THE UNIVERSITY OF MICHIGAN

COLLEGE OF ENGINEERING
Department of Aerospace Engineering

ON THE MODELLING OF SYSTEMS FOR IDENTIFICATION

Part II. Time-Varying Systems

William L. Root

ORA Project 011306

supported by:

UNITED STATES AIR FORCE
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
AIR FORCE SYSTEMS COMMAND
GRANT NO. 72-2328
WASHINGTON, D. C.

administered through:

OFFICE OF RESEARCH ADMINISTRATION ANN ARBOR

May 1973

ABSTRACT

Certain Banach spaces, denoted $\mathcal{L}^p_T$, of equivalence classes of functions of a real variable are introduced and investigated. These are to be used as system input and output spaces for systems operating for all time. Some basic properties of causal systems with finite memory are established. The concept of forming time-interval truncations of a time-varying system is formalized and investigated. Trajectories of such truncations are studied. It is proved that under certain reasonable conditions, the trajectories of a class of systems are generated by a strongly continuous semigroup of linear operators. It is also shown that $\mathcal{E}$-representations of the truncations can be generated by an induced semigroup. Thus, the evolution in time of classes of, in general, nonlinear, time-varying systems, is described in the framework of a linear dynamical theory.

INTRODUCTION TO PART 2

A "system" as defined in Part 1 is simply an input space, an output space, and a mapping carrying inputs into outputs. In Part 1 some abstract structure and representation theory is established for classes of systems for which the system mappings are bounded and continuous, and for which certain conditions are satisfied by the input and output spaces and by the class of mappings itself. In Part 2 the interest is in systems and classes of systems where the inputs and outputs are functions of time. Again there is no restriction to linear or to time-invariant systems. The chief emphasis is on causal systems, and indeed on causal systems with bounded memory.

The emphasis on causal systems needs no justification, but there might be a question raised as to why one should consider systems with bounded memory. The primary answer is: the bounded memory condition turns out to fit very conveniently in the mathematical structure used, and since almost any system of interest has a decaying memory it can be approximated as well as desired by a system with finite memory. The goal in this work is to set up approximate models or representations of classes of systems to be used in identification, so approximation is permissible. It is certainly realistic to stipulate that the observation periods for both inputs and outputs be of finite duration, and this requirement influences the mathematical structure chosen.

The first Section after this introduction is devoted to setting up and investigating certain function spaces which are appropriate for modelling input and output spaces for systems; the second to establishing basic facts about causal and bounded memory systems. In the third Section the concept of trajectories of time-limited truncations of systems is developed. The time-limited truncations are, roughly speaking, observable portions of a system which is operating for all time. In the fourth Section families of trajectories associated with a class of systems are considered. Under certain circumstances, these trajectories can be generated by semigroups of linear operators.

The results of Part 1 are not used explicitly in Part 2 till near the end of the fourth Section. However the work of Part I influences what is done in Part 2 throughout. Some of the material of Sections I and III appeared, with only partial proofs, in the conference paper [1] .

I. SOME FUNCTION SPACES FOR INPUTS AND OUTPUTS

We want to be able to treat systems for which the inputs and outputs are functions of time, real or vector-valued with finitely many components, and extending for infinite time. We do not want it to be required that the inputs and outputs must always die out in some sense in the infinite future or infinite past. Hence, function spaces, such as the L_p spaces with $1 \le p < \infty$, which have the property that their constituent functions all get arbitrarily small (in some sense) as $t \to \pm \infty$ will usually not be suitable for modelling admissible collections of inputs or outputs. The spaces of bounded or essentially bounded functions are satisfactory on this score, and sometimes we shall use the space of bounded continuous functions on R^1 with the sup norm. However, there are certain non-standard function spaces that are suitable and especially convenient, and which will be used customarily. These are spaces of functions that are uniformly local-time L_2 provided with one of a family of norms to be given in the definition below. These spaces are only Banach spaces, but the local L_2 character is advantageous. Since it is quite as easy to define such spaces more generally using a local-time L_p property, $1 \le p < \infty$, we do so, even though the local L_2 spaces are the ones chiefly desired.

Let y be either a real-valued function of a real variable, or a vector-valued function that has finitely many real-valued components. In the second case, $|y(t)|$ will denote the Euclidean norm of the vector y(t). Define the operator P_t by

$$[P_t y](s) = y(s), \quad s \le t$$
$$= 0 \quad, \quad s > t \qquad (1)$$

As usual, let $L_p(A)$ denote the Lebesgue space of p-integrable N-vector-valued functions on the measurable set $A \subset R^1$. It will always be assumed that $1 \le p < \infty$. The L_p norm is written $\|\cdot\|_p$. Let $\mathcal{L}_o^{(p)}$ be the space of all functions y that satisfy the following condition: $y \in \mathcal{L}_o^{(p)}$ iff for any T, $0 < T < \infty$, there is a positive number $K = K(T,y)$ such that $\|(P_{t+T} - P_t)y\|_p \le K$ for all $t \in R^1$. Obviously $\mathcal{L}_o^{(p)}$ is a linear space over the real numbers under the usual addition and scalar multiplication of functions. It is made into a normed linear space by the assignment of a norm:

$$\|y\|_T^{(p)} = \sup_t (\|(P_{t+T} - P_t)y\|_p) \qquad (2)$$

where T is an arbitrary fixed number, $0 < T < \infty$. We call the resulting normed linear space $\mathcal{L}_T^{(p)}$.

Proposition 2.1 $\mathcal{L}_T^{(p)}$ is a Banach space if its elements are interpreted to be the equivalence classes of functions in $\mathcal{L}_o^{(p)}$ that are equal a.e. Lebesgue.

Proof: It is immediately verifiable that $\mathcal{L}_T^{(p)}$ is indeed a normed linear space, so it remains only to show it is complete. Consider a particular set of half-open, half-closed intervals $(kT, (k+1)T]$ where k is any integer, $-\infty < k < \infty$. Since for any $(t, t, t+T] \subset (kT, (k+1)T] \cup ((k+1)T, (k+2)T]$ for some k, we have

$$\sup_t \|(P_{t+T} - P_t)y\|_p \le \sup_k \|(P_{(k+2)T} - P_{kT})y\|_p$$

$$\le 2 \sup_k \|(P_{(k+1)T} - P_{kT})y\|_p$$

or, $\|y\|_T^{(p)} \le 2 \sup_k \|(P_{(k+1)T} - P_{kT})y\|_p \le 2\|y\|_T^{(p)}$. Now let $\{y_n\}$ be a sequence such that for any $\epsilon > 0$, $\|y_m - y_n\|_T^{(p)} \le \epsilon$ whenever $m, n \ge n_o(\epsilon)$. Then, for any integer k

$$\|(P_{(k+1)T} - P_{kT})(y_n - y_m)\|_p \le \epsilon \quad \text{for} \quad m, n \ge n_o(\epsilon).$$

Since $L_p(kT, (k+1)T]$ is complete, there is a $y^{(k)} \in L_p(kT, (k+1)T]$ which is the limit of the sequence $\{(P_{(k+1)T} - P_{kT})(y_n)\}$, regarded as a sequence of functions on $(kT, (k+1)T]$. Let y be the equivalence class of functions on R^1 that are equal a.e. on each interval $(kT, (k+1)T]$ to any function representing $y^{(k)}$. Then $y \in \mathcal{L}_T^{(p)}$, and

$$\|y - y_n\|_T^{(p)} \le 2 \sup_k \|(P_{(k+1)T} - P_{kT})(y - y_n)\|_p \quad .$$

Since,

$$\|(P_{(k+1)T} - P_{kT})(y - y_n)\|_p \le 2\epsilon \quad \text{for} \quad n \ge n_o(\epsilon)$$

and all k, $\|y - y_n\|_T^{(p)} \le 4\epsilon$ for $n \ge n_o(\epsilon)$. Thus y is the limit of y_n.†

Henceforth, unless there is some particular reason to be precise, we shall refer to the "functions in $\mathcal{L}_T^{(p)}$" or the "functions in $L_p(A)$" instead of to the elements, which are properly equivalence classes of functions. Some elementary properties of the spaces $\mathcal{L}_T^{(p)}$ are noted in the next proposition and succeeding remarks.

<u>Proposition 2.2</u> Let N, the number of (real) components of the vector-valued functions under consideration, be fixed. Let T_1, T_2 be any

positive numbers. Then,

(a) $\mathcal{L}_{T_1}^{(p)}$ and $\mathcal{L}_{T_2}^{(p)}$ are comprised of the same elements, and the norms on these two spaces are equivalent.

(b) If $p > q$, then any function belonging to $\mathcal{L}_T^{(p)}$ belongs also to $\mathcal{L}_T^{(q)}$. Also, convergence in $\mathcal{L}_T^{(p)}$ implies convergence in $\mathcal{L}_T^{(q)}$.

<u>Proof:</u> The elements of $\mathcal{L}_{T_1}^{(p)}$ and $\mathcal{L}_{T_2}^{(p)}$ are the elements of $\mathcal{L}_o^{(p)}$ (for the fixed N in question). Suppose $T_1 < T_2$ and m is an integer such that $m\,T_1 > T_2$. Then for $y \in \mathcal{L}_o^{(p)}$, $\|y\|_{T_1} \le \|y\|_{T_2} \le m\,\|y\|_{T_1}$.

Part (b) follows from Holder's inequality. In fact,

$$\begin{aligned}\|y\|_T^{(q)} &= \sup_t \{ \|(P_{t+T} - P_t)\, y\|_q \} \\ &\le \sup_t \{ \|(P_{t+T} - P_t)\, y\|_p \cdot T^{\frac{p-q}{pq}} \} \\ &= T^{\frac{p-q}{pq}} \|y\|_T^{(p)} \quad . \quad \dagger\end{aligned}$$

Let $\mathcal{M}(\Delta)$ denote the set of functions in $\mathcal{L}_o^{(p)}$ which vanish a.e. outside the interval $\Delta = (a,b)$, where $-\infty \le a < b \le +\infty$. Then clearly $\mathcal{M}(\Delta)$ is a closed linear subspace of $\mathcal{L}_T^{(p)}$, for any $0 < T < \infty$. If a and b are finite, $\mathcal{M}(\Delta)$ may be identified with either $L_p(\Delta)$ or with a closed linear subspace of $L_p(R^1)$, and the 1:1 correspondence in either case is a linear homeomorphism. If $T = b - a$, the correspondence is isometric.

We denote the operations of translation by c to the left or right, respectively, by L_c and R_c; i.e., $(L_c u)(t) = u(t+c)$. L_c and R_c are linear operations on $\mathcal{L}_o^{(p)}$ which pressure norm in any $\mathcal{L}_T^{(p)}$, and $L_c = R_{-c} = R_c^{-1}$. The following identities hold for any function u defined on R^1 and any real numbers a, b, c :

$$L_c (P_b - P_a) u = (P_{b-c} - P_{a-c}) L_c u \qquad (3)$$

$$R_c (P_{b-c} - P_{a-c}) u = (P_b - P_a) R_c u \quad .$$

It will sometimes be convenient in order to avoid an awkward locution to apply L_c or R_c to elements of an $L_p(\Delta)$, for some finite interval Δ. When this is done it will always be intended that $L_p(\Delta)$ be identified with $\mathcal{M}(\Delta)$, as above, so that the operation is defined. Care must be taken of course to ensure that this operation is meaningful.

Compactness of input spaces is required for much of the structure described in Part I. Because that structure is to be applied to what follows, compactness in some form will again often appear as a requirement, but usually not as the condition that an input space be a compact subset of an $\mathcal{L}_T^{(p)}$ space. A weaker condition is appropriate, one which says that ordinary compactness only hold locally in time (not local compactness). This notion is formalized, and it is proved below that there is an abundance of subsets of $\mathcal{L}_T^{(p)}$ which have this property along with certain other desirable properties.

A subset $\mathcal{A}$ of $\mathcal{L}_o^{(p)}$ is T-compact if $(P_{t+T} - P_t)\mathcal{A}$ regarded as a subset of $L_p(t, t+T]$ is compact for every t. Relative T-compactness is defined correspondingly.

Proposition 2.3 If $\mathcal{A}$ is a compact subset of $\mathcal{L}_T^{(p)}$ it is T-compact, but the converse is not necessarily true. Also, if $\mathcal{A}$ is T-compact for a positive number T, it is T_1-compact for any other positive number T_1.

Proof: Obvious. †

Another property of input sets that will be essential is the following: a subset $\mathcal{U}$ of $\mathcal{L}_o^{(p)}$ will be said to have the projection property, denoted (P), if $u \in \mathcal{U}$ implies that $P_t u$, $(I - P_t)u$ and $(P_t - P_s)u$ belong to $\mathcal{U}$ for any real numbers s and t. Of course, if s = t, $(P_t - P_s)u = 0$, so in particular the zero function must belong to $\mathcal{U}$.

Proposition 2.4 Let $\mathcal{U}_o$ be any compact subset of $L_p(0,T]$. Then there exists a set $\mathcal{U} \subset \mathcal{L}_T^{(p)}$ with the following properties:

1) $(P_T - P_o)\mathcal{U} \supset \mathcal{U}_o$ (using the identification explained previously between $L_p(0,T]$ and $\mathcal{M}(0,T]$)

2) $\mathcal{U}$ has property (P)

3) $\mathcal{U}$ is T-compact

4) $\mathcal{U}$ is invariant under time shift; thus if $u \in (P_{t+T} - P_t)\mathcal{U}$ then $L_t u \in (P_T - P_o)\mathcal{U}$, and vice versa.

Proof: We first enlarge $\mathcal{U}_o$ so as to have a set that is closed under the projections $(P_t - P_s)$, $0 \le s, t \le T$. Let $\mathcal{U}_o'$ be the subset of $L_p(0,T]$ consisting of all functions u' satisfying

$$u' = (P_t - P_s)u \quad , \quad u \in \mathcal{U}_o \quad , \quad 0 \le s, t \le T$$

a.s. $\mathcal{U}_o'$ is compact in $L_p(0,T]$. In fact, let $\{u_n'\}$ be an infinite sequence of elements of $\mathcal{U}_o'$; $u_n' = (P_{t_n} - P_{s_n})u_n$. Form a subsequence of the positive integers, $\{n_i\}_i$, such that $t_{n_i} \to t_o$, $s_{n_i} \to s_o$, $\|u_{n_i} - u_o\|_p \to 0$ as $i \to \infty$. Put $u_o' = (P_{t_o} - P_{s_o}) u_o$. Then

$$\|u_{n_i} - u'_o\|_p \le \|(P_{t_{n_i}} - P_{s_{n_i}})u_{n_i} - (P_{t_{n_i}} - P_{s_{n_i}})u_o\|_p$$

$$+ \|(P_{t_{n_i}} - P_{s_{n_i}})u_o - (P_{t_o} - P_{s_o})u_o\|_p$$

$$\le \|u_{n_i} - u_o\|_p + \|(P_{t_{n_i}} - P_{t_o})u_o\|_p$$

$$+ \|(P_{s_o} - P_{s_{n_i}})u_o\|_p$$

which is arbitrarily small for sufficiently large i. $\mathcal{U}'_o$ is closed under the projections $(P_t - P_s)$ since $(P_t - P_s)u'$, $u' \in \mathcal{U}'_o$, can always be written $(P_{t_1} - P_{s_1})u$, $u \in \mathcal{U}_o$, for some t_1 and s_1.

Now construct a subset $\tilde{\mathcal{U}}$ of $\mathcal{L}_T^{(p)}$ as follows: let the elements $u \in \tilde{\mathcal{U}}$ be defined by

$$\begin{aligned}
u(t) &= u_o(t) , & 0 < t \le T \\
&= u_{-1}(t+T) , & -T < t \le 0 \\
&= u_1(t-T) , & T < t \le 2T \\
&\cdots \\
&= u_k(t-kT) , & kT < t \le (k+1)T \\
&\cdots
\end{aligned}$$

where the u_k are any sequence of elements from $\mathcal{U}'_o$. Put $\mathcal{U} = \bigcup_{0 \le \eta \le T} L_\eta \tilde{\mathcal{U}}$. Obviously $\mathcal{U} \supset \mathcal{U}_o$ and $\mathcal{U} \subset \mathcal{U}_T^{(p)}$.

$\mathcal{U}$ has property (P) since $\tilde{\mathcal{U}}$ does and translation does not affect the property.

$\mathcal{U}$ is invariant under time shift. By the way it is constructed, $\mathcal{U}$ is invariant under shifts which are integer multiples of T. Since any shift can be decomposed into a shift by NT for some integer N, and a shift L_η, $0 \le \eta < T$, $\mathcal{U}$ is invariant for any shift.

$\mathcal{U}$ is T-compact. It is sufficient to prove that $(P_T - P_o)\mathcal{U}$ is compact. Let $\{z_n\}$ be an infinite sequence contained in $(P_T - P_o)\mathcal{U}$. By the construction given, each z_n must be of the form:

$$z_n = (I - P_o) L_{\eta_n} u_n + P_T R_{T - \eta_n} u'_n$$

where u_n, $u'_n \in \mathcal{U}'_o$, and $0 \le \eta_n \le T$. Let n_i be a subsequence such that $\|u_{n_i} - u_o\|_p \to 0$, $\|u'_{n_i} - u'_o\|_p \to 0$ and $\eta_{n_i} \to \eta$, where u_o and u'_o belong to $\mathcal{U}'_o$ and $0 \le \eta \le T$. There is such a subsequence because of the compactness of $\mathcal{U}'_o$ (and of the interval $[0, T]$). Then

$$\lim_{i \to \infty} z_i = (I - P_o) L_\eta u_o + P_T R_{T - \eta} u'_o$$

$$= L_\eta (I - P_\eta) u_o + R_{T - \eta} P_\eta u'_o .$$

In fact, since $\|L_\alpha u - u\|_p \to 0$ as $\alpha \to 0$, it follows that $L_{\eta_{n_i}} u_i \to L_\eta u_o$ and $R_{T - \eta_{n_i}} u'_i \to R_{T - \eta} u'_o$ by the triangle inequality. The limit element belongs to $(P_T - P_o)\mathcal{U}$ by the definition of $\mathcal{U}$. †

T-compactness cannot be replaced by compactness here. In fact it is trivially verifiable that if $\mathcal{U}_o$ has even two distinct elements, then any $\mathcal{U}$ satisfying 1) and 4) in the theorem is not compact, whether it satisfies 2) or not. Indeed, $\mathcal{U}$ need only be invariant under shifts by integer multiples of T to make compactness impossible: with no loss of generality let $\mathcal{U}_o$ consist of the functions $f_o(t) = 0$, $0 \le t < T$, and $f_1(t) = 1$, $0 \le t < T$. Then it is sufficient to observe that the sequence

$\{u_n\}$, u_n defined by $u_n(t) = 1$, $nT \leq t < (n+1)T$, $u_n(t) = 0$ for all other $t \in R^1$, has no limit point.

The construction given in the proof of Proposition 2.4 depends on T and happens to give a class $\mathcal{U}$ that includes all the periodic functions of period T generated by $\mathcal{U}_o'$. However, we remark that if $\mathcal{U}$ is a subset of $\mathcal{L}_o^{(p)}$ which is T_1-compact, shift-invariant and has property (P), then it is T_2-compact and, of course, still shift-invariant with property (P). Thus, in modelling a system, an input space $\mathcal{U}$ can be chosen with the desirable properties listed without any consideration being given to the value of T to be used.

Clearly the bounded continuous functions on R^1 (denoted* $\mathcal{B}_c$) are contained in all $\mathcal{L}_o^{(p)}$, and convergence in the uniform norm of B_c implies convergence in $\mathcal{L}_T^{(p)}$ for any p and any T. It is easy to see also that the functions of B_c are not dense in any $\mathcal{L}_T^{(p)}$. We give a characterization of the closed subspace of $\mathcal{L}_T^{(p)}$ which is generated by $\mathcal{B}_c$.

Proposition 2.5 For any $y \in \mathcal{L}_o^{(p)}$, let $y^{(k)} \in L_p[0,T]$ be defined by: $y^{(k)}(t) = y(t+kT)$, $0 \leq t < T$, for all integers k. Then, a necessary and sufficient condition that y can be approximated in $\mathcal{L}_T^{(p)}$ norm by functions from $\mathcal{B}_c$ is that the functions $[y^{(k)}(t)]^p$ be uniformly integrable.

Proof: (Sufficiency). Let $B_k(b) = \{t \leq [0,T) : |y^{(k)}(t)|^p > b\}$. The condition that the $[y^{(k)}]^p$ are uniformly integrable is that, given any $\eta > 0$, there exists $b > 0$ such that $\int_{B_k(b)} |y^{(k)}(t)|^p \, dt \leq \eta$ for all integers k.

*To be consistent, what is here denoted $\mathcal{B}_c$ should be $\mathcal{F}(R^1, R^n)$, but it seems less confusing to introduce a new symbol for this special case.

Let $\epsilon > 0$ be given, and put $\eta = \left(\frac{\epsilon}{2}\right)^p$. Let $y_b^{(k)}(t) = y^{(k)}(t)$ whenever $|y^{(k)}(t)| \le b^{1/p}$ and equal to zero otherwise. For each k, there is a function $f^{(k)}$ defined on $[0, T]$ which is continuous on $[0, T]$, and satisfies $|f^{(k)}(t)| \le b^{1/p}$ and $\|y_b^{(k)} - f^{(k)}\|_p \le \eta$. Then,

$$\|y^{(k)} - f^{(k)}\|_p \le \|y^{(k)} - y_b^{(k)}\|_p + \|y_b^{(k)} - f^{(k)}\|_p$$

$$\le \left[\int_{B_k(b)} |y^{(k)}(t)|^p \, dt\right]^{1/p} + \eta \le \eta^{1/p} + \eta \le \epsilon$$

when $\epsilon < 2$. Now if $f \in \mathcal{B}_c$ is the function formed by piecing together the $f^{(k)}$,

$$\|y - f\|_T^{(p)} = \sup_t \|(P_{t+T} - P_t)(y - f)\|_p$$

$$\le \sup_k \|(P_{(k+2)T} - P_{kT})(y - f)\|_p$$

$$\le 2\epsilon \ .$$

(Necessity). Suppose that $\|y - f_n\|_T^{(p)} \to 0$ as $n \to \infty$, where the $f_n \in \mathcal{B}_c$. It follows that $\|y^{(k)} - f_n^{(k)}\|_p \to 0$ uniformly in k. Suppose further that the $[y^{(k)}]^p$ are not uniformly integrable; we shall obtain a contradiction. Then, for some $\epsilon > 0$, $\epsilon < \frac{1}{3}$, and for every real number $b > 0$, no matter how large, there is an integer $k' = k'(b, \epsilon)$ such that

$$\int_{B_{k'}(b)} |y^{(k')}(t)|^p \, dt > 3\epsilon \quad .$$

For the same ϵ, let n be a fixed integer so large that $\|y^{(k)} - f_n^{(k)}\|_p \le \epsilon$ for all k. We have,

$$\|y^{(k)} - f_n^{(k)}\|_p \ge \left[\int_{B_k(b)} |y^{(k)}(t) - f_n^{(k)}(t)|^p \, dt\right]^{1/p}$$

for any b, for all k. Put $K_n = \sup_t |f_n(t)|$, and $b = 2K_n$. Then, for $k = k'(b, \epsilon)$,

$$|y^{(k')}(t) - f_n^{(k')}(t)| \geq \Big|\, |y^{(k')}(t)| - |f_n^{(k')}(t)| \,\Big|$$

$$= |y^{(k')}(t)| - |f_n^{(k')}(t)| \geq |y^{(k')}(t)| - K_n \quad \text{for} \quad t \in B_{k'}(b) .$$

Hence,

$$\|y^{(k')} - f_n^{(k')}\|_p \geq \left[\int_{B_k(b)} (|y^{(k')}(t)| - K_n)^p \, dt\right]^{1/p}$$

$$\geq \left| \left[\int_{B_{k'}(b)} |y^{(k')}(t)|^p \, dt\right]^{1/p} - \left[\int_{B_{k'}(b)} K_n^p \, dt\right]^{1/p} \right|$$

$$\geq (3\epsilon)^{1/p} - K_n \, [\mu(B_{k'}(b))]^{1/p}$$

where $\mu(B)$ is the Lebesgue measure of B. But since

$$\epsilon \geq \|y^{(k')} - f_n^{(k')}\|_p \geq \left[\int_{B_{k'}(b)} (2K_n - K_n)^p \, dt\right]^{1/p}$$

$$= K_n \, [\mu(B_{k'}(b))]^{1/p}$$

and since $3\epsilon \geq (3\epsilon)^p$, we have

$$\|y^{(k')} - f_n^{(k')}\|_p \geq 3\epsilon - \epsilon = 2\epsilon$$

which yields a contradiction. †

Clearly, if $y \in \mathcal{L}_o^{(p)}$ satisfies the condition of Proposition 2.5 for $T > 0$ it also satisfies the condition for any other $T' > 0$. Since the value of T is immaterial, we can denote the class of all $y \in \mathcal{L}_o^{(p)}$ which satisfy the condition by $\mathcal{L}_{ou}^{(p)}$. The functions belonging to $\mathcal{L}_{ou}^{(p)}$, or more properly the usual equivalence classes of such functions, belong to $\mathcal{L}_T^{(p)}$ for any

$T > 0$, and as a subset of $\mathcal{L}_T^{(p)}$ this class is denoted $\mathcal{L}_{Tu}^{(p)}$. An immediate corollary of Proposition 2.5 is:

Proposition 2.6 $\mathcal{L}_{Tu}^{(p)}$ is a closed linear subspace of $\mathcal{L}_T^{(p)}$, and is the smallest closed linear subspace containing $\mathcal{B}_c$.

Proof: Obvious.†

With reference to Proposition 2.4 it may be noted that since the set $\mathcal{U}_o$ is a compact subset of $L_p[0, T]$ the functions belonging to $\mathcal{U}_o$ are uniformly integrable. Then the construction given for $\mathcal{U}$ guarantees that the functions belonging to $\mathcal{U}$ satisfy the hypothesis of Proposition 2.5. Hence $\mathcal{U} \subset \mathcal{L}_{Tu}^{(p)}$.

II. PRELIMINARIES ON CAUSAL AND BOUNDED MEMORY TRANSFORMATIONS

Let $\mathcal{U}$ be a metric space whose elements are either functions of a real variable t (time) or are equivalence classes of such functions that are equal a.e. Lebesgue. Correspondingly, let $\mathcal{Y}$ be a Banach space of functions or equivalence classes of functions of t. If both $\mathcal{U}$ and $\mathcal{Y}$ have property (P), the properties of causality and bounded memory for a mapping F from $\mathcal{U}$ into $\mathcal{Y}$ can be defined, and in the usual way: F is <u>causal</u> if $P_t F(u) = P_t F P_t(u)$ for all t and all $u \in \mathcal{U}$; F has <u>bounded memory (d)</u> if $(I - P_t) F(u) = (I - P_t) F (I - P_{t-d})(u)$ for all t and all $u \in \mathcal{U}$. Note that the same symbol, P_t, is being used to denote the linear projection on the past in both $\mathcal{U}$ and $\mathcal{Y}$, but this should cause no confusion.

<u>Proposition 2.7</u> If F is a mapping from $\mathcal{U}$ into $\mathcal{Y}$ that is causal and has bounded memory (d), then for every T >0,

$$(P_{t+T} - P_t) F(u) = (P_{t+T} - P_t) F (P_{t+T} - P_{t-d}) (u) \tag{4}$$

for all t and all $u \in \mathcal{U}$.

Conversely, if equation (4) is satisfied for some T >0 and all t and all $u \in \mathcal{U}$, then F is causal and has bounded memory (d).

<u>Proof:</u> The assertions appear to be obvious. However a proof is given in Appendix A, where the algebraic properties of the P_t are isolated and are used precisely.[†]

In the class of bounded continuous mappings $\mathcal{F} = \mathcal{F}(\mathcal{U}, \mathcal{Y})$, let $\mathcal{F}^o$ denote the subclass of causal mappings, and let $\mathcal{F}^o_d$ denote the subclass of causal mappings with bounded memory (d). Henceforth we only consider metric function spaces $\mathcal{U}$ that have property (P), and $\mathcal{Y}$ can always be chosen to be either one of the $\mathcal{L}^{(p)}_T$ or $\mathcal{B}$, both of which have property (P). Hence $\mathcal{F}^o$ and $\mathcal{F}^o_d$ are defined. In some instances, however, where $\mathcal{B}$ could be used for $\mathcal{Y}$ it may be convenient to take $\mathcal{Y}$ to be a subspace of $\mathcal{B}$ that does not possess property (P), e.g., the subspace of bounded continuous functions $\mathcal{B}_c$. This is alright, because in this situation where the elements of $\mathcal{Y}$ are functions (not equivelance classes of functions) the definition of causality may be replaced by:

F is causal if, for all t and all u,

$$[Fu](s) = [FP_t u](s) \qquad \text{for all} \quad s \le t.$$

An equivalent condition is the apparently weaker statement:

$$[Fu](s) = [FP_s u](s) \qquad \text{for all } s \text{ and all } u.$$

In fact, suppose the second condition holds. Since $\mathcal{U}$ has property (P), $P_t u \in \mathcal{U}$ for all $u \in \mathcal{U}$. Take $t > s$, then

$$[F(P_t u)](s) = [FP_s(P_t u)](s) = [FP_s u](s)$$

and

$$[Fu](s) = [FP_s u](s).$$

Hence,

$$[Fu](s) = [FP_t u](s) \qquad \text{for all} \quad s \le t.$$

Analogous statements hold for the case of bounded memory.

<u>Proposition 2.8</u> $\mathcal{F}^o$ and $\mathcal{F}^o_d$ are closed linear subspaces of $\mathcal{F}$.

<u>Proof:</u> $\mathcal{F}^o_d$ is obviously linear; we need to prove it is closed. First, let

us note the following. If $y \in \mathcal{Y} = \mathcal{L}_T^{(p)}$, then by definition

$$\|y\| = \sup_t \|(P_{t+T} - P_t)\, y\|_p \quad , \quad y \in \mathcal{Y} \quad .$$

On the other hand if $y \in \mathcal{Y} = \mathcal{B}$, then

$$\|y\| = \sup_t |y(t)| = \sup_t \|(P_{t+T} - P_t)\, y\|_{\mathcal{B}} \quad ,$$

where

$$\|(P_{t+T} - P_t)\, y\|_{\mathcal{B}} = \sup_{t \le s \le t+T} |y(s)|$$

is the norm in $\mathcal{B}$ of the truncation of y to $[t, t+T]$. Thus in either case, $\|y\| = \sup_t \|(P_{t+T} - P_t) y\|$, where the norm on the right side is appropriately interpreted.

Now suppose that $F_n \in \mathcal{F}_d^o$ and $\lim_n F_n = F$ where $F \notin \mathcal{F}_d^o$. Put $\Delta_t = P_{t+T} - P_t$ and $\Delta'_t = P_{t+T} - P_{t-d}$. Then

$$\|F_n - F\| = \sup_{u \in} \sup_t \|\Delta_t (F_n u - F u)\|$$

$$= \sup_u \sup_t \|\Delta_t F_n \Delta'_t u - \Delta_t F u\|$$

$$\ge \sup_u \sup_t \Big| \|\Delta_t F_n \Delta'_t u - \Delta_t F \Delta'_t u\| - \|\Delta_t F \Delta'_t u - \Delta_t F u\| \Big|$$

For some t_o and u_o, $\|\Delta_{t_o} F u_o - \Delta_{t_o} F \Delta'_{t_o} u_o\| \ge \alpha$, $\alpha > 0$, whereas for sufficiently large n_o ,

$$\|\Delta_{t_o} F_n (\Delta'_{t_o} u) - \Delta_{t_o} F (\Delta'_{t_o} u)\|$$

$$\le \|F_n (\Delta'_{t_o} u) - F (\Delta'_{t_o} u)\| \le \alpha/2 \quad , \quad n \ge n_o$$

Hence $\|F - F_n\| \geq \alpha/2$, $n \geq n_o$, which is a contradiction. The proof for $\mathcal{F}^o$ is similar. †

If F is any mapping from $\mathcal{U}$ into $\mathcal{B}$, then one can reasonably define the causal part of F, denoted F^o, and the causal and bounded-memory (d) part of F, denoted F^o_d , by

$$[F^o u](t) = [F P_t u](t) , \quad \text{for all } t, \quad \text{all } u \in \mathcal{U}$$

$$[F^o_d u](t) = [F (P_t - P_{t-d}) u](t) , \quad \text{for all } t, \quad \text{all } u \in \mathcal{U} .$$

For the rest of this Section, we assume $\mathcal{Y} = \mathcal{B}$.

<u>Proposition 2.9</u> Let $\mathcal{U}$ have the property that for any $u, u' \in \mathcal{U}$ and any s,t, $d[(P_t - P_s)u , (P_t - P_s)u'] \leq d[u , u']$. Let $F \in \mathcal{F}$. Then a sufficient condition that $F^o \in \mathcal{F}^o$ and $F^o_d \in \mathcal{F}^o_d$ is that F is uniformly continuous on $\mathcal{U}$.

<u>Proof:</u> That F^o_d is causal and has bounded memory (d) is shown by a simple verification. F^o_d is a mapping into $\mathcal{B}$ and is bounded; in fact, for any u and t,

$$|[F^o_d u](t)| = |[F (P_t - P_{t-d}) u] (t)|$$

$$\leq \|F[(P_t - P_{t-d}) u]\| \leq \|F\| .$$

To show that F^o_d is continuous, choose $\epsilon > 0$ arbitrarily. Let $\delta > 0$ be small enough that if u', u satisfy $d[u', u] < \delta$, then $\|Fu' - Fu\| \leq \epsilon/2$. Take any such pair u', u , then there is a t_o such that

$$\|F^o_d u' - F^o_d u\| \leq |[F^o_d u'](t_o) - [F^o_d u](t_o)| + \epsilon/2$$

$$= |[F (P_{t_o} - P_{t_o-d}) u'](t_o) - [F (P_{t_o} - P_{t_o-d}) u](t_o)| + \epsilon/2$$

$$\leq \|F[(P_{t_o} - P_{t_o-d}) u'] - F[(P_{t_o} - P_{t_o-d}) u]\| + \epsilon/2$$

$$\leq \epsilon/2 + \epsilon/2 = \epsilon$$

by the uniform continuity. Hence F_d^o is continuous, and indeed uniformly continuous. The same sort of argument shows that $F^o \in \mathcal{F}$. †

It is to be noted that $F \in \mathcal{F}$ does not by itself necessarily imply that $F^o \in \mathcal{F}^o$, or $F_d^o \in \mathcal{F}_d^o$; i.e., a continuous mapping from $\mathcal{U}$ into $\mathcal{Y}$, where $\mathcal{U}$ and $\mathcal{Y}$ satisfy the conditions of Proposition 2.9, does not necessarily have a continuous causal part, nor a continuous causal and bounded memory part. The condition of uniform continuity is perhaps the most obvious condition that guarantees the continuity of F^o and F_d^o. The following is an example where F_d^o is not continuous, even though $F \in \mathcal{F}$.

Consider the set $\mathcal{E}$ of real-valued functions on R^1 described as follows. Each $u \in \mathcal{E}$ is of the form, for some τ_1, τ_2, $-\infty \leq \tau_1 < \tau_2 \leq \infty$,

$$\begin{aligned} u(t) &= 0 \quad , \quad t \leq \tau_1 \\ &= u(\tau_2) \quad , \quad \tau_1 < t \leq \tau_2 \\ &= 0 \quad , \quad \tau_2 < t \end{aligned}$$

where $-1 \leq u(\tau_2) \leq 1$, and where the convention is made that if $\tau_1 = -\infty$, $u(t) = u(\tau_2)$, $t \leq \tau_2$; and correspondingly, if $\tau_2 = +\infty$, $u(t)$ has a constant value for all $t > \tau_1$. Thus the constant functions, and the functions that are constant except for a single step up from zero or down to zero are included. Obviously $\mathcal{E} \subset \mathcal{B}$, has property (P) and is T-compact.

Let F, a mapping from $\mathcal{E}$ into $\mathcal{B}$, be defined as follows:

$$\begin{aligned} [Fu](t) &= 0 \quad , \quad t \leq \tau_1 \\ &= \phi(u(\tau_2), \tau_2) \quad , \quad \tau_1 < t \leq \tau_2 \\ &= 0 \quad , \quad \tau_2 < t , \end{aligned}$$

where

$$\phi(a,\tau) = |a| \;, \quad \text{if } |\tau| \le 1$$

$$= |\tau| \cdot |a| + 1 - |\tau| \;, \quad \text{if } 1 < |\tau| < \frac{1}{1 - |a|}$$

$$= 0 \;, \quad \text{if } |\tau| \ge \frac{1}{1 - |a|} \;.$$

It will be noted that F carries all the constant functions in $\mathcal{E}$, and in fact all the functions in $\mathcal{E}$ with $\tau_2 = +\infty$, into zero. It is readily verified that F is a bounded continuous mapping from $\mathcal{E}$ (regarded as a metric subspace of $\mathcal{B}$) into $\mathcal{B}$. Actually, the range of F is contained in $\mathcal{E}$.

Now, let $u(t) \equiv 1$ and $u_n(t) \equiv 1 - \frac{1}{n}$, $n = 1, 2, \cdots$. The functions u and $u_n \in \mathcal{E}$, and $u_n \to u$ in $\mathcal{E}$. Consider $F^o u$,

$$[F^o u](t) = [F P_t u](t)$$

$$= u(t) = 1 \;, \quad |t| < 1$$

$$= |t| \, u(t) + 1 - |t| = 1 \;, \quad |t| < \frac{1}{1-1} = \infty$$

i.e., $[F^o u](t) \equiv 1$. On the other hand,

$$[F^o u_n](t) = [F P_t u_n](t)$$

$$= 1 - \frac{1}{n} \;, \quad |t| \le 1$$

$$= |t| \, (1 - \frac{1}{n}) + 1 - |t| = 1 - \frac{|t|}{n} \;, \quad 1 < |t| < n$$

$$= 0 \;, \quad |t| > n \;.$$

Thus $F^o u_n$ does not converge to $F_o u$ in $\mathcal{B}$, although of course $[F^o u_n](t) \to [F^o u](t)$ for each t. Hence F^o is not a continuous mapping. For these particular u and u_n, $F_1^o u = F^o u$, and $F_1^o u_n = F^o u_n$, so it follows that F_1^o is not continuous either.

The following very simple result gives some justification for introducing the concepts of causal part and of bounded-memory causal parts of mappings, at least when the intended use of these mappings is for approximation.

Proposition 2.10 Let F and $F_d^o \in \mathcal{F}$. If for some $\alpha > 0$ there is a $G \in \mathcal{F}_d^o$ such that $\|F - G\| \leq \alpha$, then $\|F - F_d^o\| \leq 2\alpha$. The corresponding statement is true for F^o and $G \in \mathcal{F}^o$.

Proof: For any $\epsilon > 0$ there is a $u \in \mathcal{U}$ and a t_o such that

$$\|G - F_d^o\| \leq |[Gu](t_o) - [F(P_{t_o} - P_{t_o-d})u](t_o)| + \epsilon/2$$

$$= |[G(P_{t_o} - P_{t_o-d})u](t_o) - [F(P_{t_o} - P_{t_o-d})u](t_o)| + \epsilon/2$$

$$\leq \|G - F\| + \epsilon/2 \quad .$$

Hence $\|G - F_d^o\| \leq \|G - F\|$, and $\|F - F_d^o\| \leq 2\alpha$. †

III. FINITE-TIME-INTERVAL PROJECTIONS OF SYSTEMS AND THEIR TRAJECTORIES

The general situation to be discussed next is the following. The kind of system in question consists of an input space $\mathcal{U}$ of functions of time, an output space $\mathcal{Y}$ of functions of time, and a continuous bounded mapping F from $\mathcal{U}$ into $\mathcal{Y}$. The mapping F may or may not be causal and of finite memory, but there is some emphasis on the case where it is. Such a system operates for infinite time. We want to look at pieces of the system corresponding to finite observation intervals for both input and output, and at the relations among such pieces and between them and the entire system. Each real number t can be taken to be the epoch of an observation interval. If the observation intervals are of fixed duration, then as t changes, a trajectory of comparable finite-time systems is generated by the original system. The elementary properties of these trajectories are investigated in this Section.

It is assumed for the remainder of the paper that $\mathcal{U}$ is a subset of $\mathcal{L}_o^p$ for some fixed p, $1 \leq p < \infty$, and that it is T-compact, shift-invariant and has property (P). $\mathcal{U}$ is to be regarded as a metric subspace of $\mathcal{L}_{T+d}^p$ for some T and $d > 0$ as given. Since all $\mathcal{L}_T^p$ spaces with the same p are topologically equivalent, changing T and d changes only the metric on $\mathcal{U}$; it does not affect the T-compactness. We have then always

$$\|u\| = \sup_t \|(P_{t+T} - P_{t+T-d})u\|_p \quad , \quad u \in \mathcal{U}$$

$\mathcal{Y}$ is always either an $\mathcal{L}_T^p$ space or $\mathcal{B}$, the Banach space of bounded functions on R^1 with the uniform norm, or a closed linear subspace of one of these. In the propositions of this Section, whenever $\mathcal{Y}$ is to be one of some particular class of spaces that fact is stated; otherwise it may be any of the spaces just indicated. One slight technical annoyance is that sometimes it is desirable to take $\mathcal{Y}$ to be $\mathcal{B}_c$, the bounded continuous functions regarded as a subspace of $\mathcal{B}$, but this space quite obviously does not have property (P), which is usually needed. It is not always satisfactory just to replace $\mathcal{B}_c$ with $\mathcal{B}$ in every statement, but it will be clear that when necessary, $\mathcal{B}_c$ can be imbedded in $\mathcal{B}$ in order to make the calculations meaningful. $\mathcal{F} = \mathcal{F}(\mathcal{U}, \mathcal{Y})$ is the family of bounded continuous mappings from $\mathcal{U}$ into $\mathcal{Y}$ made into a Banach space with the sup norm, as before. Thus,

$$\|F\| = \sup_{u \in \mathcal{U}} \sup_{t} \|(P_{t+T} - P_t)\, F(u)\|_{\mathcal{Y}} \quad , \quad F \in \mathcal{F}$$

in all cases, where the norm on the right is the L_p norm or the uniform norm as appropriate.

We now introduce notations for the finite-time pieces of a system. Let $T > 0$ and $d > 0$ be given. Put

$$\mathcal{U}'_{t,T} \overset{d}{=} (P_{t+T} - P_{t-d})\,\mathcal{U} \tag{5}$$

$${}'_{t,T}\; u' \overset{d}{=} (P_{t+T} - P_t)\, F u' \quad , \quad u' \in \mathcal{U}'_{t,T} \tag{6}$$

Equation (6) does define a mapping on $\mathcal{U}'_{t,T}$ since u' belongs to the domain of F by property (P). Further, because of shift invariance we can write

$\mathcal{U}_T \stackrel{d}{=} L_t \, \mathcal{U}'_{t,T} = \mathcal{U}'_{t,T}$ for all t. Define $F_{t,T}$ by

$$F_{t,T} \, z \stackrel{d}{=} L_t \, F'_{t,T} \, R_t \, z \tag{7}$$

$$= L_t \, (P_{t+T} - P_t) \, F \, R_t \, z$$

$$= (P_T - P_o) \, L_t F R_t \, z \qquad , \quad z \in \mathcal{U}_T \ .$$

If $\mathcal{Y}$ has property (P), then $F_{t,T}$ is a mapping from $\mathcal{U}_T$ into $\mathcal{Y}$, and clearly it is bounded and continuous. But $F_{t,T}$ can also always be regarded as a mapping into a smaller space, denoted $\mathcal{Y}_T$. If $\mathcal{Y} = \mathcal{L}^p_T$, then $F_{t,T}$ is a bounded continuous mapping into $\mathcal{Y}_T = L_p(0, T]$, and with the same norm as if its range space is taken to be $\mathcal{Y}$. Similarly, if $\mathcal{Y} = \mathcal{B}$, $F_{t,T}$ is a bounded continuous mapping into $\mathcal{Y}_T = \mathcal{B}(0,T]$ with the same norm; even if $\mathcal{Y} = \mathcal{B}_c$, $F_{t,T}$ is a bounded continuous mapping into $\mathcal{Y}_T = \mathcal{B}_c(0,T]$, although it is not a mapping into $\mathcal{Y}$.

If T is fixed throughout a calculation we write simply F_t for $F_{t,T}$. It often avoids confusion to write

$$F_t = (P_T - P_o) \, L_t \, F \, R_t \, (P_T - P_{-d})$$

even though the projection on the right is redundant. When we are dealing with mappings F with finite memory, d is usually chosen to be the duration of the memory; however the above definitions are to be applied in the general case, whether F is causal with finite memory or not. Causality and finite memory are not to be assumed in what follows unless explicitly stipulated.

Let $\pi_t : \mathcal{F}(\mathcal{U}, \mathcal{Y}) \to \mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ be the mapping that carries F into F_t according to equation (7).

Proposition 2.11 The mapping π_t is linear and continuous; in fact $\|\pi_t F\| \le \|F\|$.

Proof: The linearity is obvious. Also

$$\|\pi_t F\| = \|F_t\| = \sup_{u \in \mathcal{U}_T} \|(P_T - P_o) L_t F R_t u\|$$

$$\le \sup_{u \in \mathcal{U}_T} \|F R_t u\| \quad \le \sup_{\mathcal{U}} \|Fz\|_{\mathcal{Y}} \quad = \|F\| ,$$

where the norm on the left is for the space $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$. †

For each t, $F_t = \pi_t F$ is an element of $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$, so as t runs through R^1 a trajectory is generated in $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ corresponding to F. If F is a time-invariant mapping this "trajectory" reduces to a single point, of course, but we are interested in time-invariant systems only as a special case. Since in general for time-varying systems these trajectories describe the evolution of the systems, we wish to investigate their properties. Note that the trajectories depend on T; however, for now, we keep T fixed arbitrarily.

Proposition 2.12 Let $\mathcal{U} \subset \mathcal{L}^q_{T'}$, $1 \le q < \infty$, and $\mathcal{Y}$ be $\mathcal{L}^p_T$, $1 \le p < \infty$, or $\mathcal{B}_c$. Then, for any $F \in \mathcal{F}(\mathcal{U}, \mathcal{Y})$ the trajectories $F_t = \pi_t F$ with values in $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ are continuous in t. Furthermore, if $\mathcal{H}$ is a compact subset of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$, then the trajectories $F_t = \pi_t F$, $F \in \mathcal{H}$, are equicontinuous functions of t.

Proof: Suppose $\mathcal{Y} = \mathcal{L}_T^p$, then

$$\|F_t - F_{t+h}\| - \sup_{u \in \mathcal{U}_T} \|(F_t - F_{t+h}) u\|_p$$

$$= \sup_{\mathcal{U}_T} \|(P_T - P_o) L_t F R_t u - (P_T - P_o) L_{t+h} F R_{t+h} u\|_p$$

$$\leq \sup_{\mathcal{U}_T} \|(P_T - P_o) L_t F R_t u - (P_T - P_o) L_{t+h} F R_t u\|_p$$

$$+ \sup_{\mathcal{U}_T} \|(P_T - P_o) L_{t+h} F R_t u - (P_T - P_o) L_{t+h} F R_{t+h} u\|_p \quad .$$

Denote the first term on the right-hand side of the inequality by I, the second by II. Then,

$$I \leq \sup_{\mathcal{U}_T} \|L_t (P_{T+t} - P_t) F R_t u - L_{t+h} (P_{T+t} - P_t) F R_t u\|_p$$

$$+ \sup_{\mathcal{U}_T} \|L_{t+h} (P_{T+t} - P_t) F R_t u - L_{t+h} (P_{T+t+h} - P_{t+h}) F R_t u\|_p$$

Denote the two terms on the right-hand side of this inequality by I_a, I_b, respectively. Then

$$I_b = \sup_{\mathcal{U}_T} \|L_{t+h} [P_{T+t} - P_{T+t+h} + P_{t+h} - P_t] \; F R_t u\|_p$$

$$\leq \sup_{\mathcal{U}_T} \|(P_{T+t} - P_{T+t+h}) F R_t u\|_p$$

$$+ \sup_{\mathcal{U}_T} \| P_{t+h} - P_t) F R_t u\|_p \quad .$$

Now, $F(R_t \mathcal{U}_T)$ is a compact subset of $\mathcal{Y}$ since $\mathcal{U}_T$ is compact in L_q (by the T-compactness of $\mathcal{U}$) and $F R_t$ is continuous. Let y_i, $i = 1, \cdots, N$, be a set of points in $\mathcal{Y}$ such that the balls of radius ε about the y_i cover $F(R_t \mathcal{U}_T)$. Then the first term in the expression dominating I_b is in turn dominated by

$$\sup_{\mathcal{U}_T} \min_{i=1,\cdots,N} \{ \| (P_{T+t} - P_{T+t+h}) (F R_t u - y_i) \|_p + \| (P_{T+t} - P_{T+t+h}) y_i \|_p \}$$

Let h be sufficiently small that

$$\| (P_{T+t} - P_{T+t+h}) y_i \|_p \leq \varepsilon$$

for all $i = 1, \cdots, N$. Then for any such h the above expression has a value $\leq 2\varepsilon$. The second term in the expression dominating I_b can be treated in the same way, so we have that for some $h_1 > 0$, $I_b \leq 4\varepsilon$ whenever $|h| < h_1$.

The term I_a can be written as

$$\sup_{\mathcal{U}_T} \| (L_t - L_{t+h}) (P_{t+T} - P_t) F R_t u \|_p \quad .$$

Since $(P_{t+T} - P_t) F R_t \mathcal{U}_T$ is a compact subset of L_p one can choose a finite set $\{z_i\}$, $i = 1, \cdots, M$, of elements of L_p such that the balls of radius ε about the z_i cover $(P_{T+t} - P_t) F R_t \mathcal{U}_T$. Then, since for h sufficiently small $\| (L_t - L_{t+h}) z_i \|_p \leq \varepsilon$ for all $i = 1, \cdots, M$, we have, very much as above, that for some $h_2 > 0$, $I_a \leq 2\varepsilon$ whenever $|h| \leq h_2$.

To bound II, we have

$$II = \sup_{\mathcal{U}_T} \| L_{t+h} (P_{T+t+h} - P_{t+h}) [F R_t u - F R_{t+h} u] \|_p$$

$$\leq \sup_{\mathcal{U}_T} \| (P_{2T+t} - P_{t-T}) [F R_t u - F R_t R_h u] \|_p$$

when $|h| \le T$. Now $\mathcal{K} = \bigcup_{|h| \le T} R_h\, \mathcal{U}_T$ is a compact subset of L_q (as in Proposition 2.4) and $(P_{2T+t} - P_{t-T})\, F R_t$ is a uniformly continuous mapping from $\mathcal{K}$ into L_p. Hence, there is an $\eta > 0$ so that

$$\| (P_{t+2T} - P_{t-T}) F R_t u' - (P_{t+2T} - P_{t-T}) F R_t u'' \|_p \le \varepsilon$$

whenever $\|u' - u''\|_q \le \eta$. Let $\{w_i\}$, $i = 1, \cdots, K$, be the centers of balls of radius η that cover $\mathcal{K}$. Then

$$II \le \sup_{\mathcal{U}_T} \{ \| (P_{t+2T} - P_{t-T}) F R_t u - (P_{t+2T} - P_{t-T}) F R_t w_i \|_p$$

$$+ \| (P_{t+2T} - P_{t-T}) F R_t w_i - (P_{t+2T} - P_{t-T}) F R_t R_h u \|_p \}$$

Let $h_3 > 0$ be small enough that $\|R_h w_i - w_i\|_q \le \eta/2$ for all $i = 1, \cdots, K$ whenever $|h| < h_3$, and temporarily fix such an h. With this fixed value of h, there is u_o so that the supremum in the inequality above is realized to within ϵ by $u = u_o$. This gives

$$II \le \| (P_{t+2T} - P_{t-T}) F R_t u_o - (P_{t+2T} - P_{t-T}) F R_t w_i \|_p$$

$$+ \| (P_{t+2T} - P_{t-T}) F R_t w_i - (P_{t+2T} - P_{t-T}) F R_t R_h u_o \|_p$$

$$+ \varepsilon \;, \quad \text{for all } i = 1, \cdots, K \;.$$

There is at least one w_i so that $\|u_o - w_i\|_q \le \eta/2$; choose such a w_i. Then the first term on the right is $\le \varepsilon$. With this particular w_i,

$$\|R_h u_o - w_i\|_q \le \|R_h u_o - R_h w_i\|_q + \|R_h w_i - w_i\|_q$$

$$= \|u_o - w_i\|_q + \|R_h w_i - w_i\|_q$$

$$\le \eta/2 + \eta/2 = \eta$$

so the second term is also $\leq \varepsilon$. Thus for $|h| < h_3$, $II \leq 3\varepsilon$. Combining these estimates gives the result that if $|h| \leq \max(h_1, h_2, h_3)$ then

$$\|F_t - F_{t+h}\| \leq I_a + I_b + II \leq 2\varepsilon + 4\varepsilon + 3\varepsilon = 9\varepsilon .$$

This proves the assertion for a single trajectory with $\mathcal{Y} = \mathcal{L}^p_T$. An inspection of the proof will show that if $F \in \mathcal{H}$, $\mathcal{H}$ a compact subset of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$, then the compact sets chosen above can each be replaced by compact sets chosen independently of F in $\mathcal{H}$. For example, the compact set $F(R_t \mathcal{U}_T)$ is replaced by $\mathcal{H}(R_t \mathcal{U}_T)$, which is a compact subset of $\mathcal{Y}$ since $\mathcal{H}$ restricted to $R_t \mathcal{U}_T$ is a compact set of bounded continuous mappings, and $R_t \mathcal{U}_T$ is a compact subset of L_q. Also the mappings $(P_{2T+t} - P_{t-T}) F R_t$ restricted to $\mathcal{K}$, $F \in \mathcal{H}$, are equicontinuous by Ascoli's theorem. These facts yield the assertion that the F_t are equicontinuous.

The proof of the assertions for the case $\mathcal{Y} = \mathcal{B}_c$ is similar, although obviously some modifications are required. The details are not given.†

Two consistency relations are introduced for the trajectories F_t. The second of these will also be used as an interpolation formula. Conditions under which they hold are given in the Proposition to follow.

$$(P_{T-\eta} - P_o) L_\eta F_t R_\eta (P_{T-\eta} - P_{-d})$$

$$= (P_{T-\eta} - P_o) F_{t+\eta} (P_{T-\eta} - P_{-d}) , \quad 0 \leq \eta \leq T \qquad (8)$$

$$F_{t+\eta} = (P_{T-\eta} - P_o)\, L_\eta\, F_t\, R_\eta\, (P_{T-\eta} - P_{-d})$$

$$+ (P_T - P_{T-\eta})\, L_{\eta-T}\, F_{t+T}\, R_{\eta-T}\, (P_T - P_{T-\eta-d}),$$

$$0 \le \eta \le T \qquad (9)$$

Proposition 2.13 i) If $F_t = \pi_t F$, $F \in \mathcal{F}(\mathcal{U}, \mathcal{Y})$, then F_t satisfies equation (8) for all t. ii) If $F \in \mathcal{F}^{o}_{d}(\mathcal{U}, \mathcal{Y})$, then F_t satisfies equation (9) for all t. iii) If H_t, H_{t+T}, $H_{t+\eta}$ are any mappings from $\mathcal{U}_T$ into $\mathcal{Y}_T$ that satisfy equation (9), then they satisfy equation (8).

Proof: The proof of i) is given by the calculation:

$$(P_{T-\eta} - P_o)\, L_\eta\, [(P_T - P_o)\, L_t\, F\, R_t\, (P_T - P_{-d})]\, R_\eta\, (P_{T-\eta} - P_{-d})$$

$$= (P_{T-\eta} - P_o)\, (P_{T-\eta} - P_{-\eta})\, L_{\eta+t}\, F\, R_{\eta+t}\, (P_{T-\eta} - P_{-d-\eta})\, (P_{T-\eta} - P_{-d})$$

$$= (P_{T-\eta} - P_o)\, L_{\eta+t}\, F\, R_{\eta+t}\, (P_{T-\eta} - P_{-d})$$

$$= (P_{T-\eta} - P_o)\, [(P_T - P_o)\, L_{t+\eta}\, F\, R_{t+\eta}\, (P_T - P_{-d})]\, (P_{T-\eta} - P_{-d})$$

$$= (P_{T-\eta} - P_o)\, F_{t+\eta}\, (P_{T-\eta} - P_{-d}) \;.$$

To prove ii) we use i) for the first term on the right side of equation (b) and make an analogous calculation for the second term. Then the right-hand side of equation (9) becomes

$$(P_{T-\eta} - P_o)\, F_{t+\eta}\, (P_{T-\eta} - P_{-d}) + (P_T - P_{T-\eta})\, F_{t+\eta}\, (P_T - P_{T-\eta-d}) \qquad (10)$$

If F is causal with bounded memory (d), then so are all the F_t, and this expression reduces to

$$(P_{T-\eta} - P_o) F_{t+\eta} + (P_T - P_{T-\eta}) F_{t+\eta} = F_{t+\eta} \quad ,$$

which proves ii). iii(can be verified immediately by substituting $F_{t+\eta}$ from equation (9) into the right-hand side of equation (8).†

The consistency condition (9), if required to hold for all t and all η, $0 \le \eta \le T$, is not quite enough to guarantee that F is causal with bounded memory (d). It does guarantee something a little weaker, and to state this we use the definition: F is weakly causal and of bounded memory (d) if for every A > 0, and for all t,

(1) $$(P_{t+A} - P_t) F (P_{t+A} - P_{t-d}) = (P_{t+A} - P_t) F (P_a - P_{t-d})$$

whenever $t + A \le a$, and

(2) $$(P_{t+A} - P_t) F (P_{t+A} - P_{t-d}) = (P_{t+T} - P_t) F (P_{t+A} - P_b)$$

whenever $b \le t - d$.

This definition rules out non-causality and non-bounded-memory (d) that depend on interactions between past and future.

Proposition 2.14 If $F \in \mathcal{F} (\mathcal{U} , \mathcal{Y})$, then equation (9) is satisfied for all T > 0, all t, and all η, $0 \le \eta \le T$, iff F is weakly causal and of bounded memory (d).

Proof: The right-hand side of equation (9) is given in different form in (10); consider the first term of (10). Since (9) is satisfied, we must have that

$$(P_{T-\eta} - P_o) F_{t+\eta} (P_{T-\eta} - P_{-d})$$
$$= (P_{T-\eta} - P_o) F_{t+\eta} (P_T - P_{-d}) \quad .$$

This may be rewritten,

$$L_{t+\eta} (P_{T+t} - P_{t+\eta}) F (P_{T+t} - P_{t+\eta-d}) R_{t+\eta}$$

$$= L_{t+\eta} (P_{T+t} - P_{t+\eta}) F (P_{T+t+\eta} - P_{t+\eta-d}) R_{t+\eta}$$

which is equivalent to

$$(P_{T+t} - P_{t+\eta}) F (P_{T+t} - P_{t+\eta-d}) = (P_{T+t} - P_{t+\eta}) F (P_{T+t+\eta} - P_{t+\eta-d}) .$$

Put $a = T+t+\eta$, $s = T+\eta$ and $A = T-\eta$. Then this is in the form of condition (1) for weak causality and bounded memory (d), and a, s and A can be given arbitrarily by choosing $\eta > 0$, t and $T > \eta$. An analogous argument applied to the second term of (10) yields condition (2). The converse follows immediately from equation (10) and the definition of $H_{t+\eta}$. †

From a family of mappings carrying $\mathcal{U}_T$ into $\mathcal{Y}_T$ it is possible under certain circumstances to synthesize a mapping from $\mathcal{U}$ into $\mathcal{Y}$. We want to be able to do this, because we want to be able to go from trajectories $\{F_t\}$ back to an overall system mapping F. The transformations ρ_s to be defined below accomplish this. If π denotes the transformation carrying F into a trajectory $\{F_t\}$, then the ρ_s are roughly inverse to π. However the situation is a little complicated in general, and ρ_s and π are inverse to each other only when F is causal with bounded memory of sufficiently short duration. These comments are made precise in what follows.

We use the notations:

$$\Delta_{n,t} = (P_{t-(n-1)T} - P_{t-nT})$$

$$\Delta'_{n,t} = (P_{t-(n-1)T} - P_{t-nT-d}) \quad ,$$

where $T > 0$, $d > 0$ are fixed. Let $\mathcal{G}$ be a bounded subset of $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ with the additional property that the $G \in \mathcal{G}$ are equicontinuous. Let $\{G_n\}$ be a sequence from $\mathcal{G}$. For any real number t, define

$$G^t = \sum_{-\infty}^{\infty} \Delta_{n,t} R_{t-nT} G_n L_{t-nT} \Delta'_{n,t} \tag{11}$$

It is clear that G^t is a mapping from $\mathcal{U}$ into $\mathcal{Y}$ if $\mathcal{Y} = \mathcal{L}^p_T$ or $\mathcal{B}$; however, see Appendix A for a formal justification of the infinite sum.

Proposition 2.15 G^t as defined by equation (11) is an element of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$, where $\mathcal{Y}$ is either $\mathcal{L}^p_T$ or $\mathcal{B}$.

Proof: Take $\mathcal{Y} = \mathcal{L}^p_T$. Given $\varepsilon > 0$, let $\delta > 0$ be such that $\|z_1 - z_2\| \leq \delta$, $z_1, z_2 \in \mathcal{U}_T$, implies $\|G_n(z_1) - G_n(z_2)\| \leq \varepsilon$ for all $n = 1, 2, \cdots$, as is possible from the hypothesis on $\mathcal{G}$. For $u_1, u_2 \in \mathcal{U}$, $\|u_1 - u_2\| \leq \delta$, one has

$$\|G^t(u_1) - G^t(u_2)\| = \sup_s \|(P_{s+T} - P_s)[G^t(u_1) - G^t(u_2)]\|_p$$

$$\leq \sup_k \| \Delta_{k,t} R_{t-kT} G_k L_{t-kT} \Delta'_{k,t}(u_1)$$

$$+ \Delta_{k+1,t} R_{t-(k+1)T} G_{k+1} L_{t-(k+1)T} \Delta'_{k+1,t}(u_1)$$

$$- \Delta_{k,t} R_{t-kT} G_k L_{t-kT} \Delta'_{k,t}(u_2)$$

$$- \Delta_{k+1,t} R_{t-(k+1)T} G_{k+1} L_{t-(k+1)T} \Delta'_{k+1,t}(u_2)\|_p$$

$$\leq 2 \|\Delta_{n,t} R_{t-nT} G_n L_{t-nT} \Delta'_{n,t}(u_1)$$

$$- \Delta_{n,t} R_{t-nT} G_n L_{t-nT} \Delta'_{n,t} (u_2) \|_p + \varepsilon$$

$$= 2 \, \| (P_T - P_o) \, [G_n L_{t-nT} \Delta'_{n,t} (u_1)$$

$$- G_n L_{t-nT} \Delta'_{n,t} (u_2)] \, \|_p + \varepsilon$$

for some n. But

$$\| L_{t-nT} \Delta'_{nt} u_1 - L_{t-nT} \Delta'_{nt} u_2 \| \le \| u_1 - u_2 \| \le \delta \quad ,$$

and $L_{t-nT} \Delta'_{nt} u \in \mathcal{U}_T$, hence

$$\| G^t (u_1) - G^t (u_2) \| \le 2 \varepsilon + \varepsilon = 3 \varepsilon \quad .$$

The boundedness of G^t follows similarly. The same proof holds for $\mathcal{Y} = \mathcal{B}$ if the L_p norms are changed to uniform norms.†

The transformation that carries the sequence $\{G_n\}$ of equicontinuous mappings into G^t is denoted ρ_t. Note that the equicontinuity condition is natural, since, when we to the other way we have that the $\pi_t F$, $F \in \mathcal{F}(\mathcal{U}, \mathcal{Y})$, are equicontinuous with respect to t.

Proposition 2.16 The transformation ρ_t is continuous in the following sense: if there are two sequences of equicontinuous mappings from $\mathcal{U}_T$ to $\mathcal{Y}_T$, $\{G_n\}$ and $\{\tilde{G}_n\}$, and $\| G_n - \tilde{G}_n \| \le \delta$ for all integers n for some $\delta = \delta(\varepsilon)$, then $\| \rho_t (\{G_n\}) - \rho_t (\{\tilde{G}_n\}) \| \le \varepsilon$.

Proof: Again take $\mathcal{Y} = \mathcal{L}^p_T$. Then,

$$\| \rho_t (\{G_n\}) - \rho_t (\{\tilde{G}_n\}) \|$$

$$\le 2 \sup_{\mathcal{U}} \| \Delta_{n,t} R_{t-nT} [G_n L_{t-nT} \Delta_{n,t} (u)$$

$$- \tilde{G}_n L_{t-nT} \Delta'_{n,t} (u)] \, \|_p + \varepsilon$$

for some n by a calculation very similar to that in the previous proof. But the right side of this inequality can be rewritten as

$$2 \sup_{\mathcal{U}} \|(G_n - \tilde{G}_n)(L_{t-nT} \Delta'_{n,t} u)\|_p + \varepsilon$$

$$= 2 \sup_{z \in \mathcal{U}_T} \|(G_n - \tilde{G}_n) z\|_p + \varepsilon$$

since, by the shift invariance of $\mathcal{U}$, any $z \in \mathcal{U}_T$ can be obtained by truncating some u by $(P_{t-(n-1)T} - P_{t-nT-\delta})$ for arbitrary t,n. Hence, if $\|G_n - \tilde{G}_n\|$ is sufficiently small for all n,

$$\|\rho_t(\{G_n\}) - \rho_t(\{\tilde{G}_n\})\| \leq 2\varepsilon + \varepsilon = 3\varepsilon \quad .$$

Again, the same proof holds for $\mathcal{Y} = \mathcal{B}$ if the L_p norms are changed to sup norms. †

Proposition 2.17 If $F \in \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, then for any t, $\{\pi_{t-nT} F\}$ is a family of equicontinuous causal mappings from $\mathcal{U}_T$ to $\mathcal{Y}_T$ with bounded memory (d), and $F = \rho_t(\{\pi_{t-nT} F\})$. Conversely, if $\{G_n\}$ is a sequence of equicontinuous causal mappings from $\mathcal{U}_T$ to $\mathcal{Y}_T$ with bounded memory (d), then $\rho_t(\{G_n\}) \in \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$ and $G_k = \pi_{t-kT} \circ \rho_t(\{G_n\})$.

Proof: The assertion that the $\pi_{t-nT} F$ are causal with bounded memory (d) is obvious; indeed all the $\pi_s F$ are causal with bounded memory (d). Further,

$$\rho_t(\{\pi_{t-nT} F\}) =$$

$$= \sum_{-\infty}^{\infty} \Delta_{n,t} R_{t-nT} [(P_T - P_o) L_{t-nT} F R_{t-nT} (P_T - P_{-d})] L_{t-nT} \Delta'_{n,t}$$

$$= \sum_{-\infty}^{\infty} \Delta_{n,t} F \Delta'_{n,t} = F \quad .$$

It is also obvious that $\rho_t(\{G_n\})$ is causal with bounded memory (d). The second inversion identity is given by the calculation:

$$\pi_{t-kT} \circ \rho_t(\{G_n\})$$

$$= P_T - P_o) L_{t-kT} \left[\sum_{-\infty}^{\infty} \Delta_{n,t} R_{t-nT} G_n L_{t-nT} \Delta'_{n,t}\right] \cdot R_{t-kT} (P_T - P_{-d})$$

$$= L_{t-kT} \Delta_{k,t} R_{t-kT} G_k L_{t-kT} \Delta'_{k,t} R_{t-kT} (P_T - P_{-d})$$

$$= (P_T - P_o) G_k (P_T - P_{-d}) = G_k \quad . \dagger$$

When F is not causal with bounded memory, the operations π and ρ_t obviously cannot be inverse to each other because some information about F is lost in the truncations given by the π_t which cannot be restored. The sense in which they are approximately inverse to each other is given in the next Proposition.

Proposition 2.18 Let $\{F_t\}$, $-\infty < t < \infty$ be a family of mappings in $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ which is bounded and in which the F_k are equicontinuous. For any fixed s consider $\{F_{s-nT}\}$, $n = \cdots, -2, -1, 0, 1, 2, \cdots$. Put

$$H_{s-nT+\eta} \stackrel{d}{=} (P_{T-\eta} - P_o) L_\eta F_{s-nT} R_\eta (P_{T-\eta} - P_{-d})$$

$$+ (P_T - P_{T-\eta}) L_{\eta-T} F_{s-(n-1)T} R_{\eta-T} (P_T - P_{T-\eta-d}),$$

$$0 \leq \eta < T \quad . \qquad (12)$$

This defines H_t for all t, and $H_{s-nT} = F_{s-nT}$. Further, define

$$H^{(1)} \stackrel{d}{=} \rho_s(\{H_{s-nT}\})$$

$$H_t^{(1)} \stackrel{d}{=} \pi_t \circ \rho_s(\{H_{s-nT}\}) = \pi_t H^{(1)}$$

$$H^{(2)} \stackrel{d}{=} \rho_s (\{\pi_{s-nT} H^{(1)}\}) = \rho_s (\{H^{(1)}_{s-nT}\})$$

Then,

i) $H^{(1)}_t = H_t$ for all t

ii) $H^{(2)} = H^{(1)}$

iii) If F_t satisfies equation (9), then $H_t = F_t$ and $H^{(1)}_t = F_t$

Proof: By the definitions, $H^{(1)}_t$ is given by

$$H^{(1)}_t = (P_T - P_o) L_t \left[\sum_{-\infty}^{\infty} \Delta_{n,s} R_{s-nT} H_{s-nT} L_{s-nT} \Delta'_{n,s} \right] R_t (P_T - P_{-d}) .$$

At most two terms from the infinite sum can contribute anything, by virtue of the projection $(P_T - P_o)$. Let k be that integer such that

$$s - kT \leq t < s - (k - 1) T \quad ,$$

and let

$$\eta = t - (s - kT) \quad .$$

Then, since $F_{s-nT} = H_{s-nT}$, the expression for $H^{(1)}_t$ above reduces to the expression for $H_t = H_{s-nT+\eta}$ given by equation (12). Thus $H^{(1)}_t = H_t$ for all t, and $H^{(2)} = H^{(1)}$ follows from this equality and from the definitions. The assertion iii) is obvious, since quation (12) becomes equation (9) if H is replaced by F. †

We conclude this Section with a simple error bound on the interpolated $H_{t+\eta}$ as given by equation (9) when H_t and H_{t+T} are in error.

Proposition 2.19 Let $\|\tilde{F}_t - F_t\| \leq \varepsilon$ and $\|\tilde{F}_{t+\eta} - F_{t+T}\| \leq \varepsilon$. Then, if $\tilde{F}_{t+\eta}$ and $F_{t+\eta}$ are each given by equation (9) in terms of $\tilde{F}_t$, $\tilde{F}_{t+T}$ and

F_t, F_{t+T}, respectively, $\|\tilde{F}_{t+\eta} - F_{t+\eta}\| \le 2\varepsilon$.

Proof: Expressing $F_{t+\eta}$, $\tilde{F}_{t+\eta}$ in terms of F_t, F_{t+T} and $\tilde{F}_t$, $\tilde{F}_{t+T}$ from equation (9) yields

$$\|\tilde{F}_{t+\eta} - F_{t+\eta}\| \le \sup_{\mathcal{U}_T} \|\tilde{F}_t R_\eta (P_{T-\eta} - P_{-d}) u - F_t R_\eta (P_{T-\eta} - P_{-d})\|$$

$$+ \sup_{\mathcal{U}_T} \|\tilde{F}_{t+T} R_{\eta-T} (P_T - P_{T-\eta-d}) - F_{t+T} R_{\eta-T} (P_T - P_{T-\eta-d})\| \quad (13)$$

Now, by the properties of $\mathcal{U}$,

$$\bigcup_{0\le\eta\le T} R_\eta (P_{T-\eta} - P_{-d}) \mathcal{U}_T \subset \mathcal{U}_T \quad .$$

Hence, $\sup_{\mathcal{U}_T} \|\tilde{F}_t R_\eta (P_{T-\eta} - P_{-d}) u - F_t R_\eta (P_{T-\eta} - P_{-d})\|$

$$\le \sup_{\mathcal{U}_T} \|\tilde{F}_t u - F_t u\| \le \varepsilon \quad .$$

The second term in the inequality (13) is also dominated by ε , by essentially the same argument.†

IV. TRAJECTORIES OF THE FINITE-TIME PROJECTIONS FOR CLASSES OF SYSTEMS

We now consider a class of systems $\mathcal{S} = (\mathcal{Y}, f, \mathcal{X}, \mathcal{U})$ in its natural representation form, $\mathcal{S}_o = (\mathcal{Y}, g, \mathcal{H}, \mathcal{U})$, where $\mathcal{U}$ is a shift-invariant, T-compact subset of $\mathcal{L}^p_{T+d}$ with property (P), and where $\mathcal{Y}$ is $\mathcal{L}^p_T$ or $\mathcal{B}$. $\mathcal{H}$ is, of course, a subset of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$; further hypotheses on $\mathcal{H}$ will be made as needed. Each $F \in \mathcal{H}$ will generate a trajectory $\{F_t\} \in \mathcal{H}_T$, whether F is causal with bounded memory less than or equal to d, or not. If $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, then each of these trajectories will yield the corresponding F through the mapping ρ. We investigate some basic properties of these families of trajectories.

Temporarily take $T > 0$ to be fixed. Let $\mathcal{M}$ be the closed linear subspace of the Banach space $\mathcal{F}(\mathcal{U}, \mathcal{Y})$ generated by $\mathcal{H}$, and let $\mathcal{M}_t = \pi_t \mathcal{M}$. $\mathcal{M}_t$ is a linear subset of $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$; its closure, $\overline{\mathcal{M}}_t$, is the closed linear subspace of $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ generated by $\pi_t \mathcal{H}$. We define $\mathcal{S}_o$ (or $\mathcal{S}$) to be a <u>linearly predictable class of systems with respect to T</u> if each mapping π_t is 1:1 from $\mathcal{M}$ onto $\mathcal{M}_t$, $t \in R^1$. When $\mathcal{S}_o$ is a linearly predictable class a prediction mapping $\theta(t,s)$ carrying H_t into H_s, $t \le s$ can be defined by

$$\theta(t,s) = \pi_s \circ \pi_t^{-1} \quad , \quad -\infty < t \, , \; s < \infty \; .$$

For each t, s, $\theta(t,s)$ is obviously a linear transformation with domain $\mathcal{M}_t$ and range $\mathcal{M}_s$.

The intuitive meaning of $\mathscr{S}_o$ being a linearly predictable class is that no two trajectories associated with the $F \in \mathscr{M}$ corresponding to $\mathscr{S}_o$ can cross or touch and be at the common point at the same time. Two trajectories can cross or touch provided the time of arrival at the common point is different for the two. A class of systems consisting of a single system ($\mathscr{H}$ has only one element) is always predictable in the sense of this definition.

We further define a <u>stationarily predictable class of systems with respect to T</u> to be a class $\mathscr{S}_o$ with the property that whenever $\pi_t F = \pi_s G$, F and $G \in \mathscr{M}$, then $\pi_{t+a} F = \pi_{s+a} G$ for all real numbers a. Intuitively, this implies that the systems F and G have trajectories which as geometrical entities are identical. Furthermore, no individual trajectory can cross itself. If the definition is weakened to read: $\pi_t F = \pi_s G$, F and $G \in \mathscr{M}$, implies $\pi_{t+a} F = \pi_{s+a} G$ for all $a \geq 0$, we call the class $\mathscr{S}_o$ a <u>future-time (f.t.) stationarily predictable class with respect to T.</u>

If either of F or G is not causal with bounded memory (d) it is obviously possible that $\pi_a F = \pi_a G$ for all a without F and G being the same. In this case $\mathscr{S}_o$ can be stationarily predictable without being linearly predictable. <u>A fortiori</u>, $\mathscr{S}_o$ can be f.t. stationarily predictable without being linearly predictable. However, if the $\mathscr{H}$ associated with $\mathscr{S}_o$ is a subset of $\mathscr{F}_d^o(\mathscr{U}, \mathscr{Y})$, so is $\mathscr{M}$. Then if for some t, $\pi_t F = \pi_t G$, it follows from stationary predictability that $\pi_a F = \pi_a G$ for all a, and hence by Proposition 2.17 that F = G. Thus, in this situation stationary predictability implies linear predictability. Under the same condition that

$\mathcal{H} \subset \mathcal{F}^{o}_{d}(\mathcal{U}, \mathcal{Y})$, if $\mathcal{S}_o$ is only f.t. stationarily predictable, the situation is complicated a little, but can be interpreted in much the same way as will be seen below.

In case $\mathcal{S}_o$ is linearly and stationarily predictable the prediction mapping $\theta(t,s)$ can be written as a function of the difference s-t only, once the domain has been defined properly. In fact, suppose to start with that $F' \in \mathcal{M}_t$ and also $F' \in \mathcal{M}_{t+a}$. Then $F' = \pi_t F$ for some $F \in \mathcal{M}$; and also $F' \in \pi_{t+a} G$ for some $G \in \mathcal{M}$.

Thus, $\theta(t,s)F' = \pi_s \circ \pi_t^{-1}(\pi_t F) = \pi_s F$ and $\theta(t+a, s+a)F' = \pi_{s+a} \circ \pi_{t+a}^{-1}(\pi_{t+a} G) = \pi_{s+a} G$. By the definition of a stationarily predictable class, $\pi_s F = \pi_{s+a} G$; hence $\theta(t,s)F' = \theta(t+a, s+a)F'$. Now (with a slight abuse of notation) let $\theta(\tau)F' = \theta(t,s)F'$, $s = t+\tau$, for all F' such that for some t, $F' \in \mathcal{M}_t$.

This definition is meaningful, because if more than one pair (t, s) satisfy the conditions they all yield the same $\theta(t,s)F'$. The domain of $\theta(\tau)$, for any τ, will now include $\bigcup_{t \in R^1} \mathcal{M}_t$; extend this by linearity to $\mathcal{N}$ = linear span $\{\bigcup_{t \in R^1} \mathcal{M}_t\}$. The family $\{\theta(\tau)\}$, $\tau \in R^1$, is now a one-parameter group of linear transformations on $\mathcal{N}$. We note that $\mathcal{N} \subset \mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$. In fact, the elements of $\mathcal{N}$ are of the form

$$F' = \sum_{n=1}^{N} \alpha_n (P_T - P_o) L_{t_n} F_n R_{t_n} (P_T - P_{-d})$$

$$= (P_T - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n} \right) (P_T - P_{-d})$$

where $\{t_1, \cdots, t_N\}$ is an arbitrary finite set of real numbers, as is also

$\{\alpha_1, \cdots, \alpha_N\}$, and $F_1, \ldots, F_N$ are each elements of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$. Since $\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n}$ is also a bounded, continuous mapping, we can denote it by $F \in \mathcal{F}(\mathcal{U}, \mathcal{Y})$. Then,

$$F' = (P_T - P_o) F (P_T - P_{-d}) \in \mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T).$$

In case $\mathcal{A}_o$ is a linearly and f.t. stationarily predictable class we can, similarly, for any $\tau \geq 0$, put $\theta(\tau) F_1 = \theta(t,s) F_1$ for all F_1 such that for some pair (t,s) with $t \geq 0$, $s - t = \tau$, it holds that $F_1 \in \mathcal{M}_t$. The domain of $\theta(\tau)$, $\tau \geq 0$, can now be extended by linearity to $\mathcal{N}_+$ = linear span $\{\bigcup_{t \geq 0} \mathcal{M}_t\}$. The family $\{\theta(\tau)\}$, $\tau \geq 0$ is now a one-parameter semigroup of linear transformations on $\mathcal{N}_+$, which is also contained in $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$.

If $\mathcal{A}_o$ is f.t. stationarily (but not linearly) predictable and $\mathcal{H} \subset \mathcal{F}_d^o(\mathcal{U}, \mathcal{Y})$, a semigroup can be established in essentially the same way. Suppose $F' = \pi_t F = \pi_t G$. Then the fact that $\pi_{t+\tau} F = \pi_{t+\tau} G$, $\tau \geq 0$, implies that F and G restricted to $(I - P_t)\mathcal{U}$ are the same mapping. We now redefine π_t^{-1} as the set function: $\pi_t^{-1}(F') = \{F : \pi_t F = F'\}$. Then $\theta(t,s)$ can again be defined as $\pi_s \circ \pi_t^{-1}$, but only, of course, for $t \leq s$. $\theta(t,s)$ is again linear on $\mathcal{M}_t$, and the development that follows for the semigroup case can be repeated exactly. In what follows we restrict attention to the semigroups of linear transformations, as being of more immediate interest than groups in modelling for system identification.

The usual linear operator norm, when it exists, of the linear transformation $\theta(\tau)$ is given by

$$|\theta(\tau)| = \sup_{F' \in \mathcal{N}_+} \left\{ \frac{1}{\|F'\|_{\mathcal{Y}_T}} \|\theta(\tau)F'\|_{\mathcal{Y}_T} \right\}$$

where the symbol $|\cdot|$ has been used to provide a reminder that this is a different kind of norm than has been used for the other mappings that have appeared. From the definition of $\mathcal{N}_+$ it follows that $\theta(\tau)$ is a bounded operator if and only if there is a number $B > 0$ such that

$$\sup_{u \in \mathcal{U}} \|(P_T - P_o)\left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau}\right)(P_T - P_{-d})u\|_{\mathcal{Y}_T}$$

$$\leq B \sup_{u \in \mathcal{U}} \|(P_T - P_o)\left(\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n}\right)(P_T - P_{-d})u\|_{\mathcal{Y}_T} \qquad (14)$$

for any positive integer N, any set of points $t_1, \cdots, t_N$ all greater than or equal to zero, any set of scalars $\alpha_1, \cdots, \alpha_N$ and any $F_1, \cdots, F_N$ belonging to $\mathcal{M}$.

This is a regularity condition on the time behavior of the mappings F. Note that, unfortunately, it is not sufficient to consider just those $F \in \mathcal{H}$, but rather all finite linear combinations of these and of their translations. If $\mathcal{H}$ is itself a subset of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$ that is invariant under time shift, then all the $\mathcal{M}_t$ are the same and the sums in the condition (14) collapse to single terms.

Using the definitions established, we can now state a basic fact, which is really a corollary to Proposition 2.12.

Proposition 2.20 Let $\mathcal{S}_o$ be such that $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, let it be f.t. stationarily predictable with respect to T, and let the $\theta(\tau)$, $\tau \geq 0$, be bounded operators. Then $\{\theta(\tau)\}$, $\tau \geq 0$, is a strongly continuous semigroup of bounded linear operators on the Banach space $\overline{\mathcal{N}}_+$, the closure of $\mathcal{N}_+$ in $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$.

Proof: It is supposed of course that the $\theta(\tau)$ are extended by continuity to $\overline{\mathcal{N}}_+$. All that has to be shown is that $\|\theta(\tau)F' - F'\| \to 0$ as $\tau \to 0$, for any $F' \in \overline{\mathcal{N}}_+$. Since $\overline{\mathcal{N}}_+ \subset \mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$, and since any $F' \in \mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ can be written $F' = (P_T - P_o)F'(P_T - P_{-d})$, it follows that F' is the image under π_o of itself, regarded as an element of $\mathcal{F}(\mathcal{U}, \mathcal{Y})$. We write, $F' = \pi_o F = F_o$. Then

$$\|\theta(\tau) F' - F'\| = \|F_\tau - F_o\| \to 0 \quad \text{as} \quad \tau \to 0$$

by Proposition 2.12.†

Clearly the hypothesis that $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$ can be replaced by the hypothesis that $\mathcal{S}_o$ is linearly predictable, and then with the other hypotheses in force the conclusion still follows. For convenience we shall refer to an $\mathcal{S}_o$ that satisfies either the conditions of Proposition 2.20 or the modified conditions just given as a linear dynamical class of systems with respect to T. This terminology is introduced with some apology since dynamical is such a widely used term; however, it seems reasonably appropriate. There is no inference, of course, that the individual systems in the linear dynamical class are linear.

Thus far, T, the length of the interval of observation of the output, has remained fixed. We now look at how the properties of the special classes of systems introduced in this Section are affected by changes in T. When T is changed, so is the norm on the input space, which is always assumed to be a subset of $\mathcal{L}^p_{T+d}$. In fact $\|u\|^{(p)}_{T'} \le \|u\|^{(p)}_{T}$, when $T' \le T$. However, as has been pointed out earlier, the membership of $\mathcal{U}$ does not depend on the value of T, nor does the topology on $\mathcal{U}$, nor do the properties of T-compactness and shift invariance. Similar statements can be made for $\mathcal{Y}$ if $\mathcal{Y} = \mathcal{L}^p_T$. If $\mathcal{Y} = \mathcal{B}$, then not even the norm on $\mathcal{Y}$ is changed. In any event, the class of mappings $\mathcal{F}(\mathcal{U}, \mathcal{Y})$ is not affected.

Proposition 2.21 If $\mathcal{S}_o$ is a linearly predictable class of systems with respect to T', then it is also linearly predictable with respect to any $T \ge T'$.

Proof: Suppose the linear mapping $\pi_t(T)$ given by

$$\pi_t(T)F = (P_T - P_o) L_t F R_t (P_T - P_{-d})$$

is singular. Then for some $F \ne 0$, $\pi_t(T)F = 0$; and for $T' \le T$,

$$(P_{T'} - P_o) [(P_T - P_o) L_t F R_t (P_T - P_{-d})u] = 0$$

for all $u \in \mathcal{U}$. Since $(P_{T'} - P_{-d}) u \in \mathcal{U}$ for all $u \in \mathcal{U}$,

$$(P_{T'} - P_o) L_t F R_t (P_{T'} - P_{-d}) u = 0$$

for all $u \in \mathcal{U}$. Hence $\pi_t(T')$ is singular, and the assertion is proved by contradiction.†

Proposition 2.22 If $\mathcal{S}_o$ is a class of systems with the property that $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, and if $\mathcal{S}_o$ is stationarily predictable with respect to T', then it is stationarily predictable with respect to any $T \ge T'$. Stationary predictability can be replaced simultaneously in hypothesis and conclusion by future-time stationary predictability.

Proof: Suppose to start with that $T' \le T \le 2T$. We need to show that the condition $\pi_t(T)F = \pi_s(T)G$, where $F, G \in \mathcal{M} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$ implies $\pi_{t+a}(T)F = \pi_{s+a}(T)G$ for all a. We note that the condition $\pi_t(T)F = \pi_s(T)G$ can be written

$$(P_T - P_o)(L_t F R_t - L_s G R_s)(P_T - P_{-d})u = 0$$

for all $u \in \mathcal{U}$. Since $(P_{T'} - P_{-d})u \in \mathcal{U}$ for all $u \in \mathcal{U}$, it follows that

$$(P_{T'} - P_o)(L_t F R_t - L_s G R_s)(P_{T'} - P_{-d})u = 0$$

for all $u \in \mathcal{U}$; i.e., $\pi_t(T')F = \pi_s(T')G$. By hypothesis, it follows that $\pi_{t+a}(T')F = \pi_{s+a}(T')G$, or,

$$(P_{T'} - P_o)L_a(L_t F R_t - L_s G R_s)R_a(P_{T'} - P_{-d})u = 0 \qquad (15)$$

for all $u \in \mathcal{U}$, and any real number a.

Now

$$(P_{2T'} - P_{T'})L_a(L_t F R_t - L_s G R_s)R_a(P_{2T'} - P_{T'-d})u$$

$$= L_{-T'}(P_{T'} - P_o)L_{a+T'}(L_t F R_t - L_s G R_s)R_{a+T'}(P_{T'} - P_{-d})R_{-T'}u = 0$$

for all $u \in \mathcal{U}$, since $R_{-T'}(u) \in \mathcal{U}$ for all $u \in \mathcal{U}$, and we can replace the a of equation (15) by a + T'. Since the mappings F and G are causal with bounded memory (d),

$$(P_{2T'} - P_o)L_a(L_t F R_t - L_s G R_s)R_a(P_{2T'} - P_{-d})u$$

$$= (P_{2T'} - P_{T'})L_a(L_t F R_t - L_s G R_s)R_a(P_{2T'} - P_{T'-d})u$$

$$+ (P_{T'} - P_o)L_a(L_t F R_t - L_s G R_s)R_a(P_{T'} - P_{-d})u \quad ,$$

which equals zero by the calculations above. It follows then by a now familiar argument that

$$(P_T - P_o) L_a (L_t F R_t - L_s G R_s) R_a (P_T - P_{-d}) u = 0$$

for all $u \in \mathcal{U}$, which is what needs to be shown. The extension to arbitrary $T \geq T'$ follows by induction. The proof for future-time stationary predictability is the same with a restricted to be ≥ 0.†

Proposition 2.23 If $\mathcal{A}_o$ is a linear dynamical class of systems with respect to T', and if it further has the property that $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, then $\mathcal{A}_o$ is a dynamical class with respect to any $T \geq T'$.

Proof: In view of the preceding Proposition, all that needs to be proved is that if $\theta_{T'}(\tau)$ is a bounded operator for all $\tau \geq 0$, then $\theta_T(\tau)$ is a bounded operator for all $\tau \geq 0$ whenever $T \geq T'$. The meaning of the subscripts T and T' on $\theta(\tau)$ is obvious. In what follows it is necessary to go back and forth between norms in $\mathcal{Y}_T$ and in $\mathcal{Y}_{T'}$, so a subscript T or T' is used. The facts that, by an obvious identification of elements, $\mathcal{Y}_{T'}$ can be thought of as a subset of $\mathcal{Y}_T$, $T' \leq T$, and that then $\|y\|_{T'} = \|y\|_T$ when $y \in \mathcal{Y}_{T'}$ are used without comment.

Again assume to start with that $T \leq 2T'$. We have,

$$\|\theta_T(\tau) F\|_T = \sup_{\mathcal{U}} \|(P_T - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau}\right) \cdot (P_T - P_{-d}) u\|_T \ , \quad \tau \geq 0 \ , \ F \in \mathcal{N}_+ ,$$

for some $F_n \in \mathcal{M} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, and some scalars α_n. Now

$$\|\theta_T(\tau) F\|_T \leq \sup_{\mathcal{U}} \|(P_{T'} - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau}\right)$$

$$\cdot (P_T - P_{-d})\, u \,\|_T + \sup \left\| (P_T - P_{T'}) \left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau} \right) \right.$$

$$\left. \cdot\ (P_T - P_{-d})\, u \right\|_T$$

$$= \sup \|A(u)\|_T + \sup \|B(u)\|_T \tag{16}$$

where the A and B are defined implicitly.

Because the $F_n \in \mathcal{F}_d^o(\mathcal{U}, \mathcal{Y})$,

$$A = (P_{T'} - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau} \right) .$$

Since A(u) is different from zero only on $[0, T']$, and since $\theta_{T'}(\tau)$ is bounded,

$$\sup_{\mathcal{U}} \|A(u)\|_T = \sup_{\mathcal{U}} \|A(u)\|_{T'} = \|\theta_{T'}(\tau) F\|_{T'} \leq |\theta_{T'}(\tau)| \cdot \|F\|_{T'}$$

$$\leq |\theta_{T'}(\tau)| \cdot \sup_{\mathcal{U}} \left\| (P_{T'} - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n} \right) u \right\|_{T'}$$

$$\leq |\theta_{T'}(\tau)| \cdot \sup_{\mathcal{U}} \left\| (P_T - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n} \right) u \right\|_T$$

$$= |\theta_{T'}(\tau)| \cdot \|F\|_T \ . \tag{17}$$

Using the fact that $T - T' \leq T'$, and also using again the fact that the F_n are causal with bounded memory (d) yields

$$\|B(u)\|_T \leq \left\| (P_T - P_{T-T'}) \left(\sum_{n=1}^{N} \alpha_n L_{t_n+\tau} F_n R_{t_n+\tau} \right) \right.$$

$$\left. (P_T - P_{T-T'-d})\, u \right\|_T$$

$$= \|C(u)\|_T \ ,$$

where C is defined implicitly.

Now, $L_{T-T'} C R_{T-T'}$

$$= (P_{T'} - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{T-T'+t_n+\tau} F_n R_{T-T'+t_n+\tau} \right) (P_{T'} - P_{-d})$$

so, by the fact that $\theta_{T'}(T - T' + \tau)$ is a bounded operator,

$$\sup_{\mathcal{U}} \| L_{T-T'} C R_{T-T'}(u) \|_{T'} \le | \theta_{T'}(T - T' + \tau) |$$

$$\cdot \sup_{\mathcal{U}} \| (P_{T'} - P_o) \left(\sum_{n=1}^{N} \alpha_n L_{t_n} F_n R_{t_n} \right) (P_{T'} - P_{-d}) \|_{T'}$$

$$= | \theta_{T'}(T - T' + \tau) | \cdot \| F \|_{T'}$$

But, $\sup_{\mathcal{U}} \| L_{T-T'} C R_{T-T'}(u) \|_{T'} = \sup \| C(u) \|_T$

Thus,

$$\sup_{\mathcal{U}} \| B(u) \|_T \le \sup_{\mathcal{U}} \| C(u) \|_T$$

$$\le | \theta_{T'}(T - T' + \tau) | \cdot \| F \|_{T'}$$

$$\le | \theta_{T'}(T - T' + \tau) | \cdot \| F \|_T \quad . \qquad (18)$$

Combining the inequalities (16), (17) and (18) yields

$$\| \theta_T(\tau) F \|_T \le (| \theta_T'(\tau) | + | \theta_T'(T - T' + \tau) |) \cdot \| F \|_T$$

for all $F \in \mathcal{N}_+$, which establishes the result when $T \le \tau T'$. This can be extended to all $T \ge T'$ by induction.[†]

If now $\mathcal{S}_o$ is a class of systems with $\mathcal{H} \subset \mathcal{F}_d^o(\mathcal{U}, \mathcal{Y})$ and is dynamical with respect to some $T' > 0$, one can put T_o equal to the infimum of all such T' and know that $\mathcal{S}_o$ is dynamical with respect to any $T > T_o$. It is to be noted that the hypothesis that $\mathcal{H} \subset \mathcal{F}_d^o(\mathcal{U}, \mathcal{Y})$ cannot be dropped in this assertion. In fact, it is not very difficult to give an example where Proposition 2.22 is violated if the mappings F are not causal with bounded memory (d); thus the semigroup property is not preserved.

If $\mathcal{S}_o$ is a linear dynamical class with respect to T and with $\mathcal{H} \subset \mathcal{F}^o_d(\mathcal{U}, \mathcal{Y})$, then it is clearly possible to deal with the discrete parameter semigroup $\{\theta^n = \theta(nT)\}$, n = 0, 1, 2, $\cdots$, and still completely describe the future of the system by virtue of the interpolation formula (9). Under certain conditions when $\mathcal{S}_o = (\mathcal{Y}, g, \mathcal{H}, \mathcal{U})$ is a linear dynamical class, the discrete parameter semigroup $\{\theta^n\}$ can be used to induce a "corresponding" semigroup $\{\tilde{\theta}^n\}$ of linear operators on the linear space spanned by the system parameter space $\mathcal{X}_1$ of an ϵ-representation of $\mathcal{S}_o$. We describe a situation in which this can be done and construct the $\tilde{\theta}^n$. The construction is not unique, as will be seen, but any $\{\tilde{\theta}^n\}$ so devised approximates $\{\theta^n\}$ in the sense to be indicated.

Let it be assumed that $\mathcal{Y}$ is $\mathcal{L}^2_T$. Write $\mathcal{S}_n = (\mathcal{Y}_T, g, \mathcal{H}_{nT}, \mathcal{U}_T)$, n = 0, 1, 2, ..., for the classes of truncated systems, where $\mathcal{H}_{nT}$ is the set of all $\pi_{nT} F$, $F \in \mathcal{H}$. By the assumption on $\mathcal{Y}$, $\mathcal{Y}_T = L_2$. Since $\mathcal{S}_o$ is a linear dynamical class, $\mathcal{H}_{nT} = \theta(nT)\, \mathcal{H}_o = \theta^n \mathcal{H}_o$. Let it further be required that $\bigcup_{n=0}^{\infty} \mathcal{H}_{nT}$ is a compact subset of $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$, and for convenience denote $\bigcup_{n=0}^{\infty} \mathcal{H}_{nT}$ by $\mathcal{G}_T$. Then each $\mathcal{S}_n$ is a subclass of $\tilde{\mathcal{S}} = (\mathcal{Y}_T, g, \mathcal{G}_T, \mathcal{U}_T)$. Since $\mathcal{U}_T$ and $\mathcal{G}_T$ are compact and $\mathcal{Y}_T = L_2$, $\tilde{\mathcal{S}}$ has a standard ϵ-representation $(\mathcal{S}_1, \phi_1)$, $\mathcal{S}_1 = (\mathcal{Y}_T, f_1, \mathcal{X}_1, \mathcal{U}_T)$ as given by Proposition 1.7 of Part 1, and ϕ_1 is linear. $\mathcal{X}_1 = \phi_1 \mathcal{G}_T$ is a subset of a finite-dimensional Euclidean space; let R^K be the Euclidean space generated by $\mathfrak{X}_1$. The representation mapping ϕ_1 as given by Proposition 1.7 is actually defined as a continuous linear map from the closed linear span of $\mathcal{G}_T$ onto R^K. Obviously the closed linear span of $\mathcal{G}_T$ is contained in $\overline{\mathcal{N}}_+$, the domain of the $\theta(\tau)$. Let $\{b_1, \cdots, b_K\}$ be elements of $\mathfrak{X}_1$ which form a basis for R^K, and denote the coordinate functionals $\{b_1^*, \cdots, b_K^*\}$, so that any element $x \in R^K$ can be written

$$x = \sum_{i=1}^{K} b_i^*(x)\, b_i \quad .$$

The idea of the construction of $\tilde{\theta}$ is that $\tilde{\theta}$ should be the composition of the mappings ψ, θ, ϕ in that order. However, this will not quite do, because $\psi(x)$, $x \in \mathfrak{X}_1$, is not necessarily contained in $\mathcal{G}_T$, and hence is not necessarily in $\mathcal{N}_+$, the domain of $\theta(\tau)$. To correct this, we construct a linear mapping $\tilde{\psi}$ which does satisfy the condition $\psi(x) \in \mathcal{G}_T$, $x \in \mathfrak{X}_1$, and which is close to ψ. Consider the continuous linear functionals on the closed linear span of $\mathcal{G}_T$ given by $b_i^* \circ \phi$, $i = 1, \cdots, K$. Let $\mathcal{E}_i$ be the null space of $b_i^* \circ \phi$. First choose an element H_1 belonging to $\mathcal{G}_T$ that does not belong to $\mathcal{E}_1$; this is possible by the definitions of $\mathfrak{X}_1$ and b_1. Then $b_1^* \circ \phi(H_1) = \alpha_1 \neq 0$. Next, choose $H_2 \in \mathcal{G}_T$, not in $\mathcal{E}_2$, and linearly independent of H_1. This can be done by virtue of the linear independence of the b_i, and yields $b_2^* \circ \phi(H_2) = \alpha_2 \neq 0$. Continue this procedure to obtain a linearly independent set $\{H_1, \cdots, H_K\}$, $H_i \in \mathcal{G}_T$, satisfying $b_i^* \circ \phi(H_i) = \alpha_i \neq 0$. Define another basis for R^K with elements in $\mathfrak{X}_1$ by

$$c_j = \sum_{i=1}^{K} [b_i^* \circ \phi(H_j)]\, b_i \quad ;$$

since each $\phi(H_j) \in \mathfrak{X}_1$, it is clear that the c_j do belong to $\mathfrak{X}_1$. Define $\tilde{\phi}$, a linear mapping from the linear span of $\{H_1, \cdots, H_K\}$ onto R^K by $\tilde{\phi}(H_i) = c_i$, $i = 1, \cdots, K$, and extending linearly. $\tilde{\phi}$ is 1:1, so we can define $\tilde{\psi} = \tilde{\phi}^{-1}$, a linear mapping from R^K onto the linear span of $\{H_1, \cdots, H_K\}$. If $\tilde{H}$ belongs to the linear span of $\{H_1, \cdots, H_K\}$ and also belongs to $\mathcal{G}_T$, then we have

$$\phi(\tilde{H}) = \phi\left(\sum_{i=1}^{K} \gamma_i H_i\right) = \sum_{i=1}^{K} \gamma_i c_i \in \mathfrak{X}_1 ,$$

thus $\tilde{\psi}$ carries any element in $\mathfrak{X}_1$ into $\mathcal{G}_T$. As was already mentioned, $\tilde{\psi}$ is not uniquely defined, except in certain cases of finite-dimensional $\mathcal{G}_T$, since the choice of $H_1, \cdots, H_K$ is not unique and the resulting linear space spanned by them is not unique.

It now follows that if $H \in \mathcal{G}_T$, then $\|H - \tilde{\psi} \circ \phi_1(H)\| \leq 2\epsilon$. In fact, since $(\mathcal{A}_1, \phi_1)$ is an ϵ-representation of $\tilde{\mathcal{A}}$, $\|H - \psi_1 \circ \phi_1(H)\| \leq \epsilon$. But $\tilde{\psi} \circ \phi_1(H)$ is an element of $\mathcal{G}_T$ and it has the same representing element as H, i.e., $\phi_1 \circ \tilde{\psi} \circ \phi_1(H) = \phi_1(H)$. Hence, $\|\tilde{\psi} \circ \phi_1(H) - \psi_1 \circ \phi_1(H)\|$ $= \|[\tilde{\psi} \circ \phi_1(H)] - \psi \circ \phi_1 \circ [\tilde{\psi} \circ \phi_1(H)]\| \leq \epsilon$, from which the assertion follows.

<u>Proposition 2.24</u> The mapping $\tilde{\theta}$ from R^K into R^K given by $\tilde{\theta} = \phi_1 \circ \theta \circ \tilde{\psi}$ is well-defined and linear. If $H_o \in \mathcal{G}_T$, then

$$\|\theta^n H_o - \tilde{\psi} \circ \tilde{\theta}^n \circ \phi_1(H_o)\|$$

$$\leq 2\epsilon \; [1 + |\theta| + \cdots + |\theta|^n] \qquad (19)$$

where $|\theta|$ denotes the norm of $\theta = \theta(T)$.

Proof: It has already been ascertained that the range of $\tilde{\psi}$ is contained in the linear span of $\mathcal{G}_T$, which in turn is contained in $\mathcal{N}_+$. So $\theta \circ \tilde{\psi}$ is defined. By definition, $\mathcal{G}_T$ is invariant with respect to θ; since θ is linear, the linear span of $\mathcal{G}_T$ is carried into itself by θ. Thus the range of $\theta \circ \tilde{\psi}$ is contained in the domain of ϕ_1, and $\phi_1 \circ \theta \circ \tilde{\psi}$ is defined as a linear transformation from R^K into itself.

If $H_o \in \mathcal{G}_T$, $\tilde{H}_o = \tilde{\psi} \circ \phi_1 (H_o) \in \mathcal{G}_T$, and $\| H_o - \tilde{H}_o \| \le 2\epsilon$, as already shown. Then $\theta \tilde{H}_o \in \mathcal{G}_T$,

$$\| \theta \tilde{H}_o - \theta H_o \| \le |\theta| \; \| H_o - \tilde{H}_o \| \le 2\epsilon \; |\theta| \quad ,$$

and

$$\| \theta H_o - \tilde{\psi} \circ \phi_1 \circ \theta \circ \tilde{\psi} \circ \phi_1(H) \| \le 2\epsilon + 2\epsilon \; |\theta| \quad .$$

The inequality (19) follows by induction.†

Only linear predictability and associated ideas have been considered in this Section. However, it probably should be noted, although the fact is obvious, that a class of systems could be described as predictable in a wider sense. Indeed, if $\{T_n\}$, $n = 1, 2, \cdots$, is any sequence of mappings from $\mathcal{F}^o_d(\mathcal{U}_T, \mathcal{Y}_T)$ into $\mathcal{F}^o_d(\mathcal{U}_T, \mathcal{Y}_T)$ so that the images under these mappings satisfy the conditions of Proposition 2.15, then the class is "predictable" in an obvious sense.

V. REMARKS

It will be noticed that, for what has been labeled a linear dynamical class of systems, a structure has been described that is analogous to the usual state-variable formulation of a linear system. In fact, we can write either

$$\begin{cases} F_t = \theta(t)\, F_o \\ y_t = F_t\, u_t \end{cases}$$

or

$$\begin{cases} F_{nt} = \theta(T)\, F_{(n-1)T} \\ y_{nT} = F_{nT}\, u_{nT} \end{cases}$$

where $u_t = (P_{t+T} - P_{t-d})\, u$, $y_t = (P_{t+T} - P_t)\, y$. The first equation in either case corresponds to the state equation for a linear, time-invariant unforced system, and the second to a time-varying observation equation--actually a linear observation equation, since $F_t\,(u_t)$ for fixed u_t defines a linear mapping from $\mathcal{F}(\mathcal{U}_T, \mathcal{Y}_T)$ into $\mathcal{Y}_T$. It follows that the identification problem, when there is noise added, is thereby analogous to the problem of estimating state in a linear system when there is additive noise. A study of identification of $F \in \mathcal{F}$ along the lines of this analogy will be made in a future report. A practical difficulty is, of course, that in modelling many real problems involving rapid time variation the transformations $\theta(\tau)$ cannot be known; but this is simply to say that a rapidly time-varying system is not identifiable if there is no information about the future time variation.

The characterization of system trajectories in terms of strongly continuous semigroups of linear operators obviously suggests the application of some of the elaborate theory of such semigroups to further study of the structure of these classes of systems, but this is a matter for future work.

APPENDIX

PROJECTIONS ON PAST AND FUTURE

The projections P_t used in this paper are defined by

$$
\begin{aligned}
[P_t f](s) &= f(s) \quad , \quad s \le t \\
&= 0 \quad , \quad s > t
\end{aligned}
\tag{A1}
$$

where f is a function on R^1. This definition is still meaningful if f is an element of a space for which the elements are equivalence classes of functions equal a.e. Lebesgue, for then it is applied to each representative of the equivalence class. Most of the operations involving these projections are intuitively clear from the definition. Here and there, however, one may want a formal proof of an identity involving these projections. If one if going to the trouble to provide such proofs, it seems as if the properties that are used might as well be axiomatized, particularly since this does not involve much effort. Then generalizations are at least possible. There is nothing new in thus generalizing the notions of past and future, of course; see, e.g., [3] , [4] , and [5] . However it is not the intent in this paper really to pursue any notion of generalized time; so we do not build on theory established in the references cited, but merely develop some simple results ad hoc. These results are more than sufficient for what is needed here.

For the remainder of this Appendix, the operators P_t are not to be taken as defined in Section I unless such an interpretation is specifically indicated, but are to be considered abstractly as operators belonging to a family according to the following definition.

Definition A1. Let $\mathcal{Z}$ be a linear space. Let $\{P_t\}$, $-\infty \le t \le \infty$, be a parametrized family of operators on $\mathcal{Z}$ (that is, mappings from $\mathcal{Z}$ into $\mathcal{Z}$) such that the following conditions are satisfied:

1) $P_{-\infty} = 0$ (the zero operator); $P_{+\infty} = I$ (the identity operator)

2) $P_t P_s = P_s P_t$ for all t, s

3) If $t \le s$, $P_t P_s = P_t$

4) P_t is linear on $\mathcal{Z}$

5) If $(P_b - P_a) y = (P_b - P_a) z$ for arbitrarily large positive numbers b and arbitrarily large negative numbers a where y and z are elements of $\mathcal{Z}$, then $y = z$.

Then $\{P_t\}$ will be called a family of generalized-time projections on $\mathcal{Z}$ (g.t. projections).

Proposition A1 Let $\mathcal{Z}$ be any $\mathcal{L}_T^p$ space, or any $L^p(R^1)$ space, or $\mathcal{B}$, or any closed linear subspace of one of these. Let $\{P_t\}$ be the family of projection operators defined by equation (1), or the extension of (1) to equivalence classes of functions. Then $\{P_t\}$ is a family of g.t. projections on the space in question.

Proof: Obvious verifications.†

The projection property (P) as defined in Section I is still a meaningful concept when applied to g.t. projections on a subset of $\mathcal{Z}$. Let $\mathcal{Z}_1$ and $\mathcal{Z}_2$ be linear spaces with families of g.t. projections $\{P_t\}$ and $\{Q_t\}$, respectively. Let $\mathcal{U}$ be a subset of $\mathcal{J}_1$ with property (P), and let F be a mapping from $\mathcal{U}$ into $\mathcal{J}_2$. As in the special case, F is said to be <u>causal</u> if $Q_t F(u) = Q_t F P_t (u)$ for all t and all $u \in \mathcal{U}$; F has <u>bounded memory (d)</u> if $(Q_\infty - Q_t) F(u) = (Q_\infty - Q_t) F (P_\infty - P_{t-d}) (u)$ for all t and all $u \in \mathcal{U}$.

<u>Proposition A2 (Proposition 2.7)</u> If F is a mapping from $\mathcal{U}$ into $\mathcal{J}_2$ that is causal and has bounded memory (d), then for every $T > 0$,

$$(Q_{t+T} - Q_t) F(u) = (Q_{t+T} - Q_t) F (P_{t+T} - P_{t-d}) (u) \tag{A4}$$

for all t and all $u \in \mathcal{U}$.

Conversely, if equation (A4) is satisfied for some $T > 0$ and all t and all $u \in \mathcal{U}$, then F is causal and has bounded memory (d).

Proof: We prove first that causality and bounded memory (d) imply the property (A4). For any $u \in \mathcal{U}$, any real number t and any $T > 0$,

$$(Q_{t+T} - Q_t) F u = (Q_{t+T} - Q_t) Q_{t+T} F u$$

$$= (Q_{t+T} - Q_t) Q_{t+T} F P_{t+T} u$$

$$= (Q_{t+T} - Q_t) F P_{t+T} u$$

$$= Q_{t+T} (Q_\infty - Q_t) F (P_{t+T} u)$$

$$= Q_{t+T} (Q_\infty - Q_t) F (P_\infty - P_{t-d}) (P_{t+T} u)$$

$$= (Q_{T+t} - Q_t) F (P_{t+T} - P_{t-d}) u \quad .$$

Only conditions 1), 2), 3) and 4) of definition (A1) and the properties of causality and bounded memory have been used.

Now suppose that (A4) is satisfied. We prove causality. Let $b > t$ be positive and $a < t$ be negative. Then,

$$(Q_b - Q_a) Q_t F u = (Q_b - Q_a) \sum_{k=0}^{K} (Q_{t-kT} - Q_{t-(k+1)T}) F u$$

for any K such that $t = (K+1) T < a$. By (A4) this is equal to

$$(Q_b - Q_a) \left[\sum_{k=0}^{K} (Q_{t-kT} - Q_{t-(k+1)T}) F (P_{t-kT} - P_{t-(k+1)T-d}) \right] u$$

$$= (Q_b - Q_a) \left[\sum_{k=0}^{K} (Q_{t-kT} - Q_{t-(k+1)T}) F (P_{t-kT} - P_{t-(k+1)T-d}) P_t \right] u$$

$$= (Q_b - Q_a) \left[\sum_{k=0}^{K} (Q_{t-kT} - Q_{t-(k+1)T}) F \right] P_t u$$

$$= (Q_b - Q_a) Q_t F P_t u$$

Hence, by condition 5) of the definition, $Q_t F u = Q_t F P_t$.

The proof that F has bounded memory is completely analogous.†

Let $\{z_k\}$ be an arbitrary sequence of elements belonging to $\mathfrak{Z}$, and let $\{\Delta_k = P_{t_k} - P_{t_{k-1}}\}$ be a sequence of differences of g.t. projections where the $\{t_k\}$, $\cdots$, -2, -1, 0, 1, 2, $\cdots$, satisfy $t_k < t_{k+1}$ and $\lim_{k \to -\infty} t_k = \infty$, $\lim_{k \to -\infty} t_k = -\infty$. In Section III infinite sums of the form

$$\sum_{k=1}^{\infty} \Delta_k z_k$$

are used. These sums have no meaning as far as the structure given by definition A1 is concerned, and some further condition is necessary. It is sufficient to require:

6) Corresponding to every $\{z_k\}$, $z_k \in \mathfrak{Z}$, and $\{\Delta_k\}$, $k = 1, 2, \cdots$, where the Δ_k are as defined above, there exists a $z \in \mathfrak{Z}$ with the property

$$(P_b - P_a) z = (P_b - P_a) \sum_{k=-K_1(a)}^{K_2(b)} \Delta_k z_k$$

for all $b \geq a$, where K_1 and K_2 are any integers large enough that the interval $(a, b]$ is contained in the interval $(-t_{K_1}, t_{K_2}]$.

If condition 6) holds, the, for example,

$$z = \sum_{-\infty}^{\infty} \Delta_k z \quad , \qquad \text{and} \qquad P_t z = \sum_{k=0}^{\infty} (P_{t-kT} - P_{t-(k+1)T}) z \; .$$

Also, expressions of the kind

$$\sum_{k=-\infty}^{\infty} (Q_{t-kT} - Q_{t-(k+1)T}) F_k (P_{t-kT} - P_{t-(k+1)T-d}) u$$

are defined, where each F_k is a mapping from $\mathcal{U}$ into $\mathcal{Z}_2$, as above.

It is clear that if $\mathcal{Z}$ is any $\mathcal{L}_T^p$ space, or $\mathcal{B}$ (but not, of course, L^p), and the $\{P_t\}$ are ordinary time projections then condition 6) holds.

REFERENCES

1. Root, W.L. Approximate representations of causal systems with bounded memory. Appears in Techniques of Optimization. A.V. Balakrishnan, ed., Academic Press, pp. 51-64 (1972). (Proceedings of Fourth IFIP Colloquium on Optimization Techniques, Los Angeles, 1971.)

2. Root, W.L. On the modelling of Systems for Identification. Part 1. ε-Representations of classes of systems. Technical Report, Univ. of Michigan, 1972.

3. Porter, W.A. Some circuit theory concepts revisited. Internat. J. of Control 12 (1970), 11. 433-488.

4. DeSantis, R.M. and Porter, W.A. On time-related properties of nonlinear systems. SIAM J. Appl. Math., V. 24, No. 2 (1972), pp. 188-206.

5. Saeks, R. Resolution space. Operators and systems. Lecture Notes in Economics and Mathematical Systems. No. 82. Springer-Verlag.