega_i \quad (15)$$

This formula has an interesting physical interpretation: the frequency components of the output of a nonlinear device are sums and harmonics of the frequencies comprising the input. Equation 15 defines the magnitudes of these contributions to the spectrum of the output. If we write

$$G_n^*(\omega_1, \dots, \omega_n) = H_n(\omega_1, \dots, \omega_n) \prod_{i=1}^n F(\omega_i) \quad (16)$$

we can conclude that there exists a function

$$g_n^*(t_1, \dots, t_n) = \int \dots \int G_n^*(\omega_1, \dots, \omega_n) \prod_{i=1}^n e^{j\omega_i t_i} d\omega_i \quad (17)$$

which is such that

$$g_n(t) = g_n^*(t, t, \dots, t) \quad (18)$$

at least formally; but noting that transform inversions are, in general, valid only almost everywhere, we are not assured that Eq. 18 really defines the output.

2. Analytic Systems in Cascade

Suppose that we have an input $e(t)$ to an analytic system \mathcal{X} whose output is $f(t)$, and that this $f(t)$ is then taken as the input to another system \mathcal{H} whose output is $g(t)$. We seek a description of the equivalent system, the system \mathcal{L} whose input is $e(t)$ and whose output is $g(t)$. This turns out to be an analytic system.

Mathematically, we have

$$g(t) = \sum_{n=0}^{\infty} g_n(t) = \sum_{n=0}^{\infty} h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n f(t - \tau_i) d\tau_i \quad (19)$$

$$f(t) = \sum_{m=0}^{\infty} f_m(t) = \sum_{m=0}^{\infty} k_m(\sigma_1, \dots, \sigma_m) \prod_{j=1}^m e(t - \sigma_j) d\sigma_j \quad (20)$$

and we want to find functions ℓ_p that are such that

$$g(t) = \sum_{p=0}^{\infty} \gamma_p(t) = \sum_{p=0}^{\infty} \ell_p(\eta_1, \dots, \eta_p) \prod_{r=1}^p e(t - \eta_r) d\eta_r \quad (21)$$

We obtain

(X. STATISTICAL COMMUNICATION THEORY)

$$g_n(t) = \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} \int \cdots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n f_{m_i}(t - \tau_i) d\tau_i \quad (22)$$

We then select, for each p , those terms that are of p^{th} degree in $e(t)$; that is, we define

$$g_{np}(t) = \sum_{\sum_{i=1}^n m_i=p} \int \cdots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n f_{m_i}(t - \tau_i) d\tau_i \quad (23)$$

and recombine these to form

$$\gamma_p(t) = \sum_{n=0}^{\infty} g_{np}(t) \quad (24)$$

and we find that Eq. 21 will hold if we have

$$\ell_p(\eta_1, \dots, \eta_p) = \sum_{n=0}^{\infty} \sum_{\sum_{i=1}^n m_i=p} \int \cdots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^n k_{m_i}(\eta_{r_i+1} - \tau_i, \dots, \eta_{r_i+m_i} - \tau_i) d\tau_i \quad (25)$$

where $r_i = r_{i-1} + m_{i-1}$, and $r_1 = 0$. In other words, the variables η are numbered consecutively in the order in which they appear when the k_{m_i} are written in order of increasing i . The second summation is taken over all permutations of all n numbers m_1, \dots, m_n whose sum is p . We obtain for the corresponding Fourier transforms

$$L_p(\omega_1, \dots, \omega_p) = \sum_{n=0}^{\infty} \sum_{\sum_{i=1}^n m_i=p} H_n(\Omega_1, \dots, \Omega_n) \prod_{i=1}^n K_{m_i}(\omega_{r_i+1}, \dots, \omega_{r_i+m_i}) \quad (26)$$

where Ω_i is the sum of the m_i arguments of K_{m_i} . These results follow rigorously from the absolute integrability of the h_n and k_m functions.

Equations 25 and 26 cannot be evaluated exactly if we are dealing with infinite series. If, however, there are no terms in \mathcal{K} of degree higher than M and no terms in \mathcal{H} of degree higher than N , these equations can be evaluated, and there are no terms in \mathcal{L} of degree higher than $P = MN$.

M. B. Brilliant

References

1. V. Volterra, *Lecons sur les Fonctions de Lignes* (Paris, 1913).
2. N. Wiener, Average of an analytic functional, *Proc. Nat. Acad. Sci.* 7, 253-260 (1922).
3. N. Wiener, Differential space, *J. Math. Phys.* 2, 131-174 (1923).
4. N. Wiener, Response of a nonlinear device to noise, Report 129, Radiation Laboratory, M.I.T., April 6, 1942.
5. S. Ikehara, A method of Wiener in a nonlinear circuit, Technical Report 217, Research Laboratory of Electronics, M.I.T., Dec. 10, 1951.
6. R. Deutsch, On a method of Wiener for noise through nonlinear devices, *IRE Convention Record*, Part 4, 1955, pp. 186-192.

B. INVARIANCE OF CORRELATION FUNCTIONS UNDER NONLINEAR TRANSFORMATIONS

In 1952, Bussgang (1) demonstrated the useful and interesting result that if one of a pair of stationary time series with gaussian probability density functions is amplitude-distorted in a nonlinear no-memory device, then the crosscorrelation function after distortion is proportional to the crosscorrelation function before distortion. We shall call this the "invariance property." In the nonstationary case, the constant of proportionality may be dependent on the instantaneous time in the time series which passes through the nonlinear device. In this report, the crosscorrelation function is defined as in Eq. 10. This definition is sometimes called the "covariance function." A no-memory device is one for which the output at any one instant of time is dependent only on the input at that same instant. Such a system is shown in Fig. X-1, in which

$$y_2(t) = f[x_2(t), t] \quad (1)$$

In 1955, Barrett and Lampard (2) generalized the invariance property to the extent that the second-order probability density function of the pair of (nonstationary) time series $x_1(t)$ and $x_2(t)$ need only be expressible in the form of a single series as

$$p(x_1, t_1; x_2, t_2) = p_1(x_1, t_1) p_2(x_2, t_2) \sum_{n=0}^{\infty} a_n(t_1, t_2) \theta_n^{(1)}(x_1, t_1) \theta_n^{(2)}(x_2, t_2) \quad (2)$$

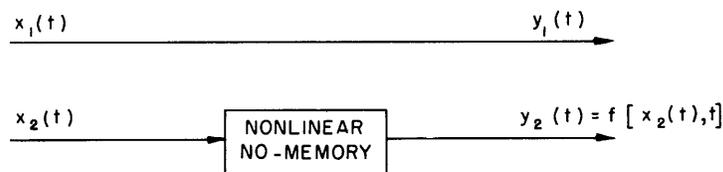


Fig. X-1. Nonlinear no-memory transformation of one input process.

(X. STATISTICAL COMMUNICATION THEORY)

This class of second-order probability density functions will be denoted by Λ . The corresponding first-order probability density functions are

$$p_1(x_1, t_1) = \int p(x_1, t_1; x_2, t_2) dx_2 \quad (3)$$

and

$$p_2(x_2, t_2) = \int p(x_1, t_1; x_2, t_2) dx_1 \quad (4)$$

[All integrals in this report are over the range $(-\infty, +\infty)$.]

The two sets of orthonormal polynomials, $\{\theta_n^{(1)}(x_1, t_1)\}$ and $\{\theta_n^{(2)}(x_2, t_2)\}$, were constructed to satisfy

$$\int p_1(x_1, t_1) \theta_m^{(1)}(x_1, t_1) \theta_n^{(1)}(x_1, t_1) dx_1 = \delta_{mn} \quad (5)$$

and

$$\int p_2(x_2, t_2) \theta_m^{(2)}(x_2, t_2) \theta_n^{(2)}(x_2, t_2) dx_2 = \delta_{mn} \quad (6)$$

However, Barrett and Lampard restricted their nonlinear devices to a class that we shall call θ , with the property that $f[x_2(t), t]$ be expressible as a series in the second set of orthonormal polynomials:

$$f[x_2(t), t] = \sum_{n=0}^{\infty} b_n(t) \theta_n^{(2)}(x_2, t) \quad (7)$$

The present report demonstrates two theorems.

THEOREM 1. If the second-order probability density function of the pair of input time series $x_1(t)$ and $x_2(t)$ belongs to the class Λ , then the invariance property holds for any time-varying nonlinear no-memory device, and not merely for the class θ .

Denote by P the class of time-varying, nonlinear, no-memory devices that can be expressed as power series:

$$f[x_2(t), t] = \sum_{n=0}^{\infty} c_n(t) x_2^n(t) \quad (8)$$

We then have Theorem 2.

THEOREM 2. The condition

$$\overline{x_1(t_1) x_2^n(t_2)} - \overline{x_1(t_1)} \overline{x_2^n(t_2)} = d_n(t_2) \left[\overline{x_1(t_1) x_2(t_2)} - \overline{x_1(t_1)} \overline{x_2(t_2)} \right] \quad \text{for all } n \quad (9)$$

where $\{d_n(t_2)\}$ is a sequence of real constants at any t_2 , is sufficient for the invariance property to hold for all nonlinear devices in P.

PROOF OF THEOREM 1. The output correlation function of the system of Fig. X-1 is defined as

$$\Phi_{12}(t_1, t_2) = \overline{\{y_1(t_1) - \overline{y_1(t_1)}\} \{y_2(t_2) - \overline{y_2(t_2)}\}} \quad (10)$$

$$= \overline{\{x_1(t_1) - \overline{x_1(t_1)}\} \{f[x_2(t_2), t_2] - \overline{f[x_2(t_2), t_2]}\}} \quad (11)$$

$$= \iint [x_1 - \mu_1^{(1)}(t_1)] [f(x_2, t_2) - \lambda^{(2)}(t_2)] p(x_1, t_1; x_2, t_2) dx_1 dx_2 \quad (12)$$

where

$$\mu_1^{(1)}(t_1) = \overline{x_1(t_1)} \quad (13)$$

and

$$\lambda^{(2)}(t_2) = \overline{f[x_2(t_2), t_2]} \quad (14)$$

Assuming now $p(x_1, t_1; x_2, t_2) \in \Lambda$ (see Eq. 2), substituting in Eq. 12, and interchanging integration and summation, we have

$$\begin{aligned} \Phi_{12}(t_1, t_2) &= \sum_{n=0}^{\infty} a_n(t_1, t_2) \left[\int (x_1 - \mu_1^{(1)}(t_1)) p_1(x_1, t_1) \theta_n^{(1)}(x_1, t_1) dx_1 \right] \\ &\quad \times \left[\int (f(x_2, t_2) - \lambda^{(2)}(t_2)) p_2(x_2, t_2) \theta_n^{(2)}(x_2, t_2) dx_2 \right] \end{aligned} \quad (15)$$

$$= \frac{\phi_{12}(t_1, t_2)}{\sigma_2(t_2)} \int [f(x_2, t_2) - \lambda^{(2)}(t_2)] p_2(x_2, t_2) \theta_1^{(2)}(x_2, t_2) dx_2 \quad (16)$$

where

$$\phi_{12}(t_1, t_2) = \overline{[x_1(t_1) - \overline{x_1(t_1)}] [x_2(t_2) - \overline{x_2(t_2)}]} = \overline{x_1(t_1) x_2(t_2)} - \overline{x_1(t_1)} \overline{x_2(t_2)} \quad (17)$$

Also, since

$$\int \lambda^{(2)}(t_2) p_2(x_2, t_2) \theta_1^{(2)}(x_2, t_2) dx_2 = 0 \quad (18)$$

we have

$$\Phi_{12}(t_1, t_2) = C_f(t_2) \phi_{12}(t_1, t_2) \quad (19)$$

(X. STATISTICAL COMMUNICATION THEORY)

where

$$C_f(t_2) = \frac{1}{\sigma_2^2(t_2)} \int f(x_2, t_2) p_2(x_2, t_2) [x_2 - \overline{x_2(t_2)}] dx_2 \quad (20)$$

and is dependent only on the nonlinear device and first-order statistics of $x_2(t)$. (For details, see ref. 2.) This completes the proof of Theorem 1.

PROOF OF THEOREM 2. For the class P,

$$\Phi_{12}(t_1, t_2) = [x_1(t_1) - \mu_1^{(1)}(t_1)] \left[\sum_{n=0}^{\infty} c_n(t_2) x_2^n(t_2) - \sum_{n=0}^{\infty} c_n(t_2) \overline{x_2^n(t_2)} \right] \quad (21)$$

$$= \sum_{n=0}^{\infty} c_n(t_2) \left[\overline{x_1(t_1) x_2^n(t_2)} - \overline{x_1(t_1)} \overline{x_2^n(t_2)} \right] \quad (22)$$

By using Eq. 17, we obtain

$$\Phi_{12}(t_1, t_2) = \sum_{n=0}^{\infty} c_n(t_2) d_n(t_2) \phi_{12}(t_1, t_2) \quad (23)$$

Therefore,

$$\Phi_{12}(t_1, t_2) = C_f(t_2) \phi_{12}(t_1, t_2) \quad (24)$$

where

$$C_f(t_2) = \sum_{n=0}^{\infty} c_n(t_2) d_n(t_2) \quad \text{Q. E. D.} \quad (25)$$

The validity of condition 9 is easily checked from the second-order probability density function of x_1 and x_2 , and avoids entirely the determination of the orthonormal polynomials $\{\theta_n^{(1)}(x_1, t_1)\}$ and $\{\theta_n^{(2)}(x_2, t_2)\}$, and subsequent determination of the validity of Eq. 2. However, the simplicity of Eq. 9 as a condition is limited by the present proof to nonlinear devices belonging to P. Generalization of the class P will be made in the following Quarterly Progress Report.

If $x_1(t)$ instead of $x_2(t)$ were passed through the nonlinear device, subscripts 1 and 2 would have to be interchanged in all the previous work. In particular, condition 9 would become

$$\overline{x_1^n(t_1) x_2(t_2)} - \overline{x_1^n(t_1)} \overline{x_2(t_2)} = d_n(t_1) \phi_{12}(t_1, t_2) \quad \text{for all } n \quad (26)$$

which is possibly a different restriction from condition 9. Either, both, or neither of

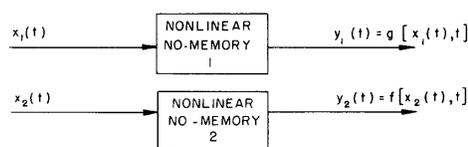


Fig. X-2. Nonlinear no-memory transformations of both input processes.

conditions 9 and 26 may be true in the general case, and care must be taken that the nonlinear device be in the corresponding lead for validity of the invariance property. It might be pointed out that satisfaction of both conditions 9 and 26 does not imply that the invariance property holds for the system in Fig. X-2. Without going into detail, the sufficient condition for the invariance property to hold in the system of Fig. X-2 is

$$\overline{x_1^m(t_1) x_2^n(t_2)} - \overline{x_1^m(t_1)} \overline{x_2^n(t_2)} = c_{mn} \phi_{12}(t_1, t_2) \quad \text{for all } m, n \quad (27)$$

where the invariance property is now interpreted as

$$\Phi_{12}(t_1, t_2) = C_{fg} \phi_{12}(t_1, t_2) \quad (28)$$

and C_{fg} is independent of both t_1 and t_2 . This is, indeed, a very stringent requirement. However, the class of density functions satisfying condition 27 is not void, since all second-order probability density functions of independent variables x_1 and x_2 are in this class.

A. H. Nuttall

References

1. J. J. Bussgang, Crosscorrelation functions of amplitude distorted gaussian signals, Technical Report 216, Research Laboratory of Electronics, M. I. T., March 26, 1952.
2. J. F. Barrett and D. G. Lampard, An expansion for some second-order probability distributions and its applications to noise problems, Trans. IRE, vol. IT-1, pp. 10-15 (March, 1955).

C. REPRESENTATION AND SYNTHESIS OF CERTAIN LINEAR OPERATORS

It will be shown that linear time-invariant operators may be thought of as vectors in a linear vector space. Under suitable conditions, which will be specified, given sets of operators then play the part of coordinate axes, so that an arbitrary operator in the space is represented as a linear combination (vector sum) of the given set of operators. The coordinate system can be chosen in such a way that this form of representation

(X. STATISTICAL COMMUNICATION THEORY)

affords an intuitive understanding of the mode of operation of the given operator, and immediately suggests a way of synthesizing it.

I. Introduction

The germ of the idea is contained in M. V. Cerrillo's explanation of the operation of linear prediction or extrapolation. The first-order prediction kernel (1), shown in Fig. X-3a, is thought of as the sum of a pure transmission kernel (Fig. X-3b) and

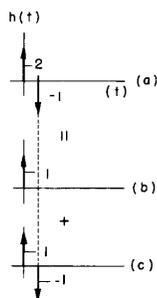


Fig. X-3. Decomposition of a prediction kernel.

a differentiation kernel (Fig. X-3c). From this it becomes clear that the kernel of Fig. X-3a, in itself not very meaningful intuitively, produces just the first two terms of the Taylor series of the input shifted forward by one unit. This method of presentation shows why a network with the impulse response of Fig. X-3a produces an output which approximates the input predicted one unit ahead.

More generally, the method shows two things: (a) how intuitively meaningful network operation can be made by the simple expedient of regarding the given network as the sum of other networks whose individual operation is intuitively meaningful;

and (b) that if a given operator (or impulse response) is to be synthesized physically, we can accomplish this immediately if we can decompose the desired operator into a sum of other operators each of which we already know how to synthesize. It appears, therefore, that it might be useful to know how to decompose any arbitrary operator or network, if the decomposition takes the form of a linear combination of elementary operators specified a priori. This last reservation is necessary for two reasons: (a) if the object of performing the decomposition is to aid us in understanding the operation of a given network, we must reserve some freedom of choice of the elements into which the decomposition is to be made, to insure that those elements will be meaningful to us; and (b) if the object is the physical synthesis of an operator, we must have some freedom in the choice of the decomposition elements, so that only those elements occur which we already know how to synthesize. In fact, we may have a set of operators already available in physical form (for example, analog computer elements), in which case we will wish the synthesis to be carried out in terms of the available elements.

Can our requirements be satisfied? We shall consider the question in the three following parts:

1. Under what conditions can an arbitrary time-invariant operator be represented as a linear combination of a given set of operators without having anything left over?
2. Assuming the first question is answered, by what method, short of cut-and-try,

can we determine in what proportions to take the operators of the given set in order to have their sum equal to the desired operator?

3. Are these proportions unique? Can the operator be decomposed in more than one way?

These questions can be answered very simply.

2. Decomposition of Singular Networks into Elementary Operations

Our discussion will be restricted to the decomposition of singular networks, but, as was shown in reference 2, this is really no restriction, since any network can be considered equivalent (i. e., approximately equal, in some sense) to a singular network of sufficiently high, finite degree N . It was also shown in reference 2 that, by using the results of numerical analysis, any time-invariant linear operator can be represented by a singular network. Approximations are necessary only in going from operators and networks to singular networks. From that point on, our discussion is exact.

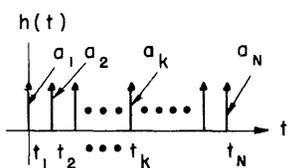


Fig. X-4. A singular impulse response.

Consider the singular response $h(t)$ shown in Fig. X-4. The only information contained in $h(t)$ is the set of impulse areas $\{a_k\}$ and the set of corresponding impulse positions $\{t_k\}$. We summarize this information by writing

$$h(t) = \sum_{k=1}^N a_k \delta(t - t_k) \quad (1)$$

The symbol $\delta(t - t_k)$, when it occurs in a network impulse response, stands for the instruction: "delay by t_k seconds." If we think of this instruction as an operator, and if we denote it by τ_k , then Eq. 1 can be represented by

$$h(t) = \sum_{k=1}^N a_k \tau_k \quad (2)$$

One possible interpretation of Eq. 2 is to think of $h(t)$ as a vector in N -space, with the delay operators τ_k acting as the unit vectors of a coordinate system, and with the a_k representing the components of $h(t)$ along the various directions. This simple point of view solves our problem, and if we regard the set of impulse-positions $\{t_k\}$ as fixed we can show easily that the set of all possible N -tuplets $\{a_k\}$, together with addition and multiplication by a scalar defined as usual, forms a linear vector space, so that

(X. STATISTICAL COMMUNICATION THEORY)

our interpretation is possible.

In the discussion that follows, let us suppose, for simplicity, that the set of fixed impulse positions is uniformly spaced from $t_1 = 0$ to $t_N = T$. Vectors will be denoted by lower-case Greek letters, and their components (scalars) by lower-case Roman letters.

Using the results of reference 2, we represent any linear, time-invariant operator by an expression of the form

$$\omega = \sum_k w_k \tau_k$$

If we are given a set of operators $\{\omega_j\}$, in which the j^{th} operator is given by

$$\omega_j = \sum_k w_{kj} \tau_k \quad (3)$$

then we can, under certain conditions, use the set $\{\omega_j\}$ as a new coordinate system for the original vector space. In fact, if the set of N coordinate vectors $\{\omega_j\}$ spans the space, we can express any vector in the space in terms of the $\{\omega_j\}$ (ref. 3), so that we can write

$$h(t) = \sum_k a_k \tau_k = \sum_j a'_j \omega_j \quad (4)$$

Equation 4 has precisely the desired form.

In terms of vector algebra, the questions asked at the beginning of this report are answered immediately. Question 1 has been answered already: if the $\{\omega_j\}$ span the space (i. e., form a linearly independent set of N vectors), any vector in the space can be expressed exactly in terms of its components along the ω -system. Spanning the space is also the necessary and sufficient condition that there exist a nonsingular transformation A connecting the τ -system with the ω -system. The matrix of the transformation, $[A]$, is specified by the N^2 coefficients w_{kj} in Eq. 3, that is, $[A] = [w_{kj}]$. Since Eq. 3 gives the coordinate vectors of the ω -system in terms of those of the τ -system, any vector with components $\{a_k\}$ in the τ -system will have components $\{a'_k\}$ in the ω -system given by

$$\begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_k \\ \vdots \\ a'_N \end{bmatrix} = [A]^{-1} \times \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \\ \vdots \\ a_N \end{bmatrix} \quad (5)$$

(X. STATISTICAL COMMUNICATION THEORY)

certainly linearly independent, since, as shown in Eqs. 6, each successive vector occupies an additional dimension of the space. Thus our set $\{\omega_k\}$ constitutes a basis, and, therefore, it can be used to represent any vector in the space.

3. Synthesis

The synthesis problem is trivial. The decomposition process results in the sum

$$\sum_k a'_k \omega_k$$

and, since the $\{\omega_k\}$ were presumably chosen with synthesis in mind, the physical embodiment of each ω_k is known, and the problem is solved.

In our present terminology, we can regard the synthesis process as a coordinate rotation from the τ -system, in which the problem is obscure, to a suitably chosen ω -system in which the solution is obvious.

R. E. Wernikoff

References

1. M. V. Cerrillo, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., Oct. 15, 1956, p. 66.
2. R. E. Wernikoff, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1956, p. 44.
3. P. R. Halmos, Finite Dimensional Vector Spaces (Princeton University Press, Princeton, N. J., 1955).

D. CONVERGENCE PROPERTIES OF SEQUENCES OF SINGULAR NETWORKS

Let $A(t)$ be the step-response of a network, and assume that it is bounded and that it becomes constant after some finite time T . Under these conditions, it is possible to construct a sequence $\{A_n(t)\}$ of approximating simple functions that converges uniformly to $A(t)$. The networks corresponding to the step responses $A_n(t)$ are singular networks. Let $f(t)$ be an input function used to excite the networks $A(t)$ and $A_n(t)$ simultaneously, and let $g(t)$ and $g_n(t)$ be the corresponding outputs. Then, as was shown in reference 1,

$$g(t) = \lim_{n \rightarrow \infty} g_n(t)$$

In this report we shall study briefly the mode of convergence of $\{g_n(t)\}$ to $g(t)$, and describe the restrictions that must be placed on the inputs $f(t)$ to make the convergence uniform.

Consider an arbitrary step response $A(t)$ which vanishes for $t < 0$ and is constant

(X. STATISTICAL COMMUNICATION THEORY)

for $t \geq T$ (where T is some finite time), and let $A(t)$ be bounded for $t \in (0, T)$. Let the sequence of approximating simple functions $\{A_n(t)\}$ be constructed in such a way that $A_n(t) \equiv A(t)$ for $t < 0$ and for $t \geq T$. Since $\{A_n(t)\}$ converges to $A(t)$ uniformly, it is always possible to write $|A(t) - A_n(t)| < \epsilon_n$, in which ϵ_n may be arbitrarily small if n is sufficiently large.

Now we shall study the sequence of outputs from the networks $A_n(t)$. If $f(t)$ starts at $t = -\infty$ (precisely the same analysis can be carried out for inputs starting at $t = 0$),

$$g(t) = f(-\infty) A(\infty) + \int_0^{\infty} A(\sigma) \frac{df(t-\sigma)}{d(t-\sigma)} d\sigma$$

and

$$g_n(t) = f(-\infty) A_n(\infty) + \int_0^{\infty} A_n(\sigma) \frac{df(t-\sigma)}{d(t-\sigma)} d\sigma$$

Therefore,

$$|g(t) - g_n(t)| = \left| f(-\infty)[A(\infty) - A_n(\infty)] + \int_0^{\infty} [A(\sigma) - A_n(\sigma)] \frac{df(t-\sigma)}{d(t-\sigma)} d\sigma \right|$$

Because of the way we constructed $\{A_n(t)\}$,

$$[A(\infty) - A_n(\infty)] = 0$$

and

$$[A(\sigma) - A_n(\sigma)] = 0 \quad \text{for } \sigma > T$$

so that

$$\begin{aligned} |g(t) - g_n(t)| &= \left| \int_0^T [A(\sigma) - A_n(\sigma)] \frac{df(t-\sigma)}{d(t-\sigma)} d\sigma \right| \\ &\leq \int_0^T |A(\sigma) - A_n(\sigma)| \left| \frac{df(t-\sigma)}{d(t-\sigma)} \right| d\sigma \end{aligned}$$

Remembering that we can make $|A(\sigma) - A_n(\sigma)| < \epsilon_n$, we obtain

$$|g(t) - g_n(t)| < \epsilon_n \int_0^T \left| \frac{df(t-\sigma)}{d(t-\sigma)} \right| d\sigma$$

or, making the change of variable $\xi = t - \sigma$, we have

(X. STATISTICAL COMMUNICATION THEORY)

$$|g(t) - g_n(t)| < \epsilon_n \int_{t-T}^t \left| \frac{df(\xi)}{d\xi} \right| d\xi \quad (1)$$

which is the desired result.

The result shows clearly that uniform convergence can be obtained if

$$\int_{t-T}^t \left| \frac{df(\xi)}{d\xi} \right| d\xi \quad (2)$$

has a uniform bound. Several types of restrictions can be placed on the class of admissible inputs $f(t)$ to insure the existence of this bound. For example, a sufficient (although not necessary) condition for the existence of the bound is to consider as admissible inputs only the class of functions with bounded first derivative. In this case, we can write $\left| \frac{df}{d\xi} \right| \leq Q$ (Q a finite constant) and $|g - g_n| < \epsilon_n(QT)$. To see that this requirement is not necessary, note that a step-function input would not satisfy it, and yet, for a step input, the value of the integral 2 never exceeds the height of the step.

Integral 2 is just the variation of $f(t)$ on the interval $(t-T, T)$. Therefore, any input function that is of uniformly bounded variation on every interval of length T produces the desired uniform convergence.

Let the uniform bound on the variation be denoted by V . Then $|g - g_n| < \epsilon_n V$. To get an estimate of the rapidity of convergence with increasing n , we must get an estimate on the size of ϵ_n . Suppose that the step response $A(t)$ has a bounded first derivative, $\left| \frac{dA}{dt} \right| < M$ (M a finite constant). If we let $A_n(t)$ consist of n steps equally spaced along the t -axis at intervals $\delta = T/n$, which is the worst possible way of constructing the approximating sequence, then it is easy to show that $\epsilon_n \leq M\delta = M(T/n)$. Therefore,

$$|g - g_n| < \frac{MVT}{n} = \frac{K}{n}$$

so that even in this worst possible case, $|g - g_n| \rightarrow 0$ as $1/n$.

R. E. Wernikoff

References

1. R. E. Wernikoff, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1956, p. 44.

E. TIME-LIMITED AND BAND-LIMITED FUNCTIONS

We present three short proofs of the fact that it is impossible for a function to be simultaneously time-limited and band-limited. That is, it is impossible to construct a Fourier transform pair $[f(t), F(\omega)]$, not equivalent to zero, which has the property that

1. $f(t)$ vanishes identically everywhere, except on a t -interval of finite length; and simultaneously

2. $F(\omega)$ vanishes identically everywhere, except on an ω -interval of finite length.

PROOF 1. We quote the Paley-Wiener Theorem (1):

Let $\phi(\omega)$ be a real non-negative function, not equivalent to zero, defined for $-\infty < \omega < \infty$, and of integrable square in this range. A necessary and sufficient condition that there should exist a real- or complex-valued function $f(t)$, defined for $-\infty < t < \infty$, vanishing for $t \geq t_0$ for some number t_0 , and such that the Fourier transform $F(\omega)$ of $f(t)$ should satisfy $|F(\omega)|^2 = \phi(\omega)$, is that

$$\int_{-\infty}^{\infty} \frac{|\log \phi(\omega)|}{1 + \omega^2} d\omega < \infty \quad (1)$$

We notice that if $F(\omega) \equiv 0$ over any ω -set of positive measure (in particular, over any ω -interval of positive length), then integral 1 is not finite. Therefore, according to this theorem, if $F(\omega)$ occupies only a finite band, and is zero outside of that band, $f(t)$ cannot vanish even over a half-line $t \geq t_0$, let alone be time-limited.

PROOF 2. Proof by contradiction. Suppose that there does exist a time-limited function $f(t)$ with spectrum $F(\omega)$, which is such that the part of the spectrum which does not vanish identically is entirely contained in the ω -interval $(-W, W)$. We assume that $f(t)$ vanishes identically outside some interval of length T , which we can take, without loss of generality, to be the interval $(0, T)$. Then, according to the Shannon Sampling Theorem (2), $f(t)$ is given by

$$f(t) = \sum_n a_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}, \quad a_n = f\left(\frac{n}{2W}\right)$$

However, since $f(t)$ is limited to $(0, T)$, the only samples that are not zero are those taken in $(0, T)$. There are $N = 2TW$ such samples, so that

$$f(t) = \sum_{n=0}^N a_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}$$

That is, $f(t)$ is given by a finite sum of $\frac{\sin x}{x}$ functions. But now, the assumption that $f(t)$ is limited to $(0, T)$ implies, for example, that

$$\sum_{n=0}^N a_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)} \equiv 0 \quad \text{for all } t \in (-\infty, 0)$$

This requirement states that the tails of a finite number of $\frac{\sin x}{x}$ functions have to

(X. STATISTICAL COMMUNICATION THEORY)

combine in such a way that they cancel each other completely over a whole interval. Since the $\frac{\sin x}{x}$ functions are linearly independent over any interval, there cannot exist a set of nonvanishing coefficients a_n that satisfy this requirement. Therefore, there does not exist any function, not equivalent to zero, limited simultaneously in time and in frequency.

PROOF 3. From the argument in Proof 2, we have that any function $f(t)$ which is limited simultaneously in time and in frequency must be expressible as a finite linear combination of $\frac{\sin x}{x}$ functions. That is, it must have the form

$$f(x) = \sum_{k=1}^N a_k \frac{\sin(x - x_k)}{(x - x_k)}$$

If we extend $f(x)$ to a function $f(z)$ of the complex variable $z = x + jy$, we notice that each term

$$\frac{\sin(z - x_k)}{(z - x_k)}$$

is an analytic, entire function, so that $f(z)$, which is a finite linear combination of such functions, is also analytic and entire.

Now we have the well-known result (3): If a function is analytic in a region, and vanishes along any segment of a continuous curve in the region, then it must vanish identically. Since $f(z)$ is supposed to vanish identically over certain intervals of the real line (corresponding to the regions in which $f(t)$ vanishes), $f(z)$ must vanish identically over the whole complex plane. In particular, then, $f(x)$ must vanish over the whole real line. Since the set of functions

$$\left\{ \frac{\sin(x - x_k)}{(x - x_k)} \right\}$$

is orthogonal over $(-\infty, \infty)$, the requirement $f(x) \equiv 0$ for $-\infty < x < \infty$ can only be satisfied by setting $a_k = 0$ ($k = 1, 2, \dots, N$).

Thus, any function that is limited simultaneously in time and in frequency is equivalent to zero.

R. E. Wernikoff

References

1. R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Vol. 19 (American Mathematical Society Colloquium Publications, New York, 1934), p. 16 (Theorem 12).
2. C. E. Shannon, *A Mathematical Theory of Communication*, Monograph B-1598 (Bell Telephone System Technical Publications, New York, 1948), p. 51.
3. E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, London, 1939), Section 2.61.

F. ANALOG PROBABILITY DENSITY ANALYZER

Additional experimental work with the probability density analyzer (1) resulted in modification of the series diode slicer circuit and development of a clipper integrator circuit to average the pulse output of the slicer circuit. Operation of the resultant distribution analyzer was checked by finding the probability distribution for a sine wave.

In addition to the 60-mc carrier frequency which is switched by the slicer circuit diodes, it is necessary to introduce a signal voltage and a scanning voltage to the slicer circuit. In the previous circuit arrangement (2) the signal and scanning voltages were introduced by separate cathode followers. The signal voltage remained at a constant level, and the scanning voltage varied the level of the diode bias during the scanning operation. This required that the signal and scanning cathode followers produce a linear output over a voltage range of 150 volts, which is difficult to obtain in practice. This problem led to maintaining the slicer diode bias at a constant level and introducing both the signal and scanning voltages through the same cathode follower, as is shown in Fig. X-5. With this arrangement, the same portion of the cathode-follower characteristics is used for each of the analyzed amplitude levels, and the cathode follower distortion at other levels is not important.

When the pulse output produced by the slicer circuit is averaged, the resultant voltage is proportional to the probability of the particular signal amplitude selected by the scanning voltage. The pulse output voltage of the detector is not suitable for averaging by a resistor-capacitor integrator, since the detector diode would be biased by the integrator voltage. It is necessary to integrate the pulse output of a current source if the integrator voltage is to be proportional to the average of the pulses. A pentode amplifier is used for this purpose, since the plate current is independent of plate voltage for a large range of plate voltages. The circuit of Fig. X-6 uses two sharp cutoff pentodes as a current source for the resistor-capacitor integrator and also as a clipper

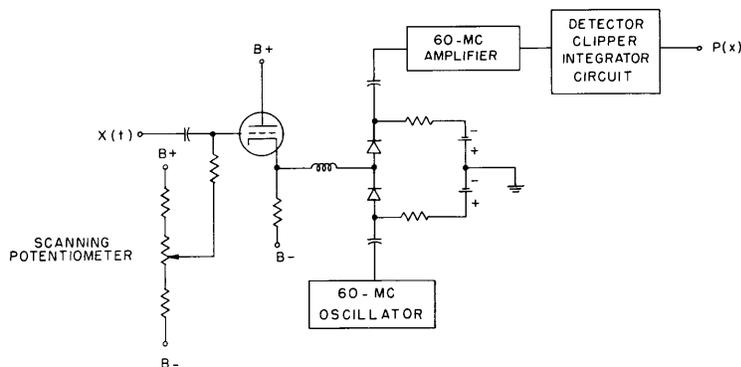


Fig. X-5. Detector, clipper integrator circuit.

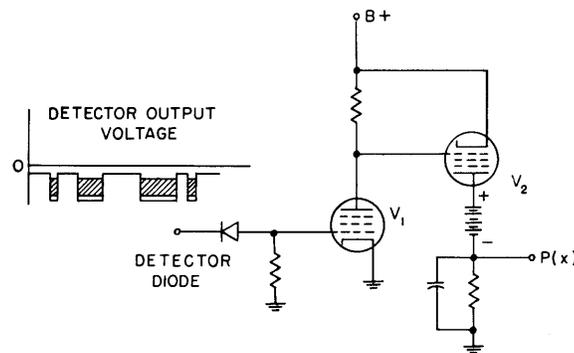


Fig. X-6. Analog probability density analyzer.

to remove any variations of pulse amplitude.

Clipping of large values of the detector pulse output is necessary to remove variations of pulse amplitude caused by oscillator drift or changes of amplifier gain. These variations of amplitude are clipped by pentode V_1 , which is driven below cutoff during peaks of the detector output. Pentode V_2 performs a similar operation for small values of detector output, since it is cut off during periods when there is no pulse output from the slicer circuit. The clipping of small values of the detector output is required in order to remove a small constant output voltage caused by stray coupling through the slicer circuit. The portion of the detector output that is averaged by the resistor-capacitor integrator is, therefore, the shaded area of the detector output waveform shown in Fig. X-6. The integrator voltage is obtained at ground potential by using a separate plate-voltage source for V_2 and by placing the integrator circuit at ground rather than at the plate of V_2 .

Operation of the analog probability density analyzer with the slicer circuit and clipper integrator circuit described above, was checked by finding points on the amplitude distribution of a sine wave for 50 amplitude channels. The experimental points show no deviation from the calculated curve for frequencies less than 25 kc, which is consistent for frequency-response calculations based on the pulse response of the band-pass amplifier. Exact comparison of experimental results with calculated values will be made, when the output is plotted by a pen recorder synchronized with the scanning voltage.

H. E. White

References

1. H. E. White, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., April 15, 1956, p. 70.
2. A. G. Bose, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., Oct. 15, 1955, p. 58.

G. ON THE SYNTHESIS OF LINEAR SYSTEMS FOR PURE TRANSMISSION, DELAYED TRANSMISSION, AND LINEAR PREDICTION OF SIGNALS[†]

III. CONVERGENCE PHENOMENA OF THE SOLUTIONS. EFFECT OF THE DISPERSION OF IMPULSES

3.0 OBJECTIVE OF THIS CHAPTER

In section 0.51, we indicated that an acceptable excitation function $\phi(t)$ can be decomposed into three component functions: a continuous function $\phi_1(t)$ of class C° in the interval $0 \leq t < \infty$; a discontinuous function $\phi_2(t)$ in the form of a step function which presents jumps of finite height on a countable set of isolated points in the interval $0 \leq t < \infty$; a singular function $\phi_3(t)$ which is zero for almost all points in the interval $0 \leq t < \infty$. The exceptional points form a countable set of isolated points where $\phi_3(t)$ has the same singular functions as $\phi(t)$. That is to say, $\phi_3(t)$ is composed of the impulses, doublets, and so forth, of $\phi(t)$.

3.01 The problem of convergence which we now propose to investigate can be described as follows. We consider the general problem of synthesis. First, we take the component $\phi_1(t)$ of the class C° associated with $\phi(t)$. With this component we determine the measures a_0, \dots, a_m of the distribution of windows, as we did in section 2.2, of the previous chapter. Now we propose to investigate the response of this system to the component excitations $\phi_2(t)$ and $\phi_3(t)$.

The study will be organized as follows:

- a. response of the system in the vicinity of an isolated impulse of $\phi(t)$;
- b. response of the system in the vicinity of an isolated jump of $\phi(t)$;
- c. response of the system to the excitation $\phi(t)$ in the interval $0 \leq t < t_\ell$, where t_ℓ is the duration of the life of the window distribution.

The responses of the systems to cases a, b, and c, constitute an intrinsic behavior of the system, producing convergence phenomena which may exceed the tolerance limits prescribed by the set $[\epsilon_j]$ mentioned in Chapter II. The result of this study will demonstrate that the apertures a_k , of the set $[a_k]$, are centered on the points which form the sets on which $\phi(t)$ is discontinuous, on which $\phi(t)$ is singular, and in the interval

[†](Continued from the Quarterly Progress Report of Oct. 15, 1956.) Translated by R. E. Wernikoff from the Spanish – with some corrections and additions that were made, particularly in the sections dealing with error analysis, by Mr. Wernikoff in cooperation with Dr. Cerrillo.

[Editor's note: This material, which was published under the title "Sobre la Sintesis de Sistemas Lineales para la Transmision sin Retraso, Retrasada, y Prediccion Lineal de Señales," in Revista Mexicana de Fisica, is an application of the theory given in Technical Report 270, by Dr. Cerrillo, "On Basic Existence Theorems. Part V. Window Function Distributions and the Theory of Signal Transmission" (to be published). The direct connection of the present paper with the work of the Statistical Communication Theory group and other groups in the Laboratory led to its translation, by Mr. Wernikoff, and its presentation here.]

(X. STATISTICAL COMMUNICATION THEORY)

$0 \leq t < t_\ell$. We shall also show that the minimum aperture is necessarily equal to t_ℓ , the life of the distribution.

3.02 In a finite system, especially one with lumped elements, impulses are dispersed. The impulse is transformed into a pulse of aperture a that is different from zero, and it may have lateral oscillations. For this reason, a system function $S_c(t)$ is not formed by an ideal distribution of a number of impulses of zero duration. The type of system function that we shall now discuss is illustrated in Fig. 1(I-1.4). In this chapter we propose to study the following question:

Suppose that the function $S_c(t)$ is not formed by impulses distributed over a set of measure zero, but instead, that each impulse is dispersed into a pulse of aperture δ containing an area equal to the measure a_k , $k = 0, 1, \dots, m$ of each impulse. Reference is made to Fig. 1(III-3.02) in which the dispersion in the distribution of impulses of Fig. 1(II-2.1) is shown. Now we ask, "What modification should be made or interpretation given in expression 2(II-2.22) which determines the measures a_k of a window distribution in the solution of a given problem?" As a result of this study, it will be shown that the expression 2(II-2.22) also produces the areas of the windows in the dispersed kernel, with the additional condition that the modulus of oscillation of the continuous component $\phi_1(t)$ of $\phi(t)$ be small in the interval t_ℓ of the life of the distribution $S_c(t)$.

3.1 RESPONSE IN THE VICINITY OF AN ISOLATED IMPULSE OF $\phi(t)$

We consider the i^{th} impulse of the component $\phi_3(t)$ associated with $\phi(t)$. Let p_i be the measure of this impulse. The distribution function $\beta(t)$ associated with this impulse is a step of height p_i , occurring at time t_i . The response of the system characterized by the system $S_c(t)$ will be given by the convolution integral that has already been described. The answer, $r_i(t)$, will be given by

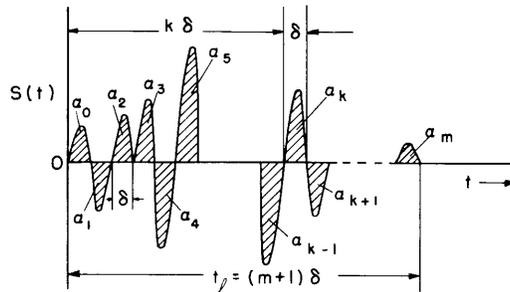


Fig. 1(III-3.02). Dispersion of the impulse distribution shown in Fig. 1(II-2.1).

$$r_i(t) = \int_0^t S_c(t-\tau) d\beta_i(\tau) = p_i S_c(t - t_i) \quad 1(\text{III-3.1})$$

The response is, therefore, the system function itself displaced by a time t_i from the origin and amplified p_i units. The duration of the response will evidently be t_ℓ units of time.

In virtue of the singular character of the impulse of excitation, the result is equally correct when $S_c(t)$ is formed by dispersed impulses. Figure 1(III-3.1) is a sketch, not to scale, of this last type of response.

Let t_i and t_{i+1} be the times of occurrence of two successive impulses of $\phi_3(t)$. If $t_{i+1} - t_i \geq t_\ell$, then the responses to each impulse do not overlap. For $t_{i+1} - t_i < t_\ell$, the response is obtained by superposition of the individual responses.

3.2 RESPONSE IN THE VICINITY OF AN ISOLATED JUMP OF $\phi_2(t)$

In this case the effect of the function $S_c(t)$ is different but similar, depending on whether or not this function is singular. With reference to Figs. 1(II-2.1) and 1(III-3.02), if we apply the convolution integral, we obtain immediately the responses in both cases. Let J_k be the measure of the jump of the function $\phi_2(t)$, occurring at the time t_k . Then

$$R_k(t) = \int_0^t \phi_2(t-\tau) da(\tau) = J_k \int_0^{t-t_k} da(\tau)$$

since $\phi_2(t) = 0$ for $t < t_k$, and $\phi_2(t) = J_k$ for $t > t_k$.

Figure 1(III-3.2) shows both answers. The answer in the form of a step refers to

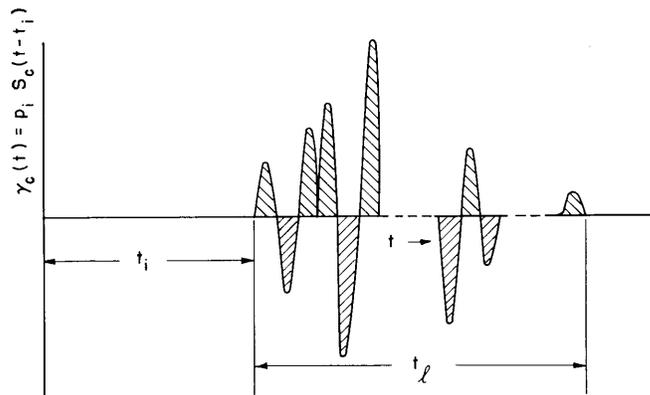


Fig. 1(III-3.1). Response of the system to an impulse of measure p_i applied at $t = t_i$.

(X. STATISTICAL COMMUNICATION THEORY)

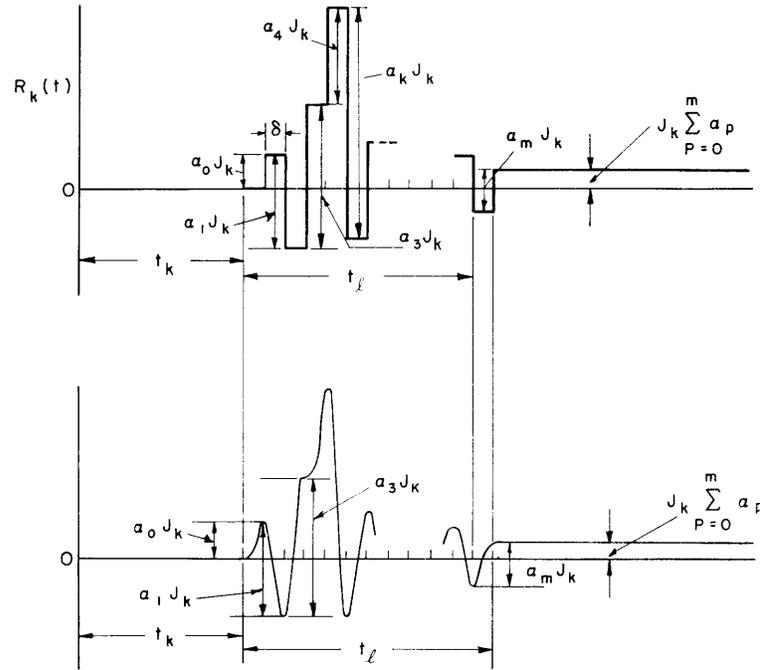


Fig. 1(III-3.2). Response of the distribution $S_C(t)$, for singular and nonsingular cases, to an isolated jump of $\phi_2(t)$ of height J_k applied at $t = t_k$.

the singular kernel. The answer for the continuous kernel is also continuous and rests on the steps of the step function that corresponds to the answer in the singular case. Note that this answer starts out at zero for $t < t_k$ and fluctuates in the interval $(t_k, t_k + t_l)$, remaining constant for $t \geq t_k + t_l$.

The response corresponding to each jump should be added, taking into account its displacement in time. When two successive jumps of $\phi_2(t)$ occur with a time difference between them that is greater than t_l , then the fluctuations caused by the first jump do not interfere with those of the second jump, except for a constant value.

3.3 DETERMINATION OF THE MEASURES OF THE WINDOWS OF NONZERO APERTURE

We propose to evaluate the measures a_0, \dots, a_m of a distribution of windows for the case in which each window covers an aperture $\delta > 0$. It will be shown that expressions 2(II-2.22) remain the same in the case of dispersion when the following condition is satisfied:

$\Omega(\delta)$ of $\phi(t)$ is small compared with the variation of $S_C(t)$
in the same interval.

Strictly speaking, there exists another condition which will not be repeated because

it has already been accepted in the statement of the synthesis problem formulated in section 0.6. This condition consists in admitting as solutions functions $\gamma^*(t)$ that are similes of $\gamma(t)$. We have to prove, then, that, within certain approximations which are consistent with the substitution of $\gamma^*(t)$ for $\gamma(t)$, the evaluation of the measures of the windows in the case of impulses is the same as the evaluation in the case of windows of nonzero aperture.

The formulas 2(II-2.22) are solutions for $t \geq t_\ell$. Here we consider the following situation: 1. Determination of the a 's for $t \geq t_\ell$. 2. Effect of the kernel so determined in the time $0 \leq t < t_\ell$. The second situation covers problem c of section 3.01. We shall use a graphical illustration of the process of evaluation that shows clearly the mechanism of formation of the function $\gamma^*(t)$ which results from the continuous component $\phi_1(t)$ of $\phi(t)$.

3.31 Since $\phi_1(t)$ and $S_c(t)$ are now, by hypothesis, continuous functions, the convolution integral exists in the sense of Riemann. We shall begin by indicating graphically the operation of the convolution integral, a process which is well known. Let $\phi_1(t)$ belong to the class C° . Suppose, for the moment, that the function $S_c(t)$ is a known window distribution with finite life t_ℓ . We shall consider the following cases:

1. $t < 0$
2. $0 < t < t_\ell$
3. $t > t_\ell$

to indicate the mechanism of formation of the integral

$$\gamma(t) = \int_0^t \phi_1(t-\tau) S_c(\tau) d\tau \quad 1(\text{III-3.31})$$

Figure 1(III-3.31) shows the position of the factors of the integrand for $t < 0$. Naturally, $\gamma(t)$ is necessarily zero in this interval. The hypothesis of slow oscillation of $\phi_1(t)$ in the interval δ is expressed graphically by the smooth form of the curve in the figure. Figure 2(III-3.31) shows the situation for $t = 0$. The value of $\gamma(t)$ is still zero. Figure 3(III-3.31) shows the situation of the functions of the integrand in the interval $0 < t < t_\ell$. We already have a contribution to $\gamma(t)$, but not all windows of $S_c(t)$ contribute to the value of $\gamma(t)$. Figure 4(III-3.31) shows the situation of the functions of the integrand in the interval $t > t_\ell$. All the windows of $S_c(t)$ now contribute to the formation of the integral. Henceforth, all the windows produce contributions, but only the part of $\phi(t-\tau)$ which is on top of $S_c(t)$ produces a contribution. The effect of this partial use of $\phi(t)$ produces certain indications in the synthesis problem, as will be seen later.

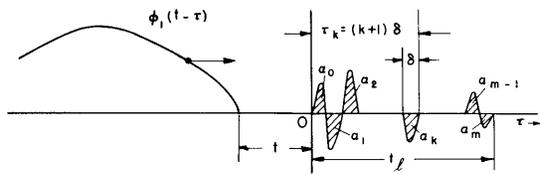


Fig. 1(III-3.31). The integrand of expression 1(III-3.31) for $t < 0$.

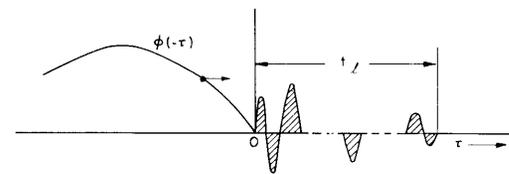


Fig. 2(III-3.31). The integrand of expression 1(III-3.31) for $t = 0$.

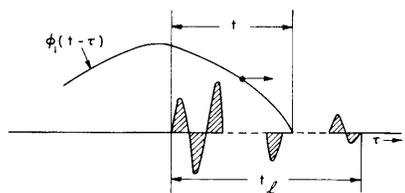


Fig. 3(III-3.31). The integrand of expression 1(III-3.31) for $0 < t < t_l$.

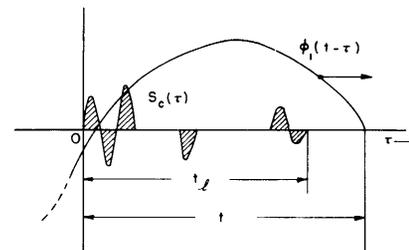


Fig. 4(III-3.31). The integrand of expression 1(III-3.31) for $t_l < t < \infty$.

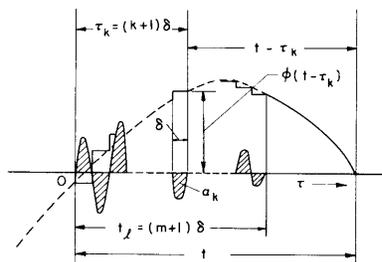


Fig. 1(III-3.32). Substitution of a step function for $\phi_1(t)$.

3.32 EVALUATION OF THE MEASURES a FOR $t > t_\ell$

The hypothesis of slow variation of $\phi_1(t)$ allows us to suppose that $\phi_1(t)$ remains approximately constant in each interval δ . This is equivalent to substituting the step function indicated in Fig. 1(III-3.32) for $\phi_1(t)$. The hypothesis, which allows the substitution of $\gamma^*(t)$ for $\gamma(t)$, makes possible this approximation whenever the tolerances specified by the set $[\epsilon_j]$, given as a datum of synthesis, are not exceeded. The value of the integral 1(III-3.31) is obtained immediately by using the step function that rests on $\phi_1(t)$. Let us call $\gamma^*(t)$ the value of the convolution integral obtained by this substitution. We have

$$\begin{aligned} \gamma^*(t) &= \sum_{k=0}^m \phi_1(t - \tau_k) \int_{k\delta}^{(k+1)\delta} S_c(\tau) d\tau \\ &= \sum_{k=0}^m a_k \phi[t - (k+1)\delta] \quad t > t_\ell \end{aligned} \quad 1(\text{III-3.32})$$

We see immediately that the approximate solution 1(III-3.32), which gives us $\gamma^*(t)$, is equal to the expression which produces the singular kernel formed by a distribution of impulses of zero aperture. See expression 2(II-2.1).

Now, expanding the function $\phi_1[t - (k+1)\delta]$, which is now considered to be of class $C^{(m)}$, and forming the function $\gamma^*(t)$ that is associated with Eq. 1(III-3.32), we see that this function is identical with the function $\gamma^*(t)$ that is extracted from expression 3(II-2.11). This proves the equivalence of an impulsive kernel and a continuous kernel in the sense of the formulation of the synthesis problem given in section 0.6. We can conclude from this that formula 2(II-2.22), which determines the measures of the windows of $S_c(t)$, is equally applicable to the continuous kernel, with the condition of slow variation of $\phi_1(t)$ in the interval δ .

3.4 RESPONSE OF THE SYSTEM DETERMINED BY THE a 's IN THE INTERVAL $0 < t < t_\ell$

This answer is obtained easily if we neglect the fine structure and just consider the basic behavior. Figure 3(III-3.31) immediately allows us to make this evaluation by substituting a step function for $\phi(t)$. We have

$$\gamma^*(t) = \sum_{k=0}^q a_k \phi[t - (k+1)\delta] \quad 0 < t < t_q \quad 1(\text{III-3.4})$$

An expansion similar to the one in section 2.11 now produces the expression

(X. STATISTICAL COMMUNICATION THEORY)

$$\begin{aligned} \gamma^*(t) \approx & \phi(t) \sum_{k=0}^q a_k - \phi^{(1)}(t) \sum_{k=0}^q \mu_k a_k + \frac{\phi^{(2)}(t)}{2!} \sum_{k=0}^q \mu_k^2 a_k + \dots \\ & + (-1)^m \frac{\phi^{(m)}(t)}{m!} \sum_{k=0}^q \mu_k^m a_k \quad 0 < t \leq t_q \leq t_\ell \end{aligned} \quad 2(\text{III-3.4})$$

Of course, this solution does not coincide with expression 2(II-2.11) until $t_q = t_\ell$. Therefore, in the interval $0 < t < t_\ell$, the system function $S_c(t)$ determined by expression 2(II-2.22) does not necessarily reproduce the desired response $\gamma^*(t)$ for the output of the system. This implies that the interval belongs, in general, to the aperture set $[a_j]$.

3.5 OTHER INTERVALS OF NON-REPRESENTATION

We have indicated that neighborhoods of width t_ℓ , centered on the points of discontinuous or impulsive behavior of $\phi(t)$, and the initial interval $0 < t < t_\ell$, compose part of the aperture set $[a_j]$ of non-representation.

In addition, other intervals of non-representation exist that depend on the continuous behavior of $\phi(t)$ when the window distribution has elements of nonzero aperture. Among these intervals of non-representation we have the following two:

1. Suppose that $\phi(t)$ is identically zero on a set of disjoint intervals separated by distances greater than t_ℓ . Then the beginnings and endings of these intervals belong to the set $[a_j]$.

2. Suppose that $\phi(t)$ has intervals of the order of magnitude δ , in which $\phi(t)$ has a rapid variation comparable to that of $S_c(t)$.[†] Then these intervals belong to the set $[a_j]$.

Heretofore, we have spoken of slow and rapid variation of $\phi(t)$. A quantitative measure of the rapidity of these variations will be given in a later chapter.

IV. THEOREM OF CLASS CONSERVATION WITH DEGENERATE KERNELS AND FINITE LIFE

4.0 OBJECTIVE OF THIS CHAPTER

A basic question proposed in the synthesis problem formulated in section 0.6 requires the investigation of the possible class connections that may exist between the excitation function and the response of a finite, linear, passive, four-terminal system.

[†]The very fast variations of $\phi(t)$ can be compared to jumps, impulses, and so forth; hence this statement.

In fact, the basic understanding of the proposed synthesis problem resides fundamentally in the determination of these connections between the classes of excitation functions and response functions. For example, in the transmission problem it is necessary to preserve both the class of the function and the function itself.

The study of this determination of classes is very difficult, at least for the author. The use of singular distributions of the impulsive type allows us some insight into this problem. For example, if we consider kernels of finite life, expression 2(II-2.1), which is valid for $t > t_\ell$, strongly suggests that the linear, finite, passive, four-terminal system conserves the class of functions, since the response is expressed by a finite linear combination of functions of the class of the excitation function. This inference, however, is not completely correct because of the existence of a set of apertures $[a_j]$, in which the behavior of $\gamma(t)$ is not necessarily represented by the expression 2(II-2.1).

In some special cases, especially with degenerate kernels of finite life, it is possible to establish the conservation of class between excitation and response functions, when the former are continuous. The object of this chapter is to give certain results which tend to clarify the problem of class preservation.

4.1 DEGENERATE CONTINUOUS KERNELS. FINITE LIFE

Let us consider the convolution integral

$$\gamma(t) = \int_0^t \phi(t-u) S_c(u) du \quad 1(\text{IV-4.1})$$

and let us suppose that the kernel $\phi(t-u)$, which is formed by the excitation, and also the system function $S_c(t)$, are continuous in the interval $0 < t < \infty$. In addition, $S_c(t)$ will be supposed to be a finite distribution of windows, of duration t_ℓ . The hypothesis that the kernel is degenerate allows us to write

$$\phi(t-u) = \sum_{p=1}^n A_p \beta_p(u) \lambda_p(t) \quad 2(\text{IV-4.1})$$

The initial excitation is obtained by setting $u = 0$, so that

$$\phi(t) = \sum_{p=1}^n A_p \beta_p(0) \lambda_p(t)$$

Since the initial excitation is not zero, it follows that not all products $A_p \beta_p(0)$ are zero.

Degenerate kernels appear in many practical and important types of excitation. For example, excitations formed by exponential functions

(X. STATISTICAL COMMUNICATION THEORY)

$$\phi(t) = \sum_{p=0}^n A_p e^{\phi_p t}$$

where A_p and ϕ_p are real or complex constants.

Substituting expression 2(IV-4.1) in the integral 1(IV-4.1), we obtain

$$\gamma(t) = \sum_{p=0}^n A_p \lambda_p(t) h_p(t) \tag{3(IV-4.1)}$$

where

$$h_p(t) = \int_0^t S_c(u) \beta_p(u) du$$

Let us now introduce the condition of finite life for the distribution $S_c(t)$. That is to say,

$$S_c(t) \begin{cases} \equiv 0 & t < 0 \\ \neq 0 & 0 < t < t_\ell \\ \equiv 0 & t_\ell < t < \infty \end{cases}$$

Then

$$h_p(t) \begin{cases} \equiv 0 & t < 0 \\ \neq 0 & 0 < t < t_\ell \\ = h_p(t_\ell) = \text{constant} & t_\ell < t < \infty \end{cases}$$

We conclude that

$\gamma(t)$ is of the same class as $\phi(t)$ for $0 < t_\ell < t$

$\gamma(t)$ is not necessarily of the same class as $\phi(t)$ for $0 < t < t_\ell$

since $h_p(t)$ is a function of time.

4.2 EXTENSION TO DEGENERATE KERNELS. INFINITE LIFE

The previous results can be extended to the case of a system function formed by a window distribution with infinite life, when certain additional conditions that will be given below are satisfied. We also consider that n is finite. The expressions of the previous case are valid, making $t_\ell \rightarrow \infty$. With the condition that

$$h_p(t_\ell) \rightarrow c_p < \infty \quad \text{as } t_\ell \rightarrow \infty$$

we have

$$\gamma(t) = \sum_{p=1}^n A_p c_p \lambda_p(t)$$

and we again preserve the class. (It should be noted that the idea of class conservation is a steady-state concept. Questions of class conservation can be considered only after all transient phenomena have died out, so that, in general, conservation relations obtain only after $t \rightarrow \infty$.)

4.3 SOLUTION OF THE PROPOSED SYNTHESIS PROBLEM WHEN CLASS CONSERVATION IS STIPULATED. FINITE LIFE

We indicate here a formal solution to the synthesis problem that was proposed in section 0.6, when we stipulate class conservation between the excitation and response functions, which are here supposed to be continuous.

Let us write, as before,

$$\phi(t-u) = \sum_{p=1}^n A_p \beta_p(u) \lambda_p(t) \quad \text{kernel}$$

$$\phi(t) = \sum_{p=1}^n A_p \beta_p(0) \lambda_p(t) \quad \text{excitation}$$

The response can be given by

$$\gamma(t) = \sum_{p=1}^n B_p \lambda_p(t)$$

in virtue of class conservation. It suffices to set

$$A_p h_p(t_\ell) = B_p$$

and, in order to simplify, we shall write

$$h_p(t_\ell) = \text{constant} = c_p < \infty$$

Therefore, the problem is reduced to determining the function $S_c(t)$ in such a manner that

$$c_p = \int_0^{t_\ell} S_c(u) \beta_p(u) du$$

It will be sufficient to form the function $S_c(t)$ as an orthogonal expansion of functions

(X. STATISTICAL COMMUNICATION THEORY)

$\beta_p(t)$ in the interval $(0, t_f)$. By a well-known theorem, we have as the solution

$$S_c(t) = \sum_{p=1}^n c_p \beta_p(t)$$

A discussion of this type of solution is beyond the scope of this paper. For this reason, it is presented here only as a formal solution.

V. TRANSMISSION WITHOUT DELAY

5.0 OBJECTIVE OF THIS CHAPTER

We propose to study in some detail the system function $S_c(t)$, already obtained, for the transmission without delay of a signal $\phi(t)$. We construct window distributions of nonzero apertures, since these represent the solution of linear, passive, and finite systems.

By solution of transmission without delay we shall understand a response of the form $\phi^*(t)$ caused by an excitation $\phi(t)$ when the system function $S_c(t)$ is determined from the continuous component of $\phi(t)$.

5.1 GENERAL FORMULA

For convenience we shall repeat the formula, already given in section 2.3, for the measures a_j of the appropriate window distribution. These measures are given by

$$a_j = (-1)^m \frac{\prod_{p=0}^m \mu_p}{\prod_{p=0}^m (\mu_j - \mu_p)} \quad j = 0, 1, \dots, m \quad 1(V-5.1)$$

In this formula μ_j represents the distance from the origin to the end of the j^{th} window. Although the formula was obtained specifically for the case in which the apertures of the windows are equal, it is easy to see that it is equally valid when the windows do not have equal apertures, in which case μ_j is interpreted as the distance from the origin to the end of the j^{th} window. The width of this window will then be $\mu_j - \mu_{j-1} = a_j$. The case of windows of constant aperture is easily synthesized by the system shown in Fig. 2(I-1.51). For this reason we shall discuss here window distributions with constant aperture.

5.2 WINDOW DISTRIBUTION OF CONSTANT APERTURE

In this case

$$\mu_k = (k+1)\delta \quad 1(V-5.2)$$

Substituting these values in expression 1(V-5.1), and remembering the definition of Newton's binomial coefficients, we obtain

$$a_j = (-1)^j \frac{(m+1)!}{(j+1)! (m-j)!} = (-1)^j \binom{m+1}{j+1} \quad 2(V-5.2)$$

Let us now consider windows with aperture $\delta > 0$. Then the a_j represent the areas of the pulses that represent each window. Since the pulses have equal apertures, their areas are proportional to their heights, when we suppose that the pulses have similar shapes.

As we have seen, the window distribution that forms the system function is defined only by the areas of the windows and not by their specific shapes. We shall use a conventional form of window in their representation. As will be shown in this chapter, the major effect of the windows is attributable to their area. Their specific form only influences the fine structure of the response function.

5.3 DISTRIBUTIONS CORRESPONDING TO m VARIABLES. ENVELOPE

We calculate here the window distributions for different values of m . These distributions are indicated in Fig. 1(V-5.3). The width δ is, for the moment, irrelevant. The sketches of the windows are not necessarily to scale. It should be noted that the distribution is formed according to the table

				1				
				2	-1			
			3	-3	1			
		4	-6	4	-1			1(V-5.3)
	5	-10	10	-5	1			
6	-15	20	-15	6	-1			
7	-21	35	-35	21	-7	1		

where the initial term of the binomial coefficients, 1, is always omitted. The suppression of this term is important, because if it is added to the table we shall have a no-pass system which will reject the signal $\phi(t)$ instead of transmitting it. These kernels will be discussed later.

(X. STATISTICAL COMMUNICATION THEORY)

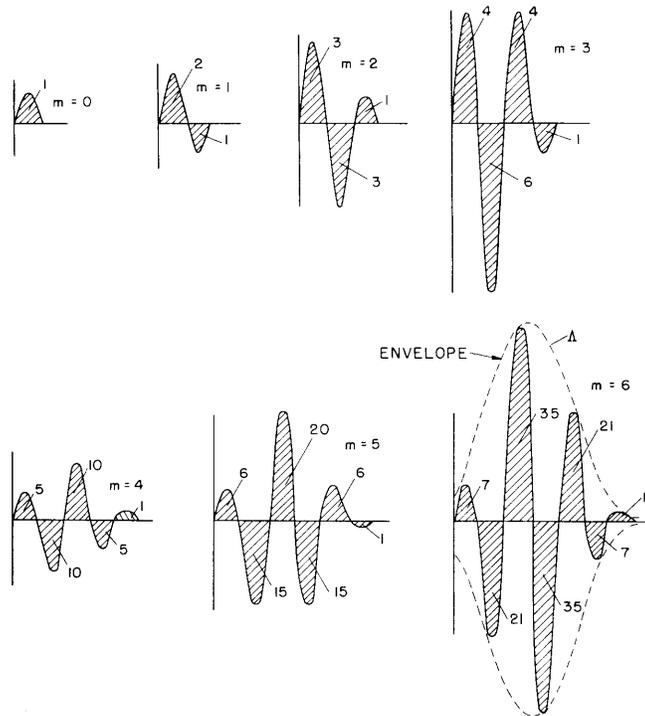


Fig. 1(V-5.3). Distributions associated with pure transmission.

Finally, we shall give the expression of the envelope function of the peaks of the distributions. It is easy to prove that this envelope is

$$\Lambda(x) = \pm \frac{m-x+1}{B(m-x, x)} \quad 0 \leq x \leq m \quad 2(V-5.3)$$

where $B(m-x, x)$ is the beta function normalized to the aperture δ .

5.4 INTERPRETATION OF RESULTS

We propose to interpret these distributions in terms of the mechanism of their operation in order to produce the transmission of a function $\phi(t)$.

The graphical method is simple and clear. Let us take $m = 0$. The distribution consists of a single window of area one and aperture δ , placed at the origin. Figure 1(V-5.4) shows the integrand of the convolution integral. Note, in the figure, that the condition of slow variation of $\phi(t)$ in the interval δ is satisfied. It is clear that

$$\gamma(t) = \int_0^\delta \phi(t-u) S_c(u) du \quad 1(V-5.4)$$

and

$$\gamma^*(t) = \phi(t-\delta) \int_0^\delta S_c(u) du = \phi(t-\delta)$$

when we take the step indicated in the figure. We see that

1. $\gamma^*(t)$ presents a delay δ with respect to $\phi(t)$.
2. As we make the aperture smaller and smaller, $\gamma^*(t)$ approaches $\phi(t)$, which is the limit reached as $S_c(t)$ becomes an impulse at the origin.

The situation indicated for $m = 0$ represents the classical solution to the transmission problem. The transmission for $m > 0$ constitutes the generalization introduced by this paper.

5.41 The adoption of the term "window" may be justified by the effect in the integral of the pulse of $S_c(t)$ in Fig. 1(V-5.4), since it is equivalent to taking the value $\phi(t-\delta)$ of the function $\phi(t)$ "seen" through the "window" of the kernel. In American usage, the function $S_c(t)$ is also called "scanning function" because of the scanning effect on $\phi(t)$.

5.42 To interpret the mechanism of these distributions when $m \geq 1$, it is convenient to study, first, the effect of just one window of this distribution, let us say the k^{th} , which is displaced $k\delta$ units from the origin. Figure 1(V-5.42) shows the integrand which produces the k^{th} window's contribution to $\gamma^*(t)$. Following the same line of reasoning as before, we find that the contribution of the k^{th} window is

$$\gamma_k^*(t) = a_k \phi[t - (k+1)\delta] \quad 1(\text{V-5.42})$$

It follows that the k^{th} window of the kernel produces the excitation function delayed by $(k+1)\delta$ units and amplified a_k times.

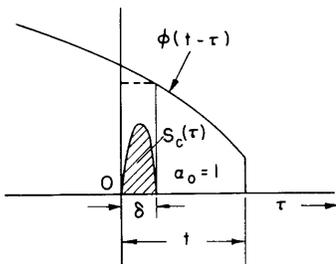


Fig. 1(V-5.4). Operation of $S_c(t)$ for $m = 0$.

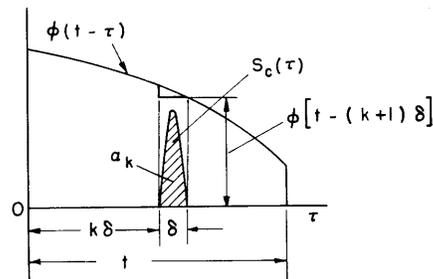


Fig. 1(V-5.42). Operation of $S_c(t)$ for $m = k$.

(X. STATISTICAL COMMUNICATION THEORY)

5.43 We can immediately write the effect of the whole window distribution that corresponds to the general case $m > 0$:

$$\gamma^*(t) = \sum_{k=0}^m a_k \phi[t - (k+1)\delta] \quad 1(V-5.43)$$

We see, then, that the response function $\gamma^*(t)$ is formed by a weighted sampling, with weights $a_0 \dots a_m$, of the function $\phi(t)$ in the interval from $t - (m+1)\delta$ to t .

It is also worth while to interpret the expression 1(V-5.43) from another point of view, one which is very useful in the theory of linear prediction. Let us form the following associations:

With pure transmission, without delay, let us associate the idea of "present."

With delayed transmission let us associate the idea of "past."

With advanced transmission let us associate the idea of "future."

In this language the expression 1(V-5.43), which produces pure transmission, can be interpreted by saying that the system forms a present which is composed of a weighted sampling of the past and a term from the present.

5.44 To close this section, let us compare the distributions corresponding to $m = 0$ and $m > 0$ from the point of view of section 5.43. The expressed idea now becomes precise if the window distribution is considered as a distribution of impulses.

For the case $m = 0$, the distribution contains only one impulse of measure one at the origin. The response function $\gamma(t) = \phi(t)$ is obtained simply by considering the function $\phi(t)$ at the precise instant t .

When $m > 0$, the distribution contains an impulse of measure a_0 at the origin and other impulses displaced from the origin. The pulse or window that corresponds to a_0 produces a contribution from the present of $\phi(t)$ at the time t . For the reasons given before, we attempt to interpret pure transmission as an "intrinsic present," $m = 0$, and "present with experience," $m > 0$. Accordingly, m is a measure of experience in the sense used here.

When the window distribution does not consist of impulses, then the first window produces a response delayed by δ units. That is to say, $\phi(t)$ at the time t is determined by the value of the function in the interval $t - \delta$ and t . This is interpreted by saying that the "perception" is not instantaneous. Thus δ could be called the minimum perception time.

5.5 CONVERGENCE TO $\phi(t)$

We now consider the convergence of the function $\gamma^*(t)$ to $\phi(t)$ in the special case of pure transmission. The discussion and results of Chapter III can be transposed without

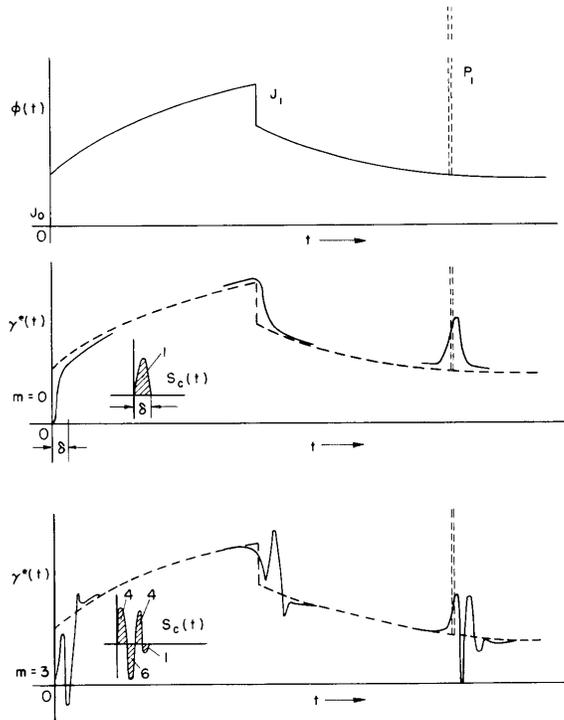


Fig. 1(V-5.5). Examples of convergence at jumps and impulses for pure transmission with $m = 0$ and $m = 3$.

difficulty to the case of pure transmission, whence we shall not make an extended study. The important point here is to observe the mode of convergence to $\phi(t)$ of solutions given by window distribution with $m = 0$ and $m > 0$. A simple graphical illustration is sufficient to clarify the situation. Let us take, for example, $m = 0$ and $m = 3$. The extension to other values of m is obvious. Let us also take as a typical excitation function the one indicated in Fig. 1(V-5.5), which contains a continuous component, a jump, and an impulse. The results of this study are clearly shown in the figure.

The lives of the functions $S_c(t)$ that are used in Fig. 1(V-5.5) are not equal. When the distributions have equal lives, transitions are produced which last for approximately equal times. This is true when $\phi(t)$ has as a component a continuous function with oscillations that are slow relative to the perception δ . For $m > 0$, there are always oscillations in the intervals which form the aperture set $[a_j]$.

Because of space limitations we do not illustrate the transitions in the case when $\phi(t)$ has doublets, triplets, and so forth, as singular components. In the case of a doublet it is easily shown that the graph of the transition phenomenon is the first derivative of the distribution $S_c(t)$ divided by the width of the doublet. In general, the transition caused by an n -tuple of small width equal to Δt is given approximately by

(X. STATISTICAL COMMUNICATION THEORY)

$$\text{Response} \approx \frac{\frac{d^n}{dt^n} S_c(t)}{(\Delta t)^n} \quad 1(V-5.5)$$

Because of the small value of Δt and because of the rapid variation of $S_c(t)$, these responses represent oscillations of very large amplitude.

5.6 DETERMINATION OF THE ERROR

Let us consider the error expression given by formula 1(II-2.8). Considering formula 2(V-5.2) and a known property of the binomial coefficients, we arrive immediately at the result

$$\sum_{p=0}^m (p+1)^{m+1} a_p = \sum_{p=0}^m (-1)^p (p+1)^{m+1} \binom{m+1}{p+1} = (-1)^m (m+1)! \quad 1(V-5.6)$$

Therefore,

$$|\xi_{(m), k}| \approx \Omega_k^{(m)}(\delta) \delta^m \quad 2(V-5.6)$$

which gives the error incurred in the k^{th} interval of continuity of the excitation function $\phi(t)$ as a function of the order m of the window distribution, the aperture δ of a window, and the modulus of oscillation of the m^{th} derivative of the excitation in the k^{th} interval of continuity mentioned above. Formula 2(V-5.6) indicates that the error decreases very rapidly with decreasing aperture δ . It can also be shown that the error decreases with increasing m . The intrinsic transmission, $m = 0$, is associated with the error

$$\xi_{(0), k} = \Omega_k^{(0)}(\delta)$$

An application of this formula to the design of a linear system will be given in a later section.

We shall close this section by stating important properties of the binomial coefficients, properties that can be proved immediately by using the results of this chapter. Since pure transmission is characterized by the values

$$y_p = \begin{cases} 1 & p = 0 \\ 0 & p = 1, 2, \dots, m \end{cases}$$

we obtain

$$\sum_{k=0}^m (-1)^k (k+1)^q \binom{m+1}{k+1} = \begin{cases} 1 & \text{when } q = 0 \\ 0 & \text{when } q = 1, 2, \dots, m \\ (-1)^m (m+1)! & \text{when } q = m + 1 \end{cases}$$

5.7 LAPLACE TRANSFORM OF $S_c(t)$ FOR PURE TRANSMISSION

The transforms will be determined for the case in which the graphs of the windows are similar curves. Let us consider a window of measure one and of given form located at the origin. Let us call this function $h(t)$, and denote by $H(s)$ the corresponding Laplace transform. Let us consider a function $S_c(t)$ formed by a distribution of similar windows of measures a_0, \dots, a_m . Let

$$\mathcal{L}[S_c(t)] = S_c(s) \quad (s \text{ is a complex variable})$$

then

$$S_c(s) = H(s) \sum_{p=0}^m a_p e^{-p\delta s} \quad 1(V-5.7)$$

For the case of pure transmission, using formula 2(V-5.2), we have

$$S_c(s) = H(s) \sum_{p=0}^m (-1)^p \binom{m+1}{p+1} e^{-p\delta s} \quad 2(V-5.7)$$

To sum this expression, it is enough to transform it as follows:

$$\begin{aligned} S_c(s) &= -H(s) e^{\delta s} \sum_{p=0}^m (-1)^{p+1} \binom{m+1}{p+1} e^{-(p+1)\delta s} \\ &= H(s) e^{\delta s} \{1 - (1 - e^{-\delta s})^{m+1}\} \\ &= H(s) e^{\delta s} \left\{ 1 - 2^{m+1} e^{-\frac{m+1}{2}\delta s} \sinh^{m+1} \left(\frac{\delta s}{2} \right) \right\} \end{aligned} \quad 3(V-5.7)$$

The previous expression takes on a simple form when the function $S_c(t)$ is formed by impulses with the distribution indicated in Fig. 1(II-2.1). In this case

$$H(s) = e^{-s\delta}$$

and

$$S_c(s) = 1 - 2^{m+1} e^{-\frac{m+1}{2}\delta s} \sinh^{m+1} \left(\frac{\delta s}{2} \right) \quad 4(V-5.7)$$

This last expression is very useful for illustrating the mechanism of pure transmission in the frequency domain.

(X. STATISTICAL COMMUNICATION THEORY)

5.8 FREQUENCY SPECTRUM FOR PURE TRANSMISSION

We shall begin by determining the spectrum associated with the expression 4(V-5.7). We have, setting $s = i\omega$

$$S_c(\omega) = 1 - \left[2 \sin \frac{\delta\omega}{2} \right]^{(m+1)} \exp \left[-\frac{i}{2} (m+1)(\delta\omega - \pi) \right] \quad 1(V-5.8)$$

We see immediately that the constant term, which is equal to 1, is responsible for pure transmission without delay. The variable term appears because of the substitution of $\gamma^*(t)$ for $\gamma(t)$. The components of the spectrum are

$$\begin{aligned} |S_c(\omega)|^2 &= 1 + \left(2 \sin \frac{\delta\omega}{2} \right)^{2(m+1)} - 2^{m+2} \sin^{m+1} \frac{\delta\omega}{2} \cos \left[\frac{m+1}{2} (\delta\omega - \pi) \right] \\ \tan \Psi_c(\omega) &= \frac{\left(2 \sin \frac{\delta\omega}{2} \right)^{m+1} \sin \left[\frac{m+1}{2} (\delta\omega - \pi) \right]}{1 - \left(2 \sin \frac{\delta\omega}{2} \right)^{m+1} \cos \left[\frac{m+1}{2} (\delta\omega - \pi) \right]} \end{aligned} \quad 2(V-5.8)$$

where $\Psi_c(\omega)$ is the phase angle.

5.9 LOW-PASS FILTER WITHOUT DELAY

The preceding expressions, which give us the frequency spectrum $S_c(s)$, are too complicated to be interpreted readily. We can, however, see that, for relatively small values of $\omega\delta$, the function $S_c(\omega)$ is practically constant and real, passing frequencies in

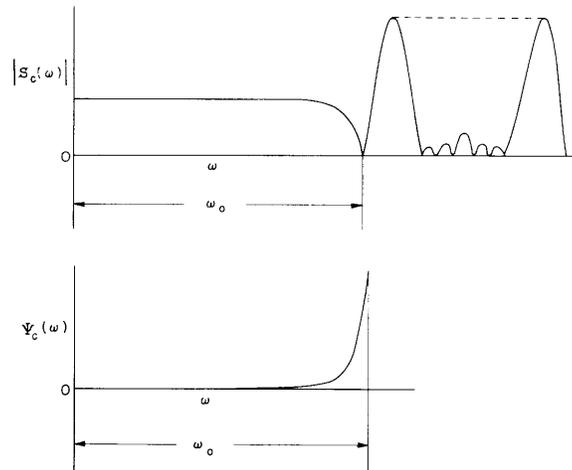


Fig. 1(V-5.9). Typical form of the functions $|S_c(\omega)|$ and $\Psi_c(\omega)$ of a pure transmission system. Note the low-pass filter effect.

the range $0 < \omega\delta < 0.1$ almost without delay. Figure 1(V-5.9) shows (not to scale) the typical and fundamental form of the functions $|S_c(\omega)|$ and $\Psi_c(\omega)$. We note, in particular, that a linear passive system that transmits without delay, in the sense of this chapter, behaves fundamentally like a low-pass filter, in which transmission is produced without appreciable phase change.

A filter of this type can be easily synthesized by the means indicated in Fig. 2(I-1.51). It is sufficient to adjust the amplifiers and inverters in such a way that a window distribution whose measures are given by formula 2(V-5.2) is obtained. The expression which determines the bandwidth ω_0 will be given later.

5.10 COMPLEMENTARY FILTERS

Let us now take a distribution consisting of $m + 2$ windows and use the $m + 1$ measures

$$a_j = (-1)^j \binom{m+1}{j+1}$$

that correspond to the windows for pure transmission, and let us add to these windows a window of measure +1 as the initial window. This distribution is represented by the table

$$\begin{array}{cccccc}
 & & & & 1 & -1 \\
 & & & & 1 & -2 & 1 \\
 & & & 1 & -3 & 3 & -1 \\
 & 1 & -4 & 6 & -4 & 1 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots
 \end{array} \tag{V-5.10}$$

Let us call β_g , $g = 0, 1, \dots, m + 1$, the measures of the new distribution. We have

$$\sum_{g=0}^{m+1} \beta_g = 1 - \sum_{p=0}^m (-1)^p \binom{m+1}{p+1} = 0 \tag{V-5.10}$$

in virtue of a known property of the binomial coefficients. This result shows immediately that a system characterized by this distribution cannot transmit, since it is equivalent to setting

$$y_0 = \sum_{g=0}^{m+1} \beta_g = 0$$

In virtue of the property

(X. STATISTICAL COMMUNICATION THEORY)

$$y_p = \sum_{g=0}^{m+1} (g+1)^p \beta_g = 0 \quad (p = 0, 1, \dots, m-1)$$

$$y_m = \sum_{g=0}^{m+1} (g+1)^m \beta_g \neq 0$$

we see that this system does not transmit the first $m - 1$ derivatives either, and will only start transmitting derivatives from the m^{th} order derivative.

We propose to show that a linear system so formed constitutes the filter which is complementary to the filter formed by the system indicated in section 5.9. It will be sufficient to determine the frequency spectrum. The Laplace transform of the distribution 2(V-5.10) is

$$G_c(s) = H(s) e^{sP} (1 - e^{-s\delta})^{m+1}$$

$$F(s) = H(s) e^{sP} \left(2 \sinh \frac{s\delta}{2}\right)^{m+1} e^{-\left(\frac{m+1}{2}\right)\delta s} \quad 3(V-5.10)$$

where

$$F(s) = \mathcal{L}[\text{system function}]$$

For a singular distribution like the one indicated in Fig. 1(II-2.1) we have

$$H(s) e^{sP} = 1$$

so that

$$F(s) = \left(2 \sinh \frac{s\delta}{2}\right)^{m+1} e^{-\left(\frac{m+1}{2}\right)\delta s} \quad 4(V-5.10)$$

If we add Eqs. 4(V-5.1) and 4(V-5.7), we obtain

$$S_c(s) + F(s) = 1 \quad 5(V-5.10)$$

which demonstrates that the filters associated with the window distributions that are given by formulas 1(V-5.10) and 2(V-5.2) are complementary. The complementary filter rejects in the band ω_o , which is the passband of the original filter, the function $\phi(t)$ and its first $m - 1$ derivatives. From the point of view of the frequency domain, the complementary filter is a highpass filter.

5.11 EXAMPLE OF THE TRANSMISSION OF A PERIODIC WAVE

The example we are about to give illustrates in a simple manner various aspects of the transmission problem and clarifies many of the concepts that will be introduced.

From the practical point of view, the problem has some importance.

Let us suppose that we want to transmit a periodic wave without delay and that we want to transmit all harmonics up to order n . The error of transmission should be less than some preassigned fixed number ϵ . For simplicity, we shall suppose that the wave which is to be transmitted is continuous. Let

$$f(t) = A_0 + \sum_{n=1}^{\infty} A_n \sin[n\omega_0 t + \phi_n] \quad 1(V-5.11)$$

Consider the transmission of the n^{th} harmonic,

$$f_n(t) = A_n \sin[n\omega_0 t + \phi_n] \quad 2(V-5.11)$$

Let us consider the m^{th} derivative of this harmonic. Its maximum modulus of oscillation is

$$\Omega^{(m)} = 2A_n (n\omega_0)^m \quad 3(V-5.11)$$

Then we must choose the numbers m and δ in such a way that the condition

$$2A_n (n\omega_0)^m \delta^m \leq \epsilon \quad 4(V-5.11)$$

is satisfied. The upper limit will be

$$(n\omega_0 \delta)^m = \frac{\epsilon}{2A_n} \quad 5(V-5.11)$$

or

$$(\delta f_n)^m = \frac{\epsilon}{2A_n (2\pi)^m}$$

where f_n is the maximum frequency to be transmitted. The tolerance ϵ is most naturally expressed as a fraction of the amplitude of the wave to be transmitted. That is, $\epsilon = k A_n$, so that

$$(\delta f_n)^m = \frac{k}{2(2\pi)^m} \quad 6(V-5.11)$$

Let T_n be the period $1/f_n$ of the n^{th} harmonic. It is easy to see that

$$T_n \geq 4\delta \quad 7(V-5.11)$$

Let us choose $\delta f_n = \frac{1}{8}$, so that

$$\delta = \frac{1}{8f_n} \quad 8(V-5.11)$$

(X. STATISTICAL COMMUNICATION THEORY)

Therefore,

$$\left(\frac{1}{4}\right)^m = \frac{k}{2\pi^m}$$

and so

$$m = \frac{\log \frac{k}{2}}{\log \frac{\pi}{4}} \tag{V-5.11}$$

As a numerical example, let us suppose $k = 0.02$ and $f_n = 10^4$ cps. Then

$$\delta = \frac{1}{8 \times 10^4} \text{ seconds}$$

$$m \approx 10$$

Formulas 8(V-5.11), 9(V-5.11), and 2(V-5.2) determine completely the window distribution that is capable of transmitting the wave under consideration. The corresponding linear system is obtained by adjusting the amplifiers and inverters of Fig. 2(I-1.51).

M. V. Cerrillo

[Chapters 6, 7, and 8 will be published in the Quarterly Progress Report, April 15, 1957.]