SIAM J. CONTROL AND OPTIMIZATION
Vol. 28, No. 1, pp. 1-33, January 1990

001

INVITED EXPOSITORY ARTICLE

This paper is another in the continuing series of expository papers that were invited by the editors. These papers undergo the same refereeing procedure as do research papers submitted directly by the authors, although the refereeing guidelines are modified to suit the largely expository nature of the paper. Due to the rapid recent technical development of a number of areas in control and optimization, many of the seminal papers are quite specialized and are readily accessible to a limited group of experts only. Moreover, the original motivations and practical importance of the ideas are sometimes difficult to find in the mathematical development. The purpose of these papers is to bring the ideas, techniques, and applications of a few selected areas to the attention of a wider audience, so that their basic importance can be more easily and widely appreciated.

CONTROLLABILITY OF NONLINEAR DISCRETE-TIME SYSTEMS: A LIE-ALGEBRAIC APPROACH*

BRONISLAW JAKUBCZYK† AND EDUARDO D. SONTAG‡

Abstract. This paper presents a geometric study of controllability for discrete-time nonlinear systems. Various accessibility properties are characterized in terms of Lie algebras of vector fields. Some of the results obtained are parallel to analogous ones in continuous-time, but in many respects the theory is substantially different and many new phenomena appear.

Key words. controllability, Lie algebras of vector fields, nonlinear systems, discrete time

AMS(MOS) subject classifications. 93C10, 93C55, 93B05

1. Introduction. This paper deals with questions of controllability for discrete-time nonlinear systems

$$x(t+1)=f(x(t), u(t)) \tag{1}$$

for which the control variables u and state variables x take continuous values. Systems of the type (1) but with discrete-valued states and controls have long been studied in automata and sequential machine theory, but the continuous case has only recently become the subject of serious investigation as far as controllability properties are concerned. Our objective here is to survey a number of known results and to present new characterizations involving geometric ideas.

The study of controllability questions for the better known continuous-time analogue of (1), the differential equation

$$\dot{x}(t)=\phi(x(t), u(t)), \tag{2}$$

has been the subject of a concentrated research effort, as documented, for instance, in the survey papers [2] and [7], the text [8], and the exposition [35]. It is known, for instance, that the set *accessible* from any given state x^0, that is to say, the set of points reachable from x^0, contains a smooth submanifold of the state space and is in turn contained in a submanifold of the same dimension. Thus, for instance, the cusp in Fig. 1 cannot be an accessible set for any system of the type (2). More interestingly perhaps, this dimension can be computed from the rank of certain matrices formed

* Received by the editors January 11, 1988; accepted for publication (in revised form) March 14, 1989.
† Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-950 Warsaw, Poland. This work was done while the author was a Visiting Professor at Rutgers, The State University of New Jersey.
‡ Rutgers Center for Systems and Control, Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903. E-mail address: sontag@math.nutgens.edu. This research was supported in part by United States Air Force grant 85-0247 and National Science Foundation grant DMS-8803396.

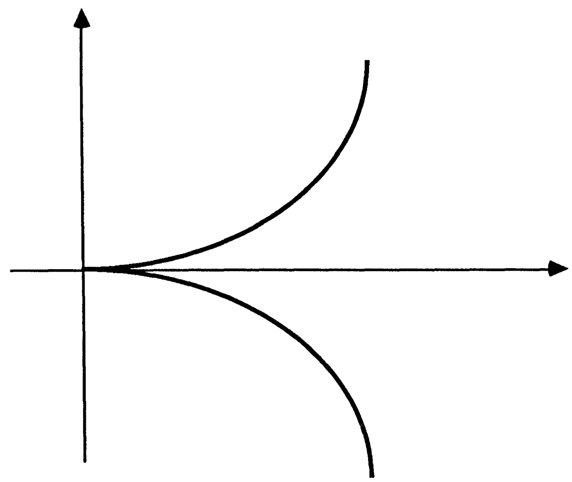

FIG. 1. *Impossible reachable set.*

by taking iterated Jacobians of the various vector fields $\phi(\cdot, u)$ evaluated at the state x^0. These Lie-theoretic characterizations are "direct" in that they do not involve integration of the differential equation, and they are closely related to more classical geometric material related to Frobenious' theorem.

(Certain technical hypotheses are of course required for the validity of the above and other assertions that we will make here; for purposes of providing an informal introduction we shall not make them precise yet; however, as a general rule, real-analyticity of f and ϕ and the assumption that states and controls take values in Euclidean space $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, are more than sufficient.)

Discrete control systems (1) are of interest for various reasons. Of course in many areas difference equation models are more natural than differential equations, but our interest has been motivated more by the problem of modeling physical systems under digital control via *sampling*. Recall that sampling is the process under which the state of a continuous time system is measured at discrete instants, and control actions are taken also at discrete instants. Figure 2 illustrates a typical approach to computer control. A discrete-time algorithm observes the state (or more generally, the outputs)

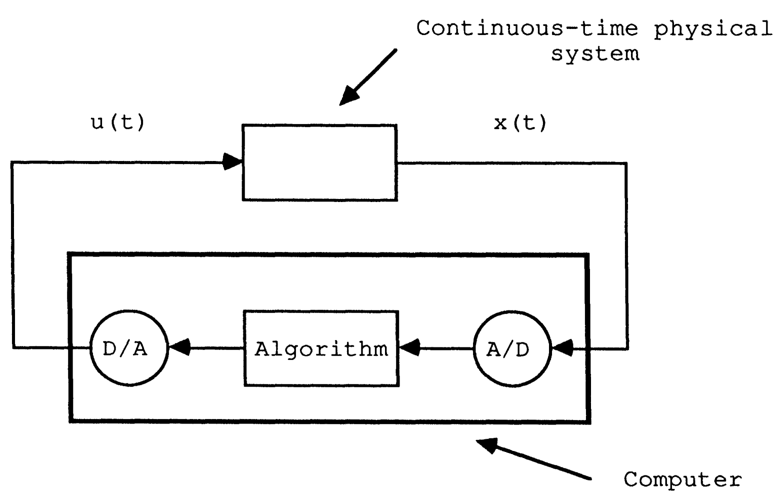

FIG. 2. *Digital control configuration.*

of a physical system, through an analogue-to-digital converter. Typically this observation is made at periodic time instants $\delta, 2\delta, \cdots$. On the basis of this observation the controller decides upon a control value u to be applied during the next period of length δ. This value is converted to analogue form and is held constant during that next period. So the controls applied to the physical system are restricted to be δ-sampled controls, constant on intervals $[k\delta, (k+1)\delta]$ (Fig. 3). The main point here is that, as far as the control algorithm is concerned, the physical system is a discrete-time system described by an equation of type (1), where $f(x, u)$ is the solution of the differential equation (2) at the end of an interval of length δ assuming that the initial state was x and control was held constantly equal to u.

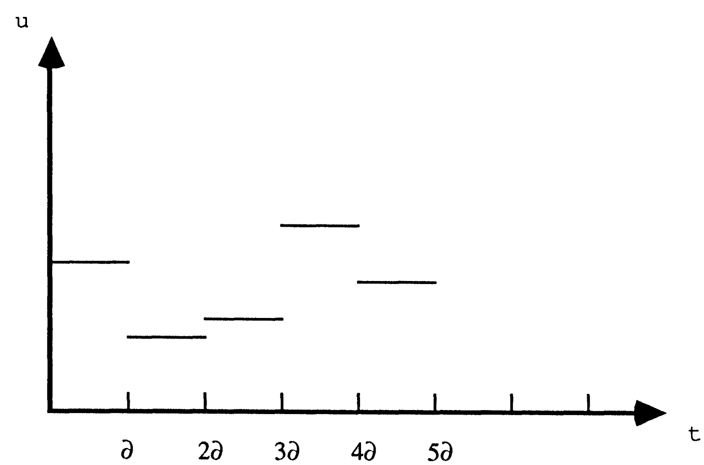

FIG. 3. *∂-sampled control.*

This description of sampling is oversimplified in many respects. For instance, analogue/digital conversion involves a quantization of the values of x into a discrete number of steps. Constant controls values may be smoothed out by a filter before being applied to the system. Multirate strategies, in which the sampling period is varied in a fixed set, may also be used. And the time involved in the algorithm actually computing the value of the control is sometimes nontrivial and must be included in the model as well. But even without these complications, the study of discrete-time control systems appears naturally.

Another area in which results from discrete-time nonlinear control theory are of importance is in the study of Markovian systems (1). There, the variables $u(t)$ are random, and together with the transitions f they characterize the probabilistic behavior of the process $x(\cdot)$. Accessibility conditions play a central role in establishing the existence and smoothness properties of equilibrium distributions; see for instance [15] and [16].

Yet another source of discrete-time control systems, related to but different from sampling, arises when numerically approximating the solution of a system (2). For instance, a Euler approximation with stepsize h gives the recursion

$$x(t+1) = x(t) + h\phi(x(t), u(t)).$$

These motivations notwithstanding, discrete-time systems have been studied much less than their continuous counterparts, and it has long been felt that their properties may diverge considerably from those of the latter. Regarding control and observation problems, the paper [26] and the monograph [27] considered various aspects of discrete-time systems defined by polynomial evolution equations. However, the general theory remained, until recently, much weaker than that possible in the more classical continuous time case, for which a large body of knowledge, as described above, is now available.

One of the main difficulties in the general discrete-time case is due to the possible noninvertibility of the one-step transition maps

$$x \mapsto f(x, u),$$

which means that semigroups tend to appear where groups would appear in the continuous case, so less algebraic structure is available. Accessible sets with singularities such as the curve in Fig. 1 can then easily appear.

An important observation, however, is that—due to the time-reversibility of finite-dimensional differential equations—for those discrete-time systems that arise through sampling these transition maps, obtained by integrating (2) over an interval of length δ with control $\equiv u$, *are* invertible. More precisely, each of these maps is a diffeomorphism (possibly not everywhere defined) of the state space. This is analogous to the situation in classical dynamical system theory, where one studies time-one diffeomorphisms and Poincaré maps associated to differential equations. Invertible discrete-time systems are often also obtained in numerical schemes for discretizing continuous-time models, if mesh sizes are chosen small enough.

In this paper we shall restrict our attention to *invertible* systems, for which the maps $f(\cdot, u)$ are assumed to be diffeomorphisms. For such systems we derive several characterizations of accessibility and we study the geometric structure of accessible sets. As an example, we provide a theorem that shows that, at least from equilibrium states, a picture such as that in Fig. 1 can never hold for these sets. (Precise statements of results are given later.) As with continuous-time systems, we also give Lie-theoretic characterizations of accessibility. These characterizations have the advantage that they do not require the computation of arbitrary iterates of the transition map, save for those iterates corresponding to just one value of the control value set.

The basic fact that underlies our approach is that one has an analogue for difference equations of the infinitesimal information obtained in the continuous-time case by taking derivatives with respect to time. One uses here derivations *with respect to control values.* This idea can be traced back to the paper [9], the first to deal in detail with general invertible discrete nonlinear control systems, although in the context of realization theory rather than controllability problems. For the latter, and for the source of the closest related material to that presented here, the credit goes to Fliess and Normand-Cyrot ([3], [25]), who originally proposed the definition in this manner of Lie algebras associated to discrete-time systems. This is analogous to associating a Lie algebra action to any given Lie group action. Other work along those lines was carried out in [11], [32], [17], [29], and related papers. A particularly important line of work is that pursued in [18], [20], [22], as well as by other authors (see, e.g., [5]), who have shown how to frame a large number of problems of control design (decoupling, noninteracting control, immersion, and so forth) in this geometric formalism; we shall not deal with such questions in this paper, however. For other recent references on geometric discrete-time control, see, for instance, the following papers as well as references given there: [1], [6], [10], [12], [14], [19], [24], [28].

We close this introduction with the precise statement of a simplified version of one of our main results to illustrate the nature of our contribution. Assume that the system (1) is analytic, in the sense that f is analytic, and invertible, meaning that each of the maps

$$f_u = f(\cdot, u): \mathbb{R}^n \to \mathbb{R}^n$$

is a global diffeomorphism of $\mathbb{R}^n$ for each control value u; for simplicity assume further that the control values are arbitrary real numbers, $u \in \mathbb{U} := \mathbb{R}$.

Denote by f_0^k the kth power of f_0 with respect to composition, and define the following vector fields depending on u:

$$X_u^+(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_u^{-1} \circ f_{u+v}(x),$$

$$X_u^-(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_u \circ f_{u+v}^{-1}(x),$$

and more generally for each integer k and for $\sigma = \pm$, $f_u^+ = f_u$, $f_u^- = f_u^{-1}$,

$$(\mathrm{Ad}_0^k X_u^\sigma)(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_0^{-k} \circ f_u^{-\sigma} \circ f_{u+v}^{\sigma} \circ f_0^k(x),$$

where $-\sigma = -, +$ if $\sigma = +, -$, respectively. These vector fields were introduced in [11], [17], [20], and [21].

In analogy with standard continuous time notions of accessibility, we call the system (1) *forward accessible* from the state $x^0 \in \mathbb{R}^n$ if its attainable set from x^0 has a nonempty interior. Similarly, we say that (1) is *backward accessible* from x^0 it its backward attainable set from x^0, the set of points controllable to x^0, has a nonempty interior. Finally, we say that the system is *forward-backward accessible* or *transitive* from x^0 if its orbit through this state (the smallest positive and negative-invariant set containing x^0) has a nonempty interior. The orbit turns out to be a submanifold, so forward-backward accessibility is equivalent to this orbit being an open subset of the state space.

By an equilibrium state x^0 we mean one that satisfies $f(x^0, 0) = 0$. Part (c) of the following theorem had already been stated in [11] (see also Theorem 7 in [20]) but parts (a) and (b) are totally new. The theorem is a specialization to analytic systems and equilibrium states of much more general results to be discussed later.

THEOREM 1. *The following statements hold for any analytic system* (1) *and equilibrium state* x^0:

(a) *System* (1) *is forward accessible from* x^0 *if and only if*

$$\dim \mathrm{Lie}\,\{\mathrm{Ad}_0^k\, X_u^+ | k \geqq 0,\, u \in \mathbb{U}\}(x^0) = n.$$

(b) *System* (1) *is backward accessible from* x^0 *if and only if*

$$\dim \mathrm{Lie}\,\{\mathrm{Ad}_0^k\, X_u^- | k \leqq 0,\, u \in \mathbb{U}\}(x^0) = n.$$

(c) *System* (1) *is forward-backward accessible from* x^0 *if and only if*

$$\dim \mathrm{Lie}\,\{\mathrm{Ad}_0^k\, X_u^\sigma | k \in \mathbb{Z},\, u \in \mathbb{U},\, \sigma = \pm\}(x^0) = n.$$

It is an easy corollary of this theorem that all three conditions (forward, backward, and forward-backward accessibility) coincide for analytic systems and equilibrium initial states. This gives a generalization of the well-known Chow Theorem in the continuous-time theory. More generally, the dimension of the corresponding (forward, etc.) accessible sets are given by the dimensions of the above subspaces, from which it follows that the (forward) accessible set is an open subset of a manifold (the orbit);

therefore, the cusp in Fig. 1 cannot be a forward accessible set. Later we give an example for which this cusp appears as the union of three orbits, corresponding to the origin and each of the two smooth branches.

Note that the conditions in Theorem 1 involve iterated compositions of transitions corresponding to only *one* control—arbitrarily taken as the zero control. The "naive" conditions that one can give based on the implicit function theorem for the above accessibility properties, reviewed below, would involve compositions of all transition mappings, as well as, for backward and forward-backward accessibility of their (possibly hard to compute) inverses. Moreover, in the particular case when the system has, for instance, the form

$$x(t+1)=x(t)+g(x(t), u(t))$$

with $g(x, 0)\equiv 0$, the "Ad's" become all the identity and no compositions at all need be computed.

In this paper, we present an exposition, including complete proofs, of the known transitivity (positive and negative-time accessibility) facts, as well as of new results for the substantially different (positive-time) forward accessibility problem. We also clarify the relationship between a large number of forward and/or backward controllability notions. Another topic studied is the role played by various continuous time systems derived mathematically from the original discrete time model, and we show how to view the more classical results for continuous-time systems as a particular case (essentially when "time" is thought of as a control) of our theory. Finally, we provide an application of our accessibility characterizations to the sampled control of continuous systems; the resulting explicit eigenvalue condition, which generalizes the classical (linear system) sampling theorem, illustrates the power of the techniques developed. An illustrative example is included towards the end of the paper, which ends with a brief description of the alternative approach due to Normand-Cyrot.

2. Basic definitions. We start by introducing basic notation and definitions. As stated previously, time takes integer values, $t\in\mathbb{Z}$. We introduce the following notations for the effect of shift operators:

$$x^+(t)=x(t+1) \quad \text{and} \quad x^-(t)=x(t-1).$$

In this way we can write equation (1) in the more compact form, with $f^+=f$

$$x^+=f^+(x, u), \qquad x(t)\in\mathbb{X}, u(t)\in\mathbb{U}.$$

The state set $\mathbb{X}$ is a connected differentiable manifold of dimension n. To simplify the notation we first assume that the control is scalar, meaning that $\mathbb{U}$ is a subset of $\mathbb{R}$ contained in the closure of its interior,

$$\mathbb{U}\subset \text{clos int}\,\mathbb{U},$$

such that $0\in\mathbb{U}$. Later we show how to generalize everything to the case where $\mathbb{U}$ is a subset of a more general manifold.

The system is *of class* C^k if the manifold $\mathbb{X}$ is of class C^k, Hausdorff, second countable, and the function $f:\mathbb{X}\times\mathbb{U}\to\mathbb{X}$ is of class C^k, meaning, to be precise, that there exists a C^k extension of f to an open neighborhood of $\mathbb{X}\times\mathbb{U}$ in $\mathbb{X}\times\mathbb{R}$. When $k=\infty$ we say simply *smooth*; for $k=\omega$, *analytic.*

Associated to each such system there is a family of maps

$$f_u=f(\cdot, u): \mathbb{X}\to\mathbb{X}, \qquad u\in\mathbb{U}.$$

DEFINITION 2.1. The system (1) is *invertible* if for each u in an open neighborhood of $\mathbb{U}$ the map f_u is a global diffeomorphism of $\mathbb{X}$.

Invertibility can be weakened in various ways. For instance, many results can be obtained under the assumption of *local invertibility at* x, meaning that for each $u \in \mathbb{U}$ f_u is a local diffeomorphism at x, i.e., rank $(\partial f_u/\partial x)(x) = n$, or the assumption that this holds for every state, *local invertibility* of the system. The paper [10] shows how a condition called *submersibility* is in fact enough to define many of the concepts that we use in this paper.

To any invertible system one can associate an *inverse* or *reversed-time* system with equations

$$x^- = f^-(x, u), \tag{3}$$

where $f^-(x, u) = f_u^{-1}(x)$. By the implicit mapping theorem, this is again of class C^k, and its inverse is the original system.

Unless otherwise stated, every system appearing in this paper will be assumed to be invertible. Furthermore, until § 6, *controls are scalar.*

The maps f_u and their inverses f_u^{-1} can be considered as "one step forward maps" (respectively, "one step backward maps"). If we apply a sequence of controls $u_1, \cdots, u_k$ then we obtain the composition of these maps denoted by

$$f_{u_k \cdots u_1} = f_{u_k} \circ \cdots \circ f_{u_1}. \tag{4}$$

Allowing backward as well as forward steps we obtain a larger family of maps

$$f^{\varepsilon_k \cdots \varepsilon_1}_{u_k \cdots u_1} = f^{\varepsilon_k}_{u_k} \circ \cdots \circ f^{\varepsilon_1}_{u_1}, \tag{5}$$

where each of $\varepsilon_1, \cdots, \varepsilon_k$ takes a value ± 1.

We shall denote by $A_k^+(x)$ the set of points attainable from x in k forward steps, and by $A^+(x)$ the set of points attainable from x in any nonnegative number of forward steps. Replacing forward steps by backward steps we obtain other sets, $A_k^-(x)$ and $A^-(x)$, which consist of points controllable to x in k steps, and controllable to x in any nonnegative number of steps, respectively. Finally, the set of points attainable from x in any number of positive and negative steps is called the *orbit* of x and is denoted by $A(x)$.

DEFINITION 2.2. The system (1) is *forward* (*backward*) *accessible from* x if its attainable set $A^+(x)$ (respectively, $A^-(x)$) has a nonempty interior. It is called *transitive from* x (*or forward-backward accessible from* x) if its orbit $A(x)$ has a nonempty interior (and so it is necessarily open).

Finally, the system is *forward* (*backward*) *accessible* if it is forward (backward) accessible from any $x \in \mathbb{X}$, and it is called *transitive* if it is transitive from any $x \in \mathbb{X}$.

Observe that there is a straightforward criterion for accessibility of the discrete time system, based on the rank of the following map. For each fixed state x and integer k define

$$\psi_{k,x}(\mathbf{u}) := f_{u_k \cdots u_1}(x),$$

where $\mathbf{u} = (u_1, \cdots, u_k)$ takes values in the kth Cartesian product $\mathbb{U}^k$. Notice that the attainable set $A_k^+(x)$ is by definition equal to the image of this map. The following proposition says that this set is of nonempty interior if and only if the linearization along some trajectory starting from x is controllable.

PROPOSITION 2.3. *Let* (1) *be smooth. For any fixed x and k, the interior of the attainable set* $A_k^+(x)$ *is nonempty if and only if*

$$\sup\left\{\operatorname{rank}\frac{\partial}{\partial \mathbf{u}}\,\psi_{k,x}(u),\,\mathbf{u}\in\mathbb{U}^k\right\}=n$$

and thus

$$\sup\left\{\operatorname{rank}\frac{\partial}{\partial \mathbf{u}}\,\psi_{k,x}(u),\,\mathbf{u}\in\mathbb{U}^k,\,k\geqq 1\right\}=n$$

is necessary and sufficient for forward accessibility of system (1) *from x.*

Proof. If there is a point $\mathbf{u}$ at which the rank of the map $\psi_{k,x}$ is equal to n, we may assume without loss of generality that $\mathbf{u}$ is in the interior of $\mathbb{U}$, because of the hypothesis that $\mathbb{U}\subset \operatorname{clos}\operatorname{int}\mathbb{U}$. It then follows from the implicit function theorem that the image of this map has a nonempty interior. Thus, the attainable set $A_k^+(x)$ has a nonempty interior. (Only that the system is of class C^1 is used for this implication.)

Conversely, if the rank of the map $\psi_{k,x}$ is less than n at each $u\in\mathbb{U}$, then every element of $A_k^+(x)$ is a critical value of $\psi_{k,x}$ as a map defined on an open subset of $\mathbb{R}^k$. It follows by Sard's theorem that the image of $\mathbb{U}$ under this map is of empty interior and is of measure zero under the measure induced by any Riemann metric on $\mathbb{X}$ (the Euclidean metric in $\mathbb{R}^n$). Therefore, the attainable set $A_k^+(x)$ must have an empty interior and it is even of measure zero.

The second statement follows from the first because a countable union of sets of measure zero again has measure zero. □

REMARK 2.4. Since the orbit $A(x)$ is the (countable) union of the images of the maps (5) we can use an analogous argument to give a criterion for transitivity from x, using the maps (5) rather than (4) to define a family of maps playing the role of the $\psi_{k,x}$'s.

The above proposition and remark might appear to give satisfactory criteria for forward accessibility and transitivity. Unfortunately, this is not the case. Although for simple systems they may be used to decide whether a given system is forward accessible or not, for more complicated sytems explicitly computing the functions $\psi_{k,x}$ may be highly nontrivial, since composition is hard to deal with computationally. As an example, consider for instance the problem of obtaining a general formula for the nth composition of the quadratic function $g(x)=ax^2+bx+c$ with itself or that of computing the function $\psi_{k,x}$ if $f(x,u)=g(x)+xu$. The problem becomes even more serious in the case of deciding the transitivity of the system, as this requires also finding the inverse maps f_u^{-1} needed for computing the composed maps (5). One approach here is to develop a calculus for these compositions, as in the work of Monaco and Normand-Cyrot; see the last section. But in any case, even for classes such as that of bilinear systems, Proposition 2.3 doesn't seem to provide much useful information regarding accessibility properties.

Also, from a purely theoretical point of view, Proposition 2.3 is of little interest. This is because it gives too limited an insight into the geometry of our systems and it provides an even more limited tool for their study. The maps appearing in the criteria do not have much algebraic and geometric structure.

The main aim of the next section is to introduce a sort of "infinitesimal description" of the discrete-time system. This is done by introducing certain vector fields associated to it. By doing so we immediately get a powerful tool and a rich algebraic and geometric structure based on the Lie product of vector fields. In particular, the accessibility properties of the system can be studied using natural Lie algebras of vector fields associated to the system. The idea of introducing vector fields corresponding to infinitesimal perturbations of control values is a natural generalization of the concept of actions of Lie groups, and it was originally proposed in the context of nonlinear control in [3]. These vector fields also find natural applications in the study of controllability properties and the feedback linearizability of sampled systems ([29], [12]).

3. Vector fields associated to the system. We associate the following four families of vector fields to our discrete time system (1), one vector field for each $u \in \mathbb{U}$:

$$X_u^+(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_u^{-1} \circ f_{u+v}(x),$$

$$X_u^-(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_u \circ f_{u+v}^{-1}(x),$$

$$Y_u^+(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_{u+v}^{-1} \circ f_u(x),$$

$$Y_u^-(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_{u+v} \circ f_u^{-1}(x).$$

The partial derivatives here are well defined in the interior of $\mathbb{U}$; therefore, they are also uniquely defined on the boundary of $\mathbb{U}$ because of continuity. The geometric

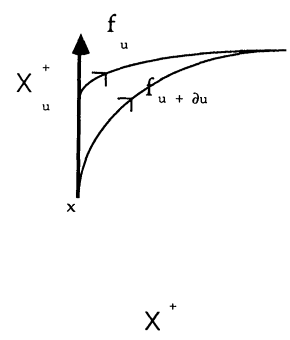

FIG. 4 (a)

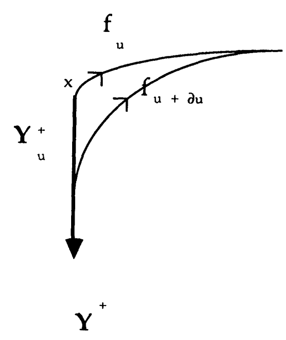

FIG. 4 (c)

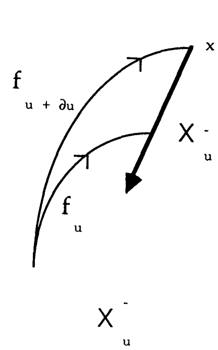

FIG. 4 (b)

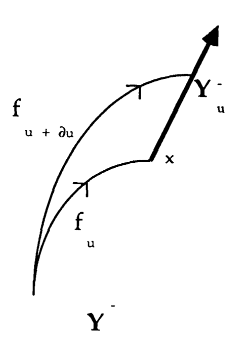

FIG. 4 (d)

meaning of these vector fields is illustrated by Fig. 4, and the interrelations between them are explained in the next proposition. These vector fields were also introduced in [17], [20], and [21], using somewhat different terminology. The last section will explain the relation between the different notations.

The special case in which the function f happens to correspond to the flow of a vector field Z, that is, $f(x, u) = \exp(uZ)$, will be important later when discussing continuous time systems within our framework. In that case all of the above vector fields are in fact independent of u, and they provide the same information about the system. This is because by the semigroup property of flows it holds that $f_{u+v} = f_u \circ f_v = f_v \circ f_u$, so that $X_u^+ = -X_u^- = Z = -Y_u^+ = Y_u^-$. These equalities help us to understand why the continuous time theory is considerably simpler than the discrete one.

Note that applying these definitions to the inverse system (3) instead of system (1) gives the same vector fields except that the pluses are changed for minuses and vice versa.

Given a vector field Y and a control value u, we can define another vector field from Y by applying a change of coordinates given by the diffeomorphism f_u,

$$(\mathrm{Ad}_u\, Y)(x) = (\mathrm{d}f_u(x))^{-1} Y(f_u(x)).$$

Here $\mathrm{d}f_u$ stands for the differential of f_u with respect to x. Using the diffeomorphisms (4), we may also define

$$(\mathrm{Ad}_{u_k \cdots u_1}\, Y)(x) = (\mathrm{d}f_{u_k \cdots u_1}(x))^{-1} Y(f_{u_k \cdots u_1}(x)),$$

and, applying the even more general family of diffeomorphisms (5),

$$(\mathrm{Ad}_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1}\, Y)(x) = (\mathrm{d}f_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1}(x))^{-1} Y(f_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1}(x)). \tag{6}$$

Clearly, the operators "Ad" so defined are linear operators acting on vector fields Y, and we have that

$$\mathrm{Ad}_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1}\, Y = \mathrm{Ad}_{u_1}^{\varepsilon_1} \cdots \mathrm{Ad}_{u_k}^{\varepsilon_k}\, Y. \tag{7}$$

(Note the reversal of indices.) We will use the abbreviated notation $\mathrm{Ad}_0^k\, Y$ for $\mathrm{Ad}_{0 \cdots 0}\, Y$ with $u = 0$ repeated k-times, if $k > 0$, and for $\mathrm{Ad}_{0 \cdots 0}^{-1 \cdots -1}\, Y$, if $k < 0$. Additionally, $\mathrm{Ad}_0^0\, Y = Y$. With this notation we have that

$$(\mathrm{Ad}_0^k\, X_u^+)(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_0^{-k} \circ f_u^{-1} \circ f_{u+v} \circ f_0^k(x)$$

(see Fig. 5) and, more generally,

$$(\mathrm{Ad}_{u_k \cdots u_1} X_{u_0}^+)(x) = \frac{\partial}{\partial v}\bigg|_{v=0} f_{u_k \cdots u_1}^{-1} \circ f_{u_0}^{-1} \circ f_{u_0+v} \circ f_{u_k \cdots u_1}(x).$$

Since our system is assumed to be invertible, we could apply all definitions to the inverse system (3) instead of (1). Then all the pluses in the superscripts change for minuses and Ad_u changes for Ad_u^{-1}, and vice versa. Therefore, we will have the following fact, which we shall use repeatedly.

REVERSION PRINCIPLE. *Any general property of systems of the type* (1) *that can be expressed in terms of the above defined vector fields is preserved if we change the pluses in the superscripts for the minuses and each* Ad_u *for* Ad_u^{-1}, *and vice versa.*

Remark 3.1. Some of the above defined vector fields can be equivalently defined as follows:

$$X_u^+(x) = (\mathrm{d}f_u(x))^{-1} \frac{\partial}{\partial u} f_u(x),$$

$$(\mathrm{Ad}_{u_k \cdots u_1}\, X_{u_0}^+)(x) = (\mathrm{d}f_{u_k \cdots u_1}(x))^{-1} X_{u_0}^+(f_{u_k \cdots u_1}(x)).$$

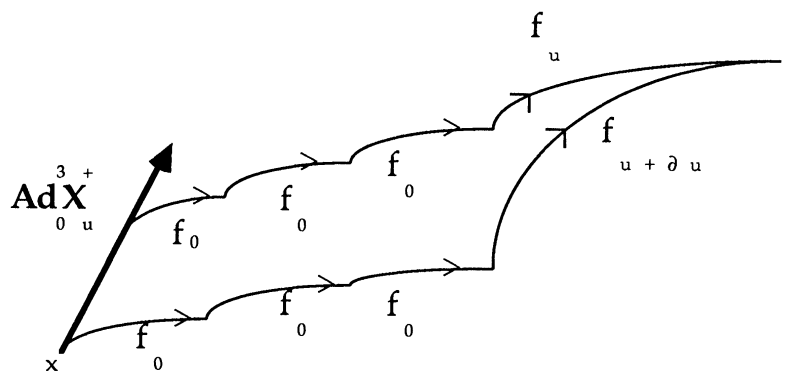

FIG. 5. $\mathrm{Ad}_0^3 X_u^+$.

Since the inverses f_u^{-1} do not appear, the right-hand sides now make sense for *locally* invertible systems. Those of our results that can be stated exclusively in terms of the above vector fields will also hold for locally invertible systems. Furthermore, criteria stated in their terms can be checked without computing the inverse of any diffeomorphism; only matrix inversions are required. For instance, take the system with $\mathbb{X} = \mathbb{R}$, $\mathbb{U} = [-1, 1]$, and equations

$$x^+ = x^3 + 2x + u \sin x.$$

Since for each fixed value of u the right-hand side is strictly increasing, this is an invertible system. We obtain here that

$$X_u^+(x) = \frac{\sin x}{3x^2 + 2 + u \cos x}$$

in the natural coordinates.

The basic interrelations between the vector fields X_u^+, X_u^-, Y_u^+, Y_u^- are given by the following proposition.

PROPOSITION 3.2. *The following equalities hold for each* $u \in \mathbb{U}$.

(*a*) $X_u^+ = -Y_u^+$, $X_u^- = -Y_u^-$.

(b) $X_u^+ = -\mathrm{Ad}_u X_u^-$, $Y_u^+ = -\mathrm{Ad}_u Y_u^-$.

Proof. To prove (a), we differentiate with respect to u the equality

$$f_u^{-1} \circ f_u(x) = x$$

and we get

$$Y_u^+(x) + X_u^+(x) = 0.$$

The second equality in (a) follows from the first by the reversion principle.

On the other hand, differentiating with respect to v the equality

$$f_u^{-1} \circ f_{u+v}(x) = f_u^{-1} \circ f_{u+v} \circ f_u^{-1} \circ f_u(x)$$

we get $X_u^+ = \mathrm{Ad}_u Y_u^-$, which together with (a) gives (b). The proof of the last equality now follows by the reversion principle. □

Later in the paper it will be very useful to have a formula for the derivative with respect to u of a vector field Y transformed by the diffeomorphism f_u. It was noted in [25], [11], [17], [18] that this derivative can be easily expressed via the above introduced vector fields and the Lie bracket; in fact, the next two propositions appear as the first steps in the proof of Theorem 3 on page 26 of [17] and of Lemma 3 in [18].

Here and further we shall use the standard notation $[Y, Z]$ for the Lie bracket of the vector fields Y and Z which, in $\mathbb{R}^n$, is given by $[Y, Z] = \partial Z/\partial x\ Y - \partial Y/\partial x Z$. We also denote $\operatorname{ad} Z(Y) = [Z, Y]$ and the kth iteration of the operator $\operatorname{ad} Z$, $\operatorname{ad}^k Z(Y) = \operatorname{ad} Z \cdots \operatorname{ad} Z(Y)$. The flow of the vector field Y is denoted by $\exp(tY)$.

PROPOSITION 3.3. *The following equalities hold for any vector field Z and any $u \in \mathbb{U}$:*

$$\frac{\partial}{\partial u} \operatorname{Ad}_u Z = \operatorname{ad} X_u^+(\operatorname{Ad}_u Z)$$

and

$$\frac{\partial}{\partial u} \operatorname{Ad}_u^{-1} Z = \operatorname{ad} X_u^-(\operatorname{Ad}_u^{-1} Z).$$

Proof. It is enough to prove each of the equalities locally, so we shall assume that we are in $\mathbb{R}^n$. We have that

$$\begin{aligned}
\frac{\partial}{\partial u} \operatorname{Ad}_u Z &= \frac{\partial}{\partial t}\bigg|_{t=0} \frac{\partial}{\partial u} f_u^{-1} \circ \exp(tZ) \circ f_u(x) \\
&= \frac{\partial}{\partial t}\bigg|_{t=0} \left(\frac{\partial}{\partial u} f_u^{-1}\right) \circ f_u \circ f_u^{-1} \circ \exp(tZ) \circ f_u(x) \\
&\quad + \frac{\partial}{\partial t}\bigg|_{t=0} \mathrm{d}(f_u^{-1} \circ \exp(tZ) \circ f_u)(x)(\mathrm{d}f_u(x))^{-1} \frac{\partial}{\partial u} f_u(x) \\
&= (\partial Y_u^+/\partial x)(x) \operatorname{Ad}_u Z(x) + (\partial \operatorname{Ad}_u Z/\partial x)(x) X_u^+(x) \\
&= [X_u^+, \operatorname{Ad}_u Z](x),
\end{aligned}$$

where we use the equality $X^+ = -Y^+$.

The second equality follows from the first by the reversion principle, replacing f_u by f_u^{-1}. □

In the next proposition and in the rest of the paper we shall use the following notational convention. Given a family of vector fields $\{Y_\alpha | \alpha \in A\}$, we denote by $\operatorname{Lie}\{Y_\alpha | \alpha \in A\}$ the Lie algebra generated by this family of vector fields and by $\operatorname{Lie}\{Y_\alpha | \alpha \in A\}(x)$ the subspace of the tangent space at x generated by the vector fields in this Lie algebra.

PROPOSITION 3.4. *For analytic systems and connected $\mathbb{U}$,*

$$\operatorname{Ad}_0^k X_u^+(x) \in \operatorname{Lie}\{\operatorname{Ad}_0^{k+1} X_u^- | u \in \mathbb{U}\}(x)$$

and

$$\operatorname{Ad}_0^{-k} X_u^-(x) \in \operatorname{Lie}\{\operatorname{Ad}_0^{-k-1} X_u^+ | u \in \mathbb{U}\}(x)$$

for each $x \in \mathbb{X}$, each $u \in \mathbb{U}$, and each integer k.

In the proof of this proposition we shall use the following lemma. This lemma is in fact about identities on free Lie algebras; we give a somewhat informal statement to avoid having to introduce considerably more machinery.

LEMMA 3.5. *For any $r \geqq 0$ there are coefficients $a_1, \cdots, a_p \in \mathbb{R}$ and $b_1, \cdots, b_q \in \mathbb{R}$ independent of x and u such that*

$$\mathrm{Ad}_u \left(\frac{\partial^r}{\partial u^r} X_u^- \right) = \sum a_i Y_i \text{ where } Y_i \in \mathrm{Lie} \left\{ X_u^+, \frac{\partial}{\partial u} X_u^+, \cdots, \frac{\partial^r}{\partial u^r} X_u^+ \right\} = M_u^{r,+},$$

$$\mathrm{Ad}_u^{-1} \left(\frac{\partial^r}{\partial u^r} X_u^+ \right) = \sum b_i Z_i \textit{ where } Z_i \in \mathrm{Lie} \left\{ X_u^-, \frac{\partial}{\partial u} X_u^- \cdots, \frac{\partial^r}{\partial u^r} X_u^- \right\} = M_u^{r,-}.$$

Moreover, these coefficients, as well as the expressions of each Z_i and Y_i in terms of the generators of the corresponding Lie algebra of vector fields, are independent of the particular system.

Proof. From Proposition 3.2 it follows that the assertions are true for $r = 0$. Assume that the first of them is true for $r = k$. From Proposition 3.3 it follows that

$$\frac{\partial}{\partial u} \mathrm{Ad}_u \frac{\partial^k}{\partial u^k} X_u^- = \mathrm{ad}\, X_u^+ \,\mathrm{Ad}_u \frac{\partial^k}{\partial u^k} X_u^- + \mathrm{Ad}_u \frac{\partial^{k+1}}{\partial u^{k+1}} X_u^-. \tag{8}$$

In general for parametrized vector fields A_u, B_u we have that

$$\frac{\partial}{\partial u} [A_u, B_u] = \left[\frac{\partial}{\partial u} A_u, B_u \right] + \left[A_u, \frac{\partial}{\partial u} B_u \right].$$

Thus it follows from the induction assumption that the left side term in (8) is a linear combination of elements in $M^{k+1,+}$ and so is the first term on the right. Therefore, the second element on the right is a linear combination of elements in $M_u^{k+1,+}$ and the assertion is true for $r = k+1$.

The second part of the proposition follows from the first and the reversion principle. □

Proof of Proposition 3.4. In the proof we shall use the following corollary to the Taylor formula for an analytic, vector valued function g defined on a connected set $\mathbb{U}$ containing the origin: $\mathrm{span}\,\{g(u) | u \in \mathbb{U}\} = \mathrm{span}\,\{g^{(i)}(0) | i \geqq 0\}$. We have

$$\begin{aligned} \mathrm{span}\,\{\mathrm{Ad}_0^k X_u^+, u \in \mathbb{U}\}(x) &= \mathrm{Ad}_0^k \,\mathrm{span} \left\{ \frac{\partial^r}{\partial u^r} \bigg|_{u=0} X_u^+, r \geqq 0 \right\}(x) \\ &\subset \mathrm{Ad}_0^k \left(\mathrm{Ad}_0 \,\mathrm{Lie} \left\{ \frac{\partial^r}{\partial u^r} X_u^- \bigg|_{u=0}, r \geqq 0 \right\} \right)(x) \\ &= \mathrm{Lie}\,\{\mathrm{Ad}_0^{k+1} X_u^- | u \in \mathbb{U}\}(x). \end{aligned}$$

Here the inclusion follows from Lemma 3.5 (apply Ad_u to both sides of the second equation and then evaluate at $u = 0$); the first and the third equality follow from Taylor's formula.

The second assertion of the proposition is a consequence of the first and the reversion principle. □

Note that it is not claimed in Proposition 3.4 that, for instance, X_u^+ is in the Lie algebra generated by the vector fields $\mathrm{Ad}_0 X_u^-$. The statement pertains only to the equality of the associated distributions, that is, of the tangent spaces at each point.

4. Acessibility criteria. To state our criteria we shall need the following families of vector fields:

$$\Gamma^+ = \{\mathrm{Ad}_{u_k \cdots u_1} X_{u_0}^+ | k \geqq 0, u_0, \cdots, u_k \in \mathbb{U}\},$$

$$\Gamma^- = \{\mathrm{Ad}_{u_k \cdots u_1}^{-1 \cdots -1} X_{u_0}^- | k \geqq 0, u_0, \cdots, u_k \in \mathbb{U}\},$$

$$\Gamma = \{\mathrm{Ad}_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1} X_{u_0}^{\sigma} | k \geqq 0, u_0, \cdots, u_k \in \mathbb{U}, \varepsilon_1, \cdots, \varepsilon_k = \pm 1, \sigma = \pm\}.$$

As previously, for a family of vector fields Δ, we denote by Lie $\{\Delta\}$ the Lie algebra of vector fields generated by Δ, by $\Delta(x)$ the linear space spanned by the vectors at x given by the vector fields in Δ, and by Lie $\{\Delta\}(x)$ the linear space of tangent vectors at x given by the vector fields in the Lie algebra.

The following theorem gives criteria for accessibility of smooth systems. It will be one of the main results of this paper.

THEOREM 2. *The following properties hold for any smooth system* (1).

(a) *The system is forward accessible if and only if any of the following two equivalent conditions hold*:

$$\dim \Gamma^+(x) = n \quad \forall x \in \mathbb{X}, \quad \textit{or} \quad \dim \text{Lie}\,\{\Gamma^+\}(x) = n \quad \forall x \in \mathbb{X}.$$

(b) *The system is backward accessible if and only if any of the following two equivalent conditions hold*:

$$\dim \Gamma^-(x) = n \quad \forall x \in \mathbb{X}, \quad \textit{or} \quad \dim \text{Lie}\,\{\Gamma^-\}(x) = n \quad \forall x \in \mathbb{X}.$$

(c) *The system is transitive if and only if any of the following two equivalent conditions hold*:

$$\dim \Gamma(x) = n \quad \forall x \in \mathbb{X}, \quad \textit{or} \quad \dim \text{Lie}\,\{\Gamma\}(x) = n \quad \forall x \in \mathbb{X}.$$

To state a stronger version of our result, valid for analytic systems, we need the following Lie algebras of vector fields:

$$L^+ = \text{Lie}\,\{\text{Ad}_0^k\, X_u^+ | k \geqq 0,\, u \in \mathbb{U}\},$$

$$L^- = \text{Lie}\,\{\text{Ad}_0^k\, X_u^- | k \leqq 0,\, u \in \mathbb{U}\},$$

$$L = \text{Lie}\,\{\text{Ad}_0^k\, X_u^\sigma | k \in \mathbb{Z},\, u \in \mathbb{U},\, \sigma \in \{+, -\}\}.$$

The following inclusions are evident:

$$L^+ \subset \text{Lie}\,\Gamma^+, \quad L^- \subset \text{Lie}\,\Gamma^-, \quad L \subset \text{Lie}\,\Gamma.$$

In terms of this data, we now state another one of our main results. As remarked earlier, the transitivity case had been stated before ([11], [20]). Even for that case, however, we believe that this paper contains the first complete proof.

THEOREM 3. *The following properties hold for any analytic system* (1) *with connected* $\mathbb{U}$:

(a) *The system is forward accessible if and only if*

$$\dim L^+(x) = n \quad \textit{for any } x \in \mathbb{X}.$$

(b) *The system is backward accessible if and only if*

$$\dim L^-(x) = n \quad \textit{for any } x \in \mathbb{X}.$$

(c) *The system is transitive if and only if*

$$\dim L(x) = n \quad \textit{for any } x \in \mathbb{X}.$$

Remark 4.1. As a consequence of Proposition 3.4, if we were to take in the definition of the Lie algebra L only $\sigma = +$, or alternatively, only $\sigma = -$, a smaller set of vector fields may result, but the conclusions in the theorem would hold equally well.

There is a pointwise version of the above results. An *equilibrium point* $x_0 \in \mathbb{X}$ is one such that $f(x_0, 0) = 0$.

THEOREM 4. *The following properties hold, if* $\mathbb{U}$ *is connected*:

(a) *A smooth system* (1) *is transitive from x if and only if* $\dim \Gamma(x) = n$ (*equivalently*, $\dim \text{Lie}\,\{\Gamma\}(x) = n$). *An analytic system* (1) *is transitive from x if and only if* $\dim L(x) = n$.

(b) *An analytic system* (1) *is forward* (*respectively, backward*) *accessible from an equilibrium point* x_0 *if and only if* $\dim L^+(x_0)=n$ (*respectively*, $\dim L^-(x_0)=n$).

The proofs of all these results are given later after we develop some further theory.

The second part of Theorem 4 will be strengthened as a consequence of the following proposition.

PROPOSITION 4.2. *If the system is analytic,* $\mathbb{U}$ *is connected, and* x_0 *is an equilibrium point, then*

$$L^+(x_0)=L^-(x_0)=L(x_0).$$

Proof. Since $L^+(x_0)\subset L(x_0)$, it is enough to show that $L^+(x_0)$ has the same dimension as $L(x_0)$ to conclude that they are equal. Pick a basis of the latter and assume that the elements in the basis involve vector fields of the form $\mathrm{Ad}_0^k X_u^+$, with the possible k bounded below by the integer k^*. (Recall Remark 4.1 to the effect that we may always assume that $\sigma=+$ in the definition of L.) Applying the operator

$$\mathrm{Ad}_0^{-k^*}$$

to these vector fields, we obtain vector fields in L^+. As x_0 is an equilibrium point, the operator $\mathrm{Ad}_0^{-k^*}$ preserves the tangent space at x_0 and we obtain a set of linearly independent vectors in $L^+(x_0)$, as desired. The argument for L^- follows by the reversion principle. □

The above theorem and proposition immediately imply the following corollary.

COROLLARY 4.3. *Assume that the system is analytic,* $\mathbb{U}$ *is connected, and* x_0 *is an equilibrium point. Then forward accessibility from* x_0, *backward accessibility from* x_0, *and transitivity from* x_0 *are all equivalent properties.*

We will prove the above theorems by splitting them into (somewhat stronger) sufficiency and necessity results.

Define the following families of vector fields:

$$X_u^{+,i}=\frac{\partial^i}{\partial u^i}X_u^+,\qquad X_u^{-,i}=\frac{\partial^i}{\partial u^i}X_u^-.$$

THEOREM 5. *The following statements hold for any smooth system* (1).

(a) *If*

$$\dim \mathrm{Lie}\,\{\Gamma^+\}(x)=n \quad \textit{for all } x\in\mathbb{X}, \tag{9}$$

then the system is forward accessible.

(b) *If* x_0 *is an equilibrium point and if*

$$\dim \mathrm{Lie}\,\{\mathrm{Ad}_0^k X_0^{+,i} \mid k\geqq 0,\ i\geqq 0\}(x_0)=n, \tag{10}$$

then the system is forward accessible from x_0.

(c) *The same statements hold for backward accessibility if we replace* Γ^+ *for* Γ^- *and* $X_u^{+,i}$ *for* $X_u^{-,i}$.

Proof. (a) Let us fix an $x\in\mathbb{X}$. Let p and $v_1^*,\cdots,v_p^*$ be such that the rank of the Jacobian of the map

$$(v_1,\cdots,v_p)\mapsto f_{v_p\cdots v_1}(x) \tag{11}$$

is maximal (over all $p\geqq 0$ and $v_1,\cdots,v_p\in\mathbb{U}$) at $v_1^*,\cdots,v_p^*$. Because $\mathbb{U}\subset \operatorname{clos}\operatorname{int}\mathbb{U}$, we may assume that these are in the interior of $\mathbb{U}$. Let W be a neighborhood of $(v_1^*,\cdots,v_p^*)$ on which this rank is maximal and such that the image S of W under the above map is a submanifold. Since $S\subseteq A^+(x)$, it is enough to show that the dimension of S is equal to n, from which it will follow that S is an open subset of $\mathbb{X}$.

We now prove that each vector field of the type $\mathrm{Ad}_{u_k\cdots u_1} X^+_{u_0}$ is tangent to S. It will follow then that all the Lie brackets of these vector fields are tangent to the submanifold S. This, together with assumption (9), will imply that S is of dimension n.

Assume that the vector

$$\mu := (\mathrm{Ad}_{u_k\cdots u_1} X^+_{u_{k+1}})(y)$$

is not tangent to S at $y = f_{v_p\cdots v_1}(x)$, for some $u_1, \cdots, u_{k+1}$ (for convenience we denote u_0 by u_{k+1} now) and some $(v_1, \cdots, v_p) \in W$. Again, we may assume that these are all in the interior of $\mathbb{U}$. Thus

$$\mu = \frac{\partial}{\partial v}\bigg|_{v=0} (f_{u_k\cdots u_1})^{-1} \circ f^{-1}_{u_{k+1}} \circ f_{u_{k+1}+v} \circ f_{u_k\cdots u_1}(y)$$

is not tangent to S and therefore also

$$\frac{\partial}{\partial v}\bigg|_{v=0} f_{u_{k+1}+v} \circ f_{u_k\cdots u_1} \circ f_{v_p\cdots v_1} = (\mathrm{d}f_{u_{k+1}\cdots u_1})(y)\mu$$

is not tangent to the submanifold $f_{u_{k+1}\cdots u_1}(S)$. But this means that the rank of the Jacobian map of the mapping

$$(v_1, \cdots, v_p, u_1, \cdots, u_{k+1}) \to f_{u_{k+1}\cdots u_1 v_p \cdots v_1}(x)$$

is at least $\dim S + 1$ for this sequence $v_1, \cdots, v_p, u_1, \cdots, u_{k+1}$, contradicting maximality of the rank. It follows that the vector field $\mathrm{Ad}_{u_k\cdots u_1} X^+_{u_{k+1}}$ must indeed be tangent to S.

(b) The idea of this part of the proof is the same as in part (a) except that now the rank assumption is made at one point only. Thus, we have to construct the manifold S in a neighborhood of x_0 so that n linearly independent vector fields in the Lie algebra (10) are linearly independent in this neighborhood and tangent to this manifold.

Let V be a coordinate neighborhood of x_0 such that there are n vector fields in the Lie algebra (10) which are linearly independent on V. Suppose that these vector fields involve only $k \leqq k^*$. Let $V_\varepsilon \subset V$ denote the open ball of radius ε centered at x_0. Fix δ so that $V_\delta \subset V$ and denote by r_ε the supremum of the possible ranks of those maps (11) with $p \geqq 1$ and $x = x_0$ for which all the points of the trajectory

$$x_i = f_{v_i\cdots v_1}(x_0), \qquad i = 1, \cdots, p,$$

lie in V_ε. Note that r_ε is nondecreasing with ε. Let $r = \inf\{r_\varepsilon | 0 < \varepsilon < \delta\}$ and let $\varepsilon^* := \sup\{\varepsilon | r = r_\varepsilon\}$. Note that $\varepsilon^* > 0$. Take $0 < \sigma < \varepsilon^*$ such that all trajectories starting from V_σ stay in V_{ε^*} for the next k^*+1 steps, under the constant control $u = 0$. Let the corresponding supremum of ranks defining $r_\sigma = r$ be achieved at p and $(v_1^*, \cdots, v_p^*)$.

We define our manifold S as previously, where W is a neighborhood of $(v_1^*, \cdots, v_p^*)$ such that all trajectories corresponding to controls in W lie in V_σ. By an analogous argument as for (a) we see that the vector fields $\mathrm{Ad}_{u_k\cdots u_1} X^+_{u_0}$ are tangent to S, provided that $k \leqq k^*$ and $u_0, \cdots, u_k$ are close enough to zero so that our trajectory does not leave V_{ε^*}, and so the rank cannot increase over r (cf. the definition of σ). Taking $u_1 = \cdots = u_k = 0$ and the derivative $(\partial/\partial u_0)^i$ at u^* we conclude that the vector fields $\mathrm{Ad}_0^k X_0^{+,i}$ are tangent to S. Therefore, their Lie brackets must be tangent to S, also. Because of our choice of the neighborhoods, there are n linearly independent vector fields among those Lie brackets and so S is an open subset of $\mathbb{X}$.

Statement (c) follows from (a) and (b) and the reversion principle. □

The above proof, part (a), gives a somewhat stronger result, actually, which we state below for further use.

COROLLARY 4.4. *If $y \in \mathbb{X}$ is a point forward reachable from x with maximal rank (in the sense of the ranks of maps (11)), then the condition* $\dim \mathrm{Lie}\{\Gamma^+\}(y) = n$ *implies that the system (1) is forward accessible from x.*

We are now ready to establish a converse to Theorem 5.

THEOREM 6. (a) *If system (1) is of class C^1 and forward accessible from x, then*

$$\dim \Gamma^+(x) = n.$$

(b) *If system (1) is analytic, forward accessible from x, and $\mathbb{U}$ is connected, then*

$$\dim L^+(x) = n.$$

(c) *Analogous results hold for backward accessibility with Γ^+, L^+ replaced by Γ^-, L^-.*

Remark 4.5. The case when $\mathbb{U}$ is a nonconnected subset of $\mathbb{R}$ can also be treated. Assume that $\mathbb{U}$ is a disjoint union of connected subsets of $\mathbb{R}$, each of which is in the closure of its interior. Then (b) also holds but we have to choose a subset $\mathbb{U}_0 \subset \mathbb{U}$ which has at least one point in each of these sets. Then

$$L^+ = \mathrm{Lie}\{\mathrm{Ad}_{u_k \cdots u_1} X_u^+ \mid k \geqq 0, u \in \mathbb{U}, u_1, \cdots, u_k \in \mathbb{U}_0\}$$

must be used in this case as the definition of L^+.

Proof of Theorem 6. (a) If the system is accessible, then it follows from Proposition 2.3 that, for some $k \geqq 1$ the rank of the map $\psi_{k,x}$ is equal to n at some point. This means that the following vectors span an n-dimensional space, for some sequence $u_1, \cdots, u_k$:

$$\frac{\partial}{\partial u_i} f_{u_k \cdots u_1}(x), \qquad i = 1, \cdots, k.$$

Hence, also the vectors

$$(\mathrm{d}f_{u_k \cdots u_1}(x))^{-1} \frac{\partial}{\partial u_i} f_{u_k \cdots u_1}(x)$$

which can be equivalently written as

$$\left.\frac{\partial}{\partial v}\right|_{v=0} f^{-1}_{u_{i-1} \cdots u_1} \circ f^{-1}_{u_i} \circ f_{u_i + v} \circ f_{u_{i-1} \cdots u_1}(x) = \mathrm{Ad}_{u_{i-1} \cdots u_1} X^+_{u_i},$$

$i = 1, \cdots, k$, span an n-dimensional space and statement (a) follows.

(b) The proof will be based on a reduction to continuous time systems, as done in [29] for the transitivity problem. A different proof, not involving such a reduction, is provided in a later section. If our system is accessible from x, then it follows from Proposition 2.3 that there exists a k such that the rank of the map

$$(u_1, \cdots, u_k) \mapsto f_{u_k \cdots u_1}(x)$$

is equal to n at some point $(u_1^*, \cdots, u_k^*)$, and so its image contains an open set V. Then $W = f_0^{-k}(V)$ is also open and $x \in W$. We will show that W is contained in the orbit through x of the Lie algebra L^+ (cf. [34]), which we denote by $\mathrm{Orb}_{L^+}(x)$. This will imply that the orbit is of dimension n and from a theorem of Nagano ([23], [34]) it will follow that $\dim L^+(x) = n$.

Let $y \in W$. We will show that $y \in \mathrm{Orb}_{L^+}(x)$ by showing the equivalent fact: $x \in \mathrm{Orb}_{L^+}(y)$. We have that

$$x = f^{-1}_{u_1} \circ \cdots \circ f^{-1}_{u_k} \circ f^k_0(y) = g_{1,u_1} \circ \cdots \circ g_{k,u_k}(y),$$

where

$$g_{i,u_i} = f_0^{-i+1} \circ f_{u_i}^{-1} \circ f_0^i.$$

Denote $y_k = y$, and

$$y_{i-1} = g_{i,u_i}(y_i), \qquad i = k, \cdots, 1.$$

We have that $y_0 = x$. It is enough to show that $y_{i-1} \in \mathrm{Orb}_{L^+}(y_i)$, for $i = 1, \cdots, k$.

Denote

$$\gamma(u) = f_0^{-i+1} \circ f_u^{-1} \circ f_0^i(y_i).$$

Then, for $u \in [0, u_i]$, γ is a curve in $\mathbb{X}$ joining y_i with y_{i-1}; its tangent vector at u is

$$\frac{\partial}{\partial u}\gamma(u) = \frac{\partial}{\partial v}\bigg|_{v=0} f_0^{-i+1} \circ f_{u+v}^{-1} \circ f_u \circ f_0^{i-1}(\gamma(u)) = \mathrm{Ad}_0^{i-1}\, Y_u^+(\gamma(u)).$$

As $\gamma(0) = y_i$ and $\mathbb{U}$ is connected, it follows that $y_{i-1} = \gamma(u_i)$ belongs to the orbit through y_i of the family of vector fields $\mathrm{Ad}_0^{i-1} Y_u^+$, $u \in \mathbb{U}$. Since $Y_u^+ = -X_u^+$, it then follows that y_{i-1} belongs to the orbit through y_i of the family $\mathrm{Ad}_0^{i-1} X_u^+$, $u \in \mathbb{U}$. □

Remark 4.6. If $\mathbb{U}$ is not connected, then the result still holds with the modified definition of the Lie algebra L^+ as given in the remark following Theorem 6. The necessary modifications in the above proof are as follows. We choose elements $v_1, \cdots, v_k \in \mathbb{U}_0$ so that v_i belongs to the same connected component of $\mathbb{U}$ as u_i^*. Then we define

$$W = f_{v_k}^{-1} \circ \cdots \circ f_{v_1}^{-1}(V).$$

Then we have that

$$x = g_{1,u_1} \circ \cdots \circ g_{k,u_k}(y), \quad g_{i,u_i} = f_{v_{i-1}\cdots v_1}^{-1} \circ f_{u_i}^{-1} \circ f_{v_i} \circ f_{v_{i-1}\cdots v_1}.$$

Finally, we take the curve

$$\gamma(u) = f_{v_{i-1}\cdots v_1}^{-1} \circ f_u^{-1} \circ f_{v_i\cdots v_1}(y_i),$$

with u in the interval joining u_i and v_i. Differentiation with respect to u now gives the vector fields in the modified Lie algebra L^+ as defined in the remark following Theorem 6.

To obtain criteria for transitivity using Theorems 5 and 6, we may apply the following trick which reduces the transitivity problem to the forward accessibility problem.

Define $\mathbb{U}^\pm$ as the disjoint union of two copies of $\mathbb{U}$ denoted by $\mathbb{U}^+$ and $\mathbb{U}^-$. Consider a system

$$x^+ = f^\pm(x, u), \quad x(t) \in \mathbb{X}, \quad u(t) \in \mathbb{U}^\pm = \mathbb{U}^+ \cup \mathbb{U}^- \tag{12}$$

where $f^\pm(x, u) = f(x, u)$ if $u \in \mathbb{U}^+$ and $f^\pm(x, u) = f^-(x, u) = f_u^{-1}(x)$ if $u \in \mathbb{U}^-$. As the control set $\mathbb{U}^\pm$ has two components, we define its Lie algebra of our new system L^+ using the definition in Remark 4.6 with $\mathbb{U}_0 = \{0^+, 0^-\}$, where $0^+ \in \mathbb{U}^+$ and $0^- \in \mathbb{U}^-$ are two copies of $0 \in \mathbb{U}$. Of course, there is no difficulty in embedding the new control set again in the reals. The following proposition is then clear.

PROPOSITION 4.7. (a) *The Lie algebra L^+ of the system* (12) *is equal to the Lie algebra L of the original system* (1).

(b) *The family of vector fields* Γ^+ *for system* (12) *is equal to the family* Γ *defined by system* (1).

(c) *The forward accessible set of system* (12) *is equal to the orbit of system* (1).

We may now complete the proofs of all the theorems in this section.

Proof of Theorem 2. Statement (a) follows immediately from Theorems 5 and 6, part (a). Statement (b) follows analogously from part (c) of these theorems. Finally, statement (c) is the consequence of statement (a) via the above reduction of the transitivity problem to the forward accessibility problem and Proposition 4.7. □

Proof of Theorem 3. Statement (a) follows from Theorem 5 (a) and the inclusion $L^+ \subset \text{Lie}\{\Gamma\}$ (sufficiency), and from Theorem 6(b). Statement (b) follows analogously from statements (c) of these theorems. Finally, statement (c) is the consequence of statement (a) via the above reduction trick and Proposition 4.7. □

Proof of Theorem 4. (a) In the smooth case the "if" part follows from Corollary 4.4 by the above reduction procedure and Proposition 4.7 as, for system (12) the point x is attainable from itself with full rank. The analytic case follows from the smooth case by the inclusion $L(x) \subset \text{Lie}\{\Gamma\}(x)$.

The "only if" part follows from Theorem 6 and Proposition 4.7 via the above reduction.

(b) The "only if" part is the consequence of Theorem 6. To prove the "if" part suppose that there are n linearly independent vectors in $L^+(x_0)$. Each of them can be taken in the form

$$\text{ad}\,(\text{Ad}_0^{k_1} X_{u_1}^+) \cdots \text{ad}\,(\text{Ad}_0^{k_{p-1}} X_{u_{p-1}}^+)(\text{Ad}_0^{k_p} X_{u_p}^+)(x_0). \tag{13}$$

If we take the partial derivatives of these vectors with respect to $u_1, \cdots, u_p$ at zero, we obtain vectors which appear in the Lie algebra in (10). From the Taylor formula it follows then that the rank condition in (10) is also satisfied and Theorem 5 implies the result. □

5. Nonaccessible systems. In this section we will briefly discuss nonaccessible and, more generally, nontransitive systems. The following "orbit theorem" is crucial in understanding such systems. The theorem has a long history starting with results of Chow, Nagano [23], Sussmann [34], and Stefan [33] in the continuous time case. In the discrete time case, analogous results to those in continuous time were provided in [9], [32], [11], and [29], the latter containing also a proof of a more abstract result dealing with a general notion of action on manifolds. These papers should be consulted for details of the proof, which we omit.

THEOREM 7. *Any orbit* $A(x)$ *of the smooth system* (1) *is an immersed submanifold of* $\mathbb{X}$ *with at most countably many connected components, whose tangent space is given by*

$$T_yA(y) = \Gamma(y)$$

at each $y \in A(x)$. *In the analytic case we have that*

$$T_yA(y) = L(y)$$

holds also.

As the attainable set from x lies in the orbit from x, there is no chance for forward or backward accessibility from x if there is no transitivity from x (that is, the orbit is not of full dimension). In this case it is reasonable to ask whether the attainable set has a nonempty interior in the orbit. In the case of analytic continuous time systems the answer is always positive, as proved by Sussmann and Jurdjevic [36]. The following theorem generalizes this result to discrete time systems.

THEOREM 8. *If x_0 is an equilibrium point of an analytic system* (1), *then each of the attainable sets $A^+(x_0)$ and $A^-(x_0)$ has a nonempty interior in the orbit $A(x_0)$.*

Proof. If we restrict our system to the orbit then the problem reduces to proving that the system is forward (backward) accessible from x_0, if it is transitive from x_0. But this follows immediately from Theorem 4 and Proposition 4.2. □

Remark 5.1. The above theorem provides an analogue of what is sometimes called the *positive form of Chow's lemma* for continuous time systems. In fact, the proof is related to that of the continuous time case. However, there is an interesting subtlety that appears here. Contrary to the continuous situation, it is *not* true now that the assumption that x_0 is an equilibrium state can be relaxed. In the paper [29, Remark 9.15], an example is given of an analytic system on $\mathbb{X}=\mathbb{R}$, with $\mathbb{U}=\mathbb{R}$, and a state $x \in \mathbb{X}$ such that $A(x)=\mathbb{X}$, but the system is not forward accessible from this x. In fact, the system in question arises from the sampling of a continuous time system.

We now give the basic outline of how such an example arises. A real-analytic function of one variable

$$g(x)$$

is first constructed, with the property that

$$|g'(x)| \leqq 1 \text{ for all } x \in \mathbb{R}$$

and whose zeros are exactly at the nonnegative integers $0, 1, 2, \cdots$. Now the system is given by equations

$$x^+ = 1 + x + ug(x)$$

with

$$\mathbb{U} = (-1, 1)$$

as control value set. Observe that this system is indeed invertible, since for each fixed u the right-hand side is a strictly increasing function of x. Furthermore, for each x the set

$$\{x, 1+x, 2+x, \cdots\}$$

is included in $A^+(x)$. When x is a nonnegative integer, this is *precisely* $A^+(x)$, while for any other x one can reach an open set in one step, and hence $A^+(x)$ is of dimension 1. Since each nonnegative integer x can be reached from, say, -1, it follows that $A(x)=A(-1)$ has dimension 1, so by connectedness, *the orbit through each point is all of* $\mathbb{X}=\mathbb{R}$, even though $A^+(0)$, $A^+(1), \cdots$ are discrete.

These remarks probably mean that the notion of transitivity is in the discrete time case too weak to be of interest.

The following families of vector fields will help us to better understand the geometry of the attainable sets $A^+(x)$ and $A^-(x)$ and, in particular, to estimate their dimensions. Define

$$\Delta_k^+ = \{\mathrm{Ad}_0^i\, X_u^+ | 0 \leqq i \leqq k-1,\ u \in \mathbb{U}\}, \qquad L_k^+ = \mathrm{Lie}\, \Delta_k^+,$$

and

$$\Delta_k^- = \{\mathrm{Ad}_0^{-i}\, X_u^- | 0 \leqq i \leqq k-1,\ u \in \mathbb{U}\}, \qquad L_k^- = \mathrm{Lie}\, \Delta_k^-.$$

For any family of vector fields Δ, let $\mathrm{Orb}_\Delta(x)$ denote the orbit of this family passing through x. This orbit has a natural structure of immersed second countable submanifold ([34], [33]). Further, the orbit of Lie Δ coincides with the orbit of Δ.

PROPOSITION 5.2. *For any smooth system with connected control set* $\mathbb{U}$ *we have that*

$$A_k^+(x) \subset \mathrm{Orb}_{\Delta_k^-}(y), \quad \textit{for any} \quad y \in A_k^+(x),$$

and

$$A_k^-(x) \subset \mathrm{Orb}_{\Delta_k^+}(y), \quad \textit{for any } y \in A_k^-(x).$$

Proof. It is enough to prove the first inclusion as the second will follow from the reversion principle. It is also enough to show this inclusion for any particular y in $A_k^+(x)$, since for any other y this will be implied by the general equality $\mathrm{Orb}_\Delta(y) = \mathrm{Orb}_\Delta(z)$ for any $z \in \mathrm{Orb}_\Delta(y)$. Our argument will be similar to that used in the proof of Theorem 6(b). Take

$$y = f_0^k(x)$$

and

$$z = f_{u_k \cdots u_1}(x) = f_{u_k} \circ \cdots \circ f_{u_1} \circ f_0^{-k}(y).$$

We have to show that $z \in \mathrm{Orb}_{\Delta_k^-}(y)$. The point z can be written in a different way as

$$z = g_{k,u_k} \circ \cdots \circ g_{1,u_1}(y),$$

where

$$g_{i,u} = f_0^{k-i} \circ f_u \circ f_0^{-k+i-1}, \qquad i = 1, \cdots, k.$$

Taking $z_0 = y$, $z_i = g_{i,u_i}(z_{i-1})$, $i = 1, \cdots, k$, it is enough to show that $z_i \in \mathrm{Orb}_{\Delta_k^-}(z_{i-1})$. Consider the curve $\gamma_i(u) = g_{i,u}(z_{i-1})$, which joins z_{i-1} with z_i when $u \in [0, u_i]$. The tangent vectors to this curve are given by

$$\frac{\partial}{\partial u}\gamma_i(u) = \frac{\partial}{\partial v}\bigg|_{v=0} f_0^{k-i} \circ f_{u+v} \circ f_u^{-1} \circ f_0^{i-k}(\gamma_i(u))$$
$$= \mathrm{Ad}_0^{i-k}\, Y_u^-(\gamma_i(u)) = -\mathrm{Ad}_0^{i-k}\, X_u^-(\gamma_i(u)).$$

As $-k+1 \leqq i-k \leqq 0$, it follows that the above curve lies in the orbit of the family Δ_k^- and the proof is complete. □

From the above proposition we immediately conclude the following necessary conditions for accessibility.

COROLLARY 5.3. *If an analytic system with connected* $\mathbb{U}$ *is forward accessible from* x, *then*

$$\dim L^-(y) = n \quad \textit{for any } y \in A^+(x).$$

Similarly, if it is backward accessible from x, *then*

$$\dim L^+(y) = n \quad \textit{for any } y \in A^-(x).$$

Proof. The first statement follows directly from the first inclusion in Proposition 5.2 and the inclusions

$$\bigcup_{k>0} A_k^+(x) = A^+(x), \quad \mathrm{Orb}_{\Delta_k^-}(x) \subset \mathrm{Orb}_{L^-}(x).$$

The second statement follows analogously. □

We now turn to yet another reason why our Lie algebras of vector fields emerge in studying controllability properties of discrete time systems. We will consider our system in another (time-dependent) system of coordinates. This is basically the same as the "local" dynamics defined in the references [18] and [20] in the context of invariant distributions for nonlinear discrete-time systems.

Consider the usual system $x(t+1)=f(x(t), u(t))$ and introduce the time-dependent change of variables

$$x(t)=f_0^t(z(t)),$$

where f_0^t is the tth power of f_0 (in the sense of composition). In the new coordinates our system becomes time-dependent and takes the form

$$z(t+1)=g(t, z(t), u(t)), \tag{14}$$

where

$$g(t, z, u)=f_0^{-t-1}\circ f_u\circ f_0^t(z).$$

What is simpler about the new system is that it has the "doing nothing" option, as $g(t,\cdot,0)=\mathrm{id}$. As a consequence, if the control set $\mathbb{U}$ is connected then so are the attainable sets of system (14): $A^+(x)$, $A^-(x)$, and $A(x)$. In that case the next point on the trajectory, $z(t+1)$, can be connected with the previous one, $z(t)$, by the smooth curve $\gamma(u)=g(t, z(t), u)$, where $u\in[0, u(t)]$ if $u(t)>0$ and $u\in[u(t), 0]$ if $u(t)<0$. As

$$\begin{aligned}\partial\gamma/\partial u(u)&=\partial g/\partial u(t, z(t), u)\\&=\frac{\partial}{\partial v}\bigg|_{v=0}(f_0^{-t-1}\circ f_{u+v}\circ f_u^{-1}\circ f_0^{t+1}(\gamma(u)))\\&=\mathrm{Ad}_0^{t+1}\,Y_u^-(\gamma(u)),\end{aligned}$$

we see that the point $z(t)$ lies in the orbit through $z(t+1)$ of the family of vector fields $\mathrm{Ad}_0^{t+1}\,Y_u^-$, $u\in\mathbb{U}$. Since $Y_u^-=-X_u^-$, it follows by induction that for $t\leqq -1$ any point $z(t)$ on a trajectory of system (14) starting from $z(0)$ lies in the orbit through $z(0)$ of the family of vector fields Δ_k^-, where $k=-t$ and so also in the orbit through $z(0)$ of the Lie algebra L_k^-. By the reversion principle, or by the above argument applied for $t>0$, it also follows that any point $z(t)$ of any trajectory of system (14) starting from $z(0)$ lies in the orbit through $z(0)$ of the family of vector fields Δ_k^+, with $k=t$, and so also in the orbit through $z(0)$ of the Lie algebra L_k^+.

Because of our change of coordinates $x(t)=f_0^t(z(t))$ it follows that a point $x(t)$ on any trajectory of the original system (1) starting from x_0, lies in the image under the map f_0^k of the orbit $\mathrm{Orb}_{\Delta_k^+}(x_0)$ if $t=k>0$ (respectively, the image of $\mathrm{Orb}_{\Delta_k^-}(x_0)$, if $t<0$, $k=-t$). Thus, we have the following proposition.

PROPOSITION 5.4. *If the control set $\mathbb{U}$ is connected then, for any $k>0$, we have the inclusions*

$$A_k^+(x)\subset f_0^k(\mathrm{Orb}_{\Delta_k^+}(x))=f_0^k(\mathrm{Orb}_{L_k^+}(x))$$

and

$$A_k^-(x)\subset f_0^{-k}(\mathrm{Orb}_{\Delta_k^-}(x))=f_0^{-k}(\mathrm{Orb}_{L_k^-}(x)).$$

The orbits of discrete time systems can be expressed via the orbits of the Lie algebra L according to the formula

$$A(x)=\bigcup_{k\in\mathbb{Z}} f_0^k(\mathrm{Orb}_L(x)).$$

Proof. The first two inclusions follow from the argument above. It also follows from the above consideration that the vector fields in L are tangent to the orbit $A(x)$ (cf. Theorem 7). Thus, $\mathrm{Orb}_L(x)\subset A(x)$. As the maps f_0^k preserve the orbit $A(x)$ and the family of vector fields L, it follows that the inclusion "$\supset$" holds. On the other hand, the computation preceding the proposition also shows that any two points which can be joined by a (forward or backward) step of the discrete time system can also be joined by a trajectory of a continuous time system

$$\dot{x} = h(x, u), \quad \text{where } h(x, u) = \mathrm{Ad}_0^k X_u^+(x)$$

and a (forward or backward) jump by f_0. It is well known that each trajectory of a continuous time system lies in a single orbit of this system. It follows then that any trajectory of the above system lies in an orbit of the family of vector fields L, and so the inclusion "$\subset$" follows. □

The relation between the inclusions in Propositions 5.2 and 5.4 can be further clarified by the following relation between the Lie algebras L_k^+ and L_k^-.

PROPOSITION 5.5. *For an analytic system the distributions spanned by the Lie algebras L_k^+ and L_k^- are related by the change of coordinates given by the diffeomorphism f_0^k, i.e.,*

$$(\mathrm{Ad}_0^k L_k^-)(x) = L_k^+(x), \quad \textit{and} \quad (\mathrm{Ad}_0^{-k} L_k^+)(x) = L_k^-(x) \quad \forall x \in \mathbb{X}.$$

Proof. Since the operator Ad_0 is a homomorphism of the Lie algebra of vector fields, it follows that

$$\mathrm{Ad}_0^k L_k^- = \mathrm{Lie}\,\{\mathrm{Ad}_0^i X_u^- | 1 \leqq i \leqq k\}.$$

From Proposition 3.4 it follows that

$$(\mathrm{Ad}_0^i X_u^-)(x) \in \mathrm{Lie}\,\{\mathrm{Ad}_0^{i-1} X_u^+ | u \in \mathbb{U}\}(x) \quad \forall x.$$

Thus, all the vector fields $\mathrm{Ad}_0^i X_u^-$, $i = 1, \cdots, k$ are tangent to the orbit of the Lie algebra L_k^+ and so

$$(\mathrm{Ad}_0^k L_k^-)(x) \subset L_k^+(x) \quad \forall x \in \mathbb{X}. \tag{15}$$

The reversion principle and the above inclusion yield

$$(\mathrm{Ad}_0^{-k} L_k^+)(x) \subset L_k^-(x) \quad \forall x \in \mathbb{X}.$$

Applying the operator Ad_0^k to both sides of the above inclusion gives the converse inclusion to (15) and proves the first equality in the proposition.

The second equality follows from the first and the reversion principle. □

6. Nonscalar controls. All our previous results can be extended, without difficulties, to the case of multidimensional controls. The basic modification needed is that, whenever derivatives with respect to u are used in the scalar control case, partial derivatives with respect to the components of u should be used in the multicontrol case.

We assume that the control set $\mathbb{U}$ is a subset of $\mathbb{R}^m$ and satisfies the assumption $\mathbb{U} \subset \mathrm{clos\,int}\,\mathbb{U}$. Additionally, we assume that any two points in the same connected component of $\mathbb{U}$ can be joined by a smooth curve lying entirely in $\mathrm{int}\,\mathbb{U}$ (except of endpoints, possibly). We denote $u = (u^1, \cdots, u^m)$ and $v = (v^1, \cdots, v^m)$.

The vector fields X_u^+ defined at the beginning of § 3 should now be redefined as follows:

$$X_{u,i}^+(x) = \frac{\partial}{\partial v^i}\bigg|_{v=0} f_u^{-1} \circ f_{u+v}(x),$$

one for each $i = 1, \cdots, m$. Analogously, we define $X_{u,i}^-$, $Y_{u,i}^+$ and $Y_{u,i}^-$.

The Lie algebras Γ^+, Γ^- and Γ are now defined as

$$\Gamma^+ = \{\mathrm{Ad}_{u_k \cdots u_1} X_{u_0,i}^+ | k \geqq 0, 1 \leqq i \leqq m, u_0, \cdots, u_k \in \mathbb{U}\},$$

$$\Gamma^- = \{\mathrm{Ad}_{u_k \cdots u_1}^{-1 \cdots -1} X_{u_0,i}^- | k \geqq 0, 1 \leqq i \leqq m, u_0, \cdots, u_k \in \mathbb{U}\},$$

$$\Gamma = \{\mathrm{Ad}_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1} X_{u_0,i}^\sigma | k \geqq 0, 1 \leqq i \leqq m, u_0, \cdots, u_k \in \mathbb{U}, \varepsilon_1, \cdots, \varepsilon_k = \pm 1, \sigma = \pm\}.$$

We also redefine the Lie algebras L^+, L^-, and L as follows. We choose a subset $\mathbb{U}_0 \subset \mathbb{U}$ which has at least one point in each connected component of $\mathbb{U}$. In particular, if the set $\mathbb{U}$ is connected and $0 \in \mathbb{U}$ we can take $\mathbb{U}_0 = \{0\}$. We define

$$L^+ = \text{Lie}\,\{\text{Ad}_{u_k \cdots u_1} X^+_{u,i} | k \geqq 0, 1 \leqq i \leqq m, u \in \mathbb{U}, u_1, \cdots, u_k \in \mathbb{U}_0\},$$

$$L^- = \text{Lie}\,\{\text{Ad}^{-1 \cdots -1}_{u_k \cdots u_1} X^-_{u,i} | k \geqq 0, 1 \leqq i \leqq m, u \in \mathbb{U}, u_1, \cdots, u_k \in \mathbb{U}_0\},$$

$$L = \text{Lie}\,\{\text{Ad}^{\varepsilon_k \cdots \varepsilon_1}_{u_k \cdots u_1} X^\sigma_{u,i} | k \geqq 0, 1 \leqq i \leqq m, u \in \mathbb{U}, u_1, \cdots, u_k \in \mathbb{U}_0, \varepsilon_1, \cdots, \varepsilon_k = \pm 1, \sigma = \pm\}.$$

THEOREM 9. *With the above definitions of the Lie algebras Γ^+, Γ^-, Γ, L^+, L^-, and L, all the theorems stated in the preceding two sections remain true.*

The proof of the multicontrol versions are completely analogous to the scalar case. The main modifications needed are the replacement of derivatives with respect to u by partial derivatives with respect to the components of u, and the replacement of parameterizations of curves by u with parameterizations by components of u. We leave the details to the reader.

7. From discrete time to continuous time systems. In this section we have two goals. The first is the description of one manner in which the study of continuous time systems can be reduced to that of discrete time systems. The second is the development of a technique, based on expansions of the previously defined families of vector fields, which gives added power to the use of these vector fields and their associated Lie algebras. As an illustration of the use of this technique, we provide a short proof of part (b) of Theorem 6 which is independent of Nagano's theorem and of the orbit theorem. In this manner, not only does the discrete time theory become independent of continuous time techniques, but in fact it becomes itself a basis for the accessibility theory for the latter, via the reduction also described here.

To show how continuous-time systems can be viewed as a special case of discrete time systems, we consider a continuous-time system of the form

$$\dot{x} = h(x, v), \tag{16}$$

where $x(t) \in \mathbb{X}$ and $v(t) \in V$ is the control. We assume that the controls are piecewise constant (this assumption does not affect the controllability properties of the system we are studying). For the convenience of having all the maps defined everywhere we assume that our system is complete. We introduce the discrete-time system

$$x^+ = f(x, u), \quad x(t) \in \mathbb{X}, u(t) \in \mathbb{U} = \mathbb{R}_+ \times V, \quad \mathbb{R}_+ = [0, \infty), \tag{17}$$

where $u = (t, v)$ and $f(x, u) = \exp(th(\cdot, v))(x)$. In this way, going forward by time t with a constant control v for the continuous-time system corresponds to a forward step using the control $u = (t, v)$ for the discrete time system. Analogously, going backward by time $-t$ with the control v corresponds to a backward step with $u = (t, v)$. This implies that the forward (respectively, backward) attainable sets as well as the orbits of both systems (16) and (17) coincide. Thus both systems have identical controllability properties.

It is convenient to endow V with the discrete topology. The set $\mathbb{U} = \mathbb{R}_+ \times V$ can be viewed then as the disjoint union of copies of $\mathbb{R}_+$. We compute the Lie algebras L^+, L^-, and L corresponding to system (17) according to the remark following Theorem 6. We choose the subset $\mathbb{U}_0 = \{(0, v) | v \in V\} \subset \mathbb{U}$. Then $f_0 = \text{id}$ and we can easily compute that

$$X^+_u = h(\cdot, v) = -X^-_u, \quad \text{for } u = (t, v).$$

Strictly speaking, the present set $\mathbb{U}$ is not an allowable control set, since it is not a subset of $\mathbb{R}^m$. However, the arguments in previous sections can be repeated as long as we use in the definition of X_u^+ and X_u^- only differentiation with respect to t but not differentiation with respect to v. Finally, we obtain

$$L^+ = L^- = L = \text{Lie}\,\{h(\cdot, v) | v \in V\}.$$

Our aim now is to prove a discrete time version of the well-known Baker–Campbell–Hausdorff expansion formula for a vector field Y transformed by the flow of a vector field Z:

$$\text{Ad}_u\, Y = \sum_{k=0}^{\infty} \frac{1}{k!} \text{ad}^k\, Z(Y).$$

This is classical when Ad_u corresponds to $f_u = \exp(uZ)$, for which $X_u^+ = Z = -X_u^-$. Assume now that f is of the general form $f = f(x, u)$; we wish to generalize the above formula.

LEMMA 7.1. *For analytic f and Y we have the following expansions, for u sufficiently close to zero,*

$$\text{Ad}_u\, Y = \sum_{k=0}^{\infty} \int_0^u \int_0^{v_1} \cdots \int_0^{v_{k-1}} \text{ad}\, X_{v_1}^+ \cdots \text{ad}\, X_{v_k}^+\, \text{Ad}_0\, Y\, dv_k \cdots dv_1,$$

$$\text{Ad}_u^{-1}\, Y = \sum_{k=0}^{\infty} \int_0^u \int_0^{v_1} \cdots \int_0^{v_{k-1}} \text{ad}\, X_{v_1}^- \cdots \text{ad}\, X_{v_k}^-\, \text{Ad}_0^{-1}\, Y\, dv_k \cdots dv_1,$$

where the series converge pointwise at each $x \in \mathbb{X}$.

If f and Y are of class C^∞ only, then we have the formula

$$\text{(18)}\qquad \text{Ad}_u^{\pm 1}\, Y = \sum_{i=0}^{k} \int_0^u \int_0^{v_1} \cdots \int_0^{v_{i-1}} \text{ad}\, X_{v_1}^\pm \cdots \text{ad}\, X_{v_i}^\pm\, \text{Ad}_0^{\pm 1} Y\, dv_i \cdots dv_1 + R_k,$$

where

$$R_k = \int_0^u \int_0^{v_1} \cdots \int_0^{v_k} \text{ad}\, X_{v_1}^\pm \cdots \text{ad}\, X_{v_{k+1}}^\pm\, \text{Ad}_{v_{k+1}}^{\pm 1}\, Y\, dv_{k+1} \cdots dv_1.$$

(Note the subscript "0" in $\text{Ad}_0^{\pm 1}\, Y$ in each of the above formulas except for the one for the reminder term R_k.)

In order to prove the above lemma we shall first prove the following estimate. Below we shall denote by $|\psi|$ the absolute value of ψ, if ψ is a scalar, and the "max" norm $|\psi| = \max\{|\psi_1|, \cdots, |\psi_n|\}$, if ψ is a vector $\psi = (\psi_1, \cdots, \psi_n)$.

LEMMA 7.2. *Let x be a point in $\mathbb{R}^n$. If $Y_0, \cdots, Y_k$ are real analytic vector fields on a subset of $\mathbb{R}^n$ containing x that have complex analytic continuations (denoted by the same letters) to the closed polydisc $D = D_{x,r} = \{z \in \mathbb{C}^n \| z_1 - x_1| \leqq r, \cdots, |z_n - x_n| \leqq r\}$, then*

$$\text{(19)}\qquad |\text{ad}\, Y_k \cdots \text{ad}\, Y_2(Y_1)(x)| \leqq \sup_{z\in D} |Y_k(z)| \cdots \sup_{z \in D} |Y_1(z)| (2/r)^{k-1} k^k.$$

Proof. Before we prove the estimate in the lemma, we shall derive the following estimate. Let ϕ be a real analytic function which has a complex analytic extension to the polydisc D. Then the iterated derivative of ϕ along the vector fields $Y_1, \cdots, Y_k$ can be estimated by

$$\text{(20)}\qquad |Y_k \cdots Y_1 \phi(x)| \leqq \sup_{z \in D} |\phi(z)| \sup_{z\in D} |Y_1(z)| \cdots \sup_{z\in D} |Y_k(z)| (k/r)^r.$$

To prove this estimate we use a method of Sussmann [34] (proof of Lemma 4.2) which reduces the problem to Cauchy inequalities. Consider the complex analytic vector fields $z_1 Y_1, \cdots, z_k Y_k$ defined on D, where $z_1, \cdots, z_k$ are complex parameters in the unit disc $\{z \| z| \leqq 1\}$. Let $\exp(tz_i Y_i)$ denote the flow of $z_i Y_i$ in $\mathbb{C}^n$. Then

$$\phi \circ \exp(t_1 z_1 Y_1) \circ \cdots \circ \exp(t_k z_k Y_k)(x)$$

is a well-defined analytic function on the unit polydisc $|z_1| \leqq 1, \cdots, |z_k| \leqq 1$, if

$$|t_i| \leqq r(k \sup |Y(z)|_{z\in D})^{-1}, \qquad i = 1, \cdots, k \tag{21}$$

(as the concatenation of the trajectories of $z_1 Y_1, \cdots, z_k Y_k$ starting from x does not leave D if $t_1, \cdots, t_k$ satisfy the above inequalities). From the Cauchy inequality we obtain then that the iterated derivative at the origin of this function with respect to $z_1, \cdots, z_k$ is estimated by the supremum of this function on the unit polydisc. This gives the inequality

$$|(t_k Y_k) \cdots (t_1 Y_1)\phi(x)| \leqq \sup_{z\in D} |\phi(z)|.$$

If we take the maximal values of $t_1, \cdots, t_k$ in the inequalities (21), the above gives (20).

The estimate in (20) gives the inequalities

$$|Y_{i_k} \cdots Y_{i_1}\phi_i(x)| \leqq \sup_{z\in D} |Y_1(z)| \cdots \sup_{z\in D} |Y_k(z)| k^k r^{-k+1}, \tag{22}$$

for $\phi_i = x_i$ and $i_1, \cdots, i_k$ any permutation of $1, \cdots, k$. These inequalities imply the estimate in (19) as the left-hand side of this estimate can be replaced by the components of the vector field given by ad $Y_k \cdots$ ad $Y_1 \phi_i$ and each such component consists of 2^{k-1} terms of the form as in (22) (this follows from the definition of the Lie bracket as a commutator). □

Proof of Lemma 7.1. Integration of the first equation in Proposition 3.3 between 0 and u gives

$$\text{Ad}_u Y = \text{Ad}_0 Y + \int_0^u \text{ad } X_v^+(\text{Ad}_v Y)\, dv.$$

Replacing $\text{Ad}_v Y$ on the right by this expression yields

$$\text{Ad}_u Y = \text{Ad}_0 Y + \int_0^u \text{ad } X_v^+(\text{Ad}_0 Y)\, dv + \int_0^u \int_0^{v_1} \text{ad } X_{v_1}^+ \text{ad}_{v_2}^+ (\text{Ad}_{v_2} Y)\, dv_2\, dv_1.$$

Repeating such a replacing k times gives the "+" case of formula (18). The "−" case follows by the reversion principle.

To prove the first formula of the lemma we shall now use the estimate in Lemma 7.2. Our families of vector fields, X_u^+ and $\text{Ad}_u Y$, are analytic with respect to x and u. Let us fix an $x \in \mathbb{X}$. Then, there exist an $r > 0$ and a u_0 such that both families have complex analytic extensions to the complex polydisc D in $\mathbb{C}^n$, with the (real) center at x and radius r, for all $u \in [0, u_0]$. Denote

$$C = \sup_{z\in D, u\in[0,u_0]} |X_u^+(z)|, \qquad D = \sup_{z\in D, u\in[0,u_0]} |\text{Ad}_u Y(z)|.$$

Lemma 7.2 gives the following estimate for $R_k(x)$ with, if $u \in [0, u_0]$,

$$|R_k(x)| \leqq (2/r)^k (k+1)^{k+1} C^{k+1} D \int_0^u \int_0^{v_1} \cdots \int_0^{v_k} dv_{k+1} \cdots dv_1$$

$$= CDu_0 (2Cu_0/r)^k \frac{(k+1)^{k+1}}{(k+1)!}.$$

From Stirling's formula,

$$\lim_{k\to\infty} (2\pi k)^{1/2} e^k k^k (k!)^{-1} = 1,$$

it follows then that $R_k(x)$ tends to zero as k tends to infinity. This implies that the first series in the lemma converges.

The second formula follows from the first by the reversion principle. □

Both expansions in Lemma 7.1 can be combined to obtain a more general expansion. In order to have a compact expression for this expansion we introduce the following notation. Define the following linear operators acting on vector fields Y or, more generally, on smooth families of vector fields $Y_{(\cdot)}$ depending on $u \in \mathbb{U}$:

$$\text{ad}\, I_u^{\sigma,i} Y_{(\cdot)} := \int_0^u \int_0^{v_1} \cdots \int_0^{v_{i-1}} \text{ad}\, X_{v_1}^\sigma \cdots \text{ad}\, X_{v_i}^\sigma Y_{v_i}\, dv_i \cdots dv_1,$$

and $\text{ad}\, I_u^{\sigma,0} Y_{(\cdot)} = Y_u$, where σ is either $+$ or $-$. With this notation, formula (18) in Lemma 7.1 takes the form

$$\text{Ad}_u^{\pm 1}\, Y = \sum_{i=0}^{k} \text{ad}\, I_u^{\pm,i}\, \text{Ad}_0^{\pm 1}\, Y + \text{ad}\, I_u^{\pm,k+1}\, \text{Ad}_{(\cdot)}^{\pm 1}\, Y.$$

Finally, using analogous techniques as above, one can also establish the forward/backward version of the above.

LEMMA 7.3. *If f and the vector field Y are analytic, then the following expansion holds*:

$$\text{Ad}_{u_k \cdots u_1}^{\varepsilon_k \cdots \varepsilon_1} Y = \sum_{i_1 \geqq 0, \cdots, i_k \geqq 0}^{\infty} \text{ad}\, I_{u_k}^{\sigma_k, i_k}\, \text{Ad}_0^{\varepsilon_k} \cdots \text{ad}\, I_{u_1}^{\sigma_1, i_1}\, \text{Ad}_0^{\varepsilon_1}\, Y,$$

where σ_j is the sign of ε_j, $j = 1, \cdots, k$, and the series converges pointwise for small enough u's.

From this we can draw the following conclusions.

COROLLARY 7.4. *If the system is analytic and $\mathbb{U}$ is connected, then*

$$L^+(x) = \Gamma^+(x), \quad L^-(x) = \Gamma^-(x), \quad L(x) = \Gamma(x),$$

for any $x \in \mathbb{X}$.

Again, the result is valid also in the nonconnected case provided that one modifies the definitions of the Lie algebras as explained in Remark 4.5.

Because of Corollary 7.4, part (b) is equivalent to part (a) in Theorem 6. This provides the promised direct proof of part (b) of Theorem 6.

8. Sampling. In this section, we explain briefly how some of our results can be applied to the sampling problem. More details are given in the conference paper [31]. For other related facts about sampling, the reader should consult [19] and [21].

When a continuous-time system is digitally controlled, decisions are often restricted to be taken at fixed times $0, \delta, 2\delta, \cdots$; $\delta > 0$ is the *sampling time*. Under what is often called zeroth-order hold sampled control, the resulting situation can be modeled through the constraint that the inputs applied be constant on intervals of length δ. It is thus of interest to characterize the preservation of basic system properties when the controls are so restricted. For controllability, this problem motivated the results in the classical paper of Kalman, Ho, and Narendra [13]. This studied the case of linear systems and established that controllability when sampling at intervals of length δ is preserved if $\delta(\lambda - \mu)$ is not of the form $2k\pi i$ for any pair of distinct eigenvalues of the A matrix. The dual version of this result, for observability, is basically the classical

Nyquist–Shannon sampling theorem from digital signal processing, and is often summarized by the statement that controllability (or observability) is preserved provided that one samples at more than twice the natural frequencies of the system. We sketch here how a similar result can be obtained for certain nonlinear systems, using the accessibility conditions given above. This is an improvement over the result in [30], where only the case of bilinear systems was treated, and more importantly, where only transitivity conditions were obtained.

Let $\tilde{\Sigma}_d$ denote the class of all continuous time systems Σ of the type

$$\dot{x} = Fx + \sum_{i=1}^{m} u_i g_i(x), \tag{23}$$

where F is an n by n matrix and the coordinates of all the g_i are polynomials of degree at most d. For instance, $\tilde{\Sigma}_0$ is the class of all linear systems (the g_i's are constant vectors), while $\tilde{\Sigma}_1$ is the class of bilinear systems. Here $x(t) \in \mathbb{R}^n$ and $u_i(t) \in \mathbb{R}$ for each t; n is the dimension of the system, m the number of independent controls. We shall study controllability properties of (23) from the initial state $x_0 = 0$. Nonequilibrium initial states can also be studied, but we restrict ourselves to the equilibrium case, always reducible to $x_0 = 0$, for simplicity. We let $f(x) = Fx$ be the linear vector field corresponding to the matrix F.

We shall say that the *natural frequencies* of the system (23) are the imaginary parts of the eigenvalues of F, and let $\Omega(\Sigma, 0)$, or just Ω, be the set of these numbers (counted with multiplicities). Note that since F is real, $-\omega \in \Omega$ whenever $\omega \in \Omega$. For each nonnegative integer j we denote by $\mathcal{B}_j$ the set of all linear combinations

$$\frac{1}{k} \sum_{i=1}^{n} \rho_i \omega_i \tag{24}$$

with k any nonzero integer, $\omega_1, \cdots, \omega_n$ the natural frequencies, and the ρ_i's nonnegative integers satisfying

$$\sum_{i=1}^{n} \rho_i = 2j + 2.$$

Note that if λ is the largest of the ω_i (equivalently, the largest absolute value of these), each element of $\mathcal{B}_j$ is in magnitude bounded by $(2j+2)\lambda$.

Denote the set of states of the continuous time system Σ that can be reached from 0 in time $T > 0$, using arbitrary (measurable locally integrable) controls $u(\cdot)$ by A^T. We shall say that the system (23) is (*forward*) *accessible from* 0 if A^T has nonempty interior for some $T > 0$. Let $\omega > 0$ be any real number. We shall say that Σ is *ω-accessible from* 0, or *accessible under sampling at frequency ω from* 0, if the set of states A^T_ω reachable from 0 in time T using controls sampled at that frequency has a nonempty interior. A control $u(\cdot)$ defined on an interval $[0, T]$ is said to be sampled at frequency ω (in radians/sec) if and only if T is an integer multiple of $\delta := 2\pi/\omega$, say $T = r\delta$, and there are vectors

$$v_1, \cdots, v_r$$

such that $u(t) \equiv v_i$ on the interval $[(i-1)\delta, i\delta)$. Thus accessibility under sampling corresponds to forward accessibility for a discrete time system derived from the corresponding Σ and ω. With this definition it is clear that ω-accessibility for even a single ω implies accessibility. The following theorem from [31] provides a converse to this fact. The corollary is immediate from the theorem and the discussion given above about the largest frequency λ.

THEOREM 10. *Assume that $\Sigma \in \tilde{\Sigma}_d$ is accessible from 0. If $\omega > 0$ is not in $\mathcal{B}_j$ for any $j \leqq d$, then Σ is also ω-accessible.*

COROLLARY 8.1. *Accessibility is preserved under sampling for systems in $\tilde{\Sigma}_d$ provided that the sampling frequency be larger than $2d+2$ times the largest natural frequency of the system.*

The reader is referred to [31] for the details of the reduction of the above theorem to the results given earlier in this paper. However, we wish to at least sketch this reduction here. For each fixed δ, the vector fields X^+ can be explicitly described using a Lie expansion formula ([4], see also [25], and especially [19], [21]):

$$X_{0,i}^+ = \frac{\partial}{\partial \varepsilon}\bigg|_{\varepsilon=0} e^{-\delta f} e^{\delta(f+\varepsilon g_i)}(x).$$

(We will be interested here only on the case $u=0$.) Under suitable assumptions, which are satisfied for the class of systems considered here, this can also be written as

$$\theta_\delta(\operatorname{ad} f)(g_i),$$

where as earlier $\operatorname{ad} f$ is the operator $\operatorname{ad} f(h) = [f, h]$ and for each fixed real number δ, θ_δ is the entire function

$$\theta_\delta(z) := \frac{e^{\delta z} - 1}{z}.$$

Finally, one also has a formal expression, for each fixed δ,

$$\operatorname{Ad}_0 = e^{\delta \operatorname{ad} f}.$$

This expression can be made rigorous when acting on polynomial vector fields such as those that appear in the classes $\tilde{\Sigma}_d$. Thus the Lie algebra L^+, for each fixed δ, contains the Lie algebra $\tilde{L}^+$ generated by the vector fields

$$\{\theta_\delta(\operatorname{ad} f)(g_1), \cdots, \theta_\delta(\operatorname{ad} f)(g_m), e^{\delta \operatorname{ad} f}\theta_\delta(\operatorname{ad} f)(g_1), \cdots, e^{\delta \operatorname{ad} f}\theta_\delta(\operatorname{ad} f)(g_m), \cdots,$$
$$e^{\delta k \operatorname{ad} f}\theta_\delta(\operatorname{ad} f)(g_1), \cdots e^{\delta k \operatorname{ad} f}\theta_\delta(\operatorname{ad} f)(g_m), \cdots\},$$

which equals the span of the vector fields

$$\{g_1, \cdots, g_m, [f, g_1], \cdots, [f, g_m], \cdots, \operatorname{ad}^k f(g_1), \cdots, \operatorname{ad}^k f(g_m), \cdots\}$$

when δ is as in Theorem 10 (see [31] for details). It follows that $\tilde{L}^+$ coincides with the strong accessibility Lie algebra associated to the original continuous time system, which has full rank at the origin due to the accessibility assumption. Then Theorem 4 gives the desired result.

9. An example. Consider the following invertible polynomial system with $\mathbb{X} = \mathbb{R}^3$.

$$\begin{aligned} x^+ &= x(z^2+1)^2 \\ y^+ &= y(z^2+1)^3 \\ z^+ &= z+u, \end{aligned} \tag{25}$$

where we are using the superscript + to denote time shift, and we denote coordinates as (x, y, z). Calculating, we obtain that $X_u^+ = -2z(z^2+1)^{-1}Z - X_u^-$ and $X_u^- = (0, 0, -1)'$, where Z is the vector field

$$\begin{pmatrix} 2x \\ 3y \\ 0 \end{pmatrix},$$

for each $u \in \mathbb{R}$. Since the basic vector fields X_u^+ and X_u^- turn out to be independent of u in this example, we drop the subscripts u from now on. Further,

$$\mathrm{Ad}_0 X^+ = \begin{pmatrix} -8xz(z^2+1)^{-1} \\ -12yz(z^2+1)^{-1} \\ 1 \end{pmatrix} = 2X^+ + X^-, \tag{26}$$

from which it follows that

$$\mathrm{span}\,\{X^+, \mathrm{Ad}_0 X^+\} = \mathrm{span}\,\{X^+, X^-\}.$$

The identity $\mathrm{Ad}_0 X^- = -X^+$ (cf. Proposition 3.2(b)) implies that

$$\mathrm{Ad}_0^2 X^+ = 2\,\mathrm{Ad}_0 X^+ - X^- \in \mathrm{span}\,\{X^+, X^-\},$$

so the linear span of the set of all generators of L^+, $\{\mathrm{Ad}_0^k X^+, k \geqq 0\}$, coincides with the span of X^+ and X^-. Similarly, applying Ad_0^{-1} to both sides of (26),

$$\mathrm{Ad}_0^{-1} X^- = X^+ - 2\,\mathrm{Ad}_0^{-1} X^+ = X^+ + 2X^-,$$

so the span of the $\{\mathrm{Ad}_0^k X^-, k \leqq 0\}$, the generators of L^-, is again the same. Finally,

$$[X^+, X^-] = 2(1-z^2)(z^2+1)^{-2} Z,$$

from which it follows that $\{X^-, X^+, [X^+, X^-]\}$ and $\{X^-, Z\}$ span the same C^∞ submodule of vector fields. The latter set is involutive, and we conclude that, for this example,

$$L^+ = L^- = L.$$

Thus the orbits have dimension 2 through each point except at those points with $x = y = 0$, where Z vanishes, and there the dimension is 1. The tangent spaces are given by the vectors $\partial/\partial z$ and $2x\partial/\partial x + 3y\partial/\partial y$. The forward and backward accessible sets contain open subsets of each orbit, by the equality of these Lie algebras.

Of course, in this very simple example one can analyze the system directly. The initial states (x_0, y_0, z_0) with $x_0 = y_0 = 0$ are such that the only possible directions of movement are those in which z changes, as is clear from the equations (25), consistently with the above conclusion about tangent spaces. The points where exactly one of x_0 or y_0 is nonzero are also easy to analyze. Take now a point with both x_0 and y_0 nonzero. Consider the set C consisting of all points (x, y, z) with

$$y_0^2 x^3 = x_0^3 y^2.$$

This is the cross product of a cusp with a line. The forward accessible set consists of all (x, y, z) in C with sign $y = \text{sign } y_0$ for which $|x| \geqq |x_0|$ and $|y| \geqq |y_0|$. The backward accessible has both these inequalities reversed, and the orbit consists of the branch of C with just sign $y = \text{sign } y_0$. Note how each such set C, an algebraic variety, can be stratified into three submanifolds, which turn out to be its singular set (the orbit of $(0, 0, 0)$), the orbit of (x_0, y_0, z_0), and the orbit of the "conjugate" point $(x_0, -y_0, z_0)$. See Fig. 6 for a picture of a typical cross-section with constant z.

Thus in this example both the forward-accessible set and the orbit from each point are open subsets of an irreducible algebraic variety. More generally, similar behavior may be expected when dealing with invertible polynomial systems and equilibrium initial states. We conjecture that the orbit is an open subset of the *quasi-reachable* set in the sense of [26] and [27]. This is an algebraic variety, and it can be computed explicitly, via Jacobians of the n-step transition map. Note that polynomial invertible systems may exhibit highly nonlinear behavior, such as in the case $x^+ = x^3 + x + u$,

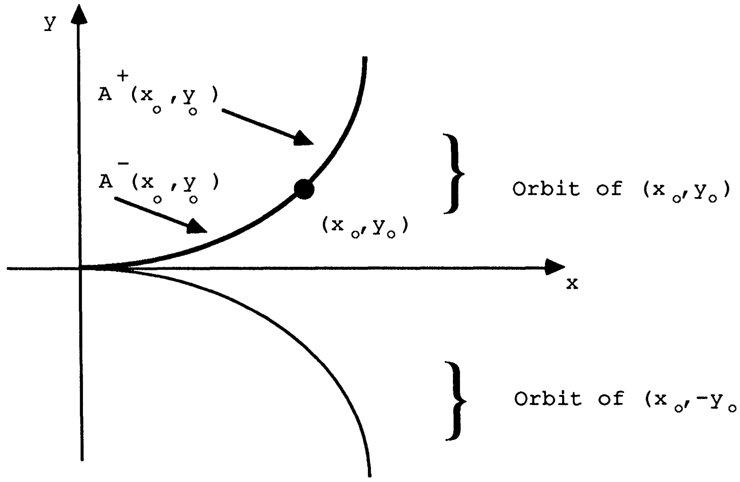

FIG. 6. *Forward and backward accessible sets in example* $(x_0, y_0 > 0)$.

where the inverse of the transition mapping is not even rational. We plan to study such systems in greater depth in the future.

10. An alternative formalism. We now briefly describe the formalism due to Monaco and Normand-Cyrot; the thesis [25] and the papers [17]-[22], as well as the references given there, should be consulted for details.

Their approach is based on the introduction of certain operators and the formal relations that these satisfy. As a first step, one writes the system equations as

$$x^+ = x + f(x, u)$$

so that the new "f" is our $f(x, u) - x$. Thus now f indicates what the increment is, rather than the new state, making things more analogous to differential equations. (This is similar to the introduction of the forward difference operator in numerical analysis.)

For simplicity we shall assume again that inputs are scalar, and also that $\mathbb{X} = \mathbb{R}^n$. Thus we may identify functions $F: \mathbb{R}^n \to \mathbb{R}^n$ (in particular, the functions $F = f(\cdot, u)$) with vector fields, in the usual coordinate system for $\mathbb{R}^n$,

$$F = \sum_{i=1}^{n} F_i(\cdot) \frac{\partial}{\partial x_i}.$$

We will work purely formally, since the intent is merely to point out the relations with the alternative notations in the papers mentioned above. Formally then, one introduces the operators on smooth functions

$$L_F^{\otimes k} := \sum_{i_1, \cdots i_k = 1}^{n} F_{i_1}(\cdot) \cdots F_{i_k}(\cdot) \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}}$$

and the complete series

$$\Delta_F := I + \sum_{k \geqq 1} \frac{1}{k!} L_F^{\otimes k}.$$

Now one can obtain similar series for compositions and inverses of the dynamics map.

Further, the vector fields that we use can be expressed then as

$$X_u^+(x) = \frac{\partial}{\partial v}\bigg|_{v=0} \Delta_{f(\cdot,u+v)} \circ \Delta_{f(\cdot,u)}^{-1}(\mathrm{Id})|_x,$$

$$X_u^-(x) = \frac{\partial}{\partial v}\bigg|_{v=0} \Delta_{f(\cdot,u+v)}^{-1} \circ \Delta_{f(\cdot,u)}(\mathrm{Id})|_x,$$

$$Y_u^+(x) = \frac{\partial}{\partial v}\bigg|_{v=0} \Delta_{f(\cdot,u)} \circ \Delta_{f(\cdot,u+v)}^{-1}(\mathrm{Id})|_x,$$

$$Y_u^-(x) = \frac{\partial}{\partial v}\bigg|_{v=0} \Delta_{f(\cdot,u)}^{-1} \circ \Delta_{f(\cdot,u+v)}(\mathrm{Id})|_x,$$

$$\mathrm{Ad}_0^k\, X_u^\sigma(x) = \frac{\partial}{\partial v}\bigg|_{v=0} \Delta_{f(\cdot,0)}^{k}\Delta_{f(\cdot,u+v)}^{\sigma} \circ \Delta_{f(\cdot,u)}^{-\sigma}\Delta_{f(\cdot,0)}^{-k}(\mathrm{Id})|_x,$$

and many properties of these vector fields can be obtained from the corresponding expansions.

The reader is directed to the above references for details on how these expansions can be very useful in studying, among others, problems of disturbance decoupling, sampling, Volterra expansions, linearization, and realization.

REFERENCES

[1] A. Arapostathis, B. Jakubczyk, H.-G. Lee, S. I. Marcus, and E. D. Sontag, *The effect of sampling on linear equivalence and feedback linearization*, submitted.

[2] R. Brockett, *Nonlinear systems and differential geometry*, Proc. IEEE, 64 (1976), pp. 61–72.

[3] M. Fliess and D. Normand-Cyrot, *A group-theoretic approach to discrete-time nonlinear controllability*, Proc. IEEE Conf. Decision and Control, Albuquerque, NM, 1981.

[4] R. Goodman, *Lifting vector fields to nilpotent Lie groups*, J. Math. Pures et Appl., 57 (1978), pp. 77–86.

[5] J. W. Grizzle, *Controlled invariance for discrete-time nonlinear systems with an application to the decoupling problem*, IEEE Trans. Automat. Control, AC-30 (1985), pp. 868–874.

[6] J. W. Grizzle and P. V. Kokotovic, *Feedback linearization of sampled-data systems*, IEEE Trans. Automat. Control, to appear.

[7] R. Hermann and A. J. Krener, *Nonlinear controllability and observability*, IEEE Trans. Automat. Control, 22 (1977), pp. 728–740.

[8] A. Isidori, *Nonlinear Control Systems: An Introduction*, Springer-Verlag, Berlin, 1985.

[9] B. Jakubczyk, *Invertible realizations of nonlinear discrete time systems*, Proc. Princeton Conf. Inf. Sci. and Sys., 1980, pp. 235–239.

[10] ———, *Feedback linearization of discrete-time systems*, Systems Control Lett., 9 (1987), pp. 411–416.

[11] B. Jakubczyk and D. Normand-Cyrot, *Orbites de pseudo-groupes de diffeomorphismes et commandabilité des systemes non linéaires en temps discret*, C.R. Acad. Sci. Paris, 298-I (1984), pp. 257–260.

[12] B. Jakubczyk and E. D. Sontag, *The effect of sampling on feedback linearization*, Proc. IEEE Conf. Decision and Control, Los Angeles, Dec. 1987, pp. 1374–1379.

[13] R. E. Kalman, Y. C. Ho, and K. S. Narendra, *Controllability of linear dynamical systems*, Contr. Diff. Eqs., 1 (1963), pp. 189–213.

[14] H. G. Lee, A. Arapostathis, and S. I. Marcus, *On the linearization of discrete time systems*, Internat. J. Control, 45 (1987), pp. 1103–1124.

[15] S. P. Meyn, *Ergodic theorems for discrete time stochastic systems using a generalized stochastic Lyapunov function*, SIAM J. Control Optim., 27 (1989), pp. 1409–1439.

[16] A. Mokkadem, *Orbites de semi-groupes de morphismes réguliers. Systèmes non linéaires en temps discret.*, Forum Math., to appear.

[17] S. Monaco and D. Normand-Cyrot, *Développements founctionels pour les systèmes non linéaires en temps discret*, report from University of Rome, 1984, submitted for publication.

[18] ———, *Invariant distributions for nonlinear discrete-time systems*, Systems Control Lett., 5 (1984), pp. 191–196.

[19] S. MONACO AND D. NORMAND-CYROT, *On the sampling of a linear control system*, Proc. IEEE Conf. Decision and Control, Ft. Lauderdale, Florida, 1985, p. 1095.
[20] ———, *Nonlinear systems in discrete-time*, in Algebraic and Geometric Methods in Control Theory, M. Fliess and M. Hazewinkel, eds., Reidel, Dordrecht, 1986, pp. 411–430.
[21] ———, *A Lie exponential formula for the nonlinear discrete time functional expansion*, in Theory and Applications of Nonlinear Control Systems, C. Byrnes and A. Lindquist, eds., North Holland, Amsterdam, 1986.
[22] ———, *Invariant distributions under sampling*, in Theory and Applications of Nonlinear Control Systems, C. Byrnes and A. Lindquist, eds., North Holland, Amsterdam, 1986.
[23] T. NAGANO, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan, 18 (1966), pp. 398–404.
[24] H. NIJMEIJER, *Observability of autonomous discrete-time nonlinear systems: a geometric approach*, Internat. J. Control, 36 (1982), pp. 867–874.
[25] D. NORMAND-CYROT, *Théorie et Pratique des Systèmes Non Linéaires en Temps Discret*, Thèse de Docteur d'Etat, Université Paris-Sud, 1983.
[26] E. D. SONTAG AND Y. ROUCHALEAU, *On discrete-time polynomial systems*, J. Nonlinear Analysis, 1 (1976), pp. 55–64.
[27] E. D. SONTAG, *Polynomial Response Maps*, Springer-Verlag, Berlin, New York, 1979.
[28] ———, *A concept of local observability*, Systems Control Lett., 5 (1984), pp. 41–47.
[29] ———, *Orbit theorems and sampling*, in Algebraic and Geometric Methods in Nonlinear Control Theory, M. Fliess and M. Hazewinkel, eds., Reidel, Dordrecht, 1986, pp. 441–486.
[30] ———, *An eigenvalue condition for sampled weak controllability of bilinear systems*, Systems Control Lett., 7 (1986), pp. 313–316.
[31] ———, *A Chow property for sampled bilinear systems*, in Analysis and Control of Nonlinear Systems, C. I. Byrnes and B. Saeks, eds., North Holland, Amsterdam, 1988, pp. 483–492.
[32] E. D. SONTAG AND H. J. SUSSMANN, *Accessibility under sampling*, Proc. IEEE Conf. Decision and Control, Orlando, Florida, Dec. 1982.
[33] P. STEFAN, *Attainable sets are manifolds*, preprint, University of Wales, 1973.
[34] H. J. SUSSMANN, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc., 180 (1973), pp. 171–188.
[35] ———, *Lie brackets, real analyticity, and geometric control*, in Differential Geometric Control Theory, R. W. Brockett, R. S. Millman, and H. J. Sussmann, eds., Birkhauser, Boston, 1983.
[36] H. J. SUSSMANN AND V. JURDJEVIC, *Controllability of nonlinear systems*, J. Differential Equations, 12 (1972), pp. 95–116.