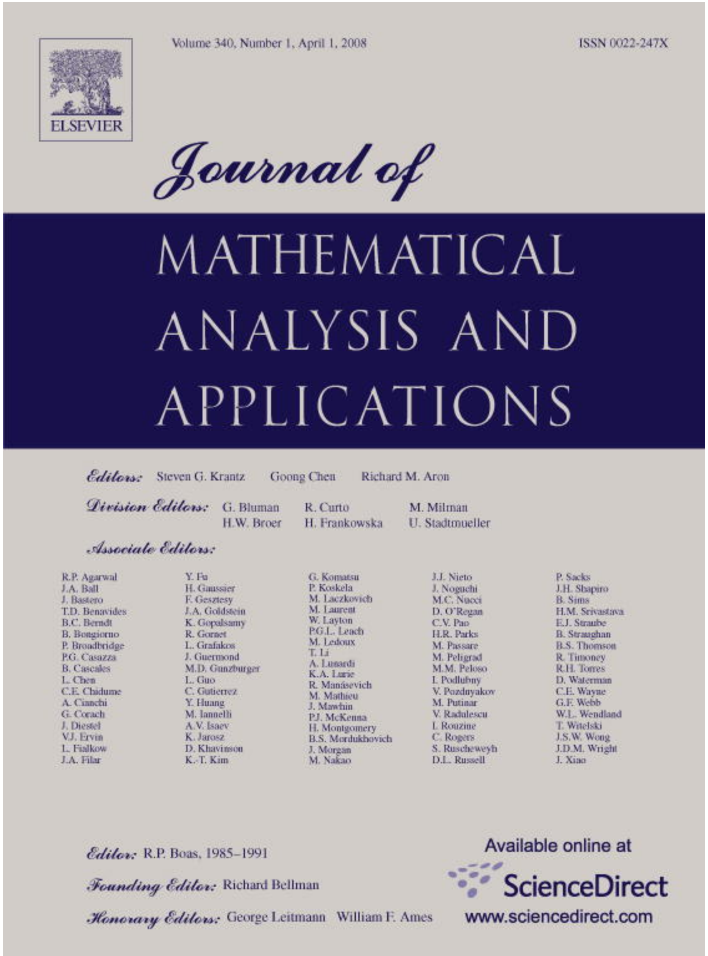
Volume 340, Number 1, April 1, 2008
ISSN 0022-247X
ELSEVIER
Journal of
MATHEMATICAL ANALYSIS AND APPLICATIONS
Editors: Steven G. Krantz Goong Chen Richard M. Aron
Division Editors: G. Bluman H.W. Broer R. Curto H. Frankowska M. Milman U. Stadtmueller
Associate Editors:
Editor: R.P. Boas, 1985–1991
Founding Editor: Richard Bellman
Honorary Editors: George Leitmann William F. Ames
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
J. Math. Anal. Appl. 340 (2008) 116–125
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

On the outputs of linear control systems

Joel H. Shapiro

Department of Mathematics and Statistics, Portland State University, Portland, OR 97207, USA

Received 10 May 2007

Available online 24 August 2007

Submitted by Richard M. Aron

Abstract

This paper studies autonomous, single-input, single-output linear control systems on finite time intervals. The object of interest is the *output operator* $\mathcal{O}$, which associates to each input function and initial state vector the corresponding system output. *Main result*: If the system has relative degree $r < \infty$, then for any "admissible" Banach space $\mathscr{U}$ of inputs, $\mathcal{O}$ is a bounded operator taking $\mathscr{U} \times \mathbb{C}^n$ onto the "Sobolev space" of complex functions $f \in C^{(r-1)}([0, T])$ for which the $(r-1)$-order derivative $f^{(r-1)}$ is absolutely continuous, with $f^{(r)} \in \mathscr{U}$. This completes recent results of Jönsson and Martin [Ulf Jönsson, Clyde Martin, Approximation with the output of linear control systems, J. Math. Anal. Appl. 329 (2007) 798–821] who showed that if the system is minimal and $\mathscr{U}$ is either $L^2([0, T])$ or $C([0, T])$, then $\mathcal{O} : \mathscr{U} \times \mathbb{C}^n \to \mathscr{U}$ has dense range.

Keywords: Linear control system; Volterra operator; Sobolev space

1. Introduction

This paper completes recent work initiated by Jönsson and Martin [1] who, for autonomous, single-input, single-output (SISO) linear control systems that are both controllable and observable, studied the "output operator" $\mathcal{O}$, which associates to each input function $u : [0, T] \to \mathbb{C}$ and initial state $x_0 \in \mathbb{C}^n$ the corresponding system output. They showed that for input space $\mathscr{U} = L^2([0, T])$ or $C([0, T])$ the operator $\mathcal{O}$ maps $\mathscr{U} \times \mathbb{C}^n$ to $\mathscr{U}$, boundedly, and (their main point) *with dense range*.

The work below exploits the quasinilpotence of Volterra convolution operators to obtain the following characterization of the range of $\mathcal{O}$ for a very general class of input spaces that are "admissible" in a sense to be described shortly:

If an autonomous SISO linear system with state space $\mathbb{C}^n$ has finite relative degree r, then for any admissible Banach space $\mathscr{U}$ of inputs on $[0, T]$, the output operator $\mathcal{O}$ takes $\mathscr{U} \times \mathbb{C}^n$ boundedly onto the "Sobolev space" $W^{(r)}(\mathscr{U})$ consisting of functions in $f \in C^{(r-1)}([0, T])$ with $f^{(r-1)}$ absolutely continuous and $f^{(r)} \in \mathscr{U}$.

E-mail address: shapiroj@pdx.edu.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2007.08.032

Here "boundedly" means that $\mathcal{O}$ is continuous when viewed as a mapping from $\mathscr{U} \times \mathbb{C}^n$ to $W^{(r)}(\mathscr{U})$, and "relative degree" is, roughly speaking, the number of times you have to differentiate the output in order to see the input (details to follow in Section 2.3).

Two important cases deserve special mention: if $\mathscr{U} = L^p([0,T])$ with $p \geqslant 1$, then $W^{(r)}(\mathscr{U})$ is the classical Sobolev space $W^{p,r}([0,T])$, while if $\mathscr{U} = C([0,T])$, then $W^{(r)}(\mathscr{U}) = C^{(r)}([0,T])$. Other examples of admissible input spaces $\mathscr{U}$ include Orlicz spaces and more generally, as we will see in Section 2.5, many Banach function spaces.

The definition of "admissible input space" shares the next section with the required control- and operator-theoretic preliminaries. Then follows the proof of the main theorem, after which the concluding section records some extensions to function spaces over semi-closed and semi-infinite intervals, along with some further connections with the work of [1].

2. Preliminaries

Following Jönnson and Martin [1], the systems considered here all have the form

$$\dot{x} = Ax + bu, \tag{1}$$

$$y = cx, \tag{2}$$

where A is an $n \times n$ (complex) matrix, b is an n-dimensional column vector (i.e., $b \in \mathbb{C}^n$), c is an n-dimensional row vector, and u is a complex-valued integrable "input" defined on a finite interval $[0,T]$. The matrices in (1), (2) above will always be assumed nonzero. For solutions $x:[0,T] \to \mathbb{C}^n$ of the differential equation (1), the value $x(t)$ reports the "state" of the system at time t, and (2) gives a scalar function y that is the observed "output."

2.1. *Notation and terminology*

(a) The triple of matrices (A,b,c) will be used as a shorthand notation for the system given by Eqs. (1) and (2).
(b) The usual superscript convention "$f^{(k)}$" will be used to denote the kth order derivative of a function f, with the "zeroth order derivative" being the function itself.
(c) The term "isomorphism" between two Banach spaces, say X and Y, will refer to a linear transformation of X onto Y that is one-to-one, bounded (i.e., continuous), and whose inverse is—necessarily, thanks to the Open Mapping Theorem—continuous.
(d) An *isometry* S of a Banach space X, with norm $\|\cdot\|_X$ into another Banach space Y, with norm $\|\cdot\|_Y$, is a linear map that preserves norms: $\|Sx\|_Y = \|x\|_X$ for each $x \in X$. Such a map is clearly bounded, and if $S(X) = Y$ it is an isomorphism of X onto Y—an "isometric isomorphism."
(e) The terms: "operator," "linear operator," "bounded operator," "continuous linear operator" ... all mean the same thing.

2.2. *Output operators*

For a SISO control system (A,b,c), the *output kernel* is the entire function $K:\mathbb{C} \to \mathbb{C}$ defined by

$$K(z) = ce^{zA}b \quad (z \in \mathbb{C}). \tag{3}$$

The variation-of-constants formula implies that for each initial state $x_0 \in \mathbb{C}^n$ and input function u integrable on $[0,T]$, the corresponding output $y = cx \overset{\text{def}}{=} \mathcal{O}(u,x_0)$ is given by

$$\mathcal{O}(u,x_0)(t) = ce^{tA}x_0 + V_K u(t) \quad (0 \leqslant t \leqslant T), \tag{4}$$

where V_K is the *Volterra convolution operator*

$$V_K u(t) = \int_{\tau=0}^{t} K(t-\tau)u(\tau)\,d\tau \quad (0 \leqslant t \leqslant T). \tag{5}$$

More generally (but not nearly as generally as possible), if K is *any* entire function we will call the operator V_K, defined by (5) for functions integrable on $[0, T]$, a *Volterra convolution operator with kernel* K. The kernel $K \equiv 1$ produces the standard Volterra operator V,

$$Vu(t) = \int_{\tau=0}^{t} u(\tau)\,d\tau,$$

henceforth called simply … "the Volterra operator."

2.3. Relative degree

To say the system (A, b, c) has *finite relative degree* means that there exists a positive integer k such that $cA^{k-1}b \neq 0$. The least such k is called the *relative degree* of the system, written $r = r(A, b, c)$. By the Cayley–Hamilton theorem this positive integer must be $\leqslant n$; for this reason we will refer to "relative degree n" as "full relative degree." If the system does not have finite relative degree, it is said to have *infinite relative degree*: $r(A, b, c) = \infty$.

By (3) we have for the output kernel K of the system, $K^{(j)}(0) = cA^j b$ for each positive integer j, hence to say that, for example, "the system (A, b, c) has relative degree 1" means that $K(0) = cb \neq 0$, and to say its relative degree r is "finite and > 1" means that the r relations below hold:

$$K^{(j)}(0) = cA^j b = 0 \quad (0 \leqslant j \leqslant r-2) \quad \text{and} \quad K^{(r-1)}(0) = cA^{r-1}b \neq 0. \tag{6}$$

To understand how the notion of relative degree occurs in the output identification problem being considered here, note that upon differentiating both sides of (5) we obtain, for any input u integrable on $[0, T]$,

$$DV_K u(t) = \int_{\tau=0}^{t} \dot{K}(t-\tau)u(\tau)\,d\tau + K(0)u(t) \quad (0 \leqslant t \leqslant T),$$

where differentiation with respect to t is denoted on the left by "D" (to emphasize its "operator-ness"), and on the right by a raised dot (to save space). Thus if the system has relative degree 1, it just takes one differentiation of the output formula (4) for the input to appear explicitly on the right-hand side.

If the system has finite relative degree $r > 1$, then (6) shows that, on $L^1([0, T])$,

$$D^j V_K = V_{K^{(j)}} \quad \text{for} \quad j < r, \tag{7}$$

and

$$D^r V_K = V_{K^{(r)}} + K^{(r-1)}(0)I. \tag{8}$$

Since $K^{(r-1)}(0) \neq 0$ we see from (7) and (8) that it takes exactly r differentiations of (4) to make the input u appear explicitly.

The case of infinite relative degree is truly pathological; it occurs precisely when the output kernel (3) is identically zero. To see this one need only note that in the MacLaurin expansion of this kernel the coefficient of z^j is $cA^j b/j!$, which is zero for each j if and only if $r(A, b, c) = \infty$. Thus "infinite relative degree" means that the input has no effect.

For the remainder of this paper, "relative degree" means "finite relative degree."

For the concept of relative degree in a nonlinear setting, see, e.g., [9].

The next result, undoubtedly well known, will be needed for the proof of our main theorem (Theorem 5 below).

Proposition 1. *If the system (A, b, c) has relative degree r, then the set of row vectors $\{c, cA, \ldots, cA^{r-1}\}$ is linearly independent.*

Proof. Suppose, for the sake of contradiction, that there exist scalars γ_j $(0 \leqslant j \leqslant r-1)$, not all zero, such that

$$\sum_{j=0}^{r-1} \gamma_j c A^j = 0. \tag{9}$$

Let k be the largest index such that $\gamma_k \neq 0$. Rewrite (9) as

$$\gamma_k c A^k = \sum_{j=0}^{k-1} \gamma_j c A^j,$$

let $\delta = r - 1 - k$, and multiply both sides of the last equation on the right by $A^\delta b$. The result is

$$\gamma_k c A^{r-1} b = \sum_{j=0}^{k-1} \gamma_j c A^{j+\delta} b.$$

Since $j + \delta \leqslant r - 2$, every summand on the right is zero, thus (since $\gamma_k \neq 0$) forcing the contradiction $cA^{r-1}b = 0$. Thus all the coefficients γ_j in (9) must have been zero, which completes the proof. □

2.4. *Observability*

The system (A, b, c) is called *observable* whenever different initial conditions lead to different outputs, or what is equivalent: *the "partial output operator"* $\mathcal{O}(0, \cdot) : \mathbb{C}^n \to \mathbb{C}^n$ *is one-to-one*. The fundamental characterization of this concept, due to Kalman [3] (see also [4, Theorem 2, p. 289] or [8, Theorem 3, p. 713]), expresses it in terms of the $n \times n$ *observability matrix* M_{obs} whose jth row is cA^{j-1} $(1 \leqslant j \leqslant n)$:

The system (A, b, c) *is* observable *if and only if* M_{obs} *has rank* n.

In particular, if the system is observable, then $M_{\text{obs}} b \neq 0$ for any nonzero $b \in \mathbb{C}^n$, hence *observable systems have finite relative degree.* In the (partial) converse direction, Proposition 1 insures that *systems of full relative degree are observable*.

2.5. *Admissible input spaces*

An *admissible input space* will be a Banach space $\mathscr{U}$ of (a.e. equivalence classes of) integrable functions on $[0, T]$ that obeys the following axioms:

(I-1) The constant functions belong to $\mathscr{U}$.
(I-2) The Volterra operator V is a bounded linear operator on $\mathscr{U}$.
(I-3) If $f \in \mathscr{U}$, then $|f| \in \mathscr{U}$.
(I-4) If $f, g \in \mathscr{U}$ and $|f| \leqslant |g|$ a.e. on $[0, T]$, then $\|f\| \leqslant \|g\|$, where $\|\cdot\|$ is the norm of $\mathscr{U}$.

Here are some examples of admissible input spaces:

(a) $L^p = L^p([0, T])$ for $1 \leqslant p \leqslant \infty$. The spaces decrease in size as p increases; in particular all are contained in L^1. It is an easy matter to show that V maps L^1 boundedly into $C([0, 1])$, so by a simple comparison of norms, it maps L^p boundedly into itself for each p.
(b) $C([0, T])$, those continuous complex-valued functions on $[0, T]$, endowed with the "uniform norm":

$$\|f\|_\infty = \max_{0 \leqslant t \leqslant 1} \bigl|f(t)\bigr|.$$

(c) *Banach function spaces acted upon by* V. By a *Banach function space* on $[0, T]$ we will mean a linear subspace $\mathscr{B}$ of (a.e. equivalence classes of) functions integrable on $[0, T]$ that contains the constant functions and is endowed with a norm $\|\cdot\|$ such that:

(B1) $\mathscr{B}$ is a Banach space in the norm $\|\cdot\|$.

(B2) If $f \in \mathscr{B}$ and g is measurable on $[0,T]$ with $|g| \leqslant |f|$ a.e. on $[0,T]$, then $g \in \mathscr{B}$ and $\|g\| \leqslant \|f\|$.
(B3) $\mathscr{B}$ contains all constant functions.

We further assume:

(B4) $V(\mathscr{B}) \subset \mathscr{B}$.

It follows from (B2) that convergence in the norm of B implies convergence in measure [5, Proposition 2.6.3, p. 116]. This, along with the Closed Graph Theorem and (B4), shows that the Volterra operator V is bounded on B; thus B is an admissible input space.

The L^p spaces on $[0,T]$ for $1 \leqslant p \leqslant \infty$ are Banach function spaces of the type considered above. More generally, so are the *Orlicz spaces* L^φ *on* $[0,T]$ defined for convex, increasing functions on φ on $[0,\infty)$ with $\varphi(0)=0$ (see, for example, [6]). Here the L^p spaces correspond to the special case $\varphi(t)=t^p$. It is also possible to have Banach function spaces of functions locally integrable on $[0,T)$ that are larger than $L^1([0,T])$, but on which V acts boundedly, and for which all our arguments work. However, in order to minimize distractions it seems best to relegate such generalizations to the final section.

2.6. What is in $\mathscr{U}$?

Suppose $\mathscr{U}$ is an admissible input space on $[0,T]$. By axiom (I-1) the constant function 1 belongs to $\mathscr{U}$ and by (I-2) V operates on $\mathscr{U}$. Thus $t^j/j! = V^j 1 \in \mathscr{U}$ for each positive integer j, i.e. *$\mathscr{U}$ contains all the polynomial functions on* $[0,T]$. More is true:

$\mathscr{U}$ contains all functions represented by MacLaurin series with radius of convergence $> T$.

To prove this, suppose $u(t) = \sum_{j=0}^\infty a_j t^j$ is such a function. For integers p,q with $0 \leqslant p \leqslant q$ we know that the polynomial $\sum_{j=p}^q a_j t^j$ lies in $\mathscr{U}$, and so can make the following estimates of its $\mathscr{U}$-norm:

$$\left\| \sum_{j=p}^q a_j t^j \right\| \leqslant \sum_{j=p}^q |a_j| \|t^j\| \leqslant \sum_{j=p}^q |a_j| T^j \|1\|,$$

where the last inequality comes from axioms (I-1) and (I-4). Since the numerical series $\sum_j |a_j| T^j$ converges by our radius-of-convergence hypothesis, the above inequalities show that the MacLaurin series representing u is Cauchy in the norm of $\mathscr{U}$ hence, by completeness, $u \in \mathscr{U}$.

2.7. Sobolev spaces

Given an admissible input space $\mathscr{U}$ on $[0,T]$ and an integer $k>0$, by the *Sobolev space of order k based on* $\mathscr{U}$ we will mean the space $W^{(k)}(\mathscr{U})$ of functions $f \in C^{(k-1)}([0,T])$ (the $(k-1)$-times differentiable functions on $[0,T]$) for which $f^{(k-1)}$ is absolutely continuous on $[0,T]$, with $f^{(k)} \in \mathscr{U}$. Since the definition of "admissible" requires that $\mathscr{U} \subset L^1([0,T])$, the requirement that $f^{(k)}$ belong to $\mathscr{U}$ will usually be a stronger one than absolute continuity. The norm

$$\|f\|_W = \sum_{j=0}^{k-1} \left| f^{(j)}(0) \right| + \left\| f^{(k)} \right\|_{\mathscr{U}} \quad \left(f \in W^{(k)} \right) \tag{10}$$

is easily seen to make $W^{(k)}(\mathscr{U})$ into a Banach space (see below).

Examples. If $\mathscr{U} = L^p([0,T])$ with $1 \leqslant p \leqslant \infty$, then $W^{(k)}(\mathscr{U})$ is the classical Sobolev space $W^{p,k}([0,T])$. If $\mathscr{U} = C([0,T])$, then $W^{(k)}(\mathscr{U})$ is just $C^{(k)}([0,T])$.

Of particular importance will be the closed subspace $W_0^{(k)}(\mathscr{U})$ consisting of functions in $W^{(k)}(\mathscr{U})$ which vanish, along with their derivatives of orders $< k$, at the origin. Note that V^k is an isometry taking $\mathscr{U}$ onto $W_0^{(k)}(\mathscr{U})$. This shows that $W_0^{(k)}(\mathscr{U})$ is a Banach space, and since V maps $\mathscr{U}$ into itself, $W_0^{(k)}(\mathscr{U}) \subset \mathscr{U}$.

Upon letting $\mathscr{P}_j$ denote the space of polynomial functions on $[0, T]$ of degree $\leqslant j$, we see that $W^{(k)}(\mathscr{U})$ can be written as the algebraic and topological direct sum of $\mathscr{P}_{k-1}$ and $W_0^{(k)}(\mathscr{U})$. Thus $W^{(k)}(\mathscr{U}) \subset \mathscr{U}$ and is a Banach space in the norm given by (10).

2.8. What is in a Sobolev space?

In Section 2.6 we noted that each admissible input space on $[0, T]$ contains all functions represented by MacLaurin series with radius of convergence $> T$. In the course of proving Theorem 5 below we will need something more

Each such function belongs to $W^{(k)}(\mathscr{U})$ for each positive integer k.

Indeed, given such a function u, the MacLaurin series of each derivative has the same radius of convergence as that of u, and so each derivative is continuous on $[0, T]$, and, by Section 2.6, belongs to $\mathscr{U}$. Thus, by definition, u belongs to $W^{(k)}(\mathscr{U})$, for each k.

3. Output characterization

In this section we prove that "output ranges are Sobolev spaces." Everything depends on the Volterra integration operator V, which by hypothesis takes any admissible input space boundedly into itself. The crucial property of V is this:

Proposition 2. *V is quasinilpotent on any admissible input space.*

Note. To say that V is quasinilpotent means that its spectrum $\sigma(V)$ is the singleton $\{0\}$, i.e. that $V - \lambda I$ is invertible for every complex number $\lambda \neq 0$. The quasinilpotence of V on L^p spaces or $C([0, T])$ is well known, and the proof below follows a standard pattern.

Proof. For $u \in \mathscr{U}$ we have for each positive integer k,

$$V^k u(t) = \int_{\tau=0}^{t} \frac{(t-\tau)^{k-1}}{(k-1)!} u(\tau)\, d\tau \quad (0 \leqslant t \leqslant T), \tag{11}$$

hence simple estimates yield for each $t \in [0, T]$,

$$\left|V^k u(t)\right| \leqslant \frac{T^{k-1}}{(k-1)!} \big(V|u|\big)(t). \tag{12}$$

Since $u \in \mathscr{U}$, axiom (I-3) implies that $|u| \in \mathscr{U}$ and, by (I-4), u and $|u|$ have the same norm. Thus (I-4) and (12) yield (since both $V^k u$ and $V|u|$ belong to $\mathscr{U}$)

$$\left\|V^k u\right\| \leqslant \frac{T^{k-1}}{(k-1)!} \left\|V\big(|u|\big)\right\| \leqslant \frac{T^{k-1}}{(k-1)!} \|V\| \|u\|,$$

hence

$$\left\|V^k\right\| \leqslant \frac{T^{k-1}}{(k-1)!} \|V\| \quad (k = 1, 2, \ldots). \tag{13}$$

From this we see that

$$\lim_{k\to\infty} \left\|V^k\right\|^{1/k} = 0 \tag{14}$$

so $\sigma(V) = \{0\}$ by the spectral radius formula [2, Problem 88, pp. 48, 150, and 232]. □

The next step is to transfer the quasinilpotence of V to more general Volterra convolution operators. The result below is standard for $\mathscr{U} = L^p([0, T])$ or $C([0, T])$, even for kernels much more general than the ones we are using here (see [2, Problems 186–187, pp. 98–99, 298–299], for example).

Proposition 3. *Suppose $\mathscr{U}$ is an admissible input space and K is an entire function. Then the Volterra convolution operator V_K is bounded and quasinilpotent on $\mathscr{U}$.*

Proof. (a) *Boundedness.* To say K is entire means that its MacLaurin series $\sum_{j=0}^{\infty} \frac{K^{(j)}(0)}{j!} z^j$ converges for each $z \in \mathbb{C}$. The convergence is absolute, and uniform on each compact subset of $\mathbb{C}$. By (13) this guarantees the convergence of the numerical series $\sum_{k=0}^{\infty} K^{(j)}(0)\|V^{j+1}\|$, and therefore of the operator series

$$\sum_{k=0}^{\infty} K^{(j)}(0) V^{j+1} \tag{15}$$

in the norm of bounded operators on $\mathscr{U}$. Thus the sum of this series is a bounded operator on $\mathscr{U}$; I claim it is V_K.

Indeed, for each $u \in \mathscr{U}$ and $t \in [0, T]$ we have from the defining equation (5) for V_K, upon replacing K by its power series expansion and using uniform convergence on $[0, T]$ to interchange the order of summation and integration

$$V_K u(t) = \sum_{j=0}^{\infty} K^{(j)}(0) \int_{\tau=0}^{t} \frac{(t-\tau)^j}{j!} u(\tau)\, d\tau = \sum_{j=0}^{\infty} K^{(j)}(0) V^{j+1} u(t),$$

where the last equality follows from (11). This identifies the sum of the series (15) to be V_K, as desired.

(b) *Quasinilpotence.* From the definition (5) of V_K we see that for each $t \in [0, T]$,

$$\left|V_K u(t)\right| \leqslant \|K\|_\infty V|u|(t),$$

where $\|K\|_\infty = \max_{0 \leqslant \tau \leqslant T} |K(\tau)|$. A straightforward induction now shows that for each positive integer k,

$$\left|V_K^k u\right| \leqslant \|K\|_\infty^k \big(V^k|u|\big)$$

at each point of $[0, T]$. Thus if $u \in \mathscr{U}$ it follows (just as in the proof of Proposition 2) from (I-4) and the fact that $V_K^k(\mathscr{U}) \subset \mathscr{U}$ that

$$\left\|V_K^k u\right\| \leqslant \|K\|_\infty^k \left\|V^k\big(|u|\big)\right\| \leqslant \|K\|_\infty^k \left\|V^k\right\| \|u\|,$$

that is,

$$\left\|V_K^k\right\| \leqslant \|K\|_\infty^k \left\|V^k\right\|.$$

This, along with the previous estimate (14) for the Volterra operator shows that $\lim_{k\to\infty} \|V_K^k\|^{1/k} = 0$, hence $\sigma(V_K) = \{0\}$, again by the spectral radius formula. □

Remark. In both Propositions 2 and 3, one could circumvent the spectral radius formula by observing that whenever an operator S on a Banach space has the property that $\lim_k \|S^k\|^{1/k} = 0$, then for each nonzero complex number λ the *Neumann series* $\sum_{k=0}^{\infty} \frac{S^k}{\lambda^{k+1}}$ converges in the operator norm to a (necessarily) bounded operator, easily identified as the inverse of $\lambda I - S$. The Neumann series is, of course, at the heart of the proof of the spectral radius formula.

The next result, which is the major step in the proof of the main theorem, identifies the range of $V_K(\cdot) = \mathcal{O}(\cdot, 0)$.

Theorem 4. *Suppose the autonomous linear SISO system (A, b, c) has relative degree r. Let K denote its output kernel* (3). *Then for any admissible input space $\mathscr{U}$, the Volterra convolution operator $V_K(\cdot) = \mathcal{O}(\cdot, 0)$ is an isomorphism taking $\mathscr{U}$ onto $W_0^{(r)}(\mathscr{U})$.*

Proof. Recall from Section 2.3 that r is the least positive integer (necessarily $\leqslant n$) for which $K^{(r-1)}(0) = cA^{r-1}b \neq 0$.

Since K is an entire function, so are all its derivatives, and therefore by Proposition 3 the operator $V_{K^{(r)}}$ is quasinilpotent on $\mathscr{U}$. Since $K^{(r-1)}(0) \neq 0$, this quasinilpotence combines with Eq. (8) to guarantee that $D^r V_K$ is an isomorphism of $\mathscr{U}$ onto itself.

Now V^r is an isometry taking $\mathscr{U}$ onto $W_0^{(r)}(\mathscr{U})$ (see Section 2.7; this is essentially the *definition* of $W_0^{(r)}(\mathscr{U})$). Thus the operator $V^r D^r V_K$ is an isomorphism taking $\mathscr{U}$ onto $W_0^{(r)}(\mathscr{U})$. Furthermore (7) shows that, for each $u \in \mathscr{U}$, the image function $V_K u$ and all its derivatives of order $< r$ vanish at the origin. Consequently $V^r D^r$ is the identity on $V_K(\mathscr{U})$, hence $V_K = V^r D^r V_K$, so V_K takes $\mathscr{U}$ isomorphically onto $W_0^{(r)}(\mathscr{U})$. □

The main theorem now emerges as a consequence of Theorem 4.

Theorem 5. *Suppose the autonomous linear SISO system (A, b, c) has relative degree r. Then the output operator $\mathcal{O}$ takes $\mathscr{U} \times \mathbb{C}^n$ (boundedly) onto $W^{(r)}(\mathscr{U})$.*

Proof. By Theorem 4 and Section 2.8 we know that $\operatorname{ran} \mathcal{O} \subset W^{(r)}(\mathscr{U})$, so the only issue here is surjectivity. For this, fix $y \in W^{(r)}(\mathscr{U})$; the goal is to find $x_0 \in \mathbb{C}^n$ and $u \in \mathscr{U}$ such that $\mathcal{O}(u, x_0) = y$.

To this end, note that, by Proposition 1, the $r \times n$ principal submatrix of the observability matrix has rank r. Thus there is a vector $x_0 \in \mathbb{C}^n$ such that $M_{\mathrm{obs}} x_0$ has the successive derivatives $y^{(j)}(0)$ for $0 \leqslant j \leqslant r-1$ as its first r ($\leqslant n$, recall) entries. Thus

$$y^{(j)}(0) = cA^j x_0 = \frac{d^j}{dt^j} ce^{tA} x_0 \Big|_{t=0} \quad (j = 0, 1, \ldots, r-1),$$

so the function $y - ce^{tA}x_0$, which, by Section 2.8, belongs to $W^{(r)}(\mathscr{U})$, has derivatives of order 0 through $r-1$ vanishing at the origin, and therefore belongs to $W_0^{(r)}(\mathscr{U})$. By Theorem 4 this function is therefore in the range of V_K,

$$y - ce^{tA}x_0 = V_K u$$

for some $u \in \mathscr{U}$, hence

$$y = ce^{tA}x_0 + V_K u = \mathcal{O}(u, x_0),$$

which establishes the desired result: every $y \in W^{(r)}(\mathscr{U})$ lies in the range of the output operator. □

The proof of Theorem 5 establishes, by a different method, the following result from [1, Theorem 2.2, pp. 803–805]:

Corollary 6. *The output operator for the autonomous SISO linear system is one-to-one if and only if the system has full relative degree.*

In light of Theorem 5 this can be rephrased:

For an autonomous SISO linear system, the following are equivalent:
(a) *The system has full relative degree.*
(b) *The output operator is one-to-one.*
(c) *For any admissible input space $\mathscr{U}$, the output operator is an isomorphism taking $\mathscr{U} \times \mathbb{C}^n$ onto $W^{(n)}(\mathscr{U})$.*

4. Closing remarks

4.1. Extensions to locally integrable functions

The requirement that admissible input spaces be contained in $L^1([0, T])$ can be relaxed somewhat. Here are two possibilities.

(a) *Finite intervals.* Suppose, for example, that $w : [0, T) \to [0, \infty)$ is an integrable, decreasing function, and $d\mu(t) = w(t)\,dt$. Then $L^1(\mu)$ consists of a.e. equivalence classes of measurable functions which, while not necessarily integrable on $[0, T)$ (example: $w(t) = T - t$), are at least *locally integrable* in the sense that they are integrable on each closed subinterval. In particular, the Volterra operator V can be defined on $L^p(\mu)$ for $1 \leqslant p < \infty$ ($L^\infty(\mu)$

is just ordinary L^∞, so let us exclude it here), and thanks to the decreasing-ness of w, it is easy to check that V is actually a bounded operator on $L^p(\mu)$.

Thus $L^p(\mu)$ satisfies the input space axioms (I-1)–(I-4) of Section 2.5, with $[0,T]$ replaced by $[0,T)$, and integrability replaced by local integrability. In this slightly extended setting the arguments of Section 3 go through *verbatim*, provided we relax the definition of the Sobolev space $W^{(k)}(\mathscr{U})$ slightly, requiring the $C^{(k-1)}$-differentiability of the functions involved only on the semi-open interval $[0,T)$, and the absolute continuity of $f^{(k)}$ just "locally," i.e., only on each closed subinterval of $[0,T)$.

(b) $T=\infty$. The results obtained in Section 3 can be extended in a natural way to input spaces defined over the semi-infinite interval $[0,\infty)$. For example, if $1\leqslant p<\infty$ let $L^p_{\mathrm{loc}} = L^p_{\mathrm{loc}}([0,\infty))$ be the space of a.e. classes of measurable functions u on $[0,\infty)$ for which $|u|^p$ is integrable over every finite subinterval. For $s>0$ let $\|\cdot\|_s$ be the seminorm defined on L^p_{loc} by

$$\|u\|_s^p = \int_0^s \bigl|u(t)\bigr|^p\,dt \quad \bigl(u\in L^p_{\mathrm{loc}}\bigr). \tag{16}$$

Then the collection of all such seminorms makes L^p_{loc} into a complete, metrizable, locally convex linear topological space—the "projective limit" of the spaces $L^p([0,s])$.

The fact that all the operators considered here are bounded in each seminorm is (more than) enough to guarantee that they are continuous on L^p_{loc}. From this it is a simple matter to check that if, for $\mathscr{U}=L^p_{\mathrm{loc}}$, one defines the Sobolev space $W^{(k)}(\mathscr{U})$ in a manner analogous to that of Section 4.1(a), but now endowing it with its natural projective limit topology, then the results of Section 3 easily yield corresponding ones for the semi-infinite setting.

4.2. $\mathcal{O}:\mathscr{U}\times\mathbb{C}^n\to\mathscr{U}$

In [1] it is proved that if the autonomous linear SISO system (A,b,c) is both controllable and observable, and if $\mathscr{U}$ is $L^2=L^2([0,T])$ or $C=C([0,T])$, then the output operator $\mathcal{O}$, viewed now as a bounded operator $\mathscr{U}\times\mathbb{C}^n\to\mathscr{U}$ has dense range, and in fact for $\mathscr{U}=L^2$, even $V_K(\cdot)=\mathcal{O}(\cdot,0):L^2\to L^2$ has dense range.

To see how this plays out in the narrative developed here, recall from Section 2.7 that any admissible input space $\mathscr{U}$ contains all the polynomials, hence $W_0^{(k)}(\mathscr{U})=V^k(\mathscr{U})$ contains the linear span of the monomials t^j for $j\geqslant k$, and $W^{(k)}(\mathscr{U})$ again contains all the polynomials. Thus Theorems 5 and 4, respectively, tell us that in $\mathscr{U}$: the closure of $\mathcal{O}(\mathscr{U}\times\mathbb{C}^n)$ contains the closure of the polynomials, and the closure of $\mathcal{O}(\mathscr{U}\times\{0\})=V_K(\mathscr{U})$ contains the linear span of the monomials t^j for $j\geqslant r$.

Thus $\mathcal{O}(\mathscr{U}\times\mathbb{C}^n)$ will be dense in $\mathscr{U}$ whenever the polynomials are dense in $\mathscr{U}$; in particular this is true for $\mathscr{U}=C$ and $\mathscr{U}=L^p$ with $1\leqslant p<\infty$. For $\mathscr{U}=L^\infty$ the range of the output operator is (norm) dense in C rather than in L^∞, but density in L^∞ can be restored by taking that space in its weak-star topology.

Similarly, the range of V_K, viewed now as an operator on $\mathscr{U}$, contains the closure in $\mathscr{U}$ of the subspace spanned by the monomials of degree larger than r. By the Muntz–Szasz theorem (see [7, §15.25, pp. 312–315], for example) this subspace is uniformly dense in C_0, the continuous functions on $[0,T]$ that vanish at the origin. Thus, for example, regardless of the relative degree of the system, *$V_K(C)$ is dense in C_0*,[1] and since C_0 is dense in L^p for $1\leqslant p<\infty$ we see that for this range of p, *$V_K(L^p)$ is dense in L^p*.

Acknowledgments

I wish to thank Professor Gerardo Lafferriere of Portland State University for many interesting conversations and constructive suggestions regarding the material of this paper. I also wish to thank Professor Paul Bourdon of Washington and Lee University, for advice on earlier drafts.

[1] This rectifies a minor oversight in [1, Theorems 3.8 and 3.9], where it is erroneously asserted that C is, for general r, the direct sum of the closure of $V_K(C)$ with the polynomials of degree $\leqslant r$.

References

[1] Ulf Jönsson, Clyde Martin, Approximation with the output of linear control systems, J. Math. Anal. Appl. 329 (2007) 798–821.
[2] Paul R. Halmos, A Hilbert Space Problem Book, second ed., Springer-Verlag, 1982.
[3] Rudolph E. Kalman, On the general theory of control systems, in: Proc. of the First IFAC Congress, 1, Butterworths, London, 1960, pp. 481–491.
[4] David G. Luenberger, Introduction to Dynamic Systems, John Wiley & Sons, 1979.
[5] Peter Meyer-Nieberg, Banach Lattices, Springer-Verlag, 1991.
[6] M.M. Rao, Z.D. Ren, Theory of Orlicz Spaces, Marcel Dekker, Inc., 1991.
[7] Walter Rudin, Real and Complex Analysis, third ed., McGraw–Hill, 1987.
[8] William J. Terrell, Some fundamental control theory I: Controllability, observability, and duality, Amer. Math. Monthly 106 (1999) 705–719.
[9] William J. Terrell, Some fundamental control theory II: Feedback linearization of single input nonlinear systems, Amer. Math. Monthly 106 (1999) 812–828.