STATE UNIVERSITY OF NEW YORK AT STONY BROOK

COLLEGE OF ENGINEERING

Report No. 101

APPLICATIONS OF GENERALIZED FUNCTIONS TO NETWORK THEORY

by

A. H. Zemanian

December, 1967

An invited lecture presented to the Summer School on Circuit Theory - 1968

Prague, Czechoslovakia

June 28-July 12, 1968

REPORT No. 101

APPLICATIONS OF GENERALIZED FUNCTIONS TO NETWORK THEORY

by

A. H. Zemanian

DECEMBER, 1967

An invited lecture presented to the Summer School on

Circuit Theory - 1968

Prague, Czechoslovakia

June 28-July 12, 1968

APPLICATIONS OF GENERALIZED FUNCTIONS TO NETWORK THEORY

A. H. Zemanian

State University of New York at Stony Brook

Stony Brook, N. Y. 11790, U. S. A.

1. Introduction

It has been some twenty years now since L. Schwartz developed his theory of distributions (Schwartz [1]). Its importance in pure mathematics was instantly recognized, and a large body of mathematical literature on the subject has since been published. A good part of this literature is concerned with the applications of distributions and other generalized functions to the theory of partial differential equations (Friedman [1], Gelfand and Shilov [1], and Hormander [1]), and as a consequence the latter theory has received a strong impetus in its developments. In the mathematical sciences the most notable application of distribution theory has been to quantum field theory (Bremermann [1], de Jager [1], Guttinger [1], Streater and Wightman [1], and Vladimirov [1], wherein it has been crucially important. Turning to the subject at hand, namely, network theory, we feel it is fair to say that distribution theory has been and will continue to be of value in certain areas, most notably in the axiomatic foundations of linear system theory, but is useless in other areas, such as nonlinear system theory.

Our original intent in writing this paper was to discuss most of the applications of generalized functions to network theory, but it rapidly became clear that the paper would be excessively long if this objective was adhered to. Because of this we shall examine only a few parts of the subject and will, in fact, devote most of the discussion to the axiomatic foundations of linear one-port theory. In this introduction we shall merely point to the literature where other applications of generalized functions to network theory are given. No claim is being made that the following literature citations exhaust the pertinent bibliography.

In certain problems the use of generalized functions is essential; the results obtained could not have been developed without them. This occurs, for example, in the time-domain theory of linear n-ports, as is indicated in Secs.

4 and 5, and in obtaining a frequency-domain criterion for the causality of active networks, a subject we discuss in Sec. 6. This is also the case in the theory of the generalized Bode equations and in the characterization of various broad classes of systems by their real frequency behavior. See Arsac [1] (Chapter 7), Beltrami [1], Beltrami and Wohlers [1] (Chapter III), [2], Bremermann [1] (Chapter 11), Guttinger [1], Lauwerier [1], Schwartz [3], Taylor [1], Tillmann [1], and Wohlers [1]. Generalized functions are also used in an essential way in the analysis and synthesis of time-varying networks; see, for example, Anderson and Newcomb [1], Dolezal [1], [2], [3], Newcomb [2], Newcomb and Anderson [1], and Spaulding and Newcomb [1].

In other subjects various classical problems, which had been solved in terms of classical mathematics, become open problems once again when they are reformulated in terms of generalized functions. An example of this is the classical time-domain approximation problem which requires the construction of a function that can be realized as the response of a specified type of network and that approximates a given signal in some sense. If the given signal is a generalized function, the classical solutions become inapplicable and substantial modifications in the approximation theory must be made in order to make the problem tractable again. For two examples of this, see Zemanian [3] (Sec. 9.10), [4], [5].

Similarly, it is well-known that certain types of time-varying networks can be analyzed by means of the Mellin, Hankel, and K transformations (Aseltine [1], Gerardi [1]). Here again, these classical transformations become inapplicable if the signals within the network are generalized functions. One must now resort to the corresponding generalized integral transformations; see Zemanian [3] (Secs. 4.5 and 6.9).

2. Some Notations and Results from the Theory of Distributions

The symbols and terminology used in this paper follow that employed in Zemanian [1], and we refer the reader to that source for a more detailed discussion of the definitions used here. D and D′ denote the conventional Schwartz spaces of testing functions of compact support and distributions respectively. Also, S is the space of testing functions of rapid descent and

S' the space of distributions of slow growth. E' is the space of distributions of compact support. The usual topologies for these spaces are understood. We shall use the terms "distributions" and "generalized functions" in different ways: A generalized function is any continuous linear functional on some testing-function space as defined in Zemanian [3], Sec. 2.4. On the other hand, a distribution is a special type of generalized function and is, namely, any member of D'. For a generalized function to be a distribution it is not sufficient that its domain contain D and its restriction to D be linear and continuous on D; it must also be such that its behavior as a generalized function is completely determined by its restriction to D. The dual of B (see Schwartz [1], vol. II, p. 56) and Z' (the space of Fourier transforms of all $f \in D'$) are spaces of generalized functions but not spaces of distributions.

R denotes the real line and throughout this paper t, τ, x, σ, and ω are variables in R; also, p, u, and v are complex variables with $p = \sigma + i\omega$. $[a, b]$ and (a, b) denote respectively a closed and an open interval on the real line with endpoints a and b, $a < b$; the notations $[a, b)$ and $(a, b]$ are defined similarly. Supp f denotes the support of either a conventional function or distribution f. $f^{(n)}$ denotes the nth derivative of f. $\check{f}$ is the transpose of f; i.e., $\check{f}(t) = f(-t)$. A smooth function is one having continuous derivatives of all orders at all points of its domain.

D_R (respectively, D_R' denotes the space of smooth functions (respectively, distributions) on R whose supports are bounded on the left. D_R' is not the dual of D_R; also, $D_R \subset D_R'$. D_+' is the space of distributions on R whose supports are bounded on the left at the origin. Thus, $f \in D_+'$ if and only if $f \in D'$ and supp $f \subset [0, \infty)$. A sequence $\{\varphi_n\}$ converges in D_R if and only if there is a fixed real number $T > -\infty$ such that supp $\varphi_n \subset [T, \infty)$ for all n and, for each nonnegative integer k, $\{D^k\varphi_n\}_n$ converges uniformly on every compact subset of R. Similarly, a sequence $\{f_n\}$ converges in D_R' if and only if, for some fixed real number $T > -\infty$, supp $f_n \subset [T, \infty)$ for all n and $\{f_n\}$ converges in D'. D is dense in D' and D_R'.

At various places in this work we will be concerned with a conventional or generalized function y in some space A that is defined either on the real line or on the real plane. To indicate that y is defined on the real line we write $y = y(t) \in A|_t$, and to indicate that it is defined on the real plane we write $y = y(t, \tau) \in A|_{t, \tau}$. A similar notation is used when t and τ are replaced by the complex variables u and v and y is a conventional or generalized function on either the one-dimensional or two-dimensional complex euclidean spaces.

A standard form in distribution theory is Schwartz's kernel representation (Schwartz [2]) which is defined in the following way.
Let $y = y(t, \tau) \in D'|_{t, \tau}$ and let $v = v(\tau) \in D|_\tau$.
Then, $y \circ v = y(t, \tau) \circ v(\tau)$ denotes a distribution in $D'|_t$ defined as follows:
For any $\varphi = \varphi(t) \in D|_t$,

$$\langle y \circ v, \varphi \rangle \triangleq \langle y(t, \tau), v(\tau)\, \varphi(t) \rangle$$

Thus, $v \mapsto y \circ v$ is a mapping of D into D' and is, in fact, linear and continuous from D into D'. The converse also happens to be true: every continuous linear mapping of D into D' has a kernel representation. Under suitable restrictions on y this mapping can be extended from D onto wider spaces of distributions (say, onto A). But, in the latter case the right-hand side of the above definition may not possess a sense. (The extension is made via the continuity of the mapping on A and the denseness of D in A.) We shall make use of these facts later on.

Still more symbols will be introduced in Sec. 6 as the need arises.

3. Why Bother with Distributions?

One reply is that distributions arise quite naturally as the responses of systems that contain perfect differentiators, such as the pure capacitor under a voltage excitation. In classical analysis any signal that is not differentiable at certain times cannot be an allowable input to such a system. An exception can be made for signals with ordinary discontinuities by introducing the delta functional and its derivatives, which in itself testifies to the need for distributions or other types of generalized functions. If one wishes to allow an arbitrary number of perfect differentiators in various types of systems and wants to analyze these systems in a classically rigorous manner, he will have to restrict the allowable signals to those functions having continuous derivatives of sufficiently high order. Under a distributional analysis he is free to introduce any locally integrable signal or, more generally, any distributional signal.

But, even if one does restrict all signals to smooth functions, the use of distribution theory is still advantageous. Indeed, an input-output system can be viewed as a mapping of one class of smooth functions into another such class. Under certain linearity and continuity assumptions this mapping can be represented in the time-domain by means of a distributional kernel (Schwartz [2]), and one can then employ all the results connected with Schwartz's kernel theorem. This result applies to time-varying as well as time-invariant systems. From a physical point of view the kernel $y(t, \tau)$ for the system can be interpreted as the response of the system to the delta functional $\delta(t - \tau)$ applied at time $t = \tau$.

That the kernel is in general a distribution and not a function is illustrated by a fixed inductor L whose kernel is $\delta^{(1)}(t - \tau)$. But, since the first derivative of the delta functional is such a familiar object, let's present another system whose kernel is a singular distribution.

Consider the infinite Foster-type network shown in Fig. 1, where we assume that the voltage v is the input and the current i is the output. The response $y(t)$ of this system to $v(t) = \delta(t)$ is formally

$$y(t) = a_0 1_+(t) + 2 \sum_{n=1}^{\infty} a_n 1_+(t) \cos n t$$

where

$$L_0 = \frac{1}{a_0}$$

$$L_n = \frac{1}{2a_n}, \quad C_n = \frac{1}{n^2 L_n}, \quad n = 1, 2, 3, \ldots$$

$$1_+(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{2} & t = 0 \\ 1 & t > 0 \end{cases}$$

A sufficient condition for the series for $y(t)$ to converge in the distributional sense and for the infinite network of Fig. 1 to have a sense as the limit of a sequence of finite Foster networks is that the coefficients a_n be such that

$$\sum_{n=1}^{\infty} n^{-2} |a_n| < \infty$$

[In fact, this condition is necessary as well as sufficient for the series for $y(t)$ to converge in D'.] In the special case where $a_n = 1$ for all $n = 0, 1, 2, \ldots$, $y(t)$ can be shown to be equal to the delta train:

$$y(t) = \pi\, \delta(t) + 2\pi \sum_{m=1}^{\infty} \delta(t - 2\pi m).$$

For other permissible choices of the a_n, $y(t)$ can be made equal to other distributions that are not conventional functions. The kernel for the system $v \mapsto i$ of Fig. 1 is $y(t-\tau)$.

Still other systems containing an infinite collection of ideal elements whose time or frequency domain characterizations require the use of generalized functions are given in Zemanian [6]. It is worth noting that the frequency-domain descriptions of the two systems in Zemanian [6] lie outside the realm of distributional as well as classical mathematics and require instead the theory of generalized functions. Another situation like this is presented in Sec. 6.

We end this apologia for the introduction of generalized functions into network theory by merely referring to a cogent argument devised by Liverman [1]. When a voltage v is represented by a conventional function $v(t)$, it is implicitly being asserted that at some precise instant of time the precise value $v(t)$ of voltage is known. This is a fiction. The perception of any physical variable requires some kind of measurement. Any measurement of $v(t)$ will contain two random errors, one in the value of the voltage and one in the determination of the time t at which the measurement is being made. Upon taking into account the presence of such errors, one is naturally lead to a functional description of v. Namely, v is more properly described as a functional on the subspace P of all probability density functions in D. Liverman proved (among other things) that the values that any $f \in D'$ assigns to P determine f on all of D; in other words, we need only test a given distribution on the space P in order to determine it. This result is not only intuitively satisfying, but will probably be quite useful.

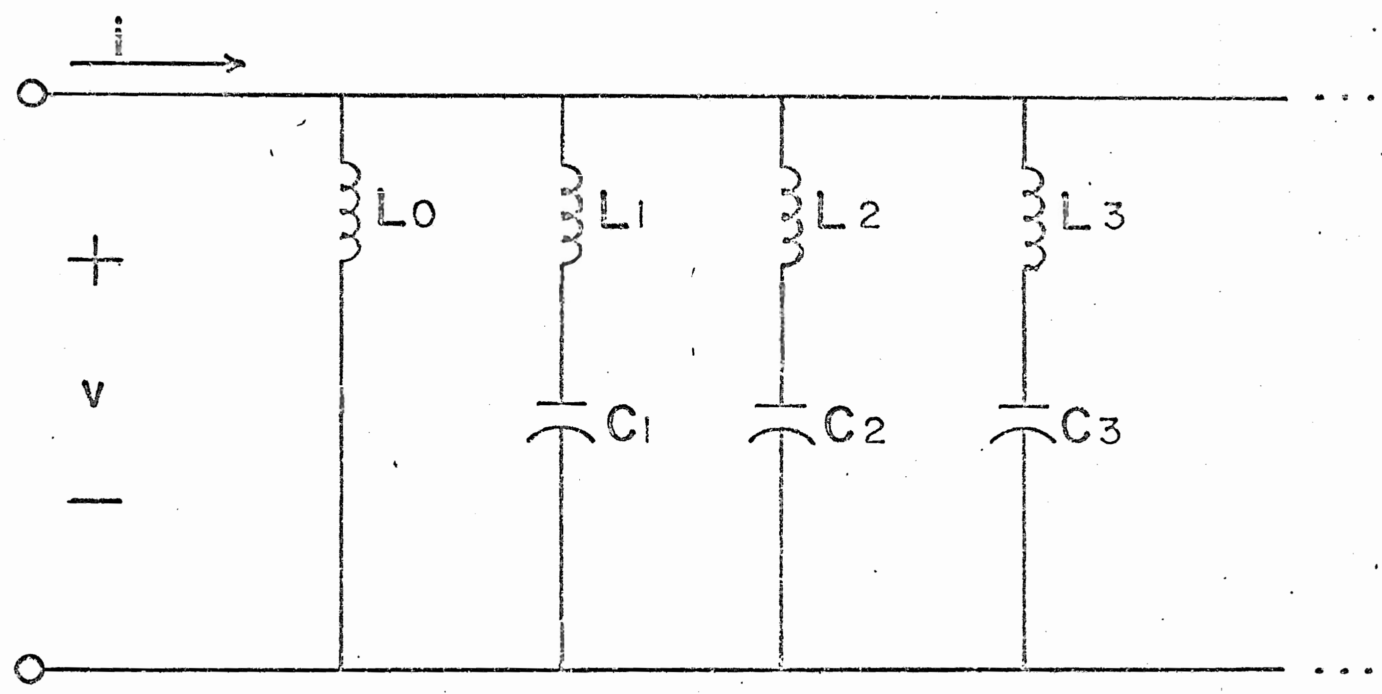

FIG. I

4. An Axiomatic Approach to One-ports: The Admittance Formulism

We shall discuss one-ports rather than n-ports since almost all the important ideas can be discussed in this context. The n-port theory follows essentially the same development as that for the one-port theory, but requires a more complicated notation (see Beltrami [1], Beltrami and Wohlers [1], Newcomb [1], and Zemanian [2]. We also assume that all voltages, currents, and scattering parameters (in the time domain) are real since nothing is gained by allowing complex quantities. We view a one-port N as an operator mapping voltages into currents and possessing certain properties as described by the postulates stated below. These postulates are similar but not the same as those used by Konig and Meixner [1] and Zemanian[1], Chapter 10.

P1. N is a single-valued mapping of D into D′.

As is indicated here, we at first restrict the domain of N to the space D. Later on, N will be extended in a unique way onto various spaces of distributions.

P2. N is linear on D.

This means that given any two real numbers α and β and any two functions v_1 and v_2 in D, we have

$$N(\alpha v_1 + \beta v_2) = \alpha N v_1 + \beta N v_2 .$$

P3. N is continuous from D into D′.

That is, if $v_n \to v$ in D as $n \to \infty$, then $N v_n \to Nv$ in D'.

These first three postulates allow us to invoke Schwartz's kernel theorem (Schwartz [2]) to show that N has a kernel representation on D.

Theorem 1: N satisfies P1, P2, and P3 if and only if there exists a unique kernel $y = y(t, \tau) \in D' \mid_{t,\tau}$ such that $N = y\bullet$ on D. (That is, $Nv = y \bullet v$ for all $v \in D$.)

We call y the admittance of the one-port.

P4. N is time-invariant on D.

To explain this, let σ_x be the shifting operator defined on any conventional function or distribution f(t) by $\sigma_x f(t) = f(t + x)$ where x is any real number. Then, the postulate means that N commutes with σ_x whenever N operates on D (i.e., $\sigma_x Nv = N\sigma_x v$ for all $v \in D$ and all x.)

Under this additional postulate the kernel representation becomes a convolution representation (Schwartz [1], vol. II, pp. 53-54); that is, $y(t, \tau)$ becomes $y(t-\tau)$ where $y(t)$ is now a member of $D'|_t$. In particular, we have

Theorem 2: N satisfies P1 through P4 if and only if there exists a unique $y = y(t) \in D'|_t$ such that $N = y *$ (i.e., $Nv = y * v = \langle y(\tau), v(t-\tau)\rangle$ for all $v \in D$).

We can now extend N via its convolution representation onto the space E' of distributions of compact support. That is, for any $v \in E'$, Nv is defined as $y * v \in D'$. Because D is dense in E', this extension of N is unique: There cannot be another continuous linear mapping of E' into D' that agrees with N on D but differs from N on some other member of E'. Now, y can be identified as the unit impulse response of N; that is, $y = N\delta = y * \delta$ where δ denotes the delta functional. Furthermore, if the distribution y happens to be suitably restricted, we can extend N onto still larger spaces of distributions via the convolution representation. For example, if $y \in D'_R$, then N can be extended onto D'_R, and, if $y \in E'$, then N can can be extended onto all of D'. These extensions are also unique because D is dense in D'_R as well as in D'.

In the case where N does not satisfy P4 the same sort of unique extension can be made via the kernel representation: $N = y\bullet$. Let's be more precise. Assume that A is a complete countably multinormed or a complete countable-union space of smooth functions on R (see Zemanian [3], Chapter 1), and let A' be the weak dual of A. Suppose that A and A' possess the following four properties:

I. $D \subset A$; convergence in D implies convergence in A; D is dense in A.

II. If $\psi \in A$ and if λ is a smooth function on R such that it and each of its derivatives are bounded on R (i.e., $\lambda \in B$), then $\lambda\psi \in A$.

III. D is dense in A'.

IV. For certain (but not necessarily all) $w \in D'$, the operator $w\bullet$ is defined on A' and is a continuous linear mapping of A' into D' (see Schwartz [2], pp. 224-225).

Condition I implies that A' is a subspace of D' and that convergence in A' implies convergence in D'. We will use condition II in a subsequent proof. It is conditions III and IV that allow us to extend $N = y\bullet$ onto A' if y happens to be one of the w indicated in IV. From now on we shall always assume that N has been extended through the right-hand side of $N = y*$ onto every such space of distributions that satisfies the above four conditions with $y = w$. Condition III implies that this extension is unique in the aforementioned sense. We also assume that N is extended no further.

P5. N is passive on D.

This means that, for every $v \in D$, $i \triangleq Nv$ is locally integrable (i.e., Lebesgue integrable on every bounded interval), and in addition, for every real finite number t, we have that

$$\int_{-\infty}^{t} v(x)i(x)dx \geq 0 \tag{1}$$

Note that, if N satisfies P1 through P4, then, for any $v \in D$, $i = y * v$ is smooth and therefore locally integrable.

Definition of causality: Let v_1 and v_2 be distributions in the domain of an operator W mapping D' or a subset of D' into D', and let $i_1 = Wv_1$ and $i_2 = Wv_2$. W is said to be causal (or to satisfy causality) if the condition $v_1(t) = v_2(t)$ on $-\infty < t < t_o$ implies that $i_1(t) = i_2(t)$ on $-\infty < t < t_o$ and if this property holds for all real values of t_o.

The equalities herein are understood in the sense of equality in D'. Also, a causal operator is clearly single-valued. Another obvious result that we shall need later on is

Lemma 1: If N is a causal mapping of D_R into D' and is passive on D, then, for all $v \in D_R$, $i \triangleq Nv$ is locally integrable and (1) holds for all $t < \infty$ (i.e., N is passive on D_R as well).

We can characterize causality in the following way:

Theorem 3: Let N satisfy P1, P2, and P3; then, (the extended) N is causal if and only if the support of $y = y(t, \tau)$ is contained in the half-plane $\Lambda \triangleq \{(t, \tau): t \geq \tau\}$. Next, assume in addition that N satisfies P4; then, (the extended) N is causal if and only if the support of $y = y(t)$ is contained in the interval $0 \leq t < \infty$.

Proof: The second sentence follows from the first one through the equation $y(t, \tau) = y(t-\tau)$. To prove the first sentence, let $v \in D$ be such that $v(\tau) \equiv 0$ for $\tau < t_o$, and let $\varphi \in D$ be such that supp $\varphi \subset (-\infty, t_o)$. Then, the support of $v(\tau)\ \varphi(t)$ is contained in the half-plane $\{(t,\tau): t < \tau\}$.

Moreover,

$$\langle Nv, \varphi \rangle = \langle y \circ v, \varphi \rangle = \langle y(t,\tau), v(\tau)\, \varphi(t) \rangle$$

The only way that the right-hand side can equal zero for every such v and φ is that supp $y(t, \tau) \subset \Lambda$. This proves the "only if" part of the first sentence.

This also proves that $i(t) \triangleq (Nv)(t) = 0$ distributionally for $t < t_o$ whenever supp $y(t, \tau) \subset \Lambda$, $v \in D$, and $v(t) \equiv 0$ on $-\infty < t < t_o$. We shall show that the same result holds when v is any distribution in the domain of N that is distributionally equal to zero on $-\infty < t < t_o$. Indeed, by the way N was extended, v is a member of some space of distributions A' on which N is a continuous linear mapping into D' and in which D is dense. Given any $\epsilon > 0$, choose a sequence of functions $v_n \in D$ such that $v_n \to v$ in A' as $n \to \infty$ and supp $v_n \subset (t_o - \epsilon, \infty)$ for all n. (That this can be done follows from condition II above and the fact that supp v is contained in $t_o \leq t < \infty$.) By our previous result, $i_n \triangleq Nv_n = 0$ distributionally on $-\infty < t < t_o - \epsilon$. But, $i_n \to i \triangleq Nv$ in D' since N is continuous on A'. Therefore, $i = 0$ distributionally on $-\infty < t < t_o - \epsilon$. Since $\epsilon > 0$ was arbitrary, this is true on $-\infty < t < t_o$.

The "if" part of the first sentence now follows from the linearity of N. Q.E.D.

A remarkable fact discovered by Youla, Castriota, and Carlin [1] (see also Youla [1]) states essentially that linearity and passivity imply causality. For one-ports possessing kernel representations, this fact can be stated as follows.

Theorem 4: If N satisfies P1, P2, P3, and P5, then N is causal.

For a proof, see Zemanian [1], pp. 301-303. (The proof of lemma 2 on p. 301 requires now a modification to make it applicable to kernel representations; the argument needed is that given in the first paragraph of the proof of theorem 3 in this paper.)

Theorems 1, 3, and 4 imply that, if N satisfies P1, P2, P3, and P5, then D'_R is contained in the domain of (the extended) N.

The next theorem characterizes, in six different ways, one-ports that satisfy P1 through P5.

Theorem 5: If N satisfies P1 through P5, then $y \triangleq N\delta$ (the extended N is understood here) satisfies the following six equivalent conditions. Conversely, if $y \in D'$ satisfies any one of the following conditions, then $N \triangleq y *$ satisfies P1 through P5.

1. The Laplace transform Y of y is a positive-real function.
2. y has the representation:

$$y(t) = \alpha\delta^{(1)}(t) + 1_+(t) \int_{-\infty}^{\infty} (1 + \eta^2) \cos \eta t \, dH(\eta) \tag{2}$$

where α is a real nonnegative number, $1_+(t)$ is the unit step function defined in Sec. 3, and $H(\eta)$ is a real nondecreasing bounded function on $-\infty < \eta < \infty$.

3. y also has the representation:

$$y(t) = \alpha\delta^{(1)}(t) + \beta \, 1_+(t) + p(t) \, 1_+(t) + \frac{d^2}{dt^2} \{1_+(t)[p(0)-p(t)]\}. \tag{3}$$

Here, α and β are real nonnegative numbers and

$$p(t) = \int_{-\infty}^{\infty} \cos \eta t \, dM(\eta) \tag{4}$$

where $M(\eta)$ is a real nondecreasing bounded function continuous at the origin.

4. $y = \alpha\delta^{(1)} + y_0$ where α is a real nonnegative number, and y_0 is a member of D'_+ and a distribution of zero C-order; moreover, the even part of y, namely:

$$y_e(t) = \tfrac{1}{2} [y(t) + y(-t)]$$

is a nonnegative-definite distribution.

5. $y \in D'_+$, and, for every real nonnegative number c and every $\varphi \in D$,

$$\langle y(t)e^{-ct}, \varphi(t) * \varphi(-t)\rangle \geq 0 \tag{5}$$

6. Let $\tilde{y}$ denote the Fourier transform of y. Then,

$$\tilde{y}(\omega) = i\omega \left[\alpha + \int_{-\infty}^{\infty} d\, H(\eta)\right] + (1 + \omega^2) \left[\pi \, H^{(1)}(\omega) - i \, H^{(1)}(\omega) * \mathrm{Pv} \frac{1}{\omega}\right] \tag{6}$$

where α and H are defined in condition 2 and $\mathrm{Pv}\ 1/\omega$ is the standard pseudofunction arising from Cauchy's principal value.

Remarks: Let us make some explanations about each of these conditions.

Condition 1: A positive-real function W(p) is a function that is defined on the open right-half p plane $\{p: \mathrm{Re}\ p > 0\}$ and satisfies there the following conditions:

(i) W(p) is analytic.

(ii) W(p) is real whenever p is real.

(iii) Re $W(p) \geq 0$.

For a proof of this part of the theorem, see Zemanian [1] Chapter 10.

Condition 2: The integral in (2) must be interpreted in the generalized sense since it doesn't converge in general in the conventional sense. In particular, the second term on the right-hand side of (2) is a distribution of slow growth and the value that it assigns to any $\varphi \in S$ can be shown to be equal to

$$\int_0^\infty dt \int_{-\infty}^\infty \varphi(t) \cos \eta t \, dH(\eta) + \int_0^\infty dt \int_{-\infty}^\infty \varphi^{(2)}(t) (1 - \cos \eta t \, dH(\eta))$$

Again, see Zemanian [1], Chapter 10.

Condition 3: This representation is due to Konig and Meixner [1]. See also Konig [1] and Meixner [1].

Condition 4: This representation is proven in Konig and Zemanian [1]. By the C-order of a distribution we mean the least nonnegative integer r for which the $(r + 2)$th-order primitives of y_0 are continuous functions. Also, a distribution f is said to be nonnegative-definite if, for every testing function φ in D, $\langle f, \varphi * \check{\varphi} \rangle \geq 0$ where $\check{\varphi}(t) \triangleq \varphi(-t)$.

Condition 5: This is proven in Wohlers and Beltrami [1], p. 168.

Condition 6: This condition is also due to Beltrami and Wohlers [1], pp. 86-89. It can be viewed as a generalized Bode equation characterizing positive-real functions. These authors also show that $\tilde{y}(\omega)$ is the limit as $\sigma \to 0+$ of the Laplace transform $Y(\sigma + i\omega)$ of y in the sense of convergence in the space S'.

5. Two Other Axiomatic Approaches to One-ports: The Scattering Formulism.

We turn now to the scattering formulism for one-ports. This is obtained by defining two new variables as follows:

$$a = \tfrac{1}{2} (v + i) \tag{7}$$

$$b = \tfrac{1}{2} (v - i) \tag{8}$$

a and b can be physically interpreted as incident and reflected waves. If N satisfies P1 through P4, then $i = Nv = y * v$, and

$$a = \tfrac{1}{2} (\delta + y) * v$$

$$b = \tfrac{1}{2} (\delta - y) * v$$

If, in addition, N satisfies P5, then $\delta + y$ possesses an inverse in the convolution algebra D_R'. That is, there exists a unique member of D_R', which we denote by $(\delta + y)^{*-1}$ such that

$$(\delta + y)^{*-1} * (\delta + y) = \delta$$

Indeed, the Laplace transform of $\delta + y$ is $1 + Y(p)$, which is positive-real. Consequently, $[1 + Y(p)]^{-1}$ is also positive-real and therefore the Laplace transform of a unique member of D_R', namely, $(\delta + y)^{*-1}$.

These results show that, if N satisfies P1 through P5, then, for every $v \in D_R'$ and $i + Nv$ and for a and b given by (7) and (8), we have

$$b = s * a \tag{9}$$

where

$$s = (\delta - y) * (\delta + y)^{*-1} \tag{10}$$

Equation (10) is the scattering representation, and s is called the scattering parameter for N. Wohlers and Beltrami [1], p. 168, have pointed out that the representation (9) contains all $a \in D_R'$ in its domain because for every $a \in D_R'$ there exists a unique $v \in D_R'$ such that $2a = v + Nv$. We can establish this by again taking Laplace transforms and invoking the positive-reality of $1 + Y(p)$ as above.

Actually, we can arrive at a scattering formulism in another way and indeed obtain greater generality if we employ a different set of postulates. The postulates, which we now present, are a modified form of those suggested by Youla, Castriota, and Carlin [1] and are very similar to those used by Newcomb [1].

Q1. <u>N is a single-valued or multivalued mapping of a subset of D' into D'</u>.

We have not as yet specified the domain of N. A rather restricted domain for N is implied by the next postulate.

In contrast to postulate P1, Q1 allows N to have more than one response $i \in D'$ to any given $v \in D'$ in the domain of N. An example of a multivalued N is the ideal short circuit. Its domain contains only one voltage, the zero distribution; but, it can respond with any current in D'. The short circuit was prohibited under postulate P1. (Multivaluedness occurs more commonly among n-ports. The ideal transformer is an example of a multivalued 2-port.)

Q2. N <u>is uniquely solvable from</u> D <u>into</u> D_R' <u>with solutions in</u> D_R.

By this we mean that, given any $e \in D$, there exists a unique $v \in D_R'$ which satisfies

$$e = v + Nv$$

and that this v will be a member of D_R. The last equation signifies that a unit resistor has been connected in series with the one-port and that e is the voltage applied to the resulting series circuit; this is illustrated in Fig. 2.

Postulate Q2 implicitly specifies a rather restricted domain for N and also restricts the range values. In particular, let C(v, i) denote the set of all pairs v, i appearing as solutions of the equations:

$$e = v + Nv, \qquad i + Nv$$

as e traverses D. In symbols,

$$C(v, i) = \{v, i: \quad e = v + Nv, \quad i + Nv, \quad e \in D\} .$$

By Q2, every such v, i is a pair of elements in D_R. (When N is the ideal short circuit, v is always the zero distribution, whereas i can be any member of D because in this case $i = e \in D$.) We will assume for the moment that the domain of N is restricted to the set of v's appearing in C(v, i) and that the range of N is also restricted in accordance with C(v, i). Subsequently, the domain and range of N will be extended by means of a kernel representation for the so-called "augmented one-port."

{A remark: As is indicated by the preceding paragraph, a more natural way of viewing a multivalued one-port is not as an operator mapping voltages v into currents i, but rather as a graph (i.e., as a binary relation prescribing allowable pairs v, i). In this regard, see Newcomb [1].}

Fig. 2 and Q2 suggest that we can define a new operator on any $e \in D$ by $i = N_a e$ where $i = Nv = e - v$, $v \in D_R'$. N_a is called the augmented one-port or the augmented operator. Moreover, N_a turns out to be a single-valued mapping of D into D_R' with its range in D_R. Indeed, by Q2, given any $e \in D$, v and therefore $i \triangleq Nv = e - v$ are unique members of D_R' that are both in D_R, which verifies our assertion.

Because N_a is single-valued from D into D_R', we can define it as a single-valued operator from D into D' simply by prohibiting $N_a e$ (for any $e \in D$) from having any other values in D'. Henceforth, we adopt this convention. The physical significance of this is that we are requiring that the one-port N_a be initially at rest. For example, consider the network of Fig. 3. We have

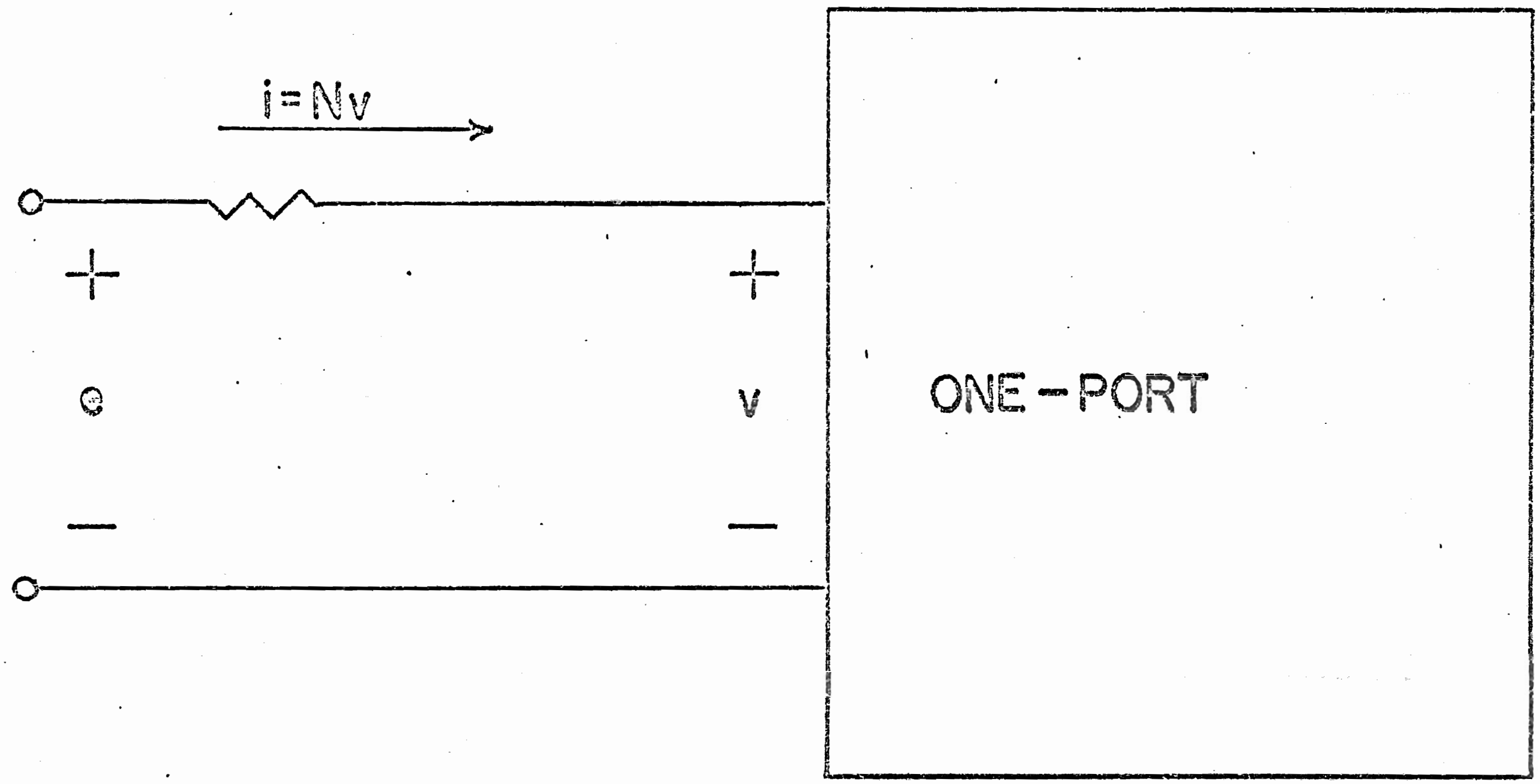

FIG. 2

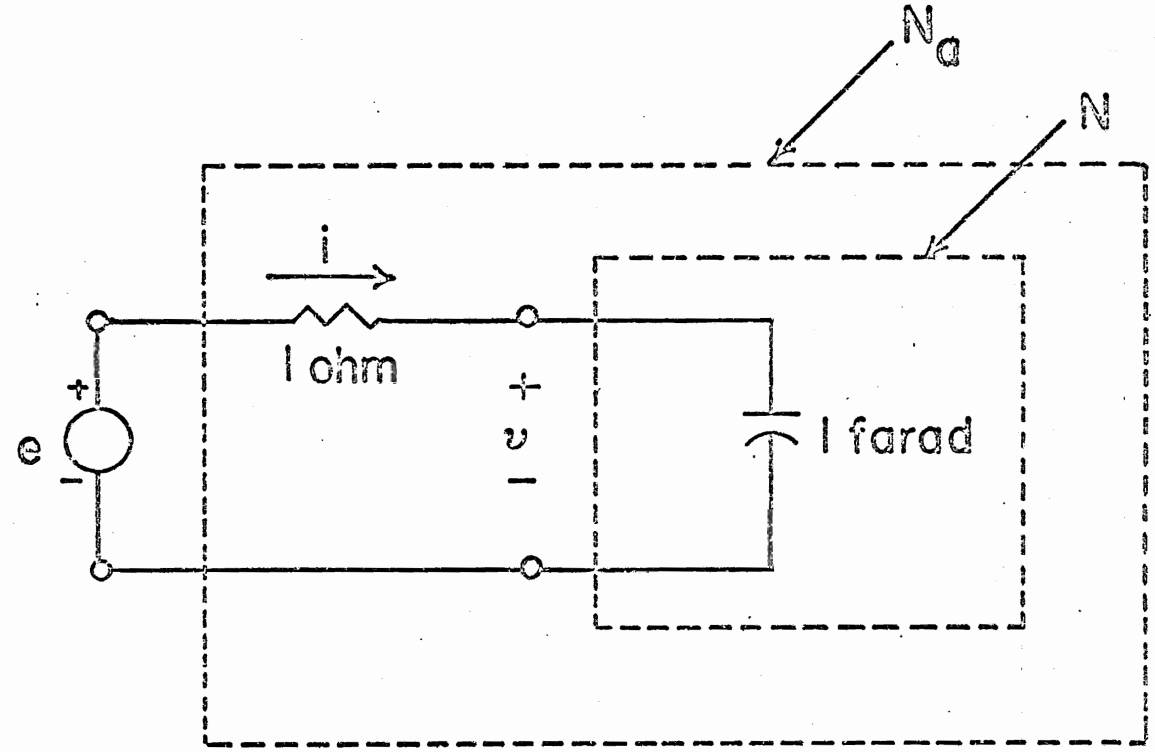

FIG. 3

that

$$e = v + i = v + v^{(1)}$$

Every free oscillation of this network is of the form:

$e = 0$, $v = ce^{-t}$, where c is any constant. It follows that the one-port N satisfies Q2 because the only solution of $0 = v + v^{(1)}$ in D'_R is the zero distribution (i.e., if $c \neq 0$, then $ce^{-t} \notin D'_R$). On the other hand, $0 = v + v^{(1)}$ has an infinity of solutions in D', namely, $v = ce^{-t}$ with c arbitrary. Hence, N_a is not single-valued as a mapping of D into D'. But, if we add the additional condition that N_a be initially at rest (so that $c = 0$), then N_a becomes a single-valued mapping of D into D'.

Q3. N <u>is</u> <u>linear</u> <u>on</u> C(v, i).

This means that, if v_1, i_1 and v_2, i_2 and two pairs in C(v, i) and if α and β are real numbers, then $\alpha v_1 + \beta v_2$, $\alpha i_1 + \beta i_2$ is another pair in C(v, i).

It readily follows from Q2 and Q3 that N_a is linear on D; that is, N_a satisfies P2, as well as P1.

Q4. N <u>is</u> <u>passive</u> <u>on</u> C(v, i).

This means that, for every pair v, i in C(v, i), (1) holds for all finite t; that is,

$$\int_{-\infty}^{t} v(x)\, i(x)\, dx \geq 0$$

A useful result can now be established: Under the preceding four postulates, N_a is a continuous mapping of D into D'. Indeed, let n be any positive integer. For any $e_n \in D$, $e_n = v_n + Nv_n$, and $i_n = Nv_n$, Q2 implies that both v_n and i_n are members of D_R. Hence, all the integrals in the following equation exist for each finite value of t.

$$\int_{-\infty}^{t} e_n^2\, dx = \int_{-\infty}^{t} v_n^2\, dx + \int_{-\infty}^{t} i_n^2\, dx + 2\int_{-\infty}^{t} v_n\, i_n\, dx \qquad (11)$$

Furthermore, the last term is nonnegative according to Q4, and so are all the others. Assume now that $e_n \to 0$ in D as $n \to \infty$. The left-hand side of (11) tends to zero as $n \to \infty$. Therefore, each term on the right-hand side does too. Next, let $\varphi \in D$ with supp $\varphi \subset [\tau, T]$, $-\infty < \tau < T < \infty$. Then,

$$|\langle i_n\ \varphi\rangle| = \left|\int_{-\infty}^{T} i_n\, \varphi\, dx\right| \leq \left[\int_{-\infty}^{T} i_n^2\, dx \int_{-\infty}^{T} \varphi^2 dx\right]^{\frac{1}{2}} \to 0 \qquad n \to \infty$$

which verifies that N_a satisfies P3.

Moreover, when N satisfies Q1 through Q4, N_a also satisfies P5 since for any $e \in D$, we have

$$\int_{-\infty}^{t} ei\ dx = \int_{-\infty}^{t} vi\ dx + \int_{-\infty}^{t} i^2 dx \geq 0 .$$

We have just seen that the augmentation technique indicated in Fig. 2 allows us to start with a one-port N which is neither single-valued nor continuous and to obtain from it a one-port N_a which is. Moreover, even though N may not contain all of D in its domain, N_a will. Furthermore, we have

Theorem 6: If N satisfies Q1 through Q4, then N_a satisfies P1, P2, P3, and P5. Moreover, N_a is causal and has a kernel representation $N_a = y_a \bullet$ where $y_a = y_a(t, \tau) = D'|_{t, \tau}$.

Proof: We have already established the first statement; we remind the reader that the satisfaction of P1 by N_a is a consequence of our convention concerning the single-valued definition of the range of N_a in D'. The second statement follows immediately from theorems 1 and 4.

y_a is called the augmented admittance of N.

As was done in the admittance formulation, we can now extend N_a via the kernel representation $N_a = y_a \bullet$ onto every space of distributions A' that satisfies the conditions stated in the preceding section. This extension is unique in the aforementioned sense, and we henceforth assume that it has been made. This automatically extends the operator N, under the constraint that it has been augmented as in Fig. 2, into a mapping from a subset of D' into D' defined by

$$i = N\ (e - i)$$

where $i = N_a e = y_a \bullet e$.

We add one more postulate:

Q5. N is time-invariant on C(v, i).

This means that, if v, i is a pair in C(v, i) and if σ_x is the shifting operator as before, then $\sigma_x v$, $\sigma_x i$ is also a pair in C(v, i), whatever be the value of x. Since i = Nv, this implies that $\sigma_x Nv = N \sigma_x v$ for every v, Nv in C(v, i).

If N satisfies Q1, Q2, and Q5, then N_a satisfies P4 (i.e., N_a is time-invariant on D). To show this, we start from $e = v + Nv$, $e \in D$. Thus, $\sigma_x e = \sigma_x v + \sigma_x Nv = \sigma_x v + N \sigma_x v$, and this decomposition is unique by Q2.

Therefore, for all $e \in D$,

$$\sigma_x N_a e = \sigma_x Nv = N \sigma_x v = N_a \sigma_x e$$

which verifies our assertion.

Under the additional postulate Q5, theorem 6 becomes

Theorem 7: If N satisfies Q1 through Q5, then N_a satisfies P1 through P5. Moreover, N_a is causal and has a convolution representation $N_a = y_a *$ where $y_a = y_a(t) \in D'|_t$.

We are at last ready to derive the scattering formulism for a one-port N that satisfies Q1 through Q4. We have that

$$i = y_a \bullet e = y_a \bullet (v + i)$$

Therefore,

$$y_a \bullet v = (\delta - y_a) \bullet i$$

[Here, δ denotes the kernel $\delta(t - \tau)$, so that $\delta \bullet i = i$.]
Now, substitute the quantities: $v = a + b$, $i = a - b$. This yields

$$b = s \bullet a \tag{12}$$

where

$$a = \tfrac{1}{2}(v + i) \ , \ b = \tfrac{1}{2}(v - i),$$

and s is the kernel

$$s = \delta - 2\, y_a \tag{13}$$

s is the time-variable scattering parameter for N.

Theorem 8: If N satisfies Q1 through Q4, then N has the scattering formulism (12) and (13), where y_a is the augmented admittance indicated in theorem 6. Moreover, supp $s(t, \tau) \subset \{(t, \tau): t \geq \tau\}$. If in addition N satisfies Q5, then

$$b = s * a$$

where $s = s(t) \in D'|_t$ and supp $s \subset [0, \infty)$. (14)

Proof: Invoke theorems 3, 6, and 7.

The restriction on the support of s implies that D'_R is contained in the domain of the operator $s\bullet$.

As one of the major conclusions of this section, we have that the P postulates imply the Q postulates. Indeed, assume that N satisfies P1 through P5. Then, N obviously satisfies Q1. That it satisfies Q2 can be established through an argument due to Wohlers and Beltrami [1], p. 168. Q3 and Q5 are also clearly satisfied since $i = Nv = y * v$ where $y \in D'_+$. This also shows that N is a causal mapping of D_R into D_R, and it now follows from lemma 1 and P5 that N satisfies Q4.

Let us list some of the properties of the operator M = s * assuming that N satisfies Q1 through Q5. In view of theorem 8 and the properties of distributional convolution, we have the following:

S1: M is a single-valued mapping of D into D′.

S2: M is linear on D.

S3: M is continuous from D into D′.

S4: M is time-invariant on D.

S5: M is causal on D.

S6: M is weakly passive on D

Property S5 means that the causality property defined in the preceding section holds whenever the elements a in the domain of M are restricted to D.

Property S6 means that, given any a ∈ D, the integral:

$$\int_{-\infty}^{\infty} vi \; dt = \int_{-\infty}^{\infty} (a^2 - b^2) \; dt \tag{15}$$

exists and is nonnegative. That M = s * truly satisfies S6 can be shown as follows. For any a ∈ D, we have that $b = Ma = s * a \in D_R$ since $s \in D'_+$. Hence, $v = a + b \in D_R$ and $i = a - b \in D_R$. Also, v, i is a pair in C(v, i) since 2a = v + i. By Q4 and the fact that $vi = a^2 - b^2$, we have that for all finite t

$$\int_{-\infty}^{t} a^2 \; dx \geq \int_{-\infty}^{t} b^2 \; dx \geq 0.$$

The left-hand side converges as $t \to \infty$, and therefore the middle term does too. Hence, (15) exists and is nonnegative.-

The concept of weak passivity was used by Raisbeck [1].

Wohlers and Beltrami [1] have shown that the theory of one-ports can be derived by using the S properties as postulates on the operator M that relates the reflected wave b to the incident wave a. Indeed, S1 through S4 read precisely the same as P1 through P4. Therefore, S1 through S3 imply that M has a kernel representation: M = s∘, and S1 through S4 imply that M has a convolution representation: M = s *. The additional postulate S5 and the argument in the first paragraph of the proof of theorem 3 show that s satisfies the same restrictions on its support as does y in that theorem. We can now extend M onto larger spaces of distributions (for example, onto D'_R) via these representations and this extension must be unique as explained before.

Theorem 3 and its proof show that, if M satisfies S1 through S4, and is also passive in the sense that

$$\int_{-\infty}^{t} (a^2 - b^2)\, dx \geq 0 \qquad a \in D,\ -\infty < t < \infty \tag{16}$$

then M is causal. It is also weakly passive on D. Conversely, Wohlers and Beltrami [1], p. 167, have shown that S1 through S6 imply that M is passive in the sense of (16). Thus, assuming that S1 through S4 are satisfied, we can conclude that postulates S5 and S6 are equivalent to the single assumption that M is passive in the sense of (16) However, using S5 and S6 is preferable since causality and weak passivity are independent assumptions. Indeed, that S5 does not imply S6 follows from the example $s = 2\delta$. That S6 does not imply S5 follows from the example $s(t) = \delta(t + c)$ where c is a positive number (see Wohlers and Beltrami [1], p. 165).

By itself, postulate S6 allows one-ports that up to finite instances of time have emitted more energy than they have received, but will ultimately (i.e., as $t \to \infty$) absorb at least as much energy as they have emitted. Such one-ports are prohibited by either P5 or Q4.

It is worth noting that, when N satisfies P1 through P4, the nonnegativity of (15) for all $v \in D$ and the causality of N do not imply the passivity of N according to P5. In this regard consider the example, $N = -\delta^{(1)}*$.

It may also be argued, especially when the one-port is a part of a wave propagation system, that the basic physical variables are the incident wave a and the reflected wave b. If this point of view is accepted, then the axioms that ought to be used are the S postulates and not the P or Q postulates.

Operators M that satisfy the S postulates can be characterized as follows:

<u>Theorem 9</u>: <u>If</u> M <u>satisfies postulates</u> S1 <u>through</u> S6, <u>then</u> $S \triangleq M\,\delta$ (<u>as always, the extended</u> M <u>is understood here</u>) <u>satisfies the following three equivalent conditions</u>. <u>Conversely, if</u> $s \in D'$ <u>satisfies any one of the following conditions, then</u> $M \triangleq s\,*$ <u>satisfies</u> S1 <u>through</u> S6.

1. <u>The Laplace transform</u> S(p) <u>of</u> s <u>is a bounded-real function</u>. [S(p) is called bounded-real if on the open right-half p plane {p: Re p > 0} we have that S(p) is analytic, S(p) is real whenever p is real, and $|S(p)| \leq 1$.]

2. $s \in D'_+ \cap D'_{L_2}$, and, for every $\varphi \in D$,

$$\langle \delta - s * \check{s}, \varphi * \check{\varphi} \rangle \geq 0$$

where $\check{f}(t) \triangleq f(-t)$. [For a definition of D'_{L_2}, see Schwartz, vol. II, pp.55-56.]

3. Let $\tilde{s}$ denote the Fourier transform of s. Then, $\tilde{s}(\omega)$ is a conventional function such that $|\tilde{s}(\omega)| \leq 1$ almost everywhere on $-\infty < \omega < \infty$, $\tilde{s}(-\omega)$ is equal to the complex conjugate of $\tilde{s}(\omega)$, and

$$\tilde{s}^{(i)}(\omega) = \frac{1}{i\pi} s^{(i)}(\omega) * Pv \frac{1}{\omega}$$

where Pv $1/\omega$ is again the standard pseudofunction arising from Cauchy's principal value.

This theorem is proven in Beltrami and Wohlers [1], pp. 89-93, and Beltrami [1]. The third condition presents another example of a generalized Bode equation.

We have seen that the P postulates imply the Q postulates which in turn imply the S postulates. On the other hand, we cannot go from the Q postulates to the P postulates, the ideal short circuit being an example that verifies this assertion. However, condition 1 of theorem 9, and a theorem of Youla, Castriota, and Carlin [1], pp. 116-117, assert that thd S postulates imply their form of the Q postulates.

Under the present formulation we can get from the S postulates to the Q postulates in the following way. Under the S postulates we have $b = s * a$ or $v - i = s * (v + i)$ whenever $v + i \in D'_R$. This can be rearranged into

$$(\delta - s) * v = (\delta + s) * i \tag{17}$$

To verify Q1, we shall show that there exists at least one pair of distributions v, i that satisfies (17). Two cases arise.

(i) $S(p) \not\equiv -1$ for Re $p > 0$. Then, by the bounded reality of S(p) and the maximum-modulus theorem, $1 + S(p) \neq 0$ at every point in the right-half p plane, and

$$\frac{1 - S(p)}{1 + S(p)} \tag{18}$$

is a positive-real function. Let $y \in D'_R$ be its inverse Laplace transform. Then, for all $v \in D'_R$, we have that $i = y * v \in D'_R$ satisfies (17). Indeed, upon substituting this i into (17), we get

$$(\delta - s) * v = (\delta + s) * y * v.$$

Since D'_R is a convolution algebra with no divisors of zero, the last equation can hold for all $v \in D'_R$ if and only if $\delta - s = (\delta + s) * y$, which is seen to be true by taking Laplace transforms again.

(ii) $S(p) \equiv -1$ for Re $p > 0$. Thus, $\delta + s$ is the zero member 0 of D'_R and hence of D', and (17) becomes $2\delta * v = 0 * i$. Therefore, v must be the zero member of D' whereas i can be any member of D'. This degenerate case corresponds to the ideal short circuit.

In both cases, we have obtained a one-port N that satisfies Q1. Moreover, in case (i) we see from theorem 5 (condition 1) and the fact that

$$\frac{1 - S(p)}{1 + S(p)}$$

is positive-real whenever $S(p)$ is bounded-real that N satisfies the P postulates and therefore the Q postulates. In the degenerate case (ii) it is obvious that N satisfies the Q postulates.

We can also conclude at this time that the only possible multivalued one-port satisfying the Q postulates is the short-circuit. (On the impedance basis, it would be the open-circuit.)

In the case of n-ports ($n > 1$) however, there are many operators exhibiting multivaluedness. When the n-port is treated on an admittance basis as above, it is multivalued when and only when the n x n matrix $1_n + S(p)$ is singular everywhere in the right-half plane Re $p > 0$. Here, 1_n denotes the n x n unit matrix. On the impedance basis, the n-port is multivalued when and only when the n x n matrix $1_n - S(p)$ is singular everywhere in the half-plane Re $p > 0$.

6. A Frequency-domain Characterization for the Causality of Active One-ports.

If a one-port N (or more generally any input-output system) satisfies postulates P1 through P4, then, according to theorem 2, $N = y *$ for some $y \in D'|_t$. Under these conditions a criterion for causality is provided by orem 3, which states that N is causal if and only if supp $y \subset [0, \infty)$. es there exist a corresponding frequency-domain characterization of causality?

Under the additional assumption that y is Laplace-transformable in Schwartz's sense (Schwartz [4]), we can apply a well-known criterion (see, for example, Zemanian [1], Sec. 8.4) to obtain

<u>Theorem 10: Let N satisfy P1 through P4, let $y = N\delta$ be Laplace-transformable in Schwartz's sense, and let $Y(p)$ be the Laplace transform of y. Then, necessary and sufficient conditions for N to be causal are that there be a half-plane Re $p > a$ on which $Y(p)$ is analytic and satisfies</u>

$$|Y(p)| \leq P(|p|) \tag{19}$$

<u>where P is some polynomial.</u>

However, this result is too restrictive since there exist active systems for which y is not Laplace-transformable in Schwartz's sense. For example, this is the case when $y(t) = 1_+(t) \exp t^2$. (See, also, Zemanian [6].) A criterion is needed that is applicable when y is any distribution. This can be obtained by modifying a recent result due to Zielezny [1] and combining it with theorem 3 above.

We again refer the reader to Zemanian [1] for more information concerning the symbols used here. Our notation differs somewhat from that of Zielezny, as does our definition of the Fourier transform $\tilde{f}$ of any $f \in D'$, which is

$$\langle \tilde{f}, \varphi \rangle \triangleq \langle f, \tilde{\varphi} \rangle$$

where $\varphi \in Z$, $f \in D'$, and $\tilde{\varphi} = F\varphi \in D$. F denotes the Fourier transformation and F^{-1} its inverse. For the testing functions we have

$$\tilde{\varphi}(t) = F\varphi = \int_{-\infty}^{\infty} \varphi(u)\, e^{-iut}\, du$$

$$\varphi(u) = F^{-1}\tilde{\varphi} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi(t)\, e^{iut}\, dt$$

Z is the space of inverse Fourier transforms of testing functions in D and Z_+ is that subspace of Z whose members have the form $F^{-1}\varphi$ where $\varphi \in D$ and supp $\varphi \subset (0, \infty)$. Also, Z' is the space of Fourier transforms of the members of D'. In this section, u, v, and z denote complex variables. We specify that Z' is a space of generalized functions on either the one-dimensional or two-dimensional complex euclidean spaces by using the notation $Z'|_u$ or $Z'|_{u,v}$ respectively. $Z|_u$ and $Z|_{u,v}$ carry similar meanings.

If $\tilde{f}, \tilde{g} \in Z'|_u$, the direct product $\tilde{f}(u) \odot \tilde{g}(v)$ is defined as a member of $Z'|_{u,v}$ by

$$\langle \tilde{f}(u) \odot \tilde{g}(v), \varphi(u)\psi(v) \rangle \triangleq \langle \tilde{f}, \varphi \rangle \langle \tilde{g}, \psi \rangle \tag{20}$$

where $\varphi, \psi \in Z|_u$. That (20) suffices as a definition follows from the fact that testing functions of the form $\varphi(u)\psi(v)$ span a dense subspace of $Z|_{u,v}$. $\tilde{f}(u+v) \odot \tilde{g}(v)$ is a member of $Z|_{u,v}$ obtained from $\tilde{f}(u) \odot \tilde{g}(v)$ by the substitutions: $u \mapsto u + v$, $v \mapsto v$. Zielezny showed that

$$\tilde{f}(u + v) \otimes \tilde{g}(v) = F[f(t) \odot g(\tau - t)] \qquad (21)$$

where now F denotes the two-dimensional Fourier transformation and $f(t) \odot g(\tau - t)$ is that member of $D'|_{t, \tau}$ obtained from the customary direct product $f(t) \odot g(\tau)$ by making the substitutions: $t \mapsto t$, $\tau \mapsto \tau - t$.

If $h(u, v) \in Z'|_{u, v}$ and $\psi(v) \in Z|_v$, then the kernel representation $h(u, v) \circ \psi(v)$ is a member of $Z'|_u$ defined by

$$\langle h(u, v) \circ \psi(v), \varphi(u)\rangle \triangleq \langle h(u, v), \varphi(u)\, \psi(v)\rangle \qquad \varphi \in Z|_u \qquad (22)$$

We can characterize the members of Z_+ as follows:

Lemma 2: $\psi \in Z_+$ [i.e., $\psi = F^{-1}\tilde{\psi}$ where $\tilde{\psi} \in D$ and supp $\tilde{\psi} \subset (0, \infty)$] if and only if ψ is an entire function and there exist two real numbers a and b with $0 < a < b < \infty$ and a sequence of positive numbers $\{C_k\}_{k=0}^{\infty}$ such that for each k

$$|v^k \psi(v)| \leq C_k e^{-\text{Im}\, v\, a} \qquad \text{Im}\, v \geq 0$$

and

$$|v^k \psi(v)| \leq C_k e^{-\text{Im}\, v\, b} \qquad \text{Im}\, v \leq 0$$

Also, we let $F E'_+$ denote the space of Fourier transforms of those $f \in E'$ satisfying supp $f \in [0, \infty)$; that is, $\tilde{f} \in F E'_+$ if and only if $\tilde{f} = F f$ and $f \in D'_+ \cap E'$.

Lemma 3: $\tilde{f} \in F E'_+$ if and only if $\tilde{f}$ is an entire function satisfying the following two conditions:

(i) There exists a polynomial P such that

$$|\tilde{f}(u)| \leq P(|u|) \qquad \text{Im}\, u \leq 0$$

(ii) There exists a real number $T > 0$ and a polynomial Q such that

$$|\tilde{f}(u)| \leq Q(|u|)\, e^{\text{Im}\, u\, T} \qquad \text{Im}\, u \geq 0$$

Proofs for lemmas 2 and 3 can be extracted from Zemanian [1], Secs. 7.6 and 8.4.

Now, the special form of Zielezny's result that leads to the criterion we seek is given by

Theorem 11: Let $y \in D'$. Also, let $\psi \in Z_+$ have no zero at the origin (i.e., $\psi(0) \neq 0$]. Then, supp $y \subset [0, \infty)$ if and only if $\tilde{y} \triangleq Fy$ is such that

$$[\tilde{y}(u + v) \circledcirc \tilde{1}_+(v)] \circ \psi(v) \in F\, E'_+ \tag{23}$$

Here, $\tilde{1}_+$ is the Fourier transform of 1_+ and is defined as a functional on Z by

$$\langle \tilde{1}_+, \theta \rangle = \int_{-\infty - ci}^{\infty - ci} \frac{\theta(z)}{iz}\, dz$$

where c is any real positive number and $\theta \in Z$. Note that (23) is an assertion concerning spaces Z_+ and $F\, E'_+$ of entire (and not generalized) functions. The idea behind Zielezny's proof of theorem 11 is that $f \in D'_+$ if and only if $f \in D'$ and $\lambda f \in D'_+ \cap E'$ where $\lambda(t)$ is some fixed smooth function that is identically equal to a nonzero constant on a neighborhood of $(-\infty, 0]$ and identically equal to zero for all sufficiently large t.

We combine theorems 3 and 11 to obtain the principal result of this section:

Theorem 12: Let N satisfy P1 through P4, and let $y = N\delta$. Then, N is causal if and only if $\tilde{y} \triangleq F y$ satisfies (23) for one $\psi(v)$ in Z_+ that has no zero at $v = 0$.

Dr. A. Csurgay has pointed out to the author that for certain studies it is important to know whether $t = 0$ is the smallest essential point of $y \triangleq N\delta$. That is, is it true that inf supp $y = 0$? This matter is resolved by the following corollaries.

Corollary 10a: Assume that the conditions of theorem 10 are satisfied. Then, inf supp $y = 0$ if and only if for each $\tau = 0$ there does not exist any polynomial M for which Y(p) is bounded according to

$$|Y(p)| \le M(|p|)\, e^{-\mathrm{Re}\, p\, \tau}$$

on some half-plane $\mathrm{Re}\, p > b$.

Corollary 11a: Assume that the conditions of theorem 11 are satisfied. Then, inf supp $y = 0$ if and only if for each $\tau > 0$ there does not exist a polynomial M for which

$$|[\tilde{y}(u + v) \circledcirc \tilde{1}_+(v)] \bullet \psi(v)| \le M(|u|)\, e^{\mathrm{Im}\, u\, \tau} \qquad \mathrm{Im}\, u \le 0.$$

References

Anderson, B. D. O., and R. W. Newcomb

[1] "Functional Analysis of Linear Passive Networks," Internat. J. Engr. Sci., (to appear).

Arsac, J.

[1] "Fourier Transforms and the Theory of Distributions," Prentice-Hall, Englewood Cliffs, New Jersey, 1966.

Aseltine, J. A.

[1] "Transforms for Linear Time-varying Systems," Dept. of Elect. Engr., University of California at Los Angeles, Rep. No. 52-1, 1951.

Beltrami, E. J.

[1] "Linear Dissipative Systems, Nonnegative Definite Distributional Kernels, and the Boundary Values of Bounded-real and Positive-real Matrices," J. Math. Anal. and Appl., vol. 19, pp. 231-246, 1967.

Beltrami, E. J., and M. R. Wohlers

[1] "Distributions and the Boundary Values of Analytic Functions," Academic Press, New York, 1966.

[2] "Distributional Boundary Value Theorems and Hilbert Transforms," Arch. Ratl. Mech. and Analysis, vol. 18, pp. 304-309, 1965.

Bremermann, H.

[1] "Distribution, Complex Variables, and Fourier Transforms," Addison-Wesley Publishing Co., Reading, Massachusetts, 1965.

de Jager, E. M.

[1] "Applications of Distributions in Mathematical Physics," Mathematical Centre Tracts 10, Mathematisch Centrum, Amsterdam, 1964.

Dolezal, V.

[1] "Dynamics of Linear Systems," Publishing House of the Czechoslovak Academy of Sciences, Prague, 1964.

[2] "Über eine Klasse Linearer Operatoren," Časopis pro pěstování matematiky, vol. 86, pp. 200-232, 1961.

[3] "On Linear Passive n-ports with Time-varying Elements, J. SIAM Appl. Math., vol. 15, pp. 1018-1029, 1967.

Friedman, A.

[1] "Generalized Functions and Partial Differential Equations," Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

Gelfand, I. M., and G. E. Shilov

[1] "Generalized Functions," Academic Press, vol. 3, New York, 1967.

Gerardi, F. R.

[1] "Applications of Mellin and Hankel Transforms to Networks with Time-varying Parameters," I. R. E. Trans. on Circuit Theory, vol. CT-6, pp. 197-208, 1959.

Guttinger, W.

[1] "Generalized Functions in Elementary Particle Physics and Passive System Theory: Recent Trends and Problems," J. SIAM Appl. Math., vol. 15, pp. 964-1000, 1967.

Hörmander, L.

[1] "Linear Partial Differential Operators," Springer-Verlag, Berlin, 1963.

König, H.

[1] "Lineare Dissipative Systeme," Colloque sur L'Analyse Fonctionnelle, Centre Belge de Rechercher Mathématiques, pp. 71-79, Belguim, 1961.

König, H., and A. H. Zemanian

[1] "Necessary and Sufficient Conditions for a Matrix Distribution to have a Positive-real Laplace Transform," J. Soc. Indust. Appl. Math., vol. 13, pp. 1036-1040, 1965.

König, H., and J. Meixner

[1] "Lineare Systeme und Lineare Transformationen," Math. Nach., vol. 19, pp. 265-322, 1958.

Lauwerier, H. A.

[1] "The Hilbert Problem for Generalized Functions," Arch. Ratl. Mech. and Anal., vol. 13, pp. 157-166, 1963.

Liverman, T. P. G.

[1] "Physically Motivated Definitions of Distributions," J. SIAM Appl. Math., vol. 15, pp. 1048-1076, 1967.

Meixner, J.

[1] "On the Theory of Linear Passive Systems," Arch. Ratl. Mech. and Anal., vol. 17, pp. 278-296, 1964.

Newcomb, R. W.

[1] "Linear Multiport Synthesis," McGraw-Hill Book Co., New York, 1966.

[2] "The Foundations of Network Theory," The Institute of Engineers, Electrical and Mechanical Trans., vol. EM-6, pp. 7-12, Australia, 1964.

Newcomb, R. W., and B. D. O. Anderson

[1] "A Distributional Approach to Time-varying Sensitivity," J. SIAM Appl. Math., vol. 15, pp. 1001-1010, 1967.

Raisbeck, G.

[1] "A Definition of Passive Linear Networks in Terms of Time and Energy," J. Appl. Phys., vol. 25, pp. 1510-1514, 1954.

Schwartz, L.

[1] "Théorie des Distributions," vols. I, II, Hermann, Paris; 1957 and 1959.

[2] "Théorie des Noyaux," Proc. Int. Congress of Mathematicians Cambridge, Massachusetts, pp. 220-230, 1950.

[3] "Causalite et Analyticité," Anais da Academia Brasiliera de Ciencias, vol. 34, pp. 13-21, 1962.

[4] "Transformations de Laplace des Distributions," Seminaire Math., Université Lund, tome supplémentaire dedié à M. Riesz, pp. 196-206, 1952.

Spaulding, D. A., and R. W. Newcomb

[1] "The Time-variable Scattering Matrix," Proc. I. E. E. E., vol. 53, pp. 651-652, 1965.

Streater, R. F., and A. S. Wightman

[1] "PCT, Spin and Statistics, and All That," W. A. Benjamin, New York, 1964.

Taylor, J. G.

[1] "Dispersion Relations and Schwartz's Distributions," Ann. of Phys., vol. 5, pp. 391-398, 1958.

Newcomb, R. W., and B. D. O. Anderson

[1] "A Distributional Approach to Time-varying Sensitivity," J. SIAM Appl. Math., vol. 15, pp. 1001-1010, 1967.

Raisbeck, G.

[1] "A Definition of Passive Linear Networks in Terms of Time and Energy," J. Appl. Phys., vol. 25, pp. 1510-1514, 1954.

Schwartz, L.

[1] "Théorie des Distributions," vols. I, II, Hermann, Paris; 1957 and 1959.

[2] "Théorie des Noyaux," Proc. Int. Congress of Mathematicians Cambridge, Massachusetts, pp. 220-230, 1950.

[3] "Causalite et Analyticité," Anais da Academia Brasiliera de Ciencias, vol. 34, pp. 13-21, 1962.

[4] "Transformations de Laplace des Distributions," Seminaire Math., Université Lund, tome supplémentaire dedié à M. Riesz, pp. 196-206, 1952.

Spaulding, D. A., and R. W. Newcomb

[1] "The Time-variable Scattering Matrix," Proc. I. E. E. E., vol. 53, pp. 651-652, 1965.

Streater, R. F., and A. S. Wightman

[1] "PCT, Spin and Statistics, and All That," W. A. Benjamin, New York, 1964.

Taylor, J. G.

[1] "Dispersion Relations and Schwartz's Distributions," Ann. of Phys., vol. 5, pp. 391-398, 1958.

Tillmann, H. G.

[1] "Randverteilungen Analytischer Funktionen und Distributionen," Math. Zeitschr., vol. 59, pp. 61-83, 1953.

Vladimirov, V. S.

[1] "Methods of the Theory of Functions of Several Complex Variables," M. I. T. Press, Cambridge, Massachusetts, 1966.

Wohlers, M. R.

[1] "A Distributional Study of the Real Frequency Behavior of Passive Systems," Proc. Allerton Conference on Circuit and System Theory, October 20-22, 1965, University of Illinois, Urbana, Illinois.

Wohlers, M. R., and E. J. Beltrami

[1] "Distribution Theory as the Basis of Generalized Passive-Network Analysis," I. E. E. E. Trans. on Circuit Theory, vol. CT-12, pp. 164-170, 1965.

Youla, D. C.

[1] "Physical Realizability Criteria," I. R. E. International Convention Record, Part 2, pp. 181-192, 1960.

Youla, D. C., L. J. Castriota, and H. J. Carlin

[1] "Bounded Real Scattering Matrices and the Foundations of Linear Passive Network Theory," I. R. E. Trans. on Circuit Theory, vol. CT-6, pp. 102-124, 1959.

Zemanian, A. H.

[1] "Distribution Theory and Transform Analysis," McGraw-Hill Book Co., New York, 1965.

[2] "An n-port Realizability Theory Based on the Theory of Distributions," I. E. E. E. Trans. on Circuit Theory, vol. CT-10, pp. 265-274, 1963.

[3] "Generalized Integral Transformations," (to appear).

[4] "Orthonornal Series Expansions of Certain Distributions and Distributional Transform Calculus," J. Math. Anal. Appl., vol. 14, pp. 263-275, 1966.

[5] "The Time-domain Synthesis of Distributions," I. E. E. E. Trans. on Circuit Theory, vol. CT-11, pp. 487-493, 1964.

[6] "Paper Networks," I. E. E. E. Trans. on Circuit Theory, vol. CT-12, pp. 425-426, 1965; and vol. CT-13, pp. 110-111, 1966.

Zieleźny, Z.

[1] "On Fourier Transforms of Distributions with One-Sided Bounded Carriers," Studia Mathematica, vol. 25, pp. 73-80, 1964.