Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Linear Algebra and its Applications 424 (2007) 540–554
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

When are linear differentiation-invariant spaces differential?

Vakhtang Lomadze

Institute of Mathematics, Tbilisi 0193, Georgia

Received 15 December 2005; accepted 21 June 2006
Available online 12 March 2007
Submitted by P. Fuhrmann

Abstract

It is shown that a linear differentiation-invariant subspace of a C^{∞}-trajectory space is differential (i.e., can be represented as the kernel of a linear constant-coefficient differential operator) if and only if its McMillan degree is finite.

Keywords: Linear system; Transfer function; Convolution function; Zero initial condition

1. Introduction

Let k be the field of real or complex numbers, s an indeterminate, $\mathcal{U}$ the space of all infinitely differentiable k-valued functions of the nonnegative real variable, and let q be a fixed positive integer.

The paper is concerned with the following question: When a linear differentiation-invariant subspace of $\mathcal{U}^q$ can be described via an equation of the form $R(\partial)w = 0$, where R is a polynomial matrix (with q columns) and ∂ is the differentiation operator? This natural question was posed by Willems (see [7,8]), and we try here to give a brief answer to it.

Let O be the ring of proper rational functions (in s), and let t denote the "uniformizer" s^{-1}. The space $\mathcal{U}$ has a natural O-module structure: Given $g \in O$ and $\xi \in \mathcal{U}$, we define

$$g\xi = \sum_{n \geqslant 0} b_n \int^n \xi,$$

E-mail address: loma@rmi.acnet.ge

0024-3795/$ - see front matter
doi:10.1016/j.laa.2006.06.031

where b_n are the coefficients in the expansion of g at infinity and $\int^n$ stands for the n-fold iteration of the integration operator with itself. The series converges uniformly on $[0, X]$ for each $X > 0$. Indeed, we can find $r > 0$ so that $\sum |b_n| r^n = B < +\infty$, and consequently $|b_n| < Br^{-n}$ for all $n \geqslant 0$. Letting now $M = \sup_{0 \leqslant x \leqslant X} |\xi(x)|$, we have

$$\forall x \in [0, X], \quad \sum_{n \geqslant 0} |b_n| \left| \left(\int^n \xi \right)(x) \right|$$
$$= |b_0||\xi(x)| + \sum_{n \geqslant 1} |b_n| \left| \int_0^x \frac{(x-u)^{n-1}}{(n-1)!} \xi(u)\,\mathrm{d}u \right|$$
$$\leqslant |b_0|M + \sum_{n \geqslant 1} |b_n| M \int_0^x \frac{(x-u)^{n-1}}{(n-1)!}\,\mathrm{d}u$$
$$\leqslant \sum_{n \geqslant 0} |b_n| M \frac{x^n}{n!} \leqslant \sum_{n \geqslant 0} Br^{-n} M \frac{x^n}{n!} \leqslant BM\exp(X/r).$$

It is remarkable that $\mathscr{U}$ is torsion free. (This immediately follows from the fact that the integration operator is injective and the fact that every proper rational function is represented as $t^n u$ with $n \geqslant 0$ and invertible $u \in O$.) Let $L : k \mapsto \mathscr{U}$ be the canonical map embedding numbers into constant functions. For $g \in O$, we define the (inverse) Laplace transform $L(g)$ to be the function $gL(1)$, i.e., the analytic function

$$x \mapsto \sum_{n \geqslant 0} b_n \frac{x^n}{n!} \quad (x \geqslant 0),$$

where b_n are as above. The functions $L(g)$ will be called exponential functions. (In the case $k = \mathbb{C}$ these are precisely finite linear combinations of functions $x^n e^{\lambda x}$, where $n \in \mathbb{Z}_+$ and $\lambda \in \mathbb{C}$.)

Define a transfer function as a submodule $T \subseteq O^q$ such that O^q/T is torsion free, i.e., a subset of the form GO^m, where m is a nonnegative integer and G is a left invertible proper rational matrix of size $q \times m$. This notion is a natural generalization of the classical notion of transfer function. (Indeed, up to componentwise partition $k^q \simeq k^m \oplus k^p$, a transfer function is the graph of a classical transfer function $u \mapsto Au (u \in O^m)$, where A is a proper rational matrix of size $p \times m$.) A submodule $T \subseteq O^q$ gives rise to a submodule $T\mathscr{U} \subseteq \mathscr{U}^q$ consisting of all finite sums of trajectories of the form $g\xi$ $(g \in T, \xi \in \mathscr{U})$. Notice that if G is a generating matrix of T, then $T\mathscr{U} = G\mathscr{U}^m$, where m is the column number of G. It is interesting to note that the correspondence $T \mapsto T\mathscr{U}$ is one-to-one. We think of the distinguished modules $T\mathscr{U}$ as zero initial condition trajectory modules (ZICTMs).

It can be shown without difficulty that if $\mathscr{S}$ is a linear differentiation-invariant subspace of $\mathscr{U}^q$, then the set

$$T = \{g \in O^q | g\mathscr{U} \subseteq \mathscr{S}\}$$

is a transfer function. We call it the transfer function of $\mathscr{S}$, and we regard trajectories in $T\mathscr{U}$ as zero initial condition trajectories of $\mathscr{S}$. We define the McMillan degree of $\mathscr{S}$ as its dimension modulo $T\mathscr{U}$, i.e., the dimension of $\mathscr{S}/T\mathscr{U}$. The space $\mathscr{S}/T\mathscr{U}$ itself is called the initial condition (or state) space. We define a linear system to be a linear differentiation-invariant subspace with finite McMillan degree.

Not surprisingly, the kernel of a linear constant-coefficient differential operator is a linear system. The main result of this paper (namely, Theorem 3) states that the converse also is true. To prove this result we consider a canonical k-linear bilinear form $k[s]^q \times \mathscr{U}^q \to k$ defined by the formula

$$\langle f, \xi\rangle = (f^{\mathrm{tr}}(\partial)\xi)(0).$$

("tr" stands for the transpose.) If $\mathscr{S}$ is a linear system, then clearly $\mathscr{S}^\perp$ is a submodule of $k[s]^q$. It is trivial that every submodule has an "image representation", and letting E be such a representation of $\mathscr{S}^\perp$, the idea is that a "kernel representation" of $\mathscr{S}$ should be $R = E^{\mathrm{tr}}$. In deriving the result helpful roles will be played by the "Riemann–Roch formula" and the "key lemma" (Lemma 8). The key lemma gives a duality relation between transfer functions and, what we call, convolution functions. (Convolution functions are certain linear subspaces of $k[s]^q$, which play in the paper just an auxiliary role; they are connected with submodules as ZICTMs are connected with linear systems.) This immediately leads to a relation between ZICTMs and convolution functions. We apply the Riemann–Roch formula to compute some dimensions. This computation allows then to extend the relation above to a one between linear systems and submodules of $k[s]^q$.

Concluding the introduction, it seems worthwhile to point out that the paper is self-contained.

2. Mikusinski functions

We let $\mathscr{M}$ be the fraction space of $\mathscr{U}$. Elements of $\mathscr{M}$ are called Mikusinski (or generalized) functions. Every Mikusinski function can be written as a ratio ξ/t^n, where $\xi \in \mathscr{U}$ and $n \geqslant 0$. (This is because every $\neq 0$ element in O, as already remarked, is a power of t modulo invertible elements.) Of course $t^n \cdot \xi/t^n = \xi$, and this means that every generalized function is a quantity that after "integrating" sufficiently many times becomes an ordinary function.

Remark. It is Mikusinski's idea to define generalized functions as ratios (see [5]). This is a nice idea.

We identify $\mathscr{U}$ with its image in $\mathscr{M}$ under the canonical map $\xi \mapsto \xi/1$. It is obvious that

$$\mathscr{U} \subset s\mathscr{U} \subset s^2\mathscr{U} \subset \cdots \quad \text{and} \quad \mathscr{M} = \cup s^n\mathscr{U}.$$

The homomorphism L can be uniquely continued to a $k(s)$-linear map $k(s) \to \mathscr{M}$, and we shall use the same letter L to denote it. We call elements of $L(sk[s])$ purely impulsive functions.

The Newton–Leibniz formula can be rewritten as $s\xi = \xi' + s\xi(0)$. Using induction argument, one easily deduces the Taylor formula

$$s^n\xi = \xi^{(n)} + s^n\xi(0) + \cdots + s\xi^{(n-1)}(0).$$

The following says that every Mikusinski function has the "regular" part and the purely impulsive part.

Lemma 1. $\mathscr{M} = \mathscr{U} \oplus L(sk[s])$.

Proof. Follows from Taylor's formula. □

We shall need the following

Lemma 2. *Let R be a polynomial matrix of size $p \times q$. Then*

$$\mathrm{Ker}\, R(\partial) = \{\xi \in \mathscr{U}^q \,|\, R\xi \in L(sk[s]^p)\}.$$

Proof. Let $R = R_0 s^n + R_1 s^{n-1} + \cdots + R_n$, and let $\xi \in \mathscr{U}^q$. Using Taylor's formula, we have

$$R\xi = R(\partial)\xi + [sI_p \dots s^n I_p]\begin{bmatrix} R_{n-1} & R_{n-2} & \dots & R_0 \\ R_{n-2} & R_{n-3} & & 0 \\ \vdots & & & \vdots \\ R_0 & 0 & \dots & 0 \end{bmatrix}\begin{bmatrix} \xi(0) \\ \xi'(0) \\ \vdots \\ \xi^{(n-1)}(0) \end{bmatrix}.$$

We see that $R(\partial)\xi$ is equal to the regular part of $R\xi$, and the lemma follows. □

The following two elementary examples illustrate how Mikusinski functions work.

Example 1. Let $r = a_0 s^n + a_1 s^{n-1} + \cdots + a_n$ be a polynomial with $a_0 \neq 0$, and let $x_0, \dots, x_{n-1} \in k$. Consider the Cauchy problem

$$\begin{cases} a_0 x^{(n)} + a_1 x^{(n-1)} + \cdots + a_n x = 0; \\ x(0) = x_0, \dots, x^{(n-1)}(0) = x_{n-1}. \end{cases}$$

Applying the Taylor formula, we can rewrite this as

$$rx = L(f),$$

where f is a polynomial given by the formula

$$f = [s \dots s^n]\begin{bmatrix} a_{n-1} & a_{n-2} & \dots & a_0 \\ a_{n-2} & a_{n-3} & & 0 \\ \vdots & & & \vdots \\ a_0 & 0 & \dots & 0 \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}.$$

Multiplying both sides of this equation by $1/r$, we obtain

$$x = \frac{1}{r}L(f) = L\left(\frac{f}{r}\right).$$

Notice that f/r is a proper rational function, and so the solution is an exponential function (as it should be of course).

Example 2. Let r be as in the previous example, and let $\xi \in \mathscr{U}$. Consider the Cauchy problem

$$\begin{cases} a_0 x^{(n)} + a_1 x^{(n-1)} + \cdots + a_n x = \xi; \\ x(0) = 0, \dots, x^{(n-1)}(0) = 0. \end{cases}$$

Applying the Taylor formula, we can rewrite this as

$$rx = \xi.$$

Multiplying both sides of this equation by $1/r$, we obtain

$$x = \frac{1}{r}\xi = \frac{t^n}{a_0 + a_1 t + \cdots + a_n t^n}\xi.$$

3. Algebraic preliminaries

Let D be a nonsingular rational matrix of size p. The number $-\mathrm{ord}_\infty(\det D)$ is called the Chern number of D and is denoted by $\mathrm{ch}(D)$. (We remind that the order at infinity of a rational function u/v with $u, v \in k[s]$ and $v \neq 0$ is defined to be $\deg(v) - \deg(u)$.) We define the dual of D as $D^* = (D^{-1})^{\mathrm{tr}}$. The cohomology spaces are defined as

$$H^0(D) = sk[s]^p \cap DO^p \quad \text{and} \quad H^1(D) = k(s)^p/(k[s]^p + tDO^p).$$

One can easily compute that

$$\dim H^0(s^n I_p) = \max\{np, 0\} \quad \text{and} \quad \dim H^1(t^n I_p) = \max\{np, 0\},$$

where n is an arbitrary integer. It immediately follows from these formulas that the spaces $H^0(D)$ and $H^1(D)$ have finite dimension. (Indeed, for sufficiently large n, $DO^p \subseteq s^n O^p$ and $t^n O^p \subseteq DO^p$. Hence, $H^0(D) \subseteq H^0(s^n I_p)$ and there is a surjective linear map $H^1(t^n I_p) \to H^1(D)$.)

We shall need the following nice formula ("Riemann–Roch formula")

$$\mathrm{ch}(D) = \dim H^0(D) - \dim H^1(D).$$

To prove it, choose $n \geqslant 0$ so large that $DO^p \subseteq s^n O^p$, and consider the diagram

$$\begin{array}{ccccccccc} 0 & \to & sk[s]^p \oplus DO^p & \to & sk[s]^p \oplus s^n O^p & \to & s^n O^p/DO^p & \to & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \to & k(s)^p & = & k(s)^p & \to & 0 & \to & 0 \end{array}$$

The diagram commutes and has exact rows. Applying the snake lemma (see, for example, Proposition 2.10 in [1]) and the facts that

$$k(s)^p/(sk[s]^p + s^n O^p) = \{0\} \quad \text{and} \quad k(s)^p/(sk[s]^p + DO^p) \simeq H^1(D),$$

we get an exact sequence

$$0 \to H^0(D) \to H^0(s^n I_p) \to s^n O^p/DO^p \to H^1(D) \to 0.$$

The space $s^n O^p/DO^p \simeq O^p/t^n DO^p$ has dimension equal to $\mathrm{ord}_\infty(t^n \det D) = np - \mathrm{ch}(D)$, and the formula follows.

Let us say that two nonsingular rational matrices D_1 and D_2 are similar if there exists a biproper matrix B such that $D_2 = D_1 B$. Notice that if this is the case, then D_1 and D_2 have the same Chern number and the same cohomologies.

Remark. There is a close link between similarity classes of nonsingular rational matrices and vector bundles over the projective line (see [4]), and this explains the terminology above.

If X and Y are k-linear spaces such $X \subseteq Y$, we write $[Y : X]$ to denote the codimension of X in Y.

Lemma 3. *Let V be a $k(s)$-linear space of finite dimension, and let M and N be submodules in V over $k[s]$ and O, respectively. The following conditions are equivalent:*

(a) *M and N have full rank;*
(b) *$[V : (M + N)]$ is finite.*

Proof. Let r denote the dimension of V.

(a) ⇒ (b) Take an isomorphism $\phi : V \to k(s)^r$ so that $\phi(M) = k[s]^r$. Then $\phi(N) = DO^r$ for some nonsingular rational matrix D, and therefore $V/(M + N) \simeq H^1(D)$.

(a) ⇐ (b) Say that M is not of full rank. Let i denote its rank and put $j = r - i$. Take an isomorphism $\phi : V \to k(s)^r$ so that $\phi(M) = k[s]^i \oplus 0$ $(\subset k[s]^i \oplus k[s]^j)$ and choose $n \geqslant 1$ so that $\phi(N) \subseteq s^n O^r$. We then have a surjective linear map

$$V/(M+N) \to k(s)^r/((k[s]^i \oplus 0) + s^n O^r).$$

It remains now to notice that

$$k(s)^r/((k[s]^i \oplus 0) + s^n O^r) = k(s)^i/(k[s]^i + s^n O^i) \oplus k(s)^j/s^n O^j = k(s)^j/s^n O^j$$

has infinite dimension. □

Lemma 4. *Let R be a full row rank polynomial matrix of size $p \times q$. Then there exists a nonsingular rational matrix D satisfying the following equivalent conditions*:

(a) *$D^{-1}R$ is a right invertible proper rational matrix*;
(b) *$RO^q = DO^p$.*

The matrix D is uniquely determined up to similarity.

Proof. Clearly RO^q is a full rank O-submodule in $k(s)^p$. Hence, $RO^q = DO^p$ for some nonsingular rational matrix D. It is obvious that saying that $D^{-1}R$ is a right invertible proper rational matrix is equivalent to saying that $D^{-1}RO^q = O^p$, i.e., $RO^q = DO^p$.

Assume that D_1 and D_2 satisfy the condition. Then $D_1O^p = D_2O^p$, and therefore $O^p = D_1^{-1}D_2O^p$. It follows that $D_1^{-1}D_2$ is biproper. □

Lemma 5. *Let E be a full column rank polynomial matrix of size $q \times p$. Then there exists a nonsingular rational matrix D satisfying the following equivalent conditions*:

(a) *ED is a left invertible proper rational matrix*;
(b) *$EDO^p = Ek(s)^p \cap O^q$.*

The matrix D is uniquely determined up to similarity.

Proof. This can be deduced easily from the previous lemma. (A direct proof is possible, and we leave it to the interested reader.) □

4. Convolution and transfer functions

Given a rational subspace $V \subseteq k(s)^q$, we shall write V_- to denote the set of the polynomial parts of all elements in V.

Lemma 6. *Let $M \subseteq k[s]^q$ be a submodule and $V \subseteq k(s)^q$ a rational subspace such that $M \subseteq V$. The following conditions are equivalent*:

(a) *V is the fraction space of M*;
(b) *$[V_- : M]$ is finite.*

Proof. Consider the canonical map $V \to V_-/M$, which certainly is surjective. Its kernel is equal to $M + (V \cap tO^q)$. Indeed, assume that $x + ty \in V$, where $x \in k[s]^q$ and $y \in O^q$, goes to zero. Then we must have $x \in M$. Because $M \subseteq V$, we also must have $y \in V$, and so $x + ty \in M + (V \cap tO^q)$. Thus, we have a canonical isomorphism

$$V/(M + (V \cap tO^q)) \simeq V_-/M.$$

Using Lemma 3, we complete the proof. □

Any subset $C \subseteq k[s]^q$ of the form $C = V_-$, where V is a $k(s)$-linear subspace of $k(s)^q$, will be referred to as a convolution function. (It can be shown easily, using the previous lemma, that V_- is uniquely determined by V.) The convolution function of a submodule $M \subseteq k[s]^q$ is defined to be V_-, where V is the fraction space of M. By the lemma above, $[V_- : M] < +\infty$. The following says that this property uniquely characterizes the convolution function of a module.

Corollary 1. *If M is a submodule and C a convolution function such that $M \subseteq C$ and $[C : M] < +\infty$, then necessarily C is the convolution function of M.*

Proof. Let V be the fraction space of M, and let W be a rational subspace such that $C = W_-$. Then

$$W_-/M \oplus V_-/M \to (W + V)_-/M$$

clearly is surjective, and consequently $[(W + V)_- : M] < +\infty$. Using now the previous lemma, we find that $W + V = V$. Hence, $W \subseteq V$. Because V is the least rational subspace containing M, we conclude that $W = V$. □

Lemma 7. *Let E be a full column rank polynomial matrix of size $q \times p$, and let D be a nonsingular rational matrix satisfying the conditions of Lemma 5. Letting $M = Ek[s]^p$ and $C = (Ek(s)^p)_-$, we then have*

$$[C : M] = -\mathrm{ch}(D).$$

Proof. The matrix E induces a canonical linear map $H^0(D) \to H^0(I_q)$, which must be injective because E has full column rank. It follows that $H^0(D) = 0$. Hence, by the Riemann–Roch formula, $\mathrm{ch}(D) = -\dim H^1(D)$. Further, there is (see the proof of Lemma 6) a canonical isomorphism

$$C/M \simeq Ek(s)^p/(M + Ek(s)^p \cap tO^q).$$

This completes the proof, because the right hand side is isomorphic to $H^1(D)$. □

We call a transfer function any subset $T \subseteq O^q$ of the form $T = V \cap O^q$, where V is a $k(s)$-linear subspace of $k(s)^q$. (This definition is equivalent to that given in Introduction.) The dimension of V is called the input number of T. It should be noted that the correspondence $V \mapsto V \cap O^q$ is one-to-one. (This is because V is equal to the fraction space of $V \cap O^q$). If T is a transfer function with input number m, then T can be written as $T = GO^m$, where G is left invertible proper rational matrix of size $q \times m$. If G_1 and G_2 are two generating matrices, then they are equivalent in the sense that $G_2 = G_1 B$ for some biproper rational matrix B.

Given a proper rational function g, we let $g(\infty)$ be its value at infinity and g^σ its backward shift. (If $g = b_0 + b_1 t + b_2 t^2 + \cdots$, then $g(\infty) = b_0$ and $g^\sigma = b_1 + b_2 t + \cdots$). Define a canonical k-bilinear form

$$k[s]^q \times O^q \to k, \quad \langle f, g\rangle = (f^{\mathrm{tr}}(\sigma)g)(\infty), \tag{1}$$

which clearly is nondegenerate. For a k-linear subspace X in $k[s]^q$ or O^q, we let $X^\perp$ denote the orthogonal of X with respect to this bilinear form.

Given a $k(s)$-linear subspace $V \subseteq k(s)^q$, we set

$$V^\circ = \{f \in k(s)^q \mid f^{\mathrm{tr}} g = 0 \ \ \forall g \in V\}.$$

Obviously V° also is a $k(s)$-linear subspace, and $V^{\circ\circ} = V$. The following lemma, which relates convolution and transfer functions to each other, will play a key role. (For convenience, we postpone its proof to Appendix A.)

Lemma 8 (Key lemma). *Let V be a $k(s)$-linear subspace in $k(s)^q$. Then*

$$(V \cap O^q)^\perp = (V^\circ)_- \quad \textit{and} \quad (V_-)^\perp = V^\circ \cap O^q.$$

Corollary 2. *If C is a convolution function, then $C^{\perp\perp} = C$; likewise, if T is a transfer function, then $T^{\perp\perp} = T$.*

5. Linear systems

Given a transfer function T, let $T\mathscr{U}$ denote the submodule of $\mathscr{U}^q$ generated by all columns of the form $g\xi$, where $g \in T$ and $\xi \in \mathscr{U}$. Remark that if G is a generating matrix of T, then $T\mathscr{U} = G\mathscr{U}^m$; in other words, letting $g_1, \ldots, g_m$ denote the columns of G, then every element $\xi \in T\mathscr{U}$ can be (uniquely) written as

$$\xi = g_1\xi_1 + \cdots + g_m\xi_m$$

with $\xi_1, \ldots, \xi_m \in \mathscr{U}$. We remark also that $T\mathscr{U}$ is the image under the canonical homomorphism $T \otimes \mathscr{U} \to O^q \otimes \mathscr{U} = \mathscr{U}^q$.

It is interesting to note that the correspondence $T \mapsto T\mathscr{U}$ is one-to-one. Indeed, let T be a transfer function and let $\{g_1, \ldots, g_m\}$ be its basis. Because O^q/T is torsion free (and therefore free), we can find $h_1, \ldots, h_p \in O^q$ such that $\{g_1, \ldots, g_m, h_1, \ldots, h_p\}$ is a basis of O^q. Any element of $\mathscr{U}^q$ is uniquely represented then as

$$g_1\xi_1 + \cdots + g_m\xi_m + h_1\zeta_1 + \cdots + h_p\zeta_p.$$

This belongs to $T\mathscr{U}$ if and only if

$$\zeta_1, \ldots, \zeta_p = 0,$$

and belongs to $L(O^q)$ if and only if

$$\xi_1, \ldots, \xi_m, \zeta_1, \ldots, \zeta_p \in L(O).$$

We see that $T\mathscr{U} \cap L(O^q) = L(T)$, and hence

$$T = L^{-1}(T\mathscr{U} \cap L(O^q)).$$

Proposition 1. *Let $\mathscr{S}$ be a linear subspace in $\mathscr{U}^q$ that is invariant with respect to the differentiation operator. Then the set*

$$T = \{g \in O^q \mid g\mathscr{U} \subseteq \mathscr{S}\}$$

is a transfer function (called the transfer function of $\mathscr{S}$).

Proof. Obviously, T is a submodule (in O^q). Choose any its generating matrix G, and assume that it is not left invertible. Then the scalar matrix $\bar{G}$ is not of full column rank. (The bar here denotes the canonical homomorphism from O to $k = O/tO$.) This means that the columns $g_1, \dots, g_m$ of G are linearly dependent modulo tO^q. Say that

$$g_m \equiv a_1 g_1 + \cdots + a_{m-1} g_{m-1} \pmod{tO^q},$$

where $a_1, \dots, a_{m-1} \in k$. Then there exists a column $h \in O^q$ such that

$$g_m = a_1 g_1 + \cdots + a_{m-1} g_{m-1} + th.$$

Certainly $h \neq 0$. We claim that $h \in T$. Indeed, let ξ be an arbitrary function. Then $th\xi \in \mathscr{S}$ (because $th \in T$). Using the invariance property of $\mathscr{S}$, we have $h\xi = (th\xi)' \in \mathscr{S}$. The claim is proved.

The columns $g_1, \dots, g_{m-1}, h$ generate T, and they must form a basis (since their number is m). But $\mathrm{diag}(1, \dots, 1, t)$ is not biproper, and therefore $\{g_1, \dots, g_{m-1}, g_m\}$ can not be a basis. The contradiction shows that T must be a transfer function. □

Given a linear differentiation-invariant subspace $\mathscr{S}$ with transfer function T, we call $\mathscr{S}/T\mathscr{U}$ the initial condition space of $\mathscr{S}$. If ξ is a trajectory in $\mathscr{S}$, then its image in $\mathscr{S}/T\mathscr{U}$ is called the initial condition of ξ. The cardinality $[\mathscr{S} : T\mathscr{U}]$ is called the McMillan degree. We shall see in the next section that the solution sets of linear constant-coefficient differential equations have finite McMillan degree. The following examples show that, in general, the McMillan degree is not finite.

Example 3. The space $\mathscr{S} = k[x]^q$, i.e., the space of all polynomial trajectories, clearly is differentiation-invariant. Obviously,

$$L^{-1}(\mathscr{S}) = k[t]^q.$$

It is clear that the transfer function is $\{0\}$, and so the space has infinite McMillan degree.

Example 4. Let $n \geqslant 0$, and let $\mathscr{S} = \{\xi \in \mathscr{U}^q | \forall i \geqslant n,\ \xi^{(i)}(0) = 0\}$. Clearly $\mathscr{S}$ is differentiation-invariant. We have

$$L^{-1}(\mathscr{S} \cap L(O^q)) = \{f \in k[t]^q | \deg f \leqslant n-1\}.$$

The only transfer function contained in the above set is $\{0\}$, and so the transfer function of our space is $\{0\}$. It follows that the McMillan degree is infinite.

Lemma 9. *Let $\mathscr{S}$ be a linear subspace in $\mathscr{U}^q$. There may exist only one transfer function T such that*

$$T\mathscr{U} \subseteq \mathscr{S} \quad \text{and} \quad [\mathscr{S} : T\mathscr{U}] < +\infty.$$

Proof. Suppose that there are two such transfer function T_1 and T_2, and put $T = T_1 + T_2$. (Notice that T may not be a transfer function, but $T\mathscr{U}$ still is defined.) Clearly, we have $[T\mathscr{U} : T_i\mathscr{U}] < +\infty$. From this and from the exact sequence

$$0 \to T_i\mathscr{U} \to T\mathscr{U} \to T/T_i \otimes \mathscr{U} \to 0,$$

which is obtained by tensoring the exact sequence $0 \to T_i \to T \to T/T_i \to 0$ with $\mathscr{U}$, it follows that $T/T_i \otimes \mathscr{U}$ has finite dimension. We see that T/T_i must be a torsion module, and hence T_i has the same fraction space as T. We conclude that each T_i is equal to $V \cap O^q$, where V is the fraction space of T. □

By a linear (dynamical) system we shall understand a linear differentiation-invariant subspace of $\mathscr{U}^q$ that has finite McMillan degree.

Proposition 2. *Let $\mathscr{S}$ be a linear system with transfer function T. Then*

$$\mathscr{S} \subseteq T\mathscr{U} + L(O^q);$$

in other words, there always exists in $\mathscr{S}$ an exponential trajectory with a given initial condition.

Proof. Take any $\xi \in \mathscr{S}$. Modulo $T\mathscr{U}$ the trajectories $\xi, \xi', \xi'', \dots$ are linearly dependent. It follows that there exist an integer $n \geqslant 1$ and elements $a_1, \dots, a_n \in k$ such that

$$\xi^{(n)} + a_1\xi^{(n-1)} + \cdots + a_n\xi \equiv T\mathscr{U}.$$

This means that our trajectory ξ satisfies the differential equation

$$x^{(n)} + a_1 x^{(n-1)} + \cdots + a_n x = \xi_0$$

with $\xi_0 \in T\mathscr{U}$. In view of Example 2, a particular solution of this equation is

$$t^n(1 + a_1 t + \cdots + a_n t^n)^{-1}\xi_0,$$

which certainly belongs to $T\mathscr{U}$. Further, in view of Example 1, ξ differs from this particular solution by an exponential trajectory. The proof is complete. □

6. Linear differential operators

Let R be a full row rank polynomial matrix of size $p \times q$. A nonsingular matrix D satisfying the conditions of Lemma 4 is called a denominator of R. The module $T = \{w \in O^q | Rw = 0\}$ is called the transfer function; the space $X = sk[s]^p \cap RO^q = H^0(D)$ is called the initial condition (or state) space; the Chern number of D is called the McMillan degree. It is easily seen that the McMillan degree is equal to the dimension of the state space. Indeed, the matrix R induces a canonical linear map $H^1(I_q) \to H^1(D)$, which must be surjective, because R has full row rank. Hence, $H^1(D) = 0$, and the statement follows from the Riemann–Roch formula.

Remark. The above concept of states is, in principle, the same as Fuhrmann's classical one [2]. Indeed, with notation of [2], we have $H^0(D) = sS_D$.

Example 5. Assume that $q = 1$, and let r be as in Example 1. The initial condition space of r is

$$X = sk[s] \cap rO = sk[s] \cap s^n O = \oplus_{1 \leqslant i \leqslant n} ks^i.$$

On the other hand, according to the textbooks, the initial condition space of the equation $r(\partial)w = 0$ is k^n. The two definitions agree with each other; namely, there is a canonical isomorphism $k^n \simeq X$ given by

$$x \mapsto [s \ \cdots \ s^n]Ax,$$

where A denotes the triangle matrix from Example 1.

The linear subspace Ker $R(\partial)$ is easily seen to be differentiation-invariant.

Theorem 1. *The McMillan degree of* Ker $R(\partial)$ *is finite* (*and is equal to that of* R).

Proof. Put $\mathscr{S} = \operatorname{Ker} R(\partial)$. According to Lemma 2,

$$\mathscr{S} = \{\xi \in \mathscr{U}^q \mid R\xi \in L(sk[s]^p)\}.$$

Consider the canonical linear map $\mathscr{S} \to L(sk[s]^p)$ (which is determined by the homomorphism $R : \mathscr{U}^q \to \mathscr{M}^p$). The image of this map is equal to

$$\begin{aligned} L(sk[s]^p) \cap R\mathscr{U}^q &= L(sk[s]^p) \cap D\mathscr{U}^p = L(sk[s]^p) \cap L(k(s)^p) \cap D\mathscr{U}^p \\ &= L(sk[s]^p) \cap D(Lk(s)^p \cap \mathscr{U}^p) = L(sk[s]^p) \cap DL(O^p) = L(X). \end{aligned}$$

So, we have a canonical surjective linear map $\mathscr{S} \to X$.

Consider now the exact sequence

$$0 \to T \to O^q \xrightarrow{R} DO^p \to 0.$$

The module $\mathscr{U}$ is torsion free (and hence flat). Therefore tensoring this sequence by $\mathscr{U}$, we get an exact sequence

$$0 \to T \otimes \mathscr{U} \to \mathscr{U}^q \xrightarrow{R} D\mathscr{U}^p \to 0.$$

Replacing $T \otimes \mathscr{U}$ by $T\mathscr{U}$, we obtain an exact sequence

$$0 \to T\mathscr{U} \to \mathscr{U}^q \xrightarrow{R} D\mathscr{U}^p \to 0.$$

This immediately implies that the kernel of the canonical map $\mathscr{S} \to X$ is equal to $T\mathscr{U}$, and consequently we have an exact sequence

$$0 \to T\mathscr{U} \to \mathscr{S} \to X \to 0.$$

This shows immediately that the transfer function of $\mathscr{S}$ is the same as that of R. This shows also that the initial condition space of $\mathscr{S}$ is canonically isomorphic to that of R. □

We shall need the following:

Lemma 10. *Let $R = E^{\mathrm{tr}}$, where E is a full column rank polynomial matrix of size $q \times p$. Then the transfer function of R is equal to $C^{\perp}$, where C is the convolution function of E.*

Proof. Given $u \in k(s)^p$ and $v \in k(s)^q$, we have $(Eu)^{\mathrm{tr}} v = u^{\mathrm{tr}} R v$. From this evident formula it immediately follows that

$$(Ek(s)^p)^\circ = \{v \in k(s)^q \mid Rv = 0\}.$$

Applying now the key lemma, one completes the proof. □

7. Main theorems

We have a canonical k-bilinear form

$$k[s]^q \times \mathscr{U}^q \to k, \quad \langle f, \xi \rangle = (f^{\mathrm{tr}}(\partial)\xi)(0). \tag{2}$$

This clearly is nondegenerate from the left (but not from the right of course). This bilinear form is related with the one defined in Section 4: If $f \in k[s]^q$ and $g \in O^q$, then $\langle f, g \rangle = \langle f, L(g) \rangle$.

Given a k-linear subspace $\mathscr{X} \subseteq \mathscr{U}^q$, we shall write $\mathscr{X}^\perp$ to denote the orthogonal of $\mathscr{X}$. (We believe that $\mathscr{X}^\perp$ can not be confused with $X^\perp$ defined earlier.)

Lemma 11. *If T is a transfer function, then $(T\mathscr{U})^\perp = T^\perp$.*

Proof. "$\supseteq$" By definition, $T = V \cap O^q$ for some $k(s)$-linear subspace $V \subseteq k(s)^q$. Let $f \in T^\perp$, and let $g \in T$ and $\xi \in \mathscr{U}$. By the key lemma, $f \in (V^\circ)_-$, and consequently $f + th \in V^\circ$ for some $h \in O^q$. We then have $(f + th)^{\mathrm{tr}} g = 0$, and therefore $(f + th)^{\mathrm{tr}} g\xi = 0$. By Lemma 2, $f^{\mathrm{tr}}(\partial)(g\xi)$ is equal to the regular part of $f^{\mathrm{tr}} g\xi$. But the latter is already regular, since it is equal to $-th^{\mathrm{tr}} g\xi$. We see that $f^{\mathrm{tr}}(\partial)(g\xi) \in t\mathscr{U}$, and so $\langle f, g\xi\rangle = 0$. Because $T\mathscr{U}$ is generated by elements of the form $g\xi$, we conclude that $f \in (T\mathscr{U})^\perp$.

"$\subseteq$" Because $T\mathscr{U} \supseteq L(T)$, we have $(T\mathscr{U})^\perp \subseteq L(T)^\perp$. Clearly $L(T)^\perp = T^\perp$, and thus $(T\mathscr{U})^\perp \subseteq T^\perp$. $\square$

Lemma 12. *Let $\mathscr{S}$ be a linear system with transfer function T. Then $\mathscr{S}^\perp$ is a submodule (in $k[s]^q$) with convolution function $T^\perp$, and the canonical bilinear form*

$$T^\perp/\mathscr{S}^\perp \times \mathscr{S}/T\mathscr{U} \to k,$$

is nondegenerate.

Proof. That $\mathscr{S}^\perp$ is a submodule follows immediately from the relationship $\langle sf, \xi\rangle = \langle f, \xi'\rangle$ (and the invariance property of $\mathscr{S}$). It is easily seen that the bilinear form is nondegenerate from the left, and therefore $T^\perp/\mathscr{S}^\perp$ is finite-dimensional. Using Corollary 1, it follows from this that $T^\perp$ is the convolution function of $\mathscr{S}^\perp$. To show that the form is nondegenerate from the right, take an arbitrary $\xi \in \mathscr{S}$ such that $\langle f, \xi\rangle = 0$ for each $f \in T^\perp$. Write $\xi = \xi_0 + L(w)$, where $\xi_0 \in T\mathscr{U}$ and $w \in O^q$. By the previous lemma, $\langle f, \xi_0\rangle = 0$ for each $f \in T^\perp$. It follows that

$$\forall f \in T^\perp, \quad \langle f, L(w)\rangle = 0.$$

Using the key lemma, we can see that $w \in T$. Hence, $\xi \in T\mathscr{U}$, and the proof is complete. $\square$

Two full row rank polynomial matrices R_1 and R_2 are said to be equivalent if there exists a unimodular matrix U such that $R_2 = UR_1$. The following is due to Schumacher [6].

Theorem 2. *Two full row rank polynomial matrices (with column number q) generate the same linear system if and only if they are equivalent.*

Proof. Let R be a full row rank polynomial matrix of size $p \times q$, and let T be its transfer function. Put $E = R^{\mathrm{tr}}$, $\mathscr{S} = \operatorname{Ker} R(\partial)$ and $M = Ek[s]^p$. We want to show that

$$\mathscr{S}^\perp = M.$$

Take $x \in M$. Then $x = Ef$ with $f \in k[s]^p$. For each $\xi \in \mathscr{S}$, we have

$$\langle x, \xi\rangle = \langle Ef, \xi\rangle = ((f^{\mathrm{tr}} R)(\partial)\xi)(0) = (f^{\mathrm{tr}}(\partial)R(\partial)\xi)(0) = \langle f, R(\partial)\xi\rangle = \langle f, 0\rangle = 0.$$

Hence $M \subseteq \mathscr{S}^\perp$. To see that in fact we have equality, consider the tower

$$M \subseteq \mathscr{S}^\perp \subseteq T^\perp.$$

Choose a denominator D of R. Then ED^* is a left invertible proper rational matrix. Using Lemma 12, Theorem 1 and Lemma 7, we get

$$[T^\perp : \mathcal{S}^\perp] = [\mathcal{S} : T\mathcal{U}] = \mathrm{ch}(D) = -\mathrm{ch}(D^*) = [T^\perp : M].$$

Therefore we indeed must have equality. The "only if" part follows because M is "representation free".

The "if" part is obvious. □

Theorem 3. *Every linear system is represented as the kernel of a linear constant-coefficient differential operator.*

Proof. Assume we have a linear system $\mathcal{S}$ with input number p. Let $M = \mathcal{S}^\perp$, and choose a full column rank polynomial matrix E such that $Ek[s]^p = M$ (minimal image representation of M). Put $R = E^{\mathrm{tr}}$. We are going to show that $\mathcal{S} = \mathrm{Ker}\, R(\partial)$.

Let T denote the transfer function of $\mathcal{S}$. By Lemma 12, $C = T^\perp$ is the convolution function of M. Thanks to the key lemma, $C^\perp = T$. It follows from Lemma 10 that T is the transfer function of R, and thus T is the transfer function of Ker $R(\partial)$ as well.

Take an arbitrary $\xi \in \mathcal{S}$, and write $\xi = \xi_0 + L(w)$ with $\xi_0 \in T\mathcal{U}$ and $w \in O^q$. Because $M \subseteq C$ and $C = (T\mathcal{U})^\perp$, $\langle x, \xi_0\rangle = 0$ for each $x \in M$. We therefore have

$$\forall x \in M, \quad \langle x, L(w)\rangle = \langle x, \xi_0\rangle + \langle x, L(w)\rangle = \langle x, \xi\rangle = 0.$$

In other words,

$$\forall f \in k[s]^p, \quad \langle Ef, L(w)\rangle = 0.$$

It follows that

$$\forall f \in k[s]^p, \quad \langle f, R(\partial)L(w)\rangle = \langle Ef, L(w)\rangle = 0.$$

Because $\langle f, R(\partial)L(w)\rangle = \langle f, R(\sigma)w\rangle$ and because (1) is nondegenerate, this implies $R(\partial)L(w) = 0$. Thus $L(w) \in \mathrm{Ker}\, R(\partial)$, and hence ξ belongs to Ker $R(\partial)$. We conclude that $\mathcal{S} \subseteq \mathrm{Ker}\, R(\partial)$.

The proof now is easily completed by dimension count. Indeed, consider the tower

$$T\mathcal{U} \subseteq \mathcal{S} \subseteq \mathrm{Ker}\, R(\partial).$$

By the proof of the previous theorem, Ker $R(\partial)^\perp = M$. Applying Lemma 12 both to $\mathcal{S}$ and Ker $R(\partial)$, we get

$$[\mathcal{S} : T\mathcal{U}] = [C : M] = [\mathrm{Ker}\, R(\partial) : T\mathcal{U}].$$

This yields $\mathcal{S} = \mathrm{Ker}\, R(\partial)$. □

Acknowledgments

The author would like to thank the referees for several valuable suggestions.

Appendix A: Proof of the key lemma

Let W_1 and W_2 be rational linear spaces of the same dimension, and assume that we are given a nondegenerate $k(s)$-bilinear form $W_1 \times W_2 \to k(s)$. There is a canonical k-linear map $k(s) \to k$ (determined by the decomposition $k(s) = sk[s] \oplus k \oplus tO$). Composing our form with this map we obtain a k-bilinear form $W_1 \times W_2 \to k$. Using the latter, for each subset $N \subseteq W_i$, define $N^{\perp}$. We state that if N is a finitely generated full rank O-submodule, say, in W_1, then $N^{\perp\perp} = N$. To see this, let us denote by p the dimension of our spaces and choose isomorphisms $\phi_1 : W_1 \simeq k(s)^p$ and $\phi_2 : W_2 \simeq k(s)^p$ so that $\phi_1(N) = O^p$ and the diagram

$$\begin{array}{ccccc} W_1 & \times & W_2 & \to & k(s) \\ \downarrow & & \downarrow & & \| \\ k(s)^p & \times & k(s)^p & \to & k(s) \end{array}$$

is commutative. (The bottom bilinear form is given by $(f, g) \mapsto f^{\mathrm{tr}} g$.) We are reduced therefore to the case when $W_1 = k(s)^p$, $W_2 = k(s)^p$ and $M = O^p$. One can check easily that in this standard case $(O^p)^{\perp} = tO^p$ and $(tO^p)^{\perp} = O^p$; hence $(O^p)^{\perp\perp} = O^p$.

We are able now to give:

Proof of the key lemma. The bilinear form (1) is extended to the canonical bilinear form $k(s)^q \times k(s)^q \to k$. We claim that with respect to this latter

$$(V \cap O^q)^{\perp} = V^{\circ} + tO^q.$$

Indeed, consider the canonical bilinear form

$$k(s)^q / V^{\circ} \times V \to k,$$

which obviously is nondegenerate, and put $N = (V^{\circ} + tO^q)/V^{\circ}$. The latter is a finitely generated full rank submodule in $k(s)^q/V^{\circ}$, and hence $N^{\perp\perp} = N$. It is easy to see that $N^{\perp} = V \cap O^q$, and so $(V \cap O^q)^{\perp} = N$. This implies our claim. Returning now to (1), we get

$$(V \cap O^q)^{\perp} = (V^{\circ} + tO^q) \cap k[s]^q.$$

The left hand side is just $(V^{\circ})_-$, and the first relation is proved.

The second relation is easy, and needs no preparation. Indeed, let $g \in O^q$ and let $f \in V$. We then have

$$\langle f, g \rangle = \langle f_-, g \rangle,$$

where f_- denotes the polynomial part of f. Hence, $g \in V^{\circ}$ if and only if $g \in (V_-)^{\perp}$. □

Appendix B: Frequency responses

We begin with the remark that if $g \in O$ and $\xi \in \mathcal{U}$, then

$$(g\xi)' = g\xi' + L(g^{\sigma})\xi(0).$$

(Indeed, it is easy to check that $(g\xi)(0) = g(\infty)\xi(0)$. In view of this, $(g\xi)' = sg\xi - sg(\infty)\xi(0)$. We therefore have

$$(g\xi)' = g(s\xi - s\xi(0)) + \xi(0)L(sg - sg(\infty)) = g\xi' + L(g^{\sigma})\xi(0).)$$

If F is a shift-invariant k-linear subspace in O^q, then the largest submodule contained in F is a transfer function. (The proof of this is the same as that of Proposition 1.) We say that F is a frequency response if its transfer function is "large enough" in the sense that has finite codimension.

Given a frequency response F with transfer function T, we set $\Sigma(F) = T\mathcal{U} + L(F)$. Using the above remark (and the equality $L(g)' = L(g^{\sigma})$), we can see that $\Sigma(F)$ is differentiation-invariant. Further, choosing a finite-dimensional linear subspace $X \subseteq F$ such that $F = T \oplus X$, we clearly have $\Sigma(F) = T\mathcal{U} \oplus L(X)$. So, $\Sigma(F)$ is a linear system.

Conversely, if $\mathscr{S}$ is a linear system, then clearly $\Phi(\mathscr{S}) = L^{-1}(\mathscr{S} \cap L(O^q))$ is a frequency response. By definition, it consists of those proper rational functions that correspond to the exponential trajectories.

It is easily seen that the mappings

$$F \mapsto \Sigma(F) \quad \text{and} \quad \mathscr{S} \mapsto \Phi(\mathscr{S})$$

are inverse to each other. It follows, in particular, that a linear system is uniquely determined by its exponential trajectories.

(If $\mathscr{X}$ is a subset of $\mathscr{U}^q$, we write $\overline{\mathscr{X}}$ to denote its topological closer.)

Proposition 3. *If F is a frequency response, then*

$$\Sigma(F) = \overline{L(F)}.$$

Proof. Let $F = T \oplus X$, and let $g_1, \dots, g_m$ be a basis of T. For each $g \in T$, we have

$$L(g) = g_1 L(a_1) + \cdots + g_m L(a_m) \quad (a_1, \dots, a_m \in O).$$

As noticed already, every $\xi \in T\mathscr{U}$ can be written uniquely as

$$\xi = g_1\xi_1 + \cdots + g_m\xi_m \quad (\xi_1, \dots, \xi_m \in \mathscr{U}).$$

Because $\overline{L(O)} = \mathscr{U}$, we see that $T\mathscr{U} = \overline{L(T)}$. Next, $L(X)$ must be closed in $\mathscr{U}^q$ as a finite-dimensional subspace. We thus have

$$\Sigma(F) = T\mathscr{U} + L(X) = \overline{L(T)} + \overline{L(X)} = \overline{L(T+X)} = \overline{L(F)}. \qquad \square$$

As a consequence we get a 1-dimensional case of Ehrenpreis–Malgrange–Palamodov approximation theorem (see [3]).

Corollary 3. *The exponential solutions of a linear constant-coefficient differential equation form a dense subset in the set of all solutions.*

Appendix C: Extension to time-series

Extension to time series is trivial. Indeed, the reader could notice that very little about C^∞ functions have been employed. Letting k be an arbitrary field and setting $\mathscr{U} = k^{\mathbb{Z}_+}(\simeq k[[t]])$, it only suffices to do the following: (1) Regard $\mathscr{U}$ as a torsion free module over O; (2) Take $L : k \to \mathscr{U}$ to be the natural embedding; (3) Check that $\mathscr{U} = t\mathscr{U} \oplus L(k)$.

References

[1] M. Atiyah, I.G. Macdonald, Introduction to Commutative Algebra, Addison-Wesley, Reading, MA, 1969.
[2] P.A. Fuhrmann, Algebraic system theory: an analyst's point of view, J. Franklin Inst. 301 (1976) 521–540.
[3] L. Hörmander, Linear Partial Differential Operators, Springer, New York, 1976.
[4] V. Lomadze, Application of vector bundles to factorization of rational matrices, Linear Algebra Appl. 288 (1999) 249–258.
[5] J. Mikusinski, Operational Calculus, Pergamon Press, London, 1959.
[6] J.M. Schumacher, Transformations of linear systems under external equivalence, Linear Algebra Appl. 102 (1988) 1–34.
[7] J.W. Polderman, J.C. Willems, Introduction to Mathematical Systems Theory, Springer, New York, 1998.
[8] J.C. Willems, Paradigms and puzzles in the theory of dynamical systems, IEEE Trans. Automat. Control 36 (1991) 259–294.