JOURNAL OF COMPUTER AND SYSTEM SCIENCES: 1, 91-116 (1967)

Foundations of the State-Space Theory of Continuous Systems. I

A. V. BALAKRISHNAN*

Department of Engineering, University of California, Los Angeles, California

ABSTRACT

Foundations of the theory of continuous systems based on the concept of states are studied with rigorous definitions, using generalized functions. Both finite- and infinite-dimensional, linear, time-invariant systems are characterized; application to Cauchy problems and distribution semigroups is presented.

1. INTRODUCTION

The purpose of this paper is to study the foundations of the state-space theory of continuous systems, and in particular, to present a precise formulation of the basic concepts and the related fundamental results. For the motivation and background of theory, reference may be made to the pioneering work of Zadeh [*1*]. So long as the systems are linear and the state-space finite-dimensional, as in the studies of Zadeh [*1*] and Kalman [*2*], one is on fairly familiar and physical foundation, and an emphasis on precision on the notions involved is not essential. However, in dealing with systems with state spaces that cannot be finite-dimensional, such as systems governed by partial differential equations, or even those described by frequency transfer functions which are not rational, the need for precise formulation becomes more apparent. In fact with the increasing range of systems amenable to optimization and their attendant complexity (see, for instance, [*3–5*] for purely mathematical aspects), the framework provided by system theory has to rest on firmer ground. This is even more true in problems involving system identification. It is in this spirit that we have undertaken to define the primary concepts both of "system" and "state". In this formulation we have also allowed the input and output to be generalized functions.

* The results reported in this paper were obtained while the author was an exchange professor at the Department of Mathematics, University of Paris, supported in part by a grant (AFOSR 700-65) from the Air Force Office of Scientific Research, Applied Mathematics Division, U. S. Air Force.

In the definitions we have tried to keep the thread of physical application as much visible as possible. However, in order to keep within the bounds of reasonable length, we have had to shorten and even delete altogether all background and motivational material. Also, the use of generalized functions has meant a different treatment even for linear systems with finite dimensional state spaces. These differences have not been emphasized, and the interested reader should consult References [*1*], [*6*], [*7*] for more detail.

As signified in the title, this is the first part of the paper and begins with the basic notation and definitions and then goes into systems that are linear and time-invariant—which here as elsewhere provide a natural starting point, and yet are sufficiently important on their own. The notation to be used (which needs some care since we are dealing with generalized functions on the half-line) is explained in Section 2, where we also include a version of the Schwartz kernel theorem for the half-line. The definitions of "system" and "state" are given in Section 3. It is of course basic that a "system" is not a "function" relating input to output, but rather a "graph". The definitions are based on the earlier ones given by Zadeh [*1*]. We have assumed that the inputs and outputs are (or have their ranges in) real or complex variables, but the generalization to finite-dimensional or Euclidean spaces presents no difficulty. A generalization to infinite-dimensional inputs and outputs may be found in [*6*], although the functions therein are not "generalized functions". Following the basic definitions, we also include some related terminology and definitions. Some conditions such as "input-continuity", for example, would appear to be minimal conditions in order that a useful theory be evolved. Nevertheless there is a loss in generality here; indeed no claim is made for complete generality in any event.

In Section 4, we study systems which are linear and time-invariant in the state-space description. We show how the reduced states may be topologized to yield a locally convex space and a strongly continuous semigroup theorem enabling us to relate state to input in a unique manner ("state–input" relations) and indicate under additional assumptions how the output can be related to the state ("state–output" relation). The results are summarized in Theorem 4.1, although the methods themselves are equally useful. The Zadeh result that finite-dimensionality of the state space implies that the system is "dynamic" (the input may be related to the output through an ordinary differential equation) is now given precise form and proof for inputs and outputs allowed to be generalized functions, clarifying in particular the input–state and state–output equations. It is shown that this result holds if we assume that the reduced state space (with its output-induced topology) is normable. This result was proved in [7] under the assumption that the outputs are ordinary functions which are Lebesgue-measurable. To explore this situation, we go on to consider an example where the state space is infinite-dimensional—a near canonical example where the inputs and outputs are given in terms of a semigroup over a Banach space, the semigroup being strongly continuous at the origin (corresponding for instance to well-posed or Cauchy problems for partial differential equations), and the relationship between the Banach space and the state space is examined. A final example in which the semigroup is not even measurable (and hence not even Laplace-transformable), namely the distribution semigroup of Lions [8]—shows how state spaces may be constructed for such systems, and how they fall under the purview of systems as defined herein. In particular, the state-space theory in turn sheds some light on distribution semigroups.

2. Notation

For each a, $-\infty \leqslant a < +\infty$, let $\mathscr{D}_a$ denote the class of infinitely differentiable (complex variable) functions with compact support in the open interval $(a, +\infty)$, endowed with the Schwartz topology, and let $\mathscr{D}_a'$ denote its dual. We shall always use the weak-star topology for it. For each a, $-\infty < a < +\infty$, let $\mathscr{D}_{a-}$ denote the class of all infinitely differentiable functions with compact support in the half-closed interval $[a, +\infty)$, the compact sets now allowed to include the origin, and again endowed with the Schwartz topology. Another way of describing $\mathscr{D}_{a-}$ is as follows: for each f in $\mathscr{D}_{-\infty}$ let us define the transformation

$$Pf = g; g(s) = f(s), \qquad s \geqslant a,$$

thus mapping $\mathscr{D}_{-\infty}$ into $\mathscr{D}_{a-}$. This is actually an *onto* mapping; in fact, for any $g(\cdot)$ in $\mathscr{D}_{a-}$ we can clearly define a function f in $\mathscr{D}_{a-\epsilon}$ for each $\epsilon > 0$ such that

$$Pf = g.$$

Furthermore, P is linear and continuous in the respective Schwartz topologies. Let us denote the dual of $\mathscr{D}_{a-}$ by $\mathscr{D}'_{a-}$. Then for any W in $\mathscr{D}'_{a-}$, we can define a continuous linear functional on $\mathscr{D}_{-\infty}$ by

$$L(f) = W(Pf),$$

and hence there is a generalized function in the usual sense (or Schwartz distribution) vanishing in the open set $(-\infty, a)$ corresponding to L, and only one of this kind. Since the representation is unique, we shall continue to use W to denote it as well. Again, $\mathscr{D}_a$ being a linear subspace of $\mathscr{D}_{a-}$,

$$M(f) = W(f), \qquad f \in \mathscr{D}_a$$

defines a continuous linear functional on $\mathscr{D}_a$, and by the Hahn–Banach theorem for locally convex spaces, the functional M can be extended to $\mathscr{D}_{a-}$. If $\tilde{W}_0$ is such an extension to $\mathscr{D}_{a-}$, we note that

$$W - \tilde{W}_0$$

must have its support reduced to the origin.

For each $a, s, -\infty < a, s, < +\infty$, we shall define a linear transformation, denoted $T(a+s; a)$, corresponding to a translation by the amount s of the time origin, mapping $\mathscr{D}_{a-}$ into $\mathscr{D}_{(a+s)-}$ by

$$T(a+s; a)f = g; g(t) = f(t-s), \qquad t \geqslant a+s. \tag{2.1}$$

This is clearly linear, continuous, and *onto*, and also one-to-one.

Next we define a translation semigroup (corresponding to "left translation") $S_1(t)$, $t \geqslant 0$ by

$$S_1(t)f = g; g(s) = f(t+s), \qquad s \geqslant a, \tag{2.2}$$

mapping $\mathscr{D}_{a-}$ into itself. This is linear and continuous for each t; and for f in $\mathscr{D}_{a-}$, $S_1(t)f$ is continuous in t, $t \geqslant 0$.

We also need a right-translation semigroup $S_r(t)$. This is defined by

$$\begin{aligned} S_r(t)f = g; g(s) &= f(s-t), \qquad s \geqslant a+t, \\ &= 0, \qquad a \leqslant s \leqslant a+t. \end{aligned}$$

which for each t is a linear continuous mapping of $\mathscr{D}_a$ into itself. Also, for each f in $\mathscr{D}_a$, $S_r(t)f$ is continuous in t, $t \geqslant 0$.

For each t, we can extend $S_r(t)$ to $\mathscr{D}_a'$ by defining for each W in $\mathscr{D}'_{a-}$:

$$(S_r(t)W)(f) = W(S_1(t)f), \qquad f \in \mathscr{D}_{a-}$$

and $S_r(t)W$ is then continuous in t in the weak-star topology of $\mathscr{D}'_{a-}$. Similarly we may extend $S_1(t)$ to $\mathscr{D}_a'$ by defining for each W in $\mathscr{D}_a'$:

$$(S_1(t)W)(\phi) = W(S_r(t)\phi), \qquad \phi \in \mathscr{D}_a,$$

and $S_1(t)W$ is now continuous in t in the weak-star topology of $\mathscr{D}_a'$.

We shall denote by C_a^∞ the class of all infinitely differentiable functions on the half-open interval $[a, \infty)$. Every element of C_a^∞ clearly defines a linear functional on $\mathscr{D}_{a-}$. The topology on C_a^∞ will be taken to be that induced on it as a subspace of $\mathscr{D}_a'$, the latter being of course taken in the weak-star topology over $\mathscr{D}_a$.

For the rest of this section we shall specialize these spaces to $a = 0$. For each W in $\mathscr{D}'_{0-}$, and f in $\mathscr{D}_{0-}$, we observe that the integral

$$\int_0^\infty S_r(t)\,Wf(t)\,dt \tag{2.3}$$

is well-defined as a Pettis integral. In fact for any g in $\mathscr{D}_{0-}$, we have

$$\int_0^\infty S_r(t)\,W(g)f(t)\,dt = W(h), \tag{2.4}$$

where

$$h = \int_0^\infty S_1(t)\, gf(t)\, dt, \tag{2.5}$$

the integral being definable as a Riemann integral, with $h \in \mathscr{D}_{0-}$. It is clear that (2.4) defines a linear (continuous) functional on $\mathscr{D}_{0-}$. We shall denote the element in $\mathscr{D}'_{0-}$ so defined by the "convolution"

$$W * f$$

The restriction of this functional to $\mathscr{D}_0$, yields an element of $\mathscr{D}_0'$. We shall denote the latter element so obtained by

$$(W * f)_0 . \tag{2.6}$$

Suppose now that f is actually an element of $\mathscr{D}_0$. Let

$$\begin{aligned} \hat{f}(s) &= f(s), \qquad s \geqslant 0, \\ &= 0, \qquad s < 0. \end{aligned}$$

Then $\hat{f}$ is an element of $\mathscr{D}_{-\infty}$. As we have noted, W can be identified with a (unique) element of $\mathscr{D}'_{-\infty}$ vanishing in the open set $(-\infty, 0)$. Let us denote this element by ${}^\infty W$. Then the convolution in the ordinary sense of ${}^\infty W$ and $\hat{f}$ defines a function in $C^\infty_{-\infty}$; in fact it vanishes in $(-\infty, +\epsilon)$ for some $\epsilon > 0$. It is not difficult to see that $(W * f)$ coincides with this function, and that in fact:

$$S_\Gamma(r)(W * f) \in C_0^\infty \qquad \text{for every} \qquad t \geqslant 0; f \in \mathscr{D}_0$$

Also:

$$S_\Gamma(t)(W * f) = W * (S_r(t) f), \qquad t \geqslant 0; f \in \mathscr{D}_0 \tag{2.7}$$

We can also verify that (2.6) defines a continuous linear transformation mapping $\mathscr{D}_{0-}$ into $\mathscr{D}_0'$, considering $\mathscr{D}_{0-}$ as a linear subspace of $\mathscr{D}'_{0-}$. For let $\phi_n \in \mathscr{D}_{0-}$ and let ϕ_n converge to an element ϕ in $\mathscr{D}_{0-}$ in the weak-star topology of $\mathscr{D}'_{0-}$. Then for each u in $\mathscr{D}_0$ we have:

$$(W * \phi_n)_0 (u) = \int_0^\infty (S_\Gamma(t)\, W)(u)\, \phi_n(t)\, dt,$$

and the function

$$S_\Gamma(t)\, W(u), \qquad t \geqslant 0$$

clearly belongs to $\mathscr{D}_{0-}$, so that

$$\lim_n (W * \phi_n)_0 (u) = \int_0^\infty S_\Gamma(t)\, W(u)\, \phi(t)\, dt = (W * \phi)_0 (u).$$

We also note the converse, being an extension of the Schwartz kernel theorem [cf. 9], which we can state as

THEOREM 2.1. *Let L denote a linear transformation of $\mathscr{D}_{0-}$ into $\mathscr{D}_0'$ such that*

(i) *for each f in $\mathscr{D}_0$, Lf is in C_0^∞ and*

$$S_\Gamma(t) Lf = L(S_\Gamma(t) f), \qquad t \geqslant 0; \tag{2.8}$$

(ii) *L is a continuous mapping, considering $\mathscr{D}_{0-}$ as a subspace of $\mathscr{D}'_{0-}$, the latter being taken in the weak-star topology. Then*

$$Lf = (W * f)_0 \qquad \text{for some } W \text{ in} \qquad \mathscr{D}'_{0-}. \tag{2.9}$$

Proof. We shall outline a proof since it is different from the one for $\mathscr{D}_{-\infty}$. First let ϕ_n be a delta-convergent sequence in $\mathscr{D}_0$ (that is, $\phi_n * u$ converges to u for every u in $\mathscr{D}_0$ in the latter's Schwartz topology; such a sequence is also known as a "regularizing" sequence). Let u be an element of $\mathscr{D}_0$. Then, because of (2.8), we have

$$L(\phi_n * u) = (L\phi_n) * u = (Lu) * \phi_n , \tag{2.10}$$

and also for every $t \geqslant 0$,

$$L(S_\Gamma(t)\phi_n * u) = (LS_\Gamma(t)\phi_n) * u = (LS_\Gamma(t)\, u) * \phi_n , \tag{2.11}$$

and hence the functions

$$L(S_\Gamma(t)\, \phi_n) * u$$

converge pointwise. For each u, let

$$\hat{u}_N(s) = u(N - s), \quad \text{where} \quad u(s) = 0 \qquad \text{for} \qquad s \geqslant N.$$

Then $\hat{u}_N$ belongs to $\mathscr{D}_0$, and letting

$$L(\phi_n) = f_n$$

we have that

$$(S_\Gamma(t) f_n)(u) = ((LS_\Gamma(t)\phi_n) * \hat{u}_N)$$

evaluated at the point N, and hence the left side converges. Again for any f in $\mathscr{D}_{0-}$, we can find ϕ in $\mathscr{D}_0$ and some $t \geqslant 0$, such that

$$f = S_1(t)\,\phi.$$

Hence

$$f_n(f) = (S_\Gamma(t) f_n)(\phi)$$

converges for every f in $\mathscr{D}_{0-}$. Hence there exists an element W in $\mathscr{D}'_{0-}$ such that

$$\lim f_n(f) = W(f).$$

Again, $\phi_n * f$ converges to f in the weak-star topology of $\mathscr{D}'_{0-}$. Hence

$$\lim_n L(\phi_n * f)(u) = (Lf)(u) \qquad \text{for } u \text{ in} \qquad \mathscr{D}_0 .$$

But again using (2.8),

$$\begin{aligned} L(\phi_n * f)(u) &= (f_n * f)(u) \\ &= \int_0^\infty (S_1(t) f_n)(u) f(t)\, dt \\ &= \int_0^\infty W(S_1(t)\, u) f(t)\, dt \\ &= (W * f)(u) \end{aligned}$$

or

$$Lf = (W * f)_0 .$$

3. Precise Definitions

In this section we define precisely the basic notions of "System" and "State" using the notation of Section 2; we also include some terminology and related definitions.

Definition of a System

For each real number a, $-\infty < a < +\infty$, let $U(a)$ and $V(a)$ denote two classes of functions, including generalized functions, on the half-closed interval $[a, \infty)$, the class $U(a)$ including $\mathscr{D}'_{a-}$ and the class $V(a)$ including $\mathscr{D}'_{a-}$. Let $S(a)$ be a subset of the cross-product space $U(a) \times V(a)$. A system (or more correctly a continuous system as opposed to a discrete system) is a collection of all these subsets, subject to the following conditions: (i) For each a, let U_a denote the projection of $S(a)$ on $U(a)$. The elements of U_a will be called "inputs starting at time a". It will be assumed that U_a includes $\mathscr{D}_{a-}$. For each element u_a of U_a, the elements in the intersection

$$(u_a \times V(a)) \cap (Sa)$$

will be called the "outputs" corresponding to the input u_a. It will be assumed that, for each u_a in $\mathscr{D}_{a-}$, the corresponding outputs are contained in $\mathscr{D}_a'$. (ii) Each element in $S(a)$ will be denoted (u_a, v_a) and called an "input–output" pair. It will be assumed that, for each input–output pair (u_a, v_a) with u_a in $\mathscr{D}_{a-}$,

$$(T(a+t; a)\, S_1(t)\, u_a,\ T(a+t; a)\, S_1(t)\, v_a)$$

is in $S(a+t)$ for each $t \geqslant 0$; and what is crucial, *that every input–output pair in* $S(a+t)$ *with* u_{a+t} *in* $\mathscr{D}_{a+t}$ *can be obtained in this manner.* (iii) For each a, let (u_a, v_a) be an input–output pair with u_a in $\mathscr{D}_a'$. Then it is possible to find a sequence of input–output pairs $(u_a{}^n, v_a{}^n)$, with $u_a{}^n$ in $\mathscr{D}_a$ such that $u_a{}^n$ and $v_a{}^n$ converge, in the weak-star topology of $\mathscr{D}_a'$, to u_a and v_a, respectively.

Definition of State Space

Let a system be defined as above, and let us retain the same notation. For each a, $-\infty < a < \infty$, let a set $\Sigma(a)$ be given, and a transformation mapping the cross-product space $\Sigma(a) \times U_a$ into $V(a)$. We shall denote this transformation by $A(a; x; u_a)$, $x \in \Sigma(a)$, $u_a \in U_a$. For each a, $\Sigma(a)$ will be called a "state space at time a", if the following two conditions hold:

(H_1): For each u_a in $\mathscr{D}_a$, let v_a be a corresponding output; then there is an element x in $\Sigma(a)$ such that

$$V_a = A(a; x; u_a)$$

and conversely, given any x in $\Sigma(a)$, and any u_a in $\mathscr{D}_a$,

$$A(a; x; u_a)$$

yields an output corresponding to u_a.

(H_2): For any a, and $t > 0$, let $C(a; t)$ denote the class of all functions in $\mathscr{D}_{a-}$ which agree on the interval $(a, a+t)$. Then there is an element y in $\Sigma(a+t)$ such that, for each u_a in $C(a; t)$ with corresponding output v_a,

$$T(a+t; a)\, S_1(t)\, v_a = A(a+t; y; T(a+t; a)\, S_1(t)\, u_a)$$

Remark. We shall use the term "system with a state-space description" to denote a system that has state spaces as defined above for every instant of time.

In the following definitions it is assumed that the system has a state-space description.

1. *State Equivalence.* Two states in $\Sigma(a)$ are said to be "equivalent", denoted $x_1 \text{ eq } x_2$, if

$$A(a; x_1; u_a) = A(a; x_2; u_a)$$

for every u_a in $\mathscr{D}_{a-}$.

Remark. It is clear that the states in $\Sigma(a)$ can be partitioned into equivalence classes under this equivalence relationship.

2. *Zero-input Equivalence.* Let 0 denote the zero-element in $\mathscr{D}_{a-}$. Two states x_1, x_2 in $\Sigma(a)$ are said to be "zero-input equivalent" if

$$A(a; x_1; 0) = A(a; x_2; 0).$$

Remark. It is clear that the states can again be partitioned into equivalence classes using zero-input equivalence as well.

3. *Input-Continuous System.* A system is said to be input-continuous if for every a and each x in $\Sigma(a)$, $A(a; x; .)$ is a continuous mapping of $\mathscr{D}_{a-}$ into $\mathscr{D}_a'$, the former being taken in the weak-star topology of $\mathscr{D}'_{a-}$.

4. *Input-Analytic System.* A system is said to be input-analytic if for every a, and each x in $\Sigma(a)$, $A(a; x; .)$ is a continuous mapping as in Definition 3, above, and $A(a, x; .)$ is an "analytic" mapping in the sense that, for each u_a, h_a in $\mathscr{D}_{a-}$, $A(a; x; u_a + \lambda h_a)$ is an entire analytic function of λ in $\mathscr{D}_a'$.

5. *Ground State.* If $\Sigma(a)$ has any element x in it such that

$$A(a; x; 0) = 0,$$

then x will be called a "ground state at time a".

Remark. It is clear that all ground states are zero-input equivalent.

6. *Time-Invariant System.* A system is said to be time-invariant if we can take

$$\Sigma(a) = \Sigma$$

for every a, and for every x in Σ,

$$A(a; x, u_a) = T(a; 0)\, A(0; x; T(0; a)\, u_a).$$

7. *Linear System.* A system is said to be linear if we can take $\Sigma(a)$ to be a linear vector space for each a, and the mapping $A(a; .; .)$ is linear on the cross-product space $\Sigma(a) \times \mathscr{D}_{a-}$, and continuous in the sense of Definition 3 above.

8. *Reduced State Space.* A state space $\Sigma(a)$ is said to be reduced (at time a) if $x_1 \text{ eq } x_2$ implies that $x_1 = x_2$.

Remarks. It is clear that if a system is linear, then zero-input equivalence implies equivalence, and that the zero-element of the vector space $\Sigma(a)$ is a ground state.

9. *Separable System.* A system is said to be separable if it has a ground state (which we shall denote by 0) and the response or output is separable: (for each a):

$$A(a; x; u_a) = A(a; 0; u_a) + A(a; x; 0), \qquad u_a \in \mathscr{D}_{a-}\,; \; x \in \Sigma(a).$$

The first term here will be called the ground-state response, and the second term the zero-input response.

Remark. It is clear that a linear system is separable; and in this case, the ground state response will be called the zero-state response as well, since the zero-state is a ground state and all ground states are equivalent.

10. *Physical Realizability.* A system is said to be physically realizable, if for each a, and each x in $\Sigma(a)$, the transformation $A(a; x; u_a)$ is "physically realizable" in the sense that if u_a^1, u_a^2 are two inputs in $\mathscr{D}_a$ which agree on the interval $[a, b]$, so do the outputs $A(a; x; u_a^1)$, $A(a; x; u_a^2)$.

Remark. It will be assumed that in what follows every system considered is physically realizable.

4. Linear Time-Invariant Systems

It is natural to begin our study of systems from the state-space point of view by first considering linear time-invariant systems. It should be noted that the concept of state was used in the very definition of linearity. We shall only consider input continuous systems; in fact even this is too general in that it is not possible to deduce structural relationships. In any event we shall require that an additional condition is satisfied by the ground-state response:

(S): For each input in $\mathscr{D}_0$, the ground-state response is in C_0^∞. This is certainly reasonable in that the response to an infinitely smooth input starting from rest should surely be expected to be smooth.

As in Section 3, linearity implies that we can decompose the output as

$$A(a; x; u_a) = Z_{\mathrm{I}}(a; x) + Z_{\mathrm{S}}(a; u_a)$$

where $Z_{\mathrm{I}}(..)$ is the "zero-input" response, and $Z_{\mathrm{S}}(..)$ the "zero-state" response. The first step is to examine the space of reduced states. Because of linearity, two states are equivalent as soon as they are zero-input equivalent. Hence, let us begin with the zero-input response. For any a, time invariance implies in particular that the state spaces do not depend on time:

$$\Sigma(a) = \Sigma, \qquad \text{say.}$$

Also

$$Z_1(a; x) = T(a; 0)\, Z_1(0; x), \qquad x \in \Sigma(a) = \Sigma(0) = \Sigma, \tag{4.1}$$

and hence it is enough to consider

$$Z_1(0; x), \qquad x \in \Sigma.$$

Let

$$\Sigma_0 = [x/Z_1(0; x) = \text{zero element of } \mathscr{D}_0'] \tag{4.2}$$

(we note that the zero-input response is the response to the zero-input function which is in $\mathscr{D}_0$, and hence the response is by assumption in $\mathscr{D}_0'$). The class of reduced states is readily verified to be a linear space; in fact, since Σ_0 is clearly a linear subspace of Σ, we have, denoting the reduced states by Σ_r,

$$\Sigma_r = \Sigma/\Sigma_0\,; \qquad \text{factor space modulo } \Sigma_0 .$$

In order not to complicate the notation unnecessarily, we shall continue to denote the reduce states by letters x, y etc., since in particular we will only be dealing with reduced states in the rest of this section. In particular, we shall continue to use the zero-state response

$$Z_1(0; x)$$

for states in Σ_r, it being understood that this is the response corresponding to any member of the corresponding equivalence class. Let

$$Px = Z_1(0; x) \qquad x \in \Sigma_r . \tag{4.3}$$

Then P is a linear (one-to-one) mapping of Σ_r into $\mathscr{D}_0'$. It should be noted that so far there has been no topology on the state spaces Σ or Σ_r. We now proceed to topologize Σ_r. For this we use the mapping P which maps Σ_r into $\mathscr{D}_0'$. Thus a set E in Σ_r is open if and only if PE is open in $P\Sigma_r$, considered as a subspace of $\mathscr{D}_0'$, the latter in the weak-star topology. We observe that Σ_r is a locally convex space in this topology. We may, and do, proceed to complete the space Σ_r in this topology, and denote the completed space by $\overline{\Sigma_r}$. We note that

$$\overline{\Sigma_r} = P^{-1}(\overline{p\Sigma_r})$$

since P can clearly be extended to the completed space $\overline{\Sigma_r}$ and is homeomorphic onto $(\overline{P\Sigma_r})$.

We shall next see how the state hypothesis H_2 leads to a semigroup of linear transformations over $\overline{\Sigma_r}$. For this, we observe that for each x in Σ_r, and $t \geqslant 0$, there is an element of Σ, and hence a unique element y of Σ_r, such that

$$T(t; 0)\, S_1(t)\, Px = A(t; y; 0) = T(t; 0)\, Z_1(0; y)$$

or

$$S_1(t)\,Px = Z_1(0; y) \qquad y \in \Sigma_\Gamma$$

Hence we may define a transformation $T(t)$ such that

$$T(t)\,x = P^{-1}S_1(t)\,Px, \tag{4.4}$$

mapping Σ_Γ into Σ_Γ. For each t, $T(t)$ is clearly linear, and because the transformation $S_1(t)$ is continuous on $\mathscr{D}_0'$ in the weak-star topology, as we have noted in Section 2, it follows that $T(t)$ can be extended to be linear and continuous on $\overline{\Sigma_\Gamma}$. Moreover, for each x in $\overline{\Sigma_\Gamma}$, $T(t)\,x$ is continuous in $t \geqslant 0$, because $S_1(t)\,Px$ is continuous in $t \geqslant 0$. Of course $T(t)$ is clearly also a semigroup. Let A denote the infinitesimal generator of this semigroup. Then A is defined on all of $\overline{\Sigma_\Gamma}$, and in fact

$$Ax = P^{-1}DPx, \tag{4.5}$$

where D denotes the differential operator over $\mathscr{D}_0'$. It is clear that for x in Σ_Γ, Ax belongs in general to $\overline{\Sigma_\Gamma}$, since DPx is in the closure of $P\Sigma_\Gamma$ in the weak-star topology of $\mathscr{D}_0'$.

Let us next consider the zero-state response. First we have, by time invariance,

$$Z_s(a; u_a) = T(a; 0)\,Z_s(0; T(0; a)\,u_a). \tag{4.6}$$

Let u be any element of $\mathscr{D}_0$; then for any $t \geqslant 0$, $S_\Gamma(t)\,u$ also belongs to $\mathscr{D}_0$, and if we let

$$h = Z_s(0; S_\Gamma(t)\,u),$$

h is in C_0^∞ by assumption(s), and (by the physical realizability hypothesis) vanishes in the interval $[0, t]$, as does

$$g = S_\Gamma(t)\,Z_s(0; u).$$

Next, by the state hypothesis H_2, we must have

$$[\text{since } h_t = T(t; 0)\,S_1(t)\,A(0; 0; S_\Gamma(t)\,u) = A(t; 0; T(t; 0)\,u)]$$

$$h_t = Z_s(t; u_t)$$

where

$$h_t(s) = h(s), \qquad s \geqslant t; \qquad\qquad u_t(s) = u(s - t), \qquad s \geqslant t,$$

and since

$$T(0; t)\,u_t = u,$$

it follows from (4.6) that

$$h_t = T(t; 0)\,Z_s(0; u)$$

and hence that

$$h(s) = g(s), \qquad s \geq t,$$

or, finally,

$$Z_s(0; S_r(t)\, u) = S_r(t)\, Z_s(0; u), \qquad u \in \mathscr{D}_0 .$$

Hence the transformation

$$Z_s(0; u),$$

mapping $\mathscr{D}_{0-}$ into $\mathscr{D}_0'$, satisfies all the conditions under which Theorem 2.1 (the extended version of the Schwartz kernel theorem) holds, and hence (cf. Section 2) there is an element W in $\mathscr{D}'_{0-}$ such that

$$Z_s(0; f) = (W * f)_0 , \qquad f \in \mathscr{D}_{0-} . \tag{4.7}$$

Let W_0 denote, as in Section 2, the restriction of W to $\mathscr{D}_0$:

$$W_0(u) = W(u), \qquad u \in \mathscr{D}_0 , \; W_0 \in \mathscr{D}_0'. \tag{4.8}$$

We shall now prove that there is an element (denoted B) in $\overline{\Sigma_r}$ such that:

$$PB = W_0 . \tag{4.9}$$

For this, let ϕ_n be a delta-convergent sequence in $\mathscr{D}_0$ such that the support of ϕ_n is contained in the interval $[0, \epsilon_n]$, $\epsilon_n \to 0$. Then, by the state hypothesis H_2 , it follows that there must be an element x_n in Σ_r such that

$$S_1(\epsilon_n)\, Z_s(0; \phi_n) = Px_n .$$

But by (4.8),

$$S_1(\epsilon_n)\, Z_s(0; \phi_n) = S_1(\epsilon_n)(W * \phi_n)_0 ,$$

and the right side of this equality is readily verified to converge in the weak-star topology of $\mathscr{D}_0'$ to W_0 . Hence the sequence x_n converges in $\overline{\Sigma_r}$ and denoting the limit by B, (4.9) is obtained.

Next we observe that the integral

$$\int_0^t T(t-s)\, Bf(s)\, ds \tag{4.10}$$

is well defined as a Pettis integral, for each f in $\mathscr{D}_{0-}$. In fact, for each u in $\mathscr{D}_0$,

$$\int_0^t (S_1(t-s)\, W_0)(u) f(s)\, ds = W_0 \left(\int_0^t S_r(t-s)\, u f(s)\, ds \right)$$

clearly defines a linear continuous functional on $\mathscr{D}_0$. Moreover, since $S_1(t)\, W_0$ is in $\overline{P\Sigma_\Gamma}$ for every $t \geqslant 0$, this functional is also in $\overline{P\Sigma_\Gamma}$, and hence (4.10) defines an element in $\overline{\Sigma_\Gamma}$.

We can now define a state-transition relation:

$$x(t) = T(t)\, x + \int_0^t T(t-s)\, Bf(s)\, ds, \tag{4.11}$$

which defines the state at time t, starting with state x at time zero, and with input f in $\mathscr{D}_{0-}$. This can be "differentiated" (in the state topology defined) to yield

$$\dot{x}(t) = Ax(t) + Bf(t), \tag{4.12}$$

where the dot indicates the derivative. It should be noted that in (4.12) no restriction is placed on the initial state at time zero.

Let us next consider the problem of relating the state function defined by (4.11) to the output. For this let us first define a (continuous) linear functional on $\overline{\Sigma_\Gamma}$ for each n by

$$L_n(x) = Px(\phi_n),$$

where ϕ_n is a delta-convergent sequence in $\mathscr{D}_0$. For each n,

$$L_n(T(t)\, x) = Px(S_\Gamma(t)\, \phi_n)$$

is continuous in t. For any $u(.)$ in $\mathscr{D}_0$,

$$\int_0^\infty L_n(T(t)\, x)\, u(t)\, dt = Px(h_n),$$

where

$$h_n = \int_0^\infty S_\Gamma(t)\, \phi_n u(t)\, dt = \phi_n * u,$$

and since h_n thus converges to u in $\mathscr{D}_0$, we have that $L_n(T(t)\, x)$ converges in the weak-star topology of $\mathscr{D}_0'$ to the zero-input response. Since the latter is in general only an element of $\mathscr{D}_0'$, pointwise convergence is not to be expected, or even meaningful. The second term in (4.11) cannot be handled in a similar manner without additional constraint, since it involves only W_0, and the zero-state response involves W. A convenient assumption is the following:

(A): Let u_n be any sequence in $\mathscr{D}_0$ such that u_n converges to an element f in $\mathscr{D}_{0-}$ in the weak-star topology of $\mathscr{D}_0'$; then $W(u_n) = W_0(u_n)$ converges, and the limit depends only on f (that is, it is independent of the particular sequence chosen).

An example where (A) is not satisfied is

$$W_0(u) = \int_0^\infty e^{1/t} u(t)\, dt, \qquad u(.) \in \mathscr{D}_0 .$$

On the other hand (A) is clearly satisfied if W_0 corresponds to a measurable function which is integrable on finite intervals.

We shall now explore some of the consequences of this assumption.

LEMMA 1. *If* (A) *holds, then for each f in $\mathscr{D}_{0-}$, we have*

$$W * f = \tilde{W} * f + \sum_0^n a_k\, d^k f/dt^k, \tag{4.13}$$

where $\tilde{W} \in \mathscr{D}_0'$ and also satisfies (A).

Proof. Let

$$W_n(f) = W_0(\phi_n * f), \qquad f \in \mathscr{D}_{0-},$$

where ϕ_n is a delta-convergent sequence in $\mathscr{D}_0$. Then since $\phi_n * f$ converges in the weak-star topology of $\mathscr{D}_0'$ to f, $W_n(f)$ converges for each f in $\mathscr{D}_{0-}$, and setting

$$\tilde{W}(f) = \lim_n W_n(f)$$

defines $\tilde{W}$ as an element of $\mathscr{D}_{0-}'$, and since it agrees with W_0 on $\mathscr{D}_0$, it satisfies (A). In fact, $\tilde{W}$ is clearly the smallest extension of W_0 to $\mathscr{D}_{0-}$. Now $\tilde{W} - W$ must have its support confined to the origin, so that

$$(W - \tilde{W})(f) = \sum_0^n a_k\, d^k f/dt^k,$$

from which (4.13) follows.

Since

$$\tilde{W}(f) = \lim_n W_0(\phi_n * f),$$

it follows that, if W_1, W_2 are two elements of $\mathscr{D}_{0-}'$ such that

$$W_1(f) = W_2(f) = \lim_n W_0(\phi_n * f),$$

then their difference $W_1 - W_2$ must be zero; and hence if

$$W * f = W_1 * f + \sum a_k\, d^k f/dt^k$$
$$= W_2 * f + \sum b_k\, d^k f/dt^k,$$

we must have $a_k = b_k$. In other words, the decomposition (4.13) is unique.

LEMMA 2. *Let* $f \in \mathscr{D}_{0-}$ *and let*

$$y(t) = \int_0^t T(t-s)\, Bf(t)\, dt. \tag{4.14}$$

Then, if (A) *holds, we have that the (continuous) functions*

$$L_n(y(t)), \qquad 0 \leqslant t$$

converge in the weak-star topology of $\mathscr{D}'_{0-}$ *to*

$$(\tilde{W} * f)_0 .$$

Proof.

$$L_n(y(t)) = W\left(\int_0^t S_{\mathrm{r}}(t-s)\,\phi_n f(s)\, ds\right). \tag{4.15}$$

For any g in $\mathscr{D}_{0-}$,

$$\begin{aligned}
\int_0^\infty L_n(y(t))\, g(t)\, dt &= \int_0^\infty f(s)\, ds \int_s^\infty g(t)\, W_0(S_{\mathrm{r}}(t-s)\,\phi_n)\, dt \\
&= \int_0^\infty f(s)\, ds \int_0^\infty g(t+s)\, W_0(S_{\mathrm{r}}(t)\,\phi_n)\, dt \\
&= \int_0^\infty f(s)\, W_0(\phi_n * S_1(s)\, g)\, ds \\
&= W_0(\phi_n * h)
\end{aligned}$$

where

$$h = \int_0^\infty S_1(s)\, gf(s)\, ds,$$

and by Assumption (A),

$$\lim_n W_0(\phi_n * h) = \tilde{W}(h) = (\tilde{W} * f)(g),$$

as required.

Combining Lemma 2 with Lemma 1, we have

LEMMA 3. *Let* $x(t)$ *be the state function defined in* (4.11). *Then* $L_n(x(t))$, *which for each* n, *is a continuous function of* t, *converges in the weak-star topology of* $\mathscr{D}_0{}'$, *and the output*

$$v = A(0;\, x;\, f) = \lim L_n(x(.)) + \sum_0^n a_k\, d^k f/dt^k. \tag{4.16}$$

In the case of the zero-state response, we can get pointwise convergence for inputs in $\mathscr{D}_0$, since by assumption $W * f$ then belongs to C_0^∞.

LEMMA 4. *Suppose $f \in \mathscr{D}_0$. Then, assuming* (A) *holds, for every $t \geqslant 0$,*

$$\lim_n (\tilde{W} * f)(S_r(t)\phi_n) = \lim L_n(y(t)). \tag{4.17}$$

Proof. Since $f \in \mathscr{D}_0$, we know that $\tilde{W} * f$ belongs to C_0^∞, and the left-side of (4.17) converges to the latter for each value of t. But for each n,

$$\begin{aligned}(\tilde{W} * f)\, S_r(t)\phi_n &= \int_0^\infty (S_r(s)\, \tilde{W})(S_r(t)\phi_n) f(s)\, ds \\ &= \int_0^t W_0(S_r(t-s)\phi_n) f(s)\, ds \\ &\quad + \int_t^\infty \tilde{W}(S_1(s-t)\phi_n) f(s)\, ds \\ &= L_n(y(t)) + \tilde{W}(f_n),\end{aligned} \tag{4.18}$$

where

$$f_n = \int_0^\infty S_1(s)\phi_n f(s+t)\, ds.$$

As we have already seen,

$$L_n(y(t)) = W_0(g_n),$$

where

$$g_n = \int_0^t S_r(t-s)\phi_n f(s)\, ds.$$

For any u in $\mathscr{D}_0$,

$$\int_0^\infty g_n(t)\, u(t)\, dt = \phi_n \left(\int_0^t S_1(s)\, uf(t-s)\, ds \right) \to \int_0^t u(s) f(t-s)\, ds,$$

or g_n converges now to the function $\hat{f}$ defined by

$$\begin{aligned}\hat{f}(s) &= f(t-s), \qquad 0 \leqslant s \leqslant t, \\ &= 0, \qquad\qquad 2 > t\end{aligned}$$

in the weak-star topology of $\mathscr{D}_0'$, $\hat{f}$ being in $\mathscr{D}_{0-}$ since f is in $\mathscr{D}_0$. Hence $L_n(y(t))$ converges to $W_0(\hat{f})$ for each t. Hence $\tilde{W}(f_n)$ converges also. But

$$\tilde{W}(f_n) = \lim_m W_0(\phi_m * f_n),$$

the function $\phi_m * \phi_n$ is given by

$$(\phi_m * \phi_n)(y) = \int_0^\infty \int_0^v \phi_m(y - \sigma)\,\phi_n(s + \sigma) f(s + t)\,ds,$$

and, for each $u(.)$ in $\mathscr{D}_0$, the double sequence

$$\int_0^\infty (\phi_m * \phi_n)(t)\,u(t)\,dt$$

converges to zero. Hence by Assumption (A) the double sequence

$$W_0(\phi_m * \phi_n)$$

converges to a limit. On the other hand, it is readily verified that for fixed m, the sequence $\phi_m * \phi_n$ actually converges to zero in the Schwartz topology of $\mathscr{D}_0$. Hence it follows that

$$\lim W(f_n) = \lim_{n,m} W_0(\phi_m * \phi_n) = \lim_{m,n} W_0(\phi_m * \phi_n) = 0.$$

Hence (4.17) follows.

Let us observe that we have proved that

$$\lim \tilde{W}(f_n) = \lim(\tilde{W} * S_1(t) f)(\phi_n) = 0$$

for any f in $\mathscr{D}_0$. Since, for any g in $\mathscr{D}_{0-}$, we can write

$$g = S_1(t) f \qquad \text{for some} \qquad f \text{ in } \mathscr{D}_0,$$

it follows that

$$\lim_n (\tilde{W} * g)(\phi_n) = 0, \qquad g \in \mathscr{D}_{0-}. \tag{4.19}$$

We can use this fact to prove

LEMMA 5. *For any t, such that* $(\tilde{W} * f)_0$, *where* $f \in \mathscr{D}_{0-}$, *coincides with an integrable function in a neighborhood of t,*

$$(\tilde{W} * f)_0\,(t) = \lim_n L_n(y(t))$$

on the Lebesgue set of the limit function in that neighborhood.

Proof. We have only to observe that on the Lebesgue set, the left side of (4.17) converges to the function. On the other hand, in (4.18), $W(f_n)$ now converges to zero from (4.19).

We may sum up these results as

THEOREM 4.1. *Given any time-invariant linear system with a state-space description, that is input-continuous and whose zero-state response has Property* (S); *it is possible to define a locally convex topology on the reduced states and a semigroup of linear continuous transformations* $T(t)$ *on the* (*complete*) *locally convex state space* X, *which is strongly continuous in* $t \geqslant 0$, *and a sequence* L_n *of linear continuous functionals on* X *such that the response*

$$A(a;\, x;\, u_a) = T(a;\, 0)\, A(0;\, x;\, T(0;\, a)\, u_a)$$

can be expressed in terms of state-input and state-output relations as follows:

$$x(t) = T(t)\, x + \int_0^t T(t-s)\, Bf(s)\, ds;\ \dot{x}(t) = Ax(t) + Bf(t)$$

where B *is a fixed element of* X, *and* $f \in \mathscr{D}_{0-}$; *letting*

$$v = A(0;\, x;\, f) \tag{4.20}$$

we have that, *if Condition* (A) *is satisfied*,

$$v = \lim_n L_n(x(t)) + \sum_0^n a_k\, d^k f/dt^k, \qquad 0 \leqslant t \tag{4.21}$$

where the limit may be taken in the weak-star topology of $\mathscr{D}_0{}'$, *and pointwise whenever the response is continuous*, *and almost everywhere in any interval where the response is Lebesgue-integrable.*

We proceed next to examine what may be looked upon as special cases of Theorem 4.1. The first and simplest case is of course the "degenerate" case where the state-space has finite dimension. We have then the Zadeh result [*1*] that the input and output are related by an ordinary linear differential equation. This was proved in [*7*] under the assumption that the outputs were Lebesgue-measurable "ordinary" functions. We shall now state and prove this in the light of our definitions, the proof in particular being immediate.

THEOREM 4.2. *Suppose the state space of a linear time-invariant system is finite-dimensional. Assume that the zero-state response satisfies Condition* (S). *Then confining ourselves to inputs and outputs starting at time zero*, (*which we may because of time variance*), *any output corresponding to an input* f *in* $\mathscr{D}_{0-}$ *must satisfy the differential equation*

$$\sum_0^p c_k\, d^k v/dt^k = \sum_0^q d_j (d^j f/dt^j). \tag{4.22}$$

For any piecewise-continuous input f, any corresponding output v is also in $\mathscr{D}'_{0-}$, *and*

$$\sum_0^p c_k D^k v = \sum_0^q d_j D^j f, \tag{4.23}$$

where D denotes the derivative operator on $\mathscr{D}_0'$. *The reduced state space being also finite-dimensional, the state–input relation can be defined as*

$$\dot{x}(t) = Ax(t) + Bf(t);\ x(0) = x, \qquad f \in \mathscr{D}_{0-} \tag{4.24}$$

in the sense that $x(t)$ *is the unique solution of this equation with the initial state being the prescribed element x, B being a fixed element of the state space. The state–output relation is*

$$v(t) = [C, x(t)] + \sum_0^n a_k\, d^k f/dt^k, \tag{4.25}$$

where C is a fixed element of the state space, and [,] *denotes inner product. For any piecewise-continuous input f, the usual "distributional" interpretation of* (4.24) *and* (4.25) *may be made.*

Proof. The state space Σ being finite-dimensional, so is the state space Σ_r, and hence

$$\Sigma_r = \overline{\Sigma_r} = X.$$

The semigroup $T(t)$ is then trivially uniformly continuous and hence

$$T(t) = \exp At, \tag{4.26}$$

where A is a linear bounded transformation on X. Hence for any x in X, we have

$$T(t)\, x = \sum_0^m \phi_k(t)\, A^k x, \tag{4.27}$$

where the $\phi_k(.)$ are (from the usual matrix exponential formulas) in C_0^∞. Hence, also

$$S_1(t)(Px) = \sum_0^m \phi_k(t)\, D^k(Px). \tag{4.28}$$

Let us denote the element Px by w. We shall now show that (4.28) implies that w is actually an element of C_0^∞. For this, let us note that

$$w(f*g) = \int_0^\infty (S_1(t)\,w)(f)\,g(t)\,dt \qquad (\text{for } \; f, g \in \mathscr{D}_0)$$

$$= \int_0^\infty \sum_0^m \phi_k(t)(D^k w)(f)\,g(t)\,dt$$

$$= \sum_0^m (D^k w)(f) \int_0^\infty \phi_k(t)\,g(t)\,dt.$$

Let f_n be a delta-convergent sequence in $\mathscr{D}_0$. Then

$$w(g) = \lim w(f_n * g) = \lim \sum_0^m (D^k w)(f_n) \int_0^\infty \phi_k(t)\,g(t)\,dt. \tag{4.29}$$

Now, the functions $\phi_k(t)$ are linearly independent over any interval of $[0, \infty)$ and by suitable choice of g, it will follow that every term in (4.29) must converge, since the number of terms therein is finite. This shows that w must coincide with a linear combination of the functions $\phi_k(t)$, and in particular, must be in C_0^∞. Hence every element in $P\Sigma_\Gamma$ is in C_0^∞. Hence so is W_0, and its Laplace transform is rational. For any element x of X, let us recall that

$$L_n(x) = (Px)(\phi_n),$$

and it follows that $L_n(.)$ converges and defines a linear functional on X. Hence there must be an element C in X such that

$$[C, x] = \lim L_n(x).$$

Moreover, the output is now continuous for any input f in $\mathscr{D}_{0-}$, so that from (4.21) we have

$$v(t) = [C, x(t)] + \sum_0^n a_k\, d^k f/dt^k. \tag{4.25}$$

In particular,

$$\left[C, \int_0^t T(t-s)\,Bf(s)\,ds\right] = \int_0^t [C, T(t-s)\,B]\,f(s)\,ds$$

$$= (\tilde{W} * f)(t), \qquad t \geqslant 0,$$

and since clearly $\tilde{W}$ coincides with W_0, it follows that

$$W_0(t) = [Ce^{A(t-s)}B], \tag{4.30}$$

where in particular we note that

$$[C, e^{A(t)}x] = 0, t \geqslant 0, \qquad \text{implies} \qquad x = 0$$

From (4.25) it is immediate that (4.22) holds, p being the degree of the minimal polynomial of A, as well as the dimension of X. Similarly (4.23) and (4.24) are easily deduced.

COROLLARY. *The statements of theorem remain valid if we merely assume that the reduced state-space X is normable.*

Proof. We have only to note that by construction X is homeomorphic with a subspace of $\mathscr{D}_0'$ in the weak-star topology, and any such subspace that is normable must be finite-dimensional (as is shown, for instance, by a ready application of the Kolmogorov theorem which states that normability is equivalent to the existence of a bounded convex symmetric neighborhood of the origin).

This Corollary should not be too surprising since the topology we have defined for the reduced state space is the weakest. For a similar result with a slightly different topology, see [7]. In order to clarify some of the relationships, we shall now consider an infinite-dimensional example which has many canonical features. We shall incidentally also indicate how a state space may be derived from given input–output description in the process; although the general problem of determining state from input–output is much more complex and will be taken up later.

We shall actually begin with a simple version first. Let E be a Banach space, and let $S(t)$ denote a semigroup of linear bounded transformations, strongly continuous at the origin; that is,

$$S(t)\, x \qquad \text{is continuous for} \qquad t \geqslant 0, x \in X. \tag{4.31}$$

Let B be a fixed element of E, and x^* a fixed element of the adjoint space. We now define an output for each u_a in $\mathscr{D}_{a-}$ by

$$v_a(t) = x^* \left(S(t-a)\, x + \int_a^t S(t-s)\, Bu_a(s)\, ds\right), \tag{4.32}$$

where the integral is a Bochner integral (in general), and x is an arbitrary element of E. It is clear that $v_a(.)$ is actually in C_a^∞. It is evident that E will serve as the set of states as defined in Section 3. We shall denote it by Σ to indicate that there is no topology on it as yet. Moreover the system is time-invariant and linear. Hence we may (and shall) take the initial time "a" to be zero.

We may proceed now to obtain the space of reduced states and the corresponding state–input/state–output relations. Thus

$$\Sigma_0 = [x \in E \mid x^*(S(t)\,x) = 0, \qquad t \geqslant 0]$$

This is of course a linear subspace of Σ (closed in E). The space Σ_{r} can be made into a Banach space if (take the Gelfand–Neumark norm)

$$\| x \|^1 = \inf \| x + \Sigma_0 \|$$

But our topology (termed the "output induced topology" in [7] and which we shall adopt to avoid confusion) is that induced by considering

$$x^*(S(t)\,x), \qquad t \geqslant 0$$

as elements of $\mathscr{D}_0'$. In this topology, Σ_{r} is not necessarily closed (in fact it is closed if and only if it is finite-dimensional). Thus in completing Σ_{r} we have introduced new states, the significance of which will be clear presently. The semigroup $T(t)$ is given by

$$P(T(t)\,x)\,x^*(S(t+\sigma)\,x) = x^*(S(\sigma)\,S(t)\,x), \qquad \sigma \geqslant 0 \tag{4.33}$$

for x in Σ_{r}. Let $\tilde{A}$ denote the infinitesimal generator of the semigroup $S(t)$. Then for any x in the domain of $\tilde{A}$,

$$P(Ax)\,x^*(S(\sigma)\,Ax), \qquad \sigma \geqslant 0.$$

Otherwise Ax is defined as the distributional derivative of $x^*(S(.)\,x)$. The element W and W_0 are the same in the present case, and avoiding the trivial case when B is an element of Σ_0, W_0 corresponds to the function

$$W_0(t) = x^*(S(t)\,B) \tag{4.34}$$

and clearly satisfies Condition (A), although not necessarily in C_0^∞. The integral

$$\int_0^t T(t-s)\,Bf(s)\,ds$$

now corresponds to the function in $\mathscr{D}_0'$:

$$x^*(S(\sigma)\,y(t)),\ \sigma \geqslant 0, \qquad \text{where} \qquad y(t) = \int_0^t S(t-s)\,Bf(s)\,ds,$$

and the latter is always an element of E for $f(.)$ in $\mathscr{D}_{0-}$. It continues to be even when $f(.)$ is not necessarily in $\mathscr{D}_{0-}$, but is rather measurable and bounded on finite intervals. But if $f(.)$ is allowed to be an element of $\mathscr{D}_0'$, it is no longer in E, necessarily, and the completed space X is necessary. Thus if we extend the inputs in (4.32) to be elements of $\mathscr{D}_0'$, then the state-space E is no longer sufficient, but X is. In many variational and control problems, the inputs have to be extended in this or a related manner. The nature of the extension of the inputs to $\mathscr{D}_0'$ is exactly as demanded in the definition of "System" (cf. Section 3).

The condition of strong continuity of the semigroup (4.31) may be dropped to make the example more general. We may thus consider the case where the semigroup is merely (strong) Lebesgue-measurable, that is

$$S(t)\,x \qquad \text{is Lebesgue-measurable for each} \qquad x \text{ in } E. \tag{4.35}$$

But in this case the Phillips–Miyadera theorem [9] asserts that $S(t)\,x$ must then be actually continuous for t positive. Hence we may define a system by

$$v_a(t) = x^*\left(S(t-a)\,x + \int_0^t S(t-s)\,Bu_a(s)\,ds\right), \tag{4.36}$$

but here the inputs have to be confined to $\mathscr{D}_a$ in order that the integral be defined. We can of course proceed to construct the reduced state-space X as before. However, since W now corresponds to the function

$$x^*(S(t)\,B), \qquad t \geqslant 0,$$

and is continuous only in $t > 0$, we need to add a condition similar to (A) to enable us to connect the state function with the output. But if we do add such a condition, we can then apply the results of Theorem 4.1.

From this point of view, we may note that we can consider the case where the semigroup is not necessarly even Lebesgue-measurable. We may then consider the "distribution semigroup" of Lions [8]. A distribution semigroup is a linear continuous mapping (denoted S) of $\mathscr{D}_0$ into the Banach space of linear bounded transformations on E. Let B be again a fixed element of E, and let x^* be as before. Then for each f in $\mathscr{D}_0$,

$$x^*(S(f)\,x), \qquad x \in E$$

defines a linear continuous functional on $\mathscr{D}_0$ or an element of $\mathscr{D}_0'$. We may thus define outputs corresponding to inputs in $\mathscr{D}_0$, analogous to (4.32) (with $a = 0$) by

$$v = x^*(S(.)\,x) + x^*(S(u)\,B). \tag{4.37}$$

The properties of a distribution semigroup (for which reference may be made to Lions [8]) relevant here are

(i) $S[\phi * \psi]\,x = S[\phi] \cdot S[\psi]\,x, \qquad x \in E, \phi\psi \in \mathscr{D}_0$;

(ii) There is a subspace L (dense in the strong topology) in E, such that for any w in $\mathscr{D}_0'$ and any ϕ in $\mathscr{D}_0$

$$S[w*\phi]\, x = S[\phi]\, \tilde{S}[w]\, x, \qquad x \in L,$$

where $\tilde{S}[w]$ is a closed linear transformation with domain L.

We shall now show that we can take L as a state space, for the input–output pairs described by (4.37), where for simplicity we shall assume that B is also in L. Let us first note that, for x in L, ϕ in $\mathscr{D}_0$,

$$S_1(t)[x*S[\phi]\, x] = x*[S[\phi]\, \tilde{S}[\delta(t)]\, x]$$

since

$$S_1(t)[x*S[.]\, x](\phi) = x*S[\phi*\delta(t)]\, x,$$

where $\delta(t)$ is the distribution corresponding to the delta-function concentrated at t. Moreover, $\tilde{S}[\delta(t)]\, x$ is again in L. Again, let $\hat{u}$ be defined so that

$$\begin{aligned} \hat{u}(s) &= u(t-s), \qquad 0 \leqslant s \leqslant t \\ &= 0 \qquad \text{otherwise for} \qquad s > t. \end{aligned}$$

Then $\hat{u}$ is an element of $\mathscr{D}_0'$, and

$$\tilde{S}[\hat{u}]\, B \in L.$$

Hence we have

$$T(t;0)\, S_1(t)\, v = T(t;0)[x*S[.](\tilde{S}[\delta(t)]\, x + \tilde{S}[\hat{u}]\, B) + x*S[g]\, B] \tag{4.38}$$

where

$$g(s) = u(s+t), \qquad s \geqslant 0.$$

This shows that L is a proper state space in the sense of the definition in Section 3, and that we thereby obtain a linear time-invariant system. Proceeding then to obtain Σ_T, and the completion $\overline{\Sigma_T}$, we note that the semigroup $T(t)$ is such that it "corresponds to"

$$\tilde{S}[\delta(t)].$$

We omit the details of the corresponding state–output relations. It may be noted parenthetically that we have as a byproduct of the state-space theory, a "representation" for the distribution semigroup[1] in terms of an ordinary (strongly continuous)

[1] According to the referee, a similar result has been obtained using completely different methods, simultaneously and independently, by D. Fusiwara (to be published in the *Journal of the Mathematical Society of Japan*, Vol. 18, 1966).

semigroup on a locally convex space. Also, this example shows the generality of the systems that may be studied using the state-space approach.

ACKNOWLEDGMENT

It is a pleasure to acknowledge the many helpful discussions with Professor J. L. Lions of the University of Paris.

REFERENCES

1. L. A. ZADEH AND C. DESOER. "Linear System Theory." McGraw-Hill, New York.
2. R. E. KALMAN. *J. SIAM (Control)* **1**, 152-192 (1963).
3. A. V. BALAKRISHNAN. *J. SIAM (Control)* **3**, 152-180 (1965).
4. H. O. FATTORINI. "Control in Finite Time of Differential Equations in Banach Space," (to be published).
5. J. L. LIONS. Sur quelques problèmes d'optimization dans les équations d'évolution linéaires de type parabolique.. *In* "Functional Analysis and Optimization," E. Caianiello, Ed. Academic Press, New York (1966).
6. A. V. BALAKRISHNAN. On the state space theory of non-linear systems. *In* "Functional Analysis and Optimization," E. Caianiello, Ed. Academic Press, New York (1966).
7. A. V. BALAKIRSHNAN. *J. Math. Anal. Appl.* **14**, 371-391 (1966).
8. J. L. LIONS. Les semi groupes distributions. *Portugal. Math.* **19**, 141-164 (1960).
9. K. YOSHIDA. "Functional Analysis." Academic Press, New York (1965).

Printed in Belgium