SIAM J. CONTROL AND OPTIMIZATION
Vol. 19, No. 5, September 1981


0363-0129/81/1905-0003 $01.00/0

# GENERIC OBSERVABILITY OF DIFFERENTIABLE SYSTEMS*

DIRK AEYELS†

**Abstract.** A dynamical system consists of a smooth vectorfield defined on a differentiable manifold, and a smooth mapping from the manifold to the real numbers. The vectorfield represents the dynamics of a physical system. The mapping stands for a measuring device by which experimental information on the dynamics is made available. The information itself is modeled as a sampled version of the image of the state trajectory under the smooth mapping. In this paper the observability of this set-up is discussed from the viewpoint of genericity. First the observability property is expressed in terms of transversality conditions. Then the theory of transversal intersection is called upon to yield the desired results. It is shown that almost any measuring device will combine with a given physical system to form an observable dynamical system, if $(2n+1)$ samples are taken and not fewer, where $n$ is the dimension of the manifold. Dually, it is shown that almost any physical system will combine with a given measuring device to form an observable dynamical system, if $(2n+1)$ samples are taken and not fewer. The analysis leads to the corollary that for nonlinear systems observability is a generic property, a fact well known for linear systems.

The relation of the theory to the study of turbulence and to control theory is explained.

**1. Introduction.** Consider a physical system with dynamics modeled by a smooth vectorfield defined on its state space, viz., a finite dimensional second countable smooth manifold with no boundary. An investigation of the qualitative behavior of the flow of smooth vectorfields has been carried out over the last two decades. In general, from a practical standpoint, the phase portrait cannot be observed directly. Instead, by means of a measuring device—modeled as a smooth function from the configuration space to the reals—experimental information on the dynamical process can be made available. This information is modeled as the image of the state trajectory under the output function, mentioned above. In this manner a mapping has been defined which assigns an output trajectory to each state trajectory. If this mapping is bijective, it makes sense to undertake a study of the state dynamics of the system, starting from experimental evidence. In control theory, a system is a pair consisting of a smooth vectorfield and a smooth output function. A system is observable if the mapping mentioned above is one-to-one.

This paper is concerned with the question whether observability is a natural assumption. Two types of problems will be considered, which in some sense are dual to each other. In the first problem an almost arbitrarily chosen smooth vectorfield is given. The question is whether the choice of a measuring device is "critical" in order to be able to investigate the flow from the output, i.e., in order to achieve observability. In the second problem an almost arbitrarily chosen measuring device is available. The question here is whether the vectorfields that can be investigated with this apparatus, i.e., the vectorfields pairing with the measuring device to form an observable system, constitute a "big" subset of the set of all smooth vectorfields. For both problems, observability turns out to be a generic property if one is allowed to sample the output trajectories $(2n+1)$ times *and not fewer.*

The notion of observability stands central in control theory. It is a necessary assumption in the reconstruction problem, i.e., the problem of recovering the state trajectory corresponding to the sampled output trajectory. Observability is also of interest in the study of turbulence and chaos, as pointed out by an anonymous referee who communicated the following problem. Assume that a dynamical system has a global attractor. Here, instead of finding the state trajectory corresponding to the

* Received by the editors May 20, 1980, and in revised form January 9, 1981.
† Department of System Dynamics, State University of Gent, Gent, Belgium.

observed output trajectory, one is asked to recover a homeomorphic picture of the global attractor or some characteristic properties of the attractor. We will return to this in the final section.

Within the context of linear, algebraic and analytic systems, some aspects of observability as a generic property have been treated in the literature [1], [2], [3]. Here we will be concerned with the general class of smooth nonlinear systems. As opposed to the more algebraic orientation in the above-mentioned references, the results in this paper are obtained as an application of parametric transversality theory.

The organization of the paper is as follows. In the second section the technical definition of observability is given and formulated as a transversality property. In § 3, the two problems mentioned above are considered. The answer to the first question is a rather immediate consequence of a parametric transversality theorem. To obtain the answer to the second question, more work is needed. As usual, verifying the transversality of the evaluation mapping, defined later, is responsible for the main part of the proof. This is harder for the second problem, where transversality has to be achieved by manipulating the vectorfields.

All this manipulation is "filtered" through the fixed output mapping. In fact, one has to show that this filtering process does not affect the transversality of the evaluation mapping. This accounts for the greater length of the proof of the result of the second problem as compared with the first problem, where this type of difficulty does not arise. In § 4 it is shown, by means of a counterexample, that in general at least $2n+1$ samples are necessary in order to achieve observability. In § 5 the paper is concluded with a few additional remarks on turbulence and control theory.

**2. Mathematical preliminaries. The observability property.** Let $X$ be a $C^k$ differentiable paracompact manifold with $k$ sufficiently high. Let $\xi \in \mathscr{X}^r(X)$, $r \geqq 1$, the space of all $C^r$ vectorfields defined on $X$. Let $h \in \mathscr{C}^r(X, \mathbb{R})$, $r \geqq 1$, the space of all $C^r$ functions mapping into $\mathbb{R}$. Both $\mathscr{X}^r(X)$ and $\mathscr{C}^r(X, \mathbb{R})$ are endowed with the Whitney $C^r$ topology. The function $h$ is called the output mapping. The results derived in this paper can be easily extended to output functions mapping into more general spaces. The proofs remain virtually unaltered. The classical set theoretic definition of observability goes as follows.

DEFINITION A. The pair $(\xi, h)$ is *observable over a time interval* $[0, T]$, $T$ a positive number, if and only if for each pair $(x, y) \in X \times X \backslash \Delta(X \times X)$ there exists a time $t \in [0, T]$ such that $h \circ \phi_t(x) \neq h \circ \phi_t(y)$. Here $\Delta(X \times X)$ is the diagonal of $X \times X$. It is remarked that in this definition, to each pair of points $(x, y)$ there might correspond a different time instant $t$.

Transversality theory is not particularly well-adapted for a discussion of Definition A, because of the infinite dimensionality of the output trajectory space. Also, considerations from engineering practice suggest the placement of a sampling device into the system description. Let $P$ be a *sample program*, i.e., a finite set of points $t_i \in [0, T]$, with $T$ given a priori; see also [1].

DEFINITION B. A system $(\xi, h)$ is *observable with respect to P*, or shortly, *P-observable*, if and only if for each $(x, y) \in X \times X \backslash \Delta(X \times X)$, there exists a $t_i \in P$ such that $h \circ \phi_{t_i}(x) \neq h \circ \phi_{t_i}(y)$.

Although stronger in general, Definition B is equivalent to Definition A for linear systems, when considering sample programs with at least $n$ points. For easy reference recall the following density theorem [4] basic to this paper. Let $A$, $X$, $Y$ be $C^r$-manifolds; $\mathscr{C}^r(X, Y)$ is the set of $C^r$ mappings from $X$ to $Y$ and $\rho: A \to \mathscr{C}^r(X, Y)$ is a map which is called a $C^r$ *representation* if and only if the *evaluation map* $\mathrm{ev}_\rho: A \times X \to Y$

with $\mathrm{ev}_\rho(a, x) = \rho(a)(x)$ is a $C^r$ map from $A \times X$ to $Y$. In the following, we write $\rho_a$ instead of $\rho(a)$ (i.e., $\rho_a : X \to Y$ is a $C^r$ map).

DENSITY THEOREM. *Let $W$ be a submanifold of $Y$. Define $\mathscr{A}_W \subset \mathscr{A}$ by $\mathscr{A}_W = \{a \in \mathscr{A} : \rho_a \pitchfork W\}$ ($\pitchfork$ is the symbol for transversality),*

*Assume that:*

(1) *$X$ has finite dimension $n$ and $W$ has finite codimension $q$ in $Y$.*

(2) *$\mathscr{A}$ and $X$ are second countable.*

(3) *$r > \max(0, n - q)$.*

(4) *$\mathrm{ev}_\rho \pitchfork W$ (this will be called transversality of the evaluation mapping).*

*Then $\mathscr{A}_W$ is residual (and hence dense) in $\mathscr{A}$.*

As for the connection of observability with transversality theory, it is remarked that observability is equivalent to injectiveness of a particular mapping. Notice also that $f : X \to Y$ is injective if and only if $(f \times f) : X \times X \to Y \times Y$ does not intersect $\Delta(Y \times Y)$ when restricted to $X \times X \backslash \Delta(X \times X)$. Under the right differentiability conditions and when $\dim(X \times X) < \operatorname{codim} \Delta(Y \times Y)$ this is equivalent to $(f \times f) \pitchfork_x \Delta(Y \times Y)$, $x \in X \times X \backslash \Delta(X \times X)$.

Given an observation interval $[0, T]$, by a sample program $P$ we mean a set with at least $(2n+1)$ different points $t_i$, $0 \leqq t_i \leqq T$ with $n = \dim X$. From now on, $P$ will be assumed to have exactly $2n+1$ points. Given $P$, define the evaluation mapping

$$\mathrm{ev}: \Delta(\mathscr{A} \times \mathscr{A}) \times (X \times X) \to \mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1},$$

with $\mathscr{A} = \mathscr{X}^r(X) \times \mathscr{C}^r(X, \mathbb{R})$ and

$$\mathrm{ev}(\xi, h, \xi, h, x, y) = \left\{ \begin{matrix} h \circ \phi_{t_1}(x) \\ \vdots \\ h \circ \phi_{t_{2n+1}}(x) \end{matrix} \right\}, \left\{ \begin{matrix} h \circ \phi_{t_1}(y) \\ \vdots \\ h \circ \phi_{t_{2n+1}}(y) \end{matrix} \right\}.$$

Here $\xi \in \mathscr{X}^r(X)$, $h \in \mathscr{C}^r(X, \mathbb{R})$, $x \in X$, $y \in X$. $\phi : X \times \mathbb{R} \to X$ denotes the flow corresponding to $\xi$.

It is clear that the natural bijection $\delta : \mathscr{A} \to \Delta(\mathscr{A} \times \mathscr{A})$ induces a $C^r$-manifold structure on $\Delta(\mathscr{A} \times \mathscr{A})$. When $X$ is compact, ev is class $C^r$ [4]. Compactness of $X$ will always be understood in the following. In the concluding remarks, the noncompact case will be treated. In the following section our main concern is showing that the evaluation mapping—with appropriate restrictions on $\mathscr{A}$—is transversal to $\Delta(\mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1})$. This will provide an answer to the problems mentioned in the introduction.

**3. Observability is generic.** In this section two results on observability as a generic property are proved. In the first result, an almost arbitrarily chosen smooth vectorfield is given. It is shown that almost all smooth output functions pair with the vectorfield to form an observable system if $(2n+1)$ samples are taken. For a motivation, one is referred to the introduction. Before announcing the theorem, we state two lemmas whose proofs are direct consequences of transversality theory.

LEMMA 1. *The subset $\mathscr{A} \subset \mathscr{X}^r(X)$, $r \geqq 1$ of vectorfields with a finite number of equilibrium points and a finite number of closed orbits with period $\leqq T$, constitutes an open and dense set of $\mathscr{X}^r(X)$.*

LEMMA 2. *Let $\xi \in \mathscr{A}$ have a finite number of equilibrium points $x_i$. The subset $\mathscr{B} \subset \mathscr{C}^r(X, \mathbb{R})$, $r \geqq 1$ of functions $h$ with $h(x_i) \neq h(x_j)$, $i \neq j$, constitutes an open and dense set of $\mathscr{C}^r(X, \mathbb{R})$.*

THEOREM 1. *Given a vectorfield $\xi \in \mathscr{A}$ and a positive real number $T$, then the set of functions $h$, belonging to $\mathscr{C}^r(X, \mathbb{R})$ such that $(\xi, h)$ is $P$-observable, is open and dense in $\mathscr{C}^r(X, \mathbb{R})$. This is true for almost any sample program $P$ of $(2n+1)$ points $t_i$, $0 \leqq t_i \leqq T$.*

*Proof.* For the proof of the density part, we will consider output functions belonging to $\mathcal{B}$. If density can be shown with respect to $\mathcal{B}$ then it is also shown with respect to $\mathcal{C}^r(X, \mathbb{R})$.

Let $\phi_t(x)$ denote the flow corresponding to $\xi$, and let $\{t_i: i=1, \cdots, 2n+1\}$ denote $(2n+1)$ different sample instants chosen from the interval $[0, T]$. Consider the evaluation mapping

$$\text{ev}: \Delta(\mathcal{B}\times\mathcal{B})\times(X\times X)\to\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1},$$

$$\text{ev}(h, h, x, y)=\begin{Bmatrix} h\circ\phi_{t_1}(x)\\ \vdots\\ h\circ\phi_{t_{2n+1}}(x)\end{Bmatrix}, \begin{Bmatrix} h\circ\phi_{t_1}(y)\\ \vdots\\ h\circ\phi_{t_{2n+1}}(y)\end{Bmatrix}.$$

This mapping is class $\mathcal{C}^r$. A pair $(\xi, h)$ is observable if and only if $\text{ev}(h, h, X\times X\backslash\Delta(X\times X))\cap\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})=\varnothing$, which is equivalent to $\text{ev}\,\pitchfork_{x,y}\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})$ on $X\times X\backslash\Delta(X\times X)$, since $\text{codim}\,\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})=2n+1>2n=\dim(X\times X)$. The application of $2n+1$ samples (or more) is fundamental to this equivalence. Notice that in order to apply the transversality density theorem, the finiteness of the codimension of $W$ in $Y$ is required (for notation, see § 2). It is by considering $P$-observality that the a priori infinite dimensional space of output curves defined on $[0, T]$ is replaced by a finite dimensional sample space. Conditions (1), (2) and (3) in the density theorem of § 2 are satisfied. In order to satisfy condition (4), we have to show that if $\text{ev}(h^*, h^*, x^*, y^*)=:(w, w)\in\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})$, then range ($D$ ev $(h^*, h^*, x^*, y^*)$) contains a complement of $T_{w,w}\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})$ in $T_w\mathbb{R}^{2n+1}\times T_w\mathbb{R}^{2n+1}$. $D$ denotes the derivative, and $T$ denotes the tangent space. When we show, by picking appropriate functions $g\in\mathcal{B}$, that

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0}(\text{ev}(h^*+\lambda g, h^*+\lambda g, x^*, y^*)),\qquad \lambda\in\mathbb{R}$$

can span $T_w\mathbb{R}^{2n+1}\times\{0\}$ or $\{0\}\times T_w\mathbb{R}^{2n+1}$, the proof will be finished. Now

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0}(\text{ev}(h^*+\lambda g, h^*+\lambda g, x^*, y^*))$$

$$=\begin{Bmatrix} g\circ\phi_{t_1}(x^*)\\ \vdots\\ g\circ\phi_{t_{2n+1}}(x^*)\end{Bmatrix}, \begin{Bmatrix} g\circ\phi_{t_1}(y^*)\\ \vdots\\ g\circ\phi_{t_{2n+1}}(y^*)\end{Bmatrix}.$$

Therefore if $x^*$ is not an equilibrium point, then it is possible to pick a $g\in\mathcal{B}$ such that

$$\begin{Bmatrix} g\circ\phi_{t_1}(x^*)\\ \vdots\\ g\circ\phi_{t_{2n+1}}(x^*)\end{Bmatrix}$$

equals any vector $\alpha\in T_w\mathbb{R}^{2n+1}\approx\mathbb{R}^{2n+1}$, a priori given, and we are done. Problems could arise if both $x^*$ and $y^*$ were equilibrium points. This cannot occur by assumption, since then $\text{ev}(h^*, h^*, x^*, y^*)\notin\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})$. When $x^*$ and $y^*$ are both on closed orbits with the same period (or $x^*$ is an equilibrium and $y^*$ is on a closed orbit) a periodic sample program (with period equal to the period of the closed orbits involved) is an obstruction to the proof of the transversality of the evaluation mapping. But this phenomenon cannot occur, due to the phrase "almost any sample program $P$" in the statement of the theorem. Indeed, by this phrase we mean to exclude the above-mentioned periodic sample programs.

The openness part of the theorem is an immediate consequence of the openness-of-transversality theorem [4]. □

*Remarks.* 1. The theorem is valid for $\xi \in \mathscr{A}$ and thus for vectorfields belonging to an open and dense subset of $\mathscr{X}^r(X)$.

2. Theorem 1 implies that the set of $P$-observable (or observable) pairs $(\xi, h)$ is open and dense in $\mathscr{X}^r(X) \times \mathscr{C}^r(X, \mathbb{R})$. This is a well-known result for linear systems. It is remarked that for linear systems the genericity of observability can be shown in a direct way by considering the algebraic characterization of observability in terms of the Kalman matrix.

We now proceed to prove a theorem dual to Theorem 1. Here an almost arbitrarily chosen smooth output function is given. It is shown that almost all smooth vectorfields pair with the output function to form an observable system if $2n+1$ samples are taken. Although the statement of this theorem is dual to the previous theorem, the proof is not. Indeed, showing that the appropriate evaluation mapping is transversal is somewhat involved. For the intuitive reason behind this, and also for a motivation of the problem one is again referred to the introduction. Before stating the theorem, we state two lemmas whose proofs follow directly from transversality theory.

LEMMA 3. *The set of functions $\mathscr{D} \subset \mathscr{C}^r(X, \mathbb{R})$, $r \geqq 1$, with a finite number of nondegenerate critical points $x_i$, and with $h(x_i) \neq h(x_j)$ $i \neq j$ constitutes an open and dense set of $\mathscr{C}^r(X, \mathbb{R})$.*

LEMMA 4. *Given $h \in \mathscr{D}$, consider the set of vectorfields $\mathscr{E} \subset \mathscr{X}^r(X)$ which is the intersection of the sets $\mathscr{E}^1$, $\mathscr{E}^2$ and $\mathscr{E}^3$ defined, respectively, by the following.*

1) *No two equilibrium points belong to the same level surface of $h$.*
2) *No two equilibrium points coincide with critical points of $h$.*
3) *No integral curve contains two (or more) critical points of $h$.*

*The set $\mathscr{E}$ is an open and dense set of $\mathscr{X}^r(X)$.*

THEOREM 2. *Given a function $h \in \mathscr{D}$ and a positive real number $T$, then the set of vectorfields $\xi$ belonging to $\mathscr{X}^r(X)$ such that $(\xi, h)$ is $P$-observable is open and dense in $\mathscr{X}^r(X)$. This is true for almost any sample program of $2n+1$ points $t_i$, $0 \leqq t_i \leqq T$.*

*Proof.* For the proof of the density part of Theorem 2, we will consider vectorfields belonging to $\mathscr{E}$. If density can be shown with respect to $\mathscr{E}$ then it is also shown with respect to $\mathscr{X}^r(X)$.

Consider the evaluation mapping

$$\mathrm{ev}: \Delta(\mathscr{E} \times \mathscr{E}) \times X \times X \to \mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1},$$

$$\mathrm{ev}\,(\xi, \xi, x, y) = \left\{ \begin{matrix} h \circ \phi_{t_1}(x) \\ \vdots \\ h \circ \phi_{t_{2n+1}}(x) \end{matrix} \right\}, \left\{ \begin{matrix} h \circ \phi_{t_1}(y) \\ \vdots \\ h \circ \phi_{t_{2n+1}}(y) \end{matrix} \right\},$$

with $h \in \mathscr{D}$, $\phi_t(x)$ the flow corresponding with $\xi$, and all $t_i$ different. The only difficulty in applying the transversality density theorem, is in proving the transversality of the evaluation mapping. The other conditions are satisfied. We have to show that, given

$$(w, w) := \mathrm{ev}\,(\xi^*, \xi^*, x^*, y^*) \in \Delta(\mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1}),$$

range $(D\,\mathrm{ev}\,(\xi^*, \xi^*, x^*, y^*))$ contains a complement of

$$T_{w,w}\Delta(\mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1}) \quad \text{in } T_w\mathbb{R}^{2n+1} \times T_w\mathbb{R}^{2n+1}.$$

We evaluate the derivative with the partial derivative rule [4]. With $\eta \in \mathbb{R}$, $\eta \in \mathscr{E}$, $\phi^\lambda$ the flow of $\xi^* + \lambda\eta$, we obtain

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathrm{ev}\,(\xi^*+\lambda\eta, \xi^*+\lambda\eta, x^*, y^*)$$

$$= \left.\frac{d}{d\lambda}\right|_{\lambda=0} \left( \begin{Bmatrix} h\circ\phi^\lambda_{t_1}(x^*) \\ \vdots \\ h\circ\phi^\lambda_{t_{2n+1}}(x^*) \end{Bmatrix}, \begin{Bmatrix} h\circ\phi^\lambda_{t_1}(y^*) \\ \vdots \\ h\circ\phi^\lambda_{t_{2n+1}}(y^*) \end{Bmatrix} \right)$$

$$(1) \qquad = \begin{Bmatrix} Dh|_{\phi_{t_1}(x^*)} \cdot \int_0^{t_1} D\phi^0_s \cdot \eta(\phi^0_{-s+t_1}(x^*))\, ds \\ \vdots \\ Dh|_{\phi_{t_{2n+1}}(x^*)} \cdot \int_0^{t_{2n+1}} D\phi^0_s \cdot \eta(\phi^0_{-s+t_{2n+1}}(x^*))\, ds \end{Bmatrix},$$

$$\begin{Bmatrix} Dh|_{\phi_{t_1}(y^*)} \cdot \int_0^{t_1} D\phi^0_s \cdot \eta(\phi^0_{-s+t_1}(y^*))\, ds \\ \vdots \\ Dh|_{\phi_{t_{2n+1}}(y^*)} \cdot \int_0^{t_{2n+1}} D\phi^0_s \cdot \eta(\phi^0_{-s+t_{2n+1}}(y^*))\, ds \end{Bmatrix}.$$

Recall that $x^* \neq y^*$, since transversality of the evaluation mapping has to be verified on $X \times X \backslash \Delta(X \times X)$. We consider three cases.

*Case* 1 Neither $x^*$ nor $y^*$ is an equilibrium point of $\xi^*$.

For each $i \in \{1, 2, \cdots, (2n+1)\}$, at least one of $Dh|_{\phi_{t_i}(x^*)}$ or $Dh|_{\phi_{t_i}(y^*)}$ is not equal to zero. Otherwise $h$ is critical in $\phi_{t_i}(x^*)$ and $\phi_{t_i}(y^*)$, $i = 1, \cdots, 2n+1$, and since $h \in \mathscr{D}$, this implies that $h(\phi_{t_i}(x^*)) \neq h(\phi_{t_i}(y^*))$, $i = 1, 2, \cdots, 2n+1$, and thus

$$\mathrm{ev}\,(\xi^*, \xi^*, x^*, y^*) \notin \Delta(\mathbb{R}^{2n+1} \times \mathbb{R}^{2n+1}).$$

Without loss of generality, we assume

$$(2) \qquad Dh|_{\phi_{t_i}(x^*)} \neq 0, \qquad i = 1, 2, \cdots, 2n+1. \qquad (2)$$

We show that for any vector $\alpha \in T_w\mathbb{R}^{2n+1}$, we can find a $C^r$ vectorfield $\eta_\alpha$ in $\mathscr{E}$ such that

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathrm{ev}\,(\xi^*+\lambda\eta_\alpha, \xi^*+\lambda\eta_\alpha, x^*, y^*) = \alpha \times \{0\}.$$

The vectorfield $\eta_\alpha$ will first be defined on the range of the integral curve from $x^*$ to $\phi_{t_{2n+1}}(x^*)$ and then extended to the whole manifold by a partition of unity argument. We denote by $\alpha_k$ the $k$th component of the vector $\alpha$. We now show how to define $\eta_\alpha$ such that the first component of

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathrm{ev}\,(\xi^*+\lambda\eta_\alpha, \xi^*+\lambda\eta_\alpha, x^*, y^*)$$

equals $(\alpha_1, 0)$.

Choose $\eta_\alpha$ such that $\eta_\alpha(\phi^0_{-s+t_1}(x^*)) = g(s) D\phi^0_{-s} \cdot \dot{x}$, $0 \leqq s \leqq t_1$. The function $g: [0, t_1] \to \mathbb{R}$ is such that $\int_0^{t_1} g(s)\, ds = 1$, and has the property that its first $r$ derivatives evaluated at 0 and $t_1$, are equal to zero. Taking into account (1), and taking $\eta_\alpha = 0$ on the range of the integral curve from $y^*$ to $\phi_{t_1}(y^*)$, one finds that the first component of the derivative of the evaluation map equals $(Dh|_{\phi_{t_1}(x^*)}\cdot \dot{x}, 0)$, which is equal to $(\alpha_1, 0)$ for a good choice of $\dot{x}$. The proof will be finished by induction. Let the vectorfield $\eta_\alpha$ be defined on the range of the integral curve connecting $x^*$ to $\phi_{t_{k-1}}(x^*)$.

The vectorfield $\eta_\alpha$ has also been defined simultaneously on the range of the integral curve connecting $y^*$ to $\phi_{t_{k-1}}(y^*)$, where, because of (2), it has been taken equal to zero. We show how to define $\eta_\alpha$ such that the $k$th component of

$$\frac{d}{d\lambda}\bigg|_{\lambda=0} \operatorname{ev}(\xi^*+\lambda\eta_\alpha, \xi^*+\lambda\eta_\alpha, x^*, y^*)$$

equals $(\alpha_k, 0)$.

Therefore, consider

$$\begin{aligned} Dh|_{\phi_{t_k}(x^*)} &\int_0^{t_k} D\phi_s^0\cdot\eta(\phi^0_{-s+t_k}(x^*))\,ds \\ &= Dh|_{\phi_{t_k}(x^*)}\left(\left(\int_0^{t_{k-1}}+\int_{t_{k-1}}^{t_k}\right)(D\phi_s^0\cdot\eta(\phi^0_{-s+t_k}(x^*))\,ds)\right) \\ &= Dh|_{\phi_{t_k}(x^*)}(\dot{y}+\dot{w}). \end{aligned}$$

Here $\dot{y}$ has been chosen already. The vector $\dot{w}$ can be chosen arbitrarily outside $\ker(Dh|_{\phi_{t_k}(x^*)})$ as indicated above. Therefore transversality of the evaluation mapping has been shown in Case 1.

*Case* 2 One of $x^*$ or $y^*$, say $x^*$, is an equilibrium point of $\xi^*$.

In this case

$$\frac{d}{d\lambda}\bigg|_{\lambda=0}(\operatorname{ev}(\xi^*+\lambda\eta, \xi^*+\lambda\eta, x^*, y^*)) = \left\{\begin{array}{c} Dh(x^*)\cdot\eta(x^*)\cdot t_1 \\ \vdots \\ Dh(x^*)\cdot\eta(x^*)\cdot t_{2n+1}\end{array}\right\}, \left\{\begin{array}{c} Dh|_{\phi_{t_1}(y^*)}\cdot\int_0^{t_1} D\phi_s^0\cdot\eta(\phi^0_{-s+t_1}(y^*))\,ds \\ \vdots \\ Dh|_{\phi_{t_{2n+1}}(y^*)}\cdot\int_0^{t_{2n+1}} D\phi_s^0\cdot\eta(\phi^0_{-s+t_{2n+1}}(y^*))\,ds\end{array}\right\}.$$

Since $\xi\in\mathscr{E}^2$, $\eta$ can be defined in $x^*$ such that the left column has an arbitrary component with arbitrary value (the other values of the components are forced). Since $\xi\in\mathscr{E}^3$, $Dh|_{\phi_{t_i}(y^*)}=0$ for at most one particular $i\in\{1,\cdots,2n+1\}$, say for $i=m$. Assume $Dh|_{\phi_{t_m}(y^*)}=0$. By constructing $\eta$ on the range of the integral curve $y^*$ to $\phi_{t_{2n+1}}(y^*)$, as in Case 1, and taking $Dh(x^*)\cdot\eta(x^*)t_m$ appropriately, it is shown that the evaluation mapping is transversal.

*Case* 3 Both $x^*$ and $y^*$ are equilibrium points of $\xi^*$.

Since $\xi\in\mathscr{E}^1$, this case need not be considered because

$$h\circ\phi_{t_i}(x^*)=h(x^*)\neq h(y^*)=h\circ\phi_{t_i}(y^*), \qquad i=1,\cdots,2n+1.$$

Hence $\operatorname{ev}(\xi^*, \xi^*, x^*, y^*)\notin\Delta(\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1})$.

For the openness part we need only refer to the openness-of-transversality theorem [4]. $\square$

*Remark.* The theorem is valid for $h\in\mathscr{D}$ and thus for functions belonging to an open and dense subset of $\mathscr{C}^r(X,\mathbb{R})$.

**4. Counterexample.** In this section we provide a counterexample to speculations that $2n$ or fewer samples might suffice to achieve observability generically. The counterexample will be constructed on $S^1$.

Given two arbitrary sample instants 0 and $t_1$ (without loss of generality, 0 is taken as the first sample instant) we will construct a vectorfield on $S^1$ with flow $\phi_t(x)$ and an output mapping $h$ from $S^1$ to $\mathbb{R}$ such that the mapping from $S^1$ to $\mathbb{R}^2$ defined by $x \to (h(x), h \circ \phi_{t_1}(x))(*)$ is not injective and such that small perturbations in the flow and the output mapping do not destroy the noninjectiveness. Notice first that there exists a smooth mapping from $S^1$ to $\mathbb{R}^2$ which twists the circle into a figure eight. The idea of the counterexample is to find a vectorfield and an output mapping which does something similar, i.e., such that the mapping (*) maps $S^1$ into a figure with transversal self-intersections.

Define a vectorfield on $S^1$ (with coordinate $\theta$ denoting the angle), by $\dot{\theta} = \pi/t_1$. The output is defined on $S^1$ considered as the closed interval [0, 1] with 0 and 1 identified. It is given by the following function, $C^\infty$-smoothed at the discontinuities

$$\begin{aligned} h &= 1, \quad 0<t<\tfrac{1}{4}, \quad \tfrac{3}{4}<t<1, \\ h &= 2, \quad \tfrac{1}{4}<t<\tfrac{3}{8}, \quad \tfrac{5}{8}<t<\tfrac{3}{4}, \\ h &= 0, \quad \tfrac{3}{8}<t<\tfrac{5}{8}. \end{aligned}$$

The mapping (*) with $\phi$ and $h$ as defined is not injective, since it maps $S^1$ into a figure with transversal crossings. The crossings are stable and thus noninjectiveness is preserved under perturbations of the vectorfield, the output mapping, and the sample times. Therefore Theorems 1 and 2 are no longer valid for sample programs consisting of $2n$ samples.

## 5. Concluding remarks.

1. In the previous theorems we have been confined to compact manifolds. The results can be extended to noncompact manifolds if "open and dense" is replaced by "residual", i.e., a countable intersection of open and dense sets, with respect to the Whitney topology. The procedure by which the extension is carried out is standard. For more details one is referred to [5].

2. This remark is related to the definition of observability. As defined in § 2 observability or $P$-observability is expressed in set theoretic terms. It is natural, when considering differentiable systems, to require not just one-to-one-ness of the relevant mapping in the definition of observability, but also some differentiability condition on this functional relationship. As far as $P$-observability is concerned, we propose to call a system $(\xi, h)$ $P$-observable over $[0, T]$ if and only if the mapping $x \to (h \circ \phi_{t_1}(x), \cdots, h \circ \phi_{t_{2n+1}}(x))$ *embeds* $X$ into $\mathbb{R}^{2n+1}$.

This definition also takes into account remarks put forward by Kalman and Sontag [6], to the effect that the inverse relation (from output curve to initial condition) should be differentiable.

Theorems 1 and 2 remain true with this stronger notion of observability. Indeed, transversality theory again provides the right framework. A proof amounts to producing the right evaluation mapping (with a codomain involving jet spaces) and then going through steps similar to those in this paper.

3. We have shown that given $2n+1$ time instants $l_i$ with $0<l_1<l_2<\cdots<l_{2n+1}$, the mapping $(h \circ \phi_{l_1}, \cdots, h \circ \phi_{l_{2n+1}})$ generically embeds $X$ into $\mathbb{R}^{2n+1}$. Notice that a state trajectory $\phi_t(p)$, $t \geqq 0$, starting in $p \in X$ at time zero is mapped into $(h \circ \phi_{t+l_1}(p), \cdots, h \circ \phi_{t+l_{2n+1}}(p))$, $t>0$. Therefore the $\omega$-limit set associated with $p$ is homeomorphic with the limit set of the curve $t \to (h \circ \phi_{t+l_1}(p), \cdots, h \circ \phi_{t+l_{2n+1}}(p))$. $x(t,p)$

Assume that the state dynamics has a global attractor; then by studying the limit set of $t \to (h \circ \phi_{t+l_1}(p), \cdots, h \circ \phi_{t+l_{2n+1}}(p))$, which is available from experiment, one obtains 'ormation on the global attractor. F. Takens, in a forthcoming paper, constructs gorithms based on experimental evidence which provide estimates of the dimension ıd the topological entropy of the attractor. This gives, in principle, a strategy for sting the Ruelle–Takens picture [7], where the onset of turbulence is caused by the esence of strange attractors.

4. Given $h \in \mathscr{D}$, and also a parametrized set of vectorfields of $\mathscr{E}$, i.e., a smooth path efined on $[0, 1]$ (or a higher dimensional interval) into $\mathscr{E}$, a similar argument to that in heorem 2 implies that for an open and dense set of paths, the vectorfield correspond- ıg with an arbitrary but a priori given point in $[0, 1]$ and the function $h$ are an bservable pair. Translated to a control context, this means that for a given output ınction and a control system, any constant control almost always distinguishes pairs of oints on the manifold. Thus, in a sense, the question whether couples of points are istinguishable or not with respect to an observed dynamical system has no bearing pon the controls available in the dynamics of the system.

**Acknowledgments.** The author is grateful to D. L. Elliott for numerous discussions nd for raising the questions which lead to this paper. The author also acknowledges everal interesting comments made by an anonymous referee.

## REFERENCES

1] E. B. LEE AND L. MARKUS, *Foundations of Optimal Control Theory*, John Wiley, New York, 1967.
2] E. SONTAG, *On the observability of polynomial systems*, I: *Finite-time problems*, this Journal, 17 (1979), pp. 139–151.
3] H. SUSSMANN, *Generic single-input observability of continuous-time polynomial systems*, IEEE Trans. Automat. Control, to appear.
4] R. ABRAHAM AND J. W. ROBBIN, *Transversal Mappings and Flows*, W. A. Benjamin, New York, 1967.
5] M. M. PEIXOTO, *On an approximation theorem of Kupka and Smale*, J. Differential Equations, 3 (1966), pp. 214–227.
6] E. SONTAG, *Polynomial Response Maps*, Springer-Verlag, New York, 1979.
[7] D. RUELLE AND F. TAKENS, *On the nature of turbulence*, Comm. Math. Phys., 20 (1971), pp. 167–192; 23 (1971), pp. 343–344.