Systems with Infinite-Dimensional State Space:
The Hilbert Space Approach

J. WILLIAM HELTON, MEMBER, IEEE

Abstract—The control of infinite-dimensional systems has received much attention from engineers and even mathematicians. Realizability although first considered in [4] has been ignored until recently. Ironically enough while state-space systems theory was developing in the early 1960's a mathematical study of scattering and of non-self-adjoint operators produced a parallel theory which was infinite dimensional "rom the beginning. When the close relationship between the two subjects became known time invariant infinite-dimensional systems theory advanced quickly and at a general level it now seems reasonably complete. This paper describes the connection between mathematical scattering and systems. It then gives a thorough treatment of infinite-dimensional time invariant continuous time systems. The last section lists recent scattering results which might be of engineering interest.

I. INTRODUCTION

VERY ROUGHLY SPEAKING, a situation which requires infinitely many parameters to specify will be modeled by a system with infinite-dimensional state space. Very early in the course of a design, one restricts attention to a few parameters which he believes are the most important. However, the resulting heavily proscribed systems will still have infinite-dimensional state space. For example, a variable coefficient transmission line has an infinitely parameterized space of capacitance, inductance, etc., distributions and any system corresponding to a transmission line will have infinite-dimensional state space. Even though there are many ways to discard all but a finite number of design parameters such as to consider only cascades of a prescribed number of uniform lines or to prescribe *a priori* the tapering of the line, all such systems still have infinite-dimensional state space. In this same vein semiconductors admit an infinite variety of doping intensities, of geometries, and of doping geometries and even the practical devices with heavily restricted geometries and doping patterns have an infinite-dimensional space of states. This applies to an enormous number of devices and though finite dimensional systems suffice as models for basic circuits, many components of a modern circuit when modeled carefully would be infinite dimensional. From this viewpoint, infinite-dimensional engineering problems possibly outnumber the intrinsically finite-dimensional ones. To reemphasize this point we mention that besides the many infinite-dimensional systems thought of more in connection with control (cf. [43]) there are transmission lines, microwave circuits, thin film devices, surface acoustic wave devices and indeed most semiconductor devices.

In fact, there are so many examples of infinite-dimensional systems that any theory general enough to deal with most of them would be too general to say much specific about any one

of them. Ideally, a complete theory would have several layers. There should be a general theory giving the structure common to many situations and then there should be a branched hierarchy of increasingly special conditions and increasingly specific theorems. Such an extensive theory will, of course, take a long time to develop. On the positive side the general theory of infinite-dimensional A, B, C, D-type linear systems now seems reasonably complete and that is what this article presents. We describe a realistic setting for many infinite-dimensional problems to which the main properties of general finite-dimensional systems adapt suitably.

Although the world abounds in important infinite-dimensional systems, of what benefit is a theory about them? Certainly the abstract considerations to be presented here are not of immediate interest to the practicing engineer. However, they may ultimately be influential. For example, the one problem common to most infinite-dimensional design situations is: How does one select a finitely parameterized approximation and how does one determine its accuracy? Problems of this type are mathematically speaking infinite dimensional and one would have little hope of developing a decent theory of finite-dimensional approximations without first having a good theory of infinite-dimensional systems. Hopefully the approach taken here addresses that need and provides a foundation for building the more specialized layers of theory which will follow. Another benefit of having an infinite-dimensional theory closely attuned to a classical mathematical physics approach (as this one is) is that developments in that area can be adapted to (and encouraged by) engineering theory; some of Section IV and the entirety of Section V are examples. Whereas, it would be highly irresponsible to urge a practicing engineer to learn the abstract theory described herein it would be equally foolish for theoretical engineers to abandon such pursuits since a unified and thorough understanding of infinite-dimensional realizability should be valuable.

Now that we have motivated a study of infinite-dimensional systems, we give some idea of what the theory entails and how it connects with other areas. The basic general principles which have emerged in finite-dimensional systems theory are the fact that two 'minimal' systems with the same frequency-response function (FRF) are isomorphic, every rational function is the FRF for some 'minimal' system, and the eigenvalues of the 'state operator' for a minimal system give the pole locations of the FRF. During the early 1960's, while this was developing a mathematical counterpart evolved independently. The mathematical theory is infinite dimensional and by combining it with the systems approach precise analogs of these basic theorems are now available in infinite dimensions. Thus there has been a recent advance in systems theory initiated by the discovery of existing mathematical work and greatly expedited by its application. A surprising fact is that not only do

the mathematical works provide techniques for the study of systems, the works themselves are reasonably close to being studies of systems. For example, the book by Nagy and Foias [49], work which was done for purely mathematical reasons, could be viewed after a translation of terms almost entirely as a treatise on infinite-dimensional discrete-time systems. If the results and methods of the book are restricted to finite dimensions they agree with the standard engineering ones for the case of 'lossless' discrete systems. A description of the relationship appears in [28].

The work of Lax and Phillips though slightly less similar in form is more similar in spirit to systems theory. This is a theory of scattering, which includes nonsymmetric scatterers as opposed to the radially symmetric ones concentrated upon in classical physics. The original work compiled in [37] treats lossless situations while [38] gives a lossy scattering theory. This works puts great emphasis on the 'state operator,' but input and output operators never appear and are replaced by a more general notion of "incoming and outgoing space." Despite these differences the author feels that Lax–Phillips scattering and infinite-dimensional systems theory are linked so closely in substance and spirit that future development of these subjects should go hand in hand. Section III is devoted to describing the precise relationship between the two subjects. An immediate consequence is that one can see how systems theory looks in the setting of classical scattering, how for example the classical objects of scattering, wave operators, and scattering operators, compare to objects in systems theory. Quite possibly these more established objects will be easier to use for infinite dimensions than controllability and observability operators. The article [28] connects Lax–Phillips scattering with systems in discrete time; while [31] does it for nonlinear situations.

Now we describe some history. Among mathematicians the first work in this direction was done by Livsic [44], [45] who began a study of operators on Hilbert space which are not self-adjoint in the early 1950's, cf. [9]. The work of Lax and Phillips on scattering theory began in the early 1960's. Also about that time DeBrange and Rovnyak and Nagy and Foias began a study of "nonunitary" operators which though similar to Livsic's in general philosophy was basically very much different. Another closely related study was done by Helson. In 1964, Adamajan and Arov [2] showed the Lax–Phillips and Nagy–Foias theories to be equivalent. In 1965, Livsic used his operator theory to derive a type of systems theory. The book received no attention in this country until connection of the above mathematical theories with systems was discovered here and generated enough interest to produce a translation [46] in 1973. The first work on infinite-dimensional systems was done in control. In realizability the first work was Balakrishnan's [6] in 1966 which described the response function for a system with distributional entries. Little happened until recently when the mathematical theory was seen to pertain; Dewilde [11], Fuhrmann [17]–[26], and Helton [27], [28]. Also around that time systems theorists Brockett and his student, Baras, became interested in the subject [5]. There was considerable subsequent work on this [6], [12], [18], [19], [29], [31], [32] and related topics [13], [20]–[26], [30]. In quite a different spirit there is the algebraic approach taken by Kamen [34]–[36] and the distribution theoretic approach of Aubin and Bensoussan [3]. More recent Russian work known to the author is the book [45] describing new directions in systems theory. Also the mathematician Potopov has

devoted his time to design of lumped circuits using the chain formalism (cf. the survey [16]), and his colleague Arov has pursued infinite-dimensional problems [1].

Amusingly enough a major problem which has troubled the field since completion of the earlier work centers on deciding what one means by an infinite dimensional system. The type of linear system treated in these works always has the form

$$\frac{d}{dt} x(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

but amusingly enough it is not at all clear what type of linear maps B and C should be. In fact this has been a major problem troubling the field since the connection between systems and operator theory was discovered. Fuhrmann and Helton avoid the problem by treating mostly discrete time systems. In treating the continuous time case Dewilde avoids B, C entirely while Baras and Brockett choose them in a mathematically convenient way which eliminates most standard examples. Thus an essential task is to find a formulation for continuous time systems which contains the main examples and to which the basic finite dimensional systems facts extend in a reasonable fashion. Ironically enough the techniques for proving these facts have (for the last three years) been much clearer than what the systems themselves should be. Even for studying discrete time systems a slightly broader set up than the one presently used might be desirable since it is necessary to embrace some common circumstances (see Section II-B).

The major accomplishment of this paper is that it gives a theory of continuous time systems which is physically realistic as well as being complete at the general level. The theory also fits together perfectly with Lax–Phillips scattering which guarantees that it contains the physical situations they have studied and is consistent with their long experience in abstracting mathematics from physical situations.

We also mention that the distributional approach of Balakrishnan and more recently Aubin and Bensoussan is very general and contains many physical situations. The work does not effectively introduce a Hilbert space structure and thus is more abstract than our approach. No isomorphism property or spectral results exist for distributional systems although the author does not know if these will come from additional work or if they require additional Banach or Hilbert space structure. Our treatment of systems might be thought of as a combined distributional and Hilbert space approach. Hilbert spaces usually arise from energy or cost considerations. Energy considerations play a fundamental role in mathematical physics which accounts for the prominence of Hilbert space there.

Our notation is standard. If H is a Hilbert space and $1 \leqslant p < \infty$, then $L^p(a, b, H)$ denotes functions $h(t)$ with values in H so that $\int_a^b \| h(t) \|^p \, dt < \infty$. A function in $L^2[-j\infty, j\infty, H]$ which is the Fourier transform of a function in $L^2[0, \infty, H]$ has an analytic continuation to the right-half plane (RHP) of complex numbers; denote the space of these functions by $H^2(H)$. The corresponding space for the left-half plane (LHP) is $\bar{H}^2(H)$. The space of all infinitely differentiable H valued functions defined on $[a, b]$ and vanishing near a and b is denoted $C_0^\infty(a, b, H)$. If H and K are Hilbert spaces $\mathcal{L}(H, K)$ denotes the space of all bounded linear operators (ones whose norm

$$\|A\| = \max_{x \in H} \frac{\|Ax\|_K}{\|x\|_H}$$

is finite) from H to K. Note a linear operator is continuous if and only if it is bounded. The space of all uniformly bounded functions on $[a, b]$ with $\mathcal{L}(H, K)$ values is $L^\infty(a, b, H, K)$ while $H^\infty(H, K)$ [resp. $\bar{H}^\infty(H, K)$] denotes those which have a uniformly bounded analytic continuation to the RHP [resp. LHP]. A one parameter semigroup $G(t)$ of linear operators on H will always satisfy a differential equation of the form $dG(t)/dt = AG(t)$ on a dense subspace $\mathfrak{D}(A)$ of H on which the linear operator A is defined. The converse also holds. A complete reference on the subject is [33], an introductory reference is [56] and an intermediate reference is [15]. A good general reference on functional analysis is [55].

II. THE SETUP

A. Continuous Time.

For our purposes a system $[A, B, C, D]$ is defined by

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

where A, B, C, D, are linear operators and $x(t)$, $u(t)$, and $y(t)$ are vector-valued functions. If the vector spaces involved are not finite dimensional, then precisely where the operators B and C map from and to is a delicate matter. One naturally wants to specify them in a nice way but it turns out that the simplest definitions exclude many interesting examples.

For example, one seemingly reasonable set up would be to have Hilbert spaces U, Y, X input, output, and state spaces, respectively. The operator A is densely defined on X, and is the infinitesimal generator of a strongly continuous semigroup e^{At} on X, the operator B maps U to X and C maps X to Y. Despite the natural appearance of this structure it is very restrictive. As Baras and Brockett [5] observe (see also [3]) the impulse response function $C e^{At} Bu$ applied to input vector u is continuous in t. This property does not hold even for lossless transmission lines; to cite the simplest example a uniform transmission line with $c = l = 1$, unit length, and with one end short-circuited has impulse response equal the dirac delta $\delta(t - 2)$ when treated in the scattering formalism (reference line $c = l = 1$). Thus a more elaborate definition of system is required in order to avoid vacuousness of the theory. Before presenting our definition of "system" we give as an example the state-space description of a lossless transmission line. This illustrates the difficulties just described and gives one something concrete to hold onto during abstract discussions. The example is like the one in [32].

Consider a transmission line with capacitance $c(x)$, inductance $l(x)$, nonvanishing and resistance $r(x)$, conductance $g(x)$ at the point x in $[0, 1]$. The energy of a current-voltage distribution $\binom{i}{v}$ in the line is

$$\left\| \binom{i}{v} \right\|_E^2 = \int_0^1 cv^2 + li^2 \, dx$$

and H_E will denote the Hilbert space of all $\binom{i}{v}$ with finite energy. The line is short-circuited at $x = 0$. To study this line we connect the $x = 1$ end to a lossless reference line having $c = l = 1$ and send signals down it toward $x = 1$. A signal moving left [resp. right] is known to have a spatial distribution at each instant of time

$$\binom{i(x)}{v(x)} = \binom{1}{-1} \alpha(x) \quad \text{resp.} \quad \binom{i(x)}{v(x)} = \binom{1}{1} \alpha(x) \ .$$

Thus a signal entering the transmission line at $x = 1$ satisfies

$$\binom{i(1)}{v(1)} = \alpha \binom{1}{-1}$$

while one leaving satisfies

$$\binom{i(1)}{v(1)} = \alpha \binom{1}{1}$$

this describes the input and output to the line. The state-space equations for this setup are:

$$\frac{\partial}{\partial t} \binom{i}{v} = A \binom{i}{v} + B\alpha$$

$$y = C \binom{i}{v} \tag{2.1}$$

where A is the unbounded operator on H_E defined by

$$A \binom{i}{v} = - \begin{pmatrix} \dfrac{r}{l} & \dfrac{1}{l} \dfrac{\partial}{\partial x} \\ \dfrac{1}{c} \dfrac{\partial}{\partial x} & \dfrac{g}{c} \end{pmatrix} \binom{i}{v}$$

acting on domain

$$\mathfrak{D}(A) = \left\{ \binom{i}{v} \text{ in } H_E \,\middle|\, \left\| A\binom{i}{v} \right\|_E \right.$$

$$\text{finite } i(1) - v(1) = 0 \text{ and } v(0) = 0 \Bigg\},$$

$$B\alpha = \delta(x - 1) \binom{1}{-1} \alpha, \quad \text{and} \quad C\binom{i}{v} = v(1) + i(1).$$

As another example, consider this same transmission line in the impedance formalism. The input is then $i(1)$ while the output is $v(1)$. The state-space equations are again (2.1), but this time

$$\mathfrak{D}(A) = \left\{ \binom{i}{v} \text{ in } H_E \,\middle|\, \left\| A \binom{i}{v} \right\|_E < \infty, v(0) = 0 = i(1) = 0 \right\},$$

$$B\alpha = \delta(x - 1) \binom{1}{0} \alpha, \quad \text{and} \quad C\binom{i}{v} = v(1).$$

Let us make an observation about the examples; while A is an unbounded operator on X as we expected, B is a distribution and C is an unbounded operator. Thus the situation is partly operator theoretic and partly distributional. This type of behavior occurs frequently and to deal with it we shall use a "rigged Hilbert space."[1] Such spaces should be familiar to followers of the Lions school; a general reference is [25].

If H and H_1 are two Hilbert spaces with $H \supset H_1$, then a third Hilbert space H' can be defined as the set of all continuous linear functionals on H_1. Since to each x in H we may assign the linear function l on H_1 defined by $l(y) = (y, x)$ for all y in H_1, the space H is regarded as being imbedded in H'. If M is an operator on H whose adjoint M^* maps H_1 into itself, then it induces an operator $\tilde{M}$ on H' according to the formula $[\tilde{M}l](y) = l(M^*y)$ for each $l \in H'$ and $y \in H_1$. The operator $\tilde{M}$

---

[1] An alternative and equivalent setting involves bilinear forms on $U \times H_1$. We could use this formalism throughout with the advantage that bilinear forms are more fashionable in current mathematical physics (see [55, ch. 8]) than rigged Hilbert space. On the other hand, the rigged Hilbert space seemed less encumbered.

is just the continuous extension of M to H, because $\widetilde{M}l_x(y) = (M^*y, x) = (y, Mx) = l_{Mx}(y)$ and so it will be denoted M. If B maps Hilbert space U into H', then its *adjoint* maps H_1 to U and is given by $(u, B^*x) = [Bu](x)$. Likewise if $C: H_1 \to U$ its adjoint sends U to H' and is given by $[C^*u](x) = (u, Cx)$. Note $B^{**} = B$ and $C^{**} = C$.

Let X, U, Y be Hilbert spaces. The closed operator A maps a dense subspace $\mathcal{D}(A)$ of X to X and generates a uniformly bounded one parameter semigroup denoted e^{At}. The expression

$$(x, y)_A = (x, y) + (Ax, Ay)$$

is meaningful for pairs of vectors x, y in $\mathcal{D}(A)$ and gives an inner product on $\mathcal{D}(A)$ which makes it a Hilbert space. Let $\mathcal{D}(A)'$ be the dual of $\mathcal{D}(A)$ and we have a rigged structure $\mathcal{D}(A)' \supset X \supset \mathcal{D}(A)$. The same can be said for $\mathcal{D}(A^*)$. Since A generates a nice semigroup A^* does, and this implies that if $\lambda > 0$ the map $(\lambda - A)^{*-1}$ sends X continuously into $\mathcal{D}(A^*)$ with $\| \;\|_{A*}$ topology. Thus $(\lambda - A)^{-1}$ is well defined on $\mathcal{D}(A^*)'$ and for w in $\mathcal{D}(A^*)'$ the linear functional $[(\lambda - A)^{-1}w](x) = w((\lambda - A)^{*-1}x)$ is continuous in the X topology. By the Riesz representation theorem it equals (x, w_0) with w_0 in X. In other words $(\lambda - A)^{-1}w = w_0$ and so $(\lambda - A)^{-1}\mathcal{D}(A^*)' \subset X$. Also $e^{A^*t}\mathcal{D}(A^*) \subset \mathcal{D}(A^*)$ (cf. [15, ch. VIII 1.5, lemma 7(b)] and is uniformly $[\,\cdot\,]_{A*}$ bounded and so e^{At} is defined on $\mathcal{D}(A^*)'$ and is uniformly bounded there.

A *system* $[A, B, C, D]$ will consist of maps, A as above, $B: U \to \mathcal{D}(A^*)'$, $C: \mathcal{D}(C) \to Y$, and $D: U \to Y$ where the domain of C satisfies $\mathcal{D}(A) \subset \mathcal{D}(C) \subset \mathcal{D}(A^*)'$, the operator B is continuous, and C is continuous on $\mathcal{D}(A)$ in the $(\,,\,)_A$ topology. A *compatible system* satisfies $(z_0I - A)^{-1}B \subset \mathcal{D}(C)$ for some z_0 with $\text{Re } z_0 > 0$. Most systems in the author's experience are compatible.

Let us pause for a moment to see what type of objects one can have for solutions $x(t)$, $y(t)$ to the system equations. Formally the state $x(t)$ arising from input function $u(t)$ is

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)} B u(s) \, ds. \qquad (2.2)$$

If $x_0 \in \mathcal{D}(A^*)'$ and $u(t)$ is a continuous U-valued function, then $e^{At}x_0$ is in $\mathcal{D}(A^*)'$, $Bu(s)$ is a continuous $\mathcal{D}(A^*)'$ valued function as is $e^{A(t-s)}Bu(s)$ and its integral is in $\mathcal{D}(A^*)'$; thus $x(t)$ is a continuous function with values in $\mathcal{D}(A^*)'$. So much for solutions to the state equation. To obtain outputs $y(t)$ from the states one needs to apply C to $x(t)$. This for systems as we have defined them is impossible because of their extreme generality: C simply cannot be applied to an arbitrary element in $\mathcal{D}(A^*)'$. The situation is not really so grim, for set $x_0 = 0$ and suppose that $u(s)$ is differentiable. Then integration by parts in the formula for $x(t)$ gives for any z_0 with $\text{Re } z_0 > 0$

$$x(t) = -\int_0^t (A - z_0)^{-1} e^{(A-z_0)(t-s)} B \frac{d}{ds} e^{z_0(t-s)} u(s) \, ds$$

$$+ (A - z_0)^{-1} B u(t) - (A - z_0)^{-1} e^{(A-z_0)t} u(0).$$

The integral is in $\mathcal{D}(A) \subset \mathcal{D}(C)$ and so, the only obstruction to having $Cx(t)$ defined lines in the last two terms. These vanish whenever support $u \subset (0, t)$ and as will be discussed later this insures that the "Hankel" operator for the system exists. In most practical situations, for example, see the transmission

line, X is a space of functions $\mathcal{D}(C)$ is a subspace of fairly smooth functions and Bu for each u in U is a "rough function." Moreover, the operator $(A - z_0)^{-1}$ is an integral operator which has a smoothing effect and $(A - z_0)^{-1}Bu$ will be smooth enough to lie in $\mathcal{D}(C)$. This guarantees that $C(A - z_0)^{-1}Bu$ and $C(A - z_0)^{-1} e^{(A-z_0)t}Bu$ exist and consequently that the output $y(t) = Cx(t) + D u(t)$ is well defined at least for any differentiable input function $u(t)$. This is the motivation behind the definition of compatible system.

There are further properties which systems frequently have. These properties of systems have been isolated by Lax–Phillips as common to all systems they study.

Meromorphic property:

The operator valued function $(zI - A)^{-1}$ is meromorphic in the whole complex plane $\mathcal{C}$.

A stronger property is

Compactness property:

There is a number $T > 0$ so that $e^{TA}(zI - A)^{-1}$ is a compact operator for $\text{Re } z > 0$.

The $A, B, C, 0$ from the transmission line example is a compatible system having $T = 0$ compactness. Note, compactness is related to stability in that any asymptotically stable system with the compactness property is exponentially stable.

The basic objects one uses to study a system are controllability, observability, and either input–output, impulse-response, or frequency-response maps. In this context they are defined as follows:

The *controllability map* $\mathcal{C}: L^1(U) \to \mathcal{D}(A^*)'$ is

$$\mathcal{C}\boldsymbol{u} = \int_0^\infty e^{At} B u(t) \, dt$$

where $\boldsymbol{u}$ is the function $u(t)$ in $L^1(U)$.

The *observability map* $\mathcal{Q}: L^1(Y) \to \mathcal{D}(A)$ is

$$\mathcal{Q}\boldsymbol{y} = \int_0^\infty [Ce^{At}]^* y(t) \, dt.$$

The FRF is

$$D + C(zI - A)^{-1} B$$

for $\text{Re } z > 0$.

The later object requires much explanation. It is well defined for a compatible system. To check this note $C(z_0I - A)^{-1}B$ is well defined and since the operator on the right-hand side of the "resolvent identity"

$$(zI - A)^{-1} - (z_0I - A)^{-1} = (z_0 - z)(z_0I - A)^{-1}(zI - A)^{-1} \qquad (2.3)$$

maps $\mathcal{D}(A^*)$ into $\mathcal{D}(A)$ we can apply C from the left and B from the right to get a natural definition of $C(zI - A)^{-1}B$. The FRF *is not necessarily defined for an arbitrary system*. The FRF is, however, given in an intuitive sense up to a constant by

$$D + (z_0 - z) C(z_0I - A)^{-1} (zI - A)^{-1}B. \qquad (2.4)$$

For many purposes the indeterminacy is unimportant. This is

because the Hankel operator for a system plays a dominant role in realizability theory and according to some thought is about the only thing one usually measures anyway; the Hankel operator for a system is determined by its FRF but is independent of any constant added to the FRF.

Now we discuss Hankel operators. If $F(jy)$ takes values in $\mathcal{L}(U, Y)$ and is uniformly bounded, then for u in $H^2(U)$ the function $F(jy) u(jy)$ is in $L^2(-\infty, \infty, Y)$ and can be written uniquely as the sum $a + b$ of functions a in $H^2(Y)$ and b in $\bar{H}^2(Y)$. Define $\text{Hankel}_F u \triangleq a$. Note that if F is constant, then $\text{Hankel}_F = 0$ so two functions differing by a constant have the same Hankel operators. If the function F given by (2.4) is uniformly bounded on RHP, then the Hankel operator for the system is Hankel_F. Now let's look at Hankel_F in the time domain. Suppose for the moment that the Fourier transform $\hat{F}$ of F which in general must be regarded as a distribution $\hat{F}(g)$ is in fact a well-behaved function. Then the Fourier transformed Hankel operator $\mathcal{H}: L^2[-\infty, 0, U] \to L^2(0, \infty, Y)$ is

$$[\mathcal{H}u](s) = \int_0^\infty \hat{F}(t+s) u(-s) \, ds.$$

If $\hat{F}$ is a distribution this definition extends naturally to give a map $\mathcal{H}: C_0^\infty(-\infty, 0, U) \to L^2(0, \infty, Y)$. The $\mathcal{H}$ formally associated with a system is

$$[\mathcal{H}u](s) = \int_0^\infty C e^{A(t+r)} B \, u(t) \, dt$$

which rigorously is defined on $C_0^\infty[-\infty, 0, U)$ through integration by parts

$$[\mathcal{H}u](r) = \int_0^\infty C(A - \rho)^{-2} \, e^{(A-\rho)(t+r)} B \, \frac{d^2}{dt^2} (e^{(t+r)\rho} \, u(t)) \, dt$$

$$(2.5)$$

and does not require F to be uniformly bounded.

The adjoint system to $[A, B, C, D]$ is $[A^*, C^*, B^*, D^*]$. The controllability [resp. observability] maps for the adjoint are the same as the observability [resp. controllability] maps for the original. Note that $\mathcal{Q}: \mathfrak{D}(A) \to L^2(0, \infty, Y)$ takes $x \in \mathfrak{D}(A)$ into the function $C e^{Ar} x$ and similarly for $\mathcal{C}$. Thus

$$\mathcal{Q}^* \mathcal{C} u = \int_0^\infty C e^{A(r+t)} B \, u(t) \, dt$$

which up to $u(r) \to u(-r)$ equals the Hankel operator for the system.

While the controllability map $\mathcal{C}$ sends $L^1(0, \infty, U)$ to $\mathfrak{D}(A^*)'$ integration by parts shows that functions in $C_0^\infty(0, \infty, U)$ map into X. If the set of vectors so obtained is dense in X the system is called *approximately controllable* or often just *controllable*. If in addition $\mathcal{C}$ is continuous it is called *continuously controllable*. If $\mathcal{C}$ maps some subspace of L^2 onto all of X, then it is *exactly controllable*. By the open mapping theorem a pseudo inverse of any continuously exactly controllable system is a bounded operator. If R is a subspace of $L^2(0, \infty, U)$ for which range $\mathcal{C}|_R$ = range $\mathcal{C}$, then for most purposes we can use $\mathcal{C}|_R$ instead of $\mathcal{C}$. We call such a subspace controlling for $\mathcal{C}$. Reachability is also used by some authors instead of the term controllability. Henceforth *observability*,

exact observability, etc. will be defined as above with $\mathcal{Q}$ replacing $\mathcal{C}$. A one parameter semigroup e^{At} of operators on X is called asymptotically stable if $e^{At}x \to 0$ for each x in X.

Remark 2.1. In the definition of system we assumed that A generated a uniformly bounded semigroup e^{At}. One could replace this by the assumption that the semigroup is bounded by $\|e^{At}\| \leqslant e^{\rho t}$ with $\rho > 0$, since a simple scale change $A - \rho$ converts this to a uniformly bounded semigroup. The FRF (at least (2.4)) will be defined on the half-plane Re $z > \rho$ rather than the RHP, and everything else works with analogous modification. In fact the main reason for taking $\rho = 0$ was to avoid carrying the subscript ρ through the whole paper.

Remark 2.2. A more general but natural definition of system is possible. Namely, the operators A and D remain as is, the operator B maps U into $\mathfrak{D}(A^{n^*})'$, while $C: \mathfrak{D}(A^m) \to Y$. With this definition everything in the section after a straightforward modification still goes through. To make sense of most formulas one simply requires what in the time domain amounts to repeated integration by parts. For example, in (2.5) just integrate by parts $n + m - 2$ more times.

B. Discrete Time

What is the correct notion of a discrete-time system? Let us consider a typical situation—discretizing the transmission line from subsection a. The method we choose is a usual one (cf. [10, Appendix C] namely, send in pulses of duration T and read values or averages of the output at time intervals of length T. The state propagation equation for the discrete version of (2.1) is

$$w(n + 1) = \tilde{A}w(n) + \tilde{B}\alpha(n)$$

where $\tilde{A} = e^{A \cdot 1}$ and

$$\tilde{B} = \int_0^1 e^{A\eta} B \, d\eta.$$

The input, output, and state space are the same as before. For each $\alpha(t)$ put into the continuous time system, the resulting state at time T is $\int_0^T e^{A(T-\eta)} B\alpha(\eta) \, dn$ and its energy and consequently $\| \ \|_E^2$ is no greater than $\int_0^T |\alpha(t)|^2 \, dt$ the energy of $\alpha(t)$. If $\alpha(t)$ is a pulse of height α, then the state is precisely $\tilde{B}\alpha$; thus $\tilde{B}$ is a bounded operator. Since A is dissipative $e^{A \cdot 1}$ has norm $\leqslant 1$. Now we turn to C. There are several reasonable ways to define it. A common one is set $\tilde{C} = C$. Thus $\tilde{C}$ is an unbounded operator on H_E. Another is to have $\tilde{C}\binom{i}{v}$ give the average value of $i - v$ over some small interval $(\beta, 1)$. If T is small and β is chosen appropriately this can be approximated in practice by taking a time average of $i(1, t) - v(1, t)$ over a sample interval of length T. Here $\tilde{C}$ is a bounded operator. The first method though seemingly simple is incompatible with energy considerations, while the second is consistent with energy considerations.

The second type of discrete system is treated in [28] and works of Fuhrmann. The first can be completely analyzed using the same techniques adapted to a setting like we use in this paper.

III. LAX–PHILLIPS SCATTERING AND SYSTEMS

A. The Lax–Phillips Model

This section begins with a summary of Lax–Phillips structure adapted somewhat to our purposes. Chapter 1 of [37] is a

boon to a physical understanding of this formalism—rather dryly presented here.

Let U, Y, X be Hilbert spaces. We define H to be the Hilbert space

$$H = L^2(-\infty, 0, U) \oplus X \oplus L^2(0, \infty, Y)$$

and $T(t)$ to be a strongly continuous uniformly bounded one parameter semigroup of operators on H which describes how information in H changes as time goes by. Let $\mathcal{T}_-(t)$ [resp. $\mathcal{T}_+(t)$] denote translation by t units to the right on $L^2(-\infty, 0, U)$ [resp. $L^2(0, \infty, Y)$]. Denote by $S_{[-a,b]}$ the subspace $L^2(-a, 0, U) \oplus X \oplus L^2(0, b, Y)$ of H; by $S_{[X,\infty]}$ the subspace $X \oplus L^2(0, \infty, Y)$, by $S_{[0,\infty]}$ the subspace $L^2(0, \infty, Y)$, etc. Let $P_{[-a, b]}$ be the orthogonal projection onto $S_{[-a, b]}$, etc. The semigroup $T(t)$ has the following properties.

(i) $T(t)g = \mathcal{T}_+(t)g$ on $L^2(0, \infty, Y)$
      $T(t)^*g = \mathcal{T}_-(t)^*g$ on $L^2(-\infty, 0, U)$.

(ii) For $a > 0$ define $Z(t) = P_{[-a, a]} T(t) P_{[-a, \infty]}$.   Then $Z(t)$ and $Z(t)^*$ are asymptotically stable.

(iii) For $a \geqslant 0$

$$\lim_{t \to \infty} P_{[-a, \infty]} T(t)^* f = 0$$

$$\lim_{t \to \infty} P_{[-\infty, a]} T(t) f = 0.$$

The number a is not critical and in practice a is chosen in an arbitrary fashion. The choice of a effects the scattering matrix but in a trivial fashion, namely, $S_b(z) = e^{(b-a)z} S_a(z)$.

A further restriction [38, eq. (1.2)] which all Lax–Phillips models satisfy is that of *inertness*:

$T_+(t) S_{[0, \infty]} = T(t) S_{[0, \infty]}$ is orthogonal to range $T(t) S_{[-\infty, X]}$

$T_-(t)^* S_{[-\infty, 0]} = T(t)^* S_{[-\infty, 0]}$ is orthogonal to range

$$T(t)^* S_{[X, \infty]}.$$

An inert model which satisfies (ii) and (iii) for a particular $a > 0$ will satisfy them for any $a > 0$. We call the setup just described a Lax–Phillips model (with reference a).

B. Example

The transmission line example fits nicely in the Lax–Phillips framework and in fact is the type of situation which motivated their theory. By reading [37, ch. 1] one can see how to do this example in the form of Section III-A. However, it is more instructive to do a typical systems example which is somewhat bizarre from the Lax–Phillips viewpoint since this underscores what is happening in general.

Consider a rod of unit length and temperature distribution $x(r, t)$ which is fixed, $x(0, t) = x(1, t) = 0$, at the ends. There is one heater in the rod which upon receipt of input α supplies heat in a smooth distribution $\alpha h(r)$, and one measuring instrument which reads off the temperature at $r = \frac{1}{2}$. The associated system is

$$X = L^2[0, 1]$$

$$U = Y = R^1$$

$$Af = \frac{d^2f}{dx^2} \text{ on } \mathfrak{D}(A)$$

$$= \left\{ f \in L^2[0, 1] : f(0) = f(1) = 0 \text{ and } \int_0^1 \left| \frac{d^2f}{dx^2} \right|^2 < \infty \right\}$$
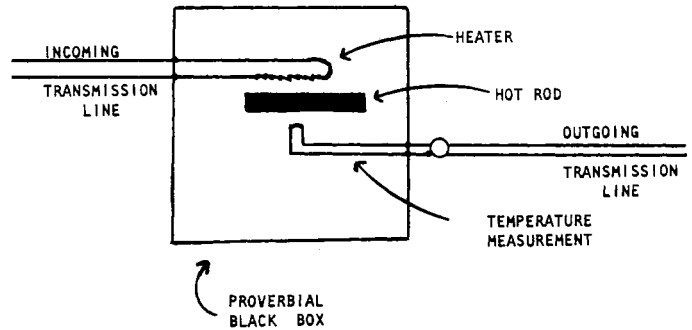

Fig. 1.

$$B\alpha = \alpha h(r)$$

$$Cx = x\left(\tfrac{1}{2}\right)$$

$$D = 0.$$

It is compatible and satisfies the compactness property. To fit the system into a Lax–Phillips model, connect an infinitely long transmission line to the heater and another one to the measuring device. Energy is supplied to the heater by sending an incoming electric signal down the input wire, and the measurement is converted into an outgoing electric signal which travels out the output wire (see Fig. 1). The space of possible incoming [resp. outgoing] signals is

$$\left\{ \begin{pmatrix} f \\ -f \end{pmatrix} \text{ with } f \text{ in } L^2(-\infty, 0, R^1) \right\}$$

$$\text{resp. } \left\{ \begin{pmatrix} f \\ f \end{pmatrix} \text{ with } f \text{ in } L^2(0, \infty, R^1) \right\}.$$

Thus the incoming and outgoing spaces are naturally identifiable with L^2 spaces and $\mathcal{H}$ is $L^2(-\infty, 0, R^1) \oplus X \oplus L^2(0, \infty, R^1)$. At a particular time in the operation of this contraption a power signal will be coming in, something will be going on in X, and a measurement signal will be going out, in other words, one would observe a vector in $\mathcal{H}$. Later t time units one would observe a vector h'. The map sending h to h' defined on all $\mathcal{H}$ by this process is $T(t)$ the time evolution semigroup. We could write it down explicitly but do not to avoid redundancy with the next section which gives a formula for doing this in general. The Lax–Phillips model here is inert. Clearly any incoming signal or state when propagated for t time units has no influence on $L^2[t, \infty, R^1]$, thus the first condition is satisfied. Intuitively the condition says that state and incoming waves cannot have influence in the outgoing space which propagates faster than the prevailing signal speed.

C. Getting a Lax–Phillips Model from a System

We just saw how a Lax–Phillips model could be associated to one particular system; now we show how to do this in general. It is done with an abstraction of the preceding construction which we call *the imbedding construction*.

Proposition 3.1: If $[A, B, C, D]$ is a system the imbedding construction associates with it a Lax–Phillips model with arbitrary reference if and only if $[A, B, C, D]$

(a) is continuously controllable and observable,
(b) has uniformly bounded FRF,
(c) has e^{At} and e^{At^*} asymptotically stable.

To describe the imbedding construction we need definitions. Let $\chi_{[0, t]}$ be the function on R^1 which equals 1 on the inter-

val $[0, t]$ and 0 off of that interval. Define $R_t : L^2(0, \infty, H) \to L^2(0, \infty, H)$ by $[R_t f](r) = \chi_{[0,t]}(r)f(t-r)$ and denote by m the reversal operator $u(r) \to u(-r)$. The space $\mathcal{H}$ in the Lax-Phillips model is obviously $L^2(-\infty, 0, U) \oplus X \oplus L^2(0, \infty, Y)$ and we must define operators $T(t)$ on $\mathcal{H}$. Since $\mathcal{H}$ is naturally decomposed into three subspaces it is natural to express $T(t)$ as an operator entried 3×3 matrix:

$$T(t) = \begin{pmatrix} \mathcal{I}_+(t) & \kappa(t) & \mu(t) \\ 0 & e^{At} & \beta(t) \\ 0 & 0 & \mathcal{I}_-(t) \end{pmatrix}$$

where $\beta(t) : L^1(-\infty, 0, U) \to \mathcal{D}(A^*)'$ is defined by

$$\beta(t)u = \int_0^t e^{A(t-r)}Bu(-r)\,dr = \mathcal{C}R_t mu$$

and $\kappa(t) : \mathcal{D}(A) \to L^2(0, \infty, Y)$ is

$$\kappa(t)x = R_t C e^{Ar}x = R_t \mathcal{Q}^*x$$

and $\mu(t) : L^2(-\infty, 0, U) \to L^2(0, \infty, Y)$ is

$$[\mu(t)u] = R_t C \int_0^r e^{A(r-\lambda)}Bu(-\lambda)\,d\lambda + R_t Dmu.$$

Properties of the system correspond to properties of $T(t)$ as follows:

System	Lax-Phillips model
$(a) + (b) + e^{At}$ uniformly bounded	$T(t)$ uniformly bounded
no restriction	(i)
(c)	(ii)
(a) + (c)	(iii)
no restriction	inertness

as we now demonstrate. Since $\|R_t\| = 1 = \|m\|$, we have $\|\beta(t)\| \leqslant \|\mathcal{C}\|$ and $\|\kappa(t)\| \leqslant \|\mathcal{Q}^*\| = \|\mathcal{Q}\|$ for all t. Conversely $\|R_t f\|_{L^2} \to \|f\|_{L^2}$ as $t \to +\infty$ and so $\max_t \|\kappa(t)\| = \|\mathcal{Q}\|$. Moreover $\cup_{t>0}$ range R_t is dense in $L^2(0, \infty, U)$, so $\max_t \|\beta(t)\| = \|\mathcal{C}\|$. Since $\mu(t)$ is R_t applied to the input–output map $\max_t \|\mu(t)\|$ equals the norm of the input–output map which in turn equals $\max_{\operatorname{Re} z > 0} \|S(z)\|$ where $S(z)$ is the FRF. This establishes the equivalence surrounding boundedness.

Lax-Phillips property (i) and inertness are built into the construction. Property (ii) is the same as (c). The second part of property (iii) says $e^{At}P_X h + \beta(t)P_{[-\infty,0]}h \to 0$ and $\mathcal{I}_-(t)P_{[-\infty,0]}h \to 0$. The last convergence is automatic and the e^{At} convergence follows from (c). Let us treat $\beta(t)u$ for fixed u in $L^2[-\infty, 0, U]$. Suppose $\epsilon > 0$, pick N so that $\|P_{[-\infty,N]}u\|_{L^2} < \epsilon$. For $t > N$

$$\beta(t)u = \int_0^N + \int_N^t e^{A(t-r)}B\,u(-r)\,dr$$

and note that the second integral is dominated by $\|\mathcal{C}\|\epsilon$. The first integral is $e^{A(t-N)}x$, where $x = \int_0^N e^{A(N-r)}Bu(-r)\,dr$ and so goes to zero as $t \to \infty$. Thus continuous controllability plus (i) gives the second part of property (ii); the first part follows similarly from continuous observability.

We still must show that $T(t)$ is a semigroup. Upon multiplying $T(t)T(w)$ and equating it to $T(t+w)$ we obtain the following identities

$$\mathcal{I}_+(t)\kappa(w) + \kappa(t)e^{Aw} = \kappa(t+w) \quad (3.2a)$$

$$e^{At}\beta(w) + \beta(t)\mathcal{I}_-(w) = \beta(t+w) \quad (3.2b)$$

$$\mathcal{I}_+(t)\mu(w) + \kappa(t)\beta(w) + \mu(t)\mathcal{I}_-(w) = \mu(t+w) \quad (3.2c)$$

which we must verify. Lets check 1; its right side is

$$\mathcal{I}_+(t)R_w C e^{Ar} + R_t C e^{A(r+w)}.$$

The first term is $Ce^{A(w-r+t)}$, for $t \leqslant r \leqslant w + c$ and 0 for all other r, while the second is $Ce^{A(t-r+w)}$ for $0 \leqslant r \leqslant t$ and so both terms sum to $Ce^{A(t+w-r)}$ for $0 \leqslant r \leqslant t+w$. This is $R_{t+w}Ce^{Ar}$ as required in 1. The adjoint of 2 has the same form as 1 and so it holds. Assume for the moment that $D = 0$. The left side of (3.2c) applied to f in $C_0^\infty(-\infty, 0, U)$ is

$$\mathcal{I}_+(t)R_w C \int_0^r e^{A(r-\lambda)}B f(-\lambda)\,d\lambda + R_t C e^{Ar}$$

$$\int_0^w e^{A(w-\lambda)}B f(-\lambda)\,d\lambda + R_t C \int_0^r e^{A(r-\lambda)}B[\mathcal{I}_-(w)f](-\lambda)\,d\lambda.$$

This equals

$$C \int_0^{w-r+t} e^{A(w-r-\lambda+t)}B f(-\lambda)\,d\lambda, \quad \text{for } r \in [t, t+w]$$

plus

$$C \int_0^w e^{A(t+w-r-\lambda)}B f(-\lambda)\,d\lambda, \quad \text{for } r \in [0, t]$$

plus

$$C \int_0^{t-r} e^{A(r-\lambda)}B f(-\lambda - w)\,d\lambda, \quad \text{for } r \in [0, t].$$

A change of variables makes the last integral

$$\int_w^{t-r+w} e^{A(t-r-\lambda+w)}B f(-\lambda)\,d\lambda.$$

This plus the middle integral is

$$C \int_0^{t+w-r} e^{A(t+w-r-\lambda)}B f(-\lambda)\,d\lambda, \quad \text{for } r \in [0, t]$$

which added to the first integral extends the formula to $[0, t+w]$. However, this is $R_{t+w} \int_0^r e^{A(r-\lambda)}B f(-\lambda)\,d\lambda$ as required. To verify (c) when D is not 0, all we must do is check that $\mathcal{I}_+(t)R_w D + R_t D \mathcal{I}_-(w) = R_{t+w}D m$. The operator D factors out to the left, $m\mathcal{I}_-(w) = \mathcal{I}_+(w)^*m$ and so the problem reduces to verifying

$$\mathcal{I}_+(t)R_w + R_t \mathcal{I}_+(w)^* = R_{t+w}$$

which is easy.

The imbedding construction can be carried out with many variations. One we shall use takes

$$\mathcal{H} = L^2(-\infty, -a, U) \oplus X_a \oplus L^2(a, \infty, Y)$$

with $a > 0$; we call it the a-imbedding.

D. Getting a system from a Lax–Phillips model

To a Lax–Phillips model for each number $a > 0$, we associate the compatible system:

- A the infinitesimal generator of the semigroup $P_{[-\infty, a]}$ $T(t) P_{[-a, \infty]}$,
- Bu is the linear functional on $\mathcal{D}(A^*)$ satisfying $[Bu](x) = (u, [P_{[-a, 0]} x](-a))$,
- Cx $[P_{[0, a]}] x(a)$, for x in $\mathcal{D}(A) + S_{[-a, a/2]}$ which we call the emitted system.

The first order of business is to clarify these definitions. In particular we need to establish that B and C are defined for those x cited. After completing that we prove the emitting and imbedding construction are inverse to each other.

Proposition 3.2: The a-imbedding construction applied to the system emitted at a from a Lax–Phillips model gives the original model and conversely.

To prove that B is well defined we need to show for x in $\mathcal{D}(A^*)$ the function $g \triangleq P_{[-a, 0]} x$ is right continuous at $-a$. Let ρ_0, ρ_- be positive smooth functions on $[-a, 0]$ with $\rho_0 + \rho_- = 1$ and $\rho_- \equiv 0$ outside of $[-a, -a/4]$. The vector x is in $\mathcal{D}(A^*)$ if and only if $P_{[-a, a]} [(T(t)^* - I)/t] x$ converges. Consider

$$P_{[-a, a]} \Delta(t) \rho_- g = P_{[-a, a]} (\Delta(t) \rho_-) g + P_{[-a, a]} (\mathcal{T}_-(t)^* \rho_-) \Delta(t) g$$

where $\Delta(t) = [(\mathcal{T}_-(t)^* - I)/t]$. The first term converges to $P_{[-a, a]} \rho_-' g$. Inertness implies that $P_{[-a, a]} (\mathcal{T}_-(t)^* \rho_-) [T(t)^* - I/t] P_{[X, a]} x = 0$ for small t, thus the second term equals $P_{[-a, a]} (\mathcal{T}_-(t)^* \rho_-) [T(t)^* - I/t] x$ which converges since $x \in \mathcal{D}(A^*)$. We have shown that $\lim_{t \to 0} P_{[-a, a]} \Delta(t) \rho_- g$ exists in $L^2[-a, 0, Y]$. In other words $\rho_- g$ is differentiable in a certain sense. It is well known (and easy to prove, cf. [53, Theorem 1.15]) that this type of differentiability implies $\rho_- g$ is continuous. Thus $P_{[-a, 0]} x$ is continuous near $-a$. A similar argument holds for $x \in \mathcal{D}(A)$ and suffices to show that C is well defined.

Proof of Proposition 3.2: The space $\mathcal{H}$ arising from the a-variation on the imbedding construction from one system $[A, B, C]$ is $\mathcal{H} = L^2(-\infty, -a, U) \oplus X_a \oplus L^2(a, \infty, Y)$ and we denote by β_s, κ_s, and μ_s the constituents of the semigroup $T_s(t)$ in this model. The space in the original Lax–Phillips model can be identified with $\mathcal{H}$ in the apparent way. ($X_a \approx L^2(-a, 0, U) \oplus X \oplus L^2(0, a, Y)$, etc.) and the semigroup $T(t)$ can be written as a 3×3 matrix with critical entries denoted β, κ, μ. We shall show that $\beta = \beta_s$, etc.; the main tool is

Lemma 3.3: Suppose that Z is a vector space and $g(s)$ is a one parameter semigroup of linear operators on Z. If for each $t, s \geqslant 0$ the map $\delta(t) : Z \to L^2(0, \infty, Y)$ satisfies

$$\mathcal{T}_+(t) \delta(s) - \delta(t) g(s) = \delta(t + s) \tag{3.3}$$

and if for each fixed z, t the L^2 function $\delta(t) z$ satisfies $\lim_{r \uparrow b} [\delta(t) z](r) = 0$, then $\delta(b) z$ is identically zero on $[0, b]$.

Proof: For fixed t, s the function $\mathcal{T}_+(t) \delta(s) z$ is identically zero on $[0, t)$ and so on that interval $\delta(t + s) z = -\delta(t) g(s) z$. Thus $\lim_{r \uparrow t} [\delta(t + s) z](r)$ exists and is 0; this is independent of s so the lemma follows.

Consider $\delta = \kappa_s - \kappa$, since both κ_s and κ satisfy (3.2a), their difference δ does also; note (3.2a) is a special case of (3.3). For x in $\mathcal{D}(A)$ the $\lim_{r \uparrow b} [\kappa_s(b) x](r)$ exists and is Cx. To complete the proof we need only show that this limit on κ equals C because the result subsequently follows from Lemma 3.3. To do this let ρ_0, ρ_+ be smooth functions on $[0, a]$ which satisfy $\rho_0 + \rho_+ = 1$ and $\rho_0 \equiv 0$ outside of $[0,$

$a/2]$. Observe using inertness that $\kappa(t)[P_{[-a, X]} x + \rho_0 P_{[0, a]} x]$ is orthogonal to range $\mathcal{T}_+(t + a/2)$, in other words it is identically zero on $[t + a/2, \infty]$. Thus $\kappa(t) x = \kappa(t) \rho_+ P_{[0, a]} x$ on $[t + a/2, \infty]$ and if $x \in \mathcal{D}(A)$ the argument used to imply continuity of $x \in \mathcal{D}(A^*)$ at $-a$ implies here that $\rho_+ P_{[0, a]} x$ is a left continuous function at a. Thus $[\kappa(t) x](r) = [\kappa(t) \rho_+ \cdot P_{[0, a]} x](r) = R_t \mathcal{T}_+(r) P_{[0, a]} x|_a$, but $\lim_{r \uparrow t} \mathcal{T}_+(t - r) P_{[0, a]} x = [P_{[0, a]} x](a) = Cx$.

Since β, β_s adjoint have the same basic form as κ, κ_s the argument above implies that $\beta = \beta_s$. Now we turn to μ and μ_s. Their difference δ satisfies (3.3) with $g(s) = \mathcal{T}_-(s)$. Inertness implies that the action of $T(t)$ on a function in $L^2(-\infty, -a, U)$ is not felt in $L^2(a, \infty, Y)$ for $2a$ time units; thus $\mu(t) = 0$ if $t < 2a$. This is also true of $\mu_s(t)$. Therefore, $\lambda(t) = 0$ for $t < 2a$ and for $a < t < 2a$ we have $\lim_{r \uparrow t} \lambda(t)(r) = 0$. On the other hand, inertness implies that $\mathcal{T}_+(t)^* \lambda(t) = 0$ which together with (3.3) gives $\lambda(t) = \mathcal{T}(t)^* \lambda(t + s)$. That is $\lambda(t + s) \cdot (r + s) = \lambda(t)(s)$; so $\lim_{r \uparrow t} \lambda(t)(r)$ is independent of t and from the above observation it is 0.

At this point, we mention an additional example. The transmission line example effectively demonstrates the system one would use to model many typical one-dimensional scattering situations. While the state operator A varies with the transmission line one is testing the operators B and C do not; they are intrinsic to the reference line. For scattering in odd dimensions the same is true and the input–output operators B and C can be written very explicitly by using directly the formulas of [38, sect. 7]. Instead of writing out the setup and formulas explicitly we just give the recipe. Here $U = Y = L^2(S^{n-1})$ where $S^{n-1} = \{(x_1, \cdots, x_n): |x_1|^2 + \cdots + |x_n|^2 = 1\}$. To define Bu plug $h(x \cdot \omega, \omega) = \delta_{-a}(x \cdot \omega) u(\omega)$ into (7.17). To define C one evaluates $\mathcal{J}d$ of (7.11) at a; note (7.11) is given in terms of (7.8) and (7.4).

E. Finer Structure

Now that we have shown how systems theory and scattering theory correspond it is natural to examine how basic objects such as the FRF the controllability and observability operators compare to basic objects in scattering theory. There are three such objects in scattering: the scattering matrix, the forward and the backward wave operators. Since systems correspond to Lax–Phillips models and any Lax–Phillips model has wave and scattering operators these objects can be associated with any system. The bulk of this section is devoted to computing them explicitly for a system; we find that they are closely related to $\mathcal{C}$, $\mathcal{O}$, and the FRF. In [28] this business was thoroughly worked out for the discrete-time case.

Before beginning the computation we mention another direction. Certain functions called 'distorted plane waves' play a fundamental role in classical scattering theory. Helton and Ralston [32] showed that for the system $[A, B, C]$ naturally associated with the lossless transmission line (as in Section II) the state-space valued function $(z - A)^{-1} Bu$[resp. $(z - A)^{-1^*} C^* y$] actually equals the classical physics incoming [resp. outgoing] distorted plane wave. (Here u, y must be chosen in a certain basis.) The $(z - A)^{-1} B$ objects are used repeatedly in studying systems and the author feels it is quite interesting that they correspond to such important objects in scattering theory. It seems reasonable clear that this correspondence will carry to most situations. One should be justified in thinking of $(z - A)^{-1} Bu$, etc., as a superposition of distorted plane waves.

Now we compute the wave and scattering operators for a given system. The basic philosophy behind scattering is to study a situation by comparing it to one where nothing happens, a free situation. Given a Lax-Phillips model as in Section III-A the appropriate free model has evolution operator denoted $S(t)$ which is the shift by t units to the right on $L(-\infty, \infty, U)$. For the sake of convenience we shall always take $U = Y$ which is no real restriction since the smaller space can be enlarged to give two equidimensional spaces which are unitarily equivalent. If the Lax-Phillips model has reference a, then let $\tilde{\rho}$ be a map of $L^2(-a, a, U)$ into $L^2(-a, 0, U) \oplus X \oplus L^2(0, a, U)$. Let ρ denote the map of $L^2(-\infty, \infty, U)$ into $\mathcal{H}$ which is the identity on $L^2(-\infty, -a, U)$ and $L^2(a, \infty, U)$ and $\tilde{\rho}$ on $L^2(-a, a, U)$. The backward wave operator is the map W_- of $L^2(-\infty, \infty, U)$ to $\mathcal{H}$ defined as

$$\lim_{t \to \infty} T(t) \rho S(t)^*.$$

This we now compute. Let $[A_1, B_1, C_1, 0]$ be the system whose Lax-Phillips model with reference a is the free system. Then $T(t)\rho S(t)^*$ written in matrix form is:

$$\begin{pmatrix} \mathcal{T}_+(t) & \kappa(t) & \mu(t) \\ 0 & e^{At} & \beta(t) \\ 0 & 0 & \mathcal{T}_-(t) \end{pmatrix} \begin{pmatrix} \mathcal{T}_+(t)^* & 0 & 0 \\ \tilde{\rho}\kappa_1(t)^* & \tilde{\rho}e^{A_1 t} & 0 \\ \mu_1(t)^* & \beta_1(t)^* & \mathcal{T}_-(t)^* \end{pmatrix}.$$

As t goes to infinity $\mathcal{T}_+(t)^*$, $\kappa_1(t)$, and $e^{A_1 t}$ go to zero (in the strong operator topology [55, ch. 6, sect. 1]. Thus we see immediately that ρ has no effect on W_-. Furthermore, since the Lax-Phillips model for $T(t)$ is inert we have $\mu(t) \mathcal{T}_-(t)^*$, $\beta(t) - \mathcal{T}_-(t)^*$, $\mathcal{T}_-(t)\mu_1(t)^*$, and $\mathcal{T}_-(t)\beta_1(t)^*$ equal zero. What remains is

$$W_- = \lim_{t \to \infty} \begin{pmatrix} \mu(t)\mu_1(t)^* & \mu(t)\beta_1(t)^* & 0 \\ \beta(t)\mu_1(t)^* & \beta(t)\beta_1(t)^* & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The main influence of a is in the middle row of $S(t)^*$ and we have seen that this row disappears in the limit. The influence of a on the rest of the calculation is merely in adjusting ranges of integration. Such bookkeeping merely confuses otherwise illuminating and straightforward computations. Therefore, we shall henceforth take $a = 0$.

Now we compute the entries of the matrix. The bottom right one is

$$\lim_{t \to \infty} \beta(t)\beta_1(t)^* = \lim_{t \to \infty} \mathcal{C}^* \chi[0, t] \mathcal{C}_1^* = \mathcal{C}\mathcal{C}_1^*.$$

Above it we have

$$\mu(t)\beta_1(t)^* x = R_t C \int_0^r e^{A(r-\lambda)} B[mR_t \mathcal{C}_1^* x](-\lambda) \, d\lambda + R_t D R_t \mathcal{C}_1^* x$$

$$= C \int_0^{t-r} e^{A(t-r-\lambda)} B C_1 e^{A_1(t-\lambda)} x \, d\lambda + \chi_{[0, t]} D \mathcal{C}_1^* x.$$

Make the change of variables $t - \lambda \to \xi$ in the integral to obtain

$$C \int_r^t e^{A(\xi-r)} B[\mathcal{C}_1^* x](-\xi) \, d\xi + \chi_{[0, t]} D \mathcal{C}_1^* x.$$

The limit as $t \to \infty$ is $\bar{\delta}_* \mathcal{C}_1^*$ where $[\bar{\delta}g](r) = C \int_{-\infty}^{-r} e^{A(-\xi - r)}$.

$Bg(\xi) \, d\xi + D$. The lower left entry is

$$\beta(t)\mu_1(t)^* g = \mathcal{C}R_t \int_\lambda^\infty B_1^* e^{A_1^*(r-\lambda)} \mathcal{C}_1^* R_t g(r) \, dr.$$

Upon writing R_t explicitly, making the change of variables $t - r \to \xi$, and sending $t \to \infty$ we obtain $\beta(t)\mu_1(t)^* g \to \mathcal{C} \int_\lambda^\infty \cdot B_1^* e^{A_1^*(\lambda - \xi)} \mathcal{C}_1^* g(\xi) \, d\xi$ which we abbreviate $\mathcal{C}\zeta_1^*$. Here δ_1 is the transfer operator for $[A_1, B_1, C_1, 0_1]$. Finally, the top corner entry is

$$\mu(t)\mu_1(t)^* g = \left[R_t C \int_0^r e^{A(r-\lambda)} B + D \right]$$

$$\cdot \int_\lambda^t B_1^* e^{A_1(s-\lambda)^*} C_1^* g(t-s) \, ds \, d\lambda.$$

Change variables twice, namely $t - s \to \xi$ and then $t - \lambda \to \eta$ to obtain

$$\left[\chi_{[0, t]} C \int_r^\infty d\eta \, e^{A(\eta - r)} B + D \right] \int_0^\eta B_1^* e^{A_1^*(\eta - \xi)} C_1^* g(\xi) \, d\xi.$$

The limit as $t \to \infty$ is $\bar{\delta}_* \delta_{1*} g$.

These expressions can be seen in a very simple way because the action of $\mathcal{C}_1^*$ and $\delta_{1*} = \delta_1^*$ are trivial. The operator W_- when written with respect to the basis

$$L^2(-\infty, 0, U) \oplus L^2(0, \infty, U)$$

$$\longrightarrow L^2(-\infty, 0, U) \oplus X \oplus L^2(0, \infty, Y)$$

is

$$W_- = \begin{pmatrix} \bar{\delta}_* & 0 \\ \mathcal{C} & 0 \\ 0 & I \end{pmatrix}.$$

The classical *forward wave operator* makes no sense for most lossy situations but natural generalizations $\lim_{t \to 0} T(t)^* \cdot S(t)^* = W_+$ or its adjoint $W_2 = \lim_{t \to \infty} S(t) T(t)$ do make sense, cf. [29] or [38]. The same procedure as above yields that W_+ in the basis $L^2(-\infty, 0, U) \oplus L^2(0, \infty, U) \to L^2(-\infty, 0, U) \oplus X \oplus L^2(0, \infty, U)$ is

$$W_+ = \begin{pmatrix} I & 0 \\ 0 & Q \\ 0 & \bar{\delta}^* \end{pmatrix}.$$

The classical scattering operator is $W_+^* W_-$ and from our computations is

$$\begin{pmatrix} \bar{\delta} & 0 \\ Q^*\mathcal{C} & \bar{\delta} \end{pmatrix}.$$

The entry $[Q^*\mathcal{C}g](\xi) = C \int_{-\infty}^0 e^{A(\xi - \eta)} B g(\eta) \, d\eta$ while the integral expression for $\bar{\delta}$ has already been given. When these are substituted into the matrix one finds that the scattering operator is

$$\left[\int_{-\infty}^{-r} C e^{A(-\xi - r)} B g(\xi) \, d\xi + D g(\xi) \right].$$

That is, the scattering operator is $\bar{\delta}$. This completes the computation of wave and scattering operators.

Now we give a brief scattering to systems dictionary.

Herein	Lax–Phillips
$L^2(0, \infty, Y)$	D_+^{ρ}
$L^2(-\infty, 0, U)$	D_-^{ρ}
$\mathcal{T}_-(t)^*$	$\mathcal{T}(t)$ or $U_0(t)$
$\mathcal{T}_+(t)^*$	$\mathcal{T}(-t)$ or $U_0(-t)$
$P_{[-\infty, \rho+s]}$ and $P_{[-\rho-s, \infty]}$	$P_+^{\rho+s}$ and $P_-^{\rho+s}$
$\mathcal{H}, T(t), e^{At}$	$\mathcal{H}, T(t), Z(t) = e^{Bt}$
implicit in the setup	translation representation with $U = Y = N$
imaginary, $j\omega$ axis	real axis

IV. The Basic Properties of Systems

There are three basic theorems which hold for finite dimensional systems: there is a controllable and observable system with a given FRF, two such systems are equivalent in a certain sense, every 'reasonable' function is the FRF for some system. These three properties when suitably interpreted hold in infinite dimensions as is described in the first three parts of this section. The fourth part shows that for meromorphic systems the poles of the FRF correspond to eigenvalues of A as a consequence of Lax–Phillips theory. It also describes eigenvalue behavior for another natural class of systems. Throughout this section we work only with continuously controllable and observable systems.

A. Canonical Decomposition of a System

We shall see that any system can be 'cut down' to a controllable observable one having the same Hankel operator (or FRF if the original system is compatible). Here is the procedure. Let $X_c = cl$ range $\mathcal{C}$ and $X_0 = cl$ range $\mathcal{Q}$ for the system $[A, B, C, D]$. Let $X_{\bar{c}}$ and $X_{\bar{0}}$ denote the orthogonal complement of X_c and X_0, respectively.

Now we describe how to restrict the system $[A, B, C, D]$ to X_c and obtain a system $[A_c, B_c, C_c, D_c]$ which acts on X_c. Since $e^{At} X_c \subset X_c$, its restriction to X_c is a semigroup and consequently has an infinitesimal generator denoted A_c whose domain $\mathfrak{D}(A_c)$ is contained in $\mathfrak{D}(A)$. Set $D_c = D$. The operator C_c is defined on $\mathfrak{D}(C) \cap X_c$ by $C_c x = Cx$. The operator C_c is $\| \ \|_{A_c}$ continuous. B_c is first defined as a linear functional on $\text{proj}_{X_c} \mathfrak{D}(A^*)$ by

$$[B_c u] (w) = [Bu] (w)$$

and then extended to w in $\mathfrak{D}(A_c^*)$ which is shown to be possible by the next two lemmas.

Lemma 4.1: If $u \in U$, there is a constant $\kappa(u)$, so that the linear functional Bu satisfies

$$|[Bu] (w)| \leqslant \kappa(u) \|\text{Proj}_{X_c} w\|_{A_c^*}$$

for each w in $\mathfrak{D}(A_c^*)$.

Proof: For $\rho < 0$ set $x_\rho(s) = \mathcal{C}(\chi_{[0,s]} e^{-\rho t} u)$; then $(x_\rho(s), x) = \int_0^s (Be^{-\rho t} u(t), e^{A^* t} x) \, dt$. Since A^* generates a uniformly bounded semigroup range $A^* - \rho = X$, see [33, Theorem 2.3.2]. If we set $x = (A^* - \rho) w$, the integral becomes $[Bu] \cdot (\{e^{A^* - \rho s} - I\}w)$. Since $\|e^{(A^* - \rho)t} w\|_{A_x} \to 0$ we get $\lim_{s \to 0} \cdot (x_\rho(s), x) = - [Bu] (w)$. Thus $x_\rho(s)$ converges weakly to a vector $x(\infty)$ which is in X_c because each $x_\rho(s)$ is. We have the estimate $|[Bu] (w)| \leqslant \|x(\infty)\| \|\text{Proj}_{X_c} (A^* - \rho)w\|$ from which the lemma follows.

Lemma 4.2: $\text{Proj}_{X_c} \mathfrak{D}(A^*)$ is contained and is dense in $\mathfrak{D}(A_c^*)$.

Proof: Write $e^{A^* t}$ as a 2×2 operator entried matrix with respect to the decomposition $X_c \oplus X_{\bar{c}}$ of X, apply it to a vector $\binom{x}{0}$ in X_c and integrate

$$f(x, s) \triangleq \int_0^s \begin{pmatrix} e^{A_c^* t} & 0 \\ g(t) & h(t) \end{pmatrix} \binom{x}{0} \, dt.$$

It is straightforward to check that $f(x, s)$ is in $\mathfrak{D}(A^*)$. Clearly $\text{Proj}_{X_c} \mathfrak{D}(A^*) \subset \mathfrak{D}(A_c^*)$ and so $p(x, s) \triangleq \text{Proj}_{X_c} f(x, s)$ is in $\mathfrak{D}(A_c^*)$. Since $e^{A_c^* t} x$ is norm continuous with limit x at $t = 0$, we get $p(x, s)/s \to x$, moreover, if $x \in \mathfrak{D}(A_c^*)$, then $A_c^* p(x, s)/s \to A_c^* x$. Thus $p(x, s)/s \to x$ in the $\| \ \|_{A_c^*}$ topology.

The first lemma guarantees that B_c once defined on $\text{Proj}_{X_c} \mathfrak{D}(A^*)$ is defined on a dense subspace of $\mathfrak{D}(A^*)$ while the second says that B_c has a continuous extension to $\mathfrak{D}(A_c^*)$.

The various properties a system might have are preserved under his cut down operation. One can check in a straightforward way that since $e^{At} X_c \subset X_c$, the operator $(z - A_c)^{-1}$ equals $(z - A)^{-1}$ restricted to X_c; moreover, $(z - A_c)^{-1} B_c = (z - A)^{-1} B$ for Re $z > 0$. This is enough to show that the function $C(z - A)^{-1} (s - A)^{-1} B$, which determines the Hankel operator for the original system equals $C_c(z - A_c)^{-1} (s - A_c)^{-1} B_c$; thus the two Hankel operators are equal. If the original system is compatible $(z - A)^{-1} B$ is contained in $\mathfrak{D}(C)$ as well as X_c; thus it is contained in $\mathfrak{D}(C_c) = \mathfrak{D}(C) \cap X_c$ and the new system is compatible. The two FRF's are clearly the same. If an operator-valued function is meromorphic or compact valued its restriction to an invariant subspace also has these properties; thus the meromorphic and compactness property are preserved.

The same type of process works with observability and projecting $[A, B, C, D]$ down to X_0. One proof relies on the fact that Q is the controllability map for the adjoint system $[A^*, C^*, B^*, D^*]$; the preceding construction applies. Then one can take adjoints to obtain $[A_0, C_0, B_0, D]$, an observable system. If one combines these two processes, one gets (by first decomposing the system $[A, B, C, D]$ into controllable and uncontrollable parts and then decomposing these systems into obserable and unobservable parts).

Theorem 4.3: There is a closed subspace X_{c0} of the state space X of a system $[A, B, C, D]$ so that the system $[A_{co}, B_{co}, C_{co}, D]$ gotten by restricting and projecting the original system to X_{co} (as described above) is controllable and observable and has the same Hankel operator as the original system. The system is compatible if the original one is and the FRF is the same for both. The meromorphic and compactness properties are also preserved.

Note: While B_{co} and C_{co} might be regarded formally as restrictions followed by projections A_{co} is defined by the relationship $e^{A_{co} t} = \text{Proj}_{X_{co}} e^{At}|_{X_{co}}$ and this is not the same thing.

B. Two Systems with the Same FRF

This section describes the extent to which a system is determined by its Hankel operator or if the system is compatible with its FRF. First it is necessary to discuss pseudo-inverses of an operator. Suppose that M is an unbounded operator with domain $\mathfrak{F}$ from one Hilbert space H_1 to another H_2, and suppose $\mathfrak{R}$ is a subspace of $\mathfrak{F}$, which does not intersect the null

space of M. We define the $\mathcal{R}$ pseudo-inverse of M, written $M_{\mathcal{R}}^{-1}$, to be the operator G with range equal to $\mathcal{R}$ so that GM (resp. MG) is the identity operator on $\mathcal{R}$ (resp. $M\mathcal{R}$). The main theorem here is

Theorem 4.4: Suppose that two controllable and observable systems $[A, B, C, D]$ and $[\dot{A}, \dot{B}, \dot{C}, \dot{D}]$ have the same Hankel operator. If the two systems are continuously controllable and observable, then there is a possibly unbounded closeable operator M with dense range such that the 'intertwining' equations

$$MA = \dot{A}M \quad MB = \dot{B} \quad \dot{C}M = C \quad D = \dot{D}$$

each hold on a dense set. One such operator M is gotten explicitly by

$$M = (\dot{\mathcal{Q}}_{\mathcal{L}}^{-1})* \; \mathcal{Q} = \dot{\mathcal{C}} \mathcal{C}_{\mathcal{R}}^{-1}$$

defined on range $\mathcal{C}$ whenever $\mathcal{R}$ and $\mathcal{L}$ are controlling and observing spaces for $\mathcal{C}, \dot{\mathcal{C}}$ and $\mathcal{Q}, \dot{\mathcal{Q}}$, respectively. Note that

$$\mathcal{R} = [\text{null } \mathcal{C} = \text{null } \dot{\mathcal{C}}]^{\perp} \text{ and } \mathcal{L} = [\text{null } \mathcal{Q} = \text{null } \dot{\mathcal{Q}}]^{\perp}$$

are one set of spaces for which the theorem applies. Clearly if both systems are continuously exactly controllable or observable, then M is bounded and invertible.

Proof: The continuous operator $\mathcal{Q}* \; e^{Ar} \mathcal{C}$ applied to u in $C_0^\infty[0, \infty, U]$ gives the function $C \; e^{As} e^{Ar} \int_0^\infty e^{At} B \; u(t) \; dt$ of s. Since unlimited integration by parts is possible, there is no problem in interchanging operators with integration to get $\mathcal{Q}* \; e^{Ar} \mathcal{C} \; u(s) = (-1)^n \int_0^\infty C \; e^{A(s+r+t)} (A - \rho)^{-n} B(e^{\rho t} u(t))^{(n)} \; dt$. When $[0, g]$ contains the support of u we may substitute $\tilde{u}(t) = u(g - t)$ for u then change variables and get

$$Q* \; e^{Ar} \mathcal{C} \; \tilde{u}(s) = (-1)^n \int_0^{g+s+r} e^{A(s+r+g-t)}$$
$$\cdot (A - \rho)^{-n} B(e^{\rho(g-t)} u(t))^{(n)} \; dt.$$

This is just the response of the system $[A, B, C, D]$ to input $e^{\rho(g-t)} u$ for values $\geq g$. Thus by hypothesis we have

$$Q* \; e^{Ar} \mathcal{C} \; = \dot{Q}* \; e^{Ar} \; \dot{\mathcal{C}}$$

for $u \in C_0^\infty$ and by continuity it holds for all $u \in L^2[0, \infty, U]$.

From this, one gets if $\mathcal{R}$ and $\mathcal{L}$ are controlling and observing for $\mathcal{C}, \dot{\mathcal{C}}$ and $\mathcal{Q}, \dot{\mathcal{Q}}$, respectively, then

$$(\dot{\mathcal{Q}}_{\mathcal{L}}^{-1})* \; \mathcal{Q} \; e^{Ar}x = e^{\dot{A}r} \; \dot{\mathcal{C}} \mathcal{C}_{\mathcal{R}}^{-1}x$$

for $x \in$ range $\mathcal{C}$. By setting $\kappa = 0$, we see that

$$(\dot{\mathcal{Q}}_{\mathcal{L}}^{-1})* \; \mathcal{Q}* = \dot{\mathcal{C}} \mathcal{C}_{\mathcal{R}}^{-1} \tag{4.1}$$

holds on range $\mathcal{C} = \mathcal{C}_{\mathcal{R}}$, regardless of which $\mathcal{L}$ and $\mathcal{R}$ are used. Define M to be this operator on range $\mathcal{C}$. The construction also works for the adjoint system and yields an operator M^+ defined on range $\dot{\mathcal{Q}}$; the two operators satisfy $(Mx, y) = (x, M^+y)$ for x, y in their respective domains. Thus M has a densely defined adjoint and by [59, ch. VII, sect. 2, Th. 3] is closable. Let $S_A = \{\omega \in \mathcal{D}(A) : \omega = \mathcal{C}u$ for a $u \in C^\infty(0, \infty, U)$ with $\dot{\mathcal{C}}u$ in $\mathcal{D}(\dot{A})\}$. Since any u in $C_0^\infty(0, \infty, U)$ will produce such a vector, S_A is dense. (even in the $\| \; \|_A$ topology) in $\mathcal{D}(A) \cap$ range $\mathcal{C}$: likewise MS_A is dense. By (4.1) the map M takes S_A into $\mathcal{D}(\dot{A})$ and so M as an unbounded operator on $\mathcal{D}(A)$ is closable. Likewise $M*$ restricted to a dense subspace S_{A*} of $\mathcal{D}(A*)$ is closable, etc.

From (4.1), we see that

$$M \; e^{Ar} = e^{\dot{A}r}M$$

on range $\mathcal{C}$ and additionally

$$MA = \dot{A}M$$

on S_A. Furthermore, $M\mathcal{C} = \dot{\mathcal{C}}$. Thus if $u \in U$, the vectors

$$\frac{1}{s} \int_0^s e^{At} Bu \; dt \triangleq b(u, s), \quad \frac{1}{s} \int_0^s e^{\dot{A}t} \dot{B}u \; dt \triangleq \dot{b}(u, s)$$

in range $\mathcal{C}$ and $\dot{\mathcal{C}}$ satisfy $Mb(u, s) = \dot{b}(u, s)$. Now $(x, b(u, s)) \to [Bu] \; (x)$ and $(x, \dot{b}(u, s)) \to [\dot{B}u] \; (x)$ and so $[MBu] \; (x) = [Bu] \cdot (M*x) = [\dot{B}u] \; (x)$ for x in S_{A*}. To see that M intertwines C and $\dot{C}$ observe that $\dot{C} \; e^{\dot{A}s}Mx = \mathcal{Q}*Mx = \mathcal{Q}*x = C \; e^{As}x$ for $x \in S_A$. Set s equal to 0 to obtain the result.

The 'note' and last line of the theorem are straightforward to prove (cf. [28 p. 29]).

Remark 4.5: If the two systems in Theorem 4.4 are embedded in a Lax–Phillips model, then $T(t)$ and $\dot{T}(t)$ are related by $GT(t) = \dot{T}(t)G$ where G in matrix form is

$$\begin{pmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & I \end{pmatrix}.$$

C. System Realizability

This subsection describes how to construct a system with preassigned FRF. An operator-valued function $F(z)$ analytic on the RHP is said to converge to 0 in the positive direction if $\lim_{x \to +\infty} \max y \|F(x + jy)\|$ is 0, it converges to F_∞ if $F(z) - F_\infty$ converges to 0. The function F is unitary inner if $F(jy)$ is a unitary operator for almost all y. Note that if F is analytic on Re $z > \rho$ and is the FRF of a compatible system $C(z - A)^{-1}B$, then the modification $A \to A - \rho - \epsilon$ of the system changes the region of analyticity for FRF to the RHP or greater. Thus as far as the realizability problem is concerned we might as well assume that the (prospective) FRF is well behaved on the RHP.

Theorem 4.5: If $F(z)$ is a uniformly bounded $\mathcal{L}(U, Y)$-valued analytic function on the RHP with a limit $F(\infty)$ in the positive direction, then there is a compatible continuously controllable and observable system satisfying $\|e^{At}\| \leq 1$ with F as its FRF, which

(i) is exactly controllable and observable if and only if Hankel$_F$ has closed range. Note such systems are always asymptotically stable.

(ii) has e^{At} and $e^{\dot{A}t}$ asymptotically stable if $F(jy)$ has either the form $U(jy) \; G(jy)$ or $G(jy) \; U(jy)$ where U is a unitary inner function on the RHP and G is the boundary value of an analytic and uniformly bounded function in the LHP.

(iii) has controllability and observability operators with isometric adjoints if F is unitary inner.

Proof: We begin informally and give an intuitive idea of the construction. Assume that F goes to zero in the positive direction. Let $\Sigma(t)$ denote the shift by t units to the left on $L^2(-\infty, \infty, Y)$. Its infinitesimal generator $\tilde{A}$ is d/dx acting on $\mathcal{D}(\tilde{A}) = \{u \in L^2(-\infty, \infty, Y):$ the derivative of u is in $L^2(0, \infty, Y)\}$. The adjoint of $\tilde{A}$ is $-d/dx$ acting on the same domain. The inverse Fourier transform of Fu is not in general a func-

tion, but it can be naturally viewed as a linear functional on $\mathfrak{D}(\tilde{A}^*) = \mathfrak{D}(A)$; namely, if $\mathcal{F}$ denotes the ordinary Fourier transform on L^2 functions, set $[\tilde{B}u](f) = \int_{-\infty}^{\infty} F(j\xi) \mathcal{F}f(\xi) d\xi$ where $f \in \mathfrak{D}(A^*)$ and B is the generalized Fourier transform of F. Note that the integral exists because $\mathcal{F}f(\xi)/(1 + |\xi|^2)$ is in $L^2(-\infty, \infty, Y)$. A simple estimate (cf. [53, Th. 7.25]) shows that every function in $\mathfrak{D}(\tilde{A})$ is continuous, thus we may define a map $\tilde{C}: \mathfrak{D}(\tilde{A}) \to Y$ by $\tilde{C}h = h(0)$ for h in $\mathfrak{D}(\tilde{A})$. The $\tilde{A}, \tilde{B}, \tilde{C}$ constitute a system. If the Fourier transform of F were a continuous function, then $[\Sigma(t)\mathcal{F}F](x) = \mathcal{F}F(t + x)$ and so the impulse response function $\tilde{C}e^{\tilde{A}t}\tilde{B}u = [\mathcal{F}Fu](t)$ is $\mathcal{F}F$, as required.

Now we turn to a procedure which works in general and rigorously. Everything is set in the frequency domain. First Fourier transform $\tilde{A}, \tilde{B}, \tilde{C}$:

$$\mathcal{F}^{-1}(z - \tilde{A}) = z - j\xi$$

$$[\mathcal{F}^{-1}\tilde{B}u](g) = \int_{-\infty}^{\infty} \mathcal{F}F(j\xi)u\,g(\xi)\,d\xi$$

$$\tilde{C}\mathcal{F}g = [\mathcal{F}g](0) = \lim_{M \to \infty} \frac{1}{\sqrt{2\pi}} \int_{-M}^{M} g(\xi)\,d\xi$$

while this is a system it is not necessarily compatible so we extend C to be defined on $\mathfrak{D}(C) = \{g : R \to Y | \mathcal{F}^{-1}g$ has finitely many poles in the RHP and $\xi \mathcal{F}^{-1}g(\xi)$ goes to zero in the positive direction$\}$. This system is compatible because $\mathcal{F}^{-1}(z - \tilde{A})^{-1}\tilde{B}u = (z - j\xi)^{-1}F(j\xi)$ is in $\mathcal{F}\mathfrak{D}(C)$ and its FRF is

$$\lim_{M \to \infty} \frac{1}{2\pi} \int_{-M}^{M} (z - j\xi)^{-1} F(j\xi)\,d\xi.$$

This is one boundary of integration over the contour $\{z : |z| = M, \text{Re } z > 0\} \cup \{z : -M \leqslant z \leqslant M\}$. The integral over the circular boundary goes to zero as $M \to \infty$ because F goes to zero in the positive direction. Thus the FRF equals the residue of the contour integral at z. In other words it is $F(z)$. If F goes to $F_\infty \neq 0$ in the positive direction set $D = F_\infty$. The system $\tilde{A}, \tilde{B}, \tilde{C}, D$ is compatible and has FRF equal to F.

The system may not be controllable or observable in any sense. Now we remedy that. The controllability operator is for each input function $u \in L^2(0, \infty, U)$ and $f \in \mathfrak{D}(A^*)$

$$[\mathcal{C}u](f) = \int_0^\infty [\tilde{B}u(t)](\Sigma(t)^*f)\,dt$$

$$= \int_{-\infty}^{\infty} (F(j\xi)[\mathcal{F}u](\xi), [\mathcal{F}f](\xi))\,d\xi.$$

By the Plancherel theorem the formula extends to all f in $L^2(-\infty, \infty, Y)$ and u in $L^2(-\infty, \infty, U)$. Thus for such u, the function $F(j\xi)[\mathcal{F}u](\xi)$ is in $L^2(-\infty, \infty, Y)$. By the closed graph theorem the system is continuously controllable. The observability map when applied to y in $L^2(0, \infty, Y]$ is $\mathcal{Q}y = \int_0^\infty \Sigma(t) \delta(x) y(t)\,dt$ which says that $\mathcal{Q}$ is the identity map. Since the system is continuously controllable and observable, Theorem 4.3 implies that the system A, B, C, D obtained by restricting and projecting $\tilde{A}, \tilde{B}, \tilde{C}, D$ to the closure of the range of the Hankel operator associated with F is compatible and has FRF equal to F.

By construction $\mathcal{C} = \text{Hankel}_F$, so the system $[A, B, C, D]$ is exactly controllable if and only if range Hankel_F is closed. As mentioned near the end of Section I the operator $\mathcal{Q}*\mathcal{C}m$ is unitarily equivalent to Hankel_F. Thus $\|\mathcal{Q}*x\| \geqslant \epsilon \|x\|$; consequently, range $\mathcal{Q}$ is closed and we have exact observability. Conversely, for any continuously exactly controllable and observable system $\mathcal{Q}*\mathcal{C}$ has closed range; thus, the associated Hankel operator has closed range. An argument like that in [28, Remark 4.2] demonstrates that continuous exact controllability [resp. observability] forces e^{A^*t} [resp. e^{At}] to be asymptotically stable.

To prove (ii), suppose that the first factorization holds. Then range $\text{Hankel}_U \supset \text{Hankel}_F$ and so if x belongs to the range of the latter it can be written $x = \text{Hankel}_U w$ for some w in $L^2(-\infty, 0, Y)$. Now $\Sigma(t)^*x = \text{Hankel}_U g(t)$ where $g(t) = P_{L^2[-\infty, 0, Y]} \Sigma(t)^*w$. Since $g(t) \to 0$ as $t \to \infty$, we have $\Sigma(t)^*x \to 0$, that is $T(t)^*x \to 0$. Note $T(t)x = P_{L^2(0, \infty, Y)} \Sigma(t)x$ automatically goes to 0. If the reverse factorization holds, then a similar argument with the adjoint system proves stability.

If $F(j\xi)$ is a unitary operator then it is straightforward to show that Hankel_F is a partial isometry. An argument like that on [28, p. 32] finishes the theorem.

D. Poles and Eigenvalues

When X is finite dimensional the poles of the FRF for a controllable and observable system are located at the eigenvalues of A, in other words, input and output information completely determine the eigenvalues of A. The importance of pole location and of this fact in finite dimensions is well known. For infinite-dimensional situations the author is more familiar with scattering theoretic implications than engineering ones (possibly because of background and possibly because efforts in the first area are more advanced). The eigenvalues of A correspond to modes which are decaying in the inside of the system and thus poles of the scattering matrix give information on this. In different cases the poles have different interpretations. Much of Section V is devoted to describing this.

When X is infinite dimensional the pole to eigenvalue correspondence is precise for exactly controllable and observable systems as we shall soon demonstrate. However, in the controllable and observable case (which is the most common) the situation is murkier. The problem is that two systems with the same FRF have operators A and $\dot{A}$ which are merely intertwined by a 1–1 densely defined M having dense range. This is a very weak relationship and it is well known that even when A and $\dot{A}$ are bounded operators no correspondence between the spectra of A and $\dot{A}$ can be inferred (cf. [49, ch. VI, sect. 4.2]). Thus if a system had spectrum in correspondence with pole location one might be able to find an M so that the equivalent system produced would have completely different spectrum.

All of this pessimism is not justified, because as we shall see the systems which come up in practice behave well. This is a consequence of the work dating back to Moeller and used for this purpose by Lax–Phillips, who prove the following.

Theorem 4.7: A compatible continuously controllable and observable system with the meromorphic property has a FRF whose poles are located exactly at the eigenvalues of A.

Proof: This is a rephrasing of [38, Th. 5.5] applied to the Lax–Phillips model in which the system is imbedded (see Section III-C).

To apply their [38, Th. 5.5], we require in addition to the meromorphic property a hypothesis on 'incoming and outgoing vectors'. Actually an immediate consequence of this additional hypothesis [38, Lemma 5.3] suffices. In systems theory language, it says that if x is an eigenvector of A [resp. A^*], then $Cx \neq 0$ [resp. $B^*x \neq 0$]. The statement about C is a weaker requirement then observability, while the one on B follows from controllability. The conclusion of their theorem says that the poles of the s-matrix are located at precisely the eigenvalues of A. The FRF for a system is just the s-matrix of the Lax–Phillips model in which it is imbedded and so Theorem 4.7 follows.

A compatible meromorphic system has a meromorphic FRF. A more general class of FRF is mathematically reasonable to consider, namely, the *pseudomeromorphic F(z)* which are those having a factorization $F(jy) = G(jy) V(jy)$ or equivalently $V_1(jy) G_1(jy)$ with the G's analytically continuable to a bounded function analytic in the LHP and the V's unitary inner on the RHP. These occurred in Theorem 4.6, part b. When U and Y are finite dimensional another definition of pseudomeromorphic is possible [13]. There is a function $H(z)$ meromorphic and of bounded type in the LHP so that $\lim_{x \to 0} H(x + jy) = F(jy)$ exists for almost all y. Thus F is meromorphic except on the jy-axis where it could have bad discontinuities. It is in these finite-dimensional circumstances where pseudomeromorphic FRF's are particularly interesting since they turn out to be precisely those function for which there is any correspondence between properties of A and properties of the FRF. Thus *they are the biggest class of functions that one would probably want to use in connection with classical A, B, C, D type systems theory*. This point is well illustrated by the type of system described in Section IV-A gotten by restricting the shift to a certain subspace. Such a restricted shift system, as Fuhrmann calls it, has a structure which can be immediately analyzed by standard mathematical theory arising from Moeller's theorem [26], [37], [49].

Theorem 4.8: The singularities of a pseudomeromorphic FRF for a restricted shift system with U and Y finite dimensional are equal precisely to the spectrum of A. The spectrum of A with Re $z < 0$ consists of isolated eigenvalues.

If the FRF is not pseudomeromorphic, then the spectrum of $e^{At} = \{z : |z| \leqslant 1\}$.

Proof: It suffices to work on the unit disk rather than the RHP since a standard transformation (Caley) connects the two. If F is a unitary inner function then the theorem is essentially Moeller's theorem (cf. [37, ch. 3, sect. 3] or [49, ch. VI, Th. 4.1]) and so our proof consists of reducing Theorem 4.8 to the unitary inner case. Suppose that F is pseudomeromorphic. Then range Hankel$_F$ is a subspace of $H^2(Y)$ whose orthogonal complement γ is invariant under multiplication by z. By the Lax–Beurling theorem there is an 'inner' function W such that $\gamma = WH^2(Y)$. The pseudomeromorphic property forces W to have V as a factor and so W is unitary inner; this is equivalent to F being pseudomeromorphic. Set $G(e^{i\theta}) = W(e^{i\theta})^*F(e^{i\theta})$. Observe if f is any function in $\overline{H}^2(U)$, then $Ff \in \gamma^\perp + \overline{H}^2(Y)$ and so W^*Ff is in $\overline{H}^2(Y)$. This forces G to belong to $\overline{H}^\infty(U, Y)$. We have $F = WG$ where by construction cl range Hankel$_F$ = range Hankel$_W$. This latter property implies that there is no nonconstant inner function W_1 which divides both W^* and G from the left, i.e., $W_1 K = W^*$ and $W_1 L = G$ for some analytic K and L. By construction e^{At} is just the shift restricted to Hankel$_F$ = Hankel$_W$. The Moeller theorem says that the spectrum of A behaves precisely as

required by the theorem but with W replacing F. Thus it remains to show that the singularities of W equal those of F. Clearly any singularity of F is a singularity of W. Note that in the nonpseudomeromorphic case W is not *unitary* inner and a refined version of the Moeller theorem gives the last statement in the theorem.

Now we show that every singularity of W is also a singularity of F. Obviously it will depend heavily on the fact that W^{-1} and G have no common (left) factor which could cause the cancelation of a singularity. This indivisibility property has the geometric interpretation that the sum S of the subspaces $W^{-1} \overline{H}^2(Y)$ and $G\overline{H}^2(U)$ is dense in $\overline{H}^2(Y)$, because $cl\,S$ being (backwards) translation invariant can by the Lax–Beurling theorem be written $cl\,S = \theta \overline{H}^2(Y)$ for θ a conjugate analytic inner function. Since S contains both $W^{-1} \overline{H}^2(Y)$ and $G\overline{H}^2(U)$ the function θ divides both W^{-1} and G from the left and, therefore, θ = constant or equivalently $cl\,S = \overline{H}^2(Y)$. Since S consists of vector functions of the form

$$n = W^{-1} c + Gd \qquad (4.2)$$

where $c \in \overline{H}^2(U)$ and $d \in \overline{H}^2(Y)$ the space $\mathcal{R} \triangleq \{c + Fd\}$ is dense in $\mathcal{R}' \triangleq W\overline{H}^2(Y)$. We shall show that this can happen only if W has no more singularities than F.

For the sake of a clear presentation let us do the dim $Y = 1$ case first; then W and G are scalar valued. Let us suppose that W has a singularity at z_0 while F does not. Note that since the singularity set for F is a closed subset of $\{z : |z| \geqslant 1\}$, z_0 is at some positive distance from any singularity of F and so F is uniformly bounded in some region about z_0. A classical factorization theorem due to Buerling (see [26]) implies that $w = m \cdot p$ where both m and p are inner, m has no singularity in common with F and p is bounded in a neighborhood of z_0. Since $W^{-1} F$ is in $\overline{H}^2(\mathcal{C})$ and m and F have no common singularities $p^{-1} F$ is in $\overline{H}^2\mathcal{C}$. Thus from (4.2) every mn is in $\overline{H}^2(\mathcal{C})$ which is impossible since the n are dense and m is a nonconstant function in $H^2(\mathcal{C})$.

After setting things up properly this same type of reasoning will apply to the dim $Y = K > 1$ case. Suppose there is a z_0 as before. Then there will be some entry, say w_{11}, of the matrix function W which is singular at z_0. The functions w_{1l} factor as before into $m_l p_l$. Now

$$\mathcal{R}_1 = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \cdot & \cdot & \cdot & \cdots \end{pmatrix} \mathcal{R}$$

is contained in

$$\mathcal{R}_1' = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \cdot & \cdot & \cdot & \cdots \end{pmatrix} \mathcal{R}'$$

which implies that for an entry f_{1l} of F the function $m_1^{-1} m_2^{-1} \cdots m_k^{-1} p_1^{-1} \cdots p_k^{-1} f_{1l}$ is in $\overline{H}^2(\mathcal{C})$ and consequently $p_1^{-1} \cdots p_k^{-1} f_{1l}$ is in $\overline{H}^2(\mathcal{C})$. Since $\mathcal{R}_1'$ is dense in $\mathcal{R}_1$, one has $w_{11} = n_1 p_1 \in clR_1'$. However, this forces $m_1 p_2^{-1} \cdots p_k^{-1}$ to be in $\overline{H}^2(\mathcal{C})$ which is impossible, since m_1 is singular at z_0 and the p_l^{-1} cannot cancel this singularity. The proof is complete.

As an aside it might be informative to observe how Theorem 4.7 follows from Theorem 4.8 for U, Y finite dimensional (and otherwise with slightly more work). Let $[A, B, C, D]$ satisfy the hypothesis of Theorem 4.7; with FRF denoted F. Let $\dot{A}$ arise from the restricted shift realization of F. By Section IV-B, there is M with $AM = M\dot{A}$. Since the restricted shift system is exactly controllable the inverse of its control-

lability operator and therefore M is a bounded operator (by the formula for M). If $\dot{A}x = \lambda x$, then $AMx = \lambda Mx$ so λ is an eigenvalue of A. Thus every pole of F is an eigenvalue of A. The reverse direction follows in an elementary fashion from $(z - A)^{-1}$ meromorphic together with the $(\lambda - A)w = 0 \Rightarrow Cw \neq 0$ and $(\bar{\lambda} - A^*)w' = 0 \Rightarrow B^*w' = 0$ property.

The subsection concludes with some historical remarks. A correspondence between the spectrum of A and poles of the FRF was widely observed once the connection between systems theory and the work of Lax–Phillips, Nagy–Foias, Helson, etc. was understood and provided many new techniques, [12], [18], [27]. The first thorough presentation for the nonboundary singularities is [12]. The pseudomeromorphic functions were studied first by mathematicians [14]. An equivalent notion called 'roomy' was introduced independently by Dewilde who asserted strongly that they were the proper class for the study of systems [11]. He also asserted that certain 'roomy' functions were appropriate for the electrical design technique called Darlington synthesis. In [13] the equivalence of 'roomy' with [14] was established and they were shown to be precisely the class for which Darlington synthesis is appropriate; the result was independently done in Russia by Arov [1]. Meanwhile Fuhrmann in his work on systems theory devoted quite a bit of attention to pseudomeromorphic functions [22], [23]. In particular he concentrated on a subclass of functions which are precisely those which can be the FRF of an exactly controllable and observable system. It was observed by Clark and Helton [28, Th. 3c.1 plus Remark 3c.2] that in the scalar discrete-time case this class has a very nice characterization and Fuhrmann extend this to the multivariable and certain real time cases [19], [21]. In the author's opinion the next necessary step in this direction to find lossy exactly controllable and observable systems which arise physically.

V. MISCELLANEOUS RESULTS IN SCATTERING THEORY

This section has a different spirit than the rest of the paper and is not related to systems *per se*. It lists some qualitative results obtained in scattering theory over the last five years. The hope is that this will benefit engineers primarily by suggesting structure which they did not know. Most surely the results as they stand will not apply without modification to someone's particular problem. However, practically all of the results are flexible and nobody knows their ultimate generality. The mathematicians involved have not bothered to do relatively straightforward generalizations because motivation was lacking. Thus if a particular fact listed herein catches a readers fancy he is urged to consult the source article for a precise statement of the result and if the context in that article is not quite right he should not dispair since the result will probably hold in many physically reasonable contexts. We emphasize that the listing here is not intended to be complete (or precise) and concentrates on work done primarily by people in the Lax–Phillips branch of scattering theory.

We begin by describing the commonly studied situations. There is *obstacle* scattering for the wave equation in $1, 3, 5, \cdots$ dimensions. Here one has an obstacle and bounces waves off of the obstacle usually in order to determine its shape. (Dimensions $2, 4, \cdots$ are highly anomalous, see [41], and will not be discussed here.) Variable coefficient scattering usually treats an equation

$$u_{tt} = Lu = a(x)\nabla \cdot A(x)\nabla u - b(x)u \qquad (5.1)$$

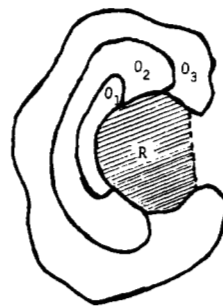Fig. 2.

where all coefficients are smooth and nonnegative—$A(x)$ is a positive definite matrix—and we assume $Lu = \Delta u$, for $|x| > R$. The objective is to study the coefficients in terms of the scattering matrix. The one-dimensional case includes the lossless transmission line, e.g., $b = 0$, $a = (1/c)$, $A\,1/l$. Very recently rudimentary Lax–Phillips scattering has been applied to sound impinging on an obstacle with a 'springy' elastic surface [8]. In one dimension this corresponds to a singly infinite constant coefficient transmission line connected to a circuit with one resister, capacitor and inductor at the finite end. We reiterate that many of the results which follow deal with or extend easily to mixtures of the first and second situation (the third has not been studied yet). In fact for many purposes an obstacle problem can be thought of as a limit of a variable coefficient problem as appropriate coefficients get infinitely large. Also many of the results will most likely have an analog for lossy situations. The preoccupation with losslessness derives from the fact that in classical physics most phenomena are energy conserving; a general 'lossy' scattering theory did not emerge until 1973 [38]. Another point is that although the one dimensional case will be emphasized through the exposition this is entirely for concreteness. Practically all structure is greatly simpler for dimension one than for higher dimensions. Thus it is regarded as trivial and frequently ignored in the works to be described; three-dimensional structure, however, is a good guide to behavior in all odd dimensions.

One general class of results concerns how close poles of a scattering matrix can come to the imaginary axis.

1) According to Beale [7], suppose that a sequence of obstacles O_n gets closer and closer to enclosing a region R (see Fig. 2) and that the Laplacian on R with zero boundary conditions has eigenvalues $e_1, e_2, \cdots$. Then each corresponding scattering matrix $s_n(z)$ has a sequence of poles $p_1^n, p_2^n, \cdots$ such that $p_k^n \to j\,e_k$. This result will almost surely apply to the variable coefficient case, say for example to a transmission line with variable capacitance having a huge well. Very crudely speaking it describes an effect that pockets of very high signal speed have on poles near the imaginary axis.

2) For which situations does the scattering matrix have all of its poles in Re $z \leqslant -\delta < 0$ for some δ? A stronger question for the corresponding system is when is e^{At} exponentially stable, i.e., $\|e^{At}\| \leqslant e^{-\delta t}$? Before describing general theorems, we list some special cases where exponential stability holds: when the obstacle is starlike (any point inside the obstacle can be joined to the origin with a line lying inside the obstacle), when the variable coefficients restricted to any ray through the origin decrease with distance from the origin. The principle behind this can be thought of in terms of geometrical optics. There are high frequency solutions of (5.1) with $b = 0$ which can be thought of as propagating along rays—called bichar-

acteristic curves of L. They are the rigorous version of 'light rays' and bend according to the laws which govern light rays going through a changing medium. For the general variable coefficient case one has exponential decay provided that any bicharacteristic curve leaves any finite region in a finite amount of time, [37, Appendix 2], [41]. This bicharacteristic condition also implies [40] that the poles of $s(z)$ are in the region

$$\operatorname{Re} z \leqslant -(a + b \log |z|)$$

for some $a > 0$. In dimension one rays can never be trapped in a finite region and so one always has exponential decay. For the obstacle case one fixes a sphere around the obstacle Ω, draws a line to the obstacle, optically reflects it, follows it to the obstacle again, optically reflects it, etc. until the ray crosses the sphere (on its way out). Let l be the length of the ray which you take to be ∞ if the path requires infinitely many reflections. Let $L(\Omega)$ be the maximum of all lengths l possibly obtained this way. The conjecture ([37, p. 155]) is that (a) if $L(\Omega) < \infty$, then e^{At} decays exponentially, (b) if $L(\Omega) = \infty$, then $\|e^{At}\| = 1$ for all t. In other words one suspects the same principle is behind both the variable coefficient and the obstacle case. A main difficulty with the present state of the art is that rays hitting the obstacle tangentially can produce some incredible mathematical pathology. Ralston [50] verified (b) while (a) has been proved for a large class of obstacles, most recently [47], [58].

If the operator e^{At} is a compact operator for t sufficiently large, then e^{At} is exponentially stable ([37, Cor. 5.1]). This stronger compactness condition is frequently satisfied, for example, if the obstacle is convex [48] or if the bicharacteristic condition holds [53].

Another class of results describes how close poles get to the origin. Given $\rho > 0$ there exists a region N in the complex plane containing the interval $[-\pi/\rho, \pi/\rho]$ so that the scattering matrix for any obstacle which is contained in the sphere about the origin of radius ρ has no poles in N; see [42].

A third class of results concerns how the scattering matrix changes when the coefficients are changed. This type of work goes under the heading of 'sensitivity analysis' in engineering literature.

1) Suppose the obstacle is starlike or that a and A decrease monotonically on rays through the origin. Then as a decreases and A increases the poles of $s(z)$ which are located on the real axis slide out the real axis. Obstacle case [37]-[42]; variable coefficient case [50].

2) An eigenvalue of $s(jw)$ equals $e^{j\delta_k(w)}$ where δ_k takes real values and is called the *phase shift*. If the obstacle increases in size, or if A, a, and b/a increase, then $\delta_k(w)$ increases. Also under starlikeness and radial monotonicity assumptions as above, the derivative $(d/dw) \delta_k(w)$ is positive. For the singly infinite transmission line (open- or short-circuited finite end) this says that for fixed frequency input the phase of the output is shifted by an amount which increases when the capacitance or inductance of the line decreases. The last fact is classical physics, the first two facts are in [32].

Quite a different topic is that of controllability and observability. It has recently come of interest to mathematicians in this group. Majda [46] studied observability of the system which arises (in the same way as the transmission line Section II) from scattering for (5.1) with $A = 1/a$ in the extension of an obstacle Ω with boundary conditions $\partial u/\partial n + \gamma(x) u_t + \sigma u = 0$ on the boundary 2Ω of the obstacle. He proves when A is constant near $\partial \Omega$ then if $\gamma(x)$ never equals 1 on $\partial \Omega$, the system is observable. In the one-dimensional case this says that if the transmission line in Section II is terminated in resistance R and has coefficients constrained as above (more than necessary), then the system associated to it in Section II is not observable if and only if R is the characteristic impedance of the line.

REFERENCES

[1] D. Z. Arov, "Darlington's method for dissipative systems," *Sov. Phys., Doklady*, vol. 16, pp. 954–956, 1972.
[2] A. Adamajan and D. Z. Arov, "On unitary couplings of semiunitary operators," *Trans. Math. Monogr. AMS*, vol. 95, pp. 75–129, 1970.
[3] J. P. Aubin and A. Bensoussan, "Models of representation of linear time invariant systems in continuous time," *MRS Tech. Rep.* 1286, Sept. 1973.
[4] A. V. Balakrishnan, "On the state space theory of linear systems," *J. Math. Anal. Appl.*, vol. 14, pp. 371–391, 1966.
[5] J. Baras and R. Brockett, "H^2 functions and infinite dimensional realization theory," *SIAM J. Control*,
[6] J. Baras, R. Brockett, and P. Fuhrmann, "State-space models for infinite dimensional systems," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 693–700, Dec. 1974.
[7] T. Beale, "Scattering frequencies of resonators," to appear in *Commun. Pure Appl. Math.*
[8] J. T. Beale and S. J. Rosencrans, "Acoustic boundary conditions," *Bull. AMS*, vol. 80, no. 6, pp. 1276–1279, Nov. 1974.
[9] M. S. Brodski, "Triangular and Jordan representations of linear operators," *Trans. Math. Monogr. AMS*, vol. 32.
[10] C. T. Chen, *Introduction to Linear Systems Theory*. New York: Holt, 1970.
[11] P. Dewilde, "Roomy scattering matrix synthesis," Univ. California, Berkeley, Tech. Rep. 1971.
[12] —, "Input–output description of roomy systems."
[13] R. G. Douglas and J. W. Helton, "Inner dilations of analytic matrix functions and Darlington synthesis," *Acta Sci. Math.*, vol. 34, pp. 61–67, 1973.
[14] R. G. Douglas, H. S. Shapiro, and A. L. Shields, "Cyclic vectors and invariant subspaces for the backward shift operator," *Ann. Inst. Fourier*, vol. 20, pp. 37–76, 1970.
[15] N. Dunford and J. Schwartz, *Linear Operators*, vol. 1. New York: Interscience, 1967.
[16] A. V. Efimov and V. P. Potapov, "J-expanding functions and their role in the analytic theory of electrical circuits," *Russ. Math. Surveys, Uspeki*, vol. 28, no. 1, 1973.
[17] P. A. Fuhrmann, "On realization of linear systems and applications to some questions of stability," to appear in *Math. Syst. Theory*.
[18] —, "Realization theory in Hilbert space for a class of transfer functions," *J. Math. Anal. Appl.*
[19] —, "Exact controllability and observability and realization theory in Hilbert space," *J. Math. Anal. Appl.*
[20] P. A. Fuhrmann, "On sums of Hankel operators," to appear in *Proc. AMS*.
[21] —, "On generalized Hankel operators induced by sums and products."
[22] —, "On Hankel operator ranges, meromorphic pseudo-continuations and factorization of operator valued analytic functions," to appear.
[23] —, "Factorization theorems for a class of bounded measurable functions," to appear.
[24] —, "On series and parallel coupling of a class of infinite dimensional systems."
[25] I. M. Gelfand, *Generalised Functions*. New York: Academic Press, pp. 1964, 1968.
[26] H. Helson, *Invariant Subspaces*. New York: Academic Press, 1964.
[27] J. W. Helton, "The characteristic functions of operator theory and electrical network realization," *Indiana J.*, vol. 22, no. 5, 1972.
[28] —, "Discrete time systems, operator models, and scattering theory," *J. Funct. Anal.*, vol. 16, no. 1, pp. 15–38, May 1974.
[29] —, "Operator theory techniques for distributed systems," in *Conf. Rec. 11th Allerton Conf.*, pp. 131–140, 1973.
[30] —, "A spectral factorization approach to the distributed stable

regulator problem; the algebraic Riccati equation," to appear in *SIAM J. Contr.*

[31] F. J. Helton and J. W. Helton, "Scattering theory for computers and other nonlinear systems," to appear.

[32] J. W. Helton and J. V. Ralston, "The first variation of the scattering matrix," to appear.

[33] E. Hille and R. S. Phillips, "Functional analysis and semigroups," *AMS Colloq.*, vol. 31.

[34] E. Kamen, "On an algebraic theory of systems defined by convolution operators," to appear in *Math Syst. Theory*.

[35] —, "Topological module structure of linear continuous time systems."

[36] —, "Applications of modern algebra to infinite dimensional systems."

[37] P. D. Lax and R. S. Phillips, *Scattering Theory*. New York: Academic Press, 1967.

[38] —, "Scattering theory for dissipative hyperbolic systems," *J. Funct. Anal.*, vol. 14, pp. 172-235, 1973.

[39] —, "Decaying modes for the wave equation in the exterior of an obstacle," *Commun. Pure Appl. Math.*, vol. XXVII, pp. 737-787, 1969.

[40] —, "A logrithmic bound on the location of the poles of the scattering matrix," *Arch. Inst. Mech. Anal.*, vol. 40, no. 4, pp. 268-280, 1971.

[41] —, "Scattering theory for acoustic equation in an even number of space dimensions," *Indiana J.*, vol. 22, pp. 101-134, 1972.

[42] —, "On the scattering frequencies of the Laplace operator for exterior domains."

[43] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*. New York: Springer, 1971.

[44] M. S. Livsic, "Operators oscillations, waves, open systems," *AMS Trans. Monogr.*, vol. 34, 1973.

[45] —, "Theory of operator nodes on Hilbert spaces," 1971, in Russian.

[46] A. Majda, "Disappearing solutions for the dissipative wave equation, to appear.

[47] C. Morawetz, "Decay for solutions of Dirchlet exterior problem for the wave equation."

[48] C. Morawetz and D. Ludwig.

[49] B. Sz-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Space*. Amsterdam, The Netherlands: North Holland, 1970.

[50] J. V. Ralston, "Solutions of the wave equation with localized energy," *Commun. Pure Appl. Math.*, vol. XXII, pp. 807-823, 1969.

[51] —, "Trapped rays in spherically symmetric media and poles of the scattering matrix," *Commun. Pure Appl. Math.*, vol. XXIV, pp. 571-582, 1971.

[52] —, "Variation of the transmission coefficient and comparison theorems for the purely imaginary polar of the scattering matrix," *Commun. Pure Appl. Math.*, vol. XXV, pp. 45-61, 1972.

[53] —, "Local decay of solutions of conservative first order hyperbolic systems in odd dimensional space," *Trans. AMS*, vol. 194, 1974.

[54] —, "On the propagation of singularities in solutions of symmetric hyperbolic partial differential equations."

[55] M. Reed and B. Simon, *Methods of Modern Mathematics and Physics: Functional Analysis*. New York: Academic Press, 1972.

[56] F. Riesz and B. Sz. Nagy, *Functional Analysis*. New York: Ungar, 1955.

[57] W. Rudin, *Functional Analysis*. New York: McGraw-Hill, 1973.

[58] W. Strauss, "Dispersion of waves vanishing on the boundary of an exterior domain," to appear.

[59] Yosida, *Functional Analysis*. New York: Springer, 1966.

Invariant Subspace Methods in Linear Multivariable-Distributed Systems and Lumped-Distributed Network Synthesis

JOHN S. BARAS, MEMBER, IEEE, AND PATRICK DEWILDE, MEMBER, IEEE

Abstract—Linear multivariable-distributed systems and synthesis problems for lumped-distributed networks are analyzed. The methods used center around the invariant subspace theory of Helson-Lax and the theory of vectorial Hardy functions. State-space and transfer function models are studied and their relations analyzed. We single out a class of systems and networks with nonrational transfer functions (scattering matrices), for which several of the well-known results for lumped systems and networks are generalized. In particular we develop the relations between singularities of transfer functions and "natural modes" of the systems, a degree theory for infinite-dimensional linear systems and a synthesis via lossless embedding of the scattering matrix. Finally coprime factorizations for this class of systems are developed.

These factorizations play an essential role in the development and show that properties of Hardy functions are of fundamental importance for this class of distributed systems as properties of rational functions are for lumped systems.

INTRODUCTION

FREQUENCY DOMAIN methods in lumped-multivariable systems have been developed in the last few years for the analysis and design of control systems [1]. These methods provided a clear understanding of the interrelations between state-space and transfer function models for systems and networks and proved to be extremely useful in practical design applications [82].

Recently, several researchers, Baras [2]-[8], Brockett [3]-[4], Dewilde [9]-[16], Fuhrmann [17]-[23], Helton [24] and [25], have been investigating a similar approach to the analysis and synthesis of distributed systems and networks. This theory applies to situations where energy considerations provide the setting of a Hilbert space for the state space of the

Comments on "Systems with Infinite-Dimensional State Space: The Hilbert Space Approach"

JAMES A. DYER

Professor Helton has written a very interesting expository paper[1] on the applications of Hilbert space techniques to systems having infinite-dimensional state spaces. I find it, however, to be rather surprising that quantum mechanics, which is the best example in all mathematical physics of the applications of Hilbert space techniques to the study of infinite-dimensional systems, is not mentioned.

Not only does quantum mechanics provide excellent examples of infinite-dimensional systems, the language currently used by many physicists to describe a quantum system is the same as that used by a systems engineer to describe an engineering system from the state viewpoint. To illustrate this, consider a single-particle quantum mechanical system. Such a system is completely described by a probability amplitude function $\phi(x, y, z, t)$. For fixed $\bar{t}$, $\phi(\bar{t}) = \phi(x, y, z, \bar{t})$ is required to be an element of the unit sphere of the Hilbert space $L_2(R^3)$. For fixed $\bar{t}$, this function is usually called the state of the system even by physicists. The propagation in time of the state is given by Schrödinger's equation

$$\frac{d\phi}{dt} = -\frac{i}{\hbar} \mathcal{H} \phi$$

$$\phi(0) = \phi_0$$

where $i = \sqrt{-1}$, $\hbar$ is $(2\pi)^{-1}$ times Planck's constant, and $\mathcal{H}$ known as the Hamiltonian operator, is a self-adjoint unbounded linear operator defined on a dense subspace of $L_2(R^3)$. It will be observed that this equation has the form of the usual state propagation equation of linear systems theory. It will be observed from this equation also that quantum mechanics concerns itself only with zero-input solutions. The quantities which a systems engineer would call outputs of this system would be called observables by a quantum physicist. An observable y of a single-particle system is related to the state of the system by the equation

$$y = \langle \phi, \mathcal{L} \phi \rangle$$

where $\langle, \rangle$ denotes the inner product on $L_2(R^3)$ and $\mathcal{L}$ is a linear self-adjoint unbounded operator defined on a dense subspace of $L_2(R^3)$. For example, if $\mathcal{H}$ is the Hamiltonian operator then the observable $E = \langle \phi, \mathcal{H} \phi \rangle$ is the energy of the system. Again, one finds here a standard type of output equation for a system having an infinite-dimensional state space and a one-dimensional output. An excellent elementary introduction to quantum mechanics from the state viewpoint is given in [1].

In the Feynman integral approach to quantum mechanics one also makes use of the concept of state transition operator. Here one writes the state $\phi(t)$ as

$$\phi(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t, \hat{x}, \hat{y}, \hat{z}) \phi_0(\hat{x}, \hat{y}, \hat{z}) d\hat{x} d\hat{y} d\hat{z},$$

where K is an L_2 valued function known as the propagator. In fact K is the impulse response of the quantum system and is the solution to Schrödinger's equation for the initial state $\phi_0(\hat{x}, \hat{y}, \hat{z}) = \delta(x - \hat{x}) \cdot \delta(y - \hat{y}) \delta(z - \hat{z})$. Since Schrödinger's equation does not have a Green's function in the usual mathematical sense, K is determined as the value of an abstract integral known as a Feynman integral. An elementary introduction to this approach to quantum theory is given in [2].

The reason for the parallelism between the usual formulations of quantum mechanics and the present formulations of systems theory is that both subjects have a common parent in classical Hamiltonian mechanics. These origins are discussed in detail in [3], part II. Since both systems theory and quantum mechanics have followed independent but approximately parallel paths since their origins it is possible that cross-fertilization between the two subjects could prove beneficial to both.

REFERENCES

[1] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, vol. III. Reading, PA: Addison-Wesley, 1965.
[2] J. B. Keller and D. W. McLaughlin, "The Feynman integral," *Amer. Math. Monthly*, vol. 82, pp. 451-465, May 1975.
[3] H. J. White and S. Tauber, *Systems Analysis*, Philadelphia: W. B. Saunders, 1969.

Reply[2] by J. W. Helton

It is pleasing to hear that my article seemed natural enough to be termed expository, since sections II, III, and IV were new. Professor Dyer gives a nice discussion of how basic quantum mechanics itself can be described as an infinite-dimensional system. Quantum mechanics was indeed not mentioned in the paper and so here are some comments on the scattering theory aspects of it.

The paper introduces a particular definition of infinite-dimensional A, B, C, D-type linear systems and in Section III proves that it is equivalent to Lax-Phillips scattering. Now their theory contains quantum mechanical scattering, although the examples they emphasize are classical. This correspondence is explained in Chapter VI, Section 4, of their book. Roughly, to a Schrodinger equation scattering problem with s-matrix S there is associated a scattering problem for a wave equation with scattering matrix Γ. Then $S(z) = \Gamma(\sqrt{z})$. Functions with branch cuts are not pseudomeromorphic in the sense of Section IV-D of the paper and so $S(z)$ is not pseudomeromorphic. Also note that most results described in the exposition, Section V, hold not only for classical but for quantum mechanical scattering. We should emphasize that when we say quantum mechanical scattering we automatically mean very special input and output operators B, C and so these comments do not address many systems based on Schrodinger's equation.

Comments on "Geometric Progression Ladder RC Networks"

L. GRUNER

Abstract—It is shown that a large class of nonrecurrent RC ladder networks including, as a special case, the geometric ladder networks discussed by Bhattacharyya and Swamy can be represented as an equivalent RCG ladder network.

Bhattacharyya and Swamy's [1] conjecture that other nonrecurrent ladder networks may be represented in terms of recurrent ladder networks is readily answered with reference to the author's earlier publications [2]-[3].

The application of Kirchhoff's law to the nth node of the ladder network of Fig. 1 shows that the voltage $V(n)$ satisfies the difference equation

$$\frac{V(n+1)}{Z(n+1)} - V(n) \left[Y(n) + \frac{1}{Z(n+1)} + \frac{1}{Z(n)} \right] + \frac{V(n-1)}{Z(n)} = 0, \quad n \geqslant 1 \tag{1}$$

where for a nonrecurrent RC network the series impedance $Z(n) = R(n)$ and the shunt admittance $Y(n) = sC(n)$.

In the special case of a recurrent ladder network, $Z(n)$ and $Y(n)$ are invariant with respect to n; in particular, if $Z(n) = R_u$ and $Y(n) = G_u + sC_u$, then replacing $V(n)$ by $V_u(n)$, the solution of the difference equation (1) is given by

$$V_u(n) = C_1 e^{n\theta} + C_2 e^{-n\theta}$$

where

$$\cosh \theta = 1 + \frac{R_u(G_u + sC_u)}{2}. \tag{2}$$