UNIVERSITY OF LEEDS

This is a repository copy of *Modeling of linear fading memory systems*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/699/

Article:
Partington, J.R. and Makila, P.M. (1996) Modeling of linear fading memory systems. IEEE Transactions on Automatic Control, 41 (6). pp. 899-903. ISSN 0018 - 9286

https://doi.org/10.1109/9.506247

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

that

$$\operatorname{conv} p(j\omega, Q) = \operatorname{conv}\{p[j\omega, q^*(\omega, \theta)] : \theta \in [0, 2\pi]\}.$$

Proof: In the proof to follow, it is convenient to use the shorthand notation $\mathbf{P}_\omega$ to denote the right-hand side above. First, note that $\operatorname{conv} p(j\omega, Q)$ is easily seen to be a subset of $\mathbf{P}_\omega$ because $q^*(\omega, \theta) \in Q$ for all $\theta \in [0, 2\pi]$. Therefore, to complete the proof, we fix some $z_0 \in \operatorname{conv} p(j\omega, Q)$ and must show that $z_0 \in \mathbf{P}_\omega$.

Proceeding by contradiction, if $z_0 \notin \mathbf{P}_\omega$, the Separating Hyperplane Theorem (for example, see [4]) guarantees that the point z_0 can be strictly separated from the closed convex set $\mathbf{P}_\omega$. Hence, there exists some nonzero complex number $\eta \in \mathbf{C}$ such that

$$\langle \eta, z_0 \rangle > \langle \eta, \mathbf{p} \rangle$$

for all $\mathbf{p} \in \mathbf{P}_\omega$. Now taking

$$\theta_0 \doteq \arg \eta$$

we divide both sides above by $|\eta|$ and obtain

$$\langle e^{j\theta_0}, z_0 \rangle > \langle e^{j\theta_0}, \mathbf{p} \rangle$$

for all $\mathbf{p} \in \mathbf{P}_\omega$. Equivalently

$$\langle e^{j\theta_0}, z_0 \rangle > \max_{\mathbf{p} \in \mathbf{P}_\omega} \langle e^{j\theta_0}, \mathbf{p} \rangle. \tag{1}$$

This is the inequality to be contradicted.

Indeed, using the fact that a linear function on $p(j\omega, Q)$ and $\operatorname{conv} p(j\omega, Q)$ has the same maximum value, we generate the chain of inequalities

$$\begin{aligned} \langle e^{j\theta_0}, z_0 \rangle &\le \max_{z \in \operatorname{conv} p(j\omega, Q)} \langle e^{j\theta_0}, z \rangle \\ &= \max_{z \in p(j\omega, Q)} \langle e^{j\theta_0}, z \rangle \\ &= \max_{q \in Q} \langle e^{j\theta_0}, p(j\omega, q) \rangle \\ &= \max_{q \in Q} f_{e^{j\theta_0}}(\omega, q) \\ &= f_{e^{j\theta_0}}[\omega, q^*(\omega, \theta_0)] \\ &= \langle e^{j\theta_0}, p[j\omega, q^*(\omega, \theta_0)] \rangle \\ &\le \max_{\mathbf{p} \in \mathbf{P}_\omega} \langle e^{j\theta_0}, \mathbf{p} \rangle. \end{aligned}$$

The proof is now complete because this inequality contradicts (1).

IV. Numerical Example

In this section, we illustrate the application of the Generalized Mapping Theorem. To this end, we consider the nonlinear uncertainty structure in the example of Section II-F and assume uncertainty bounds $q_i \in [0, 1]$ for $i = 1, 2, 3, 4$ and specific nonlinearities

$$\begin{aligned} \varphi_1(q_3) &= q_3^3 + 2q_3^2 \\ \varphi_2(q_3) &= \cos 2q_3 \\ \varphi_3(q_4) &= -(q_4 - 0.5)^2 \\ \varphi_4(q_4) &= \cos q_4. \end{aligned}$$

In Fig. 1, the convex hull of the value set is shown for frequency $\omega = 1$. For validation purposes, this figure also includes a plot of 10 000 sample points which were obtained via random Matlab evaluations of $p(j\omega, q)$; a uniform distribution over [0, 1] was used for each component q_i of q. It is interesting to note that the outward curvature of the boundary of the convex hull is consistent with the fact that the Mapping Theorem cannot be used to obtain the desired convex hull.

It is also important to recall that in a robust stability context, the convex hull can be readily exploited. That is, if $p(s, q)$ has invariant degree and $p(s, q^0)$ is stable for some $q^0 \in Q$, then satisfaction of the zero exclusion condition $0 \notin \operatorname{conv} p(j\omega, Q)$ for all $\omega \ge 0$ guarantees robust stability. For the example at hand, monicity of $p(s, q)$ guarantees invariant degree, and it is easily verified that with nominal uncertainty $q = q^0 = 0$, the polynomial $p(s, q^0)$ is stable. After carrying out a preliminary frequency sweep while plotting $\operatorname{conv} p(j\omega, Q)$, it was determined that for robust stability purposes, the critical range is $1 \le \omega \le 5$. In Fig. 2, the plot is shown with a frequency separation $\Delta\omega = 0.25$. Since zero is excluded from the convex hull at all frequencies, this family of polynomials is deemed to be robustly stable.

V. Conclusion

In this paper, the notion of mappability was introduced. This demonstrated that one can handle much larger classes of uncertainty structures than those addressed by the Mapping Theorem. This work suggests that an approach based on convexification may be fruitful for even more complicated robustness problems.

References

[1] R. A. Frazer and W. J. Duncan, "On the criteria for stability for small motions," in *Proc. Royal Society A*, vol. 124, pp. 642–654, 1929.
[2] B. R. Barmish, *New Tools for Robustness of Linear Systems.* New York: Macmillan, 1994.
[3] L. A. Zadeh and C. A. Desoer, *Linear System Theory.* New York: McGraw-Hill, 1963.
[4] R. T. Rockafellar, *Convex Analysis.* Princeton: Princeton Univ. Press, 1970.

Modeling of Linear Fading Memory Systems

J. R. Partington and P. M. Mäkilä

***Abstract*—Motivated by questions of approximate modeling and identification, we consider various classes of linear time-varying bounded-input–bounded output (BIBO) stable fading memory systems and prove some characterizations of them. These include fading memory systems, in general, almost periodic systems, and asymptotically periodic systems. We also show that norm and strong convergence coincide for BIBO stable causal fading memory systems.**

I. Introduction

Recently, an intense research effort has taken place in the emerging field of identification for robust control [21], [9], [20], [6], [17], [24], [7], [15], [14]. Both stochastic and nonstochastic approaches to

Manuscript received May 25, 1995; revised October 19, 1995 and November 3, 1995.
J. R. Partington is with the School of Mathematics, University of Leeds, Leeds LS2 9JT, U.K.
P. M. Mäkilä is with the Department of Computer Science and Electrical Engineering, Luleå University of Technology, S-971 87 Luleå, Sweden.
Publisher Item Identifier S 0018-9286(96)02821-8.

system identification have been studied in this context, often under the assumption that the true plant is a linear time invariant (LTI) system. However, real plants are often at least slightly time-varying due to wear of equipment, component aging, and other reasons.

In the present paper we shall consider modeling of causal bounded-input bounded-output (BIBO) stable linear time-varying (LTV) systems and derive results for several classes of such LTV systems. Classes of BIBO-stable LTV systems that will be studied in this paper include: fading memory (FM) LTV systems [2], [19] and various subsets of FM systems, such as almost periodic LTV systems and asymptotically periodic systems. The aim is to derive results that are important in understanding certain fundamental issues of modeling and identification of such systems rather than to derive concrete identification algorithms. One of the main questions considered here is how to obtain efficient model parameterizations for various classes of LTV systems. For a separable space such as the space of BIBO stable causal LTI systems, it is possible to use different bases to obtain useful model parameterizations. However, for LTV systems the notion of basis turns out to be less inviting. We also discuss the relevance of different convergence notions for systems (operators) and derive some results in this connection.

The methods we use are mostly those of approximation theory. Although many papers have been written on identification of certain types of LTV systems for adaptive control, approximation concepts have not received enough attention. However, Zames *et al.* [23], [10] have obtained several interesting results on identification of slowly-varying systems using such concepts.

In Section II we give some mathematical preliminaries. Fading memory systems and strongly fading memory systems are considered in Section III. Asymptotically periodic and almost periodic systems are considered in Section IV.

II. Mathematical Preliminaries

Let ℓ_∞^+ denote the space of bounded real sequences $(u(n))_{n=1}^\infty$ with norm $\|u\| = \sup_{n\geq 1} |u(n)|$, and let c_0 denote the closed linear subspace consisting of all sequences $(u(n))$ which tend to zero. Further, let c_{00} be the subspace consisting of all eventually zero sequences, so that $u \in c_{00}$ if and only if there exists $N \geq 1$ such that $u(n) = 0$ for all $n \geq N$. Equivalently $u \in c_{00}$ if and only if u is a finite linear combination of the vectors $e_1 = (1, 0, 0, \cdots)$, $e_2 = (0, 1, 0, 0, \cdots)$, $\cdots$.

We shall also require spaces indexed by the whole set of integers, so we write ℓ_∞^- for the space of bounded real sequences $(u(n))_{n\leq 0}$ indexed by the nonpositive integers. We thus have a decomposition of $\ell_\infty = \ell_\infty(Z)$ into $\ell_\infty^- \oplus \ell_\infty^+$. We write $c_{00}(Z)$ for the space of finitely supported sequences, i.e., the linear span of $(e_k)_{k\in Z}$.

Given any normed space X, the space $\ell_\infty(X)$ will be defined to be the space of all (one-sided) bounded sequences $(x_t)_{t\geq 1}$ with norm $\|(x_t)\| = \sup_{t\geq 1} \|x_t\|$. The subspace $c_0(X)$ consists of all sequences $(x_t)_{t\geq 1}$ such that $\|x_t\| \to 0$ as $t \to \infty$.

In many approximation and identification situations it is convenient to be able to parameterize the space of candidate models in a linear way. For it to be possible to approximate all possible models by models depending on only finitely many parameters, it is necessary that our model space X be separable, that is, that it possesses a countable dense subset. Often this is assumed implicitly, and *a priori* information is assumed which makes the system lie in a relatively compact (and thus separable) set of models. Equivalently, X is separable if it is the closure of an increasing union of finite-dimensional subspaces

$$X = \overline{\bigcup_{n\geq 1} X_n} \tag{1}$$

where each X_n is finite-dimensional (without loss of generality, $\dim X_n = n$ for each n) and $X_1 \subset X_2 \subset \cdots$. We call the sequence $(X_n)_{n\geq 1}$ a model set (cf. [13]). Models in X can be approximated arbitrarily closely by a suitable choice of X_n, and each model in X_n depends on only finitely many parameters. If (x_n) is a sequence in X with dense linear span, then defining X_n to be the linear span of $\{x_1, \cdots, x_n\}$ for each n gives rise to a model set.

We shall be interested in various notions of convergence of operators on normed spaces. Let X be a normed space and X^* its dual space. Recall that the norm of an operator G on X is given by

$$\|G\| = \sup_{x\in X, x\neq 0} \frac{\|Gx\|}{\|x\|}. \tag{2}$$

Let G and $(G_n)_{n\geq 1}$ be bounded operators on X. Then one says that the sequence (G_n) converges to G uniformly (or in operator norm) if

$$\|G_n - G\| \to 0 \quad \text{as } n \to \infty. \tag{3}$$

One says also that (G_n) converges to G strongly (or in the strong operator topology) if

$$\|G_n x - Gx\| \to 0 \quad \text{as } n \to \infty \tag{4}$$

for each $x \in X$. Similarly one says that (G_n) converges to G weakly (or in the weak operator topology) if

$$\phi(G_n x - Gx) \to 0 \quad \text{as } n \to \infty \tag{5}$$

for each $x \in X$ and $\phi \in X^*$. It is easy to see that uniform convergence implies strong convergence which, in turn, implies weak convergence; in general these notions are not equivalent.

As in [5] we shall consider the space $\mathcal{B}_{\mathrm{TV}}$ of strictly causal linear systems $G: \ell_\infty \to \ell_\infty^+$ written as Volterra sum operators

$$(Gu)(t) = \sum_{k\geq 1} g(t,k)u(t-k) \tag{6}$$

where for each $t \geq 1$ the kernel $g_t = (g(t,k))_{k\geq 1}$ is a real sequence. It is well known [5] that the system G is BIBO stable (that is it defines a bounded operator on ℓ_∞) if and only if the following quantity is finite:

$$\|G\| = \sup_{t\geq 1} \|g_t\|_1 = \sup_{t\geq 1} \sum_{k\geq 1} |g(t,k)|. \tag{7}$$

An important subspace of $\mathcal{B}_{\mathrm{TV}}$ is the space $\mathcal{B}_{\mathrm{TI}}$ of linear strictly causal BIBO-stable time-invariant systems. These are characterized by the property that the sequences $(g(t,k))_{k\geq 1}$ are independent of t, say, $(g(t,k))_{k\geq 1} = (h(k))_{k\geq 1}$, for each t, where $h = (h(k)) \in \ell_1$ is the unit impulse response of G.

The space $\mathcal{B}_{\mathrm{TV}}$ is too large for our purposes, since it is well known to be nonseparable. We are therefore driven to consider subspaces of $\mathcal{B}_{\mathrm{TV}}$ of which an important class is analyzed in the next section.

III. Fading Memory LTV Systems

Although fading memory operators were earlier introduced by Boyd and Chua [2], we shall follow Shamma and Zhao [18], [19] in giving a simpler definition of (uniform) fading memory which applies to LTV causal systems (i.e., the class $\mathcal{B}_{\mathrm{TV}}$).

Let X be one of the sequence spaces c_0 or ℓ_p $(1 \leq p \leq \infty)$. For $n \geq 1$, let $P_n: X \to X$ be the projection operator taking $(x_k) \in X$ to $(x_k)_{k\leq n}$.

An operator $G: X \to X$ is said to have finite memory if there is an increasing function $\phi : N \to N$ such that $(I - P_{\phi(n)})GP_n x = 0$ for all $x \in X$. (In particular $Ge_k \in \ell_\infty^- \oplus c_{00}$ for each k.) An operator is said to have fading memory if it is the norm limit of a sequence of finite memory operators.

Shamma and Zhao [18], [19] show that any bounded linear operator on X has fading memory when X is c_0 or ℓ_p, $(1 \le p < \infty)$. However, the example $G\colon \ell_\infty \to \ell_\infty$ with $(Gu)(1) = 0$ and $(Gu)(n) = u(1)$ for all $n \ge 2$ and for all $u \in \ell_\infty$ clearly does not have fading memory [19].

Shamma and Zhao [19] also introduce the notion of pointwise finite memory and fading memory, but for linear discrete-time systems defined on $X = c_0$ or ℓ_p $(1 \le p \le \infty)$ these are clearly the same as the (uniform) finite memory and (uniform) fading memory defined above.

We shall write $\mathcal{B}_{\rm FM}$ for the closed subspace of $\mathcal{B}_{\rm TV}$ consisting of fading memory operators. One is interested in properties of causal linear operators on ℓ_∞ and the following preliminary result shows that the subspace c_0 plays a key role in their study.

Lemma III.1 Let G be a causal BIBO-stable linear operator defined as in (6). Then the operator norm of G (i.e., on $\ell_\infty(Z)$) equals the operator norm of G restricted to $\ell_\infty^- \oplus c_0$ which in turn equals the norm of G restricted to $c_{00}(Z)$.

Proof: Clearly the norm of $G\colon \ell_\infty \to \ell_\infty^+$ is at least as big as the norm of its restrictions to $\ell_\infty^- \oplus c_0$ and $c_{00}(Z)$. Suppose that $u \in \ell_\infty$ is such that $\|u\|_\infty = 1$ and $\|Gu\|_\infty > K$. Then there exists an index t such that $|(Gu)(t)| > K$. That is

$$\left\| \sum_{k \ge 1} g(t,k)u(t-k) \right\| > K.$$

Clearly this last sum can be replaced by a finite sum (say, $\sum_{k=1}^{N}$) while remaining larger than K. In that case the vector $v \in c_{00}(Z)$ defined by

$$v(r) = \begin{cases} u(r), & \text{if } k - N \le r \le k-1; \\ 0, & \text{otherwise} \end{cases}$$

satisfies $\|v\| \le 1$ and $\|Gv\| > K$ which implies the result. □

We shall also require the following general result which is of interest in its own right because it means that for the purposes of approximate modeling of systems with convolution operators convergence for every input automatically guarantees norm convergence.

Theorem III.1: Let G, $(G_n)_{n \ge 1}$ be causal BIBO-stable convolution operators defined as in (6). Then $G_n \to G$ strongly (if and) only if $G_n \to G$ in norm.

Proof: Without loss of generality $G = 0$. We suppose that to obtain a contradiction $G_n u \to 0$ for all $u \in \ell_\infty$ but that $\|G_n\| \not\to 0$. By passing to a subsequence, relabeling, and scaling, we may assume without loss of generality that $\|G_n\| > 1$ for each n.

We shall use the $\vee$ notation for concatenating finite disjoint sequences while keeping them in position, so that

$$(a \vee b)(k) = \begin{cases} a(k) & \text{if } a(k) \ne 0; \\ b(k) & \text{if } b(k) \ne 0; \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

By Lemma III.1 we can find a finitely supported sequence u_1 of norm one such that $\|G_1 u_1\| > 1$. Let t_1 be any index such that $|G_1 u_1(t_1)| > 1$. By considering the kernel $(g_1(t_1,k))_{k=1}^{\infty} \in \ell_1$, we see that there is a finite set F_1 containing the support of u_1 such that $|G_1(u_1 \vee v)(t_1)| > 1$ for any $v \in \ell_\infty(Z)$ supported on the complement of F_1 and of norm at most one.

Since $G_n \to 0$ strongly, we can now find another element of the sequence, without loss of generality G_2, such that the norm of the restriction of G_2 to the subspace of ℓ_∞ consisting of all vectors supported on F_1 is at most $1/4$. In particular $\|G_2 u_1\| \le 1/4$. However, there is a finitely-supported vector $\tilde{u}_2$ of norm one such that $\|G_2 \tilde{u}_2\| > 1$. By discarding those coefficients of $\tilde{u}_2$ which lie in F_1, we obtain a vector u_2 with finite support disjoint from F_1 and such that $\|G_2(u_2)\| > 3/4$.

Thus $\|G_1(u_1 \vee u_2)\| > 1$ and $\|G_2(u_1 \vee u_2)\| > 1/2$. As above we can define a finite set $F_2 \supset F_1$, containing the support of $u_1 \vee u_2$, such that $\|G_2(u_1 \vee u_2 \vee v)\| > 1/2$ for any v supported on the complement of F_2.

We proceed inductively, obtaining now G_3 and u_3 such that $\|G_1(u_1 \vee u_2 \vee u_3)\| > 1$, $\|G_2(u_1 \vee u_2 \vee u_3)\| > 1/2$, and $\|G_3(u_1 \vee u_2 \vee u_3)\| > 1/2$.

Finally we obtain a vector $u = u_1 \vee u_2 \vee \cdots \in \ell_\infty(Z)$ such that $G_n u \not\to 0$, a contradiction. □

The fairly well-known corollary now follows.

Corollary III.1: Let G, $(G_n)_{n \ge 1}$ be LTI BIBO stable systems. Then $G_n \to G$ strongly if and only if $G_n \to G$ in norm.

In fact, weak convergence of LTI BIBO-stable systems is equivalent to norm convergence. However, in general, weak convergence does not imply norm convergence for causal BIBO-stable LTV systems. Thus there are fundamental limitations as to the achievable identification accuracy for BIBO-stable LTV systems even when there is no noise. The next result illustrates the significance of input properties.

Proposition III.1: Let G, $(G_n)_{n \ge 1}$ be LTI BIBO-stable systems restricted to $\ell_\infty^- \oplus c_0$. Then the condition that $G_n \to G$ strongly in $\ell_\infty^- \oplus c_0$ does not imply that $G_n \to G$ in norm.

Proof: Consider $(G_n)_{n \ge 1}$ with unit impulse response defined by $g_n(k) = 1/n$ for $k = 1, 2, \cdots, n$, and $g(k) = 0$ otherwise. Take any $u \in \ell_\infty^- \oplus c_0 \ne 0$. Given any $\epsilon > 0$, there then exists an integer K_ϵ such that $|u(t)| \le \epsilon$ for $t \ge K_\epsilon$. Furthermore, take $n \ge K_\epsilon \times \max(1, \|u\|_\infty/\epsilon)$. Then $|(G_n u)(t)| \le 2\epsilon$. As $\epsilon > 0$ is arbitrary, it follows that for any $u \in \ell_\infty^- \oplus c_0$, $\|G_n u\| \to 0$ when $n \to \infty$. The result follows as $\|G_n\| = 1$ for all n. □

We can now give a useful characterization of the class of all causal linear fading memory operators on ℓ_∞. Intuitively, the fading memory systems are the ones for which the outputs tend to zero whenever the inputs do.

Theorem III.2: A bounded causal linear operator $G\colon \ell_\infty(Z) \to \ell_\infty^+$ has fading memory if and only if G maps the subspace $\ell_\infty^- \oplus c_0$ into c_0.

Proof: We modify the proof of Proposition 2.3 in [18] to allow for the fact that we are allowed inputs for $t < 0$.

Let $\epsilon > 0$ be given and suppose that G does map $\ell_\infty^- \oplus c_0$ into c_0.

For each $j \ge 0$, the sequence $((I - P_n)GP_j)_{n=1}^{\infty}$ is easily seen to tend to zero strongly, since GP_j maps into $\ell_\infty^- \oplus c_0$ and hence also in operator norm. Let us choose increasing indexes $n(j)$ such that $\|(I - P_{n(0)})GP_0\| < \epsilon/2$ and $\|(I - P_{n(j)})Ge_j\| < \epsilon/2^{j+1}$ for $j \ge 1$.

Define $\tilde{G}$ by $\tilde{G}P_0 = P_{n(0)}GP_0$ and $\tilde{G}e_j = P_{n(j)}Ge_j$ for $j \ge 1$. It is clear that $\tilde{G}$ has finite memory and

$$\begin{aligned} \|(G - \tilde{G})u\| &\le \|(G - \tilde{G})P_0 u\| + \sum_{j=1}^{\infty} \|(G - \tilde{G})u_j e_j\| \\ &< \epsilon \|u\|_\infty. \end{aligned} \tag{9}$$

Conversely, since for any bounded operator G the space $G(\ell_\infty^- \oplus c_0)$ is contained in the closure of $G(\ell_\infty^- \oplus c_{00})$, it follows that if G does not map $\ell_\infty^- \oplus c_0$ into c_0, then there is an eventually zero sequence, say $v \in \ell_\infty^- \oplus \in c_{00}$, such that $Gv \notin c_0$; then the operator G clearly does not have fading memory. □

We now strengthen the notions of finite and fading memory systems.

Definition III.1: Let $M \ge 1$. A causal BIBO-stable operator $G : \ell_\infty \to \ell_\infty^+$ corresponding to a kernel $g(t,k)$ as in (6) is said to have finite memory of length M if $g(t,k) = 0$ for $k > M$. A system is said to have strongly fading memory if it is the norm limit of a sequence $(G_M)_{M \ge 1}$ where each G_M has finite memory of length M. We write $\mathcal{B}_{\rm SFM}$ for the class of strongly fading memory systems.

Clearly any system in $\mathcal{B}_{\text{SFM}}$ is a fading memory system. However, there is another neat way of characterizing systems in $\mathcal{B}_{\text{SFM}}$.

Lemma III.2: A BIBO-stable system $G = (g_t)_{t\geq 1}$ is in $\mathcal{B}_{\text{SFM}}$ if and only if the sequence $(g_t)_{t\geq 1}$ forms a relatively compact set in ℓ_1.

Proof: This follows immediately because a bounded subset $K \subset \ell_1$ is relatively compact if and only if $\lim_{n\to\infty} \sup_{f\in K} \|(I - P_n)f\| = 0$ (see [4]). □

Note that although clearly $\mathcal{B}_{\text{SFM}} \subset \mathcal{B}_{\text{FM}}$, there are operators in $\mathcal{B}_{\text{FM}}$ which are not in $\mathcal{B}_{\text{SFM}}$, for example, the finite memory operator G with $g_t = e_t$ for each $t \geq 1$.

However, for the purposes of approximation and identification, the space $\mathcal{B}_{\text{SFM}}$ is still too large since it is nonseparable. This we shall see in the next section, where we shall be interested in looking at smaller, separable, subspaces of $\mathcal{B}_{\text{FM}}$.

IV. Asymptotically Periodic Systems

We recall that an LTV causal BIBO-stable operator G is determined by a kernel $(g(t,k))_{t\geq 1, k\geq 1}$ as in (6), corresponding to a bounded sequence $(g_t)_{t\geq 1}$ of elements of ℓ_1 with $\|G\| = \sup_{t\geq 1} \|g_t\|_1$.

Thus we can regard $\mathcal{B}_{\text{TV}}$ as a space of operators induced by elements of the Banach space $\ell_\infty(\ell_1) = \ell_1 \oplus \ell_1 \oplus \cdots$ with norm $\|(h_t)\| = \sup_{t\geq 1} \|h_t\|_1$. Moreover, G is time-invariant if and only if the sequence (g_t) is a constant sequence in ℓ_1. Hence, we are interested in separable subspaces of $\ell_\infty(\ell_1)$ which contain all the constant sequences.

As before, let X be any normed space and define the left shift $S: \ell_\infty(X) \to \ell_\infty(X)$ by $(Sf)_t = f_{t+1}$, $t \geq 1$. Note that the left shift "loses" the first term. Motivated by ideas from the theory of almost periodic functions [1], [4], [8], [11], we make the following definitions. Note that although almost periodic systems are in general time-varying, there is some restriction on the extent to which they can vary in time.

Definition IV.1: A sequence $h = (h_t)_{t\geq 1} \in \ell_\infty(X)$ is said to be almost periodic if $(S^n h)_{n\geq 0}$ forms a relatively compact set. That is, given any sequence of translates $(S^{n(k)}h)$, there is a convergent subsequence. The space of almost periodic sequences in $\ell_\infty(X)$ will be denoted $\text{AP}(X)$.

A sequence $h = (h_t)_{t\geq 1} \in \ell_\infty(X)$ is said to be asymptotically periodic if h is in the closed linear span of $c_0(X)$ and the periodic sequences in $\ell_\infty(X)$. The space of asymptotically periodic sequences in $\ell_\infty(X)$ will be denoted $\text{ASP}(X)$.

We write $\mathcal{B}_{\text{AP}}$ and $\mathcal{B}_{\text{ASP}}$ for the classes of systems induced by almost periodic and asymptotically periodic kernels, $\text{AP}(\ell_1)$ and $\text{ASP}(\ell_1)$, respectively.

We are now ready to establish the strict inclusions

$$\mathcal{B}_{\text{ASP}} \subset \mathcal{B}_{\text{AP}} \subset \mathcal{B}_{\text{SFM}} \subset \mathcal{B}_{\text{FM}} \subset \mathcal{B}_{\text{TV}} \tag{10}$$

and see that only the space $\mathcal{B}_{\text{ASP}}$ is separable.

Theorem IV.1:

i) $\text{AP}(X)$ is nonseparable. Any kernel in $\text{AP}(\ell_1)$ determines an operator in $\mathcal{B}_{\text{SFM}}$, but there are strongly fading memory operators on ℓ_∞ which are not determined by almost periodic sequences.

ii) If X is separable, then $\text{ASP}(X)$ is also separable.

iii) A sequence $h = (h_j)$ is in $\text{ASP}(X)$ if and only if given $\epsilon > 0$, there are integers N, $M > 0$ such that $\|S^{N+jM}h - S^{N+kM}h\| < \epsilon$ for all j, $k \geq 0$.

iv) $\text{ASP}(X)$ is contained in $\text{AP}(X)$. Hence any kernel in $\text{ASP}(\ell_1)$ determines a strongly fading memory operator.

Proof:

i) In the scalar case $X = R$, we take an uncountable bounded set in ℓ_∞, namely $h^\theta(j) = \cos j\theta$, where θ ranges over the interval $[0, 2\pi)$. It is easy to see that each h^θ is in $\text{AP}(R)$ and $\|h^\theta - h^\phi\| \geq 1$ for $\theta \neq \phi$. Were $\text{AP}(R)$ separable, this could not happen. The general case is similar, taking $h^\theta(j) = x \cos j\theta$ where $x \in X$ is nonzero and arbitrary.

Suppose now that $h = (h_t) \in \text{AP}(\ell_1)$ determines an operator G. Then, since $(S^n h)_{n\geq 1}$ form a relatively compact set in $\ell_\infty(\ell_1)$, we see on considering the first coordinate that $(h_n)_{n\geq 1}$ forms a relatively compact set in ℓ_1 and the operator has strongly fading memory.

Clearly the operator taking e_t to $\lambda_t e_{t+1}$ for each $t \geq 1$, where (λ_t) is any bounded sequence, has strongly fading memory (indeed it is a system having finite memory of length one). However, it corresponds to a kernel $(h_t) = (0, \lambda_1 e_1, \lambda_2 e_1, \cdots)$ and so is not almost periodic in general—for example, if (λ_n) is a random binary sequence consisting of values ± 1, then with probability one, the kernel will fail to be almost periodic. The same is true whenever (λ_n) is an infinite Galois sequence [12].

ii) If X is separable, then let (x_n) be a countable dense set in X. Now $\text{ASP}(X)$ is the closed linear span of the countable family $(x_n, 0, 0, \cdots)$, $(0, x_n, 0, 0, \cdots)$, $\cdots$, $(x_n, x_n, \cdots)$, $(x_{n_1}, x_{n_2}, x_{n_1}, x_{n_2}, \cdots)$, $(x_{n_1}, x_{n_2}, x_{n_3}, x_{n_1}, x_{n_2}, x_{n_3}, \cdots)$ and is therefore separable.

iii) If h is in $\text{ASP}(X)$, then there are sequences $f \in c_0(X)$ and $g \in \ell_\infty(X)$ such that g is periodic (period M, say) and $\|h - f - g\| < \epsilon/4$. Then $S^r f \to 0$, and $S^r g$ has period M so it is clear that $\|S^{N+jM}h - S^{N+kM}h\|$ is no greater than $\|S^{N+jM}(h-f-g) - S^{N+kM}(h-f-g)\| + \|S^{N+jM}f - S^{N+kM}f\| + \|S^{N+jM}g - S^{N+kM}g\|$ which is less than $\epsilon/2 + 0 + \epsilon/2 = \epsilon$, provided that N is sufficiently large (independent of j and k).

Conversely, if given $\epsilon > 0$, there are N and M such that $\|S^{N+jM}h - S^{N+kM}h\| < \epsilon$ for all j and k, then we can define g to be M-periodic such that $g_t = h_t$ for $N+1 \leq t \leq N+M$ and $f \in c_{00}$ to be a finite sequence of length N such that $f_t + g_t = h_t$ for $1 \leq t \leq N$. It is now easy to see that $h - f - g$ has infinity norm at most ϵ.

iv) This follows easily from iii) by a diagonal argument. Given a sequence $(S^{n(k)}h)$ of translates of h, we begin by taking $\epsilon = 1/2$ and using iii) to obtain a subsequence $(S^{n(1,k)}h)$ of $(S^{n(k)}h)$ such that $\|S^{n(1,p)}h - S^{n(1,q)}h\| < 1/2$ for all p and q. Then, in general, take $\epsilon = 1/2^r$ and obtain a subsequence $(S^{n(r,k)}h)$ of $(S^{n(r-1,k)}h)$ such that $\|S^{n(r,p)}h - S^{n(r,q)}h\| < 1/2^r$ for all p and q. It is now easily verified that $(S^{n(r,r)}h)$ is a convergent subsequence of $(S^{n(k)}h)$. □

Having obtained a suitable separable subspace of $\mathcal{B}_{\text{TV}}$, it remains only to write down a suitable sequence of models whose linear span is dense in $\mathcal{B}_{\text{ASP}}$.

If we work with $\text{ASP}(\ell_1)$, then we need only find a natural countable sequence (h^n) whose closed linear span is all of $\text{ASP}(\ell_1)$. If one assumes that the most natural systems to take as a first approximation are LTI systems, then this should include a model set for the constant sequences in $\text{ASP}(\ell_1)$, e.g., the FIR systems $(e_n, e_n, e_n, \cdots)$ with e_n the natural basis of ℓ_1.

One then has to include in some order (normally by a diagonal technique) the period-two sequences, looking like $(e_n, 0, e_n, 0, \cdots)$ and $(0, e_m, 0, e_m, \cdots)$, the period-three sequences, etc., but also the null sequences $(e_n, 0, 0, \cdots)$, $(0, e_n, 0, 0, \cdots)$ and so on. One could order these basis vectors by complexity (to be defined according to what sort of model is considered to be most "reasonable").

These LTV models can then be used for input–output identification, e.g., using Chebyshev (linear programming) methods as in [12] and [16]. However, we remark briefly that there is a significant difference from the LTI case in that a single input may no longer suffice—indeed it is not hard to see that given any $u \in \ell_\infty$ there is a nontrivial system $G \in \mathcal{B}_{\mathrm{TV}}$ such that $Gu = 0$.

V. Conclusion

We have studied various classes of LTV BIBO-stable systems from the point of view of approximate modeling and proved some characterizations of them. These have included fading memory systems, strongly fading memory systems, almost periodic systems, and asymptotically periodic systems. Of these, only the space of BIBO-stable asymptotically periodic systems is separable and thus allows systematic model parameterizations. It would be important to find other natural separable subspaces of LTV systems. LTV systems exhibit fundamental limitations as to how accurately they can be identified from input–output data, cf. Tikku and Poolla [22]. Thus they provide a setting in which to understand the importance of realistic assumptions and *a priori* information about the system to be identified to obtain satisfactory modeling results.

References

[1] L. Amerio and G. Prouse, *Almost-Periodic Functions and Functional Equations*. New York: Van Nostrand, 1971.
[2] S. Boyd and L. O. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Trans. Circ. Syst.*, vol. CS-32, pp. 1150–1161, 1985.
[3] M. A. Dahleh, "Asymptotic worst-case identification with bounded noise," in *The Modeling of Uncertainty in Control Systems*, R. S. Smith and M. Dahleh, Eds. New York: Springer-Verlag, 1994, pp. 157–170.
[4] N. Dunford and J. T. Schwarz, *Linear Operators Part I: General Theory*. New York: Wiley, 1957.
[5] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.
[6] M. Gevers, "Toward a joint design of identification and control," in *Essays on Control: Perspectives in the Theory and Its Applications*, H. L. Trentelman and J. C. Willems, Eds. Boston: Birkhäuser, 1993.
[7] T. K. Gustafsson and P. M. Mäkilä, "Modeling of uncertain systems via linear programming," in *Proc. 12th IFAC World Congr.*, 1993, Sydney, pp. 293–298.
[8] Y. Katznelson, *An Introduction to Harmonic Analysis*. New York: Wiley, 1968.
[9] R. L. Kosut, G. C. Goodwin, and M. P. Polis, Guest Eds., Special issue on system identification for robust control design, *IEEE Trans. Automat. Control*, vol. 37, pp. 899–1008, 1992.
[10] L. Lin, L. Y. Wang, and G. Zames, "Uncertainty principles and identification n-widths for LTI and slowly varying systems," in *Proc. Amer. Contr. Conf.*, Chicago, 1992.
[11] P. M. Mäkilä, "H^∞-optimization and optimal rejection of persistent disturbances," *Automatica*, vol. 26, pp. 617–618, 1990.
[12] ——, "Robust identification and Galois sequences," *Int. J. Contr.*, vol. 54, pp. 1189–1200, 1991.
[13] P. M. Mäkilä and J. R. Partington, "Robust approximate modeling of stable linear systems," *Int. J. Contr.*, vol. 58, pp. 665–683, 1993.
[14] P. M. Mäkilä, J. R. Partington, and T. K. Gustafsson, "Robust identification," *Prepr. 10th IFAC Symp. Syst. Identification*, Copenhagen, 1994, vol. 1, pp. 45–63.
[15] B. M. Ninness and G. C. Goodwin, "Estimation of model quality," *Prepr. 10th IFAC Symp. Syst. Identification*, Copenhagen, 1994, vol. 1, pp. 25–44.
[16] J. R. Partington, "Interpolation in normed spaces from the values of linear functionals," *Bull. London Math. Soc.*, vol. 26, pp. 165–170, 1994.
[17] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 951–959, 1994; also in *Proc. 1992 Amer. Contr. Conf.*, Chicago.
[18] J. S. Shamma, "The necessity of the small-gain theorem for time-varying and nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1138–1147, 1991.
[19] J. S. Shamma and R. Zhao, "Fading-memory feedback systems and robust stability," *Automatica*, vol. 29, pp. 191–200, 1993.
[20] R. S. Smith and M. Dahleh, Eds., "The modeling of uncertainty in control systems," in *Proc. 1992 Santa Barbara Workshop*. London: Springer-Verlag, 1994.
[21] R. S. Smith and J. C. Doyle, "Model validation—A connection between robust-control and identification," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 942–952, 1992.
[22] A. Tikku and K. Poolla, "On the worst-case identification of slowly-varying systems," *Prepr. 10th IFAC Symp. Syst. Identification*, Copenhagen, vol. 2, pp. 127–132, 1994.
[23] G. Zames and L. Y. Wang, "Local-global double-algebras for slow H^∞ adaptation: Part I: Inversion and stability," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 130–151, 1991.
[24] T. Zhou and H. Kimura, "Time domain identification for robust control," *Syst. Contr. Lett.* vol. 20, pp. 167–178, 1993.

Local l_p-Stability and Local Small Gain Theorem for Discrete-Time Systems

Henri Bourlès

***Abstract*—The notion of local l_p-stability is defined. The relationship between this notion and Lyapunov stability is clarified. A local version of the small gain theorem is then established in the case of discrete-time systems. These results are applied to stability analysis of a nonlinear discrete-time delay system.**

I. Introduction

The small gain theorem [8], [15] plays a fundamental role in stability analysis of nonlinear systems in an input–output viewpoint. This theorem applies to discrete-time systems and to continuous-time ones as well. The type of stability which is then obtained is "l_p-stability" ($1 \le p \le \infty$); see, e.g., [13]. This approach was limited by the fact that only global results are available in the literature. This will become clear in the following discussion; consider the standard closed-loop system in Fig. 1.

Let us denote as S^n the linear space of all sequences $x = (x(0), x(1), \cdots)$, where $x(t) \in R^n$, $\forall t$; $\mathbf{G_1}$ is a causal input–output operator $S^n \to S^m$ associated with a system Σ_1 [2][1]; suppose that the l_p-gain of $\mathbf{G_1}$ is finite and is denoted as $\gamma_p(\mathbf{G_1})$. Moreover, assume that $\mathbf{G_2}$ is a causal memoryless operator $S^m \to S^n$, defined by a nonlinearity $\Phi: N \to R^m \to R^n$, as follows:

$$(\mathbf{G_2}x)(t) = \Phi(t, x(t)), \quad \forall x \in S^m,\ \forall t \in N \tag{1}$$

Manuscript received March 14, 1994; revised March 17, 1995.
The author is with the Electricité de France, 92141 Clamart Cedex, France and Ecole Normale Supérieure de Cachan, 94230 Cachan, France.
Publisher Item Identifier S 0018-9286(96)02820-6.

[1] This notion was defined in [2] in the continuous-time case. In the discrete-time case, let Σ be a system defined by a state-space realization $\eta(t+1) = f(t, \eta(t), u(t)), y(t) = g(t, \eta(t), u(t))$, where $\eta(t)$ is the state, $u(t) \in R^m, y(t) \in R^q$. Assume that $\eta(0) = 0$ (zero initial condition at initial time $t = 0$). Then, it is easy to prove by induction that for any $t \ge 0, y(t)$ can be expressed in function of $t, u(0), \cdots, u(t)$; in other words, there exists a causal operator $\mathbf{G}: S^m \to S^q$ such that $y = \mathbf{G}u$; $\mathbf{G}$ is called the input–output operator associated with Σ. If Σ is time-invariant, the initial time can be shifted without inconvenience.