

SAMPLING MODELS FOR LINEAR TIME-VARIANT FILTERS

THOMAS KAILATH

TECHNICAL REPORT 352

MAY 25, 1959

**MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS
CAMBRIDGE, MASSACHUSETTS**

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the U.S. Army (Signal Corps), the U.S. Navy (Office of Naval Research), and the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039-sc-78108, Department of the Army Task 3-99-20-001 and Project 3-99-00-000.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report 352

May 25, 1959

SAMPLING MODELS FOR LINEAR TIME-VARIANT FILTERS

Thomas Kailath

This report is based on a thesis submitted to the Department of Electrical Engineering, M.I.T., May 25, 1959, in partial fulfillment of the requirements for the degree of Master of Science.

Abstract

A large class of communication channels can be represented by linear time-variant filters. Different constraints can be imposed on these filters in order to simulate the actual operating conditions of such channels. The constraints permit the original filter to be replaced by another (simpler) filter that imitates the original filter under the operating constraints. These new filters need not resemble the physical channel at all, and need not be equivalent to the actual channel, except under the given constraints.

In this report, the constraints considered are those of finite input and output channel signals and finite channel memory. Other cases can be studied by similar methods. Methods of characterizing linear time-variant filters are investigated in order to determine the most convenient descriptions for the different constraints. These descriptions are used to obtain sampling theorems and models for the filter under the various constraints. The theorems are used to find the conditions under which a linear time-variant filter can be determined by input-output measurements only.

I. INTRODUCTION

Background of the Problem

The classical problem of electrical communications is that of securing reliable transmission of signals through a communication channel. A frequently used channel is the ionosphere. The chief obstacles to obtaining reliable signal transmission via the ionosphere are atmospheric noise and time-variant multipath propagation. The characteristics of atmospheric noise, which is largely caused by lightning discharges, have been extensively studied (1,2,3). The main defense against it is sufficiently large signal power. Atmospheric noise is an additive disturbance and therefore, in a sense, is extrinsic to the actual physical communication medium. On the other hand, multipath disturbances – manifested most commonly as selective fading, frequency distortion, and intersymbol interference – directly involve the channel itself. Multipath problems are the result of dispersive propagation by paths of different electrical (and/or physical) lengths.

Several schemes for fighting multipath propagation have been devised. The most successful at present are diversity reception (4) and the Rake system (5). The commonest forms of diversity reception are space and frequency diversity. These take advantage of the fact that signals received at different locations, or signals of slightly different frequency, do not fade synchronously. Therefore, by using two (or more) receivers and weighting the received signals appropriately, reliability is improved. The philosophy of the Rake system is to perform a continuous measurement of the multipath characteristic, which is then employed to combat the multipath propagation effectively.

The foundations of Rake were laid chiefly by the excellent communication theoretical studies of scatter-multipath channels made by R. Price (6, 7). Related work has also been done by W. Root and T. Pitcher (8) and by G. L. Turin (9). These theoretical studies, however, all suffer from the restricted nature of the channel models on which they are based. [A critique of these models and a general review of statistical multipath communication theory was recently given by G. L. Turin (10).]

The operational success of the Rake system encourages a more general study of the communication channel. We can regard the channel as a time-variant filter with additive noise superimposed on the output. In these general terms, of course, little of a specific nature can be said about the problem. But in communications systems there are certain additional constraints present: Signals are of finite time-bandwidth product, the channel is nearly linear, and so forth. Introducing these constraints into the problem should enable us to obtain restricted, but simpler, models for the filter – models that are more useful for our purposes. These models will imitate the operation of the filter under the imposed constraints but may not bear any physical resemblance to the original filter and may not imitate the filter under other operating conditions. This report is devoted to methods of obtaining such models under the constraints of linearity and limited bandwidth or limited duration of channel memory. Some properties and methods of description and analysis of such models are also studied.

II. CHARACTERIZATION OF LINEAR TIME-VARIANT NETWORKS

A time-variant (or time-varying, or time-variable) network is one whose input-output relationship is not invariant under translations in time. If, in addition, the superposition principle holds for the network, we have a linear time-variant network. In communications engineering linear time-variant networks have long been in use, especially as modulators and oscillators. Such systems, which usually contain only a single variable element, have been widely investigated, chiefly by mathematicians, and several theories and methods of solution (11, 12) — Floquet theory, Mathieu functions, the B. W. K. method, for example — have been developed, although these are not yet in the most suitable form for engineering application. In present-day communication theory, however, interest has shifted to time-variant systems of a more general nature, the behavior of which is governed by linear, and not necessarily differential, operators. L. A. Zadeh of Columbia University, who pioneered the investigation of such systems, introduced the concept of a frequency function for linear time-variant networks — a concept that has been extremely useful.

In this section we shall discuss rather briefly some methods of characterizing linear time-variant networks and describe some forms of specification that we have found useful in our analysis. The meanings and possible interpretations of the time and frequency variables that are introduced are considered. These are of service in Section III. (In this report, the terms network and filter often are used interchangeably.)

2.1 Characterization of Linear Time-Variant Networks

The most commonly used methods of characterizing a linear time-variant network, that is, of specifying its input-output relationship, are those that give (a) the differential

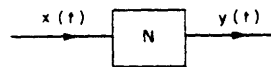


Fig. 1. Linear time-variant network, N.

equation of the network, (b) the impulse response of the network, (c) the frequency response of the network, and (d) special techniques that are valid only for particular classes of networks. In discussing these methods, we shall consider a two-terminal-pair network, N, with input $x(t)$ and output $y(t)$. (See Fig. 1.)

2.11 The Differential Equation Method

This has been the classical way of describing linear time-variant networks. We have a linear differential equation with variable coefficients relating $y(t)$ and $x(t)$:

$$\left[a_n(t) p^n + \dots + a_1(t) p + a_0 \right] y(t) = \left[b_m(t) p^m + \dots + b_1(t) p + b_0 \right] x(t) \quad (1)$$

or, more simply,

$$L(p, t) y(t) = K(p, t) x(t) \quad (2)$$

where p is the differential operator d/dt , and $L(p, t)$ and $K(p, t)$ represent respectively the left-hand and right-hand operations in Eq. 1.

Such equations have been investigated extensively by mathematicians and have been used in the study of oscillator and modulator circuits. A review of some of the techniques used for their solution, which often can only be numerical or approximate, is given by Pipes (11) and Bennett (12). This method of characterization is useful for studies of network response to a fixed input such as a sine wave or a constant. However, in modern communication theory and in modern signal theory, the emphasis is not on fixed inputs but on inputs that belong to a class of functions. Thus, for example, the input may be a member of a class of bandlimited functions or of an ensemble of random functions. In such cases, characterization by impulse response or frequency response is more suitable, although some special cases can still be handled by differential equations (13). Another reason for preferring an impulse and frequency response description is that these response functions can often be directly determined by experiment, whereas it is usually difficult, if not impossible, to so determine the differential equation of the network. However, it is clear that, theoretically, all three specifications are interrelated; a discussion of the relationships is given in reference 14.

2.12 The Impulse Response and the Frequency Response

The impulse response of a linear time-variant network is defined as $h_1(t, \tau)$, the response to an impulse input at time τ measured at time t . For a physically realizable network, $h_1(t, \tau)$ is zero for $t < \tau$.

Since the input $x(t)$ can be regarded as being composed of weighted impulses,

$$x(t) = \int_{-\infty}^t x(\tau) \delta(t-\tau) d\tau \quad (3)$$

we can write, by virtue of linearity,

$$y(t) = \int_{-\infty}^t x(\tau) h_1(t, \tau) d\tau \quad (4)$$

which, for a realizable network, can also be written

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h_1(t, \tau) d\tau \quad (5)$$

because then $h_1(t, \tau) = 0$, with $t < \tau$.

The frequency response function is defined (15) by the relation*

$$H_1(j\nu, t) = \int h_1(t, \tau) e^{-j2\pi\nu(t-\tau)} d\tau \quad (6)$$

Using Eq. 5 we can write this as

$$H_1(j\nu, t) = \frac{\text{response of the network } N \text{ to } \exp(j2\pi\nu t)}{\exp(j2\pi\nu t)} \quad (7)$$

which explains why it is called the frequency response function. By virtue of linearity, it is then easily deduced that

$$y(t) = \int H_1(j\nu, t) X(j\nu) e^{j2\pi\nu t} d\nu \quad (8)$$

In reference 15 Zadeh further defines a bifrequency function

$$\Gamma(j\nu, j\mu) = \iint h_1(t, \tau) e^{j2\pi\nu\tau} e^{-j2\pi\mu t} d\tau dt \quad (9)$$

For later comparisons, it is convenient to introduce

$$H_1(j\nu, j\mu) = \int H_1(j\nu, t) e^{-j2\pi\mu t} dt \quad (10)$$

so that (as is easily verified)

$$\Gamma(j\nu, j\mu) = H_1[j\nu, j(\mu-\nu)] \quad (11)$$

In a time-invariant network, $h_1(t, \tau)$ would be a function of $(t-\tau)$ only, and not of t and τ separately, and $H_1(j\nu, t)$ would be independent of t . Thus it would appear reasonable, in the time-variant case, to regard the frequency variable μ as corresponding to the rate of change of the system characteristics. (These points will be discussed in more detail in section 2.2.) Similarly, from Eq. 8, it would appear that the frequency variable ν is associated with input frequencies to the network.

2.13 Other Methods of Characterization

For special classes of linear time-variant networks, simpler methods of description have been suggested. For example, Aseltine (16) derived integral transforms for systems characterized by a special second-order differential equation in such a way that the frequency response function defined in terms of this transform is independent of time. For periodically varying systems, Pipes (11) developed a matrix method of solution.

* ν is in cycles per second. All frequency variables in this report will be expressed in cyclic measure. When limits of integration or summation are not explicitly indicated they are to be taken as $(-\infty, \infty)$.

A. P. Bolle (17) used the usual complex symbolism of ac circuits to solve a network with periodically varying elements. The use of complex symbolism reduces the differential equation with variable coefficients to a complex equation with constant coefficients.

It is also of interest to point out that the impulse and frequency response characterizations, based on the responses of the network to unit impulses and exponential functions, are special cases of a general technique in which a network is characterized by its responses to a set of "elementary" time functions. We shall not discuss this here; a full description can be found in reference 18. This development has also encouraged application of the theory of linear vector spaces to the theory of linear time-variant networks (19).

2.14 Other Forms for the Impulse Response and Frequency Functions

In the function $h_1(t, \tau)$ the realizability condition is that the response be identically zero for $t < \tau$. This constraint involves both the variables t and τ , and therefore is often inconvenient to use. In the alternate forms of impulse response now to be described, the realizability condition involves only one variable. Furthermore, in Section III, we shall have to impose frequency and time restrictions on the impulse and frequency responses, and this is not conveniently done with $h_1(t, \tau)$ in all cases. Thus, for example, if we have a restriction on the output frequency range of the linear time-variant network or on the duration of the impulse response of the network, it is not immediately clear how these are reflected in $h_1(t, \tau)$. The forms that we shall introduce will turn out to be more convenient for such restrictions. Another feature is that these forms will exhibit direct Fourier transform relationships between the frequency and impulse response functions.

We define

$h_2(z, \tau)$ = response to a unit impulse input at time t ,
measured at time $t = \tau + z$.

$h_3(y, t)$ = response measured at time t to a unit impulse
input at time $t - y$.

Thus z measures elapsed time, and y measures the age of the input. The realizability conditions are zero response for $z < 0$ and $y < 0$, respectively.

Of course, $h_1(t, \tau)$, $h_2(z, \tau)$, and $h_3(y, t)$ must all be related. The rules governing transformation from one form to another are given in section 2.15. They are derived from the relations $z = t - \tau = y$ between the time-domain variables z , t , τ , and y .

2.141 The Form $h_2(z, \tau)$

In terms of $h_2(z, \tau)$, the operation of the linear time-variant network can be conveniently pictured as in Fig. 2, which displays on a $z - \tau$ plane several network responses to impulse inputs at different times, T . Notice, again, that the variable z in $h_2(z, \tau)$

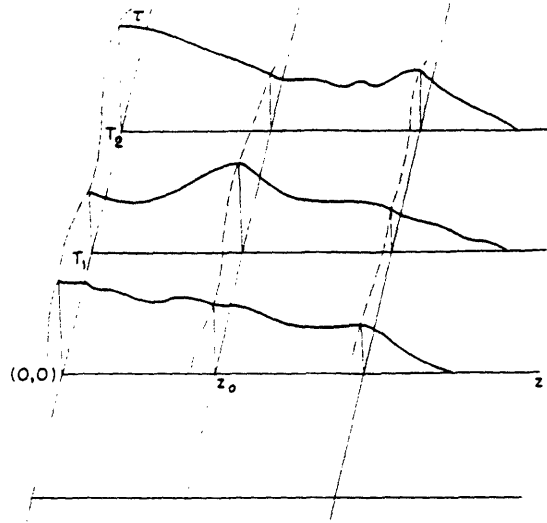


Fig. 2. The impulse response $h_2(z, \tau)$.

refers to duration of the input time function. If we had a fixed network, the response to a unit impulse at time T_1 would be the same as the response to a unit impulse at time T_2 . In terms of Fig. 2, then, it would appear that the variation of $h_2(z, \tau)$ with τ , for fixed z , would be a measure of the rate of variation of the system. We could find the Fourier transform of $h_2(z, \tau)$ with respect to τ for fixed z ,

$$H_2(z, j\mu) = \int h_2(z, \tau) e^{-j\mu\tau} d\tau \quad (12)$$

and the variable μ would be a frequency domain measure of the variation of the system. If μ were confined to low values, the system would be varying slowly; it would be varying rapidly if there were high frequencies in the μ domain. We shall see later (in section 2.21) that this interpretation of μ is consistent with the one previously given in the case of $H_1(j\nu, j\mu)$ and also with other interpretations of system variation that we shall obtain. Therefore the variable μ here has the same significance that it had in $H_1(j\nu, j\mu)$.

We may also define Fourier transforms with respect to z , keeping τ fixed:

$$H_2(j\omega, \tau) = \int h_2(z, \tau) e^{-j2\pi\omega z} dz \quad (13)$$

and

$$H_2(j\omega, j\mu) = \int H_2(j\omega, \tau) e^{-j2\pi\mu\tau} d\tau \quad (14)$$

$$= \int H_2(z, j\mu) e^{-j2\pi\omega z} dz \quad (15)$$

$$= \iint h_2(z, \tau) e^{-j2\pi\omega z} e^{-j2\pi\mu\tau} dz d\tau \quad (16)$$

We shall show later that the variable ω can be regarded as corresponding to the frequency of the output waveform.

2.142 The Form $h_3(y, t)$

Similarly for $h_3(y, t)$ [this form has been used by other writers, including Zadeh and Bendat (20)] we obtain by direct Fourier transformation the set of functions

$$H_3(j\omega, t) = \int h_3(y, t) e^{-j2\pi\omega y} dy \quad (17)$$

$$H_3(y, j\mu) = \int h_3(y, t) e^{-j2\pi\mu t} dt \quad (18)$$

$$H_3(j\nu, j\mu) = \int H_3(j\nu, t) e^{-j2\pi\mu t} dt \quad (19)$$

$$= \int H_3(y, j\mu) e^{-j2\pi\nu y} dy \quad (20)$$

$$= \iint h_3(y, t) e^{-j2\pi\nu y} e^{-j2\pi\mu t} dy dt \quad (21)$$

The last two functions are exactly equivalent to $H_1(j\nu, t)$ and $H_1(j\nu, j\mu)$. This is because, using Eq. 5, we can write

$$\begin{aligned} H_3(j\nu, t) &= \frac{\text{response of } N \text{ to } \exp(j2\pi\nu t)}{\exp(j2\pi\nu t)} \\ &= H_1(j\nu, t) \end{aligned}$$

and therefore, by virtue of Eqs. 10 and 19,

$$H_3(j\nu, j\mu) = H_1(j\nu, j\mu)$$

Thus the frequency variables for $h_3(y, t)$ have the same significance as those for $h_1(t, \tau)$, and, therefore, we have used the same symbols, ν and μ , in both cases. However, note that $H_3(j\nu, t)$ and $h_3(y, t)$ are related directly by a Fourier transform, which is not true of $H_1(j\nu, t)$ and $h_1(t, \tau)$. (Cf. Eq. 6.)

Finally, the form of the input-output relation

$$y(t) = \int_0^\infty h_3(y, t) x(t-y) dy$$

(which is derived in Appendix I) suggests that $h_3(y, t)$ can be interpreted as a weighting function by which the signal inputs in the past must be multiplied to determine their

contributions to the present output. The realizability condition, $h_3(y, t) = 0$ for $y < 0$, thus reflects the fact that the filter cannot weight portions of the input that have yet to occur.

2.15 Summary of Relations Involving the Different Forms of Impulse and Frequency Response

For convenience, we now list the relationships between the various forms of impulse and frequency response that we have introduced and also give the convolution integral formulas connecting the input and output functions. Proofs of all but the most immediately evident relations listed here are given in Appendix I.

a. Transformations between $h_1(t, \tau)$, $h_2(z, \tau)$, $h_3(y, t)$

$$h_1(t, \tau) = h_2(t-\tau, \tau) \quad (22)$$

$$h_2(z, \tau) = h_1(z+\tau, \tau)$$

$$h_2(z, \tau) = h_3(z, z+\tau) \quad (23)$$

$$h_3(y, t) = h_2(y, t-y)$$

$$h_3(y, t) = h_1(t, t-y) \quad (24)$$

$$h_1(t, \tau) = h_3(t-\tau, t)$$

b. Input-output relations – time domain

$$y(t) = \int_{-\infty}^t h_1(t, \tau) x(\tau) d\tau = \int_0^{\infty} h_1(t, t-\tau) x(t-\tau) d\tau \quad (25)$$

$$= \int_{-\infty}^t h_2(t-z, z) x(z) dz = \int_0^{\infty} h_2(z, t-z) x(t-z) dz \quad (26)$$

$$= \int_{-\infty}^t h_3(t-y, t) x(y) dy = \int_0^{\infty} h_3(y, t) x(t-y) dy \quad (27)$$

c. Transformation between $H_1(j\nu, j\mu)$, $H_2(j\omega, j\mu)$, $H_3(j\nu, j\mu)$

$$H_1(j\nu, j\mu) = H_2[j(\nu+\mu), j\mu] = H_3(j\nu, j\mu) \quad (28)$$

$$H_2(j\omega, j\mu) = H_3[j(\omega-\mu), j\mu] = H_1[j(\omega-\mu), j\mu] \quad (29)$$

$$H_3(j\nu, j\mu) = H_1(j\nu, j\mu) = H_2[j(\nu+\mu), j\mu] \quad (30)$$

d. Input-output relations – frequency domain

$$Y(j\mu) = \int_{-\infty}^{\infty} H_1[j\nu, j(\mu-\nu)] X(j\nu) d\nu \quad (31)$$

$$= \int_{-\infty}^{\infty} H_2[j\mu, j(\mu-\omega)] X(j\omega) d\omega \quad (32)$$

$$= \int_{-\infty}^{\infty} H_3[j\nu, j(\mu-\nu)] X(j\nu) d\nu \quad (33)$$

e. Interpretations of the variables: definitions of variables and their physical significance

Time Domain

- t: variable corresponding to instant of observation
- τ : variable corresponding to instant of impulse input
- z: variable corresponding to elapsed time
- y: variable corresponding to age of input

Frequency Domain

- μ : variable corresponding to system variation
- ν : variable corresponding to input frequencies
- ω : variable corresponding to output frequencies

2.2 Bandwidth Relations in Linear Time-Variant Networks

When we study the frequency behavior of linear time-variant networks the time-variant character of the network is usually evidenced by a frequency expansion, or a frequency shift, or both. Thus if we put in a sine wave of frequency ν_0 , the output is usually a band of frequencies centered about ν_0 , or a single sine wave of a different value, or a band of frequencies centered about a frequency ν that is different from ν_0 .

In the first case we can use the frequency expansion as a measure of the rate of variation of the system; a small expansion indicates a slow variation, and a large expansion a rapid variation. The amount of expansion produced may often depend on the particular frequency of the input to the networks. Therefore we shall pick the largest expansion, over all possible input frequencies, as a measure of the system variation. We shall denote this by W_s and call it the filter (system, network) bandwidth. This first case is often encountered in scatter-multipath situations.

The second case usually arises in amplitude modulation or in Doppler radar. Thus, for example, in Doppler radar, a sine wave of frequency ν_0 has a new frequency, after reflection from a body moving away from it at v meters per second,

of $\left(\frac{c+v}{c-v}\right) \nu_0$ cps. In such cases the amount of frequency shift, λ , can be regarded as a measure of the time-variation of the system. When both frequency shift and expansion are present, neither W_s nor λ by themselves will be, in general, an adequate measure of the system variation. It would be more appropriate to consider some combination of W_s and λ as a proper measure, but the nature of the combination would depend on the particular situation considered. In this report we shall be concerned only with situations of the first type – simple frequency expansion.

2.21 The Filter Variation and the Variable μ

We pointed out earlier that it seemed reasonable to associate the variable μ with the variation of the system. We shall now show that this interpretation is consistent with the notion of filter bandwidth.

Using Eqs. 26, 27, and 28, we see that if $x(t)$ is a sinusoid of frequency ν_0 – that is, if $x(t) = \exp(j2\pi\nu_0 t)$ and $X(j\nu) = \delta(\nu - \nu_0)$, then the output is given by

$$W'_s = Y(j\mu) = H_1[j\nu, j(\mu - \nu)] \quad (34)$$

$$= H_2[j\mu, j(\mu - \nu)] \quad (35)$$

$$= H_3[j\nu, j(\mu - \nu)] \quad (36)$$

Therefore W'_s is nonzero for the range of μ -values over which $H_1(j\nu, j\mu)$, $H_2(j\omega, j\mu)$, $H_3(j\nu, j\mu)$ are nonzero, and therefore the maximum μ -bandwidth of $H_1(j\nu, j\mu)$, $H_2(j\omega, j\mu)$, $H_3(j\nu, j\mu)$ is defined by $W_s = \max W'_s$ for all ν . Therefore if filter bandwidths W_s – that is, measures of the rate of variation of the system – are specified, the appropriate frequency variable to be considered is μ . Note, also, from relations 34 and 36 and the convolution formulas 31 and 33, that an input of bandwidth W_i to a linear time-variant network of filter variation bandwidth W_s results in an output of bandwidth greater than or equal to W_i but not greater than $W_i + W_s$.

2.22 Input Bandwidth and the Variable ν

We have found that

$$H_1(j\nu, t) = H_3(j\nu, t)$$

$$= \frac{\text{response of linear time-variant network to } \exp(j2\pi\nu t)}{\exp(j2\pi\nu t)}$$

Therefore, if we are interested in determining the response of the linear time-variant network to an input that is nonzero only over particular frequency ranges, we need only consider $H_1(j\nu, j\mu)$ and $H_3(j\nu, j\mu)$ for values of ν in these ranges. Thus, if input bandwidth restrictions are specified, the appropriate frequency variable to consider is ν .

2.23 Output Bandwidth and the Variable ω

We know that if we have an input, bandlimited to $(-W_i, W_i)$, for a linear time-variant network with filter variation bandwidth $2W_s$, the bandwidth of the output signal is restricted to $(-W_s - W_i, W_s + W_i)$. With this in mind, and noting the relations

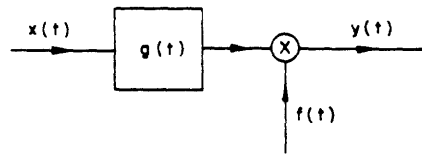
$$H_1(j\nu, j\mu) = H_2[j(\nu+\mu), j\mu]$$

$$H_2(j\omega, j\mu) = H_1[j(\omega-\mu), j\mu]$$

it seems reasonable to associate the bandwidth of the variable ω in $H_2(j\omega, j\mu)$ with the bandwidth of the output of the linear time-variant network. Furthermore, in Section III we shall find that physical networks derived on the basis of such an interpretation do actually have output bandwidths restricted by the range of ω in $H_2(j\omega, j\mu)$.

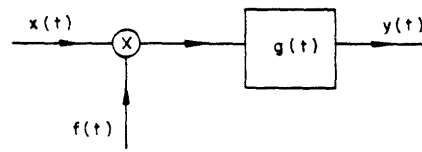
2.3 Separable Time-Variant Systems

Two forms of linear time-variant networks are worth attention, particularly because they are simple to analyze. Moreover, more complicated linear time-variant networks can frequently be built up by suitably combining these simple networks. (An illustration is given in Section III.) It has also been found that the solution to a large class of optimization problems in automatic control and communication involves such networks (20,21).



(a)

Fig. 3. Separable networks.



(b)

These networks are shown in Fig. 3. In the first network (Type I) the input $x(t)$ is passed through the linear time-invariant filter $g(t)$, and then multiplied by the function $f(t)$ to give the output $y(t)$. In the second network (Type II) the sequence is reversed.

The impulse response of the first network is given by

$$\left. \begin{array}{l} \text{or} \\ \text{or} \end{array} \right\} \begin{array}{l} h_1(t, \tau) = g(t-\tau) f(\tau) \\ h_2(z, \tau) = g(z) f(t-z) \\ h_3(y, t) = g(y) f(t) \end{array} \quad (37)$$

Similarly the impulse response of the second network is

$$\left. \begin{array}{l} \text{or} \\ \text{or} \end{array} \right\} \begin{array}{l} h_1(t, \tau) = f(\tau) g(t-\tau) \\ h_2(z, \tau) = f(\tau) g(z) \\ h_3(y, t) = f(t-y) g(y) \end{array} \quad (38)$$

Because of the form of Eqs. 37 and 38 for $h_3(y, t)$ and $h_2(z, \tau)$, these will be called separable networks.

The frequency functions for these networks also assume a simple form.

For network I:

$$\begin{aligned} H_1(j\nu, j\mu) &= G(j\nu) F(j\mu) = H_3(j\nu, j\mu) \\ H_2(j\omega, j\mu) &= G(j\omega - j\mu) F(j\mu) \end{aligned} \quad (39)$$

For network II:

$$\begin{aligned} H_1(j\nu, j\mu) &= G(j\nu - j\mu) F(j\mu) = H_3(j\nu, j\mu) \\ H_2(j\omega, j\mu) &= G(j\omega) F(j\mu) \end{aligned} \quad (40)$$

From Eq. 36 we see that network II has its output frequency range governed by $G(j\omega)$. If $G(j\omega)$ is restricted to $(-W_o, W_o)$ the output frequencies for network II, no matter what the input, will never fall outside this range.

For network I the output range is a little harder to define; but if μ is restricted to $(-W_s, W_s)$ and ν is restricted to $(-W_i, W_i)$, then, from Eq. 35, $\omega - \mu$ is restricted to $(-W_i, W_i)$ for all μ , and therefore ω is restricted to $(-W_s - W_i, W_s + W_i)$. These results will be useful in the next section.

III. SAMPLING THEOREMS FOR LINEAR TIME-VARIANT FILTERS

We have said that in general a communication channel can be regarded as a time-variant filter with constraints imposed on it. If we assume the filter to be linear it can be described conveniently by impulse and frequency response functions. Additional constraints on the filter can now be represented as constraints on these functions. The constraints presented in this section will be in the form of bandwidth or time restrictions on filter inputs and outputs. The use of sampling analysis is at once suggested, and we shall, in fact, derive appropriate sampling theorems for various sets of restrictions. Such restrictions may arise in two forms. For example, with a bandwidth constraint, it may be that the filter itself transmits only a certain range of frequencies; or we may be interested solely in the filter behavior over a particular range of frequencies. In the latter case, we may consider the actual filter to be replaced by another having the same frequency response over the specified range of interest but having zero response outside it. This situation is thus reduced to the first case. We shall, therefore, in all cases, tacitly consider only the first type of situation. That is, input or output restrictions will be suitably reflected in the impulse and/or frequency response of the linear time-variant network and we shall derive sampling results for such modified networks. These results will then hold either for arbitrary filters under the specified constraints on the input and output signals or for constrained filters under arbitrary operating conditions. Different types and sets of restrictions can be studied, but we shall consider only the following, which we think most significant. (Other cases may be studied by methods similar to those used for these cases.)

Case I: Restriction on input frequencies of signal (or filter)

Case II: Restriction on output frequencies of signal (or filter)

Case III: Restriction on filter memory, with potential limitation on range of
(a) input frequencies and (b) output frequencies.

From the discussion of the last section we find that in each case there is a most convenient form of the impulse response to use in deriving the sampling theorems. Having used this form for the derivation, we can obtain the theorems for the other forms by use of the transformations given in section 2.15. In Case I and Case II we shall consider two different situations: in one, the frequency range of interest is a lowpass region; in the other, it is a bandpass region. In none of the cases is any restriction on the filter variation necessary. However, it is often useful to consider situations in which the filter variation bandwidth is limited, to W_s , say. Therefore we shall develop theorems for both μ (the filter variation frequency variable) restricted and μ unrestricted.

3.1 Sampling Theorems for Linear Time-Variant Filters

The method of deriving sampling theorems differs according to whether the region of interest is a lowpass region or a bandpass region. In both cases, however, it is convenient to use Woodward's compact notation and method of sampling analysis (22). This method can be regarded as a translation into compact analytical form of the point of view

that regards sampling as being obtained by impulse modulation (23). With this fact in mind a physical interpretation of the steps in the following derivations is more readily seen.

We shall need a pair of definitions for studying Fourier transforms of periodic functions (22):

$$\text{rep}_T h(t) = \sum_n h(t-nT) \quad (41)$$

$$\text{comb}_T h(t) = \sum_n h(nT) \delta(t-nT) \quad (42)$$

By using a Fourier-series expansion of $\text{rep}_T h(t)$, we can derive the relation

$$\mathcal{F}\{\text{rep}_T h(t)\} = \frac{1}{T} \text{comb}_{1/T} H(jf) \quad (43)$$

and this is done in Appendix II. That is to say, if a nonperiodic function $h(t)$, which has a transform $H(jf)$, is shifted in time by all integral multiples of T , and the results are added together, the spectrum of the resulting periodic function will be obtained by picking out the values of $H(jf)$ at intervals $1/T$. Conversely,

$$\mathcal{F}\{\text{comb}_T h(t)\} = \frac{1}{T} \text{rep}_{1/T} H(jf) \quad (44)$$

Another useful pair of transforms consists of the rectangular function and its spectrum. Woodward uses the convenient notation, which we shall adopt,

$$\text{rect } t = \begin{cases} 1 & |t| < 1/2 \\ 0 & |t| > 1/2 \end{cases} \quad (45)$$

for the pulse, and

$$\text{sinc } f = \frac{\sin \pi f}{\pi f} \quad (46)$$

for its spectrum.

We now proceed to the derivation of the sampling theorems for the cases listed in section 3.1.

3.11 Sampling Theorems for Case I

In this case the frequency range of the input signals is restricted. Since we are concerned with input frequencies, the appropriate variable to consider is ν , and we might consider it in $H_1(j\nu, t)$, or in $H_3(j\nu, t)$, to be restricted to a lowpass region $(-W_i, W_i)$ or a bandpass region $(\omega_c - \frac{W_i}{2}, \omega_c + \frac{W_i}{2})$. It is rather simpler to use $H_3(j\nu, t)$ because of its direct Fourier transform relationship to $h_3(y, t)$ — a fact that is not true of $H_1(j\nu, t)$ and $h_1(t, \tau)$. We shall consider first the lowpass case.

a. In the lowpass case, ν is restricted to $(-W_i, W_i)$ and μ is either restricted to

$(-W_s, W_s)$, or is unrestricted. Then we can write (ref. 22)

$$H_3(j\nu, t) = \text{rep}_{2W_i} H_3(j\nu, t) \text{ rect } \frac{\nu}{2W_i}$$

Transforming both sides gives

$$h_3(y, t) = \text{comb}_{1/2W_i} h_3(y, t) * \text{sinc } 2W_i y$$

in which the asterisk denotes convolution. Therefore

$$\begin{aligned} h_3(y, t) &= \int \sum_n h_3(s, t) \delta\left(s - \frac{n}{2W_i}\right) \text{sinc } 2W_i(y-s) ds \\ &= \sum_n h_3\left(\frac{n}{2W_i}, t\right) \text{sinc } 2W_i\left(y - \frac{n}{2W_i}\right) \end{aligned} \quad (47)$$

This is the sampling theorem for μ unrestricted.

Next, for the variable μ we can write

$$H_3(y, j\mu) = \text{rep}_{2W_s} H(y, j\mu) \text{ rect } \frac{\mu}{2W_s}$$

and, as before, we obtain

$$h_3(y, t) = \sum_m h_3\left(y, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \quad (48)$$

Substituting for $h_3\left(y, \frac{m}{2W_s}\right)$ from Eq. 7, we then obtain the desired sampling representation:

$$h_3(y, t) = \sum_n \sum_m h_3\left(\frac{n}{2W_i}, \frac{m}{2W_s}\right) \text{sinc } 2W_i\left(y - \frac{n}{2W_i}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \quad (49)$$

If use is made of the relations between the different forms of the impulse response, this can also be written

$$h_1(t, \tau) = \sum_n \sum_m h_1\left(\frac{m}{2W_s}, \frac{m}{2W_s} - \frac{n}{2W_i}\right) \text{sinc } 2W_i\left(t - \tau - \frac{n}{2W_i}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \quad (50)$$

We next consider the bandpass case.

b. In the bandpass case, ν is restricted to the region $\left(\omega_c - \frac{W_i}{2}, \omega_c + \frac{W_i}{2}\right)$ and, as before, μ is either unrestricted or it lies in $(-W_s, W_s)$.

Here, too, we can use Woodward's (22) method of deriving bandpass sampling theorems. We go through a two-step procedure similar to the preceding one, working, however, on a complex impulse response

$$u_3(y, t) = h_3(y, t) + j\hat{h}_3(y, t)$$

where

$$\hat{h}_3(y, t) = \int \frac{h_3(r, t)}{y - r} dr \quad (51)$$

is the Hilbert transform of $h_3(y, t)$.

The final result is

$$\begin{aligned} h_3(y, t) = & \left[\sum_n \sum_m h_3\left(\frac{n}{W_i}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \text{sinc } W_i\left(y - \frac{n}{W_i}\right) \cos \omega_c\left(y - \frac{n}{W_i}\right) \right] \\ & - \left[\sum_n \sum_m \hat{h}_3\left(\frac{n}{W_i}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \text{sinc } W_i\left(y - \frac{n}{W_i}\right) \sin \omega_c\left(y - \frac{n}{W_i}\right) \right] \end{aligned} \quad (52)$$

When μ is unrestricted, this reduces to

$$\begin{aligned} h_3(y, t) = & \left[\sum_n \sum_m h_3\left(\frac{n}{W_i}, t\right) \text{sinc } W_i\left(y - \frac{n}{W_i}\right) \cos \omega_c\left(y - \frac{n}{W_i}\right) \right] \\ & - \left[\sum_n \sum_m \hat{h}_3\left(\frac{n}{W_i}, t\right) \text{sinc } W_i\left(y - \frac{n}{W_i}\right) \sin \omega_c\left(y - \frac{n}{W_i}\right) \right] \end{aligned} \quad (53)$$

Next we consider the case in which frequency restrictions are imposed on the output waveform.

3.12 Sampling Theorems for Case II

Here, recalling the discussions in Section II, the appropriate variable to use is ω , with the associated functions $h_2(z, \tau)$ and $H_2(j\omega, j\mu)$. As before, we consider a lowpass and a bandpass case.

a. In the lowpass case, the output frequency variable is restricted to $(-W_o, W_o)$, and μ may either be restricted to $(-W_s, W_s)$ or may be unrestricted. Using a procedure similar to the procedure in Case I, we have

$$H_2(j\omega, \tau) = \text{rep}_{2W_o} H_2(j\omega, \tau) \text{rect } \frac{\omega}{2W_o}$$

and transforming this, we have

$$h_2(z, \tau) = \sum_n h_2\left(\frac{n}{2W_o}, \tau\right) \text{sinc } 2W_o\left(z - \frac{n}{2W_o}\right) \quad (54)$$

which is the theorem for unrestricted μ . Similarly,

$$H_2(z, j\mu) = \text{rep}_{2W_s} H_2(z, j\mu) \text{rect } \frac{\mu}{2W_s}$$

and

$$h_2(z, \tau) = \sum_m h_2\left(z, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \quad (55)$$

Substituting in Eq. 55 from Eq. 54 yields finally

$$h_2(z, \tau) = \sum_n \sum_m h_2\left(\frac{n}{2W_o}, \frac{m}{2W_s}\right) \text{sinc } 2W_o\left(z - \frac{n}{2W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \quad (56)$$

or

$$h_1(t, \tau) = \sum_n \sum_m h_1\left(\frac{n}{2W_o} + \frac{m}{2W_s}, \frac{m}{2W_s}\right) \text{sinc } 2W_o\left(t - \tau - \frac{n}{2W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \quad (57)$$

b. In the bandpass case, ω is restricted to $\left(-\frac{W_o}{2} + \omega_c, \frac{W_o}{2} + \omega_c\right)$. The proper sampling theorems are

$$\begin{aligned} h_2(z, \tau) = & \left[\sum_n \sum_m h_2\left(\frac{n}{W_o}, \frac{m}{2W_s}\right) \text{sinc } W_o\left(z - \frac{n}{W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \cos \omega_c\left(z - \frac{n}{W_o}\right) \right] \\ & - \left[\sum_n \sum_m \hat{h}_2\left(\frac{n}{W_o}, \frac{m}{2W_s}\right) \text{sinc } W_o\left(z - \frac{n}{W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \sin \omega_c\left(z - \frac{n}{W_o}\right) \right] \end{aligned} \quad (58)$$

or

$$\begin{aligned} h_1(t, \tau) = & \left[\sum_n \sum_m h_1\left(\frac{n}{W_o} + \frac{m}{2W_s}, \frac{m}{2W_s}\right) \text{sinc } W_o\left(t - \tau - \frac{n}{W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \cos \omega_c\left(t - \tau - \frac{n}{W_o}\right) \right] \\ & - \left[\sum_n \sum_m \hat{h}_1\left(\frac{n}{W_o} + \frac{m}{2W_s}, \frac{m}{2W_s}\right) \text{sinc } W_o\left(t - \tau - \frac{n}{W_o}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \sin \omega_c\left(t - \tau - \frac{n}{W_o}\right) \right] \end{aligned} \quad (59)$$

Here, $\hat{h}_2$ and $\hat{h}_1$ are the Hilbert transforms of h_2 and h_1 with respect to z and τ , respectively. The theorems for unrestricted μ can be got from Eqs. 58 and 59 by changing $\frac{m}{2W_s}$ to τ , dropping the $\text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right)$ terms, and omitting the summations on m .

The next case involves restrictions on the memory of the filter.

3.13 Sampling Theorems for Case III

When the channel memory is limited in time, the appropriate variable to consider is y . Recall that $h_3(y, t)$ acts as a weighting function on past values of the input. Therefore if the weighting is zero after a certain range of y , say 0 to Y , for any t , channel inputs of age greater than Y seconds are "forgotten"; in other words the channel has a memory of only Y seconds. We can similarly show that, if $h_2(z, \tau) = 0$ for $z > Z$, the memory of the filter is Z seconds. However as we explained

in Cases I and II, the variables y and z are most appropriate when considering input and output frequencies, respectively. Therefore, in Case IIIa where we might wish to limit the input frequency range, we shall use the variable y and define $h_3(y, t) = 0$ for $y > Y$.

Since y is thus limited, we may apply the frequency sampling theorem (given in Appendix II), obtaining

$$\begin{aligned} h_3(y, t) &= \sum_{\ell} H_3\left(\frac{j\ell}{Y}, t\right) \exp\left(\frac{2\pi j\ell y}{Y}\right) \quad \text{for } y \text{ in } (0, Y) \\ &= 0 \quad \text{elsewhere} \end{aligned} \quad (60)$$

If, in addition, μ is restricted to $(-W_s, W_s)$ we have

$$H_3(t, j\mu) = \text{rep}_{2W_s} H_3(y, j\mu) \text{ rect } \frac{\mu}{2W_s}$$

and transforming, we obtain

$$h_3(y, t) = \sum_m h_3\left(y, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \quad (61)$$

and

$$H_3(j\nu, t) = \sum_m H_3\left(j\nu, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \quad (62)$$

Substituting into Eq. 60 we get,

$$\begin{aligned} h_3(y, t) &= \sum_{\ell} \sum_m H_3\left(\frac{j\ell}{Y}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \exp\left(\frac{2\pi j\ell y}{Y}\right) \quad \text{for } 0 \leq y \leq Y \\ &= 0 \quad \text{elsewhere} \end{aligned} \quad (63)$$

In terms of $h_1(t, \tau)$ this is

$$\begin{aligned} h_1(t, \tau) &= \sum_{\ell} \sum_m H_1\left(\frac{j\ell}{Y}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \exp\left(\frac{2\pi j\ell(t-\tau)}{Y}\right) \quad \text{for } 0 \leq y \leq Y \\ &= 0 \quad \text{elsewhere} \end{aligned} \quad (64)$$

Sampling Theorems for Case IIIb. The variable that is appropriate to considerations of output bandwidth is z . If the response $h_2(z, \tau)$ is to be zero after Z seconds, say, then z ranges from $(0, Z)$. We may note that this is effectively a restriction on the duration of the output waveform, because it requires that the output never last for more than Z seconds beyond the input.

Using, as in Case III, the frequency sampling theorems of Appendix II, we can write

$$h_2(z, \tau) = \sum_{\ell} H_2\left(\frac{j\ell}{Z}, \tau\right) \exp\left(\frac{2\pi j\ell z}{Z}\right) \quad \text{for } 0 \leq z \leq Z$$

$$= 0 \quad \text{elsewhere} \quad (65)$$

If, in addition, we assume that μ lies within $(-W_s, W_s)$ we have, finally, as in Case III,

$$h_2(z, \tau) = \sum_{\ell} \sum_m H_2\left(\frac{j\ell}{Z}, \frac{m}{2W_s}\right) \exp\left(\frac{2\pi j\ell z}{Z}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \quad (66)$$

or

$$h_1(t, \tau) = \sum_{\ell} \sum_m H_1\left(\frac{m}{2W_s}, \frac{m}{2W_s} - \frac{j\ell}{Z}\right) \exp\left(\frac{2\pi j\ell(t-\tau)}{Z}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \quad (67)$$

The methods employed in these three cases should suffice for the analysis of other types and sets of restrictions.

3.2 Discrete Models for Linear Time-Variant Filters

We can use the sampling theorems derived in the previous sections to construct discrete models for suitably restricted linear time-variant filters or, equivalently, for linear time-variant networks to be used under restricted conditions. Before proceeding to do this, however, we note that the summation over n in the various theorems need run only from 0 to ∞ , provided the (modified) networks are made approximately realizable by means of appropriate delays, because then the impulse response is zero for z or y less than zero. We say approximately because theoretically an infinite delay would be needed; in practice, however, finite delays would often suffice. We shall therefore use the range $(0, \infty)$ for n because it makes the models easier to comprehend; however, for theoretical analyses such a restriction is unnecessary. However, m and ℓ range over all the integers $(-\infty, \infty)$.

In each case the models are obtained by rearranging the appropriate sampling theorems as sums of simple separable networks of the type considered in section 2.3. Linearity is used to rearrange the resulting networks in more convenient form.

It is important to note that the models given are not the only ones that can be derived from the sampling theorems; it is often possible to recast these theorems in other forms that lead to different physical models. The only requirement is that all of these models have the required number of degrees of freedom specified by the appropriate sampling theorems.

Case I. Limited Input Frequency Range

a. Lowpass case

From Eq. 49 we can write

$$h_3(y, t) = \sum_n f_n(t) g\left(y - \frac{n}{2W_i}\right) \quad (68)$$

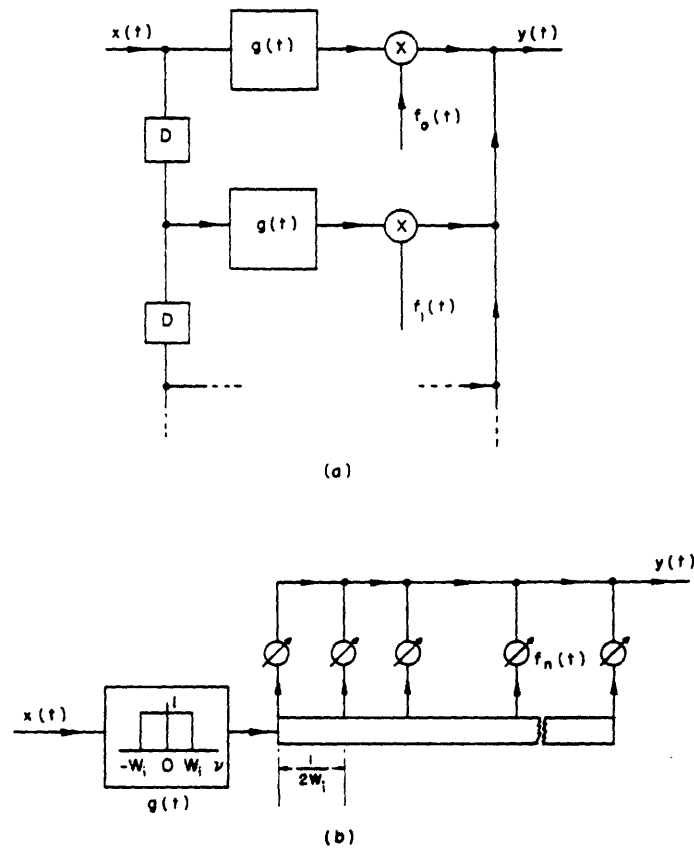


Fig. 4. A model for Case I. Lowpass restriction on input frequencies.

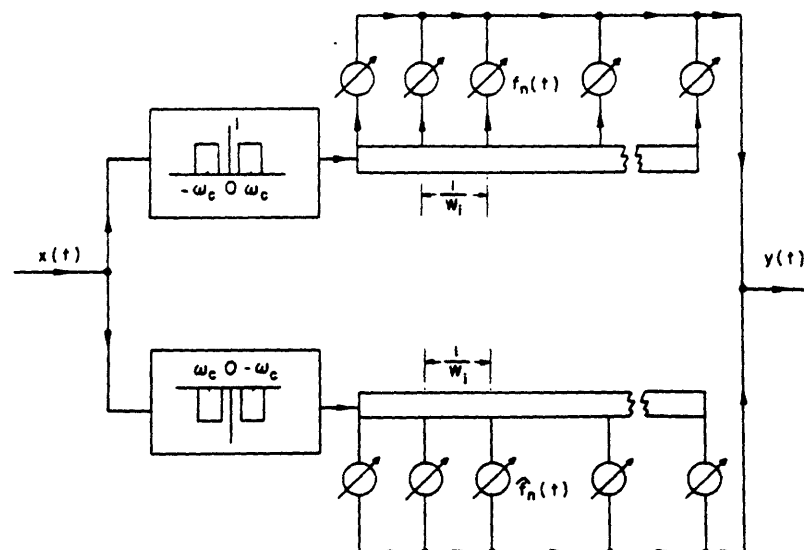


Fig. 5. A model for Case I. Bandpass restriction on input frequencies.

where

$$f_n(t) = \frac{1}{2W_i} \sum_m h_3\left(\frac{n}{2W_i}, \frac{m}{2W_s}\right) \text{sinc } 2W_s \left(t - \frac{m}{2W_s}\right)$$

and

$$g(y) = 2W_i \text{sinc } 2W_i y$$

If we now refer to section 2.3 and Fig. 3a, we see that we can synthesize $h_3(y, t)$ as shown in Fig. 4a where the boxes marked D provide pure delays in time of $1/2W_i$ seconds each. Furthermore, $g(t)$ is recognized as a filter with a flat passband from $-W_i$ to W_i . By virtue of linearity, we can transfer the $g(t)$ across the delay boxes, combine the boxes into a delay line, and obtain the model shown in Fig. 4b. The rectangular filter is assumed to have zero phase shift.

Theoretically, the delay line should have infinite length because $h_3(y, t)$ has finite bandwidth; this is indicated by the broken lines in Fig. 4b. However, since the set of sinc functions used in the sampling theorem (which may also be regarded as a series expansion) form a complete set, we see that a finite length can be used, at the cost of an error that can be made arbitrarily small by prolonging the line sufficiently. Finally, referring to the last paragraph in section 2.3, we note that if the input to our filter has bandwidth $2W_i$ and the system bandwidth is $2W_s$, then the output bandwidth with our model is not greater than $(2W_i + 2W_s)$.

b. In the bandpass case, a similar procedure leads to the model shown in Fig. 5. Here the top filter has zero phase shift; the lower filter has phase shift $\pi/2$ in the negative frequency band and $-\pi/2$ in the positive frequency band. The impulse response of the top filter is $2W_i \text{sinc } W_i t \cos \omega_c t$, and that of the lower one is $-2W_i \text{sinc } W_i t \sin \omega_c t$.

In the bandpass case we note that the sampling theorem, Eq. 48, can be rewritten in many different forms (e.g., amplitude-phase sampling theorem), and different models can be derived. Since the basic procedure is the same as that just described, we shall not consider all of these different cases.

c. If μ is unrestricted, we have

$$f_n(t) = \frac{1}{2W_i} h_3\left(\frac{n}{2W_i}, t\right) \quad \text{in Fig. 4}$$

and

$$\left. \begin{aligned} f_n(t) &= \frac{1}{2W_i} h_3\left(\frac{n}{W_i}, t\right) \\ \hat{f}_n(t) &= \frac{1}{2W_i} \hat{h}_3\left(\frac{n}{W_i}, t\right) \end{aligned} \right\} \quad \text{in Fig. 5}$$

The next case involves restrictions on the output signal.

Case II. Limited Output Frequency Range

The sampling theorems are given in section 2.21 and these are rearranged to give delay line models for this case.

a. Lowpass case

Here it is convenient to rewrite Eq. 56 as

$$h_2(z, \tau) = \sum_n f_n \left(\tau + \frac{n}{2W_0} \right) g \left(z - \frac{n}{2W_0} \right) \quad (69)$$

where

$$f_n(t) = \frac{1}{2W_0} \sum_m h_2 \left(\frac{n}{2W_0}, \frac{m}{2W_s} \right) \text{sinc } 2W_s \left(t - \frac{m}{2W_s} \right)$$

and

$$g(z) = 2W_0 \text{sinc } 2W_0 z$$

Referring to section 2.3 and Fig. 3b we note that $h_2(z, \tau)$ can be synthesized as shown in Fig. 6a. By steps similar to those used in Case I, this can be reduced to Fig. 6b.

b. The bandpass case

The model for the bandpass case is derived in similar fashion and is shown in Fig. 7. We can make the same comments on line length, filter phase shifts, frequency relations (i.e., output has bandwidth no greater than $2W_0$ no matter what the input), and different bandpass models, that we made in Case I.

c. If μ is unrestricted, we have

$$f_n(t) = \frac{1}{2W_0} h_2 \left(\frac{n}{2W_0}, t \right) \quad \text{in Fig. 6}$$

and

$$\left. \begin{aligned} f_n(t) &= \frac{1}{2W_0} h_2 \left(\frac{n}{W_0}, t \right) \\ \hat{f}_n(t) &= \frac{1}{2W_0} \hat{h}_2 \left(\frac{n}{W_0}, t \right) \end{aligned} \right\} \quad \text{in Fig. 7}$$

The next two cases do not involve delay lines, but are formed from banks of finite memory filters.

Case IIIa. Limited Memory of Channel Filter, with Potential Limitation on Input Frequency Range

In this case we can rearrange Eq. 63 in the form

$$h_3(y, t) = \sum_l f_l(t) g_l(y) \quad (70)$$

where

$$f_l(t) = \sum_m H_3 \left(\frac{j\ell}{Y}, \frac{m}{2W_s} \right) \text{sinc } 2W_s \left(t - \frac{m}{2W_s} \right)$$

and

$$g_l(y) = \exp \left(\frac{2\pi j\ell y}{Y} \right) \text{rect} \left(\frac{y}{Y} - \frac{1}{2} \right)$$

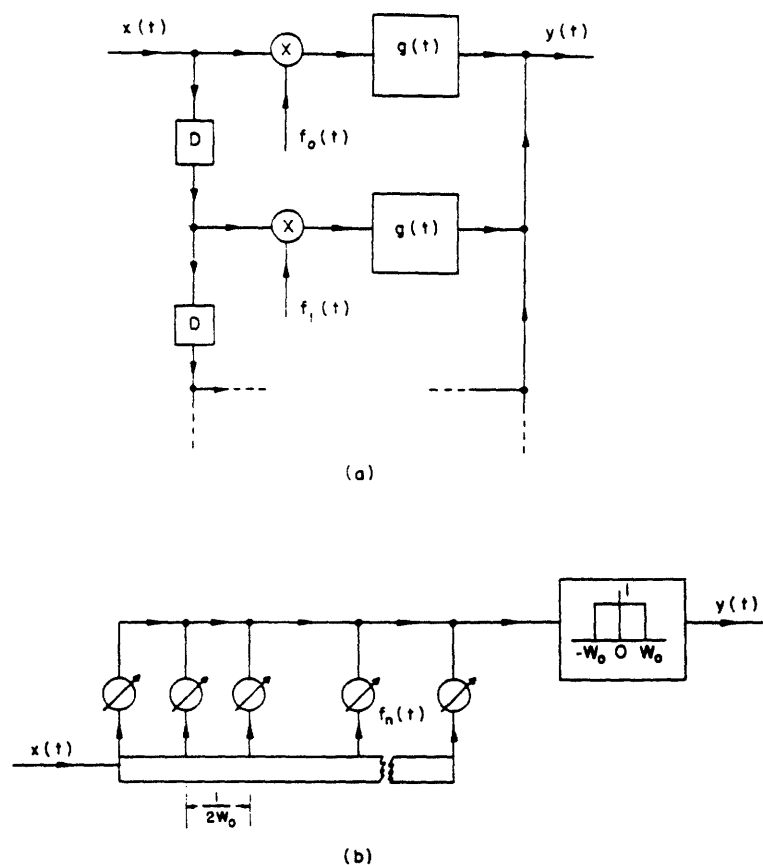


Fig. 6. A model for Case II. Lowpass restriction on output frequencies.

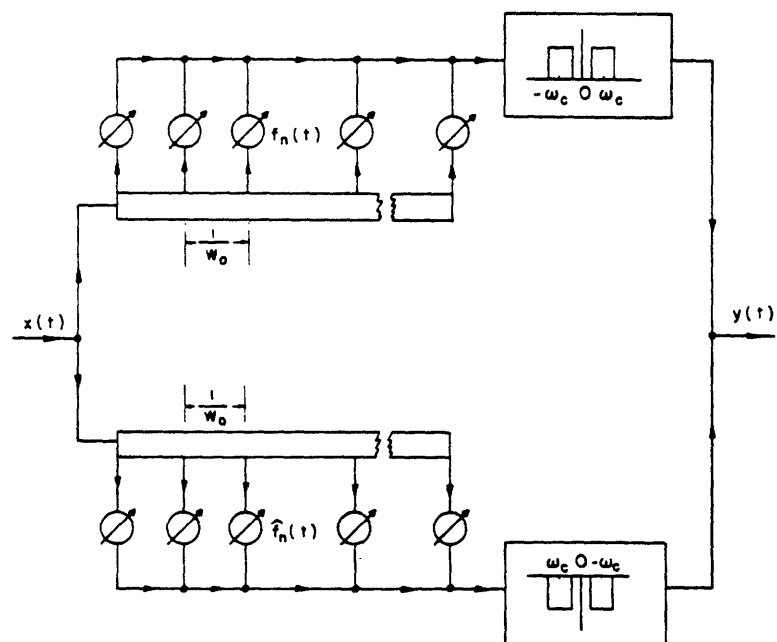


Fig. 7. A model for Case II. Bandpass restriction on output frequencies.

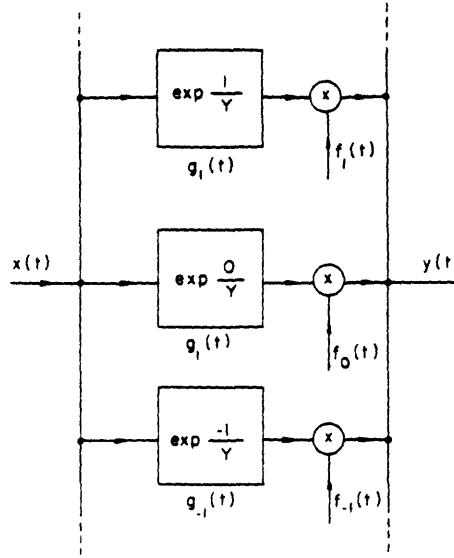


Fig. 8. A model for Case IIIa. Limited filter memory with potential restriction on input frequency range. The notation $\exp(l/Y)$ denotes a filter with impulse response $\exp[(2\pi jly)/Y] \cdot \text{rect}[(y/Y) - (1/2)]$.

Recognizing that the $g_l(y)$ represent envelope-integrating filters (with a finite integration time Y) at the frequencies l/Y and recalling the Type I network of section 2.3, we can construct a model for this case, as shown in Fig. 8. Another interpretation for the operation of the filters $g_l(y)$ is that they represent filters that continuously extract a signal that would be the l^{th} Fourier component of a periodic waveform, each period of which duplicates the last Y seconds of the input to the filters. (We might also note that such filters are used in the Kineplex system of communications (24), where they are operated as "integrate-and-quench" filters.) Theoretically we should have an infinite bank of filters, but if we impose a bandwidth restriction on the frequency range of the input signals, clearly a finite number of filters will be sufficient.

Case IIIb. Limited Memory of Filter, with Potential Limitation on Output Frequency Range

The sampling theorem, Eq. 65 can be written

$$h_2(z, \tau) = \sum_l f_l(\tau) g_l(z) \quad (71)$$

where

$$f_l(t) = \sum_m H_2\left(\frac{j\ell}{Z}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right)$$

$$g_l(z) = \exp\left(\frac{2\pi j\ell z}{Z}\right) \text{rect}\left(\frac{z}{Z} - \frac{1}{2}\right)$$

The $g_l(z)$ represent, as before, integrating filters at the frequencies l/Z , and we now

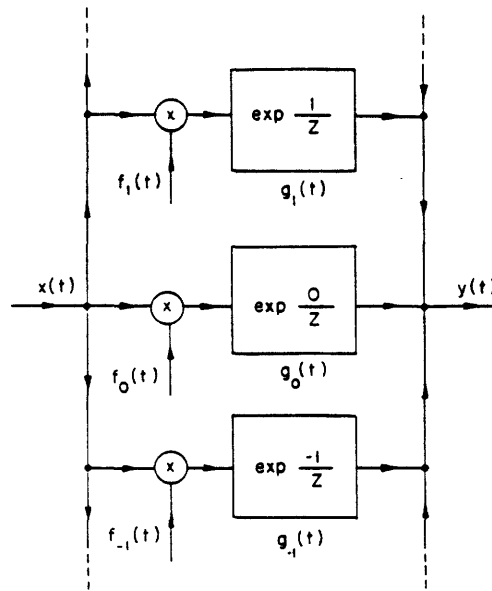


Fig. 9. A model for Case IIIb. Limited filter memory with potential restriction on output frequency range.

use networks of Type II (section 2.3) to get the model for this case, which is shown in Fig. 9. Again, theoretically we should have an infinite bank of such filters, unless we impose a bandwidth restriction on the output frequency range of the signals when a finite number will suffice.

Finally, we may point out that in both Case IIIa and Case IIIb we can combine the terms $f_{-l} g_{-l}$ and $f_{+l} g_{+l}$ to get a representation in terms of filter banks together with amplitude and phase controls.

IV. THE MEASUREMENT PROBLEM IN THE ABSENCE OF NOISE

In this section we discuss some aspects of the problem of specifying a linear time-variant filter by means of input-output measurements. That is, the linear time-variant network is regarded as a black box; the only means of getting information about it is by putting in appropriate signals and observing the corresponding output signals. In the communication situation there are various constraints on the types of signals we can use; for example, average or peak power limitations, finite bandwidth restrictions, and so forth. Furthermore, our measurements usually will be corrupted by noise. In this report we shall consider only the case of bandwidth and duration restrictions on our input and output signal waveforms and shall assume that no noise is present.

The measurement problem in the absence of noise is, of course, much simpler than that when noise is present, but the problem is still not entirely trivial. Any limitations on our ability to determine the linear time-variant network in the noiseless case will usually carry over into the noisy measurement problem. We say "usually" because in the noisy case the determination can evidently be statistical only, and the presence of noise might require us to relax some of the restrictions present in the noiseless case. Another area in which a study of the noiseless case may be useful is that of sampled-data systems, and we shall say more about this later. For simplicity, all of the analysis will be carried out for lowpass frequency regions. The extension to bandpass regions is straightforward. The constraints we shall consider in the measurement problem will be the frequency and time constraints of Section III for which we have already derived sampling theorems and models. The case considered in most detail is one in which the frequency range of the output signal is limited to, say, $(-W_0, W_0)$, which is Case II of Section III. Since other cases that can be treated similarly are discussed in Appendix III, we shall in this section speak only of Case II. We shall first deduce a condition on our linear time-variant network for a measurement of it to be possible. Then, having defined a model for the problem, we derive a matrix specification of it. (It happens that such a matrix also occurs in the study of time-variant sampled-data systems.)

This matrix is subsequently used to prove the sufficiency of the measurement condition mentioned earlier. We find that the use of the delay line model developed in Section III gives a simple physical picture of the various mathematical derivations and the results obtained.

4.1 A Necessary and Sufficient Condition for the Measurement

The fact that our available output waveform is bandlimited restricts the number of linearly independent measurements we can make per second. Because of this we would expect the rate of variation of our system to be a controlling factor in defining which systems can be determined by input-output measurements. We shall find that this is so. But first we need to formulate our problem more precisely. (Note that because of

the output restriction we are actually measuring the linear time-variant network as seen through our output-limiting filter.)

We wish, given a linear time-variant network N , to construct a model that will

produce, for inputs identical to possible inputs of N (over a given time, 0 to T seconds, say), the same outputs as N , over an output frequency range $(-W_o, W_o)$.

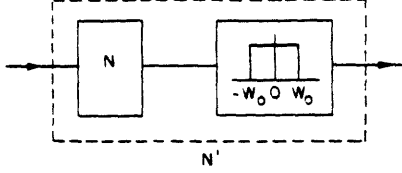


Fig. 10. The filters N and N' .

Our first step is to replace the network N with the network N' as shown in Fig. 10. N' is equivalent to N over the specified output frequency range. Also let the system bandwidth of N be $2W_s$; that is, μ lies in the range $(-W_s, W_s)$. We

shall need to assume that, for N' , the effect of an excitation dies down after some finite time, say Z . That is, if $h_2(z, \tau)$ is the impulse response of N' .

$$h_2(z, \tau) = 0 \quad z > Z$$

Mathematically, of course, since N' is bandlimited, this can never be true. However, in practice, N' would not be absolutely bandlimited, and the response would die out after some finite time. Finally, we shall assume sufficient delay to make N' realizable, as was done in section 3.2.

We can now state that to reproduce the operation of a linear time-variant network under the above conditions, with an error that can be made arbitrarily small, it is necessary and sufficient that

$$Z \leq \frac{1}{2W_s} - \frac{1}{2W_o} \quad (72)$$

or equivalently

$$W_s \leq \frac{W_o}{1 + 2ZW_o}$$

This statement is proved in two parts.

1. Necessity of the Condition

We have to specify $h_2(z, \tau)$ over a z -span Z , and a τ -span T . Over these ranges, $h_2(z, \tau)$ has $(2W_o Z + 1) \times (2W_s T + 1)$ degrees of freedom (or linearly independent values). If τ is restricted to $(0, T)$, the duration of the output waveform according to our assumptions cannot be greater than $(T + Z)$ seconds. Now in a time $(T + Z)$, since our output is bandlimited, we can obtain only

$$\frac{T + Z}{\frac{1}{2W_o}} = 2(T + Z) W_o + 1$$

independent measurements or values. Therefore we must have

$$(2W_o Z + 1)(2TW_s + 1) \leq 2(T + Z)W_o + 1$$

that is,

$$4TW_s W_o Z + 2TW_s + 2ZW_o \leq 2TW_o + 2ZW_o$$

or

$$Z \leq \frac{1}{2W_s} - \frac{1}{2W_o}$$

Therefore Eq. 72 is a necessary condition.

2. Sufficiency of the Condition

This will be shown later, in section 4.4. We shall prove sufficiency by describing an actual measurement procedure. Before doing this, however, it is necessary to investigate the relations between input and output functions in our problem.

4.2 Input-Output Relations in the Problem

The impulse response $h_2(z, \tau)$ for N' is bandlimited in both frequency variables, ω and μ . We can apply the sampling theorems of section 3.12 to get

$$h_2(z, \tau) = \sum_n \sum_m h_2\left(\frac{n}{2W_o}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(\tau - \frac{m}{2W_s}\right) \text{sinc } 2W_o\left(z - \frac{n}{2W_o}\right) \quad (73)$$

$$= \sum_n f_n\left(\tau + \frac{n}{2W_o}\right) \text{sinc } 2W_o\left(z - \frac{n}{2W_o}\right) \quad (74)$$

For convenience we shall normalize $2W_o$ to 1, so that we have

$$h_2(z, \tau) = \sum_n f_n(\tau + n) \text{sinc } (z - n)$$

or

$$h_1(t, \tau) = \sum_n f_n(\tau + n) \text{sinc } (t - \tau - n) \quad (75)$$

$$= \sum_n f'_n(\tau) \text{sinc } (t - \tau - n) \quad (76)$$

where

$$f'_n(t - n) = f_n(t)$$

Now if $x(t)$ is the input to N' and $y(t)$ the output, we have

$$\begin{aligned}
y(t) &= \int h_1(t, \tau) x(\tau) d\tau \\
&= \int \sum_n f'_n(\tau) \text{sinc}(t-\tau-n) x(\tau) d\tau
\end{aligned} \tag{77}$$

Now define

$$f'_n(t) x(t) = u_n(t)$$

with the result that

$$y(t) = \sum_n \int u_n(\tau) \text{sinc}(t-\tau-n) d\tau \tag{78}$$

This is valid provided the bandwidth of $u_n(t)$ is not greater than $2W_o$. The definition of $u_n(t)$ implies that $x(t)$ should have a bandwidth not greater than $2W_o - 2W_s$. This restriction is not unreasonable since any higher frequencies in $x(t)$ will fall outside the band $(-W_o, W_o)$ and therefore would not be considered.

Let us consider $u(t)$ and $\text{sinc } t$ as discrete time series $u(K)$, $\text{sinc}(K)$, given by their values at $K = 0, \pm 1, \pm 2, \dots$. This will give us $y(t+n)$ for all integral arguments. We have then

$$y(t) = \sum_n \sum_r u_n(r) \text{sinc}(t-n-r)$$

But because

$$\text{sinc } t = 0 \quad \text{for } t = \pm 1, \pm 2, \dots$$

$$\text{sinc } t = 1 \quad \text{for } t = 0 \tag{79}$$

this reduces (for integral values of t) to

$$y(t) = \sum_n u_n(t-n) \tag{80}$$

$$= \sum_n f'_n(t-n) x(t-n)$$

$$= \sum_n f_n(t) x(t-n) \tag{81}$$

If we pick our time origin so that $x(t) = 0, t < 0$, we can write Eq. 81 as a matrix equation

$$\underline{y} = \underline{H} \underline{x} \quad (82)$$

where the underlined letters indicate matrices. In the expanded form, Eq. 82 becomes

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ . \\ . \\ . \\ . \end{bmatrix} = \begin{bmatrix} f_0(0) & 0 & 0 & 0 & 0 & \dots & 0 \\ f_1(1) & f_0(1) & 0 & 0 & 0 & \dots & 0 \\ f_2(2) & f_1(2) & f_0(2) & 0 & 0 & \dots & 0 \\ f_3(3) & f_2(3) & f_1(3) & f_0(3) & 0 & \dots & 0 \\ . & . & . & . & . & \dots & . \\ . & . & . & . & . & \dots & . \\ . & . & . & . & . & \dots & . \\ . & . & . & . & . & \dots & . \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ . \\ . \\ . \\ . \end{bmatrix} \quad (83)$$

If the relation $f_n(t) = \sum_m h_2\left(n, \frac{m}{2W_0}\right) \text{sinc } 2W_0\left(t - n - \frac{m}{2W_0}\right)$ is used, Eq. 83 can be written

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ . \\ . \\ . \\ y(n) \\ . \\ . \end{bmatrix} = \begin{bmatrix} h_2(0,0) & 0 & 0 & 0 & . \\ h_2(1,0) & h_2(0,1) & 0 & 0 & . \\ h_2(2,0) & h_2(1,1) & h_2(0,2) & 0 & . \\ h_2(3,0) & h_2(2,1) & h_2(1,2) & h_2(0,3) & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ h_2(n,0) & h_2(n-1,1) & h_2(n-2,2) & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ . \\ . \\ . \\ x(n) \\ . \\ . \end{bmatrix} \quad (84)$$

Since $y(t)$ is bandlimited to $2W_0(=1)$, the sequence $y(0), y(1), \dots$ suffices to determine it. Thus $\underline{H}$ is a suitable specification for bandlimited linear time-variant networks of Type II of Section III. (Similar matrices can be derived for linear time-variant networks of Type I, that is, input frequency limited networks. These are derived in Appendix III.) Notice that the vanishing of all terms above the main diagonal is a consequence of our assumption of physical realizability. As stated before, this form is convenient for gaining a better physical picture of the arguments and proofs, but is not a theoretical necessity. The length Z determines the number of possible nonzero values in a column of $\underline{H}$, viz., $[Z]$, where the square brackets denote the largest integer less than Z . Thus if $Z = 2 \cdot \frac{1}{2W} = 2$, the matrix is

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \\ . \\ . \\ . \end{bmatrix} = \begin{bmatrix} h_2(0,0) & 0 & . & . \\ h_2(1,0) & h_2(0,1) & 0 & . \\ h_2(2,0) & h_2(1,1) & h_2(0,2) & 0 \\ 0 & h_2(2,1) & h_2(1,2) & h_2(0,3) \\ . & 0 & h_2(2,2) & . \\ . & . & 0 & . \\ . & . & . & . \\ . & . & . & . \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \\ . \\ . \\ . \end{bmatrix} \quad (85)$$

4.3 The Matrix $\underline{H}$

We have derived the matrix $\underline{H}$ as being sufficient to characterize the network N' under the stated conditions of operation; that is, the bandwidth of $x(t)$ is $2W_o - 2W_s$ or less. It happens that an exactly similar matrix can be used to describe a linear time-variant sampled-data system (25). In reference 25, Friedland discusses the matrix analysis of such systems, and all his results could be used in studies on our model. For example, he shows that cascading two sampled-data systems is equivalent to multiplying their matrices, paralleling two systems is the same as adding their matrices, and so forth. Although in this report we are not particularly interested in studying the properties of interconnected linear time-variant networks, we have found the results of Friedland's work on time-varying optimization theory (Wiener-Lee theory) of use in work on the noisy measurement case. However, this is not discussed here.

4.4 The Sufficiency Proof

Returning to our problem we now prove that the condition $Z \leq \frac{1}{2W_s} - \frac{1}{2W_o}$ is sufficient for a measurement of $h_2(z, \tau)$. However, the argument is somewhat long and detailed, and it makes for clearer exposition to dispose of certain details first.

1. To establish the sufficiency proof we shall use the results of our sampling theorem. This tells us that the values $h_2\left(\frac{n}{2W_o}, \frac{m}{2W_s}\right)$ for all n and m suffice to determine the behavior of $h_2(z, \tau)$. However, we are interested only in $h_2(z, \tau)$ over the interval $(0, T)$ and, therefore, we might expect that we need only the values of $h_2\left(\frac{n}{2W_o}, \frac{m}{2W_s}\right)$ in this time interval. Unfortunately, this is not true; the behavior of $h_2(z, \tau)$ is influenced by sample values outside $(0, T)$ also. Using only the values in $(0, T)$ would introduce error. However, if T were long, the error would be negligible in the middle of the interval and would be larger at the ends. This end-effect is characteristic in all applications of the sampling theorem. One way to mitigate it is to make measurements over a longer time interval (say $-\Gamma'$ to Γ'), including $(0, T)$. The behavior over the smaller interval $(0, T)$ would then be almost right. By letting Γ' tend to infinity, the

error can be made arbitrarily small for finite T . (A discussion of the end-effect in sampling is given in reference 26.) We shall now pick our time origin so that $(-\Gamma', \Gamma')$ corresponds to $(0, \Gamma)$, and talk only about this last interval in all future discussions.

2. It is convenient to assume that Z and $\frac{1}{2W_s}$ are integral multiples of $\frac{1}{2W_0}$ and, therefore, with our normalization, Z and $\frac{1}{2W_s}$ are integral. Similarly we shall assume that Γ is an integral multiple of $\frac{1}{2W_s}$. These conditions can be arranged by suitably increasing some or all of Z , Γ , and $2W_0$.

3. Under these conditions the problem is to determine the values $h_2\left(n, \frac{m}{2W_s}\right)$ over $(0, \Gamma)$. These are

$$\begin{array}{ccccccc} h_2(0, 0) & h_2\left(0, \frac{1}{2W_s}\right) & \dots\dots\dots & h_2(0, \Gamma) & & & \\ h_2(1, 0) & h_2\left(1, \frac{1}{2W_s}\right) & \dots\dots\dots & h_2(1, \Gamma) & & & \\ . & . & & & & & \\ . & . & & & & & \\ . & . & & & & & \\ h_2(Z, 0) & h_2\left(Z, \frac{1}{2W_s}\right) & \dots\dots\dots & h_2(Z, \Gamma) & & & \end{array} \quad (86)$$

If we consider the matrix $\underline{H}$, these appear in columns of the matrix, as shown, for a linear time-variant network with $Z = 2$, $\frac{1}{2W_s} = 4$, $\Gamma = 8$.

$$\left[\begin{array}{cccccccccccccccc} a & & & & & & & & & & & & & & & & \\ b & a_1 & & & & & & & & & & & & & & & \\ c & a_2 & a_4 & & & & & & & & & & & & & & \\ & a_3 & a_5 & a_7 & & & & & & & & & & & & & \\ & & a_6 & a_8 & d & & & & & & & & & & & & \\ & & & a_9 & e & a_{10} & & & & & & & & & & & \\ & & & & f & a_{11} & a & & & & & & & & & & \\ & & & & & a_{12} & a & a & & & & & & & & & \\ & & & & & & a & a & g & & & & & & & & \\ & & & & & & & a & h & a & & & & & & & \\ & & & & & & & & j & a & a & & & & & & \end{array} \right] \quad (87)$$

The letters $a, b, \dots$ stand for the values $h_2\left(n, \frac{m}{2W_s}\right)$ from the array (86). The places marked with a 's indicate sample values that are not linearly independent, but are combinations of the $a, b, \dots$. Thus $a_1, a_4, a_7, \dots$ are linear combinations of a, d , and g ; $a_2, a_5, a_8, \dots$ of a, e, h ; $a_3, a_6, a_9, \dots$ of c, f, j . It is easily verified

that a_2 is the same combination of b, e, h that a_1 is of a, d, g; the same holds for a_3 . Thus $(a_1, a_2, a_3)(a_4, a_5, a_6) \dots$ are similar in this sense. This property will be used later on.

4. In terms of matrix 87, it is easy to prove the sufficiency of our argument for $Z \leq \frac{1}{2W_0} - \frac{1}{2W_s}$. We see that, if we apply an input $x(t)$ which is 1 at $t=0, \frac{1}{2W_s}, \frac{2}{2W_s}, \dots$ and zero for all other integral t , the output would be a, b, c, 0, d, e, f, 0, g, h, j. Therefore, in general, when $Z \leq \frac{1}{2W_0} - \frac{1}{2W_s}$, an input of the form

$$\begin{array}{ccc} \begin{array}{c} \boxed{1 \ 0 \ 0 \ 0 \ 0 \dots 0} \\ t:0 \end{array} & \begin{array}{c} \boxed{1 \ 0 \ 0 \ 0 \ 0 \dots 0} \\ \frac{1}{2W_s} \end{array} & \begin{array}{c} \boxed{1 \ 0 \ 0 \ 0 \dots} \\ \frac{2}{2W_s} \end{array} \end{array} \quad (88)$$

is sufficient to determine the sample values. By taking Γ large enough, these values specify $h_2(z, \tau)$ as closely as we wish over the desired included interval, T . Thus we have proved that the condition $Z \leq \frac{1}{2W_s} - \frac{1}{2W_0}$ is sufficient to enable a measurement.

(And, in fact, the proof we have used for this also indicates that when $Z > \frac{1}{2W_s} - \frac{1}{2W_0}$ the measurement is not possible; under such conditions there will always be more unknowns than equations.) Thus, finally, we have established the condition $Z \leq \frac{1}{2W_s} - \frac{1}{2W_0}$ as both necessary and sufficient.

5. The foregoing proofs are completely general and do not depend on any particular physical model. However, since the conditions of our problem meet the requirements of Case II, Section III, we can use the delay-line model given in Fig. 6, with length equal to $(Z-1)$ seconds, to give a physical meaning to the argument. Thus the input, Eq. 88, is equivalent to feeding impulses into the delay line at times $0, \frac{1}{2W_s}, \dots$, and the condition $Z \leq \frac{1}{2W_s} - \frac{1}{2W_0}$ is equivalent to requiring that there be only one impulse on the line at any instant. Notice also that the matrix $\underline{H}$ can be obtained rather simply from this delay line model, since if the bandwidth of $x(t)$ is not greater than $2W_0 - 2W_s$, we can dispense with the output filter.

6. It is of interest to examine the condition $Z \leq \frac{1}{2W_0} - \frac{1}{2W_s}$ in some limiting cases. Notice first that if $W_s = 0$, that is, if the filter is time-invariant, there is no bound on the duration, Z , of the impulse response. And conversely if Z is infinite, the filter can be determined only if $W_s = 0$, that is, only if it is time-invariant. Secondly, if Z is zero, that is, if our filter has no memory, or the delay line has only one tap, the requirement for a measurement is $W_s \leq W_0$. And finally, if W_0 is infinite, the condition is $Z \leq \frac{1}{2W_s}$. When W_0 is infinite, there is no output frequency constraint on the original filter, and in this case Z will be the actual maximum duration of the impulse response of the filter. Then the condition implies that in the absence of any prior information about the filter, an exact determination of it, even with no additional noise present, is impossible unless $Z \cdot 2W_s \leq 1$.

7. We have shown that when $Z > \frac{1}{2W_s} - \frac{1}{2W_o}$ we cannot determine the linear time-variant network under the assumptions we have made. This follows from the fact that we have more unknowns than equations, and thus are confronted with an indeterminate situation. If, however, there exist additional constraints on the linear time-variant network, these can be reflected in equations involving the unknown sample points, and we might now have enough equations to determine all the unknowns. Thus, for example, we might have the constraint

$$\int_0^Z h_2^2(z, T) dz = K \quad (89)$$

for all T .

This additional constraint is sufficient to determine the following linear time-variant network

$$\underline{H} = \begin{bmatrix} a & & & & & & \\ b & a & & & & & \\ c & a & d & & & & \\ & a & e & a & & & \\ & & f & a & g & & \\ & & & a & h & a & \\ & & & & i & a & a \end{bmatrix} \quad (90)$$

If $x(0) = 1 = x(2) = x(4)$ and all other $x(k)$ are zero, we can find a, b, e, h, i directly and also the sums $c + d, f + g$. Now from Eq. 89 we have the additional relations

$$\left. \begin{aligned} a^2 + b^2 + c^2 &= K \\ d^2 + e^2 + f^2 &= K \\ g^2 + h^2 + i^2 &= K \end{aligned} \right\} \quad (91)$$

Since we know a, b we can solve for c and then get d from the sum $c + d$. Knowing e and d we can calculate f from Eq. 91 and then g from the sum $f + g$. Thus the linear time-variant network is determined.

Many other types of constraints may be present in any particular problem, and these often may be sufficient to determine the linear time-variant network as before, even when the condition $Z \leq \frac{1}{2W_s} - \frac{1}{2W_o}$ is violated. Sometimes, in fact, it might be useful to assign values almost arbitrarily to particular sample points and thus to obtain an estimate (albeit degraded) of the linear time-variant network. In general, however, it would be better to obtain degraded estimates on the basis of some over-all system criterion. Since such criteria are most meaningfully established for the case when noise is present, we have not investigated them here.

4.5 Measurement Conditions for Other Constraints

The discussion of the measurement problem under other frequency and time constraints is very similar to the one we have just considered. In this section we shall only quote the results that are briefly derived in Appendix III. We give necessary and sufficient conditions for the measurement under the different restrictions.

Case I. Input Frequencies Limited to $(-W_i, W_i)$

A model for this situation is shown in Fig. 5. If we define a quantity Y analogous to Z , so that $h_3(y, t) = 0$, $y > Y$, a necessary and sufficient condition is

$$Y \leq \frac{1}{2W_s} - \frac{1}{2W_i}$$

or

$$W_s \leq \frac{W_i}{1 + 2W_i Y}$$

A sufficient signal is one that consists of unit impulses at $0, \frac{1}{2W_s}, \frac{2}{2W_s}$, and so on.

Case II. Limited Output Frequency Range

This case has been discussed.

Case III. Limited Filter Memory

Models for this case have been derived in Section III and are shown in Figs. 8 and 9. In Appendix III it is shown that the measurement condition for this case is

$$Y(\text{or } Z) \leq \frac{1}{2W_s}$$

If the input or output frequencies are restricted to, say $(-W_i, W_i)$ or $(-W_o, W_o)$, the conditions are

$$Y \leq \frac{1}{2W_s} - \frac{1}{2W_i}$$

$$Z \leq \frac{1}{2W_s} - \frac{1}{2W_o}$$

The nature of these conditions in different limiting cases can be discussed as was done for Case II (section 4.4). Under these limiting conditions we would expect the results for the various constraints to agree, and in fact this is so, as can be readily verified.

V. CONCLUDING REMARKS

The development of mathematical models for communication channels has been the subject of much research since the advent of statistical communication theory. Although channels with additive random disturbances – such as the BSC and the Gaussian channel – have been studied in some detail, less work has been done on channels with nonadditive disturbances, of which multipath and scatter channels are notable examples. In particular, major work in this area has been done by R. Price (6, 7) and G. Turin (9). While their studies have yielded considerable insight into the problem and have aided the development of a successful system to combat multipath (5), the models they have used are not quite general, since several assumptions about the path structure of the channel are made. Price considers only statistically independent paths with Rayleigh-distributed strengths and known delays. In Turin's model knowledge of path delays is not required; moreover, he considers more general path statistics, but he is forced to assume that the paths are resolvable and time-invariant.

The models proposed in this report are not, as they stand, models for multipath channels, chiefly because no statistical information has been taken into account in their formulation. The determination of appropriate statistical distributions for the time-variant tap gains in our models is an interesting topic for future investigation. However we feel that a significant feature of our models is the operational, or phenomenological, point of view adopted in their derivation: Our delay-line and filter-bank models have been based on assumptions concerning the limitations of our signal-generating and measuring equipment. Thus consider, for example, the delay-line model for the situation in which the output-signal frequency range is limited: The actual channel structure may have either a discrete or a continuous structure, or have randomly varying paths, and so forth, but the model summarizes all this information into the form of a delay line with taps at fixed intervals. These models may therefore be regarded as canonical forms for the linear time-variant network under the different constraints imposed on it.

We may note in passing that the operational models we have derived are similar in form to the delay-line model used in the Rake system of communication and the filter-bank model used in the Kineplex method. Our analysis may be considered as establishing the sufficiency of such models.

The form of our sampling models would seem to be reasonable and expected – and in fact, the form of some of them was suggested in conversation by J. M. Wozencraft – but our study of the characterization of linear time-variant filters has pinned down exactly how the parameters of these models depend on the impulse and frequency response of the filters and the constraints imposed on them. It is worth noting that the modified definitions of impulse and frequency response that we have used and the definition of separable networks make the derivations and results particularly simple and intuitive. Picturing a linear time-variant network impulse response on an elapsed time-input time plane (Fig. 2) has been a simple and useful conceptual aid in this

(and some further) studies. The definition of a filter bandwidth, $2W_s$, under given operating conditions, is a compact way of characterizing the rate of variation of the network. The representation of a bivariate function or kernel $h(t, \tau)$ by a series such as $\sum_n f_n(t) g_n(\tau)$ is a useful mathematical technique, especially in the field of integral equations (27). The sampling theorems of Section III effectively do this for our impulse and frequency responses under the various constraints. Note, however, that physically meaningful series of the foregoing type are obtained only in terms of the variables z and y in the modified impulse responses, $h_2(z, \tau)$ and $h_3(y, t)$, and not for the variables t and τ in $h_1(t, \tau)$.

Finally the results and methods of our analysis in Section IV of the conditions under which our models can be determined by input-output measurements should be of use in studying communication problems in channels whose parameters cannot be assumed to be time-invariant for the duration of a signaling pulse. These results establish thresholds beyond which instantaneous measurement of the unknown channels is impossible, unless we have additional information about them.

Acknowledgment

The author is especially indebted to his thesis supervisor, Professor J. M. Wozencraft. Professor Wozencraft's generous encouragement and patient and stimulating guidance have been an excellent introduction to research.

It is also a pleasure to acknowledge helpful conversations with Professor P. Elias and Professor D. A. Huffman.

APPENDIX I

RELATIONS BETWEEN DIFFERENT FORMS OF IMPULSE RESPONSE AND FREQUENCY RESPONSE FUNCTION

These relations have been given in section 2.15, and here we explain how they are derived.

a. Transformations between impulse responses

These are almost evident from the relations

$$z = t - \tau = y$$

Thus, let us consider a unit impulse input to the network at time τ . The response of the network after z seconds is identically the same as the response at time $z + \tau$. Therefore,

$$h_2(z, \tau) = h_1(z + \tau, \tau)$$

The other relations are similarly obtained.

b. Input-output relations - time domain

We have already derived

$$y(t) = \int_{-\infty}^t h_1(t, \tau) x(\tau) d\tau$$

as Eq. 4 of section 2.12. Now let

$$\tau' = t - \tau, \quad d\tau' = -d\tau$$

and

$$y(t) = \int_0^{\infty} h_1(t, t - \tau') x(t - \tau') d\tau'$$

Change τ' to τ and we have Eq. 25. Equations 26 and 27 can be similarly derived, but they are most conveniently obtained from Eq. 25 by using the preceding transformations.

c. Transformations between frequency responses

We have

$$\begin{aligned} H_1(j\nu, j\mu) &= \iint h_1(t, \tau) e^{-j2\pi\nu(t-\tau)} e^{-j2\pi\mu t} d\tau dt \\ &= \iint h_2(t-\tau, \tau) e^{j2\pi\nu\tau} e^{-j2\pi(\nu+\mu)t} d\tau dt \end{aligned}$$

Let $t - \tau = z$; then $dt = dz$, and

$$\begin{aligned}
H_1(j\nu, j\mu) &= \int \int h_2(z, \tau) e^{-j2\pi\nu z} e^{-j2\pi\mu z} e^{-j2\pi\mu\tau} d\tau dz \\
&= H_2[j(\nu+\mu), j\mu]
\end{aligned}$$

which is Eq. 28. Equations 29 and 30 are rearrangements of Eq. 28.

d. Input-output relations – frequency domain

We have

$$\begin{aligned}
y(t) &= \int h_1(t, \tau) x(\tau) d\tau \\
&= \int h_1(t, \tau) e^{j2\pi\nu\tau} X(j\nu) d\nu d\tau \\
&= \int H_1(j\nu, t) e^{j2\pi\nu t} X(j\nu) d\nu
\end{aligned}$$

therefore

$$\begin{aligned}
Y(j\mu) &= \int \int H_1(j\nu, t) X(j\nu) e^{j2\pi\nu t} e^{-j2\pi\mu t} d\nu dt \\
&= \int H_1[j\nu, j(\nu-\mu)] X(j\nu) d\nu
\end{aligned}$$

which is Eq. 31. Equations 32 and 33 can be derived from 31 by making use of the transformations in c.

APPENDIX II

THE SAMPLING THEOREM IN THE FREQUENCY DOMAIN

Consider a time function $h(t)$ limited to $T = T_2 - T_1$. Using Woodward's method (22), we can write

$$\begin{aligned} h(t) &= \text{rep}_T h(t) \text{ rect } (t-T_0)/T \\ H(f) &= \text{comb}_{1/T} H(f) * \text{sinc } fT \exp(-2\pi j f T_0) \\ &= \sum_n H\left(\frac{n}{T}\right) \text{sinc } T\left(f - \frac{n}{T}\right) \exp\left[-2\pi j T_0\left(f - \frac{n}{T}\right)\right] \end{aligned} \quad (\text{II-1})$$

where $T_0 = (T_1 + T_2)/2$. This is one form of sampling representation. Another can be derived as follows: Transform both sides of Eq. II-1:

$$\begin{aligned} h(t) &= \sum_n H\left(\frac{n}{T}\right) \int \text{sinc } T\left(f - \frac{n}{T}\right) \exp\left[-2\pi j T_0\left(f - \frac{n}{T}\right)\right] \exp(j2\pi f t) df \\ &= \sum_n H\left(\frac{n}{T}\right) \exp\left(\frac{2\pi j n T_0}{T}\right) \int \text{sinc } T\left(f - \frac{n}{T}\right) \exp[2\pi j f (t - T_0)] df \\ &= \sum_n H\left(\frac{n}{T}\right) \exp\left(\frac{2\pi j n T_0}{T}\right) \exp\left[\frac{2\pi j n}{T}(t - T_0)\right] \text{rect } (t - T_0)/T \\ &= \sum_n H\left(\frac{n}{T}\right) \exp\left(\frac{2\pi j n t}{T}\right) \quad T_1 \leq t \leq T_2 \\ &= 0 \quad \text{elsewhere} \end{aligned} \quad (\text{II-2})$$

which we recognize as the Fourier series expansion of $h(t)$ over the interval (T_1, T_2) . This form also gives $2TW + 1$ degrees of freedom for a (W, T) function.

Proof of the Fourier Transform Relations: Eqs. 43 and 44

We first establish a preliminary result concerning the Fourier series for a periodic train of unit impulses.

Consider the periodic function

$$\begin{aligned} h(t) &= \sum_n \delta(t - nT) \\ &= \sum_n \exp\left(\frac{-2\pi j n t}{T}\right) \cdot c_n \end{aligned}$$

where

$$\begin{aligned}
c_n &= \frac{1}{T} \int_0^T h(t) \exp\left(\frac{+2\pi j n t}{T}\right) dt \\
&= \frac{1}{T} \int_0^T \sum_n \delta(t-nT) \exp\left(\frac{+2\pi j n t}{T}\right) dt \\
&= \frac{1}{T} \quad \text{for all } n
\end{aligned}$$

Therefore,

$$\sum_n \delta(t-nT) = \sum_n \frac{1}{T} \exp\left(\frac{-2\pi j n t}{T}\right)$$

Now, we have

$$\begin{aligned}
\mathcal{F}[\text{rep}_T h(t)] &= \mathcal{F}\left[\sum_n h(t-nT)\right] \\
&= H(f) \cdot \sum_n e^{-2\pi j n T} \\
&= H(f) \cdot \sum_n \frac{1}{T} \delta\left(t - \frac{n}{T}\right) \\
&= \frac{1}{T} \sum_n H\left(\frac{n}{T}\right) \delta\left(t - \frac{n}{T}\right) \\
&= \frac{1}{T} \cdot \text{comb}_{1/T} H(f)
\end{aligned}$$

which is Eq. 43.

Similarly

$$\begin{aligned}
\mathcal{F}[\text{comb}_T h(t)] &= \mathcal{F}\left[\sum_n h(nT) \delta(t-nT)\right] \\
&= \frac{1}{T} \sum_n h(t) \exp\left(\frac{-2\pi j n t}{T}\right) \\
&= \frac{1}{T} \sum_n H\left(f + \frac{n}{T}\right) \\
&= \frac{1}{T} \sum_n H\left(f - \frac{n}{T}\right) \\
&= \frac{1}{T} \text{rep}_{1/T} H(f)
\end{aligned}$$

which is Eq. 44.

APPENDIX III

MEASUREMENT CONDITIONS FOR OTHER CONSTRAINTS

These are derived by methods very similar to those we have used for Case II. A careful treatment would be as lengthy as the one given for Case II in Section IV, and therefore we shall content ourselves with a brief indication of the arguments to be used in these other cases.

Case I. Limited Input Frequency Range

A matrix relation between input and output is derived first. We have, from Eq. 68 of Section III,

$$\begin{aligned} h_3(y, t) &= \sum_n \sum_m h_3\left(\frac{n}{2W_i}, \frac{m}{2W_s}\right) \text{sinc } 2W_s\left(t - \frac{m}{2W_s}\right) \text{sinc } 2W_i\left(y - \frac{n}{2W_i}\right) \\ &= \sum_n f_n(t) \text{sinc}(y-n) \end{aligned} \quad (\text{III-1})$$

if $2W_i$ is normalized to unity. Then

$$h_1(t, \tau) = \sum_n f_n(t) \text{sinc}(t-\tau-n)$$

and

$$\begin{aligned} y(t) &= \sum_n \int f_n(t) \text{sinc}(t-\tau-n) x(\tau) d\tau \\ &= \sum_n f_n(t) \int \text{sinc}(t-\tau-n) x(\tau) d\tau \end{aligned}$$

and, therefore, for integral t , we have

$$y(t) = \sum_n f_n(t) x(t-n) \quad (\text{III-2})$$

if $x(t)$ is of bandwidth not greater than $2W_i$. This can be written

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \end{bmatrix} = \begin{bmatrix} h_3(0,0) & 0 & 0 & . & . \\ h_3(1,1) & h_3(0,1) & 0 & . & . \\ h_3(2,2) & h_3(1,2) & h_3(0,2) & . & . \end{bmatrix} \cdot \begin{bmatrix} x(0) \\ x(1) \\ x(2) \end{bmatrix} \quad (\text{III-3})$$

Notice from the relation

$$h_3(n, t) = h_2(n, t-n)$$

that this matrix is very similar to the matrix $\underline{H}$ derived in Section IV.

To find necessary and sufficient conditions on measurement with input signals limited to $(-W_i, W_i)$, we consider first the number of unknowns to be determined. Over a t -span T , and a y -span Y , we have

$$(2YW_i+1)(2W_s T+1)$$

independent degrees of freedom. The time interval over which nonzero output waveforms may be expected is $T + Y$ seconds. The bandwidth of the output is not greater than $2(W_i+W_s)$. The (apparent) degrees of freedom of the output is $2(T+Y)(W_i+W_s) + 1$. However, the degrees of freedom resulting from the frequency expansion $2W_s$ cannot be counted because they are not truly independent of the other degrees of freedom, but are calculable from the other degrees of freedom if the form of the frequency-expansion mechanism is known. Therefore the true number of degrees of freedom in this case is $2(T+Y) W_i + 1$.

Now we must have

$$(2W_i Y+1)(2W_s T+1) \leq 2(T+Y) W_i + 1$$

or

$$Y \leq \frac{1}{2W_s} - \frac{1}{2W_i} \quad (\text{III-4})$$

Sufficiency can be proved by using Eq. III-3 with inputs having $x(t) = 1$, at $t = 0, \frac{1}{2W_s}, \dots, T$, and zero at other integral t .

Case IIIa. Limited Filter Memory – Potential Restriction on Input Frequency Range
From Eq. 70 of Section III, we have

$$h_3(y, t) = \sum_{\ell} f_{\ell}(t) g_{\ell}(y)$$

and therefore, for an impulse input $x(t) = \delta(t-T)$

$$\begin{aligned} y(t) &= \int \sum_{\ell} f_{\ell}(t) g_{\ell}(t-y) x(y) dy \\ &= \int \sum_{\ell} f_{\ell}(t) g_{\ell}(t-y) x(y) dy \\ &= \sum_{\ell} f_{\ell}(t) \exp\left(\frac{2\pi j \ell t}{Y}\right) \text{rect}\left(\frac{t-T}{Y} - \frac{1}{2}\right) \end{aligned} \quad (\text{III-5})$$

$$= \begin{cases} \sum_{\ell} H_3\left(\frac{j\ell}{Y}, T\right) \exp\left(\frac{2\pi j \ell t}{Y}\right) & T \leq y \leq T + Y \\ 0 & \text{elsewhere} \end{cases} \quad (\text{III-6})$$

(assuming T/Y to be integral).

We have to find $H_3\left(\frac{j\ell}{Y}, \frac{m}{2W_s}\right)$ to determine the network. Now in Eq. 6, if we multiply $y(t)$ by $\exp - (2\pi j\ell t)/Y$ and filter out the dc term we will get $H_3\left(\frac{j\ell}{Y}, T\right)$. We can do this for all the ℓ 's at a given T . Then if $T = \frac{1}{2W_s}$ is greater than Y , we can, for different T 's, similarly isolate the corresponding $y(t)$ functions and determine $H_3\left(\frac{j\ell}{Y}, T\right)$. Therefore a sufficient condition is

$$2W_s \leq \frac{1}{Y} \quad (\text{III-7})$$

To determine if this is a necessary condition, we postulate a cut-off frequency W_i , and, calculating degrees of freedom for the output and the system, we have

$$(2W_i Y + 1)(2W_s T + 1) \leq 2(T + Y) W_i + 1 \quad (\text{III-8})$$

or

$$Y \leq \frac{1}{2W_s} - \frac{1}{2W_i} \quad (\text{III-9})$$

Now, as $2W_i \rightarrow \infty$,

$$2W_s Y \rightarrow 1 \quad (\text{III-10})$$

Case IIb. Limited Filter Memory – Potential Restriction on Output Frequency Range
We have, from Eq. 71 of Section III

$$h_2(z, \tau) = \sum_{\ell} f_{\ell}(\tau) g_{\ell}(z)$$

and therefore, for an impulse input $x(t) = \delta(t - T)$,

$$\begin{aligned} y(t) &= \sum_{\ell} \int g_{\ell}(t - z) f_{\ell}(z) x(z) dz \\ &= \sum_{\ell} f_{\ell}(T) g_{\ell}(t - T) \\ &= \sum_{\ell} H_2\left(\frac{j\ell}{Z}, T\right) \exp\left[\frac{2\pi j\ell}{Z}(t - T)\right] \text{rect}\left(\frac{t - T}{Z} - \frac{1}{2}\right) \\ &= \begin{cases} \sum_{\ell} H_2\left(\frac{j\ell}{Z}, T\right) \exp\left(\frac{2\pi j\ell t}{Z}\right) & T \leq t \leq T + Z \\ 0 & \text{elsewhere} \end{cases} \end{aligned}$$

(if T/Z is integral).

We have to find the $H_2\left(\frac{j\ell}{Z}, \frac{m}{2W_s}\right)$ to determine the network. From Eq. III-10 we see that this can be done by putting in impulses at $T = 0, \frac{1}{2W_s}, \dots$, provided the duration, Z , of the impulse response is less than $\frac{1}{2W_s}$, for we then need only multiply $y(t)$ by $\exp - (2\pi j\ell t)/Z$ and filter out the dc term to get $H_2\left(\frac{j\ell}{Z}, T\right)$. This can be done for all ℓ and for the different T .

And finally, using a degrees-of-freedom argument as in the other cases, we have

$$(2ZW_o + 1)(2TW_s + 1) \leq 2(T + Z) W_o + 1 \quad (\text{III-11})$$

or

$$Z \leq \frac{1}{2W_s} - \frac{1}{2W_o}$$

and as $W_o \rightarrow \infty$, $2W_s Z \rightarrow 1$. Since Z is equal to Y , this is, as we would expect, the same result as in Case IIIa.

If we do have restrictions on input or output frequency ranges, say $(-W_i, W_i)$ or $(-W_o, W_o)$, the conditions become

$$Y \leq \frac{1}{2W_s} - \frac{1}{2W_i} \quad (\text{III-12})$$

and

$$Z \leq \frac{1}{2W_s} - \frac{1}{2W_o} \quad (\text{III-13})$$

The necessity of these conditions is shown by Eqs. III-9 and III-11, but there appears to be a discrepancy (that is, no terms in W_i or W_o) in the sufficiency arguments. This is because in our sufficiency proofs, for finite W_i and W_o , there will be some uncertainty about the values of the impulse response at the instants T where one waveform ends and another begins. A more careful sampling-point analysis, as in Section IV, takes account of this and can be used to show that conditions III-12 and III-13 are indeed necessary and sufficient.

References

1. H. A. Thomas and R. E. Burgess, Survey of existing information and data on radio noise in the frequency range, 1-30 mc, Special Report No. 15, Radio Research Board (H. M. Stationery Office, London, England, 1947).
2. S. V. C. Aiya, Atmospheric noise interference to short-wave broadcasting, Proc. IRE 46, 580-588 (1958).
3. Ionospheric Radio Propagation, Circular 462, National Bureau of Standards, Washington, D.C., June 25, 1948.
4. S. H. Van Wambeck and A. H. Ross, Performance of diversity receiving systems, Proc. IRE 39, 256-264 (1951).
5. R. Price and P. E. Green, Jr., A communication technique for multipath channels, Proc. IRE 46, 555-570 (1958).
6. R. Price, Statistical theory applied to communication through multipath disturbances, Technical Report 266, Research Laboratory of Electronics, M.I.T., Sept. 3, 1953.
7. R. Price, Optimum detection of random signals in noise, with application to scatter-multipath communication, I, Trans. IRE, IT-2, 125-135 (1956).
8. W. L. Root and T. S. Pitcher, Some remarks on statistical detection, Trans. IRE, IT-1, 33-38 (1955).
9. G. L. Turin, Communication through noisy, random-multipath channels, Technical Report 116, Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, May 1956.
10. G. L. Turin, A review of statistical multipath communication theory, paper presented at the Second National Symposium on Global Communications, St. Petersburg, Florida, Dec. 3-5, 1958.
11. L. A. Pipes, Four methods for the analysis of time-variable circuits, Trans. IRE, CT-2, 4-11 (1955).
12. W. R. Bennett, A general review of linear variable parameter and nonlinear circuit analysis, Proc. IRE 38, 259-263 (1950).
13. A. Rosenbloom, Analysis of linear systems with randomly time-varying parameters, Proceedings of the Brooklyn Polytechnic Institute Symposium on Information Networks, pp. 145-153, January 1955.
14. K. S. Miller, Properties of impulsive responses and Green's Functions, Trans. IRE, CT-2, 26-31 (1955).
15. L. A. Zadeh, Frequency analysis of variable networks, Proc. IRE 38, 291-299 (1950).
16. J. A. Aseltine, Transforms for linear time-varying systems, Technical Report 52-1, Department of Engineering, University of California, Los Angeles, 1952.
17. A. P. Bolle, Application of complex symbolism to linear variable networks, Trans. IRE, CT-2, 32-35 (1955).
18. L. A. Zadeh, A general theory of linear signal transmission systems, J. Franklin Inst., 253, 293-312 (April 1952).
19. L. A. Zadeh, Fundamental aspects of linear multiplexing, Proc. IRE 40, 1091-1097 (1952).
20. J. S. Bendat, Principles and Applications of Random Noise (John Wiley and Sons, Inc., New York, 1958).
21. J. H. Laning and R. H. Battin, Random Processes in Automatic Control (McGraw-Hill Book Company, Inc., New York, 1956).
22. P. M. Woodward, Probability and Information Theory with Applications to Radar (McGraw-Hill Book Company, Inc., New York, 1953).

23. W. K. Linvill, Sampled-data control systems studied through comparison with amplitude modulation, Trans. AIEE 70, 1779-1788 (1951).
24. M. L. Doelz, E. T. Heald, and D. L. Martin, Binary data transmission techniques for linear systems, Proc. IRE 45, 656-661 (1957).
25. B. Friedland, Theory of time-varying sampled-data systems, Technical Report T-19/B, Electronics Research Laboratories, Columbia University, New York, April 1957.
26. D. Middleton and D. Van Meter, Detection and extraction of signals in noise from the point of view of statistical decision theory, J. Soc. Indus. Appl. Math. 3, 192-253 (December 1955); 4, 86-119 (June 1956).
27. R. Courant and D. Hilbert, Methods of Mathematical Physics (Interscience Publishers, New York, 1953), Vol. 1.

