

VII. STATISTICAL COMMUNICATION THEORY

Prof. Y. W. Lee
 Prof. A. G. Bose
 Dr. C. S. Lorens

M. Brilliant
 J. Y. Hayase
 K. L. Jordan, Jr.
 A. H. Nuttall

J. Tierney
 R. E. Wernikoff
 H. E. White

A. OUTPUT DISTRIBUTIONS OF LINEAR SYSTEMS

1. From the Viewpoint of Given Autocorrelation Functions

To obtain the probability distribution functions of the output of a linear system analytically, when the input to the system is an ergodic ensemble of strictly stationary time series, it is necessary that we have all of the autocorrelation functions of the input.

In order to show this, an input time series is denoted by $f(t, \omega)$, where ω is a continuous parameter, and it is assumed (1) that

$$f(t, \omega) f(t + \tau_1, \omega) f(t + \tau_2, \omega) \dots f(t + \tau_\nu, \omega) \quad \nu \geq 0, \tau_0 = 0 \quad (1)$$

is absolutely integrable with respect to ω . Then the ν th-order autocorrelation function of the input

$$\phi_f^*(\tau_1, \dots, \tau_\nu, \omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t, \omega) f(t + \tau_1, \omega) \dots f(t + \tau_\nu, \omega) dt \quad (2)$$

exists for almost all members of the ensemble. Since this ensemble is ergodic, Eq. 2 is equal to the integral of Eq. 1 with respect to ω over the entire ω space. Hereafter, the right side of Eq. 2 is denoted by $\phi_f^*(\tau_1, \dots, \tau_\nu)$.

The output of the linear system is strictly stationary and can be expressed as

$$g(t, \omega) = \int_0^\infty K(\sigma) f(t - \sigma, \omega) d\sigma \quad (3)$$

where $K(\sigma)$ is the impulse response of the system. We also assume that

$$g(t, \omega) g(t + \tau_1, \omega) \dots g(t + \tau_\nu, \omega) \quad (4)$$

is absolutely integrable with respect to ω ; therefore the ν th-order autocorrelation function of the output exists for almost all members of the output ensemble. The ν th-order autocorrelation function of the output is

$$\begin{aligned} \phi_g^*(\tau_1, \dots, \tau_\nu) &= \int_0^\infty K(\sigma_1) d\sigma_1 \dots \int_0^\infty K(\sigma_{\nu+1}) d\sigma_{\nu+1} \\ &\quad \times \phi_f^*(\tau_1 + \sigma_1 - \sigma_2, \dots, \tau_\nu + \sigma_1 - \sigma_{\nu+1}) \end{aligned} \quad (5)$$

(VII. STATISTICAL COMMUNICATION THEORY)

To determine the first-order probability distribution function of the output, the moments

$$a_{\nu+1} = \phi_g^*(0, \dots, 0) = \int_0^\infty K(\sigma_1) d\sigma_1 \dots \int_0^\infty K(\sigma_{\nu+1}) d\sigma_{\nu+1} \\ \times \phi_f^*(\sigma_1 - \sigma_2, \dots, \sigma_1 - \sigma_{\nu+1}) \quad \nu > 0 \quad (6)$$

$$a_1 = \int_0^\infty K(\sigma) d\sigma \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t - \sigma, \omega) dt$$

and

$$a_0 = 1$$

are required. These moments (2) determine the characteristic function $\Phi(\mu)$ of the output distribution function in a region about $\mu = 0$, for the Taylor expansion of $\Phi(\mu)$ about $\mu = 0$ is

$$\sum_{\nu=0}^{\infty} \frac{i^\nu a_\nu}{\nu!} \mu^\nu \quad (7)$$

If the infinite series in Eq. 7 converges to $\Phi(\mu)$ when $|\mu| < r$, where $r > 0$, then the characteristic function can be determined in a region about $\mu \neq 0$, $|\mu| < r$, for the Taylor expansion of $\Phi(\mu)$ about $\mu \neq 0$ is

$$\sum_{\nu=0}^{\infty} \frac{\Phi^{(\nu)}(\mu)}{\nu!} h^\nu \quad h > 0 \quad (8)$$

This shows that by means of analytic continuation $\Phi(\mu)$ can be determined over the entire range of μ . Having determined $\Phi(\mu)$, the probability distribution function of the output can be obtained by the Lévy formula (3) in any continuity interval of the output distribution function.

The higher-order probability distribution functions of the output can be similarly obtained from the information contained in Eq. 5.

This analysis applies not only to linear systems but to any system in which the output autocorrelation functions can be expressed in terms of the input autocorrelation functions.

J. Y. Hayase

References

1. N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications (Technology Press and John Wiley and Sons, Inc., New York, 1949), p. 18.
2. H. Cramér, Mathematical Methods of Statistics (Princeton University Press, Princeton, 1946), p. 176.
3. H. Cramér, op. cit., p. 93.

2. From the Viewpoint of Distribution Functions

A time-varying statistical input to a linear system produces a statistical output. This report gives the formulas for calculating the output distributions from the statistical properties of the input.

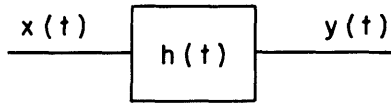


Fig. VII-1. Transfer system.

The original work, done by F. Lee, dealt with discrete approximations and stationary inputs. Subsequent work avoids the discrete approximations, produces exact results for classes of transfer functions, and is applicable to time-varying statistical inputs.

F. Lee considered the system of Fig. VII-1, in which a step approximation was made of the input $x(t)$ and the transfer function $h(t)$. The approximate output at time t is the linear sum

$$y(t) \approx \sum_{n=0}^N h_n x_n \quad (1)$$

where

$$h_n = h(nT) \cdot T \quad (2)$$

$$x_n = x(t - nT) \quad (3)$$

Thus the output distribution is given by the formula

$$W(y) = \frac{1}{h_0} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} W \left(\frac{y - \sum_{n=1}^N h_n x_n}{h_0}, x_1, \dots, x_N \right) dx_1 \dots dx_N \quad (4)$$

where $W(x_0, x_1, \dots, x_N)$ is the joint magnitude distribution of the input $x(t)$ at intervals of T into the past from time t . This formula is valid for time-varying statistical outputs.

(VII. STATISTICAL COMMUNICATION THEORY)

Reasonable step approximations cannot include effects far in the past and usually must have N large in order to get fair approximations of continuous functions.

To avoid these disadvantages of step approximations, the system transfer function can be approximated by a linear set of continuous functions

$$h(t) \approx \sum_{n=0}^N a_n h_n(t) \tag{5}$$

so that the output is the linear sum

$$y(t) \approx \sum_{n=0}^N a_n x_n \tag{6}$$

where

$$x_n = \int_0^{\infty} h_n(\sigma) x(t-\sigma) d\sigma \tag{7}$$

Then the output distribution is

$$W(y) = \frac{1}{a_0} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} W\left(\frac{y - \sum_{n=1}^N a_n x_n}{a_0}, x_1, \dots, x_N\right) dx_1 \dots dx_N \tag{8}$$

where the distribution $W(x_0, x_1, \dots, x_N)$ is the joint magnitude distribution at time t of the outputs of the set of networks $\{h_n(t)\}$ each with the input $x(t)$. That is, the distribution is experimentally obtained as shown in Fig. VII-2.

This procedure calculates the output distribution for a statistically time-varying

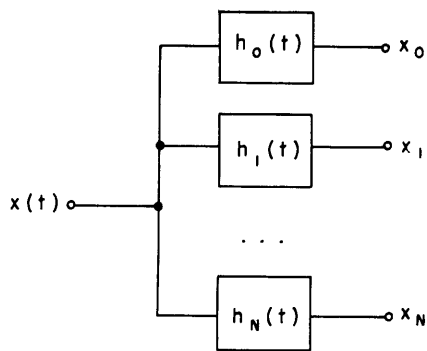


Fig. VII-2. Joint distribution.

input and avoids the step approximations by placing all of the approximation in the selection of the constants $\{a_n\}$. Notice that it is not necessary to have the set $\{h_n(t)\}$ orthogonal. The orthogonal property is only convenient in obtaining the constants $\{a_n\}$. By proper selection of the set $\{h_n(t)\}$ the convergence of the approximation is considerably improved as compared with the step approximation.

The output distribution is obtained exactly when either the transfer function is limited to a linear combination of the set $\{h_n(t)\}$ or the

input contains only nonzero components $\{x_n\}$, $n = 0, 1, \dots, N$.

A trivial case is one in which the transfer function is known within a multiplication constant.

$$h(t) = a_0 h_0(t) \quad (9)$$

Then the output distribution is

$$W(y) = \frac{1}{a_0} W_{x_0}\left(\frac{y}{a_0}\right) \quad (10)$$

This result shows that when the transfer function is known, the best method of measuring the input and calculating the output distribution consists of building the network and measuring the output distribution. This simple result is due to the fact that in all cases a physical measurement must be made or assumed. The best measurement to make is the one that measures the quantity desired.

When it is possible to measure the joint characteristic function of the distribution $W(x_0, x_1, \dots, x_N)$, simplification can be obtained. The joint characteristic function is defined as

$$\phi(u_0, u_1, \dots, u_N) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} W(x_0, x_1, \dots, x_N) \exp\left[\sum_{n=0}^N j u_n x_n\right] dx_0 dx_1 \dots dx_N \quad (11)$$

Then the characteristic function of the output distribution is

$$\phi(u) = \phi(u_0, u_1, \dots, u_N) \Big|_{u_n = a_n u} \quad (12)$$

so that the output magnitude distribution is

$$W(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(u) \exp[-j u y] du \quad (13)$$

Thus the output distribution of a linear network is conveniently calculated as a multiple integral when the distribution is measured, and as a single integral when the characteristic function is measured.

For independent components, the characteristic function is of the form

$$\phi(u_0, u_1, \dots, u_N) = \prod_{n=0}^N \phi_n(u_n) \quad (14)$$

in which each ϕ_n is measured separately at the output of the network $h_n(t)$.

(VII. STATISTICAL COMMUNICATION THEORY)

The output magnitude distribution of the linear system that has a nonstationary input can be calculated either by approximating the input and system function as step functions or approximating the system function as a linear combination of continuous functions. Extensions of this theory lead to methods of calculating correlation functions and non-linear optimum filtering.

C. S. Lorens

References

1. Fan Lee, S. M. Thesis, Department of Electrical Engineering, M. I. T., 1952.

B. FORMULATION OF FLOW GRAPHS

Proper formulation of a flow graph for electrical networks makes it possible to obtain network properties in more desirable forms. This report presents a method of flow-graph formulation, based on Ohm's law, that provides a means of making the conventional network approximations of shorting out or taking out elements.

Formulation of a flow graph involves, essentially, the selection of a consistent set of variables and the determination of the proper interrelationship. Within a given section of an electrical network a consistent set of variables is the loop currents. The loop current equations can be written in the form:

$$\pm i_1 Z_{1j} + \dots + i_j Z_{jj} + \dots \pm i_n Z_{nj} = E_j \tag{1}$$

where E_j is the effective independent voltage rise in the loop and the Z 's are the mutual and self-impedances. In flow-graph notation, Eq. 1 can be written in the form:

$$\mp \frac{i_1 Z_{1j}}{Z_{jj}} + \dots \mp \frac{i_n Z_{nj}}{Z_{jj}} + \frac{E_j}{Z_{jj}} = i_j \tag{2}$$

The flow graph is then the set of nodes that represents the loop currents, each node being multiplied by the mutual impedance and divided by the self-impedance. The flow

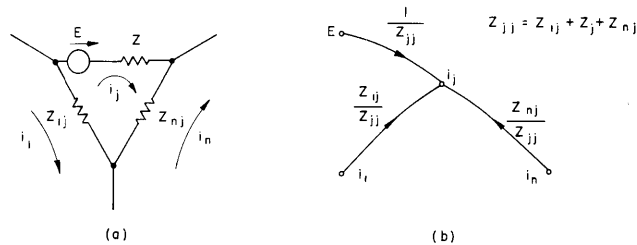


Fig. VII-3. (a) General network. (b) Flow-graph formulation for i_j .

graph of the loop current i_j in Fig. VII-3a is that of Fig. VII-3b.

This type of formulation is essentially Ohm's law

$$i_j = \frac{E_j^*}{Z_{jj}} \quad (3)$$

in which the effective voltage in each branch of the loop is the sum of the independent voltage, plus the mutual impedance, multiplied by the mutual current. The formulation places the flow graph nodes in the same geometrical relation as the corresponding loop currents in the electrical network; therefore the flow graph retains some of the geometrical properties of the network. A parallel formulation exists for node-pair voltages. A more general formulation will be found in reference 1.

Many times there is interest in formulating the flow graph so that different electric elements of the network can be shorted out or taken out. Then, it is convenient to have the element appear alone in the flow graph and in such a form that the flow-graph branch becomes zero when the electric element is shorted out or taken out.

In order to formulate a flow graph for shorting out an electric element, formulate the element on the node-pair voltage basis and the rest of the surrounding network on the loop-current basis, placing the datum voltage node at one end of the element. When an element is to be taken out, formulate the element on the loop-current basis and the surrounding network on the node-pair voltage basis.

As an example, take the flow graph of a tapped input vacuum-tube amplifier in which the effect of shorting the cathode resistor is desired. The cathode resistor is formulated with a node-pair voltage, while the remaining portion of the circuit is formulated with loop currents (Fig. VII-4a, b, c). Note that when $R_k = 0$ the flow graph reduces to a cascade graph that represents the conventional vacuum-tube amplifier.

This type of formulation, in which electric elements can be shorted out or taken out, is important, since it leads to a solution of the flow graph in which the effect of these elements can be separated from the rest of the solution. This separation is quite desirable in finding approximations, and the effect of approximations, without having to reformulate mathematically in order to recognize the effect.

As an example of separating off the effect of R_k in the network of Fig. VII-4a, the gain and input impedance are found directly from the flow graph in the following form:

$$\frac{e_o}{e_i} = \frac{\frac{-\mu R_L}{R_L + r_P} + \frac{R_k R_L}{R_1 (R_L + r_P)}}{1 + R_k \left(\frac{1}{R_1} + \frac{1 + \mu}{R_L + r_P} \right)}; \quad \frac{e_i}{i_1} = R_1 \left[1 + R_k \left(\frac{1}{R_1} + \frac{1 + \mu}{R_L + r_P} \right) \right]$$

(VII. STATISTICAL COMMUNICATION THEORY)

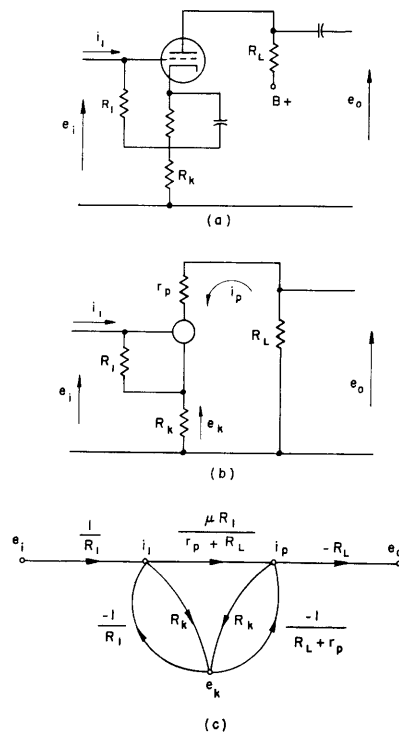


Fig. VII-4. (a) Tapped input vacuum-tube amplifier. (b) Equivalent circuit for tapped input amplifier. (c) Flow graph of tapped input amplifier.

Application of Ohm's law, with interpretations, provides a means of formulating a flow graph in such a way that the geometry of the network is preserved and approximations can be made in the solution without reformulation.

[The present study has been completed. It was submitted as a thesis in partial fulfillment of the requirements for the degree of Doctor of Science, Department of Electrical Engineering, M. I. T., July 15, 1956, and will be presented in Technical Report 317.]

C. S. Lorens

References

1. C. S. Lorens, Sc. D. Thesis, Department of Electrical Engineering, M. I. T., July 1956.

C. TIME-DOMAIN SYNTHESIS OF LINEAR SYSTEMS

A method will be presented whereby any desired impulse response, or, indeed, any desired linear time-invariant operation, can be realized to any prescribed degree of accuracy by a finite, lumped, active or passive network. The method seems practicable

because it is very simple, both in conception and in the ease with which the necessary computations and construction of apparatus can be accomplished. The method is also useful in theoretical work, in that it provides illustrations for some concepts of signal theory that are currently being studied. In fact, it was this theoretical work, rather than a specific interest in network synthesis, that provided the motivation that led to the present method.

Signal theory, per se, is the study of some of the consequences of realizing that in all physical work we must inevitably accept a finite error. This comes about because, ultimately, all of our measuring instruments have finite resolutions which prevent us from distinguishing physically things that differ by less than some finite amount. As a result, some mathematically exact concepts, as, for example, equality, do not seem to have an operational meaning. Recognition of the existence of physical limitations in accuracy should, ideally, be an integral part of the theoretical tools that we use in studying physical phenomena. Accomplishing this recognition is a problem studied in signal theory. For the purposes of this report, however, it is sufficient to be willing to agree that with all physical systems we know, there is associated a finite resolution, and that any error (i. e., difference between two things) smaller than that resolution cannot be recognized as an error, and should, therefore, not be thought of as one.

1. Reduction of the General Problem to an Equivalent Simpler One

We shall consider the problem of synthesizing an arbitrary given singular impulse response, that is, an impulse response that consists only of a finite number of impulses of finite area distributed over a finite time interval (see ref. 1). If, instead, the problem starts with the specification of an ordinary, smooth impulse response, reference 1 shows how an equivalent singular response may be obtained from the given smooth response. Similarly, if a linear, time-invariant operation is prescribed, the methods of reference 1 can be used to obtain the corresponding singular impulse response. As a result, no loss of generality is incurred by considering only the synthesis of singular impulse responses. (The reader will note that the reduction to singular responses already involves an error. However, as is shown in reference 1, this error can be made as small as desired; hence it can be made consistent with the resolution level of the system.) A possible singular impulse response is shown in Fig. VII-5b.

In reference 2 it is shown in great detail that, provided the input to the system is essentially constant within a time interval δ seconds long (in effect, provided δ is small compared with the period of the highest frequency of interest), the impulses of Fig. VII-5b may be changed to the pulses of Fig. VII-5c without appreciable change in the operation of the network, so long as the pulses have the same areas as the corresponding impulses. In fact, if the width of the pulses, δ , meets the stated requirements,

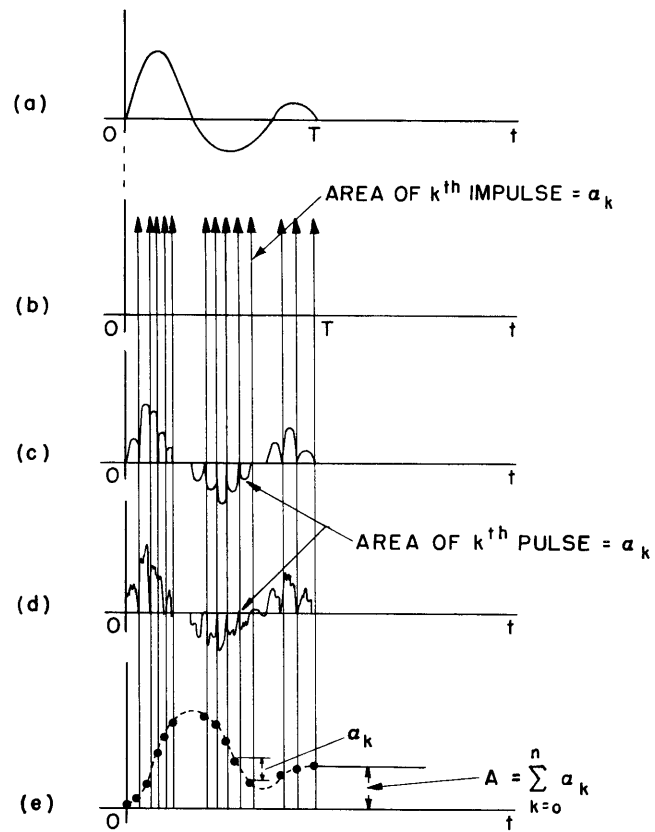


Fig. VII-5. (a) Impulse response to be synthesized; (b) equivalent singular impulse response; (c, d) equivalent dispersed responses; (e) step response corresponding to (c) or (d).

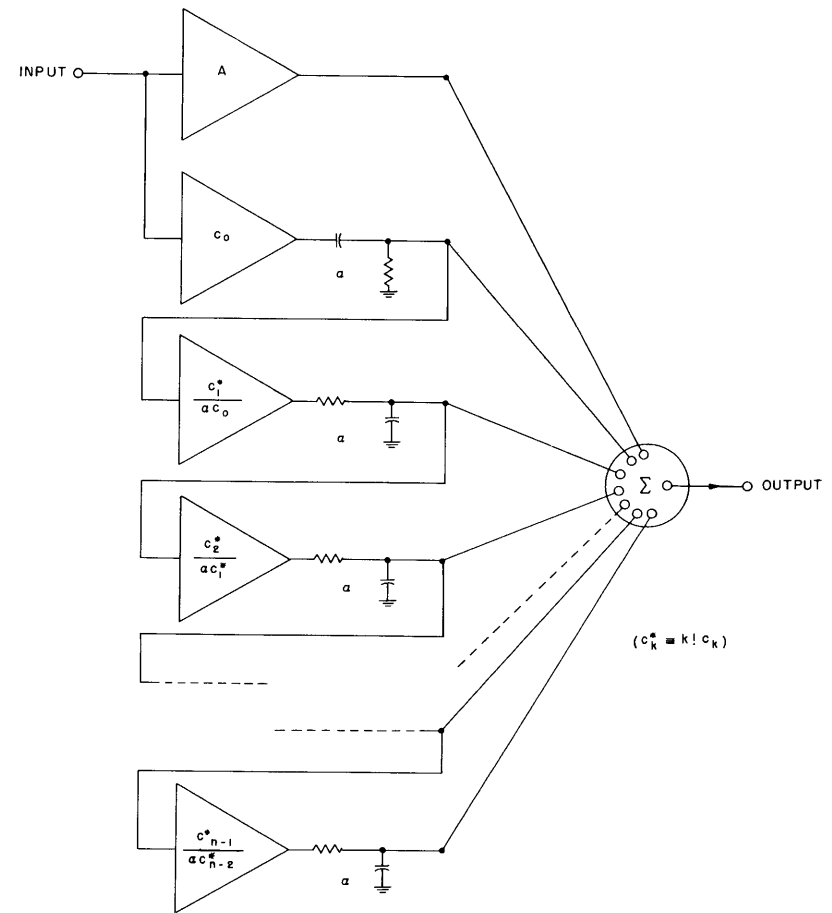


Fig. VII-6. Active network whose step-function response is $f(t)$. (For derivation see Appendix, Sec. VII-C. 3.)

the pulses may have any shape whatever (so long as they have the correct net areas) without changing the operation of the network by a distinguishable amount. Thus a second network equivalent to that of Fig. VII-5a might have the response shown in Fig. VII-5d. The reader will note that in choosing an appropriate finite δ – in fact, in admitting its existence – we again recognize and use our own limitations in the detection of errors.

We have, then, that the number of pulses and their width can be chosen in such a way that the original network with the smooth response and the network with the response shown in Fig. VII-5c or d are equivalent, in the sense that the difference between their outputs (for any input whose highest frequency does not exceed that for which the system is designed) will be smaller than the limit of resolution. Thus, from a physical point of view, the synthesis of the network with the response shown in Fig. VII-5c or d is equivalent to the synthesis of the prescribed smooth response from which the singular response of Fig. VII-5b was derived.

All arguments that require the acceptance of error have appeared in this section. From now on, the discussion is again exact.

2. The Synthesis Procedure

We shall now synthesize the reduced response of Fig. VII-5c. The crux of the method lies in noting that the only important parameters of the reduced response are the areas of its pulses, and their location; the shapes of the pulses is irrelevant. Since we cannot prevent ourselves from being conscious of pulse shapes so long as we look at the impulse response itself, we shall instead consider its time integral, the step function response of the reduced network. By this means, the only relevant data, the pulse areas and positions, are placed in evidence as points such as those shown in Fig. VII-5e. The curve between the points is shown dotted, since it is irrelevant, and the step response is assumed to remain constant for $t > T$, since the impulse response was assumed to last over only a finite time T . Our whole synthesis problem is now reduced simply to finding a curve, any curve, which passes through the n specified points for $t < T$, and becomes asymptotically constant thereafter.

This is a very simple problem, and the number of possible solutions is enormous. Since so many functions will fit the requirements, we have the freedom of selecting only those that meet the two following requirements:

1. The interpolating functions must be easily synthesizable as networks (for example, choose functions that are step responses of known networks); and
2. the undetermined coefficients of the interpolating functions must be easy to calculate.

Even with these requirements, the number of functions that can perform as desired is extremely large, and we are still left with the choice of whether the resulting network

(VII. STATISTICAL COMMUNICATION THEORY)

shall be active or passive, or what form it shall have. In general, however, it is easier to meet requirement 2 when active synthesis is employed, since, in that case, each undetermined coefficient can be adjusted independently of all the others.

As an example, we shall derive an active network whose behavior will be equivalent to that of the arbitrary impulse response shown in Fig. VII-5b. Our problem is to find a function $f(t)$, corresponding to the step response of a known active network, which will pass through the chosen n points of $s(t)$, and be substantially constant for $t > T$. A possible function is

$$f(t) = A + (c_1 + c_2 t + \dots + c_n t^{n-1}) e^{-at}$$

where a is chosen so that $f(t) \approx A$ for $t > T$, for example, $a = 5/T$. Since the only remaining requirement is that

$$f(t_i) = s(t_i) \quad (i = 1, 2, \dots, n)$$

where the t_i are the times at which the chosen points of $s(t)$ occur, our problem reduces to solving the set of simultaneous equations

$$c_1 + c_2 t_i + c_3 t_i^2 + \dots + c_n t_i^{n-1} = [s(t_i) - A] e^{at_i} \quad (i = 1, 2, \dots, n)$$

for the c 's. These equations always have a solution, because the system determinant is just the Van der Monde determinant that does not vanish so long as all the t_i are different, as they are in this case. The network which realizes $f(t)$ is shown in Fig. VII-6.

It is interesting to note that only two-element active networks are required to realize a general impulse response. RL or RC networks will do, but not LC, since it is impossible with a finite number of j -axis poles to produce an impulse response that vanishes after a finite time. The particular procedure used in this example resembles ordinary delay-line time-synthesis, with the significant difference that no delay line is involved.

A formal error analysis of these methods, though very important, has not yet been made. However, because of the types of convergence involved, it is always possible to trade number of elements for decrease in error in the final realization. The number of elements would need to become infinite to make the error approach zero, but this limiting situation is never necessary in practice, since zero error has no operational meaning. There is a question, however, about the number of elements that is necessary to attain just moderate accuracy. This has not yet been studied either in theory or in practice, but experience with the Cerrillo delay-line synthesis indicates that excellent results can be obtained with n of the order of 10 or less.

The simplicity of this synthesis procedure results from abandoning the attempt to approximate whole curves, and considering, instead, a finite number of intelligently chosen points which, by themselves, convey most of the distinguishable information

contained in the curves.

3. Appendix

We shall show briefly how the network of Fig. VII-6 can be obtained from the step-function response

$$f(t) = A + \sum_{k=0}^{n-1} c_k t^k e^{-at}$$

Using the Laplace transform relations

$$\mathcal{L}[e^{-at}] = \frac{1}{s+a}$$

$$\mathcal{L}[t^k e^{-at}] = (-1)^k \frac{d^k}{ds^k} \left(\frac{1}{s+a} \right)$$

and the identity

$$\frac{d^k}{ds^k} \left(\frac{1}{s+a} \right) = (-1)^k \frac{k!}{(s+a)^{k+1}}$$

we obtain

$$\mathcal{L}[f(t)] \equiv F(s) = \frac{A}{s} + \sum_{k=0}^{n-1} \frac{c_k k!}{(s+a)^{k+1}} = \frac{A}{s} + \sum_{k=0}^{n-1} \frac{c_k^*}{(s+a)^{k+1}}$$

where we have written $c_k^* \equiv k! c_k$.

$F(s)$ is the desired step-function response in the frequency domain. Now the transfer function $H(s)$ of a network is defined by $E_{\text{out}}(s) = H(s) E_{\text{in}}(s)$. When the input is a unit step function, $E_{\text{in}}(s) = \frac{1}{s}$. Since the desired output is $F(s)$, we have

$$H(s) = sF(s) = A + \frac{s}{s+a} \sum_{k=0}^{n-1} \frac{c_k^*}{(s+a)^k} = A + \frac{c_0^* s}{s+a} + \frac{s}{s+a} \sum_{k=1}^{n-1} \frac{c_k^*}{(s+a)^k}$$

Clearly, the given network has this transfer function.

R. E. Wernikoff

References

1. R. E. Wernikoff, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1956, p. 44.
2. M. V. Cerrillo, On the Synthesis of Linear Systems for Pure Transmission, Delayed Transmission, and Linear Prediction of Signals, Ch. III (to be published in the Quarterly Progress Report of Jan. 15, 1957).

(VII. STATISTICAL COMMUNICATION THEORY)

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

0.1 A BASIC PROBLEM IN TIME-DOMAIN SYNTHESIS OF LINEAR SYSTEMS

One of the important objectives in the synthesis of linear systems can be described as follows:

Let $\phi(t)$ and $\gamma(t)$ be two single-valued real functions of the real variable t , such that

$$\begin{aligned} \phi(t) &\equiv 0 && \text{for } -\infty < t < 0 \\ &\neq 0 && \text{for } 0 \leq t < \infty \\ \gamma(t) &\neq 0 && \text{for } -\infty < t < t_0 = \text{constant} \\ &\neq 0 && \text{for } 0 \leq t_0 \leq t < \infty \end{aligned} \tag{I-0.1}$$

Now, having prescribed two functions $\phi(t)$ and $\gamma(t)$ which satisfy these conditions but are otherwise arbitrary, we ask:

Is it possible to construct a linear, passive, and finite system whose elements are lumped or distributed but always finite in number, extension, and magnitude, and which consists of two external terminal pairs, such that when the system is excited at one pair of terminals by $\phi(t)$, the output, at the other terminal pair will be of the form $\gamma(t)$?

It should be noted that the condition of finiteness is postulated in the statement of the problem. That is, the number of elements, their extension, and their values, must be finite. Under these restrictions, the solution of the stated problem becomes difficult.

0.11 Various counterexamples can be found that indicate that the problem does not always have a solution, at least not in the form required above. A simple counterexample is the pure transmission problem, which is defined by the condition $\gamma(t) = \phi(t)$ for $0 < t < \infty$. Another one is the problem of delayed transmission, defined by $\gamma(t) = \phi(t - t_0)$ for $0 < t_0 = \text{constant} < \infty$, which also does not have a solution within the restriction of finiteness of the linear system. For obvious physical reasons, the problem of advanced transmission, or prediction, which is defined by $\gamma(t) = \phi(t + t_0)$ for $0 < t_0 = \text{constant} < \infty$, does not have a solution. The proof that the problem of transmission does not have an exact solution will be given later.

[†]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 (Vol. 4, No. 2, 1955), 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 work that is being done by the Statistical Communication Theory group and other groups in the Laboratory led to its translation, by Mr. Wernikoff, and its presentation here.]

(VII. STATISTICAL COMMUNICATION THEORY)

0.12 The problem of pure transmission and the associated problems of delayed and advanced transmission – the latter in a sense that is explained later – are of fundamental importance in the theory of communication. The nonexistence of exact solutions that satisfy the basic problem described in section 0.1 does not imply that it is impossible to find approximate solutions which, within certain tolerances, represent "acceptable" solutions. A suitable modification of the statement of the basic problem allows us to find solutions, many of them of an approximate nature, to many problems that would not have solutions under the original stipulations.

It should be noted that the condition of finiteness of the linear system will not be removed in any case, since it is impossible to construct systems composed of an infinite number of elements, or of elements with infinite extensions or values.

0.13 The principal object of this paper is to discuss and give appropriate solutions to the problem of transmission and its associated problems when the requirements of section 0.1 are suitably modified. Before introducing the modifications, we shall present a few preparatory concepts that will be needed in the following sections.

0.2 THE CONVOLUTION INTEGRAL AS A BASIC TOOL

The operation of a linear system can be characterized by the convolution integral as follows:

$$y(t) = \int_0^t \phi(t-\tau) S_c(\tau) d\tau = \int_0^t \phi(\tau) S_c(t-\tau) d\tau \quad 1(I-0.2)$$

where $S_c(t)$ represents the system function in the time domain. The function $S_c(t)$ is defined as the response of a linear system when the excitation is the unit impulse $U_0(t)$ applied at $t = 0$. Figure 1(I-0.2) is a sketch of the situation represented by the convolution integral.

The integral 1(I-0.2) will be understood to be taken in the sense of Riemann in order to illustrate better some points that will demonstrate the nonexistence of solutions of the transmission problem. Later on, the convolution integral will be considered in the sense of Stieltjes.

0.3 THE NONEXISTENCE OF SOLUTIONS OF THE PURE TRANSMISSION PROBLEM IN FINITE SYSTEMS WITH CONTINUOUS KERNEL

It will be shown that the problem of transmission does not have an exact solution, except a trivial one, when we require finiteness of the linear system in question.

0.31 It is well known that the response $S_c(t)$ of a linear, finite, and passive system to the unit impulse is a bounded, single-valued, and continuous function for $t > 0$. For

(VII. STATISTICAL COMMUNICATION THEORY)

the case of pure transmission, $\gamma(t) = \phi(t)$, we have

$$\phi(t) = \int_0^t S_c(t-\tau) \phi(\tau) d\tau \quad 1(I-0.31)$$

which is a Volterra integral equation of the first type. The theory of integral equations teaches us two things: First, that under the conditions that must be satisfied by the excitation $\phi(t)$ and because of the continuity of the kernel $S_c(t-\tau)$, the first member of 1(I-0.31) is a continuous function. Therefore, there can only exist solutions for excitations of the class C_0 , which limits the generality of the transmission problem; second, that under these circumstances the Volterra equation admits as a unique solution $\phi(t) = 0$, which is a trivial solution of the transmission problem.

0.32 There are other trivial solutions to the transmission problem. The first is that in which the system reduces to a parallel resistance, for example. The second is that in which the system is an ideal transmission line terminated in its characteristic impedance (delayed transmission). The ideal transmission line is not realizable in the strict sense of the word, because of the line dissipation.

0.33 The basic method of attack in the pure transmission and associated problems which will be employed in this paper is the convolution integral in the sense of Stieltjes:

$$\gamma(t) = \int_0^t \phi(t-\tau) dK(\tau) = \int_0^t S_c(t-\tau) dY(\tau) \quad 1(I-0.33)$$

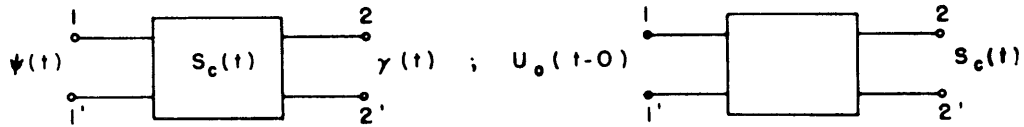
where the functions $\phi(t)$ and $S_c(t)$ may be singular, but integrable in the sense of Stieltjes. $K(t)$ and $Y(t)$ are distribution functions associated with $\phi(t)$ and $S_c(t)$, respectively.

0.4 TOLERANCE SETS AND APERTURE SETS

We shall now present some concepts that will allow us to reformulate the basic problem of section 0.1 in such a way that solutions to important problems will exist. The new formulation will allow us to solve problems, such as the pure transmission problem, that are insoluble under the original postulation of section 0.1.

0.41 We here introduce the aperture set and the tolerance set. Let $\gamma(t)$ be a function of bounded variation which satisfies the requirement of section 0.1. Figure 1(I-0.41) shows a possible $\gamma(t)$. Let us associate with $\gamma(t)$:

a. a countable set of open finite intervals $[a_k]$, such that $\bigcap [a_k] = 0$. These intervals may degenerate to points. The length of the intervals may be small, but not necessarily. We shall suppose that a_k , $k = 0, 1, 2, \dots$ represents the length of each



$U_0(t-0) = \text{UNIT IMPULSE AT } t = 0$

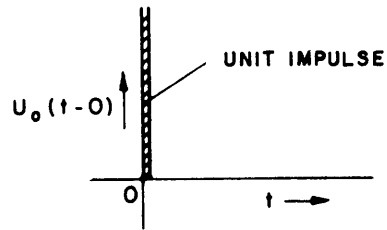


Fig. 1(I-0.2). The impulse response of a linear system.

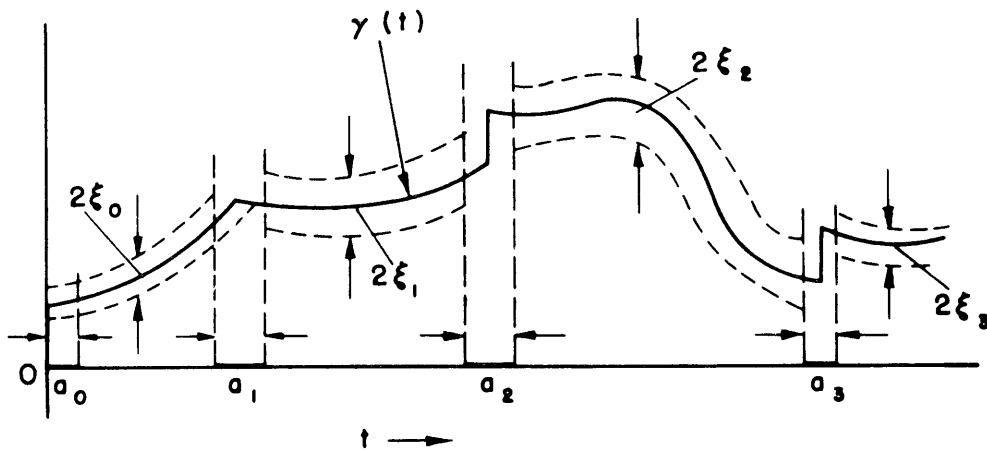


Fig. 1(I-0.41). Tolerance set and aperture set.

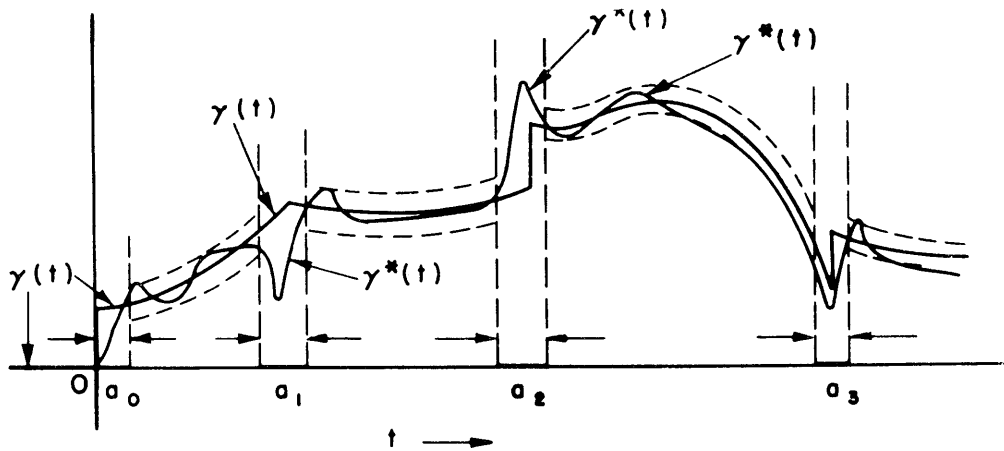


Fig. 1(I-0.5). Example of an acceptable function.

(VII. STATISTICAL COMMUNICATION THEORY)

interval. This set, which is prescribed beforehand, will be called an aperture set. The position and distribution of the intervals $[a_k]$ is not important now. Later their positions will be determined by the points of discontinuity, singularity, and so forth of the functions $\phi(t)$ and $\gamma(t)$.

Let us associate with the graph of $\gamma(t)$:

b. a countable set of small positive numbers ϵ_j , $j = 1, 2, \dots$. The values ϵ_j are used to determine neighborhoods along the length of the graph of $\gamma(t)$ in the intervals of t that are contained in the complement of $\cup [a_k]$.

0.5 Here we introduce the concept of acceptable functions $\gamma^*(t)$ associated with $\gamma(t)$. There is a function, $\gamma^*(t)$, of bounded variation, real and single-valued, of the real variable t such that:

- i. The graph of $\gamma^*(t)$ is always within the interior of the neighborhood region defined by the set ϵ_j when t is contained in the complement of $\cup [a_k]$.
- ii. The graph of $\gamma^*(t)$ may run completely outside of the neighborhood regions of $\gamma(t)$ when t is contained in $[a_k]$, $k = 0, 1, 2, \dots$.
- iii. $\gamma^*(t) = 0$ for $-\infty < t < 0$.

Condition (i) is equivalent to writing

$$|\gamma(t) - \gamma^*(t)| \leq \epsilon_j \quad j = 1, 2, \dots$$

when t is not contained in $\cup [a_k]$.

Condition (ii) indicates that the set of tolerances does not control the limits of variation of $\gamma^*(t)$ within the apertures.

The regional numerical likeness between the functions $\gamma(t)$ and $\gamma^*(t)$ is emphasized by calling $\gamma^*(t)$ the "simile" of $\gamma(t)$. The function $\gamma^*(t)$ by itself is called an acceptable function.

(In spite of the similarity between $\gamma(t)$ and $\gamma^*(t)$, we do not say that $\gamma^*(t)$ is an approximation to $\gamma(t)$, in the usual sense of this word. The function $\gamma^*(t)$ could, for example, belong to the class $L^{(\infty)}$, while $\gamma(t)$ could be of class $L^{(r)}$, $r < \infty$.)

Figure 1(I-0.5) shows an acceptable function associated with $\gamma(t)$, in the sense already explained.

0.51 We present here the class of functions $\phi(t)$, which will be considered acceptable excitations for our system.

Consider, first, the class of functions $\phi(t)$ such that:

- i. $\phi(t) \equiv 0$ for $-\infty < t < 0$
 $\neq 0$ for $0 \leq t < \infty$

and such that $\phi(t)$ is a real, single-valued function of the real variable t .

(VII. STATISTICAL COMMUNICATION THEORY)

ii. $\phi(t)$ is a regionally continuous function, class C° , for almost all values of t , $0 \leq t < \infty$. The exceptional points are formed by

a. a countable number of isolated points t_k , $k = 1, 2, \dots$, on which the function $\phi(t)$ presents simple discontinuities with finite jump;

b. a countable number of isolated points t_i , $i = 1, 2, \dots$, on which the function $\phi(t)$ has singularities (impulses, doublets, and so forth) in such a way that condition (iii) is satisfied.

iii. The function $\phi(t)$ must be integrable in the sense of Stieltjes in any finite sub-interval of time that is contained in the interval $0 \leq t < \infty$.

(Condition (iii) implies that the function $\phi(t)$ is regionally of bounded variation. This is a consequence of an elementary theorem of Stieltjes integrals. From now on it will be supposed that the reader is familiar with the elementary properties of such integrals.)

Figure 1(I-0.51) shows an example of an acceptable function $\phi(t)$, in which the singularities are simply impulses. It is apparent that the function of Fig. 1(I-0.51) can be expressed as the sum of three component functions:

- a. a function $\phi_1(t)$ continuous for all values of t in the interval $0 \leq t < \infty$;
- b. a discontinuous function $\phi_2(t)$, composed of steps, ascending and descending, on the points of discontinuity of $\phi(t)$;
- c. a singular function $\phi_3(t)$ which is zero for almost all values of t , $0 \leq t < \infty$, the

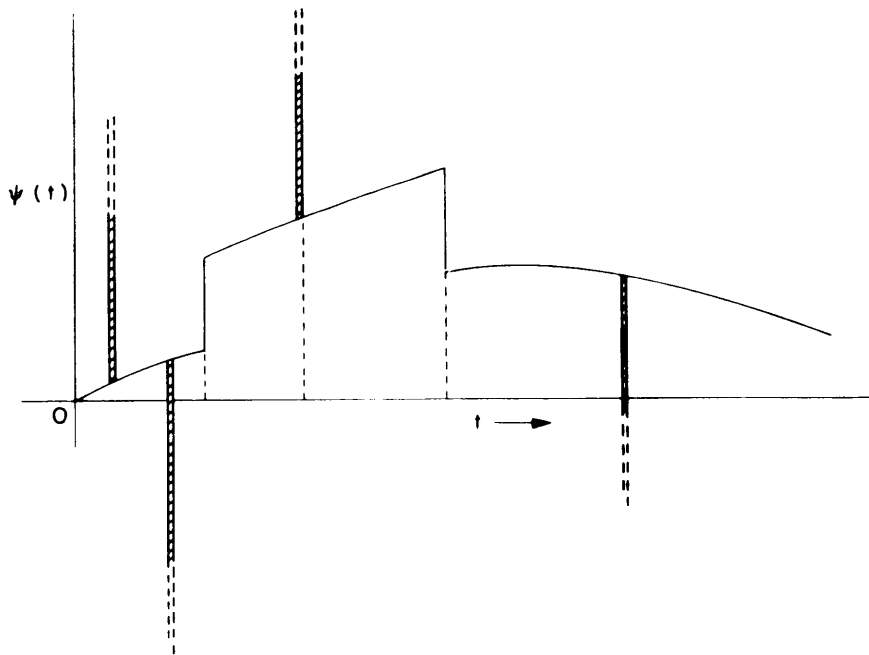


Fig. 1(I-0.51). Example of an excitation function $\phi(t)$.

(VII. STATISTICAL COMMUNICATION THEORY)

exceptional points being those where $\phi(t)$ has impulsive character.

0.52 In general synthesis problems, the class of functions $\phi(t)$ is extended to include random functions that are integrable in appropriate senses. In this paper we shall not consider such functions, since we are not concerned with transmission in the presence of noise, and so forth.

0.6 FORMULATION OF THE GENERAL PROBLEM AND OF THE TRANSMISSION PROBLEM

An adequate formulation of the general problem of the synthesis of linear systems, described in section 0.1, can now be presented in terms of the given concepts.

Given two functions $\phi(t)$ and $\gamma(t)$ that satisfy the condition prescribed in sections 0.5 and 0.51, but are otherwise arbitrary, and given, also, the sets $[\epsilon_j]$ and $[a_k]$, already described, we ask:

"Is it possible to construct a linear, finite, passive system consisting of two external terminal pairs, such that, when an excitation of the form $\phi(t)$ is applied to two of the terminals, the response will be $\gamma^*(t)$, $\gamma^*(t)$ being a simile of $\gamma(t)$?"

The object of this paper is to produce solutions to the problem of transmission, delayed transmission, and linear prediction, of a signal $\phi(t)$. The response of the system should be given by the similes $\phi^*(t)$, $\phi^*(t - t_0)$, $\phi^*(t + t_0)$, with $0 < t_0 = \text{constant}$.

0.7 CLASSICAL SITUATION OF THE TRANSMISSION PROBLEM

The transmission of a signal through a communication channel is not a new thing, either in theory or in the construction of systems that make it possible.

It is well known that, theoretically, a linear system whose response to a unit impulse applied at $t = 0$ is the same impulse at $t = 0$, has the property of transmitting without distortion or delay a signal of the form $\phi(t)$. If the response of the system is a unit impulse delayed by t_0 units of time with respect to the impulse, then the response of such a system to a function $\phi(t)$ will be of the form $\phi(t - t_0)$. The transfer function of a system that transmits with delay, but without distortion, is simply e^{-st_0} in the frequency domain. Classically, the transmission problem is solved by constructing systems whose transfer function approximates in some sense the function e^{-st_0} . Note that in the solutions given above, the kernel of the convolution integral is simply a singular function and not a continuous function, whence there is no conflict with the discussion of section 0.3.

0.71 It is now convenient to clarify the situation in relation to the classical solution of the transmission problem. In the process of this discussion, the

(VII. STATISTICAL COMMUNICATION THEORY)

following points are relevant.

1. There exist many singular kernels, in fact, an infinite number, which produce transmission without delay, delayed transmission, or linear prediction. The classical kernel, consisting of just one impulse, is a particular case of the kernels presented herein.

2. Methods will be given that allow the transition from singular kernels to continuous kernels, in such a way that the intrinsic character of the transmission is not changed, following the postulation of section 0.6.

3. The concepts and results so obtained have a basic importance in the general foundation of the theory of synthesis of linear systems.

4. The basic method for the construction of four-terminal systems that are potentially capable of producing pure, delayed, and advanced transmission of signals is given.

5. The solutions produced for linear prediction consist of kernels of finite duration, and simple relations are shown between the duration of a kernel and the time of prediction.

6. As a by-product, kernels will be given which produce the successive derivatives of a continuous function $\phi(t)$.

I. KERNELS AS DISTRIBUTIONS OF WINDOW FUNCTIONS

1.0 METHOD OF SOLUTION

The resultant of Stieltjes, or convolution integral, is the basic tool that is used to attack the problem stated in section 0.6. The method of solution consists of solving the integral equation, expressed by the convolution integral, and determining the function $S_c(t)$ which characterizes analytically the linear system. The complete solution of the synthesis problem does not end with the determination of $S_c(t)$. The following step is to construct, or synthesize, the system itself, placing in evidence its elements, their distribution, and their value. This step produces the linear network that satisfies the condition of the problem.

In the solution of the two previous problems, a central place is occupied by certain distributions of impulses on which the complete solution of the problem depends. To prepare the ground for the solution of the problem proposed in section 0.6, and in particular the problem of transmission and associated problems, we introduce and discuss these distributions.

1.1 WINDOW FUNCTIONS

We start by giving a few definitions. The term "pulse" will be used in its usual connotation: The graph of a current or voltage wave, or its equivalents, when it is

(VII. STATISTICAL COMMUNICATION THEORY)

almost unidirectional and of short duration, is called a pulse.

By a "train of pulses" we understand a finite succession of consecutive pulses of finite duration. The following definitions are motivated by the previous concepts.

DEFINITION. A symmetric pulse, in the form of a spearhead, which is of finite duration, and is always positive, or always negative, is called a pulse in the form of a "window." The function that represents this pulse, let us say $v(t)$, is called a window function, if, besides satisfying the above conditions, the function is integrable in the sense of Stieltjes to a finite limit that is not zero and remains determinate even when the duration of the pulse tends to zero. As a consequence of this definition, a window necessarily becomes an impulse when its duration tends to zero. It follows also that not all symmetric, unidirectional pulses are necessarily windows, since the definition is not satisfied if the area contained under the graph of the pulse tends to zero when its duration tends to zero. Figure 1(I-1.1) shows a few possible forms of windows displaced by a time t_0 from the origin.

In the discussion that follows it will be seen that the specific form of each window does not have a major influence on the fundamental results that will be obtained. The form of the graph of a window has a secondary influence, modifying only the fine structure, but not the basic architecture, of the results. In the course of the discussion it will be shown that the basic elements of a window function are: its aperture, its displacement from the origin, and, principally, its integral.

1.2 WINDOW DISTRIBUTIONS

The basic elements in this research are certain distribution functions formed by means of window functions. We have the following definitions:

A distribution in time of a finite number of nonoverlapping, adjacent windows is called a distribution density of windows. Figure 1(I-1.2) shows such a distribution density. Note that in this array it is not necessary that successive windows have opposite signs. The length of the interval occupied by the window array

$$\ell = \sum_{k=1}^m a_k$$

is called the measure of the support of the array.

The array of windows allows the introduction of the concept of "distribution function" associated with a specific array. Since each window is, by definition, integrable in the sense of Stieltjes, the integral of the array of windows also exists when the number of windows is finite. The Stieltjes integral between $t = t_0$ and $t = t_0 + \ell$ is called the measure of the distribution. When the distribution is formed by a finite number of windows of duration greater than zero and of finite height, the array is also integrable

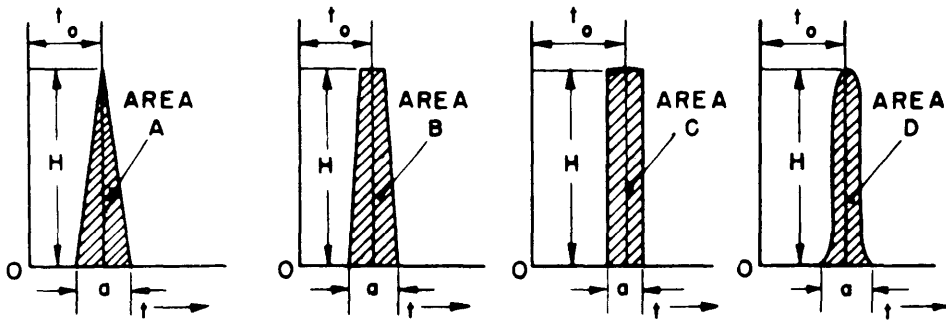


Fig. 1(I-1.1). Examples of windows.

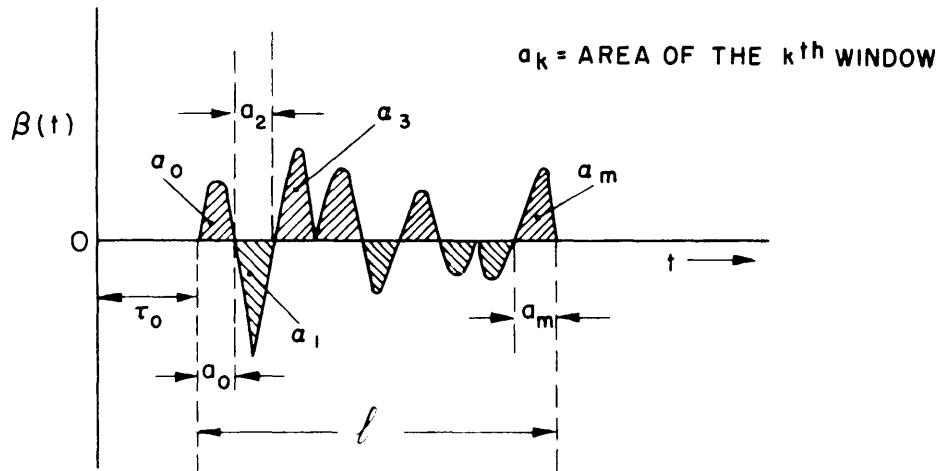


Fig. 1(I-1.2). A window array.

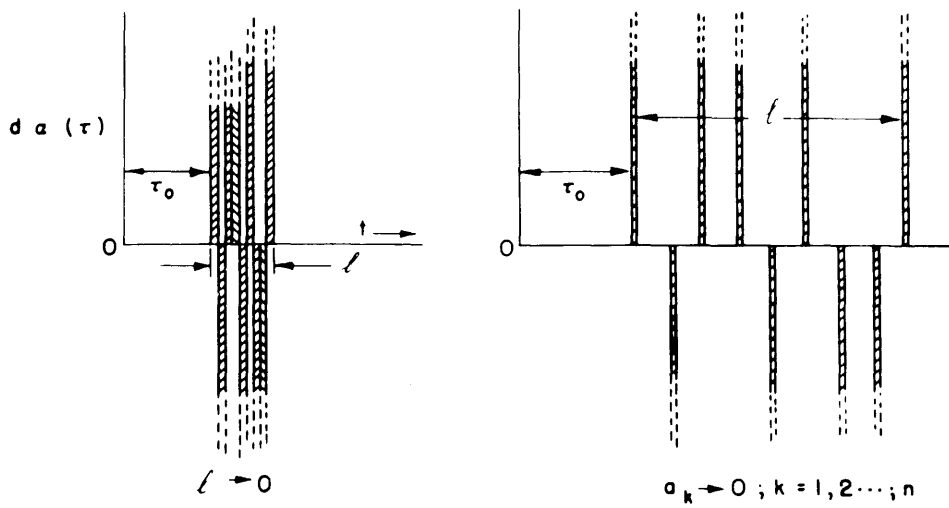


Fig. 2(I-1.2). Examples of singular distributions.

(VII. STATISTICAL COMMUNICATION THEORY)

in the sense of Riemann. We have, then,

$$\alpha(t_0, \ell) = \int_{t_0}^{t_0 + \ell} \beta(t) dt$$

The transformation $\tau = t - t_0$ allows us to write

$$\alpha(\ell) = \int_0^{\ell} \beta(\tau) d\tau$$

Let us now take a value of τ in the interval $0 < \tau < \ell$. The distribution is then defined by the expression

$$\alpha(\tau) = \int_0^{\tau} \beta(\tau) d\tau$$

The generalization of this expression to the integrals of Stieltjes is represented by the well-known notation

$$\alpha(\tau) = \int_0^{\tau} d\alpha(u)$$

where u is an auxiliary variable that disappears when the limits of the integral are substituted.

The limit of a distribution density of windows when the aperture of the windows tends to zero is an array of impulses. Two possible representations are shown in Fig. 2(I-1.2). In the first, the measure ℓ collapses to zero. In the second, the measure ℓ remains, by construction, constant. In this last case, the window function is defined as an impulse and takes the value zero on both sides of the impulse.

1.3 DISTRIBUTIONS OF INFINITE DURATION

A distribution of windows can also be defined as an array of an infinite number of windows that cover the interval $0 < t < \infty$. Distributions of this type appear in the solution of many problems of linear-system synthesis. For the purposes of this paper, these infinite distributions have very little meaning. Therefore, we shall not consider them in detail. We shall only refer to them in some limit processes.

1.4 SYSTEM FUNCTIONS IN THE TIME DOMAIN AND WINDOW ARRAYS

In this work we shall consider a system function $S_c(t)$ of a four-terminal network as being equivalent to a window array. Windows of finite duration will be our main interest.

The system function of a four-terminal linear network is, as we have already stated, the response of the system to the unit impulse. In general, this response does not have impulsive character. Only in ideal cases that are not physically realizable is the response of a system to an impulse also an impulse. In fact, the transmission of an impulse through a finite transmission system is always attended by a dispersion or opening of the pulse, sometimes accompanied by oscillations that may be small or negligible if the transmission system is properly designed. In this paper we show that systems that convert an impulse into a series of strong oscillations are also capable of transmitting with characteristics that are superior to those in which the impulse is transmitted with small dispersion. Of course, this statement refers to finite, passive systems.

Figure 1(I-1.4) shows an example of the response of a finite, linear system. Each oscillation is made equivalent to a window function and the complete oscillation is made equivalent to an array of windows. In finite systems with lumped elements "tails" appear after the principal oscillation. In this paper, the tails have a secondary effect and can be ignored, so that the function $S_c(t)$, indicated in Fig. 1(I-1.4), can be approximated by a window array with duration t_1 .

1.5 SYNTHESIS OF A LINEAR SYSTEM WHOSE RESPONSE $S_c(t)$ IS A PRESCRIBED WINDOW FUNCTION OF FINITE DURATION

We shall produce a method of synthesis of a finite, linear system whose response to a unit impulse is a window function of finite duration. We shall give a very simple and effective method of synthesis that employs vacuum tubes. The necessary equipment and concepts are well known and simple.

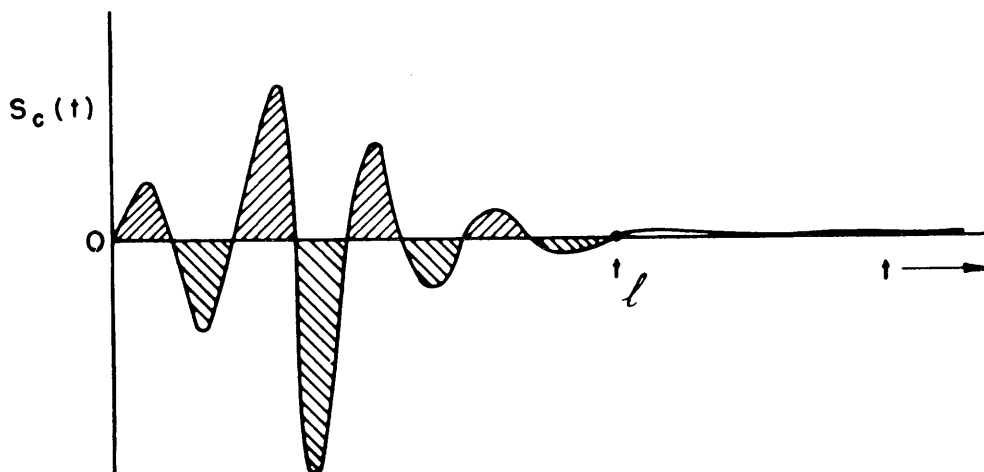


Fig. 1(I-1.4). The function $S_c(t)$ regarded as an array of window functions.

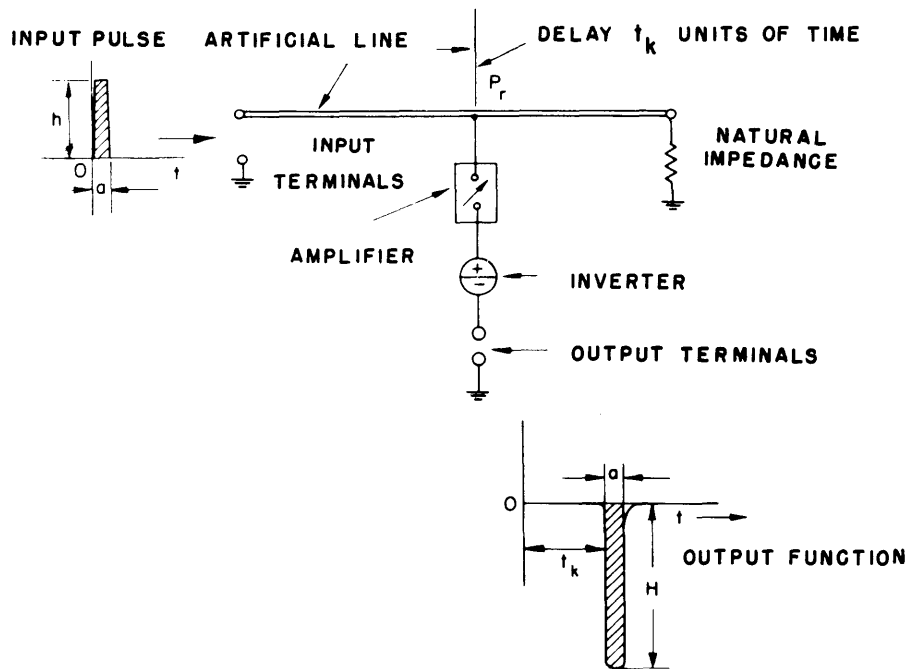


Fig. 1(I-1.51). Fundamental synthesis procedure.

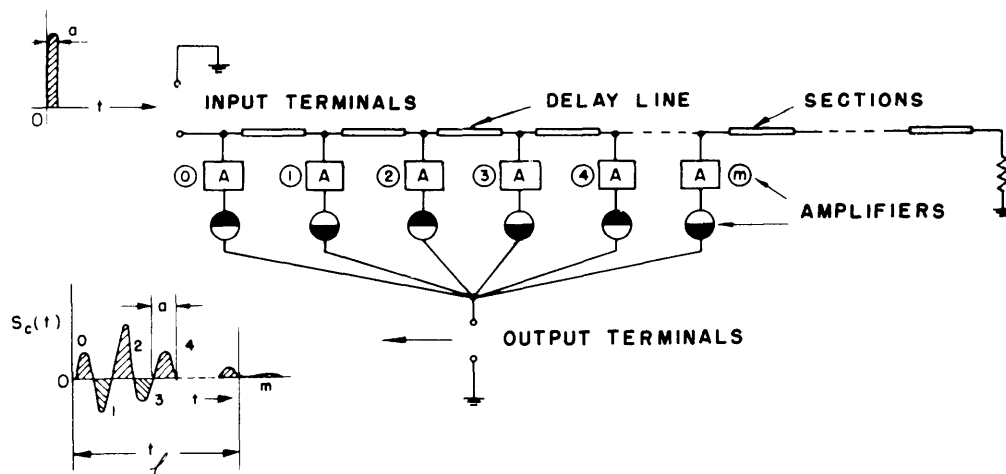


Fig. 2(I-1.51). Fundamental scheme for synthesizing a four-terminal system whose impulse response is a given function $S_c(t)$.

(VII. STATISTICAL COMMUNICATION THEORY)

The synthesis of such systems using passive elements – without vacuum tubes – can be accomplished by a complicated design procedure which cannot be given within the space allotted to this paper. A detailed discussion on this aspect of synthesis can be found in reference 1.

1.51 The basic synthesis elements employed here are: a finite artificial transmission line, linear amplifiers, and phase inverters. The fundamental disposition of the apparatus, and the basic operation of the system are illustrated in Fig. 1(I-1.51).

The artificial transmission line is terminated in its characteristic impedance in order to avoid reflections. The line is composed of a number N of equal sections connected in cascade, each section producing a delay of t_r units. Tapping at the k^{th} section produces the delay $t_k = kt_r$ unit. The maximum delay is $t_{\text{max}} = Nt_r$. In addition, the artificial line is designed to transmit a pulse of duration $t = a$ with negligible dispersion and ringing.

We tap the k^{th} section and feed a linear amplifier of variable gain g_k . The input impedance of the amplifier is practically infinite, to prevent appreciable reflections being caused by its insertion. The output terminals of the amplifier are connected to a device that allows us to invert the polarity of the received signal. Data on the design, construction, and operation of this system can be found in reference 2.

The operation of the system is as follows. Suppose, for example, that we want to synthesize a system whose response is a window of height H , negative, with aperture a , and delayed t_k units. It suffices to excite the system at the input terminals with a pulse of duration a , height h , positive. We tap the artificial line at the k^{th} section and insert the amplifier. We increase the gain of the amplifier until a pulse of height H is obtained and we set the inverter to the negative position. At the output terminals we shall then obtain the desired response.

To form the system whose response to a unit impulse is a prescribed array of windows we have only to repeat the arrangement of amplifiers and inverters, connecting them at the points of the delay line that are appropriate for producing the correct delay for each window. It is clear that the aperture of the excitation pulse should be less than, or equal to, the time t_r of delay of each section of the artificial line. The functioning of the system is obvious; we shall not give any more details. Figure 2(I-1.51) shows a sketch of such a system.

This system is designed to form an array of windows of equal apertures. By suitably selecting the aperture of the excitation pulse and the tap points on the delay line we can easily produce window arrays of variable aperture.

(VII. STATISTICAL COMMUNICATION THEORY)

II. FORMAL SOLUTION IN TERMS OF SINGULAR FUNCTIONS

2.0 PROCEDURE

We propose to solve the integral equation, given by the convolution integral, using singular functions for the function $S_c(t)$. That is to say, $S_c(t)$ will be obtained as an array of windows.

There are many ways of producing this kind of solution. The method used here was chosen to facilitate the solution of problems of pure transmission, and delayed and advanced transmission, which is our main objective. The formulation of the general problem of synthesis stipulated in section 0.6 allows the use of distributions of windows of finite duration, as will be shown in the following discussion. The solution in terms of distributions that have finite durations has an extremely important meaning in the theory of prediction which has been developed in connection with the problem of advanced transmission.

2.01 We shall consider two methods of solution. The first leads to a system of linear equations which, when solved, yield the solution to the problem. The second makes use of the theory of moments in a finite interval. Strictly speaking, the two methods are equivalent, but the forms of the solutions are better adapted to the solution of specific problems.

2.1 FIRST METHOD: EXCITATION BY FUNCTIONS OF THE CLASS C°

Let us consider the case in which the excitation function is a continuous function. This does not limit the generality of our discussion, since, from section 0.51, the acceptable functions contain, or can contain, a component of the class C° . Later, we shall study the effect of the stepped component and the impulsive component. The solution for the continuous component has, besides, theoretical as well as practical importance.

The convolution integral can be written

$$\gamma(t) = \int_0^t \phi(t-\tau) d\alpha(\tau) \quad 1(\text{II-2.1})$$

where $\alpha(\tau)$ is the distribution function associated with the distribution density $S_c(t)$, here supposed to be singular.

Let us consider a window array of finite duration.[†] The notation that we shall

[†]In the theory of linear prediction, this distribution is said to have finite "life." The reason for this terminology will be seen later.

(VII. STATISTICAL COMMUNICATION THEORY)

employ is indicated in Fig. 1(II-2.1). Successive impulses are a distance δ units of time from each other. In a linear finite system these impulses will show dispersion. We shall suppose that the aperture a of the dispersed pulse is equal to δ . Let us call a_k the measure, in the sense of Stieltjes, of the k^{th} impulse.

The Stieltjes integral 1(II-2.1) can be determined immediately and is given by

$$\gamma(t) = \sum_{k=0}^m a_k \phi[t - (k+1)\delta]; \quad t \geq t_\ell \tag{2(II-2.1)}$$

We see that the response of the system at time t is a linear combination with coefficients a_k , $k = 0, 1, 2, \dots, m$ of the values of the excitation function that lie between the times t and $t - t_\ell$, where t_ℓ is the life of the distribution of windows. It is also worthwhile to interpret Eq. 2(II-2.1) as the result of sampling with weights a_k the part of the excitation function that lies between the times t and $t - t_\ell$.

2.11 Let t_ℓ be the life of the distribution density that represents the function $S_c(t)$. Let us suppose that the excitation function $\phi(t)$ is regionally of the class $C^{(m)}$ in an interval t_ℓ . This condition is realized practically in an infinite variety of actual excitation functions. The discussion of the situation that corresponds to the exceptional points will be left to the end of this section. Under the previous assumptions, the excitation

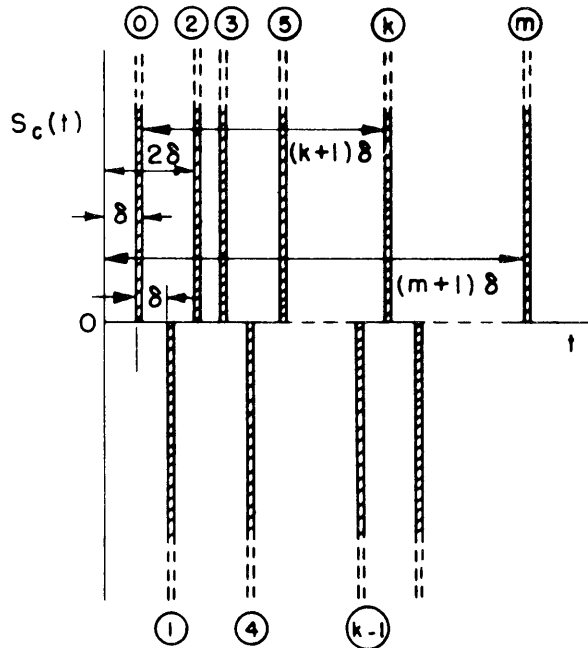


Fig. 1(II-2.1). Symbolic sketch of the singular functions which form $S_c(t)$.

(VII. STATISTICAL COMMUNICATION THEORY)

function can be expanded:

$$\phi[t - (k+1)\delta] = \phi(t) - (k+1)\delta \phi'(t) + \dots + (-1)^m \frac{(k+1)^m}{m!} \delta^m \phi^{(m)}(t) + \theta_{m+1}(t) \quad 1(\text{II-2.11})$$

where $\theta_{m+1}(t)$ is the appropriate remainder. Substituting in expression 2(II-2.1) and rearranging terms, we obtain

$$\gamma(t) = \sum_{p=0}^m \frac{y_p}{p!} \phi^{(p)}(t) + \theta_{m+1}(t) \quad 2(\text{II-2.11})$$

$$y_p = (-1)^p \sum_{k=1}^m \mu_k^p \alpha_k; \quad \mu_k = (1+k)\delta$$

Finally let us introduce the notation

$$\gamma^*(t) = \sum_{p=0}^m \frac{y_p}{p!} \phi^{(p)}(t) \quad 3(\text{II-2.11})$$

as the function associated with $\gamma(t)$. It is clear that $\gamma^*(t)$ is a simile of $\gamma(t)$ when the remainder $\theta_{m+1}(t)$ is small.

2.2 SOLUTION OF THE SYSTEM FOR α_k , $k = 0, \dots, m$

The relation 3(II-2.11) shows us that the function $\gamma^*(t)$ is a linear combination, with constant coefficients y_p , of the function $\phi(t)$ and its first m derivatives. By the nature of the synthesis problem that we are considering, the functions $\gamma^*(t)$, and $\phi(t)$ and its derivatives, are known. Therefore, the $m + 1$ constants y_0, y_1, \dots, y_m must be determined in such a way that the expression 3(II-2.11) is satisfied.

Introduction of the tolerance set ϵ_j (section 0.4) gives us some freedom in the choice of the function $\gamma^*(t)$. For example, we can choose a function $\gamma^*(t)$ which has closure in terms of the excitation functions. Under these conditions, the determination of the constant y_p becomes very easy if a set of orthogonal functions is introduced.

In this paper, our principal interest is in the problem of transmission, and its associated problems, in which the determination of the constants y_p , $p = 0, 1, 2, \dots, m$ is immediate. For example, the solution of the problem of pure transmission is obtained simply by setting

$$y_0 = 1; \quad y_1 = y_2 = \dots = y_m = 0$$

as can be seen by substituting these values in the expression 3(II-2.11), from which we obtain

(VII. STATISTICAL COMMUNICATION THEORY)

$$y^*(t) = \phi(t)$$

In the synthesis problems that are considered in this paper the y_p are easily determined. They will be determined later for special cases. The important thing now is to calculate the measures a_0, a_1, \dots, a_m that are associated with and define the function $S_c(t)$, which determines the desired linear system.

Consequently, we shall calculate the constants a_0, a_1, \dots, a_m as functions of the y_p and μ_p . We shall solve the system of $m + 1$ equations

$$y_p = (-1)^p \sum_{k=0}^m \mu_k^p a_k; \quad p = 0, 1, \dots, m \tag{II-2.2}$$

which contains $m + 1$ unknowns a_0, a_1, \dots, a_m .

2.21 The determinant of the system is the well-known Van der Monde determinant. This determinant is not zero when the elements $\mu_0, \mu_1, \dots, \mu_m$ are all different. This is our case, since

$$\mu_k = (k+1)\delta$$

Therefore, the homogeneous system 2(II-2.2) always has a solution. This establishes the existence of solutions to the synthesis problems considered in this paper.

2.22 It is convenient to introduce certain symmetric functions which simplify the solution of the system 2(II-2.2). We shall introduce the set of $m + 1$ symmetric function, associated with the unknown a_j , defined as follows

$$\left. \begin{aligned} s_m^{(j)} &= 1 \\ s_{m-1}^{(j)} &= -(\mu_0 + \mu_1 + \dots + \mu_{j-1} + \mu_{j+1} + \dots + \mu_m) \\ s_{m-2}^{(j)} &= (-1)^2 (\mu_0\mu_1 + \mu_0\mu_2 + \dots + \mu_0\mu_{j-1} + \mu_0\mu_{j+1} + \dots + \mu_0\mu_m + \\ &\quad + \dots + \mu_{m-1}\mu_m) \\ s_{m-3}^{(j)} &= (-1)^3 (\mu_0\mu_1\mu_2 + \mu_0\mu_1\mu_3 + \dots + \mu_0\mu_1\mu_{j-1} + \mu_0\mu_1\mu_{j+1} + \dots \\ &\quad + \mu_{m-2}\mu_{m-1}\mu_m) \\ &\dots \\ s_0^{(j)} &= (-1)^m \mu_0\mu_1\mu_2 \dots \mu_{j-1}\mu_{j+1} \dots \mu_m \end{aligned} \right\} \tag{II-2.22}$$

(VII. STATISTICAL COMMUNICATION THEORY)

Note that the coefficient μ_j does not enter into the formation of the symmetric functions associated with the unknown a_j . These symmetric functions have the same structure as that of the coefficients of a polynomial of the degree m whose roots are

$$\mu_0, \mu_1, \dots, \mu_{j-1}, \mu_{j+1}, \dots, \mu_m$$

In fact, the polynomial

$$G_j(\mu) = \mu^m + s_{m-1}^{(j)} \mu^{m-1} + s_{m-2}^{(j)} \mu^{m-2} + \dots + s_1^{(j)} \mu + s_0^{(j)} \quad 1(\text{II-2.22})$$

admits as roots the values μ_k , $k = 0, 1, \dots, j-1, j+1, \dots, m$. But μ_j is not a root, that is to say, $G_j(\mu_j) \neq 0$. The solution of the system of equations is given by the formula (see refs. 1 and 3):

$$a_j = \frac{y_0 s_0^{(j)} - y_1 s_1^{(j)} + y_2 s_2^{(j)} + \dots + (-1)^m y_m s_m^{(j)}}{(\mu_j - \mu_0)(\mu_j - \mu_1) \dots (\mu_j - \mu_{j-1})(\mu_j - \mu_{j+1}) \dots (\mu_j - \mu_m)}$$

$$= \frac{\sum_{p=0}^m (-1)^p y_p s_p^{(j)}}{\prod'_{p=0} (\mu_j - \mu_p)}; \quad j = 0, 1, \dots, m \quad 2(\text{II-2.22})$$

The prime on the product symbol, \prod' , shows that the factor $\mu_j - \mu_j$ has been omitted.

2.3 APPLICATION TO PURE TRANSMISSION

The case of pure transmission is given by the condition

$$\gamma^*(t) = \phi(t)$$

which implies, by virtue of expression 3(II-2.11), that

$$y_0 = 1; \quad y_1 = y_2 = \dots = y_m = 0$$

Therefore, the measures a_0, \dots, a_m of the window distribution that correspond to pure transmission are given by the formula

$$a_j = \frac{s_0^{(j)}}{\prod'_{p=0} (\mu_j - \mu_p)}$$

$$= \frac{(-1)^m \mu_1 \mu_2 \dots \mu_{j-1} \mu_{j+1} \dots \mu_m}{\prod'_{p=0} (\mu_j - \mu_p)} \quad 1(\text{II-2.3})$$

(VII. STATISTICAL COMMUNICATION THEORY)

A detailed study of the window array associated with these measures will be given in a later section.

2.4 APPLICATION TO DELAYED TRANSMISSION

This transmission is determined by the condition

$$\gamma(t) = \phi(t - t_o); \quad 0 < t_o = \text{constant} < \infty \quad 1(\text{II-2.4})$$

where t_o is the delay time. On the assumption that $\phi(t)$ is regionally at least of the class $C^{(m)}$, we have

$$\gamma(t) = \phi(t) - t_o \phi'(t) + \frac{t_o^2}{2!} \phi''(t) - \dots + (-1)^m \frac{t_o^m}{m!} \phi^{(m)}(t) + \theta_{m+1}(t)$$

Let us take as our simile function

$$\gamma^*(t) = \sum_{k=1}^m (-1)^k \frac{t_o^k}{k!} \phi^{(k)}(t) \quad 2(\text{II-2.4})$$

Comparing this expression with 3(II-2.11), we obtain immediately the values of y_o, \dots, y_m . Their expression is

$$y_p = (-1)^p t_o^p \quad 3(\text{II-2.4})$$

Substituting this expression in formula 2(II-2.22), we obtain the measures of the windows associated with the distribution corresponding to delayed transmission:

$$\begin{aligned} a_j &= \frac{t_o^m + t_o^{m-1} s_{m-1}^{(j)} + t_o^{m-2} s_{m-2}^{(j)} + \dots + t_o s_1^{(j)} + s_o^{(j)}}{(\mu_j - \mu_o)(\mu_j - \mu_1) \dots (\mu_j - \mu_{j-1})(\mu_j - \mu_{j+1}) \dots (\mu_j - \mu_m)} \\ &= \frac{G(t_o)}{\prod_{p=0}^m (\mu_j - \mu_p)} = \frac{\prod_{p=0}^m (t_o - \mu_p)}{\prod_{p=0}^m (\mu_j - \mu_p)} \quad 4(\text{II-2.4}) \end{aligned}$$

A detailed study of the window array associated with these measures a_o, \dots, a_m will be given in a later section.

2.5 APPLICATION TO ADVANCED TRANSMISSION – LINEAR PREDICTION

This transmission is characterized by the condition

$$\gamma(t) = \phi(t + t_o); \quad 0 < t_o = \text{constant} < \infty \quad 1(\text{II-2.5})$$

(VII. STATISTICAL COMMUNICATION THEORY)

Expanding in a Taylor series and putting

$$\gamma^*(t) = \sum_{k=0}^m \frac{t_0^k}{k!} \phi^{(k)}(t) \quad 2(\text{II-2.5})$$

we obtain

$$y_p = t_0^p \quad 3(\text{II-2.5})$$

The measures of the windows now turn out to be

$$a_j = \frac{G(-t_0)}{\prod_{p=0}^m (\mu_1 - \mu_p)} = (-1)^m \frac{\prod_{p=0}^m (t_0 + \mu_p)}{\prod_{p=0}^m (\mu_j - \mu_p)} \quad 4(\text{II-2.5})$$

It should be noted that expression 1(II-2.5) is valid only for $t > t_\ell$, where t_ℓ is the life of the kernel, since relations 2(II-2.1) are valid only for $t > t_\ell$. What happens for the values $0 \leq t < t_\ell$ will be discussed later.

A detailed study of the problem of linear prediction, and of the form of the system function $S_c(t)$ associated with prediction, will also be given later.

2.6 A COLLATERAL PROBLEM. CIRCUITS WHOSE RESPONSE IS THE j^{th} DERIVATIVE OF THE EXCITATION FUNCTION

Equations 2(II-2.11) and 2(II-2.22) allow the immediate determination of the measures of the window distributions associated with the system function $S_c(t)$ whose response is some given derivative, let us say the j^{th} ($0 < j \leq m$), of the excitation function.

In expression 3(II-2.11) let us set

$$y_p = \begin{cases} 0 & \text{for } p = 0, 1, \dots, k-1, k+1, \dots, m \\ p! & \text{for } p = k \end{cases} \quad 1(\text{II-2.6})$$

We have

$$\gamma^*(t) = \phi^{(p)}(t)$$

The measures of the windows will now be

$$a_j = \frac{(-1)^k y_k s_k^{(j)}}{\prod_{p=0}^m (\mu_j - \mu_p)} \quad 2(\text{II-2.6})$$

A detailed study of the distribution corresponding to these measures will be given in a later chapter.

2.7 DETERMINATION OF THE ERROR INCURRED BY USING THE FUNCTION $\gamma^*(t)$

The error committed by replacing the exact response function $\gamma(t)$ by the simile function $\gamma^*(t)$ is given by the expression

$$\xi_{(m)} = \gamma(t) - \gamma^*(t) \quad 1(\text{II-2.7})$$

Using 2(II-2.1) and the abbreviation

$$\mu_k = (k+1)\delta$$

we have

$$\gamma(t) = \sum_{k=0}^m a_k \phi(t - \mu_k) = \sum_{k=0}^m a_k \left[\sum_{j=0}^m \frac{(-\mu_k)^j}{j!} \phi^{(j)}(t) + \theta_{m+1}^{(k)}(t) \right]$$

where $\theta_{m+1}^{(k)}(t)$ is the remainder of the Taylor series after the m^{th} term, and the superscript (k) indicates that the remainder in question is the one that corresponds to the Taylor series for $\phi(t - \mu_k)$. Using definitions 2(II-2.11) and 3(II-2.11), we have

$$\gamma(t) = \gamma^*(t) + \sum_{k=0}^m a_k \theta_{m+1}^{(k)}(t)$$

so that, substituting in 1(II-2.7), the error is given by

$$\xi_{(m)} = \sum_{k=0}^m a_k \theta_{m+1}^{(k)}(t) \quad 2(\text{II-2.7})$$

From Taylor's theorem, the k^{th} remainder $\theta_{m+1}^{(k)}(t)$ is given by

$$\theta_{m+1}^{(k)}(t) = \frac{\mu_k^{m+1}}{(m+1)!} \phi^{(m+1)}(t + \eta_k \mu_k)$$

where η_k is some constant between zero and one which, in general, is different for different values of k . Therefore, finally, the general expression for the error is

$$\xi_{(m)} = \sum_{k=0}^m \frac{a_k \mu_k^{(m+1)}}{(m+1)!} \phi^{(m+1)}(t + \eta_k \mu_k) \quad 3(\text{II-2.7})$$

(VII. STATISTICAL COMMUNICATION THEORY)

Clearly, the value of the sum depends very strongly on the behavior of the $(m+1)^{-st}$ derivative of the excitation function $\phi(t)$. The sum extends over a time interval bounded, at most, by $\eta_o \mu_o = 0$ and $\eta_m \mu_m = (m+1)\delta$, so that the largest time span involved in the sum is t_ℓ seconds. If $\phi^{(m+1)}(t)$ varies slowly, so that it is essentially constant over any interval of length t_ℓ , the expression for the error can be simplified to

$$\xi_{(m)} \approx \frac{\phi^{(m+1)}(t)}{(m+1)!} \sum_{k=0}^m a_k \mu_k^{m+1} = \frac{\phi^{(m+1)}(t) \delta^{m+1}}{(m+1)!} \sum_{k=0}^m (1+k)^{m+1} a_k \quad 4(\text{II-2.7})$$

It is desirable to express the error in terms of the modulus of oscillation of $\phi^{(m)}(t)$. Since we have assumed that $\phi^{(m+1)}(t)$ is almost constant within an interval t_ℓ , it is certainly reasonable to assume that it is constant over an interval of length $\delta = \frac{t_\ell}{m+1}$. Therefore, within any interval of length δ , $\phi^{(m)}(t)$ is approximately linear, and its total change is given by $\phi^{(m+1)}(t) \delta$. Thus, the modulus of oscillation of $\phi^{(m)}(t)$ in an interval δ , $\Omega^{(m)}(\delta)$, is given, approximately, by

$$\Omega^{(m)}(\delta) \approx \left| \phi^{(m+1)}(t) \delta \right|$$

Substituting in the expression for the error, we obtain

$$\left| \xi_{(m)} \right| \approx \frac{\Omega^{(m)}(\delta) \delta^m}{(m+1)!} \left| \sum_{k=0}^m (1+k)^{m+1} a_k \right| \quad 5(\text{II-2.7})$$

When the function $\phi^{(m+1)}(t)$ oscillates rapidly within the duration t_ℓ of the kernel, formula 5(II-2.7) may not be used. However, the formula is more generally valid than is realized at first sight. The reason is simple: nothing has been said about the length t_ℓ of the life of the kernel. It is clear that if we take t_ℓ smaller and smaller, we shall eventually reach a value of t_ℓ sufficiently small so that, within that interval, the oscillation of $\phi^{(m)}(t)$ will be as small as desired, when t is within one of the regions of continuity of $\phi^{(m)}(t)$. This is a consequence of the hypothesis of regional continuity of $\phi^{(m)}(t)$.

2.71 Two important conclusions can be obtained from the error expression, 5(II-2.7):

1. For a window distribution of sufficiently small life, the regional error committed in the described synthesis process is proportional to the oscillation of the m^{th} derivative of the excitation.

2. In the regions of continuity of $\phi^{(m)}(t)$, the error committed decreases with a decrease in the life of the distribution of windows.

(VII. STATISTICAL COMMUNICATION THEORY)

2.8 CONNECTION BETWEEN THE TOLERANCE SET ϵ_j AND THE ERROR

We have the first elements that allow us to establish relations between the elements of the tolerance set $[\epsilon_j]$ and the error $\xi_{(m)}$. Let us consider the interval $0 < t < \infty$. We shall omit the following sets of measure zero:

- a. the countable set of isolated points at which the function $\phi(t)$, considered here in its general form, exhibits impulsive behavior;
- b. the similar set of points at which the function $\phi(t)$ exhibits simple discontinuities;
- c. the set of points, here supposed to be isolated and countable, at which the derivatives $\phi^{(p)}(t)$, $p = 1, \dots, m$ exhibit a discontinuity.

The function $\phi(t)$ is of class $C^{(m)}$ on the countable set of open subintervals which form the complement of the sets indicated in a, b, and c. By construction, we know that a tolerance ϵ_k corresponds to the k^{th} interval. Let us designate by $\Omega_k^{(m)}(\delta)$ the modulus of oscillation of the function $\phi^{(m)}(t)$ in the k^{th} subinterval. Then to this subinterval let us associate the error given by the expression

$$|\xi_{(m), k}| = \frac{\Omega_k^{(m)}(\delta) \delta^m}{(m+1)!} \left| \sum_{p=0}^m (p+1)^{m+1} a_p \right| \quad 1(\text{II-2.8})$$

If we now take a distribution of windows of sufficiently short life, we shall have

$$\epsilon_k \geq |\xi_{(m), k}| \quad 2(\text{II-2.8})$$

Relations 2(II-2.8) are important in the process of synthesis that is being studied because they establish simple relations between the separation of the impulses that form the distribution, the number $m + 1$ of impulses, and the oscillation of the m^{th} derivative of the excitation function. In the sections that treat the problems of pure, delayed, and advanced transmission specifically we shall calculate the error associated with each one of those problems.

Manuel V. Cerrillo

[Chapters 3, 4, and 5 will be published in the Quarterly Progress Report, Jan. 15, 1957.]

References

1. M. V. Cerrillo and E. F. Bolinder, Technical Report 246, Research Laboratory of Electronics, M.I.T., Aug. 1952.
2. C. A. Stutt, Technical Report 182, Research Laboratory of Electronics, M.I.T., May 1951.
3. M. V. Cerrillo, Technical Report 270, Research Laboratory of Electronics, M. I. T., to be published.