The Characteristic Functions of Operator Theory and Electrical Network Realization

J. WILLIAM HELTON

One of the main results in Nagy and Foias [N-F] says that any analytic function $\theta(z)$ which is defined on the unit disk and whose values are operators of norm less than or equal to one on a Hilbert space coincides with the characteristic function of some operator T on another Hilbert space $\mathbf{H}$ (Theorem VI, 3.1 [N-F]). Amazingly enough a very special case of this abstract theorem was developed independently by engineers and is one formulation of the main realizability theorem in electrical engineering. The purpose of this note is to sketch very roughly how this occurs.

The fact that a connection exists between modern operator theory and electrical engineering has been known for several years. In his book [L] Livsic worked out the connection between his approach to characteristic functions for operators and network realizability theory. Unfortunately, this book has never been translated into English and it has circulated little among American operator theorists. Also, DeWilde in [D] relates mathematics developed by Livsic and Potopov to network synthesis. There has been work by Sandburg, Sacks and Levan in a very different direction involving the original Halmos dilation for a contraction operator and the synthesis of a lossy network with a lossless network and a resistor.

This paper takes the Lax-Phillips [L-P] and Nagy-Foias viewpoint to operator theory rather than that of Livsic and the interpretation given here to the Nagy-Foias representation theorem appears to be new.

The first section is an exposition intended for operator theorists of some basic electric network theory, in the second part we relate this to operator theory. For a description of electric network theory see [W] or [N].

The author wishes to thank A. H. Zemanian for his encouragement and for discussions which were most helpful.

A brief exposition of electrical network theory. An electric network (n port) is thought of as a box with n pairs of terminals on it. For simplicity sake consider a network $\mathcal{N}$ which is a 1 port. If one connects a voltage source with voltage v across the terminals of $\mathcal{N}$, then a unique current i flows through the circuit (see Figure 1). Voltage and current are

FIGURE 1

both functions $v(t)$ and $i(t)$ of time and are assumed to be complex valued functions in $L^2[-\infty, +\infty]$. The set $\mathfrak{N}$ of all pairs $[v(t), i(t)]$ possible for a network $\mathfrak{N}$ uniquely determines $\mathfrak{N}$. We shall call $\mathfrak{N}$ the graph of the function $\mathfrak{N}$.

By comparing the network $\mathfrak{N}$ to the network consisting of a single resistor of resistance r electrical engineers arrive at an interpretation of $\frac{1}{2}[r^{-1/2}v + r^{1/2}i]$ and $\frac{1}{2}[r^{-1/2}v - r^{1/2}i]$ as incoming and outgoing signals respectively. They usually take the reference resistance r to be 1 and define a “scattering operator” S for the network $\mathfrak{N}$ by

$$S\frac{1}{2}(v + i) = \frac{1}{2}(v - i).$$

Every network which we consider here will have a “scattering operator” S which is a bounded linear map of $L^2[-\infty, +\infty]$ into $L^2[-\infty, +\infty]$. The study of scattering operators for networks is equivalent to the study of networks, at least from a “black box” point of view.

All networks which we consider will be time invariant and they will have the property that no signal comes out before time t_0 unless a signal has been put in before time t_0 . The second property is called *causality*. Mathematically, time invariance means that S commutes with translation by time, and causality means that S maps the subspace $L^2[t_0, \infty]$ of $L^2[-\infty, \infty]$ into itself. If one Fourier transform functions in $L^2[-\infty, +\infty]$ the operator induced by S on the Fourier transform space $L^2[-\infty, +\infty]$ commutes with multiplication operators and consequently is itself multiplication by a function $S(\sigma)$. Moreover, $S(\sigma)$ is the boundary value of a function $S(p)$ defined and analytic in the upper half plane. However, engineers use, and we shall use, the convention that $S(p)$ is analytic in the right half plane and that the “Fourier transform” $S(\sigma)$ of S is a function defined on the imaginary axis.

Example. Suppose that $\mathfrak{N}$ consists of a single inductor. Then by definition $v(t) = \ell(d/dt)i(t)$ where ℓ is the inductance of the inductor. Thus, the scattering operator S maps

$$\left(\ell \frac{d}{dt} + I\right)i(t) \xrightarrow{s} \left(\ell \frac{d}{dt} - I\right)i(t)$$

and consequently the scattering matrix for $\mathfrak{N}$ is

$$S(p) = (\ell p - 1)(\ell p + 1)^{-1}$$

The total energy dissipated by any network $\mathfrak{N}$ at a state $[v(t), i(t)] = a$ is

$$\mathcal{O}(a) = \operatorname{Re} \int_{-\infty}^{+\infty} v(t)i(t) dt = \frac{1}{4} \int_{-\infty}^{+\infty} \left(|v(t) + i(t)|^2 - |v(t) - i(t)|^2\right) dt.$$

This is an indefinite quadratic form on the space $\mathcal{H}$ of states. Every state n in the graph of *lossless network* $\mathfrak{N}$ has 0 energy dissipated; $\mathcal{O}(n) = 0$. Every state n in the graph of a *passive network* has non-negative energy dissipation; $\mathcal{O}(n) \geq 0$. A network is *active* if it is not passive.

Remark 1. Mathematicians began to consider Hilbert spaces with an indefinite inner product (having infinite dimensional negative part) about ten years ago. For a survey of the subject see [K]. The example given here of an indefinite inner product seems most natural. In spirit it is a bit like the one which motivated Phillip's first paper on the subject [P].

Remark 2. Actually, what is defined here is usually called a *semi-passive* network. However, the notions of *passive* and *semi-passive* coincide for causal, time invariant networks (cf. [Z]).

Now we discuss n -ports (see Figure 2). It is a simple variation on the above analysis. One measures the voltage $v_i(t)$ across the j^{th} port and the current $i_{i1}(t)$ and $i_{i2}(t)$ at the first and second terminals respectively of the j^{th} port. Engineers usually assume that $i_{i1}(t) = i_{i2}(t) = i_i(t)$ and then they work with the vector valued voltage

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{bmatrix}$$

and vector valued current

$$\mathbf{i}(t) = \begin{bmatrix} i_1(t) \\ \vdots \\ i_n(t) \end{bmatrix}.$$

Everything that has been said generalizes in an obvious way to vector valued voltages and currents. The assumption $i_{i1}(t) = i_{i2}(t)$ essentially puts a re-

striction on the type of interconnections between networks which is allowed. The preceding exposition motivates the following definitions.

Definitions. An $n \times n$ scattering matrix $S(p)$ is an $n \times n$ matrix valued function analytic on the right half plane. $S(p)$ is called *contractive* if $\|S(p)\| \leq 1$ for p in the right half plane and *lossless* if $S(p)$ is unitary when p is pure imaginary. A scattering matrix is called *rational* if $S(p)$ is a rational function of p .

It follows directly from the definitions of S and $\mathcal{O}$ above that a network $\mathcal{N}$ is passive if and only if its scattering operator S has $\|S\| \leq 1$. This is equivalent

FIGURE 2.

to saying that the scattering matrix $S(p)$ for $\mathfrak{N}$ is contractive. Similarly a network is lossless if and only if its scattering matrix is lossless.

There are three ways in which networks are connected together: series connection, parallel connection and cascade loading. We shall discuss cascade loading. Suppose that $\mathfrak{M}$ is a 2 port network and $\mathfrak{N}$ is a 1 port network. $\mathfrak{M}$ is said to be loaded or cascaded or cascade loaded with $\mathfrak{N}$, if one port of $\mathfrak{M}$ is attached to the port of $\mathfrak{N}$. In this process one port of $\mathfrak{M}$ is left free and the one port network obtained by using this free port is called the cascaded network (see Figure 3.)

Let s be the scattering operator for $\mathfrak{M}$. The operator s , recall, maps an L^2 -2 vector valued function, thought of as incoming, into an L^2 -2 vector valued function, thought of as outgoing. We write s as a matrix with operator entries:

$$(1.0) \quad s \begin{pmatrix} \text{in}_1 \\ \text{in}_2 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} \text{in}_1 \\ \text{in}_2 \end{pmatrix} = \begin{pmatrix} \text{out}_1 \\ \text{out}_2 \end{pmatrix}.$$

Similarly the scattering operator S for $\mathfrak{N}$ maps $S \text{ in} = \text{out}$. Cascading $\mathfrak{N}$ and $\mathfrak{M}$ simply sets $\text{out}_2 = \text{in}$ and $\text{in}_2 = \text{out}$, and with simple algebra one can compute that the scattering operator $\mathcal{C}_s(S)$ is

$$(1.1) \quad \mathcal{C}_s(S) = S_{11} + S_{12}S(1 - S_{22}S)^{-1}S_{21}.$$

Remark 3. Actually there is a subtlety here. Engineers sometimes assume implicitly that there is an ideal transformer added to Figure 4 in the connection between $\mathfrak{N}$ and $\mathfrak{M}$. This forces $i_{21}(t) = i_{22}(t)$.

Everything said above extends in an obvious way to the case where $\mathfrak{M}$ is a $k + n$ port and $\mathfrak{N}$ is an n port. The only change which occurs is that the

FIGURE 3.

functions in_1 , in_2 , in , out_1 , out_2 , and out must be taken to be vector valued functions in L^2 . It is possible to take Fourier transforms of $\mathfrak{s}$ and S and then apply the above argument to calculate that the scattering matrix $\mathfrak{C}_s(S)(z)$ for the cascaded network is

$$(1.2) \quad \mathfrak{C}_s(S)(z) = S_{11}(z) + S_{12}(z)S(z)(1 - S_{22}(z)S(z))^{-1}S_{21}(z).$$

So far, what we have done has been purely algebraic and formal because in both (1.2) and (1.1) we assumed that the inverse of a certain quantity exists. This may not be the case, but in interesting examples the lack of a bounded inverse causes no difficulty.

Naturally the networks which electrical engineers study are electric circuits whose components are, in the case of passive networks, capacitors, inductors, resistors, ideal transformers, coupled coils, and gyrators. The scattering matrix for such a network is always a rational function. The scattering matrix also satisfies $S(\bar{p}) = \bar{S}(p)$. The basic abstract fact of passive network realizability theory is—given a rational $n \times n$ scattering matrix there is an electric circuit $\mathfrak{N}$ whose scattering matrix is $S(p)$. The fact that every lossless scattering matrix can be realized by a network $\mathfrak{N}$, mathematically speaking, can be gotten from Theorem VI 3.1 [N-F]. The next section presents the mathematical theorem which engineers use and then interprets it. Note that there are usually many networks with the same scattering matrix.

Remark 4. Instead of associating the operator $\mathfrak{s}$ given by (1.0) with the 2 port $\mathfrak{N}$ of Figure 3 engineers [D] sometimes associate the ‘chain operator’ $\mathfrak{u}$ defined by

$$\mathfrak{u} \begin{pmatrix} \text{in}_1 \\ \text{out}_1 \end{pmatrix} = \begin{pmatrix} \text{in}_2 \\ \text{out}_2 \end{pmatrix}$$

with $\mathfrak{N}$. The chain operator naturally is associated with an operator entried matrix

$$(1.3) \quad \mathfrak{u} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

and A , B , C and D are easy to compute in terms of S_{11} , S_{12} , S_{21} and S_{22} . The advantage of this formalism is that when one connects 2 ports $\mathfrak{N}_1$, $\mathfrak{N}_2$, $\dots$, $\mathfrak{N}_k$ together in a long chain to form another 2 port $\mathfrak{N}$ the ‘chain operator’ for $\mathfrak{N}$ is just the product $\mathfrak{u}_k \cdots \mathfrak{u}_2 \mathfrak{u}_1$ of the chain operators $\mathfrak{u}_i$ for $\mathfrak{N}_i$. The disadvantage of this formalism is that $\mathfrak{u}$ is not unitary for lossless $\mathfrak{N}$. However, $\mathfrak{u}$ does satisfy $\mathfrak{u}J\mathfrak{u}^* = J$ and $\mathfrak{u}^*J\mathfrak{u} = J$ where $J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Also, $\mathfrak{u}$ is bounded if and only if $\|S_{11}\| < 1$ and $\|S_{22}\| < 1$. The operator $\mathfrak{C}_s(S)$ defined immediately above equation (1.1) can be expressed in terms of $\mathfrak{u}$ as

$$(1.4) \quad \mathfrak{F}_u(S) = \mathfrak{C}_s(S) = (AS + B)(CS + D)^{-1}.$$

Maps of the form (1.4) where $\mathfrak{U}J\mathfrak{U}^* = J$ are called 'general symplectic' and they have been investigated independently by the mathematicians Phillips [P₂], Krein-Smul'jan [K-S] and others. Work on such maps is an important part of the study of Hilbert spaces with an indefinite inner product (mentioned in Remark 1). Also, the 1-1 correspondence between J -unitary and unitary operators extends formally to a correspondence between J -expansive operators (possibly unbounded) and contraction operators. The J -expansive operators are discussed in [K-S] as a generalization of J -unitary operators.

The realizability of a scattering matrix. Let H be a complex Hilbert space with norm $\| \cdot \|$ and let $\mathfrak{L}(H)$ denote the space of bounded operators on H . Any operator in $\mathfrak{L}(H)$ with norm less than or equal to one is called a contraction. We shall assume that H is the complexification (see Chapter 9, Section 2, [S]) of a real subspace $\tilde{H}$ of H . This determines an operation which we denote by $x \rightarrow \bar{x}$ on H which we call complex conjugation. The conjugate $\bar{A}$ of an operator A in $\mathfrak{L}(H)$ is defined by $\bar{A}x = \overline{Ax}$ for all $x \in H$. An operator A in $\mathfrak{L}(H)$ is called real if $A = \bar{A}$. A corollary of Theorem VI 3.1 [N-F] is

Theorem R. *If $S(p)$ is a contraction valued function analytic in the right half plane with the properties $\|S(1)x\| < \|x\|$ if $x \in H$ and $S(\bar{p}) = \overline{S(p)}$, then there is a Hilbert space $\mathbf{H}$ and operators*

$$\begin{aligned} S_{11} : H &\rightarrow H, & S_{12} : \mathbf{H} &\rightarrow H, \\ S_{21} : H &\rightarrow \mathbf{H}, & S_{22} : \mathbf{H} &\rightarrow \mathbf{H}, \end{aligned}$$

where

$$\mathfrak{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

is a unitary map of $H \oplus \mathbf{H}$ onto $H \oplus \mathbf{H}$ such that

$$(2.0) \quad S(p) = S_{11} + S_{12}z(1 - S_{22}z)^{-1}S_{21}$$

and $z = (p - 1)/(p + 1)$. Moreover, there is a complex conjugation $\sim$ on $\mathbf{H}$, such that $\mathfrak{S}$ is real with respect to $-\oplus \sim$ on $H \oplus \mathbf{H}$. This conjugation is explicitly constructed.

Proof. Let $\theta(z) = S(p)$. Theorem VI, 3.1 [N-F] says that there is a Hilbert space $\mathbf{H}$ and a contraction operator $\mathbf{T}$ on $\mathbf{H}$ such that $\theta(z)$ coincides with the characteristic function $\theta_{\mathbf{T}}(z)$ of $\mathbf{T}$. By definition (see Chapter VI, (1.1) [N-F])

$$(2.1) \quad \theta_{\mathbf{T}}(z) = [-\mathbf{T} + zD_{\mathbf{T}^*}(I - z\mathbf{T}^*)^{-1}D_{\mathbf{T}}] |_{\mathfrak{D}_{\mathbf{T}}}$$

where $|$ denotes restriction of an operator to a subspace and where

$$\begin{aligned} D_{\mathbf{T}} &= (I - \mathbf{T}^*\mathbf{T})^{1/2}, & D_{\mathbf{T}^*} &= (I - \mathbf{T}\mathbf{T}^*)^{1/2}, \\ \mathfrak{D}_{\mathbf{T}} &= D_{\mathbf{T}}\mathbf{H} & \text{and} & \mathfrak{D}_{\mathbf{T}^*} = D_{\mathbf{T}^*}\mathbf{H}. \end{aligned}$$

For each z the operator $\theta_{\mathbf{T}}(z)$ maps $\mathfrak{D}_{\mathbf{T}}$ into $\mathfrak{D}_{\mathbf{T}^*}$, a fact which follows from the identity $\mathbf{T}D_{\mathbf{T}} = D_{\mathbf{T}^*}\mathbf{T}$ (see [N-F] I, (3.4)). The statement that $\theta_{\mathbf{T}}$ and θ coincide (see Chapter V, (2.4) [N-F]) means that there are unitary maps U and V

$$\begin{aligned} U: \mathfrak{D}_{\mathbf{T}^*} &\rightarrow H, \\ V: H &\rightarrow \mathfrak{D}_{\mathbf{T}} \end{aligned}$$

such that

$$(2.2) \quad \theta(z) = U\theta_{\mathbf{T}}(z)V.$$

Let $P(P_*)$ denote the orthogonal projection of $\mathbf{H}$ onto $\mathfrak{D}_{\mathbf{T}}$ and $\mathfrak{D}_{\mathbf{T}^*}$ respectively.

When one combines (2.1) and (2.2) one gets a representation for $\theta(z)$ and consequently for $S(p)$ of the form desired provided that

$$s = \begin{bmatrix} -UP_*\mathbf{T}V & UP_*D_{\mathbf{T}^*} \\ D_{\mathbf{T}}V & \mathbf{T}^* \end{bmatrix}$$

is a unitary map of $H \oplus \mathbf{H}$ onto $H \oplus \mathbf{H}$.

Now we prove that s is unitary. Suppose that $\begin{pmatrix} x \\ y \end{pmatrix} \in H \oplus \mathbf{H}$. Then a simple computation along with the fact that U is unitary gives us that

$$\begin{aligned} \left\| s \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 &= (P_*\mathbf{T}Vx, P_*\mathbf{T}Vx) - (P_*\mathbf{T}Vx, P_*D_{\mathbf{T}^*}y) - (P_*D_{\mathbf{T}^*}y, P_*\mathbf{T}Vx) \\ &\quad + (P_*D_{\mathbf{T}^*}y, P_*D_{\mathbf{T}^*}y) + (D_{\mathbf{T}}Vx, D_{\mathbf{T}}Vx) + (D_{\mathbf{T}}Vx, \mathbf{T}^*y) + (\mathbf{T}^*y, D_{\mathbf{T}}Vx) \\ &\quad + (\mathbf{T}^*y, \mathbf{T}^*y). \end{aligned}$$

The definition of $D_{\mathbf{T}}$ and $D_{\mathbf{T}^*}$, the identities $\mathbf{T}D_{\mathbf{T}} = D_{\mathbf{T}^*}\mathbf{T}$ and $D_{\mathbf{T}}\mathbf{T}^* = \mathbf{T}^*D_{\mathbf{T}^*}$, and the consequence $\mathbf{T}P = P_*\mathbf{T}$ of these identities, when used with the above formula yield

$$\left\| s \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 = \|x\|^2 + \|y\|^2 = \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2.$$

Thus s is an isometry. Next we prove that the range of s is dense in $H \oplus \mathbf{H}$. Suppose that

$$s \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} \ell \\ w \end{pmatrix} = 0$$

for all $x \in H$ and $y \in \mathbf{H}$. Then

$$(2.3) \quad -(UP_*\mathbf{T}Vx, \ell) + (UP_*D_{\mathbf{T}^*}y, \ell) + (D_{\mathbf{T}}Vx, w) + (\mathbf{T}^*y, w) = 0.$$

We see from setting $x = 0$ that

$$(2.4) \quad D_{\mathbf{T}^*}U^{-1}\ell + \mathbf{T}w = 0$$

and from setting $y = 0$ that

$$(2.5) \quad -V^{-1}P(\mathbf{T}^*U^{-1}\ell - D_{\mathbf{T}}w) = 0.$$

Since $\mathbf{T}^*\mathfrak{D}_{\mathbf{T}^*} \subset \mathfrak{D}_{\mathbf{T}}$, the vector $m = \mathbf{T}^*U^{-1}\ell - D_{\mathbf{T}}w$ belongs to $\mathfrak{D}_{\mathbf{T}}$ and so $V^{-1}m = 0$ implies $m = 0$. Now multiply (2.4) by $D_{\mathbf{T}^*}$, multiply m by $\mathbf{T}$, subtract the two expressions to obtain $U^{-1}\ell = 0$. Thus $\ell = 0$ and equation (2.3) reduces to

$$(D_{\mathbf{T}}Vx, w) + (\mathbf{T}^*y, w) = 0$$

for all x and y . Consequently

$$D_{\mathbf{T}}w = 0 \quad \text{and} \quad \mathbf{T}w = 0,$$

that is, $w - \mathbf{T}^*\mathbf{T}w = D_{\mathbf{T}}^2w = 0$ and $\mathbf{T}^*\mathbf{T}w = 0$. Therefore $w = 0$ and we have proved that the range of the isometry $\mathfrak{s}$ is dense in $H \oplus \mathbf{H}$. Consequently $\mathfrak{s}$ is unitary and the first part of Theorem R is proved.

Now we construct a conjugation on $H \oplus \mathbf{H}$ with respect to which the operator $\mathfrak{s}$ is real. This part of the proof is not self contained, since to make it self contained would require the reconstruction of several large parts of the book [N-F]. Thus we assume here that the reader has a good knowledge of the proof of Theorem VII, 3.1 [N-F], a knowledge of conjugation operators (see Chapter 9, Section 2 [5]), and we use the notation of [N-F] without first introducing it. Let $\mathfrak{g}$ denote the conjugation operation induced on $L^2(H)$ by —; namely

$$h(\varphi) = \sum_{-\infty}^{\infty} h_k e^{ik\varphi} \xrightarrow{\mathfrak{g}} \sum_{-\infty}^{\infty} \bar{h}_k e^{ik\varphi}$$

It is clear that the subspaces $H^2(H)$ and the orthogonal complement of $H^2(H)$ are invariant under $\mathfrak{g}$. The space $H^2(H)$ is invariant under θ and the hypothesis $\overline{\theta(\bar{z})} = \theta(z)$ says precisely that

$$\mathfrak{g}\theta = \theta\mathfrak{g}.$$

This implies that the operator Δ given by multiplication by $\Delta(t) = [I - \theta(e^{it})^*\theta(e^{it})]^{1/2}$ on $L^2(H)$ commutes with $\mathfrak{g}$. We mention at this point that an operator is real (with respect to $\mathfrak{g}$) if and only if the operator commutes with $\mathfrak{g}$. One can see from equation VI (3.29) [N-F] that the Hilbert space $\mathbf{H}$ which appears in the statement of Theorem R is

$$\mathbf{H} = [H^2(H) \oplus \overline{\Delta L^2(H)^c}] \ominus \{\theta w \oplus \Delta w : w \in H^2(H)\}$$

where $\overline{}$ denotes closure. Define $\tilde{\mathfrak{g}}$ to be the conjugation operator on $\mathbf{K} \equiv L^2(H) \oplus \overline{\Delta L^2(H)^c}$ defined by

$$\tilde{\mathfrak{g}}(u_* \oplus v) = \mathfrak{g}u_* \oplus \mathfrak{g}v \quad \text{for} \quad u_* \oplus v \in \mathbf{K}.$$

What we have done above implies that $\mathbf{H}$ and $\mathbf{H}^\perp$ are invariant under $\tilde{\mathfrak{g}}$. Thus $\tilde{\mathfrak{g}}$ commutes with the orthogonal projection $P_{\mathbf{H}}$ of $\mathbf{K}$ onto $\mathbf{H}$.

The transformation $\mathbf{U}$ of $\mathbf{K}$ onto itself defined by

$$\mathbf{U}(u_* \oplus v) = e^{it}u_* \oplus e^{it}v$$

for $u_* \oplus v \in \mathbf{K}$ is, according to Theorem VI (3.1) [N-F], the minimum unitary dilation of $\mathbf{T}$. That is,

$$\mathbf{T} = P_{\mathbf{H}} \mathbf{U}|_{\mathbf{H}}.$$

Clearly $\mathbf{U}$ commutes with $\tilde{\mathcal{G}}$, and the computation

$$\tilde{\mathcal{G}}\mathbf{T} = P_{\mathbf{H}}\tilde{\mathcal{G}}\mathbf{U}|_{\mathbf{H}} = P_{\mathbf{H}}\mathbf{U}\tilde{\mathcal{G}}|_{\mathbf{H}} = P_{\mathbf{H}}\mathbf{U}|_{\mathbf{H}}\tilde{\mathcal{G}} = \mathbf{T}\tilde{\mathcal{G}}$$

shows that $\tilde{\mathcal{G}}$ commutes with $\mathbf{T}$. Henceforth, let us write the conjugation map on $\mathbf{H}$ induced by $\tilde{\mathcal{G}}$ as $x \rightarrow \tilde{x}$. From the fact that $\mathbf{T} = \tilde{\mathbf{T}}$ we may conclude that the operator $\mathbf{S} : \mathbf{H} \oplus \mathbf{H} \rightarrow \mathbf{H} \oplus \mathbf{H}$ defined by

$$\mathbf{S} = \begin{pmatrix} -\mathbf{T} & (I - \mathbf{T}\mathbf{T}^*)^{1/2} \\ (I - \mathbf{T}^*\mathbf{T})^{1/2} & \mathbf{T}^* \end{pmatrix}$$

is $\sim \oplus \sim$ real.

The proof of Theorem R would be complete if we could show that

$$U\tilde{x} = \overline{Ux} \quad \text{and} \quad V\tilde{y} = \widetilde{Vy}$$

for every $x \in \mathfrak{D}_{\mathbf{T}}$ and $y \in H$. The explicit construction of the map V is found by combining equation VI (3.24) [N-F] which exhibits an isometry w of H onto $\mathbf{L}$ as

$$w : y \rightarrow \theta y + \Delta y \quad \text{for } y \in H$$

with equation II (1.6) [N-F] which exhibits an isometry φ of $\mathbf{L}$ onto $\mathfrak{D}_{\mathbf{T}}$ as

$$\varphi(\mathbf{U} - \mathbf{T})h = D_{\mathbf{T}}h \quad \text{for } h \in \mathbf{H}.$$

The isometry V is φw . If $y \in H$, then

$$\theta\tilde{y} + \Delta\tilde{y} = \theta\mathcal{G}y + \Delta\mathcal{G}y = \tilde{\mathcal{G}}[\theta y + \Delta y],$$

where y is thought of both as being in H and as being a constant function in $H^2(H)$. Since

$$\tilde{\mathcal{G}}D_{\mathbf{T}} = D_{\mathbf{T}}\tilde{\mathcal{G}} \quad \text{and} \quad \tilde{\mathcal{G}}(\mathbf{U} - \mathbf{T}) = (\mathbf{U} - \mathbf{T})\tilde{\mathcal{G}}$$

one gets that $V\tilde{y} = \tilde{\mathcal{G}}Vy = \widetilde{Vy}$. The isometry U is treated similarly.

In Theorem R we assumed that $\|S(1)x\| < \|x\|$ for all x in H . Such analytic functions are called purely contractive.

Proposition V (2.1) of [N-F] says that

Theorem D. *If $S(p)$ is a contraction valued analytic function, then H can be decomposed into the orthogonal direct sum of two subspaces $H = H_1 \oplus H_2$ in such a way that*

- (1) $S(p)H_j \subset H_j$ for all $\operatorname{Re} p > 0$ and $j = 1, 2$;
- (2) $S(p)$ on H_1 is purely contractive and $S(p)$ on H_2 is a constant unitary operator.

In the case where H is finite dimensional and $S(p)$ is lossless and rational, Theorems D and R taken together are precisely Theorem 3.3 in [W]. In fact, the dimension of $\mathbf{H}$ in the Nagy-Foias construction of Theorem R seems to be

the electrical engineering “Smith–McMillan” degree (c.f. Cor. 3.2 in [W]), just as one would expect from Theorem 3.3 [W]. However, when $S(p)$ is not lossless the correspondence between the electrical engineering approach and the Nagy–Foiás approach breaks down. With the Nagy–Foiás approach one needs an infinite dimensional $\mathbf{H}$ to represent any “lossy” passive $S(p)$. Physically this calls for a circuit with infinitely many components—something which is naturally anathema.

Remark 5. Engineers, when treating “lossy” networks, represent $S(p)$ as $\theta_{\mathbf{T}}((p - 1)/(p + 1)P)$ where P is some orthogonal projection onto $\mathbf{H}$. This extra freedom allows them to represent $S(p)$ with a $\mathbf{H}$ which is finite dimensional.

In order to describe the connection between Theorems D and R and network realizability theory, we need to know two things. First of all, it is at least in theory possible for engineers to build a circuit which realizes a given constant lossless scattering matrix ($S(p) = \text{constant}$). The fact that constant lossless scattering matrices are mathematically so simple, translates into the physical fact that they can always be realized by a circuit which consists only of transformers, gyrators, and wires interconnecting the ports. The second thing that we need is the fact that the scattering matrix for an inductor is $S(p) = (p - 1)/(p + 1)$ (recall the example).

For the sake of this presentation we shall take lossless constant scattering matrices to be basic and discuss them no further. Thus Theorem D reduces the realization problem for an arbitrary $S(p)$ to the purely contractive case.

Suppose that one wants to build a circuit which realizes a given purely contractive $n \times n$ scattering matrix $S(p)$ and that he has calculated the matrix $\mathbf{S}$ which appears in Theorem R. To solve the problem one first builds the $n + m$ port $\mathfrak{N}$ which has $\mathbf{S}$ as its scattering matrix (see Figure 2). Here m is the dimension of $\mathbf{H}$. Across each of the last m ports (ports associated with $\mathbf{H}$) one places a unit inductor. The scattering matrix for this circuit is seen from (1.2) to be precisely the one given by (2.0). Thus we have constructed a circuit whose scattering matrix is precisely the scattering matrix $S(p)$ that we wanted to synthesize.

Appendix. The network realizability theorem which is discussed in this article is actually a manifestation of a basic theorem in systems theory. Namely,

Theorem. Every matrix valued rational function $G(s)$ which is analytic in the unit disk can be written in the form

$$(A.1) \quad G(z) = D + zC(I - zA)^{-1}B$$

where A, B, C, D are finite dimensional matrices and $|\sigma(A)| \leq 1$.

Facts of this type were known to mathematicians since the mid-fifties. An elegant proof of this theorem which is suitable for computer implementation, was given by systems theorists B. L. Ho and independently by Youla and Tissi.

It is interesting to compare their proof with the proof of the closely related Theorem VI 3.1 [N-F]. If one expands an inner function $G(z)$ in a power series about the origin and applies the Ho construction (cf. Theorem 6.3 [C] or Ch. 10, Sec. 11 [A-F-K]) one can trace the correspondence between it and the N-F construction. In fact, the Ho and N-F constructions of the basic operator A in representation (A.1) are equivalent.

R. Kalman has proposed [A-F-K] that a certain module theoretic viewpoint gives the best approach to this branch of systems theory. Independently, two operator theorists B. Moore and E. Nordgren [M-N] [N] in studying Nagy-Foias canonical models have taken a module theoretic viewpoint very similar to the one proposed by Kalman.

REFERENCES

- [A-F-K] M. A. ARBIB, P. L. FALB, & R. E. KALMAN, *Topics in mathematical system theory*, McGraw-Hill, New York 1969.
- [C] C. T. CHEN, *Introduction to linear systems theory*, Holt, New York, 1970.
- [D] P. DEWILDE, *Roomy scattering matrix synthesis*, Technical report, University of California, Berkeley, 1971.
- [K] M. G. KREIN, *Introduction to the geometry of indefinite J -spaces and to the theory of operators in those spaces*, Amer. Math. Soc. Transl. 2, **93** (1970), 103-176.
- [K-S] M. G. Krein & J. Smul'jan, *Fractional linear transformations with operator coefficients* (Russian), Mat. Issled. **3** (1967), 64-96.
- [L] M. S. LIVSIC, *Operators, oscillations, waves, open systems*, Izdat. "Nauka", Moscow, 1966, (Russian).
- [L-P] P. D. LAX & R. S. PHILLIPS, *Scattering theory*, Academic Press, New York, 1967.
- [M-N] B. MOORE & E. NORDGREN, *On quasi-equivalence and quasi-similarity*, (to appear).
- [N] E. NORDGREN, *On quasi-equivalence of matrices over H^∞* (to appear, Acta Szeged.)
- [P₁] R. S. PHILLIPS, *The extension of dual subspaces invariant under an algebra*, Proc. Inter. Symp. on Linear Spaces, Israel, (1960), 366-398.
- [P₂] R. S. PHILLIPS, *Symplectic mappings of contraction operators*, Studia Math. **XXXI** (1968), 15-27.
- [N] R. W. NEWCOMB, *Linear multiport synthesis*, McGraw-Hill New York, 1966.
- [N-F] B. SZ-NAGY & C. FOIAS, *Harmonic analysis of operators on Hilbert space*, North Holland, Amsterdam, 1970.
- [S] M. H. STONE, *Linear transformations in Hilbert space and their applications to analysis*, A.M.S. Colloquium, 1932.
- [W] M. W. WOHLERS, *Lumped and distributed passive networks*, Academic Press, New York, 1966.
- [Z] A. H. ZEMANIAN, *A scattering formalism for the Hilbert port*, SIAM J. Appl. Math. **18**, No. 2, (1970), 467-488.

This research was partially supported by N.S.F. Grant GP-19587.

State University of New York
Stony Brook

Date communicated: NOVEMBER 12, 1971