I/O Equations for Nonlinear Systems and Observation Spaces

Eduardo D. Sontag*
Department of Mathematics, Rutgers University, New Brunswick, NJ 08903
Yuan Wang
Department of Mathematics, Florida Atlantic University, Boca Raton, FL 33431
Emails: sontag@hilbert.rutgers.edu, y_wang@acc.fau.edu

Abstract

This paper studies various types of input/output representations for nonlinear continuous time systems. The algebraic and analytic i/o equations studied in previous papers by the authors are generalized to integral and integro-differential equations, and an abstract notion is also considered. New results are given on generic observability, and these results are then applied to give conditions under which that the minimal order of an equation equals the minimal possible dimension of a realization, just as with linear systems but in contrast to the discrete time nonlinear theory.

1. Introduction

Previous papers by the authors studied the relationships between realizability of continuous-time operators defined by generating series and the existence of algebraic or analytic differential equations relating inputs and outputs. In linear systems identification practice, it is often the case that one prefilters measured signals prior to estimation, so as to eliminate noise. Various filtering procedures have been suggested in the nonlinear case; see for instance the paper [8] and other references by the same author. This motivates the question of establishing if i/o operators satisfy integral equations, or even more general types of integral/differential equations, and we study this issue. We show that such equations still imply realizability, but they are not necessary (as differential equations are). On the other hand, we present an abstract notion of equation, based on subanalytic set theory, on the basis of which a general result is indeed possible.

A question that we had not studied before is that of the *degree* of an i/o equation in comparison with the minimal possible dimension of a realization. In discrete-time, it was known for a long time (see [10]) that these numbers are in general distinct. It turns out, perhaps surprisingly, that the numbers do coincide in the continuous time case: we prove that if there is a minimal realization of dimension n then no i/o equation can have degree less than n. The proof relies on yet another characterization of observation spaces, this one different from those obtained earlier, and is based on the "universal inputs theorem" proved originally in [9] (that result, valid only on compacts, was later strengthened considerably in [14], but the weaker version suffices for our purposes). As a side benefit from the proof, one makes contact with a notion of observability proposed by Fliess, and we show that the latter coincides with the standard observability rank condition.

Some proofs are omitted for lack of space; a Technical Report with details is available.

The *i/o operators* we will deal with are those defined by convergent generating series. For definitions and properties of such i/o operators, we refer the reader to the Appendix or [18] and [19]. An important property of such an operator F is that for any analytic input u, the output $F[u]$ is also analytic.

2. Generalized Rational I/O Equations and Realizability

In this section, we shall study realizability of i/o operators. As in [19], we say an i/o operator F is *locally realizable* by an analytic system if there exist some analytic manifold $\mathcal{M}$, some $x_0 \in \mathcal{M}$, $(m+1)$ analytic vector fields $g_0, g_1, \ldots, g_m$ on $\mathcal{M}$, an analytic function $h : \mathcal{M} \to \mathbb{R}$, and some $\delta > 0$ such that for each input u with $\|u\|_\infty < 1$, $y(t) = h(x(t))$ where $x(\cdot)$ is the solution of the equations

$$x' = g_0(x) + \sum_{j=1}^{m} g_j(x)\, u_j\,, \quad x(0) = x_0\,.$$

We now study a different class of operators mapping $\mathcal{C}^\omega[0,T] \times \mathcal{C}^\omega[0,T]$ to $\mathcal{C}^\omega[0,T]$, for any fixed $T > 0$. To avoid possible confusion, we shall call such an operator a *GDIO* (*generalized differential integral operator*). For GDIO's P_1, P_2, we define $P_1 + P_2$ and $P_1 \cdot P_2$ by pointwise operations:

$$(P_1 + P_2)(\phi,\, \psi)(t) = P_1(\phi,\, \psi)(t) + P_2(\phi,\, \psi)(t),$$

$$(P_1 \cdot P_2)(\phi,\, \psi)(t) = P_1(\phi,\, \psi)(t) \cdot P_2(\phi,\, \psi)(t),$$

for all $\phi, \psi \in \mathcal{C}^\omega[0,T]$. The integral $\int P$ of an operator P is defined in the following way:

$$\left(\int P\right)(\phi,\, \psi)(t) = \int_0^t P(\phi,\, \psi)(s)\, ds, \quad \text{for } 0 \le t \le T.$$

In this work, we are interested in a special class $\mathcal{P}$ of GDIO's . To define $\mathcal{P}$, we first need to define $\mathcal{P}_k$ for each fixed nonnegative integer k. For a given k, $\mathcal{P}_k$ is defined to be the smallest set of GDIO's satisfying the following properties:

a. the derivation operators $e_i : (\phi,\, \psi) \mapsto \phi^{(i)}$ and $d_i : (\phi,\, \psi) \mapsto \psi^{(i)}$ are all in $\mathcal{P}_k$, for each $i = 0, 1, \ldots, k$,

*This research supported in part by grant AFOSR-88-0235

b. $\int P \in \mathcal{P}_k$ if $P \in \mathcal{P}_k$,

c. $P_1 + P_2,\ P_1 \cdot P_2 \in P_k$ if $P_1, P_2 \in \mathcal{P}_k$, and $cP \in P_k$ for any $c \in \mathbb{R}$ if $P \in P_k$,

d. the "constant" GDIO $(\phi, \psi) \mapsto 1\ \in \mathcal{P}_k$.

For instance, $\int_0^t \int_0^s \phi(s)\psi(\tau)\, d\tau\, ds \ \in\ \mathcal{P}_0$ and $\phi'''(t) \cdot \int_0^t \psi(s)\, ds\ \in\ \mathcal{P}_3$. It can be easily seen that P_k is in fact the smallest $\mathbb{R}$-algebra satisfying properties a and b.

We define $\mathcal{P}$ to be the union of all $\mathcal{P}_k$'s, i.e, $\mathcal{P} := \bigcup_{k=0}^{\infty} \mathcal{P}_k$.

Now we turn back to i/o operators for control systems. An i/o operator F is said to satisfy a *generalized rational i/o equation* if there exist some integer k and two operators P_0, P_1, both belonging to $\mathcal{P}_r$ for some $r < k$, such that for every analytic i/o pair of $(u,\, y)$ of F,

$$P_0(u,\, y)(t)y^{(k)}(t) = P_1(u,\, y)(t)\,, \tag{1}$$

and, furthermore, the following nondegeneracy property holds: $P_0(u,\, y) \neq 0$ for some analytic i/o pair $(u,\, y)$ of F.

Roughly speaking, the i/o operator F satisfies an equation of this type if the k-th derivative of the output, for some k, can be expressed rationally in terms of lower-order derivatives as well as integrals of inputs and outputs.

The following result is proved using techniques that generalize those employed in [18] and [19]:

Theorem 1 *If an operator F satisfies a generalized rational i/o equation, then it is locally realizable by an analytic system.* ∎

In identification, it is often the case that one prefilters signals in order to diminish the effects of noise. This sometimes can be modeled through the use of the following type of integral i/o equation:

$$E\,(u,\, u_1,\, \ldots,\, u_k y,\, y_1,\, \ldots,\, y_k) = 0,$$

where E is a polynomial function, or more generally, an analytic function. (There may also be more complicated integrals, depending on filter parameters.)

If an operator F satisfies an integral i/o equation, then it satisfies a generalized rational i/o equation. Thus we conclude, in particular:

Corollary 2.1 If an operator F satisfies an integral i/o equation, then F is locally realizable. □

In contrast to the case of differential i/o equations (in which existence of an algebraic *differential* i/o equation is equivalent to realizability by a "rational" system), it is generally not true that every operator which is realizable by a polynomial system satisfies an algebraic *integral* i/o equation. As an illustration, consider the operator

$$y(t) = F[u](t) = \exp\left(\int_0^t u(s)\, ds\right)$$

which is realized by the system:

$$x' = xu, \quad y = x, \quad x_0 = 1. \tag{2}$$

It can be seen that F satisfies no integral equation (see [17]). Note that even though F does not admit any integral equation, F does satisfy the *differential* equation $y' = yu$. This illustrates that the existence of differential equations and the existence of integral equations are not equivalent, unless much stronger conditions (such as linearity of the equations, as for linear systems) are assumed.

3. A Necessary and Sufficient Condition for Local Realizability

It has been shown that an i/o operator is locally realizable by an analytic system if it satisfies an analytic differential i/o equation (cf [19]). However, the converse fact, that any operator realizable by an analytic sytem satisfies some analytic differential i/o equation, is in general false. In this section, we provide an necessary *and* sufficient condition for realizability. The techniques used in this section are based on the subanalytic set theory developed in [15] and [4]. Throughout this section, by "analytic submanifold" we mean analytic *embedded* submanifold.

Definition 3.2 Assume that M is an analytic manifold. A subset S of M is said to be a *stratified union* of a family (whose members are called strata) $\mathcal{T}$ of connected analytic submanifolds if the following properties hold:

1. if $T \in \mathcal{T}$ is a stratum, then the closure $\bar{T}$ of T contains any strata that intersect $\bar{T}$,

2. for any stratum T_1, if a stratum $T_2 \neq T_1$ is a subset of $\bar{T}_1$, then codim $T_2 >$ codim T_1. ∎

We shall say that a subset S of an analytic manifold M is an *analytically thin* subset if S is a locally finite stratified union of analytic connected submanifolds of codimension at least 1.

Definition 3.3 An operator F is said to satisfy an *analytic constraint* if there exist an integer k, an analytic thin subset S of $\mathbb{R}^{2k+1}$, and some $\delta > 0$, such that for each i/o pair (u, y) with $\|u\|_\infty < 1$ and each $t < \delta$,

$$\left(u(t), u'(t), \ldots, u^{(k-1)}(t), y(t), y'(t), \ldots, y^{(k)}(t)\right) \in S$$

provided $u^{(k-1)}(t)$ exists. □

The following is the main result of this Section.

Theorem 2 *An i/o operator F is locally realizable if and only if F satisfies an analytic constraint.*

Sketch of the Proof. Necessity: Assume that F is realizable by a system $\Sigma = (\mathcal{M}, (g_0, \ldots, g_m), x_0, h)$ of dimension n. Take a compact neighborhood $\mathcal{U}$ of x_0 which is subanalytic. (For instance, one may assume $x_0 \in \mathbb{R}^n$ and choose $\mathcal{U}$ to be a ball $\{x : \|x - x_0\| \leq r\}$ for some small r.) Define a map $\phi : \mathcal{M} \times \mathbb{R}^{mn} \rightarrow \mathbb{R}^{n+mn+1}$ by

$$\phi : (x, \mu_0, \mu_1, \ldots, \mu_{n-1}) \mapsto (\mu_0, \mu_1, \ldots, \mu_{n-1}, h(x), y_1(x), \ldots, y_n(x))$$

where each $\mu_i \in \mathbb{R}^m$, $y_i(x) = \left.\dfrac{d^i}{dt^i}\right|_{t=0} h(x(t))$, $x(t)$ is the solution of the equations $x' = g_0(x) + \sum_{i=1}^m g_i(x)u_i(t)$, $x(0) = x$, and u is any control with the initial values $u^{(i)}(0) = \mu_i$.

Clearly, ϕ is an analytic map defined on a subanalytic set. Also, it is not hard to see that ϕ is proper on $\mathcal{U} \times \mathbb{R}^{mn}$. It then follows that the image $\mathcal{W}$ of ϕ is a subanalytic set, since $\mathcal{U} \times \mathbb{R}^n$ is a subanalytic subset of $\mathcal{M} \times \mathbb{R}^{mn}$ (cf [15] and [4]). Again by Theorem 3 in [15], one knows that there is an analytic stratification of $\mathcal{M}$ so that $\mathcal{W}$ is a union of some strata of the stratification. Notice that the preimage of $\mathcal{W}$ is $2n$-dimensional. By Sard's Theorem (cf [1]), one knows that none of the strata contained in $\mathcal{W}$ can have codimension 0. Thus we conclude that that $\mathcal{W}$ is an analytically thin set.

Now we consider the initialized system $\Sigma = (\mathcal{M}, (g_0, \ldots, g_m), x_0, h)$. For each input u with $\|u\|_\infty < 1$, there exist some $\delta > 0$ such that $x(t) \in \mathcal{U}$ for $t < \delta$. It can easily be proved that for each fixed $\tau < \delta$,

$$(u(\tau), u'(\tau), \ldots, u^{(n-1)}(\tau), y(\tau), \ldots, y^{(n)}(\tau)) = \phi(x(\tau), u(\tau), u'(\tau), \ldots, u^{(n-1)}(\tau)).$$

It then follows that

$$(u(\tau), u'(\tau), \ldots, u^{(n-1)}(\tau), y(\tau), \ldots, y^{(n)}(\tau)) \in \mathcal{W},$$

for any $\tau < \delta$, that is, F satisfies an analytic constraint.

Sufficiency: Assume F satisfies an analytic constraint, i.e., there exist some k and an analytic thin set $S \in \mathbb{R}^{2k+1}$ such that for every i/o pair (u, y) of F

$$\xi(t) := \Big(u(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k)}(t)\Big) \in S$$

for t small. Let $\mathcal{T}$ denote the family of the strata that compose S and $\rho = \min\{\operatorname{codim} T : T \in \mathcal{T}\}$. Then one can show that for any $T \in \mathcal{T}$ with $\operatorname{codim} T = \rho$, T is open relative to S.

Assume now that there exists some u so that $\xi(0) \in T$ for some T such that $\operatorname{codim} T = r \geq 1$. Then there exists some neighborhood $\mathcal{U}$ of $\xi(0)$ in $\mathbb{R}^{2k+1}$ and at least one analytic function ϕ such that $T \cap \mathcal{U} = \{q : \phi(q) = 0\}$. Since T is open relative to S, the operator F "*locally*" satisfies the analytic differential i/o equation

$$\phi(u(t), u'(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k)}(t)) = 0.$$

By using methods similar to those used in [19], one concludes that F is realizable.

We are still left with the case in which for any u bounded by 1, $\xi(0) \in T$ for some T with $\operatorname{codim} T > \rho$. In this case, one can always find sequences $\{u_j\}$ and $\tau_j \rightarrow 0$ such that, for each j, $\xi(\tau_j) \in T$ for some stratum T with $\operatorname{codim} T = \rho$. Define F_j by $F_j[u](t) := F[u_j \#_{\tau_j} u](\tau_j + t)$ for each j, where $v \#_\tau u$ denotes the function obtained by concatenating u to v at time τ. Then F_j approaches F in a suitable sense. By the above argument, F_j is realizable for each j. It then can be shown, using the fact that the Lie rank of F_j depends on j lower semi-continuously (cf [19]), that F is realizable. ∎

4. Observation Spaces and Observation Fields

In this section, we study relationships between alternative definitions of observation space and observation field, and we draw conclusions about the degree of i/o equations.

Again, we are only dealing with single input systems to make the notations simpler. Consider an *analytic system*:

$$\Sigma : \left\{ \begin{array}{rcl} x' & = & g_0(x) + g_1(x)u\,, \\ y & = & h(x)\,, \end{array} \right. \tag{3}$$

where $x \in \mathcal{M}$, an analytic manifold of dimension n, h is an analytic function defined on $\mathcal{M}$, and f, g are analytic vector fields defined on $\mathcal{M}$. If $\mathcal{M} = \mathbb{R}^n$, and the entries of g_0, g_1, and h are rational functions with no poles, then we call (3) a *rational system*.

Let $\mathcal{F}$ be the space of functions $\mathcal{M} \longrightarrow \mathbb{R}$ spanned by the Lie derivatives of h in the directions of g_0 and g_1, i.e.,

$$\mathcal{F} := \operatorname{span}_{\mathbb{R}} \left\{L_{g_{i_1}} L_{g_{i_2}} \cdots L_{g_{i_r}} h : \ r \geq 0,\ i_j = 0, 1\right\}.$$

This is the *observation space* associated to (3); see e.g. [12], Remark 5.4.2. Associated to this, let $\mathcal{O}$ be the subspace of the cotangent space defined by $\mathcal{O} := \operatorname{span}_{\mathbb{R}_x}\{d\phi : \ \phi \in \mathcal{F}\}$, where $\mathbb{R}_x$ is the field of meromorphic functions from $\mathcal{M}$ to $\mathbb{R}$.

An alternative characterization is as follows. In general, let $\phi(t, x, \omega)$ denote the state trajectory of (3) corresponding to a control ω and initial state x, defined for small t. Now for any integer k and any $\mu = (\mu_0, \ldots, \mu_{k-1}) \in \mathbb{R}^k$, we define

$$\psi_i(x, \mu) = \left.\frac{d^i}{dt^i}\right|_{t=0} h(\phi(t, x, u)) \tag{4}$$

for $0 \leq i \leq k$, where u is any control with initial values $u^{(i)}(0) = \mu_i$. The functions $\psi_i(x, \mu)$ can be expressed, -applying repeatedly the chain rule,- as polynomials in $(\mu_0, \ldots, \mu_{k-1})$ whose coefficients are analytic functions (rational functions if the system is rational) of x. For each fixed $(\mu_0, \ldots, \mu_{k-1}) \in \mathbb{R}^k$, $\psi_i(x, \mu)$ is analytic in

x. Let $\hat{\mathcal{F}} = \text{span}_{\mathbb{R}}\{\psi_k(x,\mu):\ \mu \in \mathbb{R}^k, k \geq 0\}$. The main result in [20] is that $\mathcal{F} = \hat{\mathcal{F}}$. This equality is fundamental in establishing results linking realizability to the existence of i/o equations, in [18] and [19].

A different object is obtained if one instead views the elements $\psi_i(x,\mu)$ as formal polynomials on the $\mu_0, \ldots, \mu_{k-1}$'s whose coefficients are functions. That is, let $K = \mathbb{R}(U_0, U_1, \ldots)$ be the field obtained by adjoining indeterminates $U_0, U_1, \ldots$ to $\mathbb{R}$, and let $\mathcal{F}^K$ be defined as the subspace of K_x spanned by the functions ψ_k over the field K, i.e., $\mathcal{F}^K := \text{span}_K\{\psi_i:\ i \geq 0\}$. Here $K_x = \mathbb{R}_x(U_0, U_1, \ldots)$ is the field obtained by adjoining the indeterminates $U_0, U_1, \ldots$ to $\mathbb{R}_x$, the field of meromorphic functions on $\mathcal{M}$, seen as a vector space over K. One can see the differentials of elements of K_x as rational functions in $U_0, U_1, \ldots$, whose coefficients are covector fields. Then we let $\mathcal{O}^K := \text{span}_{K_x}\{d\phi:\ \phi \in \mathcal{F}^K\}$.

The following is the main result in this Section:

Theorem 3 *For the analytic system (3),* $\dim_{\mathbb{R}_x}\mathcal{O} = \dim_{K_x}\mathcal{O}^K$.

Remark 4.4 For each $q \in \mathcal{M}$, let $\mathcal{O}(q)$ be the space obtained by evaluating the elements of $\mathcal{O}$ at q, i.e., $\mathcal{O}(q) := \text{span}_{\mathbb{R}}\{v(q):\ v \in \mathcal{O}\}$. Then Theorem 3 says that the generic rank of $\mathcal{O}(q)$ is the same as the rank of $\mathcal{O}^K$, that is, $\max_{q\in\mathcal{M}} \text{rank}_{\mathbb{R}}\mathcal{O}(q) = \text{rank}_{K_x}\mathcal{O}^K$ □

There are several immediate consequences of Theorem 3:

Corollary 4.5 For a bilinear system, $\dim_{\mathbb{R}}\mathcal{F} = \dim_K \mathcal{F}^K$.

Proof. Assume that in system (3), $g_0(x) = A_0x$, $g_1(x) = A_1x$, $h(x) = Cx$, where A_0, A_1 and C are matrices of suitable sizes. Then for each multiindex $i_1 i_2 \cdots i_r$,

$$L_{g_{i_1}} L_{g_{i_2}} \cdots L_{g_{i_r}} h(x) = C A_{i_r} A_{i_{r-1}} \cdots A_{i_1} x\,,$$

and the ψ's are also linear in x; for instance

$$\psi_2(x, \mu_0, \mu_1) = C(A_0 + \mu_0 A_1)^2 x + \mu_1 C A_1 x\,.$$

It follows that, in this case, $\dim_{\mathbb{R}}\mathcal{F} = \dim_{\mathbb{R}_x}\mathcal{O}$ and $\dim_K \mathcal{F}^K = \dim_{K_x}\mathcal{O}^K$. By Theorem 3, $\dim_{\mathbb{R}}\mathcal{F} = \dim_K \mathcal{F}^K$. ∎

Assume now that (3) is a rational system. Let $\mathcal{A}$ (respectively $\mathcal{A}^K$) be the $\mathbb{R}$-algebra (respectively K-algebra) generated by the elements of $\mathcal{F}$ (respectively $\mathcal{F}^K$). Define the observation field $\mathcal{Q}$ (respectively $\mathcal{Q}^K$) as the quotient field of $\mathcal{A}$ (respectively $\mathcal{A}^K$). Then we have the following:

Corollary 4.6 If system (3) is rational, then

$$\text{trdeg}_{\mathbb{R}}\mathcal{Q} = \text{trdeg}_K \mathcal{Q}^K\,, \tag{5}$$

where trdeg denotes transcendence degree.

Proof. Applying Theorem III.7-III of [5], one knows that

$$\text{trdeg}_{\mathbb{R}}\mathcal{Q} = \dim_{\mathbb{R}_x}\mathcal{O},\quad \text{trdeg}_K \mathcal{Q}^K = \dim_{K_x}\mathcal{O}^K\,.$$

By Theorem 3, one immediately obtains (5). ∎

Consider an analytic system (3). Fix any two states $p, q \in \mathcal{M}$ and take an input u. We say p and q are *distinguished by* u, denoted by $p \not\sim_u q$, if $h(\phi(\cdot, p, u)) \neq h(\phi(\cdot, q, u))$ (considered as functions defined on the common domain of $\phi(\cdot, p, u)$ and $\phi(\cdot, q, u)$); otherwise we say p and q cannot be distinguished by u, denoted by $p \sim_u q$. If p and q cannot be distinguished by *any* input u, then we say p and q are *indistinguishable*, denoted by $p \sim q$. If for any two states, $p \sim q$ implies $p = q$, then we that say system (3) is observable. (See [12], Chapter 5; note however that indistinguishability by a control is being defined here in a slightly different way.)

Take an open set U and any two points $p, q \in U$. If for every input u, $h(\phi(t, p, u)) = h(\phi(t, q, u))$ for each t for which $\phi(T, p, u)$ and $\phi(T, q, u)$ are both defined and in U for all $0 \leq t \leq T$, then we say that p and q are *U-indistinguishable* (see e.g. [11]).

Fix a point p. If for every neighborhood U_p there is a neighborhood $V_p \subset U_p$ so that for any $q \in V_p$, the condition that q and p are U_p-indistinguishable implies $p = q$, then we say the system (3) is *locally observable at* p. If (3) is locally observable at every point p, then we say (3) is *locally observable*. If there is an open dense set $U \subset \mathcal{M}$ such that (3) is locally observable at every point p of U, then we say (3) is *generically locally observable*. See [11] for details on local observability and related concepts such as the slightly different definition in [7].

Proposition 4.7 The analytic system (3) is generically locally observable if and only if $\dim_{K_x}\mathcal{O}^K = n$.

Proof. By Lemma 2.10 and facts (2.4) and (2.8) in [11], one knows that (3) is generically locally observable if and only if that the generic rank of $\mathcal{O}(q)$ is n. Thus Proposition 4.7 follows immediately from Theorem 1. ∎

Corollary 4.8 A polynomial system (3) is observable in the sense of [2] if and only if the system is generically locally observable.

Proof. It is shown in [2] that a polynomial system (3) is "observable" in the sense of [2] if and only if $\dim_{K_x}\mathcal{O} = n$. The Corollary then immediately follow from Proposition 4.7. ∎

Assume now that an operator F satisfies a differential i/o equation

$$E(u'(t),\ \ldots,\ u^{(k-1)}(t),\ y(t),\ y'(t),\ \ldots,\ y^{(k)}(t)) = 0\,. \tag{6}$$

Then the *order* of the equation (6) is defined to be the highest $r \leq k$ such that

$$\frac{\partial}{\partial \nu_r} E(\mu_0,\ \ldots,\ \mu_{k-1},\ \nu_0,\ \nu_1,\ \ldots,\ \nu_k)$$

is not a zero function.

For a given operator F, we define $\delta(F)$ to be the lowest possible order of an i/o equation for F. In case that there is no i/o equation for F, $\delta(F)$ is defined to be ∞.

Let $\lambda(F)$ be the Lie rank of F. It has been known that if F is realizable, then the dimension of a canonical realization for F is $\lambda(F)$, cf [3] and [13]. Here, by a canonical realization we mean a realization by an accessible and generically local observable system.

Theorem 4 *Assume F is an i/o operator. Then:*

(a) $\lambda(F) \le \delta(F)$;

(b) if there exists a rational canonical realization for F, then $\lambda(F) = \delta(F)$.

Proof. It was shown in [19] that if $\delta(F) < \infty$, then $\lambda(F) < \infty$. Thus we may assume that $\lambda(F) < \infty$. Then F is realizable by some canonical system $\Sigma = (\mathcal{M},\, x_0,\; g_0,\, g_1,\, h)$, whose dimension must be $n = \lambda(F)$.

Now assume F admits some equation (6) of order k. Assume $k < n$. Let $\mathcal{R}_{x_0}$ be the reachable set of Σ from x_0. It can be seen that for any $x \in \mathcal{R}_{x_0}$,

$$E(\mu_0,\, \dots,\, \mu_{(r-1)},\, \psi_0(x),\; \dots,\; \psi_r(x,\, \mu_0, \dots, \mu_{r-1})) = 0, \tag{7}$$

where the functions ψ_i are defined as in (4). Notice here that E depends on x analytically, and $\mathcal{R}_{x_0}$ has a nonempty interior. We conclude that (7) holds *for all* $x \in \mathcal{M}$. By the definition of k, one concludes immediately that $d\psi_k \in \operatorname{span}_{K_x}\{d\psi_0,\; d\psi_1,\; \dots,\; d\psi_{k-1}\} = \mathcal{O}_k^K$. By differentiating (6) and using the same argument repeatedly, one can show that

$$d\psi_r \in \operatorname{span}_{K_x}\{d\psi_0,\; d\psi_1,\; \dots,\; d\psi_{k-1}\} = \mathcal{O}_k^K\,,$$

for each $k \ge r$. Thus one knows that

$$\dim_{K_x} \mathcal{O}^K = \dim_{K_x} \mathcal{O}_r^K \le k < n\,. \tag{8}$$

On the other hand, from the observability of the system, we know that $\dim_{\mathbb{R}_x} \mathcal{O} = n$, which, by Theorem 3, implies that $\dim_{K_x} \mathcal{O}^K = n$. The contradiction between (8) and this shows that it is impossible for F to admit any i/o equation of order lower than n. Part (a) of the Theorem is proved.

Part (b) of the Theorem follows from the fact that if F is realizable by a rational system of dimension n, then it admits an algebraic i/o equation of order n, cf [18]. ∎

Remark 4.9 The result in Theorem 3 is *false* in general for discrete-time systems, as discussed in [10]. As a consequence of this, also the Corollaries and Theorem 4 are false in that case. To illustrate this, consider the system of dimension 3

$$\begin{aligned}
&x_1(t+1) = u(t)\,, \quad x_2(t+1) = x_3(t)\,,\\
&x_3(t+1) = x_3(t)x_1(t) + x_1(t) + x_2(t)u(t)\,,\\
&y(t) = x_3(t),
\end{aligned}$$

with initial state $x_1(0) = x_2(0) = x_3(0) = 0$, and the operator F it defines. This system is generically observable and is also accessible. However, F satisfies an equation of order 2:

$$y(t) = y(t-1)u(t-2) + y(t-2)u(t-1) + u(t-1). \blacksquare$$

Now we return to the proof of Theorem 3. To prove the Theorem, we need the following result. It is a result that was basically given in [9], with essentially the same proof (a stronger statement, that generic inputs satisfy the property, was later obtained in [14]).

Lemma 4.10 Consider an analytic system (3). For any fixed compact set $\Omega \subset \mathcal{M}$, there exists some analytic input w so that for any $p, q \in K$, if $p \sim_w q$, then $p \sim q$.

Sketch of Proof: Fix k. For each $\mu = (\mu_0,\, \dots, \mu_{k-1}) \in \mathbb{R}^k$, consider the set

$$\Delta_\mu := \{(p,\, q) \in \mathcal{M}\times\mathcal{M}:\; \psi_i(p,\, \mu) = \psi_i(q,\, \mu),\; 0 \le i \le k\}.$$

Then Δ_μ is an analytic set for each μ. By the argument used in the proof of Theorem 4.8 of [9], one can prove that there exists a minimal set in the family of sets $\{\Delta_\mu \cap \Omega\times\Omega:\;\; \mu \in \mathbb{R}^k,\;\; k \ge 0\}$. Let $\bar\mu$ be such that the set $\Delta_{\bar\mu}$ is minimal.

Claim: Any input u with the initial values $u^{(i)}(0) = \bar\mu_i,\; 0 \le i \le r$ (where r is the length of $\bar\mu$) has the desired property. Assume not. Then there exists some $p,\, q \in \Omega$ such that $\psi_i(p,\, \bar\mu) = \psi_i(q,\, \bar\mu)$ for $0 \le i \le r$, but $p \not\sim_u q$ for some u. Without loss of generality, one may assume that u is analytic and is defined on some finite interval $[0, \tau]$. Using analogous techniques to those employed in the proof of Lemma 4.1 in [19], one shows that there exists a sequence of analytic functions $\{v_j\}_{j\ge0}$ such that $\|v_j - u\|_1 \to 0$ as $j \to \infty$, and $v_j^{(i)}(0) = \bar\mu_i$ for $0 \le i \le r-1$. It then follows that $p \not\sim_{v_j} q$ for j large enough. Let v be one of the v_j's that distinguishes p and q. Then there exists some s such that $\psi_s(p, \nu) \ne \psi_s(q, \nu)$, where $\nu = (v(0),\, v'(0),\, \dots,\, v^{(s-1)}(0))$.

Since $v^{(i)}(0) = \bar\mu_i$ for $0 \le i \le r-1$, it follows that $\Delta_\nu \subset \Delta_{\bar\mu}$. However, since $(p, q) \in \Delta_{\bar\mu}$ but $(p,\, q) \notin \Delta_\nu$, one obtains the conclusion that $\Delta_\nu \cap \Omega \times \Omega$ is a proper subset of $\Delta_{\bar\mu} \cap \Omega \times \Omega$, which contradicts the minimality of $\Delta_{\bar\mu} \cap \Omega \times \Omega$. Thus for any control u with $u^{(i)}(0) = \bar\mu_i$ and any $p, q \in \Omega$, the relation $p \sim_u q$ implies $p \sim q$. ∎

Sketch of Proof of Theorem 3: It is true that $\dim_{K_x} \mathcal{O}^K \le \dim_{\mathbb{R}_x} \mathcal{O}$, since $\mathcal{F}^K \subseteq \operatorname{span}_K\{\phi:\; f \in \mathcal{F}\}$. Thus it is enough to show that

$$\dim_{K_x} \mathcal{O}^K \ge \dim_{\mathbb{R}_x} \mathcal{O}\,. \tag{9}$$

Assume now $\dim_{\mathbb{R}_x} \mathcal{O} = r$. Choose a point p_0 so that $\dim_{\mathbb{R}} \mathcal{O}(p_0) = r$. Then there exists a neighborhood U of p_0 contained in some compact set Ω such that for some suitable choice of coordinates $x = (x_1,\, x_2)$ with $x_1 \in \mathbb{R}^r,\; x_2 \in \mathbb{R}^{n-r}$, the system (3) takes the form

$$\begin{aligned}
x_1' &= g_{01}(x_1) + g_{11}(x_1)u, \quad y = h_1(x_1)\,,\\
x_2' &= g_{02}(x_1,\, x_2) + g_{12}(x_1,\, x_2)\,,
\end{aligned}$$

with the property that the rank $_{\mathbb{R}}\mathcal{O}_1 = r$, where $\mathcal{O}_1 = \text{span}_{\mathbb{R}}\{h_1, L_{g_{01}}h_1, L_{g_{11}}h_1, L_{g_{01}}L_{g_{01}}h_1, \ldots\}$. It then follows that any two points p, q in U is distinguishable if and only if $x_1^p \neq x_1^q$. By Lemma 4.10, there exists some analytic w so that $p \not\sim_w q$ if and only if $x_1^p \neq x_1^q$.

Let $\mu_i = w^{(i)}(0)$. By the techniques used in the proof of Lemma 4.10, one can show that there exists some integer k such that the condition $\psi_i(x, , \mu) = \psi_i(z, \mu)$ for $0 \le i \le k$ implies that $x \sim_w z$. Define Let $\Psi(x_1, x_2) := (\psi_0(x_1, x_2, \mu), \psi_1(x_1, x_2 \mu) \ldots, \psi_k(x_1, x_2, \mu))$. Then $\Psi(x_1, x_2) \neq \Psi(z_1, z_2)$ if and only if $x_1 \neq z_1$. It follows immediately that generically rank $J_\Psi = r$, where J_Ψ denotes the Jacobian of Ψ, which implies that $\dim_{\mathbb{R}_x} \mathcal{O}_\mu \ge r$, where $\mathcal{O}_\mu$ is the subspace of $\mathbb{R}_x$ obtained by evaluating U_i at μ_i for the elements of $\mathcal{O}^K$. Equation (9) is then proved by noticing that $\dim_{K_x} \mathcal{O}^K \ge \dim_{\mathbb{R}^x} \mathcal{O}_w \ge r$. ∎

5. References

[1] T. Bröcker and L. Lander, *Differentiable Germs and Catastrophes*, vol. 17 of London Mathematical Society Lecture Note Series, Cambridge University Press, New York, 1975.

[2] S. Diop and M. Fliess, *On nonlinear observability.* To appear.

[3] M. Fliess, *Réalisation locale des systèmes non linéaires, algèbres de Lie filtrées transitives et séries génératrices non commutatives*, Invent. Math., 71 (1983), pp. 521–537.

[4] R. M. Hardt, *Stratification of real analytic mappings and images*, Inventiones Math, 28 (1975), pp. 193–208.

[5] W. Hodge and D. Pedoe, *Methods of Algebraic Geometry,*, vol. I, Cambridge University Press, Cambridge, 1968.

[6] A. Isidori, *Nonlinear Control Systems*, Springer-Verlag, Berlin, second ed., 1989.

[7] H. Nijmeijer and A. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.

[8] A. E. Pearson, *Least squares parameter identification of nonlinear differential i/o models*, in Proceedings of the 27th Conference on Decision and Control, Austin, Texas, 1988, pp. 1831–1835.

[9] E. D. Sontag, *On the observability of polynomial systems, I: Finite-time problems*, SIAM J. Control & Optimization, 17 (1979), pp. 139–151.

[10] ———, *Polynomial Response Maps*, Springer-Verlag, Berlin-NY, 1979.

[11] ———, *A concept of local observability*, Systems & Control Letters, 5 (1984).

[12] ———, *Mathematical Control Theory, Deterministic Finite Dimensional Systems*, Springer-Verlag, New York, 1990.

[13] H. J. Sussmann, *A proof of the realization theorem for convergent generating series of finite lie rank.* Submitted.

[14] ———, *Single-input observability of continuous-time systems*, Mathematical Systems Theory, 12 (1979), pp. 371–393.

[15] ———, *Subanalytic sets and feedack control*, J. Diff. Eqs., 31 (1979), pp. 31–52.

[16] ———, *Lie brackets and real analyticity in control theory*, in Mathematical Control Theory, (C. Olech, B. Jakubczyk, and J. Zabczyk, eds.), PWN-Polish Scientific Publishers, Warszawa, 1985, pp. 515–542.

[17] Y. Wang, *Integral i/o equations and realizability*, in Proceedings of the Conference on Information Sciences and Systems, 1991, pp. 719–723.

[18] Y. Wang and E. D. Sontag, *Algebraic differential equations and rational control systems.* To appear in *SIAM J. Control and Opt.*

[19] ———, *Generating series and nonlinear systems: analytic aspects, local realizability, and i/o representations.* To appear in *Forum Math.*

[20] ———, *On two definitions of observation spaces*, Systems and Control Letters, 13 (1989), pp. 279–289.

Appendix

In this appendix, we provide basic definitions and properties of i/o operators defined by convergent generating series.

A *generating series* $c = \sum_{\iota \in I^*} \langle c, \eta_\iota \rangle \eta_\iota$, is a formal power series in the noncommutative variables $\eta_0, \eta_1, \ldots, \eta_m$ for some fixed number m, where we use the notation $\eta_\iota = \eta_{i_1}\eta_{i_2}\cdots\eta_{i_l}$ for each multiindex $\iota = i_1 i_2 \cdots i_l$. The coefficients $\langle c, \eta_\iota \rangle$ are assumed to be real.

We shall say that a power series c is *convergent* if there exist $K, M \ge 0$ such that

$$|\langle c, \eta_\iota \rangle| \le K M^k k! \quad \text{for each } \iota \in I^k, \text{ and each } k \ge 0. \tag{10}$$

For any fixed real number $T > 0$, let $\mathcal{U}_T$ be the set of all essentially bounded measurable functions $u : [0, T] \to \mathbb{R}^m$ endowed with the L^1 norm. We write $\|u\|_1$ for $\max\{\|u_i\|_1 :, 1 \le i \le m\}$ and $\|u\|_\infty$ for $\max\{\|u_i\|_\infty :, 1 \le i \le m\}$ where u_i is the i-th component of u, and $\|u_i\|_1$ is the L^1 norm of u_i, $\|u_i\|_\infty$ is the L^∞ norm of u_i. For each $u \in \mathcal{U}_T$ and each $\iota \in I^l$, we define inductively the functions $V_\iota = V_\iota[u] \in \mathcal{C}[0, T]$ by $V_{i_1\cdots i_{l+1}}[u](t) = \int_0^t u_{i_1}(s) V_{i_2\cdots i_{l+1}}(s)\, ds$, where $V_\phi = 1$ and u_i is the i-th coordinate of $u(t)$ for $i = 1, 2, \ldots, m$ and $u_0(t) \equiv 1$.

For each formal power series c in $\eta_0, \eta_1, \ldots, \eta_m$, we define a formal operator on $\mathcal{U}_T$ in the following way:

$$F_c[u](t) = \sum \langle c, \eta_\iota \rangle V_\iota[u](t). \tag{11}$$

It is known that for any $T < (Mm+M)^{-1}$, the series (11) converges uniformly and absolutely for all $t \in [0, T]$ and all those $u \in \mathcal{U}_T$ such that $\|u\|_\infty \le 1$ (cf [6]). In fact, for any L_1 input u, there exists some $\delta > 0$ such that (11) converges uniformly and absolutely on $[0, \delta)$.

For each $T > 0$, we define $\mathcal{V}_T = \{u \in \mathcal{U}_T : \|u\|_\infty < 1\}$, and we shall say that T is *admissible for c* if $T < (M(m+1))^{-1}$ for some M such that (10) holds. Then F_c is always well defined on $\mathcal{V}_T$ if T is admissible for c. We shall call F_c an *input/output operator* defined on $\mathcal{V}_T$ if T is admissible for c. Hence, every convergent power series defines an i/o map, or more precisely, one such map on each $\mathcal{V}_T$ for which T is admissible. (We often identify any two such operators, when there is no danger of confusion, dealing in effect with "germs" of such operators.) The following two properties of the i/o operators are used in this work. We refer the readers to [19] for the proof of the lemmas:

Lemma 5.11 Assume that c is a convergent power series and T is admissible for c. Then the operator $F_c : \mathcal{V}_T \to \mathcal{C}[0, T]$ is continuous with respect to the L^1 norm in $\mathcal{V}_T$ and the $\mathcal{C}^0$ norm in $\mathcal{C}[0, T]$. □

Lemma 5.12 Suppose c is a convergent series and T is admissible to c. Then $F_c[u]$ is analytic if $u \in \mathcal{V}_T$ is analytic. ∎