Systems & Control Letters 13 (1989) 279–289
North-Holland

On two definitions of observation spaces *

Yuan WANG and Eduardo D. SONTAG
Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, U.S.A.

Received 26 May 1989
Revised 7 August 1989

Abstract: This paper establishes the equality of the observation spaces defined by means of piecewise constant controls with those defined in terms of differentiable controls.

Keywords: Observation space; shuffle product; input–output system; generating series.

1. Introduction

Since their introduction in the mid 70's (see [5] and [1], as well as [7] for the discrete time analogue), observation spaces for nonlinear control systems

$$\dot{x} = f(x) + \sum u_i g_i(x), \quad y = h(x), \tag{1}$$

have played a central role in the understanding of realization theory. For the system (1), one defines the observation space $\mathscr{F}$ as the linear span of the Lie derivatives

$$L_{X_1} \dots L_{X_k} h,$$

where each X_i is either f or one of the g_i's. (Here we are taken states $x(t)$ in a manifold, f, $g_1, \dots, g_m$ vector fields, and h a function from the manifold to $\mathbb{R}$, the output map.)

It is known that many important properties of systems, such as the possibility of simulating such a system by one described by linear vector fields (the 'bilinear immersion' problem [1]), are characterized by properties of this space.

It was shown in [8] that a different type of 'observation space' is much more important when one studies questions of *input–output equations* satisfied by (1), i.e. equations of the type

$$E\left(y^{(k)}(t), \dots, y'(t),\ y(t),\ u^{(k)}(t), \dots, u'(t),\ u(t)\right) = 0 \tag{2}$$

that hold for all those pairs of functions $(u(\cdot), y(\cdot))$ that arise as solutions of (1). This alternative observation space is obtained by taking the derivatives $y(t)$, $y'(t)$, ... as functions of initial states, over all $u(t)$, $u'(t)$, ... This space is obtained by considering differentiable controls and time-derivatives, while the space previously considered is based on derivatives with respect to switching times in piecewise constant controls.

The central fact used in [8] in order to relate i/o equations to realizability is the equality of the two observation spaces defined in the above manners. This equality is fundamental not only for the results in that paper, which hold under the assumption that the spaces are finite dimensional, but also for the far more general results recently announced in [9]. However, the techniques used in [8] are based on a topological argument, involving closure in the weak topology, which does not in any way extend to the more general case of infinite dimensional observation spaces. Since the latter are the norm rather than the exception (unless the system can be simulated by a bilinear system to start with), one needs to establish the

* This research was supported in part by US Air Force Grant AFOSR-88-0235.

0167-6911/89/$3.50 © 1989, Elsevier Science Publishers B.V. (North-Holland)

equality of these two types of spaces using totally different combinatorial techniques. That is the purpose of this paper.

In the next section we provide background material on generating series. We use this formalism because in applications one does not want to restrict to systems [1] but one rather wants to treat the case of arbitrary input–output operators. Then we introduce rigorously the two spaces and establish their equality. An important role is played by an analogue of the main result in [4]. Finally we extend our results to families of operators and then give a translation of the results into the language of systems (1).

2. Generating series

Let m be a fixed integer and $I=\{0, 1, \ldots, m\}$. For any integer $k \geq 1$, we define I^k to be the set of all sequences $(i_1\, i_2\, \cdots\, i_k)$, where $i_s \in I$, $1 \leq s \leq k$. For $k=0$, we use I^0 to denote the set whose only element is the empty sequence ϕ. Let

$$I^* = \bigcup_{k \geq 0} I^k. \tag{3}$$

Then I^* is a free monoid with the composition rule:

$$(i_1\, i_2\, \cdots\, i_k)(j_1\, j_2\, \cdots\, j_l) = (i_1\, i_2\, \cdots\, i_k\, j_1\, j_2\, \cdots\, j_l).$$

If $\iota \in I^l$, then we say that the length of ι, denoted by $|\iota|$, is l.

Consider now the 'alphabet' set $P=\{\eta_0, \eta_1, \ldots, \eta_n\}$ and P^*, the free monoid generated by P, where the neutral element of P^* is the empty word, denoted by 1, and the product is concatenation. Let $P^k = \{\eta_{i_1}\eta_{i_2}\cdots\eta_{i_k} : 1 \leq i_s \leq m,\ 1 \leq s \leq k\}$ for each $k \geq 0$. We define $\mathscr{P}$ to be the $\mathbb{R}$-algebra generated by P^*, i.e., the set of all polynomials in the variables η_i's. A *power series in the noncommutative variables* $\eta_0, \eta_1, \ldots, \eta_n$ is a formal power series

$$c = \langle c, \phi\rangle + \sum_{k=1}^{\infty} \sum_{\iota \in I^k} \langle c, \eta_\iota\rangle \eta_\iota, \tag{4}$$

where $\eta_\iota = \eta_{i_1}\eta_{i_2}\cdots\eta_{i_l}$ if $\iota = i_1\, i_2\, \cdots\, i_l$, and $\langle c, \eta_\iota\rangle \in \mathbb{R}$. Note that c is a polynomial if only finitely many $\langle c, \eta_\iota\rangle$'s are non-zero. A power series is nothing more than a mapping from I^* to $\mathbb{R}$; as we shall see later, however, the algebraic structures suggested by the series formalism are very important. We use $\mathscr{S}$ to denote the set of all power series.

For c, $d \in \mathscr{S}$ and $\gamma \in \mathbb{R}$, $\gamma c + d$ is defined as the following:

$$\langle \gamma c + d, \eta_\iota\rangle = \gamma\langle c, \eta_\iota\rangle + \langle d, \eta_\iota\rangle.$$

Thus, $\mathscr{S}$ forms a vector space over $\mathbb{R}$.

We shall say that the power series c is *convergent* if

$$|\langle c, \eta_\iota\rangle| \leqslant KM^k k! \quad \text{for each } \iota \in I^k, \text{ and each } k \geq 0, \tag{5}$$

where K and M are some constants.

Let T be a fixed value of time and let $\mathscr{U}_T$ be the set of all essentially bounded functions $u:[0, T] \to \mathbb{R}^m$ endowed with the L^1 norm. We write $\|u\|_\infty$ for $\max\{\|u_i\|_\infty : 1 \leq i \leq m\}$ if u_i is the i-th component of u, and $\|u_i\|_\infty$ is the essential supernorm of u_i. For each $u \in \mathscr{U}_T$ and $\iota \in I^l$, we define inductively the functions $V_\iota = V_\iota[u] \in \mathscr{C}[0, T]$ by

$$V_\phi = 1 \quad \text{and} \quad V_{i_1 \ldots i_{l+1}}[u](t) = \int_0^t u_{i_1}(s) V_{i_2 \ldots i_{l+1}}(s)\, \mathrm{d}s, \tag{6}$$

where u_i is the i-th coordinate of $u(t)$ for $i=1, 2, \ldots, m$ and $u_0(t) \equiv 1$. It can be proved that each map

$$U_T \to \mathscr{C}[0, T], \quad u \mapsto V_\iota[u]$$

is continuous with respect to L^1 norm in $\mathscr{U}_T$, $\mathscr{C}^0$ norm in $\mathscr{C}[0, T]$.

Suppose c is convergent and let K and M be as in (5). Then for any

$$T < (Mm + M)^{-1}, \tag{7}$$

the series of functions

$$F_c[u](t) = \sum \langle c, \eta_\iota \rangle V_\iota[u](t) \tag{8}$$

is uniformly and absolutely convergent for all $t \in [0, T]$ and all those $u \in \mathscr{U}_T$ such that $\|u\|_\infty \le 1$ (cf. [3]). In fact, (8) is absolutely and uniformly convergent for all $t \in [0, T]$ provided $T\|u\|_\infty < (Mm + M)^{-1}$. For each nonnegative T, let

$$\mathscr{V}_T = \{ u \in \mathscr{U}_T \colon \|u\|_\infty < 1 \}. \tag{9}$$

We say that T is *admissible for c* if T satisfies (7). Since each operator $u \to V_\iota[u]$ is continuous, it follows that $F_c \colon \mathscr{V}_T \to \mathscr{C}[0, T]$ is continuous if T is admissible for c. We call F_c an *input–output* map defined on $\mathscr{V}_T$. Thus every convergent power series defines an i/o map. On the other hand, the power series c is uniquely determined by F_c in the following sense:

Lemma 2.1. *Suppose that c and d are two convergent power series. If* $F_c = F_d$ *on* $\mathscr{V}_T$ *for any* $T > 0$, *then* $c = d$.

Proof. It is enough to show that if c is convergent and $F_c = 0$ on $\mathscr{V}_T$ for some small T, then $c = 0$. Consider piecewise constant controls in $\mathscr{V}_T$, and use the notation

$$u = (\mu_1, t_1)(\mu_2, t_2) \cdots (\mu_k, t_k)$$

to denote the piecewise constant control whose value is μ_i in the time interval

$$\left(\sum_{j=0}^{i-1} t_j, \sum_{j=0}^{i} t_j \right)$$

where

$$\mu_j = (\mu_{1j}, \mu_{2j}, \ldots, \mu_{mj}) \in \mathbb{R}^m, \quad |\mu_{ij}| < 1, \quad 1 \le j \le k, 1 \le i \le m,$$

and $t_0 = 0$.

By assumption, for any μ_i, t_i, such that $\Sigma t_i < T$, $F_c[(\mu_1, t_1)(\mu_2, t_2) \ldots (\mu_k, t_k)](t) = 0$, where $t = \Sigma t_i$. Take $y = F_c[u]$ as a function of $\mu_1, \ldots, \mu_k$ and $t_1, \ldots, t_k$. Then

$$\left. \frac{\partial^k}{\partial t_1 \ldots \partial t_k} \right|_{t=0+} \left. \frac{\partial^s}{\partial \mu_{i_1 j_1} \ldots \partial \mu_{i_s j_s}} \right|_{\mu=0} y = 0 \tag{10}$$

for all $i_1, \ldots, i_s$, $j_1, \ldots, j_s$, where the evaluation at t^+ means that we evaluate at $t_1^+, \ldots, t_k^+$. We claim that, for $i_1, \ldots, i_s$, $j_1, \ldots, j_s$ given such that $j_r \ne j_q$ if $r \ne q$,

$$\left. \frac{\partial^k}{\partial t_1 \cdots \partial t_k} \right|_{t=0^+} \left. \frac{\partial^s}{\partial \mu_{i_1 j_1} \cdots \partial \mu_{i_s j_s}} \right|_{\mu=0} y = \langle c, \eta_{l_1} \cdots \eta_{l_k} \rangle, \tag{11}$$

where

$$l_p = \begin{cases} i_r & \text{if } k - (p-1) = j_r, \\ 0 & \text{if } k - (p-1) \notin \{ j_1, \ldots, j_s \}. \end{cases}$$

To see this, write $y(t)=\Sigma\langle c,\eta_\iota\rangle V_\iota(t)$. Then, directly from the definition (6),

$$\left.\frac{\partial^k}{\partial t_1\cdots\partial t_k}\right|_{t=0^+} y=\sum\langle c,\eta_{l_1}\cdots\eta_{l_k}\rangle\mu_{l_1k}\cdots\mu_{l_k1}. \tag{12}$$

One can see that if $\{(i_1,j_1),\dots,(i_s,j_s)\}\subseteq\{(l_1,k),\dots,(l_k,l)\}$ and $l_p=0$ for $p\notin\{j_1,\dots,j_s\}$ then

$$\left.\frac{\partial^s}{\partial\mu_{i_1j_1}\cdots\partial\mu_{i_sj_s}}\right|_{\mu=0}\mu_{l_1k}\cdots\mu_{l_k1}=1;$$

otherwise,

$$\left.\frac{\partial^s}{\partial\mu_{i_1j_1}\cdots\partial\mu_{i_sj_s}}\right|_{\mu=0}\mu_{l_1k}\cdots\mu_{l_k1}=0.$$

Combining this fact and (12), we get (11). It follows immediately that if $F_c[u]=0$ for all piecewise constant controls, then $c=0$. □

3. Observation spaces

To each monomial $\alpha=\eta_\iota$, we associate a shift operator $c\mapsto\alpha^{-1}$ defined by

$$\langle\alpha^{-1}c,\eta_\iota\rangle=\langle c,\alpha\eta_\iota\rangle\quad\text{for }\eta_\iota\in P^*.$$

Note that $\alpha_2^{-1}\alpha_1^{-1}=(\alpha_1\alpha_2)^{-1}c$. It was shown in [8] that if c is convergent and T is admissible for c, than α^{-1} is also convergent and T is also admissible for α^{-1} for any $\alpha\in P^*$. Using this notation, we get the following fundamental formula [2], which follows from the definition (6):

$$\frac{\mathrm{d}}{\mathrm{d}t}F_c[u](t)=F_{\eta_0^{-1}c}[u](t)+\sum_{j=1}^m u_j(t)F_{\eta_j^{-1}c}[u](t) \tag{13}$$

for any $u\in\mathscr{V}_T$ which is continuous.

Formula (13) implies, by induction, that if $u\in\mathscr{V}_T$ is of class $\mathscr{C}^{k-1}$, then $F_c[u]$ is of class $\mathscr{C}^k$.

In realization theory, the concept of observation spaces plays a very important role. One may defined observation spaces in two ways. Let us now introduce the first approach. To each convergent power series c, we define the observation space $\mathscr{F}_1$ to be the space spanned by all the power series $\alpha^{-1}c$ over $\mathbb{R}$, i.e.,

$$\mathscr{F}_1(c)=\operatorname{span}_{\mathbb{R}}\{\alpha^{-1}c:\ \alpha\in P^*\}. \tag{14}$$

It is well known that F_c can be realized by a bilinear system and only if $\dim\mathscr{F}_1(c)<\infty$; see e.g. [1].

To define the second type of observation spaces, we need to introduce the *shuffle product* on $\mathscr{P}$ (cf. [6]). The shuffle product on $\mathscr{P}$ is defined in the following way. First, inductively on the length of of words in P^*, we let

$$1\sqcup\!\sqcup z=z\sqcup\!\sqcup 1=z\quad\text{for any }z\in P;$$
$$wz\sqcup\!\sqcup w'z'=(w\sqcup\!\sqcup w'z')z+(wz\sqcup\!\sqcup w)z'\quad\text{for any }w,\ w'\in P^*,\ z,\ z'\in P.$$

Note that the shuffle product is commutative:

$$w_1\sqcup\!\sqcup w_2=w_2\sqcup\!\sqcup w_1\quad\text{for any }w_1,\ w_2\in P^*.$$

If $c=\Sigma\langle c,\eta_\kappa\rangle\eta_\kappa$ and $d=\Sigma\langle d,\eta_\iota\rangle\eta_\iota$ are polynomials, then

$$c\sqcup\!\sqcup d:=\sum_n\sum_{|\kappa|+|\iota|=n}\langle c,\eta_\kappa\rangle\langle d,\eta_\iota\rangle\eta_\iota\sqcup\!\sqcup\eta_\kappa.$$

The following lemma can be proved by induction on n:

Lemma 3.1. *Suppose* $w_1,\dots,w_n \in P^*$ *and* $w_i = w_i' z_i$ *with* $w_i' \in P^*$, $z_i \in P$. *Then*

$$\sum_{s=1}^{n} (w_1 ш \cdots ш w_s' ш w_n) z_s = w_1 ш w_2 ш \cdots ш w_n. \qquad \square$$

Now consider for each $q \geq 1$, the following set of $2 \times q$ matrices:

$$S_q = \left\{ \begin{pmatrix} j_1 & j_2 & \cdots & j_q \\ i_1 & i_2 & \cdots & i_q \end{pmatrix} : i_s,\ j_s \in \mathbb{Z},\ 1 \leq i_s \leq m,\ (0,1) \leq (i_1, j_1) \leq \cdots \leq (i_q, j_q) \right\}, \tag{15}$$

where '$\leq$' is the lexicographic order on the set $\{(i,j): i, j \in \mathbb{Z}\}$. For each element in S_q with $n \geq q + \Sigma j_r$, we define

$$\Gamma_{i_1 \dots i_q}^{j_1 \dots j_q}(n) = \eta_0^k ш \eta_{i_1} X^{j_1} ш \eta_{i_2} X^{j_2} ш \cdots ш \eta_{i_q} X^{j_q}|_{X=1}, \tag{16}$$

where $k = n - q - \Sigma j_s$. The evaluation is interpreted as follows: first introduce a new variable X, then perform all shuffles, and finally delete X from the result. Note that (16) is different from $\eta_{i_1} ш \eta_{i_2} ш \cdots ш \eta_{i_q}$, for example,

$$\eta_0 ш \eta_1 X|_{X=1} = \eta_0\eta_1 + 2\eta_1\eta_0 \quad \text{while} \quad \eta_0 ш \eta_1 = \eta_0\eta_1 + \eta_1\eta_0.$$

For $w \in P^*$ and $c \in \mathscr{S}$, we define $\psi_c(w) = w^{-1}c$. For any polynomial $d = \Sigma\langle d, \eta_\kappa\rangle\eta_\kappa$, we define

$$\psi_c(d) = \sum \langle d, \eta_\kappa \rangle \eta_\kappa^{-1} c.$$

Now let $X_j = (X_{1j},\dots,X_{mj})$ be m indeterminates over $\mathbb{R}$, for $j \geq 0$. For any $n > 0$, let

$$c_n(X_0,\dots,X_{n-1}) = \psi_c(\eta_0^n) + \sum_{q=0}^{n} \sum \frac{1}{s_1! \cdots s_p!} \psi_c\left(\Gamma_{i_1 \dots i_q}^{j_1 \dots j_q}(n)\right) X_{i_1 j_1} \cdots X_{i_q j_q}, \tag{17}$$

where the second sum is taken over all those elements of $\mathscr{S}_q$ such that $\Sigma j_s + q \leq n$, and where $s_1,\dots,s_p$ are integers so that

$$\begin{pmatrix} j_1 & j_2 & \cdots & j_q \\ i_1 & i_2 & \cdots & i_q \end{pmatrix} = \begin{pmatrix} \underbrace{\begin{matrix} \beta_1 \cdots \beta_1 \\ \alpha_1 \cdots \alpha_1 \end{matrix}}_{s_1} & \underbrace{\begin{matrix} \beta_2 \cdots \beta_2 \\ \alpha_2 \cdots \alpha_2 \end{matrix}}_{s_2} & \begin{matrix} \cdots \\ \cdots \end{matrix} & \underbrace{\begin{matrix} \beta_p \cdots \beta_p \\ \alpha_p \cdots \alpha_p \end{matrix}}_{s_p} \end{pmatrix}$$

and $(\alpha_1, \beta_1) < (\alpha_2, \beta_2) < \cdots < (\alpha_p, \beta_p)$. For $n = 0$, we define

$$c_0 := c.$$

We are now ready to introduce the second type of observation space associated to c, $\mathscr{F}_2(c)$. This is defined as follows:

$$\mathscr{F}_2(c) = \operatorname{span}_{\mathbb{R}}\{c_n(\mu_0,\dots,\mu_{n-1}): \mu_i \in \mathbb{R}^m,\ 0 \leq i \leq n-1,\ n \geq 0\}. \tag{18}$$

We will see below that the elements of $\mathscr{F}_2(c)$ are closely related to the derivatives of $F[u](t)$ with respect to time. A central fact that will be needed in the proof of our main result is that the coefficient of the generating series can be partitioned into infinitely many sets of finitely many elements such that the coefficient of each monomial $u_{i_1}^{(j_1)} u_{i_2}^{(j_2)} \cdots u_{i_p}^{(j_p)}$ appearing when computing the derivatives $y^{(s)}$ only depends on elements of one of these sets. This can be proved directly, but the following lemma gives a useful expression. This formula is an analogue, proved by using different techniques, of a similar formula proved for state space systems, given in the paper [4].

Lemma 3.2. *If $u \in \mathscr{V}_T$ is of class $\mathscr{C}^{n-1}$ and T is admissible for c, then we have*

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} F_c[u](t) = F_{c_n(u(t),\ldots,u^{n-1}(t))}[u](t). \tag{19}$$

Before proving this formula, we look at an example to illustrate its meaning.

Example 3.3. For $n = 2$, we have

$$\begin{aligned} c_2(X_1, X_2) &= \psi_c(\eta_0^2) + \sum_{i=1}^m \psi_c(\Gamma_i^0(2)) X_{i0} \\ &\quad + \sum_{i<j} \psi_c(\Gamma_{ij}^{00}(2)) X_{i0} X_{j0} + \sum_{i=1}^m \tfrac{1}{2}\psi_c(\Gamma_{ii}^{00}(2)) X_{i0}^2 + \sum_{i=1}^m \psi_c(\Gamma_i^1(2)) X_{i1} \\ &= (\eta_0\eta_0)^{-1}c + \sum\left((\eta_0\eta_i)^{-1}c + (\eta_1\eta_0)^{-1}c\right) X_{i0} \\ &\quad + \sum_{i<j}\left((\eta_i\eta_j)^{-1}c + (\eta_j\eta_i)^{-1}c\right) X_{i0}X_{j0} + \sum(\eta_i\eta_i)^{-1} c X_{i0}^2 + \sum \eta_i^{-1} c X_{i1}. \end{aligned}$$

Thus, for $n = 2$, formula (19) becomes:

$$\begin{aligned} y''(t) &= F_{c_2(u(t),u'(t))}[u](t) \\ &= F_{(\eta_0\eta_0)^{-1}c}[u](t) + \sum\left(F_{(\eta_0\eta_i)^{-1}c}[u](t) + F_{(\eta_i\eta_0)^{-1}c}[u](t)\right)u_i(t) \\ &\quad + \sum_{i<j}\left(F_{(\eta_i\eta_j)^{-1}c}[u](t) + F_{(\eta_j\eta_i)^{-1}c}[u](t)\right)u_i(t)u_j(t) + \sum F_{(\eta_i\eta_i)^{-1}c}[u](t)u_i^2 \\ &\quad + \sum F_{\eta_i^{-1}c}[u](t)u_i'(t). \end{aligned} \tag{20}$$

Proof of Lemma 3.2. For each $\eta_\iota \in P^*$, define $\theta_c(\eta_\iota) = F_{\eta_\iota^{-1}c}$ and for any polynomial $d = \sum\langle d, \eta_\kappa\rangle\eta_\kappa$, define

$$\theta_c(d) = \sum\langle d, \eta_\kappa\rangle\theta_c(\eta_\iota) = \sum\langle d, \eta_\kappa\rangle F_{\eta_\kappa^{-1}c}.$$

Then (19) is equivalent to

$$y^{(n)}(t) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}F_c[u](t) = \sum_{q=0}^{n}\sum_{\mathscr{S}_q}\frac{1}{s_1!\cdots s_p!}\theta_c\left(\Gamma_{i_1\cdots i_q}^{j_1\cdots j_q}(n)\right)(t)\,u_{i_1}^{(j_1)}(t)\cdots u_{i_q}^{(j_q)}(t), \tag{21}$$

in the other words, $y^{(n)}(t)$ is a polynomial in $u(t),\ldots,u^{(n)}(t)$ whose coefficients are the $\theta_c(\eta_\iota)(t)$'s, and the coefficient of $u_{i_1}^{(j_1)}(t)\cdots u_{i_q}^{(j_q)}(t)$ in $y^{(n)}(t)$ is

$$\frac{1}{s_1!\cdots s_p!}\theta_c\left(\Gamma_{i_1\cdots i_q}^{j_1\cdots j_q}(n)\right)(t). \tag{22}$$

Note that the right side of (22) can also be written as

$$\frac{1}{s_1!\cdots s_p!}\theta_c\left(\left(\eta_0^k ⧢^{s_1} \eta_{\alpha 1}X^{\beta_1} ⧢^{s_2} \eta_{\alpha_2}X^{\beta_2} ⧢ \cdots ⧢^{s_p} \eta_{\alpha_p}X^{\beta_p}\right)\Big|_{X=1}\right)(t)$$

if $u_{i_1}^{(j_1)}\cdots u_{i_q}^{(j_q)} = (u_{\alpha_1}^{(\beta_1)})^{s_1}\cdots(u_{\alpha_p}^{(\beta_p)})^{s_p}$, where

$$\begin{aligned} & w_1 ⧢^{s_1} w_2 ⧢^{s_2} w_3 ⧢ \cdots ⧢^{s_{p-1}} w_p \\ &= w_1 ⧢ \underbrace{w_2 ⧢ w_2 ⧢ \cdots ⧢ w_2}_{s_1} ⧢ \underbrace{w_3 ⧢ \cdots ⧢ w_3}_{s_2} ⧢ \cdots ⧢ \underbrace{w_p ⧢ \cdots ⧢ w_p}_{s_{p-1}}. \end{aligned}$$

We now use induction to prove the lemma. From (13) we see that the conclusion is true for $n=1$.

Suppose the conclusion is true for $n-1$. Consider the coefficient of $u_{i_1}^{(j_1)}\cdots u_{i_q}^{(j_q)}$ in $y^{(n)}$. By inducation from formula (13) it can be seen that $\Sigma j_s+q\le n$. First we assume that $\Sigma j_s+q<n$. Let $k=n-\Sigma j_s-q$. Suppose

$$u_{i_1}^{(j_1)}\cdots u_{i_q}^{(j_q)}=\left(u_{\alpha_1}^{(\beta_1)}\right)^{s_1}\cdots\left(u_{\alpha_p}^{(\beta_p)}\right)^{s_p},$$

where $(\alpha_1,\beta_1)<\cdots<(\alpha_p,\beta_p)$. Further, we assume that $\beta_r=0$ for $r\le l$. Let

$$\hat{y}_1(t)=\sum_{r=1}^{p}\frac{1}{\tau_r}\theta(w_r)(t)\left(u_{\alpha_1}^{(\beta_1)}\right)^{s_1}\cdots\left(v_{\alpha_r}^{(\beta_r)}\right)^{s_r}\cdots\left(u_{\alpha_p}^{(\beta_p)}\right)^{s_p},$$

where

$$\left(v_{\alpha_r}^{(\beta_r)}\right)^{s_r}=\begin{cases}u_{\alpha_r}^{s_r-1} & \text{if } \beta_r=0,\\ \left(u_{\alpha_r}^{(\beta_r)}\right)^{s_r-1}u_{\alpha_r}^{(\beta_r-1)} & \text{if } \beta_r\ge 1,\end{cases}$$

and $\tau_r=s_1'!\cdots s_{p'}'!$ if

$$\left(u_{\alpha_1}^{(\beta_1)}\right)^{s_1}\cdots\left(v_{\alpha_r}^{(\beta_r)}\right)^{s_r}\cdots\left(u_{\alpha_p}^{(\beta_p)}\right)^{s_p}=\left(u_{\alpha_1'}^{(\beta_1')}\right)^{s_1'}\cdots\left(u_{\alpha_{p'}'}^{(\beta_{p'}')}\right)^{s_{p'}'},$$

and

$$w_r=\begin{cases}\eta_0^k ш^{s_1}\eta_{\alpha_1}ш\cdots ш^{(s_r-1)}\eta_{\alpha_r}ш\cdots ш^{s_p}\eta_{\alpha_p}X^{\beta_p} & \text{if } \beta_r=0,\\ \eta_0^k ш^{s_1}\eta_{\alpha_1}ш\cdots ш^{s_r-1}\eta_{\alpha_r}X^{\beta_r}ш\,\eta_{\alpha_r}X^{\beta_r-1}ш\cdots ш^{s_p}\eta_{\alpha_p}X^{\beta_p} & \text{if } \beta_r\ne 0.\end{cases}$$

Note that the coefficient of $u_{i_1}^{(j_1)}\cdots u_{i_q}^{(j_q)}$ in

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{1}{\tau_r}\theta(w_r)(t)\left(u_{\alpha_1}^{(\beta_1)}\right)^{s_1}\cdots\left(v_{\alpha_r}^{(\beta_r)}\right)^{s_r}\cdots\left(u_{\alpha_p}^{(\beta_p)}\right)^{s_p}\right\}$$

is

$$\begin{cases}\dfrac{1}{s_1!\cdots(s_r-1)!\cdots s_p!}\theta(w_r\eta_r)(t) & \text{if } r\le l,\\[2ex] \dfrac{1}{s_1!\cdots(s_r-1)!\cdots s_p!}\theta(w_r)(t) & \text{if } r>l.\end{cases}$$

Let

$$y_1(t)=\hat{y}_1(t)+\frac{1}{s_1!\cdots s_p!}\theta_c\left(\Gamma_{i_1\cdots i_q}^{j_1\cdots j_q}(n-1)\right)u_{i_1}^{(j_1)}(t)\cdots u_{i_q}^{(j_q)}(t).$$

By induction assumption, the coefficient of $u_{i_1}^{(j_1)}\cdots u_{i_q}^{(j_q)}$ in $y^{(n)}(t)$ is the same as in $y_1'(t)$. Thus, this coefficient is $\theta_c(w)(t)$, where

$$\begin{aligned}w=\Bigg\{&\sum_{r=1}^{l}\frac{1}{s_1!\cdots(s_r-1)!\cdots s_p!}\left(\eta_0^k ш^{s_1}\eta_{\alpha_1}ш\cdots ш^{s_r-1}\eta_{\alpha_r}ш\cdots ш^{s_p}\eta_{\alpha_p}X^{\beta_p}\right)\eta_{\alpha_r}\\ &+\sum_{r=l+1}^{p}\frac{1}{s_1!\cdots(s_r-1)!\cdots s_p!}\eta_0^k ш^{s_1}\eta_{\alpha_1}ш\cdots ш^{s_r-1}\eta_{\alpha_r}X^{\beta_r}ш\,\eta_{\alpha_r}X^{\beta_r-1}ш\cdots ш^{s_p}\eta_{\alpha_p}X^{\beta_p}\\ &+\frac{1}{s_1!s_2!\cdots s_p!}\left(\eta_0^{k-1}ш^{s_1}\eta_{\alpha_2}ш\cdots ш^{s_p}\eta_{\alpha_p}X^{\beta_p}\right)\eta_0\Bigg\}\Bigg|_{X=1}.\end{aligned}\tag{23}$$

Notice that

$$w_1 ш^{r-1} w_2 = \frac{1}{r} \sum_{t=0}^{r-1} w_1 ш^t w_2 ш 1 ш^{r-1-t} w_2$$

and

$$\left\{ w_1 ш^{r-1} w_2 X^{\beta} ш w_2 X^{\beta-1} \right\}\Big|_{X=1} = \left\{ \left(w_1 ш^{r-1} w_2 X^{\beta} ш w_2 X^{\beta-1} \right) X \right\}\Big|_{X=1}$$
$$= \frac{1}{r} \left\{ \left(\sum_{t=0}^{r-1} w_1 ш^t w_2 X^{\beta} ш w_2 X^{\beta-1} ш^{r-1-t} w_2 X^{\beta} \right) X \right\}\Bigg|_{X=1}.$$

Applying Lemma 3.1 to (23), we get

$$w = \frac{1}{s_1! \cdots s_p!} \left\{ \eta_0^k ш^{s_1} \eta_{\alpha_1} ш \cdots ш^{s_p} \eta_{\alpha_p} X^{\beta_p} \right\}\Big|_{X=1} = \frac{1}{s_1! \cdots s_p!} \Gamma_{i_1 \cdots i_q}^{j_1 \cdots j_q}(n).$$

In the case $q + \Sigma j_s = n$, the proof is virtually the same except that $k = 0$, which leads to the fact that the coefficient of $u_{i_1}^{(j_1)} \cdots u_{i_q}^{(j_q)}$ in $y^{(n-1)}$ is 0, so the last term in (23) disappears. □

4. Main result

In last section we defined $\Gamma_{i_1 \cdots i_q}^{j_1 \cdots j_q}(n)$ and $c_n(X_0, \ldots, X_{n-1})$. One can see that $c_n(X_0, \ldots, X_{n-1})$ is a polynomial on the X_i's with coefficients belonging to $\mathscr{F}_1(c)$. Thus, $c_n(\mu_0, \ldots, \mu_{n-1})$ *is a linear combination of elements of* $\mathscr{F}_1(c)$ *for each fixed value of* $(\mu_0, \ldots, \mu_{n-1})$. Therefore,

$$\mathscr{F}_2(c) \subseteq \mathscr{F}_1(c).$$

But in fact, these two spaces are the same as we can see in the following theorem.

Theorem 1. *If c is a power series, the* $\mathscr{F}_1(c) = \mathscr{F}_2(c)$.

Proof. We have shown that $\mathscr{F}_2(c) \subseteq \mathscr{F}_1(c)$. The other direction is however much less trivial. Now for fixed positive integers k and $i_1, i_2, \ldots, i_q$ such that $1 \le i_1 \le i_2 \le \cdots \le i_q \le m$, let

$$S^k(i_1, i_2, \ldots, i_q) = \left\{ \sigma(\underbrace{0, \ldots, 0}_{k}, i_1, i_2, \ldots, i_q) \colon \sigma \in S_n \right\},$$

where $n = k + q$ and S_n is the permutation group on a set of n elements. Let

$$\Omega_k(i_1, i_2, \ldots, i_q) = \left\{ (\eta_{l_1} \eta_{l_2} \cdots \eta_{l_n})^{-1} c \colon (l_1, \ldots, l_n) \in S^k(i_1, i_2, \ldots, i_q) \right\}.$$

Then

$$\mathscr{F}_1(c) = \operatorname{span}_{\mathbb{R}} \left\{ d \in \Omega_k(i_1, i_2, \ldots, i_q) \colon k \ge 0,\ q \ge 0 \right\}.$$

Thus the theorem can be proved by showing that

$$\Omega_k(i_1, i_2, \ldots, i_q) \subseteq \mathscr{F}_2(c) \tag{24}$$

for any k, q, and $(i_1, i_2, \ldots, i_q)$. Now fix k and $(i_1, i_2, \ldots, i_q)$ and put the lexicographic order on $\Omega_k(i_1, i_2, \ldots, i_q)$ according to the order of $(l_1, l_2, \ldots, l_n)$. Write the elements of $\Omega_k(i_1, i_2, \ldots, i_q)$ ordered as $Y_1, Y_2, \ldots, Y_r$. Let

$$\hat{\Omega}_k(i_1, i_2, \ldots, i_q) = \left\{ d = \psi_c\left(\Gamma_{i_1 \cdots i_q}^{j_1 \cdots j_q}(k)\right) \colon j_s \ge 0 \right\}.$$

Then we have $\hat{\Omega}_k(i_1, i_2, \ldots, i_q) \subseteq \mathscr{F}_2(c)$. Put the lexicographic order on $\hat{\Omega}_k(i_1, i_2, \ldots, i_q)$ according to the order of $(\sum j_s, j_1, \ldots, j_q)$. Notice that for each element $d_i \in \hat{\Omega}_k(i_1, i_2, \ldots, i_q)$, there exist some positive integers a_{ij} such that

$$d_i = \sum_{j=1}^{r} a_{ij} Y_j.$$

Let A be the matrix of r columns and infinitely many rows whose (i, j)-th entry is a_{ij}, i.e., $A = (a_{ij})$.

We claim that A is of full column rank in the sense that there is no nonzero vector $v \in \mathbb{R}^n$ such that $Av = 0$. Suppose there is some $v \neq 0$ such that $Av = 0$. Construct a polynomial e in the following way:

$$\langle e, \eta_{l_1} \cdots \eta_{l_t} \rangle = 0$$

if $(l_1, \ldots, l_t) \notin S^k(i_1, i_2, \ldots, i_q)$ and

$$\langle e, \eta_{l_1} \cdots \eta_{l_t} \rangle = v_i$$

if $(l_1, \ldots, l_t) \in S^k(i_1, i_2, \ldots, i_q)$ and $(\eta_{l_1} \cdots \eta_{l_t})^{-1}c$ corresponds to the i-th element of $\Omega^k(i_1, i_2, \ldots, i_q)$. By the definitions of A and d, we know that

$$e_n(\mu_0, \ldots, \mu_{n_1}) = 0 \quad \text{for any } n.$$

Therefore,

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} F_e[u](0) = F_{d_n(\mu_0, \ldots, \mu_{n-1})}[u](0) = 0$$

for any n and any analytical control u, which implies that $F_e[u] = 0$ for any analytical controls. Since analytical controls are dense in $\mathscr{V}_T$ (under the L^1 topology), it follows that $F_e \equiv 0$. By Lemma 2.1, $e = 0$. Thus, $v = 0$, a contradiction to the assumption. Hence, A is of full column rank.

Now let $\mathscr{A}_s$ be the subspace of $\mathbb{R}^r$ spanned by the first s row vectors of A. Then

$$\mathscr{A}_1 \subseteq \mathscr{A}_2 \subseteq \cdots \subseteq \mathscr{A}_s \subseteq \cdots.$$

Since $\mathscr{A}_s \subset \mathbb{R}^r$ for any s, there exists some $s_0 > 0$ such that $\mathscr{A}_s = \mathscr{A}_{s_0}$ for every $s \geq s_0$. Let A_1 be the $s_0 \times r$ submatrix of A consisting the first s_0 rows of A. Then $A = TA_1$ for some matrix T. Therefore rank $A_1 = r$. By the construction of A_1, we know that

$$A_1 \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_r \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{s_0} \end{pmatrix}.$$

From the facts that $d_i \in \mathscr{F}_2(c)$ and A_1 is of full column rank, we get the conclusion that $Y_i \in \mathscr{F}_2(c)$ for each i, therefore, (24) holds.

Since k, q and $(i_1, i_2, \ldots, i_q)$ were arbitrary, we get the desired conclusion $\mathscr{F}_1(c) = \mathscr{F}_2(c)$. □

5. Families of series and systems

In this section we consider families of power series. Let Λ be a index set. We say that $\boldsymbol{c}$ is a *family of power series* (parameterized by $\lambda \in \Lambda$) if $\boldsymbol{c} := \{c^\lambda : \lambda \in \Lambda\}$, where c^λ is a power series for each fixed λ. A family $\boldsymbol{c}$ can also be viewed as a power series with coefficient belonging to the ring of functions from Λ to $\mathbb{R}$, i.e,

$$\boldsymbol{c} = \sum \langle \boldsymbol{c}, \eta_\iota \rangle \eta_\iota,$$

where $\langle \boldsymbol{c}, \eta_\iota \rangle : \Lambda \to \mathbb{R}$, $\langle \boldsymbol{c}, \eta_\iota \rangle(\lambda) \mapsto \langle c^\lambda, \eta_\iota \rangle$.

Let $\mathfrak{S}$ be the set of all families of power series. For $\boldsymbol{c}$, $\boldsymbol{d} \in \mathfrak{S}$ and $\gamma \in \mathbb{R}$, $\gamma \boldsymbol{c} + \boldsymbol{d}$ is defined to be the family of power series $\{\gamma c^\lambda + d^\lambda : \lambda \in \Lambda\}$. Thus $\mathfrak{S}$ forms a vector space over $\mathbb{R}$.

We say that $\boldsymbol{c}$ is a *convergent family* if each member of the family is convergent. For any monomial $\alpha \in P^*$, $\alpha^{-1}\boldsymbol{c}$ is defined to be the family $\{\alpha^{-1}c^\lambda : \lambda \in \Lambda\}$. For any $n \geq 0$, $\boldsymbol{c}_n(X_0, \ldots, X_{n-1})$ is defined to be the family

$$\{c_n^\lambda(X_0, \ldots, X_{n-1}) \colon \lambda \in \Lambda\},$$

where $X_i = (X_{i1}, \ldots, X_{im})$ are m indeterminates over $\mathbb{R}$, $i \geq 0$. Applying 3.2, we have that

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} F_{c^\lambda}[u](t) = F_{c_n^\lambda(u(t), \ldots, u^{n-1}(t))}[u](t), \tag{25}$$

for each λ.

As in the case of single power series, we associate to $\boldsymbol{c}$ two types of observation spaces in the following way:

$$\mathfrak{F}_1(\boldsymbol{c}) := \operatorname{span}_{\mathbb{R}}\{\alpha^{-1}\boldsymbol{c} \colon \alpha \in P^*\},$$

$$\mathfrak{F}_2(\boldsymbol{c}) := \operatorname{span}_{\mathbb{R}}\{\boldsymbol{c}_n(\mu_0, \ldots, \mu_{n-1}) \colon \mu_i \in \mathbb{R}^m,\ 0 \leq i \leq n-1,\ n \geq 0\}.$$

Note that $\mathfrak{F}_1(\boldsymbol{c})$ (respectively, $\mathfrak{F}_2(\boldsymbol{c})$) is formally analogous to $\mathscr{F}_1(c)$ (respectively, $\mathscr{F}_2(c)$) studied before. Using $\boldsymbol{c}$ and $\boldsymbol{d}$ instead of c and d in the proof of Theorem 1, we get the following result:

Theorem 2. *If* $\boldsymbol{c}$ *is a family of power series, then* $\mathfrak{F}_1(\boldsymbol{c}) = \mathfrak{F}_2(\boldsymbol{c})$. □

Now consider a state space system

$$\dot{x} = g_0(x) + \sum u_i g_i(x), \qquad y = h(x), \tag{26}$$

where $x(t) \in X$, a $\mathscr{C}^\omega$ manifold, $g_0, g_1, \ldots, g_m$ are $\mathscr{C}^\omega$ vector fields, and h a $\mathscr{C}^\omega$ function from X to $\mathbb{R}$. One type of observation space associated with (26) is

$$F_1 := \operatorname{span}_{\mathbb{R}}\{L_{g_{i_1}} \cdots L_{g_{i_k}} h \colon k \geq 0\}.$$

For $\mu_0, \ldots, \mu_{k-1}$ given, we let, for each $x \in X$,

$$y^{\mu_0 \cdots \mu_{k-1}}(x) := \left.\frac{\mathrm{d}^k}{\mathrm{d}t^k}\right|_{t=0} y_x(t),$$

where $y_x(t)$ is the output corresponding to initial state x and any $\mathscr{C}^\infty$ input u such that $u^{(j)}(0) = \mu_j$ for $0 \leq j \leq k-1$.

We associate to (26) a second type of observation space, as follows:

$$F_2 := \operatorname{span}_{\mathbb{R}}\{y^{\mu_1 \cdots \mu_{k-1}} \colon \mu_i \in \mathbb{R}^m,\ k \geq 0\}.$$

By a fundamental formula due to Fliess (see [3]), the input–output map of (26) can be written as

$$y(t) = F_c[u](t),$$

where the family $\boldsymbol{c}$ is defined by $\langle \boldsymbol{c}, \eta_{i_1} \cdots \eta_{i_k}\rangle = L_{g_{i_k}} \cdots L_{g_{i_1}} h$, or, equivalently, for the output corresponding to the initial state x,

$$y_x(t) = F_{c^x}[u](t),$$

where $\langle c^x, \eta_{i_1} \cdots \eta_{i_k}\rangle = L_{g_{i_k}} \cdots L_{g_{i_1}} h(x)$. Thus,

$$L_{g_{i_k}} \cdots L_{g_{i_1}} h = \langle \boldsymbol{c}, \eta_{i_1} \cdots \eta_{i_k}\rangle = \langle (\eta_{i_1} \cdots \eta_{i_k})^{-1} \boldsymbol{c}, \phi\rangle,$$

and, therefore,

$$F_1 = \{\langle \boldsymbol{d}, \phi\rangle : \boldsymbol{d} \in \mathfrak{F}_1(\boldsymbol{c})\}.$$

By (25), we know that $y_x^{\mu_0 \cdots \mu_{k-1}} = F_{c_k^x(\mu_0,\ldots,\mu_{k-1})}[u](0) = \langle c_k^x(\mu_0,\ldots,\mu_{k-1}), \phi\rangle$. Hence,

$$F_2 = \{\langle \boldsymbol{d}, \phi\rangle : \boldsymbol{d} \in \mathfrak{F}_2(\boldsymbol{c})\}.$$

So the following conclusion follows immediately from Theorem 2:

Corollary 5.1. *For the state space system* (26), $F_1 = F_2$. □

References

[1] M. Fliess and I. Kupka, A finiteness criterion for nonlinear input–output differential systems, *SIAM J. Control Optim.* **21** (1983) 721–728.

[2] M. Fliess and C. Reutenauer, Une application de l'algebra differentièlle aux systèmes reguliers (ou bilinéaires), in: A. Bensoussan and J.L. Lions, Eds., *Analysis and Optimization of Systems*, Lect. Notes. Control Informat. Sci. No. 44 (Springer-Verlag, Berlin–New York, 1982) 99–107.

[3] A. Isidori, *Nonlinear Control Systems: An Introduction* (Springer, Berlin–New York, 1985).

[4] F. Lamnabhi-Lagarrigue and P.E. Crouch, A formula for iterated derivatives along trajectories of nonlinear systems, *Systems Control Lett.* **11** (1988) 1–7.

[5] J.T. Lo, Global bilinearization of systems with controls appearing linearly, *SIAM J. Control* **13** (1975) 879–885.

[6] M. Lothaire, Combinatorics on words, in: G.C. Rota, Ed., *Encyclopedia of Mathematics and its Applications*, Vol. 2 (Addison-Wesley, Reading, MA, 1983).

[7] E.D. Sontag, *Polynomial Response Maps* (Springer, Berlin–New York, 1979).

[8] E.D. Sontag, Bilinear realizability is equivalent to existence of a singular affine differential i/o equation, *Systems Control Lett.* **11** (1988) 181–187.

[9] Y. Wang and E.D. Sontag, A new result on the relation between differential-algebraic equation and state space realizations, in: *Proceedings of Conference on Information Sciences and Systems* (Johns Hopkins University, March 1989) 143–147.