SIAM J. CONTROL AND OPTIMIZATION
Vol. 33, No. 4, pp. 1102–1126, July 1995

009

ORDERS OF INPUT/OUTPUT DIFFERENTIAL EQUATIONS AND STATE-SPACE DIMENSIONS*

YUAN WANG† AND EDUARDO D. SONTAG‡

Abstract. This paper deals with the orders of input/output equations satisfied by nonlinear systems. Such equations represent differential (or difference, in the discrete-time case) relations between high-order derivatives (or shifts, respectively) of input and output signals. It is shown that, under analyticity assumptions, there cannot exist equations of order less than the minimal dimension of any observable realization; this generalizes the known situation in the classical linear case. The results depend on new facts, themselves of considerable interest in control theory, regarding universal inputs for observability in the discrete case, and observation spaces in both the discrete and continuous cases. Included in the paper is also a new and simple self-contained proof of Sussmann's universal input theorem for continuous-time analytic systems.

Key words. control systems, input/output equations, observation spaces, universal inputs, observability

AMS subject classifications. 93B15, 93A25, 93B25, 93B27, 93B29

1. Introduction.

Previous papers by the authors (see [40], [41]) studied various relationships between realizability of continuous-time systems and the existence of algebraic or analytic input/output differential equations. These are equations of the form

$$E\left(u(t), u'(t), u''(t), \ldots, u^{(r-1)}(t), y(t), y'(t), y''(t), \ldots, y^{(r)}(t)\right) = 0 \tag{1}$$

that relate inputs $u(\cdot)$ and outputs $y(\cdot)$. Such equations, and their discrete-time analogues, are of interest in identification theory and arise also naturally in the "behavioral" approach to systems (see, e.g., [43]). They provide a natural generalization of the autoregressive moving-average representations that appear in linear systems theory, where E is linear (in that case, the Laplace transform of the equation leads to the usual transfer function).

The papers [37], [40], [41] (see also [28] for analogous work in the discrete-time case) dealt with the relationships between the existence of such equations and the possibility of realizing the corresponding input/output (i/o) operator $u(\cdot) \mapsto y(\cdot)$ by a state-space system of the type

$$x'(t) = f(x(t)) + G(x(t))u(t)\,, \quad y(t) = h(x(t)) \tag{2}$$

whose state $x(t)$ evolves in an n-dimensional manifold. (Precise definitions are given later; for the rest of the introduction we give an informal discussion. The main assumption will be that all functions appearing are analytic.) While i/o equation descriptions of type (1) are well suited to identification algorithms, state-space descriptions of type (2) are often the basis of feedback design tools and are needed for

* Received by the editors April 2, 1993; accepted for publication (in revised form) January 18, 1994.

† Department of Mathematics, Florida Atlantic University, Boca Raton, Florida 33431 (ywang@polya.math.fau.edu). This research was supported in part by National Science Foundation grant DMS-9108250.

‡ Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903 (sontag@hilbert.rutgers.edu). This research was supported in part by US Air Force grant AFOSR-91-0346.

the statement and solution of control problems. Thus, it is of great interest to study the possible relationships between the two kinds of descriptions.

A question that has not been sufficiently studied and was not addressed in [40], [41] is that of comparing the *order* r of an i/o equation (1) to the minimal possible dimension n of a realization (2). In discrete time, for the analogous equations

$$E\left(u(t-1),\, u(t-2),\, \ldots,\, u(t-r),\, y(t),\, y(t-1),\, \ldots,\, y(t-r)\right) = 0 \tag{3}$$

and systems

$$\Sigma: \quad x(t+1) = f(x(t), u(t)), \quad y(t) = h(x(t)), \quad t = 0,\, 1,\, 2, \ldots, \tag{4}$$

it was known for a long time (see [28]) that one may have $r < n$, even if the system in (4) is minimal. It turns out, perhaps surprisingly, that this cannot happen in the continuous-time case: we prove here that if there is a minimal realization of dimension n, then no i/o equation can have order less than n. Moreover, we show that the result holds true also for discrete-time systems that are *reversible*, that is, those for which the controls induce one-to-one maps on the state space (the examples in [28] were not reversible).

The results in [40], [41] depend on an important equality among observation spaces. The latter are sets of functions on the state space that are obtained by performing different kinds of "experiments" with the system and extracting infinitesimal information from the observed data. The basic fact needed was established in [39], and it related the space obtained by using piecewise constant controls (and derivatives of the output function with respect to switching times between constant pieces) to the space obtained when using differentiable inputs instead (and the corresponding jet of derivatives of the output at time zero). The new results given in this paper depend on new facts, themselves of considerable interest in control theory, regarding subspaces obtained by the application of "generic" smooth inputs.

The results in this paper were announced and their proofs sketched in the conference paper [32] (and for discrete time in [42]). To be more precise, in [32] we derived our results from an equality between observation spaces that is somewhat weaker than the corresponding one proved here; namely, instead of the current Lemma 2.1, we only had that $\mathrm{d}\mathcal{F}(x) = \mathrm{d}\mathcal{F}_\mu(x)$ for generic jets μ and for *generic* states x. This is all that is needed in order to establish the desired results on orders of i/o equations. However, while this journal version was being written, Coron [5] showed that the equality can be strengthened so that it holds for *all* (not merely generic) states (but still generic μ). Since it turns out that the stronger equality can in fact be obtained with essentially the same proof as in [32], we now present the result directly in that form. (Since we are only interested in analytic systems, we can use elementary facts from analytic geometry to present a simpler approach to the problem than in [5]; in that reference the techniques of proof are very different, as the focus is on applications to feedback control problems for smooth systems. See also [31] for remarks on applications of results of the type proved here to path planning and feedback.)

In the development of the new observation space results, we needed to extend to discrete time the well-known and fundamental theorem by Sussmann on universal inputs for distinguishability of continuous-time analytic systems. It turned out that our proof also applies in continuous time. The theorem is obtained in a fairly direct way from a stronger result, Lemma 2.1 in this work. The proof of Lemma 2.1 is very elementary and intuitive, as it does not use anything more complicated than the fact that every descending chain of sets defined by analytic equations stabilizes relative to any fixed compact. (The original proof of Sussmann's theorem relies heavily on the stratification theory of subanalytic sets, a considerably deeper set of tools. Thus one contribution of this paper is to provide an alternative and simpler proof of that important result.) In addition to its role in helping to derive the universal input theorem and our main results, Lemma 2.1 also has its own independent interest, as it provides relationships between observation spaces defined in different ways and, thus, provides connections between several different notions of observability. We also note the very recent work [33], where further results on universal inputs are presented; these results show in particular the existence of inputs that are universal uniformly over the class of all analytic systems.

Another set of results that arose naturally while studying the problems in this paper, and which are included here, deals with the relationships among various alternative notions of observability, especially those proposed in the context of the differential-algebraic approach to control theory. We are able to characterize, for instance, the notion of observability proposed in [10], [9] in terms of more standard local observability concepts.

1.1. Other related work. In addition to the references already mentioned, work by many authors is related to the topic of i/o equations and realizability; see for instance [6], [13], [37]. In particular, [7], [8] showed that one must add inequality constraints to (1) in order to obtain a precise characterization of the behavior of a state-space system, unless stronger algebraic conditions hold. In [26], [38], [4], local i/o equations were derived under nondegeneracy rank conditions, for smooth systems, under observability assumptions. The notions of observation space and algebra that we employ were introduced for discrete-time systems in [28], and their analogous continuous-time versions were given in [2], [3].

1.2. Outline of paper. In §2 we introduce continuous-time systems and a technical result on observation spaces for generic jets. Certain special cases for which stronger conclusions can be given, namely bilinear and rational systems, are also studied there. In §3, we define universal inputs and relate their properties to the results on equality of observation spaces and to the orders of i/o equations. The following section has a proof of the main technical results stated in §§2 and 3. After this, §5 provides the discrete-time results. There are also two appendices with some technical lemmas that are required by the proofs.

2. Observation spaces for continuous-time systems.

In this section we first discuss several natural ways of defining observation spaces for continuous-time systems and then explore the relationships between the different definitions.

2.1. Observation spaces. Consider an *analytic system*

$$\Sigma: \begin{cases} x'(t) &= g_0(x(t)) + \displaystyle\sum_{i=1}^{m} g_i(x(t))u_i(t)\,, \\ y(t) &= h(x(t))\,, \end{cases} \tag{5}$$

where for each t, $x(t) \in \mathcal{M}$, which is an analytic (second countable) manifold of dimension n, $h : \mathcal{M} \longrightarrow \mathbb{R}$ is an analytic function and $g_0, g_1, \ldots, g_m$ are analytic vector fields defined on $\mathcal{M}$. Controls are measurable essentially bounded maps $u : [0, T] \longrightarrow \mathbb{R}^m$ defined on suitable intervals. In general, $\varphi(t, x, u)$ denotes the state trajectory of (5) corresponding to a control u and initial state x, defined at least for small t.

In the special case in which $\mathcal{M} = \mathbb{R}^n$ and the entries of the vector fields g_i's (on the natural global coordinates) and of the function h are rational (with no real poles), we call (5) a *rational system*. If, in addition, the entries of the g_i's and h are polynomials, we call (5) a *polynomial system*.

For a given continuous-time system, let $\mathcal{F}$ be the subspace of functions $\mathcal{M} \longrightarrow \mathbb{R}$ spanned by the Lie derivatives of h in the directions of $g_0, g_1, \ldots, g_m$, i.e.,

$$\mathcal{F} := \operatorname{span}_{\mathbb{R}} \left\{ L_{g_{i_1}} L_{g_{i_2}} \ldots L_{g_{i_r}} h : \ r \geq 0, \ 0 \leq i_j \leq m \right\} . \tag{6}$$

This is the *observation space* associated with (5) (see, e.g., [30, Rem. 5.4.2]) For each $x \in \mathcal{M}$, let $\mathcal{F}(x)$ denote the space obtained by evaluating the elements of $\mathcal{F}$ at x.

For each $\alpha \in \mathcal{F}$, we may consider its differential $\mathrm{d}\alpha$, seen as a 1-form. For each $x \in \mathcal{M}$, we let $\mathrm{d}\mathcal{F}(x)$ be the space of covectors defined by

$$\mathrm{d}\mathcal{F}(x) = \{\mathrm{d}\alpha(x) : \alpha \in \mathcal{F}\} .$$

We also let $\mathrm{d}\mathcal{F}$ be $\{d\alpha : \ \alpha \in \mathcal{F}\}$ as a space of 1-forms.

A related construction is as follows. First, let $\mathbb{R}^{m,\infty} = \prod_{i=1}^{\infty} \mathbb{R}^m$ endowed with the box topology, for which a base of open sets consists of all sets of the form $\prod_{i=1}^{\infty} \mathcal{U}_i$, where each $\mathcal{U}_i$ is an open set of $\mathbb{R}^m$. A *generic* subset $\mathcal{W}$ of $\mathbb{R}^{m,\infty}$ is one that contains a countable intersection of open dense sets. It can easily be shown that with the above topology, $\mathbb{R}^{m,\infty}$ is a Baire space; thus, a generic subset is always dense.

Now for any $\mu = (\mu_0, \mu_1, \ldots)$ in $\mathbb{R}^{m,\infty}$, we define

$$\psi_i(x,\mu) = \left.\frac{d^i}{dt^i}\right|_{t=0} h(\varphi(t,x,u)) \tag{7}$$

for $i \geq 0$, where u is any C^∞ control with initial values $u^{(j)}(0) = \mu_j$. The functions $\psi_i(x, \mu)$ can be expressed—applying repeatedly the chain rule—as polynomials in the $\mu_j = (\mu_{1j}, \ldots, \mu_{mj})$ whose coefficients are analytic functions (rational functions if the system is rational) of x. Take the single-input case

$$\dot{x} = f(x) + g(x), \quad y = h(x)$$

(for simplicity of notation) as an example. The functions are

$$\begin{aligned}
&\psi_0(x, \mu) = h(x),\\
&\psi_1(x, \mu) = L_f h(x) + \mu_0 L_g h(x),\\
&\psi_2(x, \mu) = L_f^2 h(x) + \mu_0 (L_g L_f h(x) + L_f L_g h(x)) + \mu_0^2 L_g^2 h(x) + \mu_1 L_g h(x),
\end{aligned}$$

and so forth. For instance, for single-input single-output linear systems

$$x' = Ax + bu, \ \ y = cx,$$

we have,

$$\psi_l(x, \mu) = cA^l x + \sum_{i=1}^{l} \mu_{i-1} cA^{i-1} b, \quad l = 0, 1, \ldots .$$

For each fixed $\mu \in \mathbb{R}^{m,\infty}$, let $\mathcal{F}_\mu$ be the subspace of functions from $\mathcal{M}$ to $\mathbb{R}$ defined by

$$\mathcal{F}_\mu = \operatorname{span}_{\mathbb{R}} \{\psi_0(\cdot,\mu), \psi_1(\cdot,\mu), \psi_2(\cdot,\mu), \ldots\} \tag{8}$$

and let $\mathcal{F}_\mu(x)$ be the space obtained by evaluating the elements of $\mathcal{F}_\mu$ at x for each $x \in \mathcal{M}$. Let $\mathrm{d}\mathcal{F}_\mu(x)$ be the space of covectors given by

$$\mathrm{d}\mathcal{F}_\mu(x) = \{\mathrm{d}\psi(x,\mu) : \ \psi \in \mathcal{F}_\mu\}$$

for each $x \in \mathcal{M}$. For instance, for linear systems, $\mathrm{d}\psi_l(x,\mu) = cA^l$ and

$$\mathrm{d}\mathcal{F}_\mu(x) = \operatorname{span}\{c,\ cA,\ cA^2,\ \ldots\},$$

which is independent of μ (and x). We also let $\mathrm{d}\mathcal{F}_\mu$ be $\{d\psi(\cdot,\mu) : \ \psi \in \mathcal{F}_\mu\}$, seen as a space of covector fields.

Clearly, for each μ, $\mathcal{F}_\mu$ is a subspace of $\mathcal{F}$, and therefore, for each x also $\mathrm{d}\mathcal{F}_\mu(x)$ is a subspace of $\mathrm{d}\mathcal{F}(x)$. The main result in [39] says that

$$\boxed{\mathcal{F} \ = \ \sum_\mu \mathcal{F}_\mu\,.} \tag{9}$$

This equality is fundamental in establishing results linking realizability to the existence of i/o equations, in [40] and [41]. In intuitive but less rigorous terms, the equality in (9) can be interpreted as follows. We consider the successive derivatives $y(0)$, $y'(0)$, $y''(0)$, ... expressed as functions of $x(0)$ and $u(0)$, $u'(0)$, $u''(0)$, ... For particular controls $u(t)$, the $y(0)$, $y'(0)$, $y''(0)$, ... are just functions of x; taking the span of all such functions, over all possible *smooth* controls, one obtains the right-hand side of (9). On the other hand, taking all possible *piecewise-constant* instead of smooth controls and taking derivatives with respect to the times at which the controls switch values, one obtains the space $\mathcal{F}$ in the left-hand side of (9).

The following is a technical result for continuous-time systems, which will help in deriving the desired facts about i/o equations.

LEMMA 2.1. *Assume that* (5) *is an analytic system. Then there exists a generic subset* $\mathcal{W}$ *of* $\mathbb{R}^{m,\infty}$ *such that*

$$\mathcal{F}(x) = \mathcal{F}_\mu(x) \tag{10}$$

and

$$\mathrm{d}\mathcal{F}(x) = \mathrm{d}\mathcal{F}_\mu(x) \tag{11}$$

for every $x \in \mathcal{M}$ *and all* $\mu \in \mathcal{W}$.

Remark 2.2. The above conclusions are also true if instead of the box topology one uses the weak topology on $\mathbb{R}^{m,\infty}$. This is the topology for which a basis of open sets consists of all sets of the form $\prod_{i=1}^\infty \mathcal{U}_i$, where each $\mathcal{U}_i$ is an open subset of $\mathbb{R}^m$ and only finitely many of them are proper subsets of $\mathbb{R}^m$. Clearly, the weak topology is coarser than the topology used before. With this topology, $\mathbb{R}^{m,\infty}$ is again a Baire space. We will remark at the end of the proof of Lemma 2.1 in §4 that the conclusions of Lemma 2.1 also hold for the weak topology. Moreover, these conclusions can be established as consequences of a more general result about convergent generating series, that ensures there exists a generic subset $\mathcal{W}$ of $\mathbb{R}^{m,\infty}$ with the property that these jets suffice for distinguishing *all possible* convergent generating series; more details are given in [33].

Remark 2.3. The conclusions in Lemma 2.1 do not always hold for every $\mu \in \mathbb{R}^{m,\infty}$. Consider as an illustration the following bilinear system:

$$x_1' = x_2, \quad x_2' = x_2 + x_1 u, \quad y = x_2\,. \tag{12}$$

For this system, $\mathcal{F} = \operatorname{span}\{x_1, x_2\}$, thus, $\mathcal{F}(x) \neq 0$ for all $x \neq 0$. But on the other hand, we have

$$\psi_0(x,\mu) = x_2, \, . \, \psi_1(x,\mu) = x_2 + x_1\mu_0 \,,$$

and in general, $\psi_k(x,\mu) = P_k(x, \mu_0, \mu_1, \ldots, \mu_{k-2}) + x_1\mu_{k-1}$, where P_k is some polynomial. Clearly, for every $x = (x_1, x_2)$ for which $x_1 \neq 0$, one can find a solution μ recursively for the equations $\psi_i(x,\mu) = 0$ for $i > 0$. Hence, as long as $x_1 \neq 0$ and $x_2 = 0$, there exists some jet μ such that $\mathcal{F}_\mu(x) = 0$, which is therefore different from $\mathcal{F}(x)$ when $x_1 \neq 0$ and $x_2 = 0$.

2.2. Algebraic formulation. In this section, we assume for simplicity that $\mathcal{M} = \mathbb{R}^n$; we could work with more general manifolds but this would complicate notation, and in any case we will only need to apply the results given here locally. We say a function β is a *meromorphic function* if $\beta = \frac{p}{q}$, where p and q are analytic functions defined on $\mathcal{M}$, and $q \not\equiv 0$. (Note that this global definition is different from the local definition usually given; see, e.g., [17]. It will be enough for our purposes.) For each function $\alpha \in \mathcal{F}$, $\mathrm{d}\alpha$ is a covector field defined on $\mathcal{M}$. If β is a meromorphic function defined on $\mathcal{M}$, then $\beta\mathrm{d}\alpha$ is a well-defined 1-form on some open dense subset of $\mathcal{M}$ and any finite sum of such partially defined covector fields is defined on a common open dense set. Thus, we may introduce the subspace $\widehat{\mathrm{d}}\mathcal{F}$ of the cotangent space defined by

$$\widehat{\mathrm{d}}\mathcal{F} := \operatorname{span}_{\mathbb{R}_x}\{\mathrm{d}\alpha : \; \alpha \in \mathcal{F}\} \,,$$

where $\mathbb{R}_x$ is the field of meromorphic functions defined on $\mathcal{M}$. Similarly, one can define, for each $\mu \in \mathbb{R}^{m,\infty}$, the space $\widehat{\mathrm{d}}\mathcal{F}_\mu$ by

$$\widehat{\mathrm{d}}\mathcal{F}_\mu := \operatorname{span}_{\mathbb{R}_x}\{\mathrm{d}\alpha : \; \alpha \in \mathcal{F}_\mu\} \,.$$

Note that there are natural identifications $\widehat{\mathrm{d}}\mathcal{F} \simeq \mathrm{d}\mathcal{F} \otimes \mathbb{R}_x$ and $\widehat{\mathrm{d}}\mathcal{F}_\mu \simeq \mathrm{d}\mathcal{F}_\mu \otimes \mathbb{R}_x$.

Since $\mathcal{M} = \mathbb{R}^n$, we can identify elements of $\widehat{\mathrm{d}}\mathcal{F}$ with vectors

$$(\alpha_1(x), \alpha_2(x), \ldots, \alpha_n(x))$$

of meromorphic functions defined on $\mathcal{M}$. The dimension of $\widehat{\mathrm{d}}\mathcal{F}$ over $\mathbb{R}_x$ is the size of the largest matrix that can be formed out of such vectors and has full rank, i.e., has a minor that is not zero as a function. That is, $\dim_{\mathbb{R}_x} \widehat{\mathrm{d}}\mathcal{F}$ is the same as $\max_{x\in\mathcal{M}} \dim \mathrm{d}\mathcal{F}(x)$. A similar argument can be made for each $\mathrm{d}\mathcal{F}_\mu(x)$; together with Lemma 2.1, we can then conclude the following corollary.

COROLLARY 2.4. *For any analytic system, $\widehat{\mathrm{d}}\mathcal{F}_\mu = \widehat{\mathrm{d}}\mathcal{F}$ for all μ in a generic set of $\mathbb{R}^{m,\infty}$.*

Yet another object is obtained if one instead views the elements

$$\psi_i(x, U) \tag{13}$$

as rational functions (in particular polynomials), on the formal variables $U = \{U_{ij}\}$, whose coefficients are functions of x, as opposed to seeing them as functions of x for each numerical choice $U_{ij} = \mu_{ij}$. We proceed as follows. Let

$$K = \mathbb{R}\Big(\{U_{ij} : \; 1 \leq i \leq m, \, j \geq 0\}\Big)$$

be the field obtained by adjoining the indeterminates U_{ij} to $\mathbb{R}$, and let

$$K_x = \mathbb{R}_x\Big(\{U_{ij}:\ 1 \le i \le m,\ j \ge 0\}\Big)$$

be the field obtained by adjoining the indeterminates U_{ij} to $\mathbb{R}_x$. We then let $\mathfrak{F}$ be defined as the subspace of K_x spanned by the functions ψ_i over the field K, i.e.,

$$\mathfrak{F} := \operatorname{span}_K \{\psi_i :\ i \ge 0\}\ .$$

Thus, $\mathfrak{F}$ consists of finite linear combinations $\sum q_i(U)\psi_i(x, U)$, where the $q_i(\cdot)$ are rational functions on the variables $\{U_{ij}\}$. Such a linear combination can be seen as a rational function on the $\{U_{ij}\}$ whose coefficients are meromorphic functions of x (and hence also meromorphic functions) and, thus, elements of K_x. The differentials (with respect to x) of elements of K_x are viewed as rational functions in $\{U_{ij}\}$, whose coefficients are (in general, partially defined) covector fields. Finally we define

$$\widehat{\mathrm{d}\mathfrak{F}} := \operatorname{span}_{K_x}\big\{\mathrm{d}\psi :\ \psi \in \mathfrak{F}\big\}.$$

Then Lemma 2.1 implies the following corollary.

COROLLARY 2.5. *For any analytic system,* $\dim_{\mathbb{R}_x} \widehat{\mathrm{d}}\mathcal{F} = \dim_{K_x} \widehat{\mathrm{d}\mathfrak{F}}$.

Proof. Clearly $\dim_{K_x} \widehat{\mathrm{d}\mathfrak{F}} \le \dim_{\mathbb{R}_x} \widehat{\mathrm{d}}\mathcal{F}$. Conversely, $\dim_{K_x} \widehat{\mathrm{d}\mathfrak{F}} = \max_\mu \dim_{R_x} \widehat{\mathrm{d}}\mathcal{F}_\mu$. The desired conclusion then follows from Corollary 2.4. □

2.3. Bilinear and rational systems. Now consider the bilinear system

$$x' = A_0 x + \sum_{i=1}^{m} u_i A_i x,$$
$$y = cx,$$

where $A_0, A_1, \ldots, A_m$ are $n\times n$ matrices and c is an $1\times n$ matrix. For each multi-index $i_1 i_2 \ldots i_r$, where $0 \le i_j \le m$ for each $j \ge 0$,

$$L_{g_{i_1}} L_{g_{i_2}} \ldots L_{g_{i_r}} h(x) = cA_{i_r} A_{i_{r-1}} \ldots A_{i_1} x.$$

Note that ψ_i (as defined in (7)) is also linear in x for each i; for instance, in the single-input case (for simplicity of notation),

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

Thus, for the bilinear case, we have the following corollary.

COROLLARY 2.6. *For a bilinear system,*

$$\mathcal{F} = \mathcal{F}_\mu \quad \text{and} \quad \mathrm{d}\mathcal{F} = \mathrm{d}\mathcal{F}_\mu \tag{14}$$

for every μ in a generic subset of $\mathbb{R}^{m,\infty}$.

Remark 2.7. We would like to point out that this corollary does *not* hold in general. The following simple example shows that for a general nonlinear system, $\mathcal{F}$ and $\mathcal{F}_\mu$ (respectively, $\mathrm{d}\mathcal{F}$ and $\mathrm{d}\mathcal{F}_\mu$) may not be the same for *any* μ, even though the two spaces $\widehat{\mathrm{d}}\mathcal{F}$ and $\widehat{\mathrm{d}}\mathcal{F}_\mu$ are the same.

Example 2.8. Consider the system

$$x' = x^3 + x^2 u, \quad y = x.$$

It is easy to see that

$$y = x, \quad y' = x^3 + x^2 u,$$
$$y'' = 3x^5 + 5x^4 u + 2x^3 u^2 + x^2 u',$$

and in general, $y^{(k)} = (2k-1)!!x^{2k+1} + p_k(u, u', \ldots, u^{(k-2)}, x) + x^2 u^{(k-1)}$, where p_k is a polynomial in x of degree less than or equal to $2k$. It can be seen that

$$\mathcal{F} = \operatorname{span}_{\mathbb{R}} \{x,\ x^2,\ x^3, \ldots\} .$$

However, $x^2 \notin \mathcal{F}_\mu$ for any μ for the following reason. Assume that

$$x^2 = \sum_{i=0}^{k} a_i \psi_i(x, \mu)$$

for some k and some $a_0, a_1, \ldots, a_k \in \mathbb{R}$. Then $a_i = 0$ for $i \geq 2$, otherwise the degree of x in the left-hand side would be higher than 3. Thus the above equation becomes

$$x^2 = a_0 x + a_1(x^3 + x^2 \mu_0),$$

which is impossible. This shows that $\mathcal{F}_\mu \neq \mathcal{F}$ for any μ even though, in this case, $\widehat{\mathrm{d}}\mathcal{F} = \widehat{\mathrm{d}}\mathcal{F}_\mu = \operatorname{span}_{\mathbb{R}_x}\{dx\}$ for all μ.

In this example, it is also true that $\mathrm{d}\mathcal{F} \neq \mathrm{d}\mathcal{F}_\mu$ for any μ. This can be shown as follows. If $\mathrm{d}\mathcal{F} = \mathrm{d}\mathcal{F}_\mu$, then $\mathrm{d}x^2 = 2x\mathrm{d}x \in \mathrm{d}\mathcal{F}_u$. From here it would follow that $x^2 = \alpha_1(x,\mu) + \alpha_2(x,\mu) + \cdots + \alpha_l(x,\mu) + c$ for some elements $\alpha_i \in \mathcal{F}_\mu$ and some constant $c \in \mathbb{R}$. But it can be seen from the above argument that this is impossible.

Assume now that (5) is a rational system. Define $\mathcal{A}$ ($\mathcal{A}_\mu$, respectively) as the $\mathbb{R}$-algebra generated by the elements of $\mathcal{F}$ ($\mathcal{F}_\mu$, repectively). Then we define the observation field $\mathcal{Q}$ ($\mathcal{Q}_\mu$, respectively) as the quotient field of $\mathcal{A}$ ($\mathcal{A}_\mu$, respectively). For a field extension Q of $\mathbb{R}$, we use $\mathrm{trdeg}_{\mathbb{R}} Q$ to denote the transcendence degree of Q over $\mathbb{R}$. Then we have the following conclusion for rational systems, in analogy to the above conclusion about bilinear systems.

COROLLARY 2.9. *For a rational system,*

$$\mathrm{trdeg}_{\mathbb{R}}\, \mathcal{Q} = \mathrm{trdeg}_{\mathbb{R}}\, \mathcal{Q}_\mu ,$$

for each μ in a generic subset of $\mathbb{R}^{m,\infty}$.

3. Observability and universal inputs in continuous time.

Consider an analytic system (5). 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 $h(\varphi(\cdot, p, u)) \neq h(\varphi(\cdot, q, u))$ (considered as functions defined on the common domain of $\varphi(\cdot, p, u)$ and $\varphi(\cdot, q, u)$); 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 say that system (5) is *observable*. (See [30, Chap. 5].)

An input u is called a *universal* (*distinguishing*) input for system (5) if every distinguishable pair can be distinguished by u. The existence of universal inputs was first studied in [15] for bilinear systems, in [27] for analytic systems with compact state spaces, and for arbitrary analytic systems in [35] for the continuous case. In this work, we will provide a different and simpler proof of the general result in [35]. (Also, we later give a discrete-time version.) We now state the result to be proved.

For each $T > 0$, we consider $\mathcal{C}^\infty[0, T]$ endowed with the Whitney topology, that is, the topology for which a neighborhood base for each function $u(\cdot) \in \mathcal{C}^\infty[0, T]$ consists of the sets of the following form:

$$\mathcal{U}_{u,k,\delta} = \left\{ v \in \mathcal{C}^\infty[0, T] : \max_{0 \le i \le k,\ t \in [0, T]} |v^{(i)}(t) - u^{(i)}(t)| \le \delta \right\}$$

for some $k \ge 0$ and some $\delta > 0$. This is well known to be a Baire space (see [14]). By a generic subset of $\mathcal{C}^\infty[0, T]$ we mean a subset of $\mathcal{C}^\infty[0, T]$ containing a countable intersection of open dense sets.

THEOREM 3.1 (Sussmann's universal input theorem). *For any analytic system* (5), *and any fixed $T > 0$, the set of universal inputs is a generic subset of $\mathcal{C}^\infty[0, T]$.*

PROPOSITION 3.2. *There is always an analytic universal input for any analytic system.*

We will provide proofs of Theorem 3.1 and Proposition 3.2 in §4.1.

Consider the following more general class of systems:

$$x'(t) = f(x(t), u(t)), \quad y(t) = h(x(t)), \tag{15}$$

where for each t, $x(t) \in \mathcal{M}$, which is an analytic manifold of dimension n, $h : \mathcal{M} \to \mathbb{R}$ is an analytic function and $f : \mathcal{M} \times \mathbb{R}^m \to T\mathcal{M}$ is analytic and $f(x, u) \in T_x\mathcal{M}$ for each (x, u), so in particular, $f(\cdot, u)$ is an analytic vector field for each $u \in \mathbb{R}^m$. Controls are measurable essentially bounded maps: $u : [0, T] \longrightarrow \mathbb{R}^m$, for some $T = T_u > 0$. We apply the same definitions of distinguishability, observability, and universal inputs as for system (5) to system (15). One can then generalize the conclusion of Theorem 3.1 to systems of type (15) by means of the following argument. We consider the following system:

$$\begin{aligned} x'(t) &= f(x(t), z(t)), \quad z'(t) = v(t), \\ y(t) &= h(x(t)), \end{aligned} \tag{16}$$

where v is now a new control. By Proposition 5.1.11 in [30], one knows that if (x_1, x_2) is a distinguishable pair for (15), then x_1, x_2 can be distinguished by a differentiable (in fact, an analytic) control u. It then follows that for (16), the pair (ξ, ζ), where $\xi = (x_1, u(0))$ and $\zeta = (x_2, u(0))$, is distinguished by $v(t) = u'(t)$. On the other hand, if for (16) the pair $((x_1, z), (x_2, z))$ is distinguished by v, then for (15) (x_1, x_2) is distinguished by the control

$$z + \int_0^t v(s)\, ds.$$

Therefore, (x_1, x_2) is a distinguishable pair of (15) if and only if there exists some $z \in \mathbb{R}^m$ such that $((x_1, z), (x_2, z))$ is a distinguishable pair for (16) for some z. Applying Theorem 3.1 to system (16), we proved the following conclusion.

COROLLARY 3.3. *The universal inputs for system* (15) *form a generic subset of $\mathcal{C}^\infty[0, T]$, for any $T > 0$.*

3.1. Other notions of observability. In what follows, we study relationships among several alternative notions of "observability" that have been proposed by various authors.

Take an open subset $\mathcal{U}$ of $\mathcal{M}$ and any two points $p, q \in \mathcal{U}$. If for every input u, $h(\varphi(t, p, u)) = h(\varphi(t, q, u))$ for each t for which $\varphi(T, p, u)$ and $\varphi(T, q, u)$ are both defined and in $\mathcal{U}$ for all $0 \leq t \leq T$, then we say that p and q are *$\mathcal{U}$-indistinguishable* (see, e.g., [29]).

Fix a point $p \in \mathcal{M}$. If for every neighborhood $\mathcal{U}_p$ there is a neighborhood $V_p \subset \mathcal{U}_p$ such that for any $q \in V_p$ the condition that q and p are $\mathcal{U}_p$-indistinguishable implies $p = q$, then we say the system (5) is *locally observable at* p. If (5) is locally observable at every point p, then we say (5) is *locally observable*. If there is an open dense set $\mathcal{U} \subset \mathcal{M}$ such that (5) is locally observable at every point p of $\mathcal{U}$, then we say (5) is *generically locally observable*. See [29] for details on local observability and related concepts such as the slightly different definition in [26]. The following fact is an immediate consequence of Lemma 2.10 and facts (2.4) and (2.8) in [29].

PROPOSITION 3.4. *An analytic system* (5) *is generically locally observable if and only if* $\max_x \dim \mathrm{d}\mathcal{F}(x) = n$.

PROPOSITION 3.5. *Let* $\mathcal{M} = \mathbb{R}^n$ *and let* (5) *be an analytic system. Then the following are equivalent*:

(1) *The system is generically locally observable.*

(2) $\dim_{K_x} \widehat{\mathrm{d}\mathfrak{F}} = n$.

(3) $\dim_{\mathbb{R}_x} \widehat{\mathrm{d}}\mathcal{F} = n$.

Proof. The maximum dimension of $\mathrm{d}\mathcal{F}(x)$ is the same as the $\dim_{\mathbb{R}_x} \widehat{\mathrm{d}}\mathcal{F}$. This shows that (1) and (3) are equivalent; (2) is equivalent to (3) by Corollary 2.5. □

For a polynomial system, the $\psi_i(x, U)$'s (as defined in (13)) are polynomial functions of both x and U. We say that a polynomial system is *weakly algebraically observable* if each coordinate x_i is algebraically over the field $K(\{\psi_i : i \geq 0\})$ $(= \mathbb{R}(\{U_{ij}, \psi_k, i = 1, \ldots, m;\ j, k \geq 0\}))$. It follows that Σ is weakly algebraically observable if and only if $\dim_{K(x)} \widehat{\mathrm{d}\mathfrak{F}} = n$, where $K(x)$ is the field of rational functions over K. (This is proved as follows: The dimension condition is equivalent, by [18, Thm. III of III.7], to the property that the transcendence degree of $K_0 = K(\{\psi_i : i \geq 0\})$ over K should be equal to n. On the other hand, we have the inclusions $K \subseteq K_0 \subseteq K(x)$, so $\mathrm{trdeg}_K K_0 + \mathrm{trdeg}_{K_0} K(x) = n$. Thus the dimension is n if and only if $\mathrm{trdeg}_{K_0} K(x) = 0$, i.e., if and only if $K(x)$ is algebraic over K_0.) By Proposition 3.5, we have the following corollary.

COROLLARY 3.6. *A polynomial system is weakly algebraically observable if and only if the system is generically locally observable.*

The notion of weakly algebraic observability used here was called "weak observability" in [28]. The same notion was used in [10] and extended to cover implicit systems as well.

3.2. Orders of i/o equations in continuous-time case.

3.2.1. State-space systems. We say that a state-space system Σ *admits an i/o equation* such as

$$A(u(t),\, u'(t),\, \ldots,\, u^{(k-1)}(t),\, y(t),\, y'(t),\, \ldots,\, y^{(k)}(t)) = 0\,, \tag{17}$$

where A is a nonzero analytic function from $\mathbb{R}^{mk} \times \mathbb{R}^{k+1}$ to $\mathbb{R}$, if (17) holds for every initial state x, every $\mathcal{C}^k$ i/o pair (u, y) of (5), and all t such that $y(t)$ is defined. The *order* of an equation (17) is defined to be the highest $r \leq k$ such that

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

is not a zero function.

For a given system Σ, we define $\delta(\Sigma)$ to be the lowest possible order of an i/o equation that Σ admits. In the case that there is no such i/o equation, $\delta(\Sigma)$ is defined to be $+\infty$.

THEOREM 3.7. *Assume Σ is an n-dimensional analytic system defined by* (5). *If Σ is generically locally observable, then $\delta(\Sigma) \geq n$. If, in addition, Σ is a rational system, then $\delta(\Sigma) = n$.*

Proof. Let $\mathcal{U} \subseteq \mathcal{M}$ be an open subset diffeomorphic to $\mathbb{R}^n$. We consider the restriction of Σ to $\mathcal{U}$. This system is still generically locally observable, and an equation for Σ is also an equation for the restriction. So without loss of generality, we assume from now on that $\mathcal{M} = \mathbb{R}^n$.

Assume that $\delta(\Sigma) = k < \infty$ and Σ admits i/o equation (17) of order k. For each integer $i \geq 0$, let

$$A_i = \frac{\partial^i}{\partial \nu_k^i} A(\mu_0, \ldots, \mu_{k-1}, \nu_0, \nu_1, \ldots, \nu_k).$$

Claim. There exists an i, such that A_i is not an i/o equation of Σ.

We prove the claim as follows. Assume that A_i is an i/o equation of Σ for every i. Then for any fixed i/o pair (u, y) and any fixed t, it holds that

$$A_i(u(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k)}(t)) = 0$$

for all i. Thus, as a function of ν_k for these fixed values $u(t), \ldots, y^{(k-1)}(t)$, all derivatives of

$$A(u(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k-1)}(t), \nu_k) \tag{18}$$

evaluated at $\nu_k = y^{(k)}(t)$ vanish. It then follows from the analyticity of A that (18) vanishes for all values of ν_k. Let $\bar{\nu}_k$ be such that the function

$$\tilde{A}(\mu_0, \ldots, \mu_{k-1}, \nu_0, \ldots, \nu_{k-1}) := A(\mu_0, \ldots, \mu_{k-1}, \nu_0, \ldots, \nu_{k-1}, \bar{\nu}_k)$$

is not a zero function. Clearly it holds that

$$\tilde{A}(u(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k-1)}(t)) = 0 \tag{19}$$

for all i/o pairs of Σ. If one can show that $\tilde{A}$ does not depend on μ_{k-1}, then one concludes that $\tilde{A} = 0$ is an i/o equation for Σ. For this, we proceed as follows. First of all, (19) holds for all i/o pairs of (u, y) if and only if the following holds:

$$\tilde{A}(\mu_0, \ldots, \mu_{k-1}, \psi_0(x, \mu), \ldots, \psi_{k-1}(x, \mu)) = 0$$

for all $x \in \mathcal{M}$ and all μ. Note here that ψ_i defined by (7) does not depend on μ_j for $j \geq i$. It follows that for any $\bar{\mu}_{k-1}$,

$$\tilde{A}(\mu_0, \ldots, \mu_{k-2}, \bar{\mu}_{k-1}, \psi_0(x, \mu), \ldots, \psi_{k-1}(x, \mu)) = 0$$

for all x and all μ. Finally, pick $\bar{\mu}_{k-1}$ such that

$$\bar{A}(\mu_0, \ldots, \mu_{k-2}, \nu_0, \ldots, \nu_{k-1}) := \tilde{A}(\mu_0, \ldots, \mu_{k-2}, \bar{\mu}_{k-1}, \nu_0, \ldots, \nu_{k-1}, \bar{\nu}_k)$$

is not a zero function. Then $\bar{A} = 0$ is an i/o equation of order $k-1$ for Σ. This contradicts the assumption that $\delta(\Sigma) = k$. The claim is thus proved.

Now let $r \geq 1$ be the smallest number for which $A_r = 0$ is not an i/o equation for Σ. Replace A in (17) by A_{r-1}. Evaluating (17) at $t = 0$, the equation implies the identity

$$A(\mu_0, \ldots, \mu_{k-1}, \psi_0(x, \mu), \ldots, \psi_k(x, \mu)) = 0 . \tag{20}$$

Since $A_1 = 0$ is not an i/o equation of Σ, it follows that there exists some $\mu \in \mathbb{R}^{mk}$ such that

$$A_1(\mu_0, \ldots, \mu_{k-1}, \psi_0(x, \mu), \ldots, \psi_k(x, \mu)) \neq 0 , \tag{21}$$

as a function of x, and hence, by analyticity, the complement of

$$B = \{\mu \in \mathbb{R}^{mk} : \ A_1(\mu_0, \ldots, \mu_{k-1}, \psi_0(x, \mu), \ldots, \psi_k(x, \mu)) = 0 , \ \forall x\}$$

is an open dense subset of $\mathbb{R}^{mk}$.

Combining (20) and (21), one sees, for each $\mu \notin B$, that $d\psi_k(\cdot, \mu)$ is a linear combination of $d\psi_0(\cdot, \mu), \ldots, d\psi_{k-1}(\cdot, \mu)$ over $\mathbb{R}_x$. Thus $\widehat{d\mathcal{F}}_\mu^k = \widehat{d\mathcal{F}}_\mu^{k-1}$, where, for each i, $\widehat{d\mathcal{F}}_\mu^i$ is the subspace of $\widehat{d\mathcal{F}}_\mu$ spanned by $d\psi_0(\cdot, \mu), d\psi_1(\cdot, \mu), \ldots, d\psi_i(\cdot, \mu)$. Differentiating (17) with respect to time, one sees that for any $i > 1$ it holds that

$$\begin{aligned} &A_1(u(t), \ldots, u^{(k-1)}(t), y(t), \ldots, y^{(k)}(t))y^{(k+i)}(t) \\ &\quad = A_i(u(t), \ldots, u^{(k+i-1)}(t), \ldots, y(t), \ldots, y^{(k+i-1)}(t)) \end{aligned}$$

for every i/o pair (u, y) of Σ, where A_i is some analytic function. Thus, by induction, one can show that $\widehat{d\mathcal{F}}_\mu^{k+i} = \widehat{d\mathcal{F}}_\mu^{k-1}$ for all $\mu \notin B$. It then follows that $\dim_{R_x} \widehat{d\mathcal{F}}_\mu \leq k$ for all $\mu \notin \widehat{B}$, where $\widehat{B} \subset \mathbb{R}^{m,\infty}$ is defined by $\widehat{B} = B \times \mathbb{R}^m \times \mathbb{R}^m \times \cdots$.

On the other hand, by Corollary 2.4 and Proposition 3.5, one knows that $\dim_{\mathbb{R}_x} \widehat{d\mathcal{F}}_\mu = \dim_{\mathbb{R}_x} \widehat{d\mathcal{F}} = n$ for all μ in a dense (in fact, even in a generic) subset of $\mathbb{R}^{m,\infty}$. Therefore, Σ cannot admit any i/o equation of order lower than n.

If Σ is a rational system, then an easy elimination argument (based on the fact that any set of $n+1$ rational functions in n variables must be algebraically dependent; see [40] for details) shows that it admits at least one i/o equation of order n; therefore, $\delta(\Sigma) = n$. □

3.2.2. i/o operators. Next we consider i/o equations for i/o operators rather than for state-space systems. By an *i/o operator* we mean an i/o map given by a convergent generating series. For a detailed definition of i/o operators, we refer the reader to [41]. We say an i/o operator F satisfies an i/o equation (17) if every $\mathcal{C}^k$ i/o pair (u, y) of F satisfies (17).

For any given operator F, we define $\delta(F)$ to be the lowest possible order of an i/o equation for F. Again, in the case when there is no i/o equation for F, $\delta(F)$ is defined to be $+\infty$.

An operator F is said to be *realized* by an initialized analytic system

$$(\mathcal{M}, x_0, \{g_0, g_1, \ldots, g_m\}, h)$$

if every i/o pair (u, y) of F satisfies the equations

$$\begin{aligned} x'(t) &= g_0(x(t))u(t) + \sum_{i=1}^m g_i(x(t))u_i(t), \quad x(0) = x_0, \\ y(t) &= h(x(t)) \end{aligned}$$

for t small enough.

Let $\lambda(F)$ be the Lie rank of F, as defined in [11], [19], or [26]. It is well known that F is realizable if and only if $\lambda(F) < \infty$, and the dimension of any canonical realization for F is $\lambda(F)$; cf. [11] and [34]. Here, by a *canonical* realization we mean a realization by an accessible and generically locally observable system.

PROPOSITION 3.8. *Assume that F is an i/o operator. Then:*

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

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

Proof. It was shown in [41] that if $\delta(F) < \infty$, then $\lambda(F) < \infty$. Thus we may assume that $\lambda(F) < \infty$, and in this case, one knows that F is realizable by some canonical system $\Sigma = (\mathcal{M}, x_0, \{g_0, g_1, \ldots, g_m\}, h)$.

By Remark 4.2 and Lemma 4.3 in [41], one knows that F admits i/o equation (17) if and only if (17) holds at any point t at which $u^{(k-1)}(t)$ exists. Combining this fact with the accessibility of the system, one sees that F admits i/o equation (17) if and only if (20) holds for Σ for all x in an open subset $\mathcal{N}$ of $\mathcal{M}$ and for all μ. On the other hand, it can be seen that (20) holds for all $x \in \mathcal{N}$ and all μ for system Σ if and only if (17) is an i/o equation for Σ as a system restricted to $\mathcal{N}$. Applying Theorem 3.7, we obtain the desired conclusion. ∎

4. Proof of Lemma 2.1.

In this section, we will prove Lemma 2.1. We will show first that there exists a generic subset $\mathcal{W}_1$ of $\mathbb{R}^{m,\infty}$ so that

$$\mathcal{F}(x) = \mathcal{F}_\mu(x) \tag{22}$$

for all x and $\mu \in \mathcal{W}_1$ and then that there is a generic subset $\mathcal{W}_2$ of $\mathbb{R}^{m,\infty}$ so that

$$\mathrm{d}\mathcal{F}(x) = \mathrm{d}\mathcal{F}_\mu(x) \tag{23}$$

for all x and all $\mu \in \mathcal{W}_2$. Then we just let $\mathcal{W} = \mathcal{W}_1 \bigcap \mathcal{W}_2$.

Proof of first part (*equation* (22)). For system (5), let

$$\mathcal{B} := \{x \in \mathcal{M} : \mathcal{F}(x) = 0\}.$$

To prove (22), we consider, for each subset $\mathcal{N}$ of the open subset $\mathcal{M} \setminus \mathcal{B}$, the set

$$\mathcal{G}_\mathcal{N} := \{\mu \in \mathbb{R}^{m,\infty} : \Psi(x, \mu) \neq 0, \forall x \in \mathcal{N}\},$$

where $\Psi(x, \mu) = (\psi_0(x, \mu), \psi_1(x, \mu), \ldots)$.

To prove the desired conclusion, it is enough to show that $\mathcal{G}_\mathcal{N}$ is open dense whenever $\mathcal{N}$ is a compact subset of $\mathcal{M} \setminus \mathcal{B}$ (since $\mathcal{M} \setminus \mathcal{B}$ can be written as a countable union of such subsets). In the following we let $\mathcal{N}$ be a fixed compact subset of $\mathcal{M} \setminus \mathcal{B}$, and we just write $\mathcal{G}$ instead of $\mathcal{G}_\mathcal{N}$. To show that $\mathcal{G}$ is dense, we need the following fact.

Let $r > 1$ be an integer. For each fixed vector $\nu^r = (\nu_0, \nu_1, \ldots, \nu_{r-1}) \in \mathbb{R}^{mr}$, we say that $\mu = (\mu_1, \mu_1, \ldots) \in \mathbb{R}^{m,\infty}$ is an extension of ν^r if $\mu_i = \nu_i$ for each $i \in [0, r-1]$.

LEMMA 4.1. *Let $x_0 \in \mathcal{M}$ and let ν^r be a fixed vector in $\mathbb{R}^{mr}$. If $\Psi(x_0, \mu) = 0$ for every extension μ of ν^r, then $x \in \mathcal{B}$.*

The proof of the above lemma will be given in Appendix A. We now return to show that $\mathcal{G}$ is dense. Take any open subset $\mathcal{U}$ of $\mathbb{R}^{m,\infty}$; without loss of generality, we may assume that $\mathcal{U} = \mathcal{U}_0 \times \mathcal{U}_1 \times \cdots \times \mathcal{U}_l \times \cdots$, where each $\mathcal{U}_i$ is an open subset of $\mathbb{R}^m$. For each integer $r > 0$, let $\mathcal{U}^r = \prod_{i=0}^{r-1} \mathcal{U}_i$. For each $\mu^r \in \mathcal{U}^r$, define

$$\mathcal{B}_{\mu^r} := \{x \in \mathcal{N} : \Psi_r(x, \mu^r) = 0\}$$

where $\Psi_r(x, \mu^r) = (\psi_0(x, \mu), \psi_1(x, \mu), \ldots, \psi_{r-1}(x, \mu))$ for any extension μ of μ^r. Note that Ψ_r is well defined because $\psi_i(x, \mu)$ does not depend on μ_j for $j \geq i$. For each finite jet ν, $\mathcal{B}_\nu$ is an analytic subset of $\mathcal{N}$, that is, a set defined by analytic equalities. As a consequence of the Weierstrass preparation theorem (in the form given for instance in [17, Thm. 2.7, Cor. 3]), one knows that analytic subsets of a compact set satisfy a descending chain condition. That is, if $\mathfrak{A}_1 \supseteq \mathfrak{A}_2 \supseteq \cdots \supseteq \mathfrak{A}_l \supseteq \cdots$ are analytic subsets of a compact set, then there exists some $r > 0$ such that $\mathfrak{A}_j = \mathfrak{A}_r$ for all $j \geq r$. From here it follows immediately that there is a minimal element $\mathcal{B}_{\bar\nu}$ of the family $\{\mathcal{B}_\nu\}$ in the sense that $\mathcal{B}_{\bar\nu} = \mathcal{B}_\nu$ whenever $\mathcal{B}_\nu \subset \mathcal{B}_{\bar\nu}$. Assume now that $\bar\nu \in \mathcal{U}^r$ provides such a minimal element.

Claim. $\mathcal{B}_{\bar\nu} = \emptyset$.

Assume that the above claim is not true. Then there exists some $x_0 \in \mathcal{N}$ such that $\Psi_r(x_0, \bar\nu) = 0$. Pick such an x_0. By Lemma 4.1, there exists some extension μ of $\bar\nu$ such that $\Psi(x_0, \mu) \neq 0$, so there exists some $l > r$ such that $\psi_l(x_0, \mu) \neq 0$. Write

$$\mu^l = (\bar\nu_0, \bar\nu_1, \ldots, \bar\nu_{r-1}, \bar\mu_r, \ldots, \bar\mu_{l-1}) \in \mathbb{R}^{ml}.$$

Note that $(\bar\nu_0, \bar\nu_1, \ldots, \bar\nu_{r-1}) \in \mathcal{U}^r$ by construction. For these fixed $\bar\nu_0, \ldots, \bar\nu_{r-1}$ and x_0, the function $\psi_l(x_0, \mu)$ does not depend on μ_j for $j \geq l$ and is analytic in $(\mu_r, \mu_{r+1}, \ldots, \mu_{l-1})$. Since it does not vanish at $(\bar\mu_r, \ldots, \bar\mu_{l-1})$, there is also some

$$(\tilde\mu_r, \tilde\mu_{r+1}, \ldots, \tilde\mu_{l-1}) \in \mathcal{U}_r \times \mathcal{U}_{r+1} \times \cdots \times \mathcal{U}_{l-1}$$

such that, for $\tilde\nu := (\bar\nu_0, \ldots, \bar\nu_{r-1}, \tilde\mu_r, \ldots, \tilde\mu_{l-1})$, $\psi_l(x_0, \tilde\mu) \neq 0$ for any extension $\tilde\mu$ of $\tilde\nu$, and hence, $\Psi_l(x_0, \tilde\nu) \neq 0$. So $x_0 \in \mathcal{B}_{\bar\nu} \setminus \mathcal{B}_{\tilde\nu}$. Also, obviously $\mathcal{B}_{\tilde\nu} \subseteq \mathcal{B}_{\bar\nu}$, since $\tilde\nu$ is an extension of $\bar\nu$. This contradicts the minimality of $\mathcal{B}_{\bar\nu}$. So we proved that $\Psi_r(x, \bar\nu) \neq 0$ for all $x \in \mathcal{N}$, as claimed.

Take any extension $\mu \in \mathcal{U}$ of μ^r to an infinite jet. Then $\Psi(x, \mu) \neq 0$ for all $x \in \mathcal{N}$, that is, $\mathcal{G} \cap \mathcal{U}$ is not empty for any open subset $\mathcal{U}$ of $\mathbb{R}^{m,\infty}$. Since $\mathcal{U}$ was arbitrary, one concludes that $\mathcal{G}$ is dense.

To prove the openness of $\mathcal{G}$, let

$$\mathcal{G}^r = \{\mu^r \in \mathbb{R}^{mr} : \ \Psi_r(x, \mu^r) \neq 0, \ \forall x \in \mathcal{N}\}.$$

By the compactness of $\mathcal{N}$, $\mathcal{G}^r$ is open. Let $\mathcal{G}_r = \mathcal{G}^r \times \mathbb{R}^{m,\infty}$. Then $\mathcal{G}_r$ is open. Since $\mathcal{G} = \bigcup_{r=1}^\infty \mathcal{G}_r$, it follows that $\mathcal{G}$ is open.

Proof of second part (*equation* (23)). Clearly $\mathrm{d}\mathcal{F}_\mu(x) \subseteq \mathrm{d}\mathcal{F}(x)$ for all x and μ, and for each $\mu \in \mathbb{R}^{m,\infty}$, $\mathrm{d}\mathcal{F}_\mu(x) = \mathrm{d}\mathcal{F}(x)$ if and only if

$$\ker \mathrm{d}\mathcal{F}(x) = \ker \mathrm{d}\mathcal{F}_\mu(x). \tag{24}$$

We now let

$$\widehat{\mathcal{B}} = \{(x, v) \in T\mathcal{M} : \ v \in \ker \mathrm{d}\mathcal{F}(x)\}.$$

Then $T\mathcal{M} \setminus \widehat{\mathcal{B}}$ is open. Let $\widehat\Psi(x, v, \mu) = (\widehat\psi_0(x, v, \mu), \widehat\psi_1(x, v, \mu), \ldots)$, where $\widehat\psi_i(x, v, \mu) = \mathrm{d}\psi_i(x, \mu)v$. To prove the desired conclusion, it is enough to show that there exists a generic subset $\mathcal{W}$ of $\mathbb{R}^{m,\infty}$ such that for any $\mu \in \mathcal{W}$, $\widehat\Psi(x, v, \mu) \neq 0$ for all $(x, v) \notin \widehat{\mathcal{B}}$. For this, it is enough to show that for any compact subset $\widehat{\mathcal{N}}$ of $T\mathcal{M} \setminus \widehat{\mathcal{B}}$, the set

$$\widehat{\mathcal{G}}_{\widehat{\mathcal{N}}} := \{\mu \in \mathbb{R}^{m,\infty} : \ \widehat\Psi(x, v, \mu) \neq 0, \ \forall\, (x, v) \in \widehat{\mathcal{N}}\}$$

is open dense. We now fix a compact subset $\widehat{\mathcal{N}}$ of $T\mathcal{M} \setminus \widehat{\mathcal{B}}$ and write $\widehat{\mathcal{G}}$ instead of $\widehat{\mathcal{G}}_{\widehat{\mathcal{N}}}$. Similar to the proof of the first part, we need the following conclusion to prove the density property of $\widehat{\mathcal{G}}$. The proof of the conclusion will again be provided in Appendix A.

LEMMA 4.2. *For any given fixed point $(x, v) \in T\mathcal{M}$, if $\widehat{\Psi}(x, v, \mu) = 0$ for all extensions μ of ν^r, for some $\nu^r \in \mathbb{R}^{mr}$, then $(x, v) \in \widehat{\mathcal{B}}$.*

To show the density of $\widehat{\mathcal{G}}$, we take any open subset $\mathcal{U}$ of $\mathbb{R}^{m,\infty}$. Again, without loss of generality, we can assume that $\mathcal{U} = \mathcal{U}_0 \times \mathcal{U}_1 \times \cdots \times \mathcal{U}_l \times \cdots$, where each $\mathcal{U}_i$ is an open subset of $\mathbb{R}^m$. Using the same notions for μ^r and $\mathcal{U}^r$ as used before, we define

$$\widehat{\mathcal{B}}_{\mu^r} := \left\{(x, v) \in \widehat{\mathcal{N}} : \ \widehat{\Psi}_r(x, v, \mu^r) = 0\right\}$$

where $\widehat{\Psi}_r(x, \mu^r) = (\widehat{\psi}_0(x, v, \mu), \widehat{\psi}_1(x, v, \mu), \ldots, \widehat{\psi}_{r-1}(x, v, \mu))$ for any extension μ of μ^r. For each finite jet ν, $\widehat{\mathcal{B}}_\nu$ is an analytic subset of $\widehat{\mathcal{N}}$ (with the obvious analytic manifold structure on the tangent bundles). Using the same argument as before, one knows that there exists a minimal element of the family $\{\widehat{\mathcal{B}}_\nu\}$. Let $\nu \in \mathcal{U}^s$ be such that $\widehat{\mathcal{B}}_\nu$ is a minimal element.

Claim. $\widehat{\Psi}_s(x, v, \nu) \neq 0$ for all $(x, v) \in \widehat{\mathcal{N}}$.

Assume that the claim is not true. Then there exists some $(x_0, v_0) \in \widehat{\mathcal{N}}$ such that $\widehat{\Psi}_s(x_0, v_0, \nu) = 0$. By Lemma 4.2, there exists some extension μ of ν such that $\widehat{\Psi}(x_0, v_0, \mu) \neq 0$. This means there exists some $l \geq s$ such that $\widehat{\psi}_l(x_0, v_0, \mu) \neq 0$. By analyticity of $\widehat{\psi}_l$, one knows that there exists some

$$(\tilde{\mu}_s, \tilde{\mu}_{s+1}, \ldots, \tilde{\mu}_{l-1}) \in \mathcal{U}_s \times \mathcal{U}_{s+1} \times \cdots \times \mathcal{U}_{l-1}$$

such that, for $\tilde{\mu} := (\nu_0, \ldots, \nu_{s-1}, \tilde{\mu}_s, \ldots, \tilde{\mu}_{l-1})$, $\widehat{\Psi}_l(x_0, v_0, \tilde{\mu}) \neq 0$. So $(x_0, v_0) \in \widehat{\mathcal{B}}_\nu \setminus \widehat{\mathcal{B}}_{\tilde{\mu}}$. Also, obviously $\widehat{\mathcal{B}}_{\tilde{\mu}} \subseteq \widehat{\mathcal{B}}_\nu$, since $\tilde{\nu}$ is an extension of ν. This contradicts the minimality of $\widehat{\mathcal{B}}_\nu$. So we proved that $\widehat{\Psi}_r(x, v, \nu) \neq 0$ for all $(x, v) \in \widehat{\mathcal{N}}$. Noting then that for any extension μ of ν, $\widehat{\Psi}(x, v, \mu) \neq 0$ for any $(x, v) \in \widehat{\mathcal{N}}$, we conclude that $\mathcal{G} \bigcap \mathcal{U} \neq \emptyset$. This proves the density of $\mathcal{G}$.

To prove the openness of $\widehat{\mathcal{G}}$, we again let

$$\widehat{\mathcal{G}}^r = \{\mu^r \in \mathbb{R}^{mr} : \ \widehat{\Psi}_r(x, v, \mu^r) \neq 0, \ \forall (x, v) \in \widehat{\mathcal{N}}\}.$$

By compactness of $\widehat{\mathcal{N}}$, $\widehat{\mathcal{G}}^r$ is open. Let $\widehat{\mathcal{G}}_r = \widehat{\mathcal{G}}^r \times \mathbb{R}^{m,\infty}$. Then $\widehat{\mathcal{G}}_r$ is open. Since $\widehat{\mathcal{G}} = \bigcup_{r=1}^{\infty} \widehat{\mathcal{G}}_r$, it follows that $\mathcal{G}$ is open. The proof of Lemma 2.1 is then complete.

Finally, we remark that also with respect to the weak topology on $\mathbb{R}^{m,\infty}$, $\mathcal{G}_{\mathcal{N}}$ and $\widehat{\mathcal{G}}_{\widehat{\mathcal{N}}}$ are still open and dense. Density is obvious, as they are dense with respect to a stronger topology. The openness of $\mathcal{G}_{\mathcal{N}}$ and $\widehat{\mathcal{G}}_{\widehat{\mathcal{N}}}$ follows from the compactness of $\mathcal{N}$ and $\widehat{\mathcal{N}}$. Thus, the conclusions of Lemma 2.1 also hold with respect to the weak topology on $\mathbb{R}^{m,\infty}$.

4.1. Proof of Theorem 3.1. In this section, we provide a proof for Theorem 3.1.

To study the observability for system (5), we consider the system

$$\begin{aligned} \xi' &= \tilde{g}_0(\xi) + \sum_{i=1}^{m} \tilde{g}_i(\xi) u_i \,, \\ y &= \tilde{h}(\xi) \,, \end{aligned} \tag{25}$$

where

$$\xi = \begin{pmatrix} x \\ z \end{pmatrix} \in \mathcal{M} \times \mathcal{M}, \quad \tilde{g}_i(\xi) = \begin{pmatrix} g_i(x) \\ g_i(z) \end{pmatrix}, \ 0 \le i \le m,$$

and $\tilde{h}(\xi) = h(x) - h(z)$. Clearly, $x \not\sim_u z$ for system (5) if and only if $\xi \not\sim_u 0$ for system (25). Thus, to prove Theorem 3.1, it is enough to establish the following conclusion.

PROPOSITION 4.3. *Assume that for an analytic system* (5), *the i/o map induced by the zero initial state is a zero map, that is,* $h \circ \varphi(t, 0, u) = 0$ *for all* t *and all* u. *Then for any* $T > 0$, *the set*

$$\mathfrak{G} = \{u \in \mathcal{C}^\infty[0,\ T] :\ x \not\sim_u 0 \text{ if } x \not\sim 0\}$$

is a generic subset of $\mathcal{C}^\infty[0,\ T]$.

Proof. Let $\mathcal{B} := \{x :\ x \sim 0\}$. Then $\mathcal{M} \setminus \mathcal{B}$ is an open subset of $\mathcal{M}$. To prove Proposition 4.3, it is enough to show that for every compact subset $\mathcal{N}$ of $\mathcal{M} \setminus \mathcal{B}$ the set

$$\mathfrak{G}_{\mathcal{N}} = \{u \in \mathcal{C}^\infty[0,\ T] :\ x \not\sim_u 0 \text{ for all } \ x \in \mathcal{N}\}$$

is an open dense subset of $\mathcal{C}^\infty[0,\ T]$.

Note that for $u \in \mathcal{C}^\infty$, $x \not\sim_u 0$ if $\psi_i(x,\ \mu) \ne 0$ for some i, where $\mu = (\mu_0,\ \mu_1,\ \ldots) \in \mathbb{R}^\infty$ with $\mu_i = u^{(i)}(0)$, and ψ_i is as defined in (7) for each i. Also, by Theorem 3-1.5 in [19], one knows that for $x \in \mathcal{M}$, if $\mathcal{F}(x) = 0$, then $x \in \mathcal{B}$. This means that for each $x \in \mathcal{N}$, $\mathcal{F}(x) \ne 0$. Thus, by Lemma 2.1, there exists a dense subset $\mathcal{G}$ of $\mathbb{R}^{m,\infty}$ such that $\Psi(x,\ \mu) \ne 0$ for all $x \in \mathcal{N}$ and all $\mu \in \mathcal{G}$.

To complete the proof of Proposition 4.3, we need to show that $\mathfrak{G}_{\mathcal{N}}$ is an open dense subset of $\mathcal{C}^\infty[0,\ T]$.

Take $\bar{\omega} \in \mathcal{C}^\infty[0,\ T]$, and let $\mathcal{U}$ be a neighborhood of $\bar{\omega}$. Without loss of generality when showing the density of $\mathfrak{G}_{\mathcal{N}}$, we may assume that

$$\mathcal{U} = \left\{ \omega \in \mathcal{C}^\infty[0,\ T] :\ \max_{0 \le i \le k} |w^{(i)}(t) - \bar{\omega}^{(i)}(t)| < \delta,\ t \in [0,\ T] \right\},$$

for some integer $k \ge 0$ and some $\delta > 0$.

Let $\bar{\mu} := (\bar{\mu}_0,\ \bar{\mu}_1,\ \ldots)$, where $\bar{\mu}_i := \bar{\omega}^{(i)}(0)$, and let $\mathcal{W}$ be the open subset of $\mathbb{R}^{m,\infty}$ defined by

$$\mathcal{W} = \left\{ \mu \in \mathbb{R}^{m,\infty} :\ |\mu_i - \bar{\mu}_i| < \delta e^{-T},\ i \ge 0 \right\}.$$

As $\mathcal{W} \bigcap \mathcal{G} \ne \emptyset$, there exists some $\nu \in \mathcal{W}$ such that $\Psi(x,\ \nu) \ne 0$ for all $x \in \mathcal{N}$. By compactness of $\mathcal{N}$, there exists some $r > 0$ such that

$$\Psi_r(x,\ \nu^r) \ne 0, \ \text{for any } x \in \mathcal{N}. \tag{26}$$

Without loss of generality, one can always assume that $r > k$.

Now let $\bar{\omega}_0(t) := \bar{\omega}(t) - \sum_{i=0}^{r-1} \frac{\bar{\mu}_i}{i!} t^i$. Note then that $\bar{\omega}_0^{(i)}(0) = 0$ for all $0 \le i \le r - 1$.

Finally, we define

$$\omega(t) := \bar{\omega}_0(t) + \sum_{i=0}^{r-1} \frac{\nu_i}{i!} t^i.$$

Then, for $0 \le i \le k$ and $0 \le t \le T$, we have

$$|\omega^{(i)}(t) - \bar{\omega}^{(i)}(t)| \le \sum_{j=i}^{r-1} \frac{|\nu_j - \bar{\mu}_j|}{(j-i)!} t^{j-i} \le \sum_{j=0}^{\infty} \frac{|\nu_{j+i} - \bar{\mu}_{j+i}|}{j!} t^j$$
$$< \delta e^{-T} e^t \le \delta .$$

Thus, $\omega \in \mathcal{U}$.

On the other hand, (26) implies that for every $x \in \mathcal{N}$, there exists some $i \le r-1$, such that

$$\left.\frac{d^i}{dt^i}\right|_{t=0} h(\varphi(t,x,w)) \neq 0 .$$

From here it follows that $x \not\sim_\omega 0$ for every $x \in \mathcal{N}$, that is, $\omega \in \mathfrak{G}_{\mathcal{N}}$. This proves that $\mathfrak{G}_{\mathcal{N}}$ is dense.

We then conclude the proof of Proposition 4.3 by noting that the openness of $\mathfrak{G}_{\mathcal{N}}$ follows from the compactness of $\mathcal{N}$. ∎

Remark 4.4. Note that the above proof only depends on the first half of Lemma 2.1, i.e., formula (22), and the proof of (22) is fairly straightforward (though it calls upon some notions and elementary results from the theory for generating series).

Proof of Proposition 3.2. As indicated in the beginning of this section, it is enough to show the following:

> Assume that for an analytic system (5), the i/o map induced by the zero initial state is a zero map, that is, $h \circ \varphi(t, 0, u) = 0$ for all t and all u. Then there exists some analytic input u such that $x \not\sim_u 0$ for all $x \not\sim 0$.

Proof. Consider the following open subset of $\mathbb{R}^{m,\infty}$:

$$\mathcal{U} = \mathcal{U}_0 \times \mathcal{U}_1 \times \mathcal{U}_2 \times \cdots$$

where $\mathcal{U}_i = (-1, 1)$ for all $i \ge 0$. By Lemma 2.1, there is at least one jet μ in $\mathcal{U}$ such that $\mathcal{F}_\mu(x) = \mathcal{F}(x)$, from which it follows that

$$\Psi(x, \mu) \neq 0, \quad \forall x \not\sim 0. \tag{27}$$

Now let

$$u(t) = \sum_{i=0}^{\infty} \frac{\mu_i}{i!} t^i.$$

Then u is an analytic function and $u^{(i)}(0) = \mu_i$. By (27), one knows that $x \not\sim_u 0$ for all $x \not\sim 0$. ∎

5. Main results for discrete-time systems.

5. Main results for discrete-time systems. In this section, we discuss our main results for discrete-time systems.

5.1. Basic definitions for discrete-time systems.

We consider *analytic systems* as in (4), where for each t, $x(t) \in \mathcal{M}$, an analytic manifold, and $u(t) \in \mathbb{R}^m$. We assume that $h : \mathcal{M} \to \mathbb{R}^p$ and $f : \mathcal{M} \times \mathbb{R}^n \to \mathcal{M}$ are analytic. If $\mathcal{M} = \mathbb{R}^n$ and the entries of f and h are rational functions with no (real) poles, then we call (4) a *rational system*. A system Σ will be called *reversible* if $f(\cdot, u)$ is one-to-one, for each fixed $u \in \mathbb{R}^m$. (Reversible systems are a more general class than the systems usually called *invertible* in the discrete-time controllability literature, for which one makes the stronger requirement that $f(\cdot, u)$ is a diffeomorphism of $\mathcal{M}$, for each u. Invertible systems arise naturally through the sampling of continuous-time systems in digital control, by integrating flows over a sampling period; their controllability properties were studied in, among other papers, [20], [12], [21], [24], [22], [23], [1].)

For each control sequence $\omega \in \mathbb{R}^{km}$, we define $f^\omega : \mathcal{M} \to \mathcal{M}$ inductively by $f^e(x) = x$ for the empty sequence e and $f^{\omega u}(x) = f(f^\omega(x), u)$. We also let $h^\omega := h \circ f^\omega$. For $\mu = (\mu_0, \mu_1, \ldots) \in \mathbb{R}^{m,\infty}$, we let $H^\mu(x) := (h(x), h^{\mu_0}(x), h^{\mu_0\mu_1}(x), \ldots)$.

Two states p and q are said to be distinguished by $\mu \in \mathbb{R}^{m,\infty}$, denoted by $p \not\sim_\mu q$, if $H^\mu(p) \neq H^\mu(q)$. A discrete-time system is said to be *observable* if any two distinct states p and q can be distinguished by some μ. See [27] for a detailed introduction to observability and related concepts and, in particular, [25] for results on observability of discrete-time systems.

For an analytic system, we define the *observation space* of Σ as the following subspace of the space of analytic functions defined on $\mathcal{M}$:

$$\mathcal{F} = \operatorname{span}_{\mathbb{R}} \{h^\omega : \ \omega \in \mathbb{R}^{mr}, r \geq 0\}.$$

This space plays an important role in studying observability of discrete-time systems; see, e.g., [28] and [27]. See also [16] for related algebraic structures.

Associated with the above space, for each $x \in \mathcal{M}$ we let $\mathrm{d}\mathcal{F}(x)$ be the subspace of the cotangent space at x defined by

$$\mathrm{d}\mathcal{F}(x) = \{\mathrm{d}\alpha(x) : \ \alpha \in \mathcal{F}\}.$$

In analogy to the continuous-time case, we define, for each $\mu = (\mu_0, \mu_1, \ldots) \in \mathbb{R}^{m,\infty}$, the following subspace $\mathcal{F}_\mu$ of analytic functions:

$$\mathcal{F}_\mu = \operatorname{span}_{\mathbb{R}} \{h, h^{\mu_0}, h^{\mu_0\mu_1}, \ldots\}.$$

For each $\mu \in \mathbb{R}^{m,\infty}$ and each $x \in \mathcal{M}$, we also consider

$$\mathrm{d}\mathcal{F}_\mu(x) = \{\mathrm{d}\alpha : \ \alpha \in \mathcal{F}_\mu\}.$$

Clearly, $\mathcal{F} = \sum_\mu \mathcal{F}_\mu$ and $\mathrm{d}\mathcal{F}(x) = \sum_\mu \mathrm{d}\mathcal{F}_\mu(x)$ for each x. Here we will need the following result.

LEMMA 5.1. *Assume that* (4) *is reversible and observable. Then there exists a generic subset* $\mathcal{W}$ *of* $\mathbb{R}^{m,\infty}$ *such that for each* $\mu \in \mathcal{W}$,

$$\mathrm{d}\mathcal{F}(x) = \mathrm{d}\mathcal{F}_\mu(x) = \mathbb{R}^n, \tag{28}$$

for all x *in an open dense subset of* $\mathcal{M}$.

The proof will be given later; it will rely on a result about universal inputs for discrete-time systems that is presented in the next section.

Assume now that $\mathcal{M} = \mathbb{R}^n$. Still using the notation used in §2.2, we introduce

$$\widehat{\mathrm{d}}\mathcal{F} := \operatorname{span}_{\mathbb{R}_x}\{\mathrm{d}\alpha : \ \alpha \in \mathcal{F}\}, \quad \widehat{\mathrm{d}}\mathcal{F}_\mu := \operatorname{span}_{\mathbb{R}_x}\{\mathrm{d}\alpha : \ \alpha \in \mathcal{F}_\mu\}.$$

From the lemma and using an argument analogous to that used in proving Corollary 2.4, we have the following corollary.

COROLLARY 5.2. *For an analytic, reversible, and observable system,* $\widehat{\mathrm{d}}\mathcal{F}_\mu = \widehat{\mathrm{d}}\mathcal{F}$ *for all* μ *in a generic set of* $\mathbb{R}^{m,\infty}$.

5.2. Observability and universal inputs. An input sequence is said to be a *universal input* of a discrete-time system Σ if it distinguishes every distinguishable pair of Σ.

THEOREM 5.3. *Assume that* (4) *is analytic, reversible, and observable. Then the universal inputs of* (4) *form a generic subset of* $\mathbb{R}^{m,\infty}$.

Proof. First of all, we let

$$\mathcal{D} := \{(x, x) : x \in \mathcal{M}\} \subseteq \mathcal{M} \times \mathcal{M}.$$

By observability, every pair $(x, z) \in (\mathcal{M} \times \mathcal{M}) \setminus \mathcal{D}$ is a distinguishable pair of (4). For each $\nu \in \mathbb{R}^{mr}$, we let $\lambda_r(x, z, \nu) = h^\nu(x) - h^\nu(z)$, and we also let $\lambda_0(x, z) = h(x, z)$. For each $\mu = (\mu_0, \mu_1, \ldots)$, we define

$$\Lambda(x, z, \mu) = (\lambda_0(x, z), \lambda_1(x, z, \mu_0), \lambda_2(x, z, \mu_1\mu_0), \ldots).$$

To prove the desired conclusion, it is enough to show that for each compact subset $\mathcal{N}$ of $(\mathcal{M} \times \mathcal{M}) \setminus \mathcal{D}$, the set $\mathcal{G}_\mathcal{N}$ defined by

$$\mathcal{G}_\mathcal{N} := \{\mu \in \mathbb{R}^{m,\infty} : \Lambda(x, z, \mu) \neq 0, \forall (x, z) \in \mathcal{N}\}$$

is an open dense subset of $\mathbb{R}^{m,\infty}$.

For each open subset $\mathcal{U}$ of $\mathbb{R}^{m,\infty}$ given by $\mathcal{U}_0 \times \mathcal{U}_1 \times \cdots$, consider, for each $\nu \in \mathcal{U}^s = \prod_{i=0}^{s-1} \mathcal{U}_i$, the subset $\mathcal{B}_\nu$ of $\mathcal{N}$ defined by

$$\mathcal{B}_\nu = \{(x, z) \in \mathcal{N} : \Lambda_r(x, z, \nu) = 0\},$$

where $\Lambda_r(x, z, \nu) = (\lambda_0(x, z), \lambda_1(x, z, \nu_0), \ldots, \lambda_s(x, z, \nu))$. Using the same argument as that employed in the proof of Lemma 2.1, we know that there exists a minimal element $\mathcal{B}_{\bar\nu}$ of the family $\{\mathcal{B}_{\bar\nu}\}$. Suppose $\bar\nu \in \mathcal{U}^r$. We next show that $\bar\nu$ distinguishes every pair $(x, z) \in \mathcal{N}$. Assume that there would exist a pair $(x_0, z_0) \in \mathcal{N}$ such that $x_0 \sim_{\bar\nu} z_0$. Since (4) is reversible, $x_1 \neq z_1$, where $x_1 = f^\nu(x_0)$ and $z_1 = f^\nu(z_0)$. By observability of (4), one knows that there exists some $\tilde\nu \in \mathbb{R}^{ms}$ such that $x_1 \not\sim_{\tilde\nu} z_1$. Let $\widehat\nu = \tilde\nu\bar\nu$ (concatenation of sequences); then it follows that $\Lambda_{r+s}(x_0, z_0, \bar\nu\tilde\nu) \neq 0$. By the analyticity of Λ_{r+s} when fixing x_0, z_0 and $\bar\nu$, one knows that there exists some $\widehat\nu \in \mathcal{U}_r \times \cdots \times \mathcal{U}_{r+s-1}$ such that $\Lambda_{r+s}(x_0, z_0, \bar\nu\widehat\nu) \neq 0$. This implies that $(x_0, z_0) \in \mathcal{B}_{\bar\nu} \setminus \mathcal{B}_{\bar\nu\widehat\nu}$, which, in turn, implies that $\mathcal{B}_{\bar\nu\widehat\nu}$ is a proper subset of $\mathcal{B}_{\bar\nu}$, contradicting the assumed minimality of $\mathcal{B}_{\bar\nu}$. Thus, we showed that $\Lambda_r(x, z, \bar\nu) \neq 0$ for any $(x, z) \in \mathcal{N}$. Clearly, any extension μ of $\bar\nu$ in $\mathcal{U}$ is an element of $\mathcal{G}_\mathcal{N}$. This shows that $\mathcal{G}_\mathcal{N} \bigcap \mathcal{U} \neq \emptyset$ for any open subset $\mathcal{U}$ of $\mathbb{R}^{m,\infty}$. The density of $\mathcal{G}_\mathcal{N}$ is thus proved.

Again as in the proof of Lemma 2.1 for the continuous case, $\mathcal{G}_\mathcal{N}$ is open since $\mathcal{N}$ is compact. $\square$

In the statement of Theorem 5.3, we assumed more than we did in its continuous counterpart, Theorem 3.1 (and also concluded slightly less). One of the extra conditions is observability. We needed to impose this because the counterpart of Lemma 4.1 is not available in the discrete-time case. The discrete case analogy would be that any distinguishable pair is again carried to a distinguishable pair by the flow of the system, no matter which input is applied. Unfortunately, this not true in general. The following example, suggested by F. Albertini, shows that distinguishable pairs can be carried to indistinguishable pairs. (Note that this can never happen with analytic continuous-time systems.)

Example 5.4. Consider the system

$$x(t+1) = x(t) + 1, \quad y(t) = h(x(t)), \tag{29}$$

where $h(x)$ is defined by

$$h(x) = \begin{cases} \frac{\sin \pi x}{\pi x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Clearly the system is analytic and reversible. However, the distinguishable pair (0, 1) is carried to an indistinguishable pair after $t = 1$.

Proof of Lemma 5.1. To obtain the desired conclusion, it is enough to show that (28) holds in an open dense subset of $\mathcal{M}$ for every universal input μ (since universal inputs themselves form a generic subset).

Fix any universal input μ. By observability, one knows that $H(\cdot, \mu)$ is a one-to-one map. Let $k = \max_p \dim \mathcal{F}_\mu(p)$. It is sufficient to show that $k = n$. But this is an immediate consequence of Lemma B.1 (see Appendix B), applied to $\{h, h^{\mu_0}, h^{\mu_0\mu_1}, \ldots\}$ seen as a family of maps. $\square$

5.2.1. Orders of i/o equations. We say that the discrete-time system (4) *admits the i/o equation* such as that if (3) holds for all input/output pairs of (4) (for $t \geq r$ and any possible initial state $x(0)$). The *order* of the equation is r if

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

is not a zero function. For any given system Σ, we let $\delta(\Sigma)$ be the lowest possible order of an i/o equation that Σ admits. If there is no such equation, $\delta(\Sigma)$ is defined to be ∞. Following the same outline as in the proof of Theorem 3.7 but now using Lemma 5.1, we conclude as follows.

THEOREM 5.5. *Let Σ be an n-dimensional analytic system. Assume, further, that Σ is reversible and observable. Then $\delta(\Sigma) \geq n$. If, in addition, Σ is a rational system, then $\delta(\Sigma) = n$.*

Remark 5.6. The result in Lemma 5.1 is *false* if the assumption of reversability is dropped, as discussed in [28]. As a consequence of this, the above conclusions may be false without the invertibility assumption. To illustrate this, consider the following system of dimension 3:

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

This is an observable polynomial system. However, it admits an equation of order 2:

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

Note that this system is not reversible.

Appendix A. Proofs of two lemmas. In this appendix, we will prove Lemmas 4.1 and 4.2. For this, we need to recall some basic definitions and properties of i/o operators defined by convergent generating series. For a detailed study of generating series and i/o operators, we refer the reader to [41].

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 $i_1 i_2 \ldots i_k$, where $i_s \in I$ for each s. We use I^0 to denote the set whose only element is the empty sequence ϕ. Let $I^* = \bigcup_{k\geq 0}^{\infty} I^k$.

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 $\eta_\iota = \eta_{i_1}\eta_{i_2}\ldots\eta_{i_l}$ for each multiindex $\iota = i_1 i_2 \ldots i_l$. 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, $M \geq 0$ such that

$$|\langle c, \eta_\iota \rangle| \leq KM^k k! \quad \text{for each } \iota \in I^k \text{ and each } k \geq 0. \tag{30}$$

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 $\max\{\|u_i\|_1 :, 1 \leq i \leq m\}$ and $\|u\|_\infty$ for $\max\{\|u_i\|_\infty :, 1 \leq i \leq m\}$ where u_i is the ith 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

$$V_{i_1 \cdots i_{l+1}}[u](t) = \int_0^t u_{i_1}(s) V_{i_2 \ldots i_{l+1}}(s)\, ds,$$

where $V_\phi = 1$ and u_i is the ith 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:

$$F_c[u](t) = \sum \langle c, \eta_\iota \rangle V_\iota[u](t). \tag{31}$$

If the series is convergent and (30) holds, then it is known that for any

$$T < (\|u\|_\infty (Mm + M))^{-1},$$

the series (31) converges uniformly and absolutely for all $t \in [0, T]$. Let

$$\mathcal{V}_T := \{u \in L_\infty^m : \|u\|_\infty T < (Mm + M)^{-1}\}.$$

We refer the reader to [41] for the proof of the following lemmas.

LEMMA A.1. *Assume that c is a convergent power series. 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 A.2. *Suppose c is a convergent series. Then $F_c[u]$ is analytic if $u \in \mathcal{V}_T$ is analytic.*

For each convergent series c, we let, for each $\mu \in \mathbb{R}^{m,\infty}$ and each integer $i \geq 0$,

$$c_i(\mu) = \left.\frac{d^i}{dt^i}\right|_{t=0} F_c[u](t), \tag{32}$$

where u is any smooth input with $u^{(i)}(0) = \mu_i$. Note that $c_i(\mu)$ is a polynomial in μ and $c_i(\mu)$ doesn't depend on μ_j for $j \geq i$.

By Lemma 2.1 in [39], one knows that for a convergent series c, $F_c[u] = 0$ for every piecewise constant input u if and only if $c = 0$. On the other hand, it is not hard to see that for each piecewise constant function u, there exists a sequence $\{u_j\}$ of analytic functions such that $\|u_j\| \leq \|u\|$ and $u_j \to u$ as $j \to \infty$ in the L_1 norm. By Lemma A.2, one concludes that $F_c[u] = 0$ for every analytic input u if and only if $c = 0$. Since $F_c[u]$ is analytic if u is analytic, it then follows from (32) that for an analytic u with $u^{(i)}(0) = \mu_i$, $F_c[u] = 0$ if and only if $c_i(\mu) = 0$ for all $i \geq 0$. Thus we conclude that $c = 0$ if and only if $c_i(\mu) = 0$ for all μ and all i. To prove the desired conclusions, we need the following well-known fact.

LEMMA A.3. *Assume that f is a continuous function defined on $[0,\, t_0]$ for some $t_0 > 0$. Then for any given integer r and any vector $(w_0,\, w_1,\, \ldots, w_r)$, there exists a L_∞-bounded sequence of analytic functions f_j defined on $[0,\, t_0]$, such that $f_j^{(i)}(0) = w_i$ for all $i \leq r$ and f_j converges to f in the L_1 norm.*

Proof. For the given vector, let

$$\widehat{f}(t) = f(t) - \sum_{i=0}^{r} \frac{w_i t^i}{i!}.$$

Without loss of generality, one may assume that $\widehat{f}(0) = 0$. Otherwise, one can always choose a L_∞-bounded sequence of continuous functions $\widehat{f}_j$ converging to $\widehat{f}$ in the L_1 norm and such that $\widehat{f}_j(0) = 0$. Now one may apply Lemma 4.3 in [41] to $\widehat{f}$ to conclude that there exists a sequence $\tilde{f}_j$ converging to $\widehat{f}$ uniformly (hence also in L_1 norm) with the property that $\tilde{f}_j^{(i)}(0) = 0$ for all $i \leq r$. Then the functions

$$f_j(t) := \tilde{f}_j(t) + \sum_{i=0}^{r} \frac{w_i t^i}{i!}$$

give the desired sequence. □

Combining the above conclusion and Lemma A.1, one proves the following.

LEMMA A.4. *Assume that c is a convergent series and that r is an integer. Let $\bar{\mu}^r$ be a given vector in $\mathbb{R}^{mr}$. If for every extension μ of $\bar{\mu}^r$, $c_i(\mu) = 0$ for all i, then $c = 0$.*

Proof of Lemma 4.1. For analytic system (5) and for each $x \in \mathcal{M}$, we define a generating series by letting

$$\langle c^x, \eta_{i_1}\eta_{i_2}\ldots\eta_{i_r}\rangle = L_{g_{i_r}}\ldots L_{g_{i_2}}L_{g_{i_1}}h(x). \tag{33}$$

By Lemma 4.2 in [36], such a series is always convergent, and it follows from Theorem 3-1.5 in [19] that for any $\mu \in \mathbb{R}^{m,\infty}$,

$$\psi_i(x,\, \mu) = c_i(x,\, \mu), \tag{34}$$

where

$$c_i(x,\, \mu) = \left.\frac{d^i}{dt^i}\right|_{t=0} F_{c^x}[u](t)\,.$$

The conclusion of Lemma 4.1 then follows from Lemma A.4.

Proof of Lemma 4.2. For analytic system (5), instead of considering the series defined by (33), we consider, for each $(x, v) \in T\mathcal{M}$, the series defined by

$$\langle d(x,\, v),\, \eta_{i_1}\eta_{i_2}\dots\eta_{i_r}\rangle \,=\, \mathrm{d}L_{g_{i_r}}\dots L_{g_{i_2}}L_{g_{i_1}}h(x)v. \tag{35}$$

Claim. For each (x, v), the series $d(x, v)$ is a convergent series.

First of all, by Lemma 4.2 in [36], there is some constant $M_0 > 0$ such that for $g_0(x), g_1(x), \dots, g_m(x)$, $v \in T_x\mathcal{M}$, there exists some $M_0 > 0$ such that

$$\left|\mathrm{d}L_{g_{i_1}}L_{g_{i_2}}\dots L_{g_{i_r}}h(x)v\right| \,=\, \left|L_vL_{g_{i_1}}L_{g_{i_2}}\dots L_{g_{i_r}}h(x)v\right| \le M_0^{r+1}(r+1)!. \tag{36}$$

It is then not hard to see that there exist some constants K and $M > M_0$ such that

$$\left|\mathrm{d}L_{g_{i_1}}L_{g_{i_2}}\dots L_{g_{i_r}}h(x)v\right| \,\le\, KM^r r!\,,$$

for all $r > 0$. Therefore $d(x, v)$ is a convergent series for each pair (x, v).

For each smooth input u with $u^{(i)}(0) = \mu_i$, let

$$d_i(x,\, v,\, \mu) \,=\, \left.\frac{d^i}{dt^i}\right|_{t=0} F_{d(x,\,v)}[u](t)\,.$$

Then it follows from (34) that

$$\mathrm{d}\psi_i(x,\, \mu) \,=\, \mathrm{d}c_i(x,\, \mu)\,,$$

from which it follows that

$$\mathrm{d}\psi_i(x,\, \mu)v \,=\, d_i(x,\, v,\, \mu)\,.$$

Applying Lemma A.4 to the series $d(x, \mu)$, one obtains the desired conclusion of Lemma 4.2.

Appendix B. A simple consequence of the rank theorem The next result is a simple and well-known consequence of the rank theorem; we include its proof as it seems difficult to find a precise reference. (We provide a somewhat stronger form than needed, which applies in more generality, including to nonobservable systems.)

LEMMA A.5. *Assume that $\mathcal{H} = \{h_\lambda : Z \to \mathbb{R}, \lambda \in \Lambda\}$ is a family of continuously differentiable real-valued functions on an n-dimensional differentiable manifold Z, parameterized by a set Λ. Then there exists an open dense subset $Z_0 \subseteq Z$ with the following property. For each $z_0 \in Z_0$ there exist an integer $r = r(z_0)$, an open neighborhood V of z_0 in Z_0, and parameter values $\lambda_1, \dots, \lambda_r$, so that, for each parameter $\lambda \in \Lambda$,*

$$h_\lambda(z) \,=\, F_\lambda(h_{\lambda_1,\dots,\lambda_r}(z)) \;\; \forall z \in V\,,$$

where $h_{\lambda_1,\dots,\lambda_r}(z) := (h_{\lambda_1}(z), \dots, h_{\lambda_r}(z))$ and F_λ is some C^1 function from some neighborhood $\mathcal{U}$ of $h_{\lambda_1,\dots,\lambda_r}(V)$ to $\mathbb{R}$. Moreover, the rank of the differential of $h_{\lambda_1,\dots,\lambda_r}(z)$ is r at all $z \in V$ (so the nonempty fibers $h^{-1}_{\lambda_1,\dots,\lambda_r}(q)$ intersect V at submanifolds of dimension $n - r$). In particular, if it is known that $z \mapsto (h_\lambda(z), \lambda \in \Lambda)$ is one-to-one on any open subset of Z, then $r(z_0) = n$ for some $z_0 \in Z_0$.

Proof. Consider for any s and any $\lambda_1, \dots, \lambda_s$ the rank $\rho_{\lambda_1,\dots,\lambda_s}(z)$ of the differential of $h_{\lambda_1,\dots,\lambda_s}$ at z, and let $\rho(z)$ be the maximum possible value of this rank over all s and $\lambda_1, \dots, \lambda_s$. A point z is *regular* if $\rho(z)$ is constant in a neighborhood of z. The regular points form an open set by definition, and it is an easy exercise to show, by induction on $n, n-1, \ldots, 1$ that the set Z_0 of such points is also dense. Now pick any z_0 in Z_0, and let $\rho(\dot{z}_0) = r$. By definition of ρ, there are parameters $\lambda_1, \ldots, \lambda_r$ so that $\rho_{\lambda_1,\ldots,\lambda_r}(z) = r$ for all z in some neighborhood of z_0. By the rank theorem, there are local changes of coordinates in Z so that, in some neighborhood V of z_0, $h_{\lambda_i}(z) = z_i$ for $i = 1, \ldots, r$, and without loss of generality one may assume that $\rho_{\lambda_1,\ldots,\lambda_r}(z) = r$ for all z in this same V. Now pick any $\lambda \in \Lambda$. Let $f = h_\lambda$. If it were the case that $\frac{\partial f}{\partial z_j}(z)$ is nonzero for some $z \in V$ and some $j > r$, then the map $h_{\lambda_1,\ldots,\lambda_s,\lambda}$ would have rank $r+1$ at z, contradicting the choice of V. It follows that h_λ depends only on $z_1, \ldots, z_r$ on this neighborhood, as desired. □

Remark A.6. Observe that, when dealing with analytic mappings and Z connected, the rank is constant on regular points, and one could pick the elements $\lambda_1, \ldots, \lambda_r$ globally on an open dense set. Also, in general this argument shows that locally there are always n control sequences that (locally) distinguish states, even in the nonanalytic case.

REFERENCES

[1] F. Albertini and E. D. Sontag, *Discrete-time transitivity and accessibility: Analytic systems*, SIAM J. Control Optim., 31 (1993), pp. 1599–1622.

[2] Z. Bartosiewicz, *Rational systems and observation fields*, Systems Control Lett., 9 (1987), pp. 379–386.

[3] ———, *Minimal polynomial realizations*, Math. Control Signals Systems, 1 (1988), pp. 227–231.

[4] G. Conte, G. H. Moog, and A. Perdon, *Un théorème sur la représentation entrée-sortie d'un système non linéaire*, C. R. Acad. Sci. Paris, Sér I. Math., 307 (1988), pp. 363–366.

[5] J.-M. Coron, *Linearized control systems and applications to smooth stabilization*, SIAM J. Control Optim., 32 (1994), pp. 358–386.

[6] P. Crouch and F. Lamnabhi-Lagarrigue, *State space realizations of nonlinear systems defined by input-output differential equations*, in Proc. 8th Internat. Conf. Analysis Optimiz. Systems, Antibes, 1988, A. Bensoussan and J. L. Lions, eds., Berlin, 1988, Springer-Verlag, New York, pp. 138–149.

[7] S. Diop, *Elimination in control theory*, Math. Control Signals Systems, 4 (1991), pp. 17–32.

[8] ———, *Closedness of morphisms of differential algebraic sets. Application to systems theory*, Forum Math., 1 (1992), pp. 1–15.

[9] S. Diop and M. Fliess, *Nonlinear observability, identifiability, and persistent trajectories*, in Proc. 30th IEEE Conference on Decision and Control, Brighton, IEEE Publications, Piscataway, NJ, 1991, pp. 714–719.

[10] ———, *On nonlinear obervability*, in Proceedings of the first European Control Conference, C. Commault, D. Normand-Cyrot, L. J. M. Dion, M. Fliess, A. Titli, A. Benveniste, and I. Landau, eds., 1991, pp. 152–157.

[11] 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.

[12] M. Fliess and D. Normand-Cyrot, *A group-theoretic approach to discrete-time nonlinear controllability*, in Proc. IEEE Conf. Dec. Control, IEEE Publications, Piscataway, NJ, 1981, pp. 551–557.

[13] S. T. Glad, *Nonlinear state space and input output descriptions using differential polynomials*, in New Trends in Nonlinear Control Theory, J. Descusse, M. Fliess, A. Isidori, and M. Leborgne, eds., Springer-Verlag, Heidelberg, 1989, pp. 182–189.

[14] M. Golubitsky and V. Guillemin, *Stable Mapping and Their Singularities*, Springer-Verlag, New York, 1973.

[15] O. M. Graselli and A. Isidori, *Deterministic state reconstruction and reachability of bilinear control processes*, in Proc. Joint Automatic Control Conf., San Francisco, June 22–25 1977.

[16] J. W. Grizzle, *Linear algebraic framework for the analysis of discrete-time nonlinear systems*, SIAM J. Control Optim., 31 (1993), pp. 1026–1044.

[17] M. Hervé, *Several Complex Variables, Local Theory*, Oxford University Press, London, 1963.

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

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

[20] B. JAKUBCZYK, *Invertible realizations of nonlinear discrete time systems*, in Proc. Princeton Conf. Info. Sci. and Syst., 1980, pp. 235–239.

[21] B. JAKUBCZYK AND D. NORMAND-CYROT, *Orbites de pseudo groupes de diffeophismes et commandabilité des systèmes non linéaires en temps discret*, in C. R. Acad. Sci. Paris Sér. I. Math., 298, (1984), pp. 257–260.

[22] B. JAKUBCZYK AND E. D. SONTAG, *Controllability of nonlinear discrete-time systems: A Lie-algebraic approach*, SIAM J. Control Optim., 28 (1990), pp. 1–33.

[23] A. MOKKADEM, *Orbites de semi-groupes de morphismes réguliers et systèmes non linéaires en temps discret*, Forum Math., 1 (1989), pp. 359–376.

[24] S. MONACO AND D. NORMAND-CYROT, *Invariant distributions for nonlinear discrete-time systems*, Systems Control Lett., 5 (1984), pp. 191–196.

[25] H. NIJMEIJER, *Observability of autonomous discrete-time nonlinear systems: A geometric approach*, Internat. J. Control, 36 (1982), pp. 867–874.

[26] H. NIJMEIJER AND A. VAN DER SCHAFT, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.

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

[28] ———, *Polynomial Response Maps*, Springer-Verlag, Berlin, New York, 1979.

[29] ———, *A concept of local observability*, Systems Control Lett., 5 (1984), pp. 41–47.

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

[31] ———, *Universal nonsingular controls*, Systems Control Lett., 19 (1992), pp. 221–224; Errata 20 (1993), p. 77.

[32] E. D. SONTAG AND Y. WANG, *I/O equations for nonlinear systems and observation spaces*, in Proc. 30th IEEE Conf. Decision and Control, Brighton, UK, Dec. 1991, IEEE Publications, Piscataway, NJ, 1991, pp. 720–725.

[33] ———, *Orders of I/O equations and uniformly universal inputs*, in Proc. 33rd IEEE Conf. Decision and Control, Orlando, Dec. 1994, IEEE Publications, Piscataway, NJ, 1994, pp. 1270–1275. (Journal version in preparation.)

[34] H. J. SUSSMANN, *A proof of the realization theorem for convergent generating series of finite lie rank*, submitted.

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

[36] ———, *Lie brackets and local controllability: A sufficient condition for scalar-input systems*, SIAM J. Control Optim., 21 (1983), pp. 686–713.

[37] A. V. VAN DER SCHAFT, *On realizations of nonlinear systems described by higher-order differential equations*, Math Systems Theory, 19 (1987), pp. 239–275.

[38] ———, *Representing a nonlinear state space system as a set of higher-order differential equations in the inputs and outputs*, Systems Control Lett., 12 (1989), pp. 151–160.

[39] Y. WANG AND E. D. SONTAG, *On two definitions of observation spaces*, Systems Control Lett., 13 (1989), pp. 279–289.

[40] ———, *Algebraic differential equations and rational control systems*, SIAM J. Control Optim., 30 (1992), pp. 1126–1149.

[41] ———, *Generating series and nonlinear systems: Analytic aspects, local realizability and i/o represention*s, Forum Math., 4 (1992), pp. 299–322.

[42] ———, *I/O equations in discrete and continuous time*, in Proc. 31st IEEE Conf. Decision and Control, Tucson, Dec. 1992, IEEE Publications, Piscataway, NJ, 1992, pp. 3661–3662.

[43] J. C. WILLEMS, *From time series to linear systems, Part* I: *Finite dimensional time-invariant systems, Part* II: *Exact modelling, and Part* III: *Approximate modelling*, Automatica, 22, 22, 23 (1986, 1986, 1987), pp. 561–580, 675–694, 87–115.