*
*
*
*
ORBIT THEOREMS AND SAMPLING*

Eduardo D. Sontag**
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903, U.S.A.

ABSTRACT

This paper proposes a notion of smooth action on a manifold, and establishes a general integrability result for certain associated distributions. As corollaries, various classical and new results on manifold structures of orbits are established, and the main theorem on preservation of transitivity under sampling is shown to be a simple consequence.

1. Introduction

One of the basic results in control theory, due independently to [SU1] and [ST], states that, for continuous time systems, each orbit (set accessible with positive- and negative-time motions from a given starting state) has a natural structure of immersed submanifold of the state space. This structure is obtained, roughly, as follows. Given any piecewise constant control steering a state into the state ξ, this control having switches at times $t_1, ..., t_k$, tangent vectors to the orbit at ξ are obtained by taking perturbations of the t_i. (More precisely, positive- and negative-time controlled motions are used.) When phrased in terms of the integrability of an associated distribution, this generalizes classical theorems of Frobenius and Chow.

Discrete-time control systems have been studied much less than their continuous counterparts, and their properties diverge considerably from those of the latter, due mainly to the possibility of singularities; see for instance [SO]. The paper [JA] introduced the idea of studying *invertible* discrete nonlinear systems, and developed a realization theory which parallels much of the continuous time situation; further work along these lines was carried out in [FN], [NC], [SS], and related papers. Invertible systems are those for which transition maps, (one for each fixed control,) are all (local) diffeomorphisms. Invertibility is of course a priori an extremely strong assumption in the context of general discrete time systems. However, for systems that result from the *sampling* of continuous time systems, this assumption is always satisfied. For invertible discrete-time systems, it is possible to give a close analogue of the continuous time orbit theorem. Since times are discrete, it is of course not possible to take time derivatives as above. Instead, one substitutes derivations with respect to the values of the controls in each interval. (The underlying assumption being that there is some sort of manifold structure on the control value set. Precise details are given later.)

---

*To appear in *Algebraic And Geometric Methods in Nonlinear Control Theory*, M.Fliess and M.Hazewinkel (Eds.), Reidel, Dordrecht, 1986.

**Research supported in part by US Air Force Grant 85-0247

The first half of this paper introduces a framework that allows to prove an abstract orbit theorem, for "smooth actions on manifolds". This yields as consequences the above mentioned discrete-time and continuous time results. More interestingly, the theorem will also imply a number of other results, including characterizations of "zero-time" orbits of various different types, and an alternative submanifold structure in the continuous time case (the "input-topology" structure). The latter will be compared with the more classical "time-topology" structure. Certain facts that would appear to be obvious, for instance the second countability of zero-time orbits (but not of arbitrary orbits), turn out to require careful proofs. It should be pointed out that there have been many other proofs of "orbit theorems" in the literature, at various levels of generality (see e.g. [KL], [SJ], [KS]). All proofs are in principle based on the same ideas. In fact, the present approach is based on the proof in the conference paper [SS], which was in turn motivated by a general (unpublished) abstract result due to H. Sussmann, which was in turn a generalization of the proof in [SU1]. We believe that the present result strikes the right balance between generality (it appears to imply all others) and level of abstraction (it can be applied immediately to particular classes of actions), and our main contribution here in that respect is in exposition.

The second part of the paper concentrates on sampling. When a continuous time system is regulated by a digital computer, control decisions are often restricted to be taken at fixed times $0,\delta,2\delta,...$; one calls $\delta>0$ the *sampling time*. 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 [KHN], which studied the case of linear systems; more recent references are [BL], [GH]. For nonlinear systems, it appears that the problem had not been studied systematically until the paper [SS] and later conference papers by the author. As usual for nonlinear systems, it is easier to study transitivity (controllability with positive- and negative- time motions) than controllability (but, see [SO1] for various controllability results). The main result is that, for fast enough sampling, transitivity is preserved provided that the original system be "strongly" transitive in a sense to be made precise later. The proof in [SS] is based on a fixed point theorem. The same result is proved here using more elementary tools, as an almost trivial consequence of the interplay of the time- and input- topologies (see above discussion) on continuous time systems. This is probably the most natural way to understand the sampling results. For expository purposes, we have also included here a few topics that had already been covered in the above mentioned conference papers, including a more or less careful treatment of one-dimensional systems, which provide a good source of examples and counterexamples.

Acknowledgment. I wish to thank Héctor Sussmann for many discussions and suggestions during the course of this research. Part of the results presented here were initially given in our joint presentation [SS].

2. Preliminaries

We first give some differential-geometric terminology which generalizes that in the standard literature. Most is as in [SU1], [KR] and [IS]. All manifolds will be smooth (C^∞) and paracompact (hence, Hausdorff and each component is second countable). Let **M** be an arbitrary such manifold. A *submanifold* N of **M** is an immersed (not necessarily regular) submanifold.

By a *vector field* X *on* **M** we shall mean a smooth vector field (smooth section of the tangent bundle) defined on an open subset V_X of **M**; we denote by $\Xi(\mathbf{M})$ the set of all such X. Given X and Y in $\Xi(\mathbf{M})$, let $V_X \cap V_Y = V$. We then define the Lie bracket [X,Y] as the Lie bracket of X restricted to V and Y restricted to V. If V is empty, the bracket is undefined. Similarly for the sum of X and Y, and products by constants. This makes $\Xi(\mathbf{M})$ into a "pseudo-" Lie algebra (or, a "sheaf of Lie algebras"); for simplicity we shall take in this paper the term "Lie algebra" to imply only partially defined operations. Similarly, we let Diff(M) denote the set of local diffeomorphisms on **M**, with compositions only partially defined. As a general rule, if the Lie bracket of two vector fields (or the composition of two local diffeomorphisms) appears in a statement, that statement should be taken to mean "if this composition is defined, then...".

We denote by $X(\xi)$, instead of X_ξ, the value of $X \in \Xi(\mathbf{M})$ at $\xi \in \mathbf{M}$. A (possibly singular) *distribution D* on **M** is a subset of the tangent bundle T**M** with the property that

$$D(\xi) := \{v \in T_\xi\mathbf{M} \mid (\xi,v) \in D\}$$

is a subspace for each ξ. (So D is a choice of a subspace in the tangent space at each ξ in **M**.) The vector field X *belongs to D* if $X(\xi) \in D(\xi)$ for all $\xi \in V_X$. The set of vector fields belonging to D is denoted by $\{D\}_{vf}$. A subset $\Phi \subseteq \Xi(\mathbf{M})$ is *everywhere defined* if the union of the domains V_X, $X \in \Phi$, is all of **M**. For an everywhere defined set Φ, $\{\Phi\}_D$ denotes the distribution generated by Φ, i.e. the smallest distribution D for which all $X \in \Phi$ belong to D. Thus, $\{\Phi\}_D(\xi)$ is for each ξ the span of the set of vectors $\{X(\xi)$ s.t. X is defined at $\xi\}$. A distribution of the form $\{\Phi\}_D$ is called a *smooth* distribution.

For the rest of this paper, the term distribution will always mean **smooth** *distribution*.

The *rank* of D at ξ is the dimension of $D(\xi)$. Thus the constant rank case corresponds to the usual notion of (nonsingular) distribution in differential geometry. An *integral manifold* N of D is a submanifold of **M** such that $T_\xi N = D(\xi)$ for each ξ in N. An integral manifold N of D is *maximal* if it is connected and for every other integral manifold N′ of D intersecting N, N′ is an open submanifold of N. The distribution D is *integrable* iff it induces a (singular) foliation, i.e., there is a partition of **M** into maximal integral manifolds, the *leaves* of D.

The subset Φ of $\Xi(\mathbf{M})$ is *involutive* if [X,Y] is in Φ whenever (the product is defined and) X, Y are in Φ. We shall say that Φ is a *subspace* of $\Xi(\mathbf{M})$ is $rX \in \Phi$ whenever $X \in \Phi$ and $r \in \Re$, and X+Y is in Φ whenever (the sum is defined and) X, Y are in Φ. ("Presheaf of linear spaces" is probably a better terminology.) The smallest subspace containing Φ is denoted by $\{\Phi\}_{sp}$ (the "linear space generated by Φ"). The smallest involutive subspace containing Φ is denoted by $\{\Phi\}_L$ (the "Lie algebra generated by Φ"). Finally, the distribution $(\{\Phi\}_L)_D$ is the *Lie distribution generated by* Φ, and is denoted by $\{\Phi\}_{LD}$.

A (smooth) distribution D is *involutive* if $\{D\}_{vf}$ is involutive. Integrable distributions are involutive (because the vector fields belonging to D are tangent to the leaves of D, which are submanifolds).

We shall say that the subset **C** of the manifold **U** has *nice boundary* if the following property holds: for each $u \in \mathbf{C}$ there is a smooth curve

$$\gamma : [0,1] \to \mathbf{U}$$

such that $\gamma(0) = u$ and $\gamma(t)$ is in int**C** (interior of **C** with respect to **U**) for all $t>0$. (Smooth on a closed interval means smooth in a neighborhood of this interval.)

2.1. Actions

An *action* Σ is an 8-tuple ($\mathbf{M}$,$\mathbf{A}$,-,$\{\iota_a,a\in \mathbf{A}\}$,$\{\mathbf{U}_a,a\in \mathbf{A}\}$,$\{\mathbf{C}_a,a\in \mathbf{A}\}$,$\{D_a,a\in \mathbf{A}\}$,$\{g_a,a\in \mathbf{A}\}$), where:

1. $\mathbf{M}$ is a manifold (the *state space*),
2. $\mathbf{A}$ is a set,
3. - : $\mathbf{A} \to \mathbf{A}$ is a map of order 2: $-(-a)=a$,
4. For each $a\in \mathbf{A}$, $\mathbf{U}_a$ is a manifold,
5. For each $a\in \mathbf{A}$, D_a is an open subset of $\mathbf{M}\times\mathbf{U}_a$,
6. For each $a\in \mathbf{A}$, $\mathbf{C}_a$ is a subset of $\mathbf{U}_a$ with nice boundary,
7. For each $x\in \mathbf{M}$, there is some $a\in \mathbf{A}$ and some $u\in \mathbf{C}_a$ such that $(x,u)\in D_a$,
8. For each $a\in \mathbf{A}$, ι_a is a smooth map $\mathbf{U}_a \to \mathbf{U}_{-a}$ such that the composition $\iota_a\circ\iota_{-a}$ is the identity (the subscript is dropped, and $\iota_a(u)$ is written simply as $\bar{u}$, when a is clear from the context) and such that ι_a maps $\mathbf{C}_a$ into $\mathbf{C}_{-a}$ and the interior int($\mathbf{C}_a$) into int($\mathbf{C}_{-a}$), and
9. For each $a\in \mathbf{A}$, g_a is a smooth map $D_a \to \mathbf{M}$, such that:
   - $(g_a(x,u),\bar{u})$ is in D_{-a} if (x,u) is in D_a and u is in $\mathbf{C}_a$, and
   - $g_{-a}(g_a(x,u),\bar{u}) = x$ for all such (x,u).

   (So, in particular, each map $g_a(\cdot,u)$ is a local diffeomorphism.)

Various examples will be given later, including different manners of modeling continuous time systems and discrete systems. In all these examples, the sets U_a satisfy that U_a=U_{-a} and the maps ι_a are always the identity. The above definition is much easier to read in that case. However, we need the present more general definition for technical reasons, since we shall introduce various actions associated to a given action, for which the induced ι_a in general will not be identities, even if they are so for the original action.

For the rest of this section, we fix an action Σ as above. Note that, if $\mathbf{A}'$ is a subset of $\mathbf{A}$ which is invariant under "-" and satisfies property (7) then there is a *restriction* of Σ to $\mathbf{A}'$, obtained by restricting the index sets in the 7-tuple corresponding to Σ. Other ways of deriving new actions from a given one will be described below.

We let $\mathbf{A}^*$ be the free monoid on $\mathbf{A}$, that is, the set of all possible sequences ("words") of elements of $\mathbf{A}$, and identify $\mathbf{A}$ with the subset of $\mathbf{A}^*$ consisting of sequences of length 1. For any $b = (a_1,...,a_r)$ in $\mathbf{A}^*$, $-b$ is the sequence $(-a_r,...,-a_1)$, and $\mathbf{U}_b$ (respectively, $\mathbf{C}_b$) is the product of the corresponding $\mathbf{U}_a$ (respectively, $\mathbf{C}_a$), $a=a_i$. We also let g_b: $\mathbf{M}\times\mathbf{U}_b \to \mathbf{M}$ be obtained by composition. More precisely, for the empty word b, $\mathbf{U}_b = \mathbf{C}_b$ is a one-point set and g_b is the identity, and in general the sets D_b and the maps g_b are defined inductively on the length of b as follows:

$$(x,u\omega)\in D_{ab} \text{ iff } (x,u)\in D_a \text{ and } (g_a(x,u),\omega)\in D_b,$$

for u in $\mathbf{U}_a$ and ω in $\mathbf{U}_b$, and then

$$g_{ab}(x,u\omega) := g_b(g_a(x,u),\omega).$$

When a or b are clear from the context, we omit the corresponding subscripts. A concatenation notation is alternatively used to exhibit sequences in $\mathbf{U}_b$, as in "$u\omega$" above, and similarly for words in $\mathbf{A}^*$. Further, the letter "a", possibly subscripted or primed, will always denote an element of $\mathbf{A}$, while notations involving b or c stand for words in $\mathbf{A}^*$. Similarly, latin letters u,v,... will be used for elements of the sets $\mathbf{U}_a$ and greek letters ψ, χ, ω for elements of sets of the form $\mathbf{U}_b$ (that is, sequences of elements of the various $\mathbf{U}_a$). For $\omega = (u_1,\cdots,u_r)$ in $\mathbf{U}_b$, let $\bar{\omega} := (\bar{u}_r\cdots\bar{u}_1)$, an element of $\mathbf{U}_{-b}$ (note the reversed order). Then $(g_b(x,\omega),\bar{\omega})$ is in D_{-b} whenever (x,ω) is in D_b and

ω is in $\mathbf{C}_b$, and

$$g_{-b}(g_b(x,\omega),\bar{\omega}) = x \,.$$

This proves that the data **M**, $\mathbf{A}^*$, $\mathbf{U}_b$, etc., defines a new action Σ^* (taking a suitable product path, one can prove that $\mathbf{C}_b$ has nice boundary, if each $\mathbf{C}_a$ does). The main objects of study are the *orbits* of the action Σ:

$$O(x) := \{z \mid g_b(x,\omega) = z, \text{ some } b,\omega\}.$$

Note that the orbits of Σ^* coincide with the orbits of Σ. Later we shall introduce various other actions derived by restricting the set $\mathbf{A}^*$; these actions will be very valuable in studying "zero-time" accessibility and related notions.

2.2. The distribution associated to an action

Given the action Σ, certain induced vector fields will play a central role. These are defined, intuitively, as follows. Assume that $g_a(y,u) = x$. Then, if $u+\delta u$ is a perturbation of u, the state $g_a(y,u+\delta u)$ is close to x. As $\delta u \to 0$, a tangent vector at x results. If x' is another state close to x, the fact that $g_a(\cdot,u)$ is a local diffeomorphism implies that there is a y' close to y such that $g_a(y',u) = x'$. Applying the perturbation argument with the same δu, there results a vector at x'. This construction is smooth in x, and a vector field is obtained. We now make all this precise. The "tangent bundle of $\mathbf{C}_a$", for $a\in\mathbf{A}$, (where $\mathbf{C}_a$ is not necessarily a manifold,) is defined as

$$T\mathbf{C}_a := \{(u,\nu) \mid u\in\mathbf{C}_a \text{ and } \nu\in T_u\mathbf{U}_a\} \,.$$

For each (a,u,ν) with $a\in\mathbf{A}$ and $(u,\nu)\in T\mathbf{C}_a$ we define a vector field $X_{a,u,\nu}$ as follows. For any ξ such that $(\xi,\bar{u})$ is in D_{-a},

$$X_{a,u,\nu}(\xi) := \frac{\partial g_a(g_{-a}(\xi,\bar{u}),v)}{\partial v}\Big|_{v:=u}(\nu). \tag{2.1}$$

Note that $g_a(g_{-a}(\xi,\bar{u}),v)$ is the same as what can be denoted, using the word $(-a,a)\in\mathbf{A}^*$, by $g_{-a,a}(\xi,\bar{u}v)$. Equation (2.1) provides a vector at ξ. It is clear how to compute it in local coordinates (product of a Jacobian matrix by a vector). A coordinate-free interpretation is as follows. Let α be the map that sends $v\in\mathbf{U}$ into $g_a(g_{-a}(\xi,\bar{u}),v)$; this is defined in some neighborhood of u in $\mathbf{U}_a$. For any real-valued smooth map f defined in a neighborhood of x, consider the composition $\beta := f\circ\alpha$. Then $X_{a,u,\nu}(\xi)(f)$ is the evaluation $\nu(\beta)$, where we are interpreting ν as a differential operator on germs of functions at u. With this definition, it is clear that $X_{a,u,\nu}(\xi)$ is again a differential operator, and so defines a vector at ξ. From the coordinate description it follows that this is not only smooth on ξ but in fact smooth as a function of (ξ,u,ν).

If π: $\mathbf{M}\to\mathbf{M}$ is a local diffeomorphism, we denote by Ad_π the (partial) linear operator $\Xi(\mathbf{M})\to\Xi(\mathbf{M})$ corresponding to conjugation by π, more precisely:

$$Ad_\pi X(\xi) := (\pi^{-1})_*(X(\pi(\xi)))$$

for X in $\Xi(\mathbf{M})$ and ξ in **M**. (In this equation, $(\pi^{-1})_*$ denotes the differential of π^{-1} at the point $\pi(\xi)$.) Note that the domain of the vector field $Ad_\pi X$ is

$$\{\xi \mid \pi(\xi) \text{ is defined and is in } V_X\} \,.$$

If this domain is empty, $Ad_\pi X$ is undefined. If π is the (local) diffeomorphism $g_b(\cdot,\omega)$, we denote

$$Ad_\pi = Ad_{b,\omega} \,.$$

If Γ is a (pseudo-) group of (local) diffeomorphisms on **M**, and Φ is an everywhere defined subset of Ξ(**M**), we introduce the distribution

$$\mathrm{Ad}_\Gamma\Phi := \{\mathrm{Ad}_\pi X,\ \pi\in\Gamma \text{ and } X\in\Phi\}_D\ . \qquad (2.2)$$

In particular, let $\Gamma(\Sigma)$ be the group consisting of all the $g_b(\cdot,\omega)$, for $b\in\mathbf{A}^*$ and $w\in\mathbf{C}_b$, for the given action Σ, and let $\Phi(\Sigma)$ be the set of all vector fields $X_{a,u,\nu}$, for (a,u,ν) with $a\in\mathbf{A}$ and $(u,\nu)\in T\mathbf{C}_a$. We introduce the *distribution associated to* Σ,

$$\mathbf{D} = \mathbf{D}(\Sigma) := \mathrm{Ad}_{\Gamma(\Sigma)}\Phi(\Sigma)\ .$$

This is an everywhere defined distribution: for each $\xi\in\mathbf{M}$ there is by property (7) in the definition of action a pair (a,u) such that $g_{-a}(x,u)$ is defined, so $X_{a,u,\nu}$ is defined for all ν tangent at u. We consider also another distribution, the *Lie distribution associated to* Σ,

$$\mathbf{D}_L = \mathbf{D}_L(\Sigma) := \{\Phi(\Sigma)\}_{LD}\ .$$

The *rank of* Σ *at* ξ is by definition the rank of **D** at ξ; the *Lie rank of* Σ *at* ξ is the rank of $\mathbf{D}_L$.

2.3. Example: continuous time systems with time-topology

Given an arbitrary everywhere defined set of vector fields Φ, we may consider the (pseudo-) group generated by Φ,

$$\exp(\Phi) := \{\exp(tX),\ t\in\Re,\ X\in\Phi\}\ .$$

Here $\exp(tX)(\xi)$, if defined, is the solution at time t of the differential equation

$$\dot{x}(t) = X(x(t)),\ x(0) = \xi.$$

The well-posedness theorem for ode's insures that $\exp(tX)(\xi)$ is defined for an open set of pairs (t,ξ) (which depends on X). We may thus introduce the following action $\Sigma(\Phi)$, the *action associated to* Φ:

- $\mathbf{A} := \{1,-1\}\times\Phi$, with $-(\varepsilon,X) := (-\varepsilon,X)$
- $\mathbf{U}_a := \Re$ and $\mathbf{C}_a = \mathbf{U}_a$ for all $a\in\mathbf{A}$, with ι_a=identity
- $g_a(\xi,t) := \exp(\varepsilon tX)(\xi)$, if $a = (\varepsilon,X)$
- D_a is the domain of definition of g_a .

One can then consider the distribution associated to the action $\Sigma(\Phi)$. We call this distribution $\mathrm{Ad}(\Phi,\Phi)$. Let $a = (\varepsilon,X)$, and pick any u in $\Re$ and any ν in $T_u\Re$, the latter identified again to $\Re$. Then,

$$X_{a,u,\nu} = \varepsilon X\ .$$

It follows that $\{\Phi\}_D = \{\Phi(\Sigma)\}_D$, so:

$$\mathrm{Ad}(\Phi,\Phi) = \mathbf{D}(\Sigma(\Phi)) = \mathrm{Ad}_{\exp(\Phi)}(\Phi)\ .$$

Actions obtained in this way will be also refered to as *continuous time actions with time topology*, for reasons that will become clear later. They will be one of the two main types of actions to be associated to continuous time systems.

The classical orbit theorem for continuous time sytems, due independently to Stefan and Sussmann, says that $\mathrm{Ad}(\Phi,\Phi)$ is integrable, and that the orbits of $\Sigma(\Phi)$ are the leaves of this distribution. These statements will be proved later as particular consequences of the general orbit theorem (whose proof is in itself essentially that in [SU1]).

Finally, if *D* is a distribution, we let exp(*D*) be by definition exp(Φ), where $\Phi = \exp(\{D\}_{vf})$, and define Ad(*D*,*D*) as Ad(Φ,Φ) for this set Φ. The results in [SU1] also prove that, if Φ is an everywhere defined set of vector fields and if $D := \{\Phi\}_D$, then *D is integrable if and only if D is invariant*, meaning that

$$\mathrm{Ad}_{\exp(\Phi)}(D) = D\ .$$

and that a smooth distribution *is integrable iff it has the "integral manifolds property": for each* ξ *in* **M** *there is an integral manifold of D which contains* ξ*.*

2.4. Example: discrete time actions.

A *discrete-time action* is one for which:

- **A** has two elements {1,-1} and -a is (-1)a
- $\mathbf{U}_{-1} = \mathbf{U}_1 = \mathbf{U}$ is a second countable manifold, and ι is the identity
- $\mathbf{C}_{-1} = \mathbf{C}_1 = \mathbf{C}$, called the *control constraint set*.

We also denote the elements of **A** as {+,-}. These actions are associated to (invertible) discrete-time systems, to be introduced later.

2.5. Some useful formulas

The following formula will be used later; it is valid for all a∈**A**, b∈$\mathbf{A}^*$, w∈$\mathbf{C}_b$, (u,v) in $T\mathbf{C}_a$ (more precisely, if either side is defined, then the other one is defined too, and they coincide):

(2.3)

$$Ad_{b,\omega}(X_{a,u,\nu}) = \frac{\partial g_{b,-a,a,-b}(x,\omega u v\varpi)}{\partial v}\Big|_{v:=u}(\nu).$$

In general, if c∈$\mathbf{A}^*$ and ψ is in $\mathbf{C}_c$, we shall be interested in partial derivatives of $g_c(\cdot,\psi)$ with respect to components of ψ. Assume that c = b′bb″ is a factorization of c into subwords, and let ψ = χ′ωχ″ be the corresponding factorization of ψ. Then, (assuming that $x = g_c(y,\psi)$ is defined,)

$$d_b g_c(y,\psi) : T_\omega \mathbf{U}_b \to T_x \mathbf{M} \qquad (2.4)$$

is by definition the differential of $g_c(y,\chi'(\cdot)\chi'')$ at the point ω of $\mathbf{U}_b$. When b′ and b″ are empty (c=b), we often omit the subscript and write just $dg_b(y,\omega)$ or even dg(y,ω). Differentials with respect to x will be written d_x or using the $(\cdot)_*$ notation; that is, $d_x g_c(x,\omega)$ is the same as $(g_c(\cdot,\omega))_*$ evaluated at x.

Assume that the above c factors as (b′,a,-b), with a∈**A**, and consider the corresponding factorization ψ = (ω′,u,ϖ). If ν is in $T_u\mathbf{C}_a$, we can evaluate the differential in equation (2.4) at ν; there results the following formula:

$$[d_a g_c(y,\psi)](\nu) = \mathrm{Ad}_{b,\omega}(X_{a,u,\nu})\,(x)\ . \qquad (2.5)$$

Given any x∈**M** and any vector field of the type $\mathrm{Ad}_{b,\omega}(X_{a,u,\nu})$ defined at x, there is a y such that $g_c(y,\psi) = x$: just let c:= (a,-b) and ψ:= uϖ. We conclude that **D**(x) *equals the span of all the images of the maps as in equation* (2.4). (In fact, it will follow from later discussion that it equals in fact the image of just one such map.)

Consider the action Σ^*, introduced earlier, derived from the original action Σ by considering the free monoid $\mathbf{A}^*$ on **A**. This gives rise to a distribution $\mathbf{D}(\Sigma^*)$. The following result is trivial, but will be useful later:

Lemma 2.1: $\mathbf{D}(\Sigma) = \mathbf{D}(\Sigma^*)$.

Proof: By definition, $\mathbf{D}(\Sigma)$ is included in $\mathbf{D}(\Sigma^*)$, so need only prove the reverse inclusion. Since $\Gamma(\Sigma) = \Gamma(\Sigma^*)$, it is enough to show that $\Phi(\Sigma^*) \subseteq \mathbf{D}(\Sigma)$. Pick any $c \in \mathbf{A}^*$ and any χ in $\mathbf{U}_c$. The tangent space to $\mathbf{U}_c$ at χ is the direct sum of the tangent spaces to all $\mathbf{U}_a$, for the $a \in \mathbf{A}$ that appear in the factorization of c. Thus it is enough in generating $\Phi(\Sigma^*)$ to consider the following situation: c factorizes as (c′,a,-b), with corresponding factorization $\chi = \chi' u \bar{\omega}'$, ν is a tangent vector at u, and the generator is $X_{c,\chi,\nu}$. Let b′ be the word (b,-a,-c′,c′), and ω' be $(\omega,\bar{u},\bar{\chi}',\chi')$. Then formula (2.5) establishes the result.n

In formula (2.1) we could consider derivatives with respect to $\bar{u}$ instead of v. New vector fields are obtained in this way, namely

$$Y_{a,u,\nu} := \frac{\partial g_a(g_{-a}(\xi,\bar{v}),u)}{\partial v} \Big|_{v:=u}(\nu). \tag{2.6}$$

These vector fields will appear later. Computing in local coordinates, two applications of the chain rule show that in fact the same distribution would be obtained from these, because $Y_{a,u,\nu} = -X_{a,u,\nu}$. It is also easy to establish the following formula, for any (a,u,ν) with $a \in \mathbf{A}$ and $(u,\nu) \in T\mathbf{C}_a$.

$$X_{-a,\bar{u},\mu} = -Ad_{a,u}(X_{a,u,\nu}) \, , \tag{2.7}$$

where $\mu = [\partial\bar{u}/\partial u](\nu)$. Thus, in generating $\mathbf{D}$ it is redundant to include the vector field $X_{-a,\bar{u},\nu}$ when $X_{a,u,\nu}$ has been already included.

Finally, consider a word of the form c = (b,a,-b′,-a′,-b″) in $\mathbf{A}^*$ (where a and a′ are in $\mathbf{A}$), and assume that $\mathbf{U}_a = \mathbf{U}_{-a}$. Pick an $u \in \mathbf{C}_a$ and a ν in $T_u C_a$. In the next formula, ψ is a word in $\mathbf{U}_c$ of the type $(\omega,v,\bar{\omega}',\bar{v},\bar{\omega}'')$ (note the same v):

$$\frac{\partial g_{-c,c}(\xi,\bar{\psi}\psi)}{\partial v} \Big|_{v:=u}(\nu) = Ad_{b''a',\omega''u}[Ad_{b',\omega'}(X_{a,u,\nu}) - X_{a',u,\nu}](\xi). \tag{2.8}$$

2.6. Some properties of actions

The main orbit theorem will apply to arbitrary actions; however, various classes of actions have nicer properties. For ease of reference, we collect some definitions of special classes in this section. Again, Σ is a fixed given action.

Definition 2.2: The action Σ is *countable* if the index set $\mathbf{A}$ is countable (or finite) and all the manifolds U_a are second countable.

The main example of countable action is that of discrete time actions (c.f. section 2.4).

Definition 2.3: The action Σ is *connected* if for each ξ in $\mathbf{M}$ and $a \in \mathbf{A}$ there is a u_o in $\text{int}(\mathbf{C}_a)$ such that $g_a(\xi,u_o) = \xi$ and further, for any other $u \in \text{int}(\mathbf{C}_a)$ there is a smooth curve γ: $[0,1] \to \mathbf{M}$ such that $\gamma([0,1]) \subseteq \text{int}(\mathbf{C}_a)$, $\gamma(0)=u_o$, $\gamma(1)=u$, and $(\xi,\gamma(\lambda)) \in D_a$ for all λ.

Connectedness is a very strong restriction, and is the property that makes the continuous time case with time-topology so much better behaved than the general case. (However, we shall introduce a related notion later ('S-connectedness') that will be much less restrictive.) In the time-topology case, $u_o=0$ always satisfies the first property, and given any (t,ξ) such that exp(tX)(ξ) is defined, we may let $\gamma(\lambda) := \lambda t$; this satisfies the second property. For future reference we then state:

Lemma 2.4: Continuous time time-topology actions are connected.

Definition 2.5: The action Σ is *complete* if $D_a = \mathbf{M} \times U_a$ and int($\mathbf{C}_a$) is connected for each $a \in \mathbf{A}$.

Note that continuous time time-topology actions are not complete, unless the original vector fields are complete in the usual sense.

Definition 2.6: The action Σ *is an action with zero* if all the sets $\mathbf{U}_a$ are equal (say to $\mathbf{U}$), all the sets $\mathbf{C}_a$ are equal (to $\mathbf{C}$), and there exists an element 0 in int($\mathbf{C}$), $\bar{0}=0$, with the property that $g_a(\xi,0)=\xi$ for all ξ and all $a \in \mathbf{A}$.

Note that *a complete action with zero is necessarily connected*. Finally, we give two other important possible properties of actions.

Definition 2.7: The distribution *D* had *full rank at* $\xi \in \mathbf{M}$ iff its rank at ξ equals the dimension of $\mathbf{M}$. The *action* Σ has full rank at ξ iff $\mathbf{D}(\Sigma)$ does. Similarly, Σ has *full Lie rank* at ξ iff $\mathbf{D}_L(\Sigma)$ does.

Definition 2.8: A (real-) *analytic* action is one for which all the data is analytic (replace "smooth" by "analytic" in the definition of action).

3. Statement and consequences of the main orbit theorem

Theorem 3.1: Let Σ be any action. Then, the distribution $\mathbf{D} = \mathbf{D}(\Sigma)$ is integrable. Further, for each ξ in $\mathbf{M}$ the orbit $O(\xi)$ has a unique structure of (not necessarily connected or even second countable) submanifold of $\mathbf{M}$ with the property that $O(\xi)$ is an integral manifold of $\mathbf{D}$ and that for each $b \in \mathbf{A}^*$, the restriction of g_b to a map $(O(\xi) \times \text{int}\mathbf{C}_b) \cap D_b \to O(\xi)$ is smooth.

We shall prove this theorem in two steps. First, we shall consider the case in which $\mathbf{C}_a$ is a submanifold for each $a \in \mathbf{A}$. Without loss, we can (and shall) assume in that case that $\mathbf{C}_a = \mathbf{U}_a$ for all a. We call this *the manifold case*. Then we deal with the general case. For the manifold case, we shall establish the following sligthly stronger lemma:

Lemma 3.1: Let Σ be any action. Assume that $\mathbf{U}_a = \mathbf{C}_a$ for all $a \in \mathbf{A}$. Let ξ be in $\mathbf{M}$. Then $O(\xi)$ has a unique structure of submanifold of $\mathbf{M}$ such that

1. for each $b \in \mathbf{A}^*$, the (restricted) map g_b: $(O(\xi) \times \mathbf{U}_b) \cap D_b \to O(\xi)$ is smooth, and
2. for any ζ in $O(\xi)$, the dimension of $O(\xi)$ is equal to

$$r(\xi,\zeta) = \sup \{\text{rank } dg_b(\xi,\omega)\}, \tag{3.1}$$

where the sup is taken over all b and ω such that (ξ,ω) is in D_b and $g_b(\xi,\omega) = \zeta$.

The lemma will be proved in a latter section. We now show that theorem (3.1) follows from it, for the manifold case. *Claim*: the orbits of Σ are integral manifolds of $\mathbf{D}$. By formula (2.5) and part (2) of the lemma (using $\zeta=\xi$), it follows that rank of $\mathbf{D}$ at $\xi \geq \dim O(\xi)$. Further, the generators of $\mathbf{D}$ are, by part (1) of the lemma and by equation (2.5), included in the tangent space to $O(\xi)$ at

ξ. Thus $T_\xi O(\xi) = \mathbf{D}(\xi)$, as claimed. We conclude that the connected components of the possible orbits $O(\xi)$ give rise to the leaves of the integrable manifold **D**, as established as a consequence of the following lemma. This lemma can be used to prove the fact (see section 2.3) that a distribution *D* is integrable iff it has the integral manifolds property.

Lemma 3.2: Assume that the smooth curve $\gamma:(0,1)\to\mathbf{M}$ is such that (a) $\dot\gamma(t)$ is in $\mathbf{D}(\gamma(t))$ for all t, and (b) the rank of **D** is constant along $\gamma(t)$. Then the image of γ is contained in an orbit of Σ.

Proof: We claim that for each $t\in(0,1)$ there is a neighborhood V of t such that $\gamma(V)$ is included in $O(\gamma(t))$. Let $X_1,\cdots,X_r$ be vector fields such that $\{X_1(x),\cdots,X_r(x)\}$ is a basis of $\mathbf{D}(x)$, $x=\gamma(t)$. Because the rank of **D** is constant along this curve, it follows that $\{X_1(y),\cdots,X_r(y)\}$ is a basis of $\mathbf{D}(y)$, $y=\gamma(\tau)$, for τ close to t. Thus there are r smooth real functions $\rho_1,\cdots,\rho_r$, (defined in a neighborhood of t,) such that

$$\dot\gamma(\tau) = \rho_1(\tau)X_1(\gamma(\tau)) + \cdots + \rho_r(\tau)X_r(\gamma(\tau))$$

for all such τ. This can be seen as a controlled differential equation evolving in the manifold O(x). As such, there is a solution γ' of this equation, for τ near t, contained in O(x), and with $\gamma'(t)=x$. But γ' would also be a solution of this equation as an equation evolving in the manifold **M**. By uniqueness of solutions of (controlled) ode's in **M**, it follows that $\gamma = \gamma'$ is indeed contained in O(x) for τ near t. This establishes the claim. For each possible orbit O, this argument shows that $\{t|\gamma(t)\in O\}$ is open. Since the orbits are disjoint, connectedness of (0,1) implies that the range of γ can intersect at most one orbit.n

If N is any connected submanifold of **M** which is an integral manifold of **D**, and if ξ is in N, then N must be contained in $O(\xi)$, and hence in the connected component of $O(\xi)$ at ξ. This is because any y in N can be connected to ξ by a smooth curve, integral for **D**, and the above lemma concludes that $O(y) = O(\xi)$. This concludes the proof of the theorem in the manifold case, assuming lemma 3.1 is known to be true.

In the rest of this section, we show how to prove the theorem in general, assuming it has been proved in the manifold case. We need the following lemmas. Given the action Σ, we may introduce another action Σ_{int}, obtained when replacing $\mathbf{C}_a$ by $int(\mathbf{C}_a)$ for each $a\in\mathbf{A}$. All axioms are again satisfied. Let $O_{int}(x)$ be the orbit of x under this action, and let $\mathbf{D}_{int}$ be the corresponding distribution. Note that the theorem then applies to $\mathbf{D}_{int}$ (manifold case).

Lemma 3.3: $\mathbf{D}_{int} = \mathbf{D}$

Proof: Since $\mathbf{D}_{int}(\xi)$ is contained in $\mathbf{D}(\xi)$ for all ξ, it is sufficient to show that they have the same dimension. Since every $\mathbf{C}_a$ has nice boundary, $\mathbf{C}_a$ is always in the closure of its interior. Thus the generators $X_{a,u,\nu}$, for each (a,u,ν) with $a\in\mathbf{A}$ and $(u,\nu)\in T\mathbf{C}_a$, as well as the conjugates under the maps $Ad_{b,\omega}$, can be approximated by similar generators corresponding to the action Σ_{int}. In particular, any basis of $\mathbf{D}(\xi)$ can be approximated by elements of $\mathbf{D}_{int}$.n

Lemma 3.4: $O_{int}(\xi) = O(\xi)$ for all ξ.

Proof: Let $\zeta = g_b(\xi,\omega)$, with $b = a_1\cdots a_r$ and $\omega = u_1\cdots u_r$. For each i, let $\mathbf{C}_i := \mathbf{C}_{a_i}$, and let γ_i be a smooth curve $[0,1]\to\mathbf{M}$ such that $\gamma_i(t)$ is in $int(\mathbf{C}_i)$ for all t>0 and $\gamma_i(0) = u_i$. We now consider the following action Σ'. Its index set $\mathbf{A}'$ has just two elements {1,-1}. The manifolds $\mathbf{U}_1' = \mathbf{U}_{-1}'$ are both equal to (-1,1), and $\mathbf{C}_1' = \mathbf{C}_{-1}'$ is the union $(-1,0)\cup(0,1)$. The maps ι are the identity, and (with domains induced from the original action),

$g_1'(x,\lambda) := g_b(x,\gamma_1(\lambda^2)\cdots\gamma_r(\lambda^2))$
$g_{-1}'(x,\lambda) := g_{-b}(x,\gamma_r(\lambda^2)\cdots\gamma_1(\lambda^2))$

Let O′ denote orbits with respect to Σ′. Note that, from the choice of the curves γ_i, it follows that $O'(x) \subseteq O_{int}(x)$ for all x. Finally, consider the action Σ″ obtained by using instead $\mathbf{C}_1 = \mathbf{C}_{-1} = (-1,1)$ in the above description. We know that $\zeta\in O''(\xi)$; the result will follow if we can prove that ζ is also in O′(ξ). The theorem can be assumed true for both Σ′ and Σ″ (manifold case). Furthermore, a density argument as used in the previous lemma shows that the corresponding distributions **D**′ and **D**″ are equal. Thus the connected component of O′(ζ) at ζ is the same submanifold of **M** as the connected component of O″(ζ) at ζ (since it is the leaf of **D**′ through ζ). Thus there is a subset N of **M** which is a neighborhood of ζ both in the topologies of O′(ζ) and O″(ζ). For λ near 0, $g_1'(\xi,\lambda)$ is in N, by continuity of g_1' in λ for the topology of O″(ξ). (Property (2) in lemma (3.1).) Thus there exists a $\lambda \neq 0$ such that $y = g_1'(\xi,\lambda)$ is in O′(ζ); since $y\in O'(\xi)$, we conclude that $\zeta\in O'(\xi)$, as desired.n

To finalize the proof of theorem (3.1) in the "non-manifold" case, it is only necessary to *define* the manifold structure of O(x) via the structure of $O_{int}(x)$, for each x. The theorem then follows from the above two lemmas.

Remark 3.5: H.Sussmann has suggested to us a somewhat different proof of lemma 3.4. We scketch it now. Let ξ,ζ be as in the above proof, and the curves γ_i also as there. Since the elements $\gamma''(\lambda) := g_b(\xi,\gamma_1(\lambda)\cdots\gamma_r(\lambda))$ are in $O_{int}(\xi)$ for positive λ, and they approach ζ as $\lambda\to 0$, we conclude that $\dim O_{int}(\xi) \leq \dim O_{int}(\zeta)$. A symmetric argument gives the other inequality; thus **D** has the same rank at both points. (Here we are using lemma 3.3.) It follows that lemma 3.2 can be applied to the curve γ″ and the orbit $O_{int}(\xi)$ to conclude that ζ is indeed in this orbit.n

3.1. Some rank consequences

One would like to be able to conclude from theorem (3.1) that the action Σ has full rank at ξ if and only if O(ξ) is a neighborhood of ξ (in the topology of the ambient manifold **M**). However this is in general false, (counterexamples will be given later,) unless one has more a priori information on O(ξ). If the rank is full, then O(ξ), being a submanifold of the same dimension as **M**, is certainly open. But the converse need not hold. However, *if* orbits are know to be second countable, then O(ξ) cannot have lower dimension than **M** unless it has measure zero in **M** (see for instance a proof in [BC], proposition 8.5.6). Thus we would like to study when orbits are second countable. Equivalently (because of the paracompactness assumption,) we are interested in determining when O(ξ) has only countably many components (in the submanifold topology). One easy (and well known) case is that of continuous time actions with time-topology, or more generally (c.f. lemma (2.4)):

Proposition 3.6: Assume that Σ is connected. Then O(ξ) is connected (and in particular, second countable,) for each ξ. Thus O(ξ) is the leaf of **D** through ξ, and Σ has full rank at ξ if and only if O(ξ) is a neighborhood of ξ.

Proof: We may assume that we are in the manifold case, i.e. that $\mathbf{C}_a = \mathbf{U}_a$ for all $a\in\mathbf{A}$, because the orbits using $\mathbf{C}_a$ or int($\mathbf{C}_a$) are the same, and the topology is determined by the latter. Consider any ζ in O(ξ) which can be reached in one step, $\zeta = g_a(\xi,u)$. Let γ be a curve joining u with the u_o corresponding, in the definition of connected action, to this $a\in\mathbf{A}$ and ξ. Since g_a is continuous in u as a mapping into O(ξ), it follows that ζ and ξ must be in the same connected component of O(ξ). Now any ζ′ reachable in one step from ζ (hence, in two steps from ξ) is in the same component of O(ζ) = O(ξ) as ζ, and hence the same component as ξ. An inductive argument on number of steps gives the result.n

Another case where things are as desired is that of discrete time actions, or more generally:

Proposition 3.7: Assume that Σ is countable. Then O(ξ) is second countable, for each ξ. Thus, Σ has full rank at ξ if and only if O(ξ) is a neighborhood of ξ.

Proof: Note that O(ξ) is the union of countably many sets of the form $\{g_b(D_b)\}$. Each of these is the image of $g_b(\xi,\cdot)$, a smooth mapping into O(ξ) whose domain is the open subset {ω s.t. $(\xi,\omega)\in D_b$}. This open set is an open submanifold of the second countable manifold $\mathbf{U}_b$ (countability definition), and hence has only countably many components; thus its image also has countably many components.n

Let Φ = Φ(Σ), the set of all vector fields of the type $X_{a,u,\nu}$, (a,u,ν) with a∈**A** and (u,ν)∈T$\mathbf{C}_a$. Since **D** is integrable, it is invariant (recall the discussion in section 2.3,) so Ad(Φ,Φ) ⊆ Ad(**D**,**D**) = **D**. Thus, if Ad(Φ,Φ) has full rank then **D** does too. Consider the Lie distribution $\mathbf{D}_L(\Sigma) = \{\Phi\}_{LD}$; this is obtained from all possible linear combinations of iterated Lie brackets of the vector fields $X_{a,u,\nu}$. Since Ad(Φ,Φ) is involutive (because integrable,) and Φ⊆Ad(Φ,Φ), it follows that:

Proposition 3.8: The Lie distribution $\mathbf{D}_L(\Sigma)$ is included in **D**. In particular, if Σ has full Lie rank then it has full rank.n

Lemma 3.9: Let Σ be connected. Then **D** = Ad(Φ,Φ).

Proof: One inclusion is proved above; for the reverse, consider the action Σ(Φ) associated to Φ. Pick ξ in **M**. Let N be the leaf (maximal integral manifold) of Ad(Φ,Φ) through ξ. We claim that then O(ξ) is contained in N. This will imply that $T_\xi O(\xi)$ = **D**(ξ) is included in $T_\xi N$, giving the desired inclusion. By induction on the number of steps needed to reach ζ∈O(ξ), we may reduce to the problem of showing that $\zeta = g_a(\xi,u')$ is in N, for any u′ in int($\mathbf{C}_a$). Let γ be a piecewise constant curve joining u_o (corresponding to this a∈**A** and ξ) and u′, such that $g_a(\xi,\gamma(\lambda))$ is defined for each λ; such curves exist because of the assumption of connectedness and an approximation argument. Consider the curve $g_a(\xi,\gamma(\lambda))$. This joins ξ and ζ, and its derivative with respect to λ at λ =   is precisely $X_{a,u,\nu}(x)$, where

$$x = g_a(\xi,\gamma(\ )),\ u = \gamma(\ ),\ \nu = \dot\gamma(\ )\ . \tag{3.2}$$

Partition [0,1] into finitely many intervals I_i in each of which γ is constant. If y,z are endpoints of one such interval, then this argument shows that y,z are connected by an integral curve of Ad(Φ,Φ), and are therefore in the same leaf of this distribution. By induction on the intervals, ξ and ζ are in the same leaf, so ζ∈N. Thus, O(ξ)⊆N, as desired.n

Proposition 3.10: Assume that Σ is connected and *analytic*. Then **D** has full rank if and only if it has full Lie rank.

Proof: By the previous result, **D** = Ad(Φ,Φ). The vector fields Φ are analytic, so we are in the standard analytic continuous-time case, and the result is well known (see for instance [IS]), and not hard to prove. (Sketch: elements like $Ad_{\exp(tX)}(Y)$ generating Ad(Φ,Φ) can be written, for small enough t, as power series on iterated Lie brackets of X and Y, by the Baker-Campbell-Hausdorff formula, so the Lie algebra generated by Φ cannot have lower rank than Ad(Φ,Φ).)n

4. Proof of lemma 3.1

We shall need some more notation. For b in $\mathbf{A}^*$, m_b (or just m) will be the map $g_b(x,.)$, with domain $L_b := \{\omega \mid (x,\omega)\in D_b\}$. We also make the convention that a statement like "$g_b(x,\omega) = y$" will mean "(x,ω) is in D_b and $g(x,\omega) = y$". Fix an x in **M**, and let O = O(x). We establish first that

$$r(x,y) = r(x,z) \text{ for any } y,z \text{ in } O.$$

Pick b,c in $\mathbf{A}^*$ and ω,ω' such that $g_b(x,\omega) = y$, $g_c(x,\omega) = z$, and $\mathrm{rank}[dg(x,\omega)] = r(x,y)$. Introduce $e := (b,-b,c)$ and $\chi := \omega\varpi\omega'$. Since $g(x,\omega\varpi) = x$, it follows that $g(x,\chi) = z$. So $\mathrm{rank}[dg(x,\chi)] \le r(x,z)$. Let $F := g_{(-b,c)}(\cdot,\varpi\omega')$ --with domain the open set $\{x \mid (x,\varpi\omega') \in D_{(-b,c)}\}$. Since $d_xF(p)$ is a linear isomorphism for all p in the domain of F, it follows that $r(x,y) = \mathrm{rank}[dg(x,\omega)] = \mathrm{rank}[d_xF(y)\circ dg(x,\omega)] = \mathrm{rank}[d_b g(x,\chi)] \le \mathrm{rank}[dg(x,\chi)] \le r(x,z)$. A symmetric argument concludes the equality. Let r be the common value of the r(x,y).

Consider now the set S of all triples s:= (b,Q,h), where b is in $\mathbf{A}^*$ and:

$$Q \text{ is an r-dimensional embedded submanifold of } L_b, \tag{4.1}$$

$$m_b|Q\colon Q \to \mathbf{M} \text{ is injective and has differential of constant rank } r, \tag{4.2}$$

$$h\colon Q \to \Re^r \text{ is a diffeomorphism with an open subset } h(Q). \tag{4.3}$$

Fix one such s, and consider the set m(Q); this is a subset of O. The bijection m|Q induces a canonical manifold structure on this set, for which both m|Q and $\phi := h\circ(m|Q)^{-1}$ are diffeomorphisms (and such that ϕ is a chart). We now prove that, for this structure,

- (a) the inclusion i: $m(Q) \to \mathbf{M}$ has injective differential at every point, and
- (b) for any smooth structure Θ for O for which the lemma holds, the subset m(Q) is open --relative to Θ,-- and the identity map provides a diffeomorphism between the two structures.

The inclusion i factors as $m\circ j\circ(m|Q)^{-1}$, where j is the embedding of Q in L_b. Property (a) follows from the corresponding properties for its factors (for m, the properties hold on Q, which is sufficient). We now prove (b). Consider m as a map from L_b into O (with structure Θ); this map is smooth (property (1) in lemma: m is a restriction of g). So m|Q is also smooth into (O,Θ). Since the latter is a submanifold of **M**, and $\mathrm{rank}[dm|Q] = r$ (constant) as a map into **M**, this rank is also r as a map into (O,Θ). But this submanifold has dimension r, by part (2) of the lemma. Thus m(Q) is indeed open rel to Θ, and m|Q is a diffeomorphism between $(m(Q),\Theta)$ and Q, so (b) follows.

We now prove that the family of all such charts $(m|Q,\phi)$ defines a smooth (r-dimensional) manifold structure on O, and that property (1) holds. It will then follow from (a) above that this structure makes O into a submanifold of **M**, and the uniqueness statement follows from (b).

The sets m(Q) cover O: Pick any y in O and let b,ω be such that $g_b(x,\omega) = y$ and $dm(\omega) = dg(x,\omega)$ has rank r. Thus dm has maximal rank at ω, so there is an r-dimensional embedded submanifold Q of L_b, containing ω, such that equation (4.1), equation (4.2) are satisfied; replacing Q if necessary by an open subset of Q, a suitable h can be found for equation (4.3).

Compatibility: Pick any two charts $(m(Q),\phi)$ and $(m'(P),\beta)$ corresponding to (b,Q,h) and (c,P,k) respectively. Let $V := m(Q)\cap m'(P)$. We need to establish that:

(a) $\phi(V)$ is open in $\phi(m(Q))$, and
(b) $\beta\circ\phi^{-1}$ is smooth on V.

Pick an arbitrary y in V; thus there are ω,ω' in Q,P with $y = m(\omega) = m'(\omega')$. Let e:= (b,-c,c) in $\mathbf{A}^*$, and take $\chi := \omega\varpi'\omega'$. Note that $\mathrm{rank}[dm(\chi)] \geq \mathrm{rank}[d_c g(x,\chi)] = \mathrm{rank}[dg(x,\omega')] = r$. Since $dm(\chi)$ always has rank at most r, it has maximal rank at this χ. So there is an open subset Z of L_e which contains χ and such that $m_e(Z)$ is an r-dimensional embedded submanifold of **M**. Introduce the open set W [resp., W′] consisting of those υ in L_b [resp., L_c] such that $\upsilon\varpi'\omega'$ [resp., $\omega\varpi'\upsilon$] is in Z. Then ω is in W and ω' is in W′. Let

$$P' := P \cap W', \; Q' := Q \cap W.$$

Since Q is an embedded submanifold of L_b, and W is open in L_b, also Q′ is open in Q, and similarly for P,P′. Note that $m|Q'$ maps into $m_e(Z)$, and is injective with differential of constant rank r. Thus m establishes a diffeomorphism between Q′ and an open subset E of $m_e(Z)$. Similarly for $m'|P'$ and an open F in $m_e(Z)$. Note that $E \cap F \subseteq V$. Also, ω',ω are in P′, Q′ respectively, so y is in $E \cap F$. Since m|Q is injective,

$$(m|Q)^{-1}(E \cap F) = (m|Q')^{-1}(E \cap F),$$

which is then open in Q, because $E \cap F$ is open in E. So $\phi(E \cap F)$ is open in $h(Q) = \phi(m(Q))$. Thus $\phi(z)$ has a neighborhood included in $\phi(m(Q))$, and (a) follows. To prove (b), note that ϕ maps $E \cap F$ (embedded submanifold of $m_e(Z)$) diffeomorphically onto $\phi(E \cap F)$, which is open in h(Q) and contains $\phi(y)$. A similar statement holds for β. So $\beta \circ \phi^{-1}$ gives a diffeomorphism between $\phi(E \cap F)$ and $\beta(E \cap F)$, and (b) follows.

Property (1) of the lemma: We first establish that the maps m_b are smooth. Pick ω in L_b, $z = g(x,\omega)$. Since $r(x,z) = r$, there are a c and a ω' in L_c with $g(x,\omega') = z$ and $dg(x,\omega') = r$. Let e:= (b,-c,c) and $\chi := \omega\varpi'\omega'$. It will suffice to prove that m_e is smooth on some neighborhood of χ, because m_b is (in a suitable neighborhood of ω) a restriction of m_e. Note that

$$r \geq \mathrm{rank}[m(\chi)] \geq \mathrm{rank}[d_c g(x,\chi)] = \mathrm{rank}[dg(x,\omega')] = r$$

(this uses that $m(\omega\varpi') = x$). So m achieves maximal rank at χ. There is then a chart C of L_e, centered at χ, and diffeomorphic to a cube in $\mathfrak{R}^s \times \mathfrak{R}^r$, such that, if Q is the embedded submanifold corresponding to the factor $\mathfrak{R}^r$, then $\mathrm{rank}[dm(\chi)]$ is constantly r on Q and m_e is injective on Q. Let h give the corresponding diffeomorphism of Q with $\mathfrak{R}^r$. Then (e,Q,h) gives rise to a chart $(m(N),\phi)$. So $m_e|C$ is the composition of the projection onto Q and of m|Q, and is therefore smooth. To prove now that g_c is smooth as a map into O, pick any (z,ω) in D_c, z in O. Let (b,Q,h) give a chart around z. For (g,χ) in a neighborhood of (z,ω) in $(O \times U_c) \cap D_c$,

$$g_c(y,\chi) = m_{(b,c)}((m|Q)^{-1}(y),\chi), \tag{4.4}$$

so g_c is indeed smooth. This completes the proof of the lemma.n

5. S-actions

In this section and the next we introduce, for a given action Σ, two new types of associated actions Σ_S and Σ_B. These will be useful in the study of discrete and continuous time systems. Let Σ be fixed for the rest of this section.

We introduced earlier the action Σ^* whose index set is $\mathbf{A}^*$. Assume now that S is any subset of $\mathbf{A}^*$ with the property that $-s \in S$ whenever $s \in S$. Assume further that S contains all pairs of the form (-b,b), for $b \in \mathbf{A}^*$. There is a well-defined action Σ_S obtained by restricting the index set $\mathbf{A}^*$ to S. We call an action of this type an *S-action* associated to Σ. Let $\mathbf{D}_S$ be the associated distribution $\mathbf{D}(\Sigma_S)$.

Lemma 5.1: For any *S* as above, **D** = $\mathbf{D}_S$. For each $\xi\in\mathbf{M}$, $O(\xi) = \cup\{O_s(x), x\in O(\xi)\}$ (where O_s indicates orbit with respect to Σ_S).

Proof: Since *S* is a subset of $\mathbf{A}^*$ and $\mathbf{A}^*$-orbits are the same as **A**-orbits, the second statement follows. Further, by lemma (2.1) it follows that $\mathbf{D}_S$ is included in **D**. To prove the converse inclusion, pick any generator $Ad_{b,\omega}(X_{a,u,\nu})$ of **D**. We wish to prove that this is also in $\mathbf{D}_S$. Introduce the element c:= (b,-a,a,-b) of $\mathbf{A}^*$; this is in *S* by the second assumption on the form of *S*. Let ψ:= (ω,ʉ,u,ϖ), an element of $\mathbf{C}_c$. Finally, let ν' be the tangent vector at ψ having coordinates (0,0,ν,0) with respect to the decomposition of the tangent space corresponding to the product expression $\mathbf{U}_c = \mathbf{U}_b\times\mathbf{U}_{-a}\times\mathbf{U}_a X\mathbf{U}_{-b}$. By formula (2.3), it follows that the chosen generator is equal to the generator $X_{c,\psi,\nu'}$ of $\mathbf{D}_S$.n

From the rest of this section, let *S* be the set of all pairs (-b,b), $b\in\mathbf{A}^*$. We call the corresponding Σ_S the *canonical S*-action. Recall the definitions of complete and connected actions given in section (2.6). An interesting consequence of lemma 5.1 is that, if Σ is connected, (as for instance in the case of continuous time time-topology actions,) then $O_s(\xi)$ actually coincides with $O(\xi)$. This is because, by proposition 3.6, the latter is the leaf of **D** through ξ, and by the same argument for Σ_S (which is again connected), $O_s(\xi)$ is the same leaf. We introduce a more general concept:

Definition 5.2: The action Σ is *S-connected* if the canonical Σ_S is connected.

Proposition 5.3: If Σ is complete then it is *S*-connected.

Proof: In general, assume that Σ is a complete action which has the property: if $\xi\in\mathbf{M}$ and $a\in\mathbf{A}$ then there is a $u_o\in int(\mathbf{C}_a)$ such that $g_a(\xi,u_o) = \xi$. Then Σ is connected, because given any ξ, a, and u as in the definition of connected action, any curve joining u and u_o is an appropiate γ. We now show that Σ_S is of this form. Since Σ is complete, Σ_S also is (trivially). Thus is is only necessary to establish the existence of appropiate u_o's. But given any element (-b,b) in *S*, and any ξ in **M**, any $\omega\in int(\mathbf{C}_{-b,b})$, by completeness the element $g_{-b,b}(\xi,\varpi,\omega)$ is well defined, and it equals ξ.n

Corollary 5.4: Assume that Σ is a complete action. Then the leaf of **D** through ξ is $O_s(\xi)$ (orbit of canonical Σ_S). In particular, Σ has full rank at ξ if and only if $O_s(\xi)$ is a neighborhood of ξ.n

This corollary provides a characterization of the full rank property in the complete case. As a final remark, note that, for any *S*, $O_s(\xi)$ is a submanifold of $O(\xi)$ for each ξ. This is because we may always restrict the original action to $O(\xi)$, which is invariant under all $g_a(\cdot,u)$. Then theorem 3.1 may be reapplied to this restricted action, and the uniqueness statement in lemma 3.1 insures that the (sub)manifold structure obtained for $O_s(\xi)$ coincides with that obtained originally.

6. Balanced actions.

Again Σ is a fixed arbitrary action. We assume in this section that there is given a mapping σ: $\mathbf{A}\to\Re$ with the property that $\sigma(-a) = -\sigma(a)$, that the sets $\mathbf{U}_a$ are all identical, say to **U**, and that all $\mathbf{C}_a$ equal a fixed **C**. We extend σ to $\mathbf{A}^*$ by the formula $\sigma(a_1\cdots a_r) := \sum\sigma(a_i)$, with σ(empty word):= 0. This gives a monoid homomorphism $\mathbf{A}^*\to\Re$. All definitions will be with respect to a fixed such (Σ,σ).

Let π be an order 2 permutation (π^2=identity) of $\{1,\cdots,r\}$. The element $b = a_1\cdots a_r$ of $\mathbf{A}^*$ is *balanced* (with respect to π) iff $\sigma(a_i) = -\sigma(a_{\pi i})$ for all i. For such an element b, we let

$$\mathbf{C}_b^{bal} := \{(u_1 \cdots u_r) \in \mathbf{C}^r \mid u_i = \bar{u}_{\pi i}\} .$$
$$\mathbf{U}_b^{bal} := \{(u_1 \cdots u_r) \in \mathbf{U}^r \mid u_i = \bar{u}_{\pi i}\} .$$

The latter has a natural manifold structure, as an embedded submanifold of **M**, diffeomeorphic to a suitable product of **U**'s, and $\mathbf{C}_b^{bal}$ is a subset with nice boundary. We let Σ_B be the *balanced action* associated to (Σ,σ) and π. This is the action whose index set $\mathbf{A}^{bal}$ is the set of balanced $b \in \mathbf{A}^*$, and with **U** and **C** sets as above. Let $\mathbf{D}_B := \mathbf{D}(\Sigma_B)$. Note that Σ_B-orbits are contained in Σ-orbits, and that $\mathbf{D}_B$ is contained in **D**. Let

$$\mathbf{D}_o := \{Ad_\pi(X_{a,u,v}) - Ad_\rho(X_{a',u,v}) \mid \sigma(a)=\sigma(a'),\ (u,v) \in T\mathbf{C},\ \pi,\rho \in \Gamma(\Sigma)\}_D .$$

(With the understanding that only well-defined differences are considered.)

Proposition 6.1: $\mathbf{D}_B = \mathbf{D}_o$.

Proof: By formula (2.8), $\mathbf{D}_B \subseteq \mathbf{D}_o$. Consider now any $\xi \in \mathbf{M}$ and any two elements of the form $Ad_{c,\chi}X_{a,u,v}$ and $Ad_{c',\chi'}X_{a',u,v}$ defined at ξ, with σ(a)=σ(a′). Thus $\zeta := g_c(\xi,\chi)$, $\zeta' := g_{c'}(\xi,\chi')$, $g_{-a}(\zeta,\bar{u})$, and $g_{-a'}(\zeta',\bar{u})$ are all defined. Consider the word

$$d := (c,-a,a,-c,c',-a',a',-c') ,$$

which can be factored as (b,a,-b′,-a′,-b″) in the obvious way. This is a balanced word. Consider the element of $\mathbf{U}_d$ given by

$$\psi := (\chi,\bar{u},u,\bar{\chi},\chi',\bar{u},u,\bar{\chi}') .$$

It follows from formula (2.8) that the desired element is in $\mathbf{D}_B$.n

Pick now any a,a′ with σ(a)=σ(a′), and $u \in \mathbf{U}$. Let

$$\Phi_o := \{X_{a,u,v} - X_{a',u,v} \mid \sigma(a)=\sigma(a'),\ (u,v) \in T\mathbf{C}\}.$$

Let $L_o := \{\Phi_o\}_L$. Then, since $\mathbf{D}_B$ is involutive,

Lemma 6.2: $\{L_o\}_D \subseteq \mathbf{D}_B$.

Note the analogy with with proposition 3.8. The algebra L_o corresponds to the usual "0-time Lie algebra" for continuous time (time-topology) systems ([SJ]). Even if Σ is connected, Σ_B may be nonconnected; however completeness and the existence of "zero" is inherited:

Proposition 6.3: If Σ is a complete action [resp., an action with zero,] then Σ_B is complete [resp., an action with zero].

Proof: Let ξ and b balanced be given, $b = a_1 \cdots a_r$. If Σ has a zero 0, the sequence χ of r 0's is such that $g_b(\xi,\chi) = \xi$ for all ξ. So Σ_B also has a zero. (Note that χ is indeed in int($\mathbf{C}_b^{bal}$).) Take now Σ to be complete. The domain condition is certainly satisfied by Σ_B. Assume given now any $\omega = u_1 \cdots u_r$ in int($\mathbf{C}_b^{bal}$) and another such ω′. By connectedness of int(**C**) there are curves γ_i joining u_i with u_i'. We may assume that, if j=πi, then $\gamma_j(\lambda)$ is $\bar{\gamma}_i(\lambda)$. Thus $\gamma = (\gamma_1 \cdots \gamma_r)$ provides a curve joining χ to ω in int($\mathbf{C}_b^{bal}$).n

Finally, note that by an argument as in the previous section, orbits under Σ_B are submanifolds of orbits under Σ.

7. Continuous time systems

Continuous time systems are described by controlled differential equations

$$\dot{x}(t) = P(x(t),u(t)),\ t\in\Re, \qquad (7.1)$$

where the *state* x(t) belongs to a second countable (paracompact) manifold **M**, controls u(t) take values in a set **C** which is a subset with nice boundary and connected interior of a second countable (paracompact) manifold **U**, P is smooth as a mapping from **M**×**U** into the tangent bundle T**M**, and for each $u\in\mathbf{U}$, $X_u := P(\cdot,u)$ is an (everywhere defined but not necessarily complete) vector field on **M**. An *analytic* system is one for which all data is analytic.

The *orbit* O(ξ) of an $\xi\in\mathbf{M}$ is the set of points that can be reached from ξ in positive and negative times, i.e. the set of all points of the form

$$\exp[t_1X_{u_1}]\cdots\exp[t_rX_{u_r}](\xi)\ ,$$

where r is an integer, the u_i are in **C**, and the t_i are arbitrary real numbers. The *zero-time orbit* $O_o(\xi)$ is obtained by considering all points as above but with the constraint that $\sum t_i=0$.

The previous theory can be applied to the above systems in two very different ways. The first was described in section 2.3, and consists in associating the action in which

$$g_{\varepsilon,X_u}(\xi,t) = \exp[\varepsilon tX_u](\xi)\ . \qquad (7.2)$$

The alternative which motivated much of the previous discussion is to consider instead

$$g_t(\xi,u) := \exp[tX_u](\xi)\ . \qquad (7.3)$$

More precisely, we consider the action with $\mathbf{A}:=\Re$, the obvious "-", $\mathbf{U}_a$ [respectively, $\mathbf{C}_a$,] equal to **U** [resp., **C**,] for all a, the identity ι_a, and D_t the domain of the above map. We refer to an action of this type as a *continuous time action with input-topology*.

The first (time-topology) model is the one implicitely used in the literature when dealing with continuous time systems, and is the basis of the known orbit theorems. The second model (which cannot be even considered in the usual treatments, where smoothness of (7.1) in u is not necessarily assumed,) will provide the right framework for understanding sampling results.

Note that the orbit O(ξ) is the same, *as a set*, as the orbit of the corresponding continuous time action with time-topology or that of the corresponding continuous time action with input-topology. Thus the notations are consistent. We let $\mathbf{D}_T$ and $\mathbf{D}_U$ be the distributions obtained with the respective actions Σ_T, Σ_U. These are in general different, and so the manifold structures that result on O(ξ) from the two possible applications of theorem 3.1 will be in general different. For instance, in the trivial case in which $\mathbf{M} = \Re$ and $P(\xi,u)\equiv 1$ (**U** is then irrelevant), clearly O(ξ)=**M** for each ξ. Here $\mathbf{D}_T$ has full rank at every point, and O(ξ) has the topology of **M**. On the other hand, $\mathbf{D}_U$ has dimension zero at each point. The connected component of O(ξ) at ξ with the input-topology is just {ξ}; the orbit $O(\xi) = \Re$ has uncountably many components.

The set Φ is in the time-topology case just the linear span of $\{X_u,\ u\in\mathbf{C}\}$. Since the corresponding action is connected, we have that O(ξ) contains a neighborhood of ξ, relative to the topology of **M**, iff the $\mathbf{D}_T$ has full rank at ξ. We also know that in the analytic case the latter is equivalent to full Lie rank. In the input-topology case, full rank of $\mathbf{D}_U$ is sufficient but not necessary (above counterexample) for O(ξ) to be a neighborhood of ξ. The input-topology action is not countable nor, in general, connected. When the vector fields defining the system are complete in the usual sense, this action is complete and hence *S*-connected, so the leaves of the corresponding distributions are the orbits $O_s(\xi)$ (again see the above example).

In the time-topology case we can apply the material in section 6. Here σ is defined by $\sigma(\varepsilon,X) := \varepsilon$. We restate the conclusion of proposition 6.1 in this special case. Let Γ be the (pseudo-) group generated by Φ, i.e. the set of all compositions of elements of the form $\exp(tX_u)$. Then,

$$\mathbf{D}_B = \{Ad_\pi(X_u) - Ad_\rho(X_v) \mid \pi,\rho \in \Gamma \text{ and } u,v \text{ in } \mathbf{C}\}_D . \tag{7.4}$$

This action Σ_B may fail to be connected. For instance, consider the system with $\mathbf{M} = \Re^2$-{(n,-n), n=integer}, $\mathbf{U} = \mathbf{C} = \Re$, and equations

$$\dot{x} = u, \quad \dot{y} = 1-u .$$

For any control, x(t)+y(t) = x(0)+y(0)+t, and the orbit of (1/2,1/2) under Σ_B is the set of points of the form (x,-x) with x not an integer. This is not connected even in the topology of **M**. We shall prove below that these orbits are exactly the sets $O_o(\xi)$.

Remark 7.1: The pathology is due to the fact that the system (action) is not complete. If the system is complete, $O_o(\xi)$ is connected in the time-topology (and hence in the topology of **M**). This is very easy to establish directly, but also follows from the previous work. Indeed, consider the balanced action Σ_B associated to the time-topology action. If Σ is complete, Σ_B is also complete. Further, Σ has a zero, so Σ_B is a complete action with zero, and hence connected. Thus, by proposition 3.6 the orbits under Σ_B are connected. These are submanifolds of M and of the orbits under Σ_T, so they are connected in these topologies too.n

Note that the *canonical* action Σ_S (recall section 5) gives nothing new *for the time-topology* case: the obtained orbits are the same as for Σ. This is in general true for connected actions, as discussed previously. Or directly, because given any (ε,X), ξ, and any t we may write $\exp[\varepsilon tX](\xi) = \exp[\varepsilon(t+\delta)X]\circ\exp[-\varepsilon\delta X](\xi)$ for small enough δ, which exhibits $g_{\varepsilon X}(x,t)$ as an element of the form $g_{-b,b}(x,t_1t_2)$. One could also consider different balanced actions associated to the input-topology model, but we don't need to do so here.

In this section, unless otherwise stated, Σ_B *will mean* the balanced action associated to the *time-topology* model of the given continuous time system, and Σ_S *will be used exclusively* to mean the S-action associated to the continuous time *input-topology* model, with the following set *S*:

$$S := \{b \in \mathbf{A}^* \mid b = t_1 \cdots t_r, \textstyle\sum t_i = 0\} .$$

Note that *S* is invariant under "-" and that (-b,b) is in *S* whenever $b \in \mathbf{A}^*$. The corresponding Σ_S is *not* the canonical S-action associated to Σ.

Proposition 7.2: The following four sets are equal for each ξ in **M**:

- The zero-time orbit $O_o(\xi)$
- The orbit of ξ under Σ_S
- The orbit of ξ under Σ_B
- If Σ is complete, the orbit of ξ under the canonical S-action Σ_S.

Proof: From the definitions, it is clear that orbits under Σ_S equal the sets O_o, and that orbits under Σ_B and under the canonical-*S* action Σ_S are included in zero-time orbits. Consider now the proof that in the complete case the orbits under the canonical S-action are precisely the sets $O_o(\xi)$. It will be sufficient to prove that each element ζ of the form

$$\exp[t_r X_{u_r}]\cdots\exp[t_1 X_{u_1}](\xi),\ \sum t_i=0, \tag{7.5}$$

can be obtained by applying to ξ a composition g of diffeomorphisms of the type exp(tX)exp(-tX), for various X,t. More generally, assume that the times t_i in (7.5) add up to τ. We claim that ζ can be obtained as the result of the application of such a g followed by an element of the type $\exp(\tau X)$, $X=X_u$. We prove this claim by induction on r. For r=1, the result is trivial. If true for r-1, write the expression in equation (7.5) as

$$\exp[t_r X_{u_r}]\exp[\tau' X]g(\xi)\ ,$$

where $t_r+\tau' = \tau$. This equals $\exp[\tau X_{u_r}]g'(\xi)$, where $g' = \exp[-\tau' X_{u_r}]\exp[\tau' X]g$, establishing the induction step.

Consider now the remaining statement. We wish to write every expression as in equation (7.5) as a balanced expression, i.e. one such that for each exp(tX) there is a corresponding exp(-tY). (In the complete case, this follows from the previous paragraph.) We prove by induction on r that any such expression with $\sum t_i = \tau$ can be written in the form

$$\exp[s_1X_1]g_1\exp[s_2X_2]g_2\cdots g_{k-1}\exp[s_kX_k](\xi)$$

with the g_i balanced and the s_i all of the same sign and adding up to τ. (And the X_i of the form X_u.) So, if $\tau=0$, all s_i must be zero, and we have a balanced expression. When r=1, the claim is true with k=1. Assume inductively that we are given the expression

$$\exp[t_oX]\exp[s_1X_1]g_1\exp[s_2X_2]g_2\cdots g_{k-1}\exp[s_kX_k](\xi). \tag{7.6}$$

Assume without loss that $t_o\neq 0$. If t_o has the same sign as the common sign of the s_i, then the induction step is proved. Otherwise, suppose that $t_o<0$ and the s_i are all positive (the opposite case is analogous). Let σ_j be the sum of the s_i for $i=1,\cdots,j$, with $\sigma_o:=0$. If $-t_o>\sigma_k$, we may rewrite the above as $\exp[\tau X]g(\xi)$, where g is the balanced (and well defined at ξ) element

$$g = \exp[-s_1X]\exp[-s_2X]\cdots\exp[-s_kX]g'\ ,$$

and g' is the element in (7.6). If instead $-t_o>\sigma_j$ but $-t_o<\sigma_{j+1}$, let $\tau':= \sigma_j+t_o$. We may then write the above as

$$g\ \exp[(s_{j+1}+\tau')X_{j+1}]\ g_{j+1}\cdots\exp[s_kX_k](\xi),$$

where $g = \exp[\tau' X]g\exp[-\tau' X_{j+1}]$ is balanced. Note that the expression is indeed defined: from $\sigma_j+s_{j+1}>-t_o$ we conclude that $-\tau'<s_{j+1}$, and together with $\tau'<0$ we know that $s_{j+1}>s_{j+1}+\tau'>0$.n

Remark 7.3: In the above equalities, one may as well take the orbits under Σ_S and Σ_B obtained when restricting controls to int**C**. This is because, as proved earlier, orbits do not change if such a restriction is made, and the proposition applies both to the original case and the case where **C** is replaced by int**C**.

Definition 7.4: The action Σ satisfies the *strong Lie rank condition* at ξ iff Σ_B has full Lie rank at ξ.

The following definition will be more relevant to our main results.

Definition 7.5: The action Σ is *strongly transitive* at ξ iff $O_o(\xi)$ is a neighborhood of ξ.

The definition refers to the topology of the ambient manifold **M**. It turns out, however, that strong transitivity is equivalent to ξ being interior with respect to the topologies induced by Σ_S or by Σ_B. Or, in terms of the associated distributions,

Theorem 7.1: The following conditions are equivalent:

- Σ is strongly transitive at ξ
- $\mathbf{D}_S$ has full rank at ξ
- $\mathbf{D}_B$ has full rank at ξ

Definition 7.6: The *time* topology of $O_o(\xi)$ is the topology induced from the action Σ_B. The *input* topology is that induced from the action Σ_S.

The above fundamental result is a consequence of the fact, whose proof is given in the next section:

Proposition 7.7: Both the time and the input topologies on $O_o(\xi)$ are second countable.

Second countability insures that the dimension of $O_o(\xi)$ is equal to the dimension of **M**, in the strongly transitive case, for the manifold structures given by theorem 3.1 applied to each of the actions.

Recall (lemma (5.1),) that $\mathbf{D}_S$ is just **D**, where the latter is the distribution associated to the input-topology model. So (compare with formula (7.4),)

$$\mathbf{D}_S = \{Ad_\pi(X_{t,u,v}) \mid \pi\in\Gamma,\ t\in\Re,\ (u,v)\in T\mathbf{C}\}_D\ . \tag{7.7}$$

The explicit form of the vector fields appearing above is:

$$X_{t,u,v} = \frac{\partial exp[tX_v]exp[-tX_u](\xi)}{\partial v} \Big|_{v:=u}(v). \tag{7.8}$$

If P(x,u) in (7.1) is affine in u and the system is analytic, then for small t these generators can be expanded in terms of the Lie distribution appearing in lemma 6.2 (these expansions will appear later), and that fact may help to understand intuitively why theorem 7.1 is true.

Remark 7.8: It is important to note that proposition 7.7, and theorem 7.1, depend essentially, in the time-topology case, on the assumption made in defining continuous time systems that the interior of **C** is connected. Otherwise we could consider a system on $\mathbf{M} = \Re$ having, for instance, $\mathbf{U} = \mathbf{C} = (0,1)\cup(2,3)$, and the dynamics given by f(x,u) = 0 if $u\in(0,1)$ and 1 if $u\in(2,3)$. For this system, $O_o(\xi)=\Re$ for each ξ, but $\mathbf{D}_S(\xi)=\{0\}$ for each ξ, since the partial derivatives in equation 7.8 are clearly all zero.n

7.1. Second countability of 0-time orbit with both topologies.

Given a continuous time system, we consider the actions Σ_B and Σ_S. An ξ in **M** is fixed; the objective is to study the zero-time orbit $O_o(\xi)$. By remark (7.3), we may assume that **C** is open, or, changing notations, that **C** = **U**. This is because set-theoretically $O_o(\xi)$ does not change, and the topologies are determined by the interior of **C**. For each integer r and each order 2 permutation π of {1...r}, consider the function

$$\alpha(t,\omega) = \alpha_{r\pi}(t,\omega) := \exp[t_r X_{u_r}]\cdots\exp[t_1 X_{u_1}](\xi) \tag{7.9}$$

where $t=(t_1,...,t_r)$ and $\omega=(u_1,...,u_r)$. Its domain $E_{r\pi}$ is that open subset of $\Re_\pi \times \mathbf{U}^r$ where the expression is defined, and $\Re_\pi := \{t=(t_1,\cdots,t_r) \mid t_i = t_{\pi i}\}$. Note that the image of each α is included in $O_o(\xi)$, so we view α as a map into the zero-time orbit. It is essential to note, in order to understand the necessity of the arguments to follow, that this mapping is *not* continuous for either of the topologies of $O_o(\xi)$ which we are considering. However, α is continuous with respect to t when $O_o(\xi)$ is given the time topology, and is continuous with respect to ω if the input topology is considered instead (hence the names for the topologies). This is because $\alpha(t,\omega)$ equals $g_{t_1\cdots t_r}(u_1\cdots u_r)$ when g is as in (7.3) and also equals $g_{(\varepsilon_1 u_1)\cdots(\varepsilon_r u_r)}(|t_1|,\cdots,|t_r|)$ when g is as in (7.2) and $\varepsilon_i = \mathrm{sign}(t_i)$.

Every element of $O_o(\xi)$ can be written as $\alpha(t,\omega)$ for suitable r, π, ω, and t (and all $t_i \neq 0$, if desired). This is because of the equality of O_o and orbit under Σ_B. Fix now r and π.

Lemma 7.9: For each (t^o,ω^o) in $E_{r\pi}$ there exists a neighborhood N of (t^o,ω^o) in $E_{r\pi}$ such that $\alpha(N)$ is connected in the *input* topology.

Proof: For each pair of integers $1 \le i < j \le r$, we let

$$\phi_{ij}(t,\tau) := (t_1,0,t_2,0,\cdots,0,t_i,\tau,t_{i+1},0,\cdots,0,t_j,-\tau,t_{j+1},0,\cdots,0,t_r)$$
$$\chi_{ij}(\omega,u,v) := (u_1,u_1,u_2,u_2,\cdots,u_i,u,u_{i+1},\cdots,u_j,v,u_{j+1},u_{j+1},\cdots,u_r,u_r).$$

Also, let π_{ij} be the permutation of $\{1,\cdots,2r\}$ with $\pi_{ij}(2\ \ -1) = 2\pi\ \ -1$, $\pi_{ij}(2i)=2j$, $\pi_{ij}(2j)=2i$, and $\pi_{ij}k=k$ for all other k. Introduce

$$Z := \{(t,\tau,\omega,u,v) \in \Re_\pi \times \Re \times \mathbf{U}^r \times \mathbf{U} \times \mathbf{U} \mid (\phi_{ij}(t,\tau),\chi(\omega,u,v)) \in E_{2r,\pi_{ij}} \text{ for all } i<j\} .$$

The set Z is open, by well-posedness of ordinary differential equations. For the given ω^o, let C be a compact subset of $\mathbf{U}$ such that intC is path connected and all components u_o^i of ω are in intC.* Let $K := \{(t^o,0,\omega^o,u,v) \mid u,v \in C\}$. This set is compact and it is included in Z because

$$\alpha_{2r,\pi_{ij}}(\phi_{ij}(t^o,0),\chi(\omega,u,v)) = \alpha_{r\pi}(t^o,\omega^o)$$

for all u,v. Thus there is an open neighborhood V of K contained in Z. Moreover, V can be taken to be 'rectangular', meaning

$$V = \Pi(t_i^0-\delta,t_i^0+\delta) \times (-\varepsilon,\varepsilon) \times A_1 \times\cdots\times A_r \times B \times B ,$$

where B is an open set containing C, and for each i, A_i is a path-connected subset of intC which contains the corresponding u_o^i. Further, we assume that $2\delta < \varepsilon$. (The product of the intervals $(t_i^0-\delta,t_i^0+\delta)$ is understood as a subset of $\Re_\pi$.) Finally, let

$$N := \Pi(t_i^0-\delta,t_i^0+\delta) \times A_1 \times\cdots\times A_r .$$

Pick any (t,ω) and (s,ω') in N. We want to construct a path (in the input topology) connecting $\alpha(t,\omega)$ with $\alpha(s,\omega')$. We first connect $\alpha(s,\omega)$ with $\alpha(s,\omega')$. Inductively, we may assume that ω and ω' differ in only one coordinate, say the i-th. Since both u_i and u_i' are in A_i, a path connected subset of C, there exists a path γ with $\gamma(0)=u_i$, $\gamma(1)=u_i'$, and $\gamma(\lambda)$ in A_i for all λ. Composing with α (as a function of u_i) we get the desired path in $O_o(\xi)$. Note that

---

*This is the only place where the (essential) assumption that int**C** is connected is ever used.

connectedness of (interior of the) the control set is not used here; only local connectivity is used. Now consider the problem of connecting $\alpha(t,\omega)$ with $\alpha(s,\omega)$. Since α is not continuous with respect to t, this is not straightforward. Inductively, we assume that t,s differ only at the i-th coordinate. We let $j:=\pi i$, and assume $i<j$. (If i=j then the antisymmetry condition $t_i = -t_{\pi i}$ implies that both s_i and t_i must be zero, and hence equal.) Thus, s has the form

$$s = (t_1,\cdots,t_{i-1},s_i,t_{i+1},\cdots,t_{j-1},s_j,t_{j+1},\cdots,t_r)$$

with $s_i = -s_j$. Since both (t,ω) and (s,ω) are in N, we have that $|t_i-s_i|<2\delta$. Let $\tau:= s_i-t_i = t_j-s_j$, so that $\tau\in(-\varepsilon,\varepsilon)$. We can certainly write

$$\alpha(t_1\cdots t_r,\omega) = \alpha_{2r,\pi_{ij}}(\phi_{ij}(t,0),\chi(\omega,u_i,u_j)) \qquad (7.10)$$

$$\alpha(s_1\cdots s_r,\omega) = \alpha_{2r,\pi_{ij}}(\phi_{ij}(t,\tau),\chi(\omega,u_i,u_j)) \;. \qquad (7.11)$$

Claim: $\alpha(s_1\cdots s_r,\omega)$ is in the same component as

$$\alpha_{2r,\pi_{kj}}(\phi_{kj}(t,\tau),\chi_{kj}(\omega,u_k,u_j)) \qquad (7.12)$$

if $i\le k<j-1$ and in the same component as $\alpha(t,\omega)$ if $k=j-1$. We prove the claim by induction on k. For k=i, this is trivial by equation (7.11). Assume now the claim proved for k. Since ω is in $A_1\times\cdots\times A_r$, it follows that u_k and u_{k+1} are both in intC. Thus there is a path γ connecting u_k and u_{k+1}, with the image of γ contained in intC, and hence in B. Consider the path

$$\gamma'(\lambda) := \alpha_{2r,\pi_{kj}}(\phi_{kj}(t,\tau),\chi_{kj}(\omega,\gamma(\lambda),u_j)) \qquad (7.13)$$

into $O_o(\xi)$ with the time topology. This is well defined, because of the choice of the neighborhood V, and it connects the element in equation (7.12) with the the corresponding element having u_{k+1} instead of u_k. If $k+1<j$, this equals

$$\alpha_{2r,\pi_{k+1,j}}(\phi_{k+1,j}(t,\tau),\chi_{k+1,j}(\omega,u_{k+1},u_j)) \;,$$

(note the new subscripts,) because of the fact that $\exp[t_{k+1}X_{u_{k+1}}]$ and $\exp[\tau X_{u_{k+1}}]$ commute. This establishes the inductive step, and proves the first part of the claim. Applying now the same argument with u_{j-1} and u_j, the expression obtained at the end of the path is simply $\alpha(t,\omega)$, by the equality $\exp[-\tau X_{u_j}]\exp[tX_{u_j}]\exp[\tau X_{u_j}] = \exp[\tau X_{u_j}]$. This completes the proof of the lemma.n

Lemma 7.10: For each (t^o,ω^o) in $E_{r\pi}$ there exists a neighborhood N of (t^o,ω^o) in $E_{r\pi}$ such that $\alpha(N)$ is connected in the *time* topology.

Proof: The proof is very analogous to the previous one. Fix ξ in **M**. Consider now the functions

$$\phi(t,\lambda) := (\lambda t_1,(1-\lambda)t_1,\lambda t_2,(1-\lambda)t_2,\cdots,t_r,(1-\lambda)t_r)$$
$$\chi(\omega,\omega') := (u_1,u_1',u_2,u_2',\cdots,u_r,u_r') \;.$$

Let π' be the permutation of $\{1,\cdots,2r\}$ obtained from π by the formulas $\pi'(2i)=2\pi i$, $\pi'(2i+1)=2\pi i+1$. Now let

$$Z := \{(t,\lambda,\omega,\omega')\in\Re_\pi\times\Re\times\mathbf{U}^r\times\mathbf{U}^r \mid (\phi(t,\lambda),\chi(\omega,\omega'))\in E_{2r,\pi'}\} \;.$$

This is again an open set. It contains the compact set

$$K := \{(t^o,\lambda,\omega^o,\omega^o) \mid \lambda\in[0,1]\}$$

because $\alpha_{2r,\pi'}(\phi(t,\lambda),\chi(\omega,\omega')) = \alpha(t^o,\omega^o)$ and all intermediate expressions are well defined. Thus there is a neighborhood V of K in Z, of the form

$$A = \Pi(t_i^0-\delta,t_i^0+\delta) \times (-\varepsilon,1+\varepsilon) \times A_1\times\cdots\times A_r\times A_1\times\cdots\times A_r,$$

where $\delta,\varepsilon>0$ and u_o^i is in intA_i for each i. (Again, the first product of intervals should be interpreted as a subset of $\Re_\pi$.) Let

$$N := \Pi(t_i^0-\delta,t_i^0+\delta) \times A_1\times\cdots\times A_r .$$

Pick (t,ω) and (s,ω'). We want a path connecting $\alpha(t,\omega)$ and (s,ω') in the time topology. If $\omega = \omega'$, this is again easy: by induction, assume that t,s differ only at t_i, t_j, where $\pi i=j$. A path from t_i to s_i maps into the appropiate paths in $O_o(\xi)$. The interesting case is that in which we try to connect $(t,\omega)\in N$ and $(t,\omega')\in N$. In that case, consider the path

$$\gamma'(\lambda) := \alpha_{2r,\pi'}(\phi(t,\lambda),\chi(\omega,\omega')),\ 0\le\lambda\le1,$$

for this given t and ω, ω'. This is a continuous as a map into $O_o(\xi)$ with the time topology, and joins the desired points.n

We can now complete the proof of proposition 7.7. Consider the case of the input topology. The set $O_o(\xi)$ is a countable union of sets of the form $\alpha_{r,\pi}(E_{r\pi})$, so it is enough to show that each of these latter sets intersects at most countably many components of $O_o(\xi)$ with the input topology. (Note that this is not immediate, since for each component of $O_o(\xi)$, the preimages under α are *not* necessarily open, because α is not continuous.) But there is a covering of $E_{r\pi}$ by open sets N each of which maps into a connected component, by lemma 7.9. Since $E_{r\pi}$ is second countable, because **U** is, it follows that there is a countable subcover by these sets N ("Lindeloff" property), and the result follows. The time topology case is entirely analogous, using lemma 7.10 instead.n

8. Invertible discrete-time systems.

(Invertible) discrete time systems are a natural class of discrete time control systems, and where studied explicitely first by [JA]. They are described by controlled difference equations

$$x(t+1) = P(x(t),u(t)),\ t\in \mathbf{Z}, \qquad (8.1)$$

where the *state* x(t) belongs to a second countable (paracompact) manifold **M**, controls u(t) take values in a set **C** which is a subset with nice boundary and connected interior of a second countable (paracompact) manifold **U**. The map $P: D\to\mathbf{M}$ is smooth on an open subset D of $\mathbf{M}\times\mathbf{U}$, and (invertibility) for each u in **C**, $P(\cdot,u)$ is a (partial) diffeomeorphism; the set D is assumed to satisfy: for each x there is some u in **C** such that (x,u) is in D. An *analytic* system is one for which all data is analytic. We associate to each such system a discrete-time action by letting g_1 be P, and taking $g_{-1}(\cdot,u)$ to be by definition the inverse of $P(\cdot,u)$.

In this case, the vectors $X_{a,u,\nu}$ correspond either to forward motions followed by backward motions (a=-1) or viceversa (a=1). The distribution **D** consists of all conjugations of these vectors by iterated compositions of maps $P(\cdot,u)$ and their inverses. In the special case when dynamics are complete, in the sense that the maps P in (8.1) are defined and invertible for all x and u, the corresponding action is complete and hence (c.f. proposition 5.3) also *S*-complete. Applying then an argument analogous to that in the two paragraphs after equation (7.5), we have:

Proposition 8.1: For discrete-time *complete* actions, the orbits are second countable and the leaves of **D** are the 0-time sets $O_s(\xi)$ consisting of all $g_b(\xi,\omega)$ such that $\sum b_i=0$.n

Thus the leaves correspond in the complete case to "zero time orbits" in discrete time. In general, all we can say is that discrete-time systems give rise to countable actions, and hence proposition 3.7 applies.

In the analytic complete case, there is a Lie algebraic criterion for transitivity that may be easier to apply than checking the rank of **D**. It is to some extent related to the result in proposition 3.10. (Note however that in applying this criterion to sampled sytems -see next section,- it is still necessary to integrate the original continuous time system; the criterion is not a "direct" condition based on the vector fields defining the system, as one using Lie distributions would be.) This criterion, which we prove below, was first established by [JNC], based on computations in differential algebra.

Assume a discrete-time complete action Σ is given, and fix a control value u_o in int**C**, to be denoted simply by 0. Let $\pi := g_+(\cdot,0)$ and let Γ^* be the group generated by π. More generally, for each integer i>0 we associate the mappings

$$g_i(x,u) := \pi^{-i}(g_+(\pi^{i-1}(x),u))$$
$$g_{-i}(x,u) := \pi^{-i+1}(g_-(\pi^{i}(x),u)) .$$

There is then an action Σ_π defined by using these g's together with the original sets **U** and **C**. Let $O_\pi(\xi)$ be the orbit of ξ under this action. Again applying an argument analogous to that in the two paragraphs after equation (7.5), we have:

Lemma 8.2: For all ξ, $O_\pi(\xi) = O_S(\xi)$.n

Consider the distribution $\mathbf{D}_\pi$ associated to the new action. The vector fields in $\Phi(\Sigma_\pi)$ are precisely those of the form $Ad_{\pi^i}(X)$, with i an integer and X in $\Phi(\Sigma)$. The action Σ_π is connected, since $g_i(\xi,0)=\xi$ for all ξ, and any other $g_i(\xi,u)$ can be deformed to this by completeness. Thus by lemma 3.10 its distribution has full Lie rank iff it has full rank, in the analytic case. By lemma 5.1, **D** and $\mathbf{D}_\pi$ coincide. We can summarize the discussion as follows (this is, with different terminology, the result in [JNC]):

Corollary 8.3: Assume that the Σ is a complete discrete time action. Then, Σ is transitive at ξ if (and, in the analytic case, only if) Σ_π has full Lie rank at ξ.n

There is yet another sufficient condition for transitivity, not necessary even in the analytic case, which will be of interest in the context of sampling. The rest of this section studies that condition. We still assume that Σ is complete and that an element "0" has been fixed in int**C**. We let π be $g_+(\cdot,0)$, as before, and $\gamma := \pi^{-1} = g_-(\cdot,0)$. Let $e_1,\cdots,e_m$ be a basis of $T_o\mathbf{U}$. For each j = $1,\cdots,m$ we introduce the vector field

$$b_j := Y_{+,0,e_j} = -X_{+,0,e_j} ,$$

where the notation is as in equation (2.6). These vectors correspond to backward movements by 'small' u followed by forward motions by u=0. Since they are elements of **D**, it follows that if the set of vector fields

$$L := \{b_1,\cdots,b_m,Ad_\gamma(b_1),\cdots,Ad_\gamma(b_m),Ad_\gamma^2(b_1),\cdots\}_L \qquad (8.2)$$

generates a distribution of full rank at ξ, then Σ is transitive at ξ. A (rather surprising) result in sampling will be that in a certain sense this condition will be also sufficient for analytic systems. The Lie algebra L can be also generated in a different way, which will be needed later. Let $\alpha_{i,v}$ be the function $g_-(\cdot,v)^i$, for each v. For each i≥1 and each j consider

$$b_{ij}(\xi) := \frac{\partial \pi^i(\alpha_{i,v}(\xi))}{\partial v} \,|_{\,v=0}(e_j), \tag{8.3}$$

so that $b_{1j} = b_j$ for each j. An easy calculation with coordinates shows that

$$b_{i+1,j} = Ad_\gamma(b_{i,j}) + b_j$$

for all j. We conclude:

Lemma 8.4: $\{b_{ij}, j=1,\cdots,m, i\geq 1\}_L = L$.n

9. Sampling

Consider a continuous time system as in section 7. In digital control it is of interest to restrict attention to controls in equation (7.1) which are constant on intervals of length δ, where δ is a positive "sampling interval". It is of interest to determine when properties of controllability are preserved under sampling (restriction to sampled controls). There is a large literature on such issues for linear systems; see for instance [KHN], [BL], or [GH]. These results establish in particular that controllability is preserved if δ is small enough. We prove in this section a result along these lines, for transitivity, as well as a more algebraic criterion. The result is very easy to prove based on the above machinery, and generalizes that in [SS]. It is also possible to give (positive-time) controllability results, as done in [SO1], but the techniques required are different, and we do not do so here.

Assume a continuous time system (7.1) is given. We consider the associated actions Σ_B and Σ_S discussed in section 7. For each $\delta>0$ we introduce also the discrete time action Σ_δ defined by taking $g_+(\xi,u) := \exp[\delta X_u]$ and $g_-(\xi,u) := \exp[-\delta X_u]$.

Definition 9.1: The continuous time system (7.1) is δ-*transitive* at ξ iff Σ_δ is transitive at ξ. It is *sampled transitive* iff it is δ-transitive for some $\delta>0$.

The distribution associated to Σ_δ is denoted by $\mathbf{D}_\delta$. Note that the set $\Phi(\Sigma_\delta)$ is the set of all vector fields as in equation (7.8) which have $t=\pm\delta$, and $\mathbf{D}_\delta$ is obtained as in (7.7), when restricting to such t and to the subgroup of Γ generated by the $\exp[\delta X_u]$. Thus $\mathbf{D}_\delta$ is included in $\mathbf{D}_S$, and:

Proposition 9.2: If a continuous time system is sampled transitive then it must be strongly transitive.n

Before stating the basic result on sampled transitivity, which provides a strong converse to this proposition, we give a (very) easy lemma on matrices. This will be applied in a couple of places later, and also gives as a corollary the result in the appendix of [SO2]. If *A* is a set of real matrices all of size $\rho\times\sigma$, we let S(*A*) denote the subspace of $\Re^\rho$ generated by the columns of all the matrices in *A*. For any real δ and any k, $\delta\mathbf{Z}^k$ is the lattice of $\Re^k$ consisting of all $t=(t_1,\cdots,t_k)$ with t_i/δ = integer for each i. If $A(t) = A(t_1,\cdots,t_k)$ is an $\rho\times\sigma$-matrix of smooth functions of $t=(t_1,\cdots,t_k)$ defined on a connected open subset V of $\Re^k$, and if $\alpha=(\iota_1,\cdots,\iota_k)$ is a k-vector of nonnegative integers adding to r, we denote the (componentwise) α-derivative of A evaluated at t, as follows:

$$A^\alpha(t) := \partial^r A \,/\, \partial^{\iota_1} t_1 \cdots \partial^{\iota_k} t_k \,.$$

Lemma 9.3: Assume that A(t) is as above, and consider the following statements:

(1) $S(\{A(t), t\in V\}) = \Re^\rho$,
(2) For some $\Delta>0$, $S(\{A(t), t\in V\cap\delta\mathbf{Z}^k\}) = \Re^\rho$ for each $0<\delta<\Delta$,
(3) $S(\{A^\alpha(t_0), \alpha \text{ multiindex}\} = \Re^\rho$, for some t^0 in V.

Then, (1) is equivalent to (2), and (3) implies both. If A is analytic in t, all are equivalent.

Proof: We first prove that (1) is implied by (3). Assume that the space S in (1) is proper. Then there is a nonzero row vector η such that $\eta A(t)=0$ for all t. It follows that $\eta A^\alpha(t)=0$ for all t, in particular for $t=t^0$, so (3) follows. In the analytic case, the argument can be reversed.

We are only left to prove that (1) implies (2), the converse being trivial. If (1) holds, there is a finite set of matrices $A(\tau_1,\cdots,\tau_k)$ whose columns span $\Re^k$, with various choices of vectors $(\tau_1,\cdots,\tau_k)$. Appending these matrices to each other, we might as well consider the case in which there is only one such matrix (with maybe a larger σ and k). So assume that $A(\tau)$ has rank n at τ. Let $\Delta>0$ be such that A(t) is defined and has rank n whenever $|t_i-\tau_i|<\Delta$ for all i. Consider now any positive δ with $\delta<\Delta$. For each i, let s_i be an integer with $|s_i-\tau_i/\delta|<1$. Then, $A(s_1\delta,\cdots,s_k\delta)$ has rank n, as desired. Note that we have proved somewhat more: if q matrices are used to generate in (1), the same number is enough in (2).n

Theorem 9.1: If Σ is strongly transitive at ξ, then there is a $\Delta>0$ such that Σ is δ-transitive at ξ for all $0<\delta<\Delta$.

Proof: We want to show that $\mathbf{D}_\delta$ has full rank at ξ, for small δ. Since $\mathbf{D}_S$ is full rank, there are n vector fields $X_1,\cdots,X_n$ of the form $Ad_{b,\omega}(X_{a,u,v})$ such that the matrix $(X_1(\xi)\cdots X_n(\xi))$ has rank n, where we identify vectors at ξ with column vectors. Consider for each i the vector function b(t) of $t = (t_{i0},\cdots,t_{ik})$ (k depending on i) defined as follows. Assume that X_i has form $Ad_\pi(X_{\tau,u,v})$, $\tau\in\Re$, and that $\pi = \exp[\tau_1 Y_1]\cdots\exp[\tau_k Y_k]$. Then

$$b(t) := Ad_{\pi(t)}(X_{t_{i0},u,v}),$$

where $\pi(t):= \exp[t_{i1}Y_1]\cdots\exp[t_{iK}Y_k]$. This is defined for t near τ. The matrix A(t), $t = \{t_{ij}, i=1,\cdots,n, j=1,\cdots,k(i)\}$, is the matrix having the $b(t_i)$ as columns. Thus condition (1) in lemma 9.3 is satisfied (use τ's as particular values). The columns of A(t), $t\in\delta\mathbf{Z}$, are elements of $\mathbf{D}_\delta$. The result then follows from the lemma.n

Note that it is possible to refine the proof given here to conclude, as in [SS], that Δ can be chosen uniformly on neighborhoods of ξ, and hence uniformly on compacts.

9.1. A Lie condition

The above result can be complemented by sufficient conditions that give more precise estimates of those sampling periods for which transitivity is preserved. This section describes some such criteria. For simplicity, we shall assume that all continuous time systems are *complete* in this section. By a *parametrized vector field* (for short, *p.v.f.*) we shall mean a function

$$\underline{X} : \Re \to \Xi(\mathbf{M})$$

suc that $X_\tau := \underline{X}(\tau)$ is complete for each τ and $X_\tau(\xi)$ depends smoothly on $(\tau,\xi)\in\Re\times\mathbf{M}$. If $\underline{X}$ and $\underline{Y}$ are like this, $\underline{Z}:= [\underline{X},\underline{Y}]$ is by definition the parametrized vector field with $Z_\tau := [X_\tau,Y_\tau]$ for each τ. This defines a Lie algebra structure on the set of all parametrized vector fields. If $\underline{X}$ is a p.v.f., we may consider the new p.v.f. $\underline{X}'$ with

$$X_\tau'(\xi) := \frac{\partial X_\tau(\xi)}{\partial \tau} \ .$$

By commutativity (in local coordinates) of $\partial/\partial\xi$ and $\partial/\partial\tau$,

$[\underline{X},\underline{Y}]' = [\underline{X}',\underline{Y}] + [\underline{X},\underline{Y}']$ for all $\underline{X}$ and $\underline{Y}$.

Higher order derivatives $\underline{X}^{(N)}$ are defined by induction. Note that it follows from the above that

$$[\underline{X},\underline{Y}]^{(N)} = \sum_{i+j=N} \binom{N}{j} [\underline{X}^{(i)},\underline{Y}^{(j)}]. \tag{9.1}$$

If $\underline{X}$ is a p.v.f., we denote by ${}^k\underline{X}$ the new p.v.f. with ${}^kX_\tau = X_{k\tau}$. A *shift-invariant family* of parametrized vector fields is a set **F** of such $\underline{X}$, with the following property:

If $\underline{X}$ *is in* **F** *and* k *is a positive integer, then* ${}^k\underline{X}$ *is again in* **F** .

For any family of p.v.f.'s **F**, $\{\mathbf{F}\}_L$ is the Lie algebra (in the sense of the previous paragraph) generated by **F**. Note that $\{\mathbf{F}\}_L$ is again a shift-invariant family if **F** is. We consider also the new family, for each **F**:

$\mathbf{F}^\infty := \{\underline{X}^{(N)},\ N\geq 0,\ \underline{X}\in \mathbf{F}\}$.

If **F** is shift-invariant, then $\mathbf{F}^\infty$ is too. Indeed, assume given $\underline{Z} = \underline{X}^{(N)}$, and consider $\underline{Y} := {}^k\underline{X}$. Since $\underline{Y}$ is again in **F**, then ${}^k\underline{Z} = k^{-n}\underline{Y}^{(N)}$ is in $\mathbf{F}^\infty$, as desired.

Lemma 9.4: For any family **F** of p.v.f.'s, $\{\mathbf{F}_{LA}\}^\infty \subseteq \{\mathbf{F}^\infty\}_L$.

Proof: We proceed by induction on the formation of Lie brackets. Thus it is only necessary to establish that $[\psi,\omega]^{(N)}$ is in $\{\mathbf{F}^\infty\}_L$ whenever ψ and ω and all their τ-derivatives are already known to be there. But this follows from formula (9.1).n

Finally, let $\mathbf{F}^\infty_o$ be the set of all derivatives at 0 of the p.v.f.'s in **F**:

$\mathbf{F}^\infty_o := \{\ X^{(N)}_o,\ N\geq 0,\ \underline{X}\in \mathbf{F}\ \} \subseteq \Xi(\mathbf{M})$.

Lemma 9.5: For any shift-invariant family **F**, $\{\mathbf{F}_{LA}\}^\infty_o = \{\mathbf{F}^\infty_o\}_L$.

(Note that both sets represent Lie algebras of vector fields, *not* of p.v.f.'s.)

Proof: One inclusion is clear from the previous lemma, by evaluation at τ=0. We must then prove that $\{\mathbf{F}^\infty_o\}_L \subseteq \{\mathbf{F}_{LA}\}^\infty_o$. Since the latter contains $\mathbf{F}^\infty_o$, it is enough to prove that it is a Lie algebra. It is clearly a linear space. Now pick two elements there, say $X^{(i)}_o$ and $Y^{(j)}_o$, with $\underline{X}$ and $\underline{Y}$ in $\mathbf{F}_{LA}$. We must prove that their Lie bracket is in $\mathbf{F}^\infty_o$. Since **F** is a shift-invariant family, ${}^k\underline{X}$ is again in $\{\mathbf{F}\}_{LA}$. Thus it will be sufficient to show that $[X^{(i)}_o,Y^{(j)}_o]$ is a linear combination of the form

$$\sum_{k=1}^{N} r_k [{}^k\underline{X},\underline{Y}]^{(N)}_o.$$

We choose N:= i+j, and the coefficients r_k as follows. Consider the expansion

$$\sum_{k=1}^{N} r_k [{}^k\underline{X},\underline{Y}]^{(N)} = \sum_{=1}^{N} \binom{N}{} \rho \ ,$$

where ρ is the p.v.f.

$$\rho = \sum_{k=1}^{N} r_k k \ [^k(\underline{X}(\ ),\underline{Y}^{(N-\ )}].$$

Evaluating at τ=0, we obtain that

$$(\rho\ )_\tau = (\sum r_k k\ )[X_o(\ ),Y_o^{(N-\ )}].$$

By a VanDermonde argument, we may choose the sequence of reals $\{r_k\}$ so that all $\sum r_k k\ = 0$ except for the term with =i, and $=(\binom{N}{i})^{-1}$ for that term. This gives the desired equality.n

For any family **F** and each δ>0 we denote $\mathbf{F}_\delta := \{X_\delta\ ,\ \underline{X} \in \mathbf{F}\}_L \subseteq \Xi(\mathbf{M})$. Note that $\{\mathbf{F}_{LA}\}_\delta = \{\mathbf{F}_\delta\}_L$.

Lemma 9.6: If **F** is a shift-invariant family of p.v.f.'s such that the distribution $\mathbf{F}_o^\infty$ has rank n at ξ then there is a Δ>0 such that $\mathbf{F}_\delta$ has rank n at ξ for each 0<δ<Δ.

Proof: Assume that X,Y,··· is a set of n vector fields such that

$$\det\,(X_o^{(N)}(\xi),Y_o^{(M)}(\xi),\cdots) \neq 0\ . \tag{9.2}$$

Introduce the n×n matrix function of $\tau \in \Re$, $A(\tau) := (X_\tau(\xi),Y_\tau(\xi),\cdots)$. By equation (9.2), condition (3) in lemma 9.3 is satisfied. Applying the lemma, condition (2) gives a Δ such that the columns of all possible A(kδ) also generate $\Re^n$, for any 0<δ<Δ. But these columns are in $\mathbf{F}_\delta$, because **F** is shift-invariant.n

Corollary 9.7: If **F** is a shift-invariant family of p.v.f.'s such that $\{\mathbf{F}_o^\infty\}_L$ has rank n at ξ then there is a Δ>0 such that $\{\mathbf{F}_\delta\}_L$ has rank n at ξ for each 0<δ<Δ.

Proof: By lemma 9.5, $\{\mathbf{F}_{LA}\}_o^\infty$ has rank n at ξ. Applying lemma 9.6 to $\mathbf{F}_{LA}$, we conclude that $\{\mathbf{F}_{LA}\}_\delta$ has rank n for δ small enough. But as remarked earlier, this is the same as $\{\mathbf{F}_\delta\}_L$.n

The above is now applied to sampling. Assume a complete continuous time system is given. As in section 8, we assume that a control value "0" in int**C** has been fixed. For each j = 1,···,m, positive integer i, and $t \in \Re$, consider the p.v.f. $\underline{X}^{[i,j]}$ defined by:

$$(X^{[i,j]})_\tau(\xi)\ :=\ \frac{\partial exp[i\tau X_o]exp[-i\tau X_v](\xi)}{\partial v}\,|_{\,v=0}(e_j). \tag{9.3}$$

Let **F** be the family consisting of all such p.v.f.'s; **F** is shift-invariant because $\underline{X}^{[i,j]} = \underline{X}^{[i\ \ ,j]}$. Further, for each fixed δ>0, $\mathbf{F}_\delta$ is nothing else than the set of all elements $b_{i,j}$ in equation (8.3) (with respect to the action Σ_δ). It follows that Σ is δ-transitive if $\{\mathbf{F}_\delta\}_L$ has rank n at ξ, or, from corollary 9.7,

Theorem 9.2: If $\{\mathbf{F}_o^\infty\}_L$ has rank n at ξ then there is a Δ>0 such that Σ is δ-accessible at ξ whenever 0<δ<Δ.

Now consider a control-linear system, i.e. a continuous time system for which the right hand side P(x,u) of (7.1) has the form $f(x)-\sum u_i g_i(x)$, and **U** = is $\Re^m$, **C** contains the origin in its interior. We take $X_v = f - \Sigma v_i g_i$, $v \in \Re^m$, $X_o = f$, the obvious choice of e_i's, and the element "0" in int**C** as the origin. From the theory of Lie series (see [GO], or verify by direct computation), the following

Taylor expansion holds:

$$(X^{i,j})_\tau \simeq \sum_{k=1}^{\infty} i^k ad_f^{k-1}(g_j)\tau^k/k!, \tag{9.4}$$

where $ad_\alpha(\beta) := [\alpha,\beta]$. (This formula was used before in the context of sampling in [NC], where a weaker version of the corollary given below was conjectured.) It follows that

$$(X^{[i,j]})_o^{(N)} = i^N ad_f^{N-1}(g_j) , \tag{9.5}$$

so that $\{\mathbf{F}_o^\infty\}_L$ is precisely the algebra L_o that appears in lemma 6.2, because the lie algebra generated by the $ad_f^i(g_j)$ equals L_o. For any fixed $\delta>0$, introduce the Lie algebra of vector fields L_δ which is given as in equation (8.2) for the discrete time action Σ_δ . We conclude from the above:

Corollary 9.8: Consider a continuous time complete control-linear system. If $\{L_\delta\}_D$ has full rank at ξ then Σ is δ-transitive at ξ. If Σ satisfies the strong Lie rank condition at ξ then there is a $\Delta>0$ such that for each $0<\delta<\Delta$, $\{L_\delta\}_D$ has full rank at ξ.

Note that the last conclusion is a very particular case of a fact that we had already proved above (c.f. theorem 9.1). The first conclusion is rather interesting, however. Based on the discrete-time theory, there is no reason to expect the given distribution to have full rank, even if Σ is δ-transitive and analytic. But the corollary says that when dealing with sampling problems, *at least for small enough* δ this will indeed be true. Further, in particular examples it may be easier to check L_δ. Note that formally, if we let J^δ be the following (partial) linear operator on $\Xi(\mathbf{M})$:

$$J^\delta := \sum_{k=1}^{\infty} ad_f^{k-1}\delta^k/k!,$$

then $b_j = J^\delta g_j$, and

$$L_\delta = \{Ad^k_{\exp[-\delta f]} J^\delta g_j \mid k\geq 0, j=1,\cdots,m\}_L .$$

These formulas are used to provide a more explicit result for bilinear systems in [SO3].

9.2. A worked example

It seems appropiate to work out in detail a nontrivial example. We take the "rotating satellite" model (momentum control only) described by [BA], [BO]. It is not hard to show that such systems are δ-transitive (and even controllable) from all ξ, for $\delta=\delta(\xi)$ small enough, provided that they are strongly transitive to start with. (The above references give an algebraic characterization of this latter property.) For simplicity, we choose a rigid body with one symmetry axis, and take the simplest possible coefficients. This is the control-linear system in $\Re^3$ with m=1 and defining vector fields:

$$f = yz\partial/\partial x - xy\partial/\partial z,\ g = \partial/\partial x + \partial/\partial y + \partial/\partial z .$$

Thus in coordinates $f = (yz,0,-xy)'$ and g is the constant vector $(1,1,1)'$. (Prime indicates transpose.) From now on, we write everything in coordinates. We wish to find explicitely the generators $b_{i,j}$ for all δ. Consider then the p.v.f. $\underline{X} := \underline{X}^{[1,1]}$ in equation (9.3). Since the system is analytic, the following series expansion converges at least for small τ:

$$X_\tau = \sum_{N=1}^{\infty} ad_f^{N-1}(g)\tau^N/N! \,. \tag{9.6}$$

Let h be the linear vector field (z,0,-x), and consider its Jacobian

$$H := \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{matrix} \,.$$

Since g is constant, $ad^i_h(g) = (-H)^i g$. If $W = ad^i_h(g)$ then $L_W y^j = 0$ for all j and all $i \geq 1$, where L_W denotes the corresponding Lie derivative. Since f = y.h, it follows that

$$[f,g] = - yHg - h \,, \quad ad^i_f(g) = y^i(-H)^i g \quad \text{for } i \geq 2 \,.$$

So the expression in equation (9.6) becomes

$$-h\,\tau^2/2 + (1/y)\,(-H)^{(\tau y)} \,,$$

where $H^{(r)}$ is the matrix $\sum^{k=1}) r^k H^{k-1}/k!$. Using linear algebra to evaluate $(-H)^{(\tau y)}$, we conclude that we must study the following p.v.f.:

$$X_\tau = [-\tau^2 z/2+(1/y)\{\sin(\tau y)+\cos(\tau y)-1\},\tau,\tau^2 x/2+(1/y)\{\sin(\tau y)+1-\cos(\tau y)\}]'.$$

Let $\underline{Y}$ be the p.v.f. ${}^2\underline{X}$ (substitute t:=2t in above equation), and $\underline{Z}$ the p.v.f. ${}^3\underline{X}$ (substitute t:=4t). Consider now the matrix whose columns are $Y_\tau-4X_\tau$, $Z_\tau-9X_\tau$, and the Lie bracket of these. The determinant of this matrix can be computed easily with a symbolic manipulation system and one obtains

$$96\,\tau^3\,(\cos(\tau y) - 1)^3 / y^2 \,.$$

Thus the determinant is nonzero whenever τy is not a multiple of 2π. Pick now *any* $\delta>0$. For any point $(x,y,z)'$ such that δy is not a multiple of 2π, this shows that the system is δ-transitive at $(x,y,z)'$. But if (x,y,z) is now any point in $\Re^3$, we can always δ-reach from this state one with 'good' y, since the y-coordinate satisfies dy/dt = u. It follows that the system *is transitive at every* ξ *in* $\Re^3$. (Instead of the argument in terms of reaching points with 'good' y, one could search directly for more generators in order to establish the conclusion at every point.)n

9.3. The one dimensional case.

As an easy example, we consider complete continuous time analytic systems with $M = \Re$. Although elementary, this case provides some feeling for the kinds of pathologies that may occur. We let

$$N := \{(\xi,\delta) \mid \xi \in \mathbf{M} \text{ and } \mathbf{D}_\delta \text{ has full rank at } \xi\},$$

and B:= complement of N in $M\times\Re_+$. A point z in M is *invariant* if f(z,u) = 0 for all u; this is equivalent to $\mathbf{D}_T$ having rank 0 there. In that case, both {x<z} and {x>z} are invariant under the dynamics (7.1), so each of these sets gives rise to a new system (7.1) with state space again (diffeomorphic to) $\Re$. Thus B is the union of the corresponding sets B', B'' obtained from each of these, and of the set $\{(z,\delta),\ \delta>0\}$, and transitivity can be studied for each part separately. We shall assume from now on, therefore, that (7.1) has no invariant points. Call B *trivial* if B is empty or if it equals $M\times\Re_+$, and consider the δ-projection

$$C = \{\delta \mid (x,\delta)\in B, \text{ some } x\}. \tag{9.7}$$

These are the sampling periods for which (7.1) is not *globally* transitive, in the sense that $\mathbf{D}_\delta$ has full rank at all ξ. We shall prove:

Theorem 9.3: (M=$\Re$ and no invariant points.) If *B* is nontrivial, then *C* is a discrete subset of $\Re$.

In particular, the system is globally transitive for all small enough sampling times (if nontrivial). Theorem 9.3 will follow from a more detailed study of the following sets. For any two (complete) vector fields X, Y, write

$$B(X,Y):= \{(x,\delta) \mid \exp[k\delta X](x)=\exp[k\delta Y](x), \text{ all integers } k\}. \tag{9.8}$$

Take two vector fields of the form $X = X_u$ and $Y = X_v$, u,v in **C**. Assume that (x,δ) is not in *B*(X,Y), so that, for the system Σ_δ , $g_b(x,\omega) \neq g_b(x,\psi)$ for some k>0, where $\omega=u^k$, $\psi=v^k$, and $b\in \mathbf{A}^*$. Since **C** is connected, the image of $g_b(x.\cdot)$ contains a nontrivial interval. Thus Σ is transitive at x, and x is not in *B*. Conversely, assume that (x,δ) belongs to all the *B*(X,Y) of the above form. Then the orbit $O_\delta(x)$ of x under Σ_δ is included in the discrete set $\{\exp[k\delta X](x), k=\text{integer}\}$, for any fixed X, and so (x,δ) is in *B*. We conclude that

$$B = \cap\{B(X,Y),\ X=X_u, Y=X_v,\ u,v \text{ in } \mathbf{C}\}. \tag{9.9}$$

It follows that it is sufficient to prove theorem 9.3 for the sets of type *B*(X,Y). We identify vector fields with their coordinates with respect to the natural global chart in $\Re$.

Lemma 9.9: Assume that *B* is nontrivial. Then, for any X,Y as above, X(x)Y(x)>0 for all x.

Proof: An x such that f(x,u)=0 for some u is an *equilibrium point*. Let x be any such point. Since x is invariant, $f(x,v)\neq 0$ for some v in **C**. It follows that $\exp[\delta X_u](x) = x \neq \exp[\delta X_v](x)$ for all $\delta>0$, so (x,δ) is not in *B*, for any $\delta>0$. We claim that there are no equilibrium points. Indeed, assume that f(x,u) = 0 for some (x,u), and replace **C** by a compact set which contains this u and is included in the closure of the original **C**. Pick any non-equilibrium point y<x in M, and let Z:= inf{z>y | z equilibrium point}. By compactness of **C**, z is itself an equilibrium point, so $z\neq y$. Pick v,v′ such that f(z,v)=0 and f(z,v′)≠0. By definition of z, f(a,v)≠0 and f(a,v′)≠0 for all a in the interval [y,z). Compare the trajectories $\exp[tX_v](y)$ and $\exp[tX_{v'}](y)$. Assume first that f(y,v)>0. Then the v-trajectory converges to z, as $t \to \infty$, while the v′-trajectory does not. Same conclusion for f(y,v)<0 if one takes the limit as $t \to -\infty$ instead. It follows that, for every $\delta>0$, (y,δ) is not in *B*(X,Y), for $X=X_v$ and $Y=X_{v'}$, and hence also for some v,v′ in the original **C**. A similar argument holds if y>x. So the existence of an equilibrium point implies that *B* is empty, contradicting nontriviality. So f(x,u)≠0 for each x and all u, and so (recall **C** is connected) the f(x,.) indeed have constant sign.n

We are thus led to the study of the sets *B*(X,Y) with, say, X(x)>0 and Y(x)>0 for all x. Call such vector fields "positive". Conversely, any such pair {X,Y} gives rise to a system (7.1) with *B* = *B*(X,Y); this is a consequence of the following characterization, which is easy to obtain but very useful:

Lemma 9.10: Let X,Y be positive (analytic, complete) vector fields. There is then an analytic function g: $\Re \to \Re$, with derivative (dg/dt)(t)>-1 for all t and such that, for some diffeomorphism b(.),

$$g(t+k\delta)=g(t) \text{ for all integers } k \text{ iff } (b(t),\delta)\in B(X,Y), \tag{9.10}$$

for any t in $\Re$ and any $\delta>0$. Further, g is constant iff X=Y. Conversely, given any analytic g with derivative bounded below, and any (strictly increasing) diffeomorphism b, there exists a continuous time system, and in particular there are positive X,Y, such that $B = B(X,Y)$ and 9.10 holds.n

Proof: Let a(t):= exp[tX](0), b(t):= exp[tY](0), both analytic and strictly increasing. Let c:= a^{-1}, d(t):= c(b(t)). Define

$$g(t) := d(t) - t.$$

Since c(.) and d(.) are increasing, g has derivative > -1. Let x be any state, and $t_o := b^{-1}(x)$. Note that exp[tX](x) = a(c(x)+t), exp[tY](x) = $b(t_o+t)$. So these two trajectories are equal at t iff $g(t_o+t) = g(t_o)$. Further, since g(0)=0, g is constant iff g=0, which happens iff a(t) = b(t) for all t. This proves the first part of the lemma.

Conversely, assume given g and a diffeomorphism b. Multiplying g by a constant, we may assume that (dg/dt)(t) > -1/2 for all t. Let **U** = **C** = $\Re$, and introduce for each u the function $d_u(t) = (\sin^2 u)g(t)+t$; note that the derivative of d_u is >1/2, for all u. Thus $a_u(t) := b(d_u^{-1}(t))$ is well defined (and analytic). We may then introduce

$$f(x,u) := (da_u/dt)(a_u^{-1}(x)).$$

Let X:= f(·,0), X_u:= f(·,u) for u>0, and Y=f(·,1). Reversing the previous argument shows that, for any u>0, $\exp[tX_u](b(x))$ = exp[tX](b(x)) iff g(x+t) = g(x) (independent of u). For this system, then, $B(X,X_u) = B(X,Y)$ for all u>0. Thus B = B(X,Y), and 9.10 holds.

Fix now a function g satisfying the properties in lemma 9.10, and denote by B(g) the set of pairs (t,δ) with δ>0 such that g(t+kδ) = g(t) for all integers k. Also, let C(g) be the projection of B(g) in the δ-coordinate.

Lemma 9.11: Let (t,δ), (t′,δ′) be in B(g). Then,

$$|g(t)-g(t')| \le |h\delta+k\delta'| \qquad (9.11)$$

for any integers h,k such that $h\delta+k\delta' \neq 0$.

Proof: Consider any such h,k, and let r:= |hδ+kδ′|. For suitable integers a,b, r = bδ′-aδ. Without loss of generality, take m:= g(t)-g(t′) to be positive. Assume that r<m; there is then some integer s such that t′-t-m < -sr < t′-t. Let c:=as, d:=bs. We then have

$$0 < (t'+d\delta')-(t+c\delta) < m,$$

and (by hypothesis)

$$g(t+c\delta)-g(t'+d\delta') = g(t)-g(t') = m.$$

By the mean value theorem, this contradicts dg/dt>-1.n

Corollary 9.12: If δ and δ′ are rationally independent, and if (t,δ), (t′,δ′) are in B(g), then g(t) = g(t′).

Corollary 9.13: Assume that C(g) has a limit point in $\Re$. Pick (t′,δ′) and (t″,δ″) in B(g). Then g(t′) = g(t″).

Proof: We shall use the following observation twice: Assume that $\{a_i\}$ is a converging sequence of distinct real numbers, and let f be any nonzero real number. There are then (i) a subsequence $\{a_j\}$ of $\{a_i\}$, and (ii) sequences $\{b_j\}$, $\{c_j\}$ of integers, such that the numbers $e_j := b_j a_j +$

c_jf are all nonzero and $\{e_j\}$ converges to zero. [Proof: assume that $a_i \to a$. Let b_i, c_i be integers such that $b_i \neq 0$ and $|b_i a + c_i f| < 1/i$ (if a=0 use just $c_i=0$, otherwise consider the group generated by a and f). Now pick any a_j, $j=j_i$, such that the inequality is still satisfied and $e_j \neq 0$.] Assume that $\{(t_n,\delta_n)\} \subseteq B(g)$, with all δ_n distinct and converging to δ (which may be zero). Applying the above observation with f:= δ′, we conclude --for a subsequence of the (t_n,δ_n)-- that the $b_i\delta_i + c_i\delta'$ are are all nonzero and converge to 0. By lemma 9.11, $|g(t_i)-g(t')|$ also converges to 0. Taking in turn a subsequence of the $\{\delta_i\}$, and f:= δ″, we can also conclude that $|g(t_i)-g(t'')|$ converges to zero. so g(t′)=g(t″), as desired.n

Proposition 9.14: If g is nonconstant then *C*(g) is discrete as a subset of $\Re$.

Proof: Assume that there are infinitely many distinct $\delta_i \leq K$, with (t_i,δ_i) in *B*(g). By corollary 9.13, there is a constant c such that $g(t_i+k\delta_i) = c$ for all i and all integers k. Let $t_i' = t_i \bmod(\delta_i)$ such that $t_i' \in [0,K]$. Thus $g(t_i') = c$ and $\{t_i'\}$ is bounded. Since g is nonconstant and analytic, there are only finitely many t_i'. But then there are infinitely many $t_i'' := t_i' + \delta_i$ --since there are infinitely many δ_i-- and these are also bounded, with $g(t_i'')=c$. This again contradicts nonconstancy of g.n

Theorem 9.3 now follows from proposition 9.14 and lemma 9.10. Actually, we can prove somewhat more. Since *B* is analytic, each subset with constant δ also is, so *B* is the union of a discrete set and a union of lines $L_i := \{(x,\delta_i)$, x in M}. So g is periodic with period δ_i, for all i. Since periods form a subgroup, g nonconstant implies that the δ_i are integer multiples of some fixed δ>0. So the nondiscrete part of *B* is of the form

{(x,kδ), x in **M**, k = integer}.

The set *C*(g) may be rather complicated. Consider the following example. Take a sequence of numbers $\{a_n\}$ such that

$$\sum (a_n)^{-1} < 1/\pi, \text{ and} \qquad (9.12)$$

$$\cos(\pi x/a_n) > 1-2^{-n} \text{ if } x \in [-n,n]. \qquad (9.13)$$

Now let $g_n(x) := \cos(\pi x/a_n)$ and g:= (infinite) product of the g_n. This product is well defined because there is by (9.13) normal convergence on compacts, and g is indeed analytic. Further, consider its derivative

$$g' = \sum (g/g_n).g_n'.$$

Since $|g/g_n|<1$ and $|g_n'|<\pi/a_n$, also $|g'|<1$. The zeroes of g are those of its factors, i.e., the union of the sets

$$\{(t_n+ka_n), k = \text{integer}\},$$

where $t_n := a_n/2$. So all a_n are in *C*(g). If (t,δ) is in *B*(g) and δ is not rationally dependent with some a_n, then corollary 9.12 says that g(t) = 0, so δ= some a_n, a contradiction. Thus *C*(g) contains all the a_n and no other rationally independent numbers. For constructing sequences $\{a_n\}$ as above, consider the following argument: Let $\{b_n\}$ be such that $\cos(\pi x/a) > 1-2^{-n}$ whenever x is in [-n,n] and $a>b_n$ (just let b_n be such that $\cos(\pi n/b_n) > 1-2^{-n}$). Now pick any sequence $\{a_n\}$ satisfying condition (9.12). and such that $a_n>b_n$ for all n. Note that, in particular, one could choose the a_n to be rationally independent.

Remark 9.15: One of the most useful tools in the continuous time theory is (the positive form of)

Chow's theorem, which implies for analytic systems that the positive-time reachable set has nonempty interior whenever O(x) does. In fact, the term "accessibility" is used interchangeably with transitivity in that context, with the first term refering to the property of positive time reachable sets. Here, however, it may happen that $\mathbf{D}_\delta$ is full rank at ξ but that the set of points

$$A_\delta(\xi) := \{\exp[\delta X_{u_1}]\circ\cdots\circ\exp[\delta X_{u_k}](\xi) \mid u_i\in \mathbf{C}\}$$

(only positive-time motions, no exp[-δX]'s allowed,) has an empty interior. We construct an analytic continuous time system with $\mathbf{M}=\Re$ where this happens. We first obtain an analytic function $g:\Re\to\Re$ whose derivative is bounded below, and for which a pair (x,δ) satisfies the condition

$$g(x+k\delta) = g(x) \quad \text{for all } k\in \mathbf{Z}$$

iff it satisfies

$$x=2r\pi,\ \delta=2s\pi,\ r,s\in \mathbf{Z} \text{ and } s \text{ does not divide } r .$$

As above, this gives rise to a system for which $\mathbf{D}_\delta(x) = \{0\}$ iff (x,δ) is of this form. Further, assume that this g is such that, with $x_0=2\pi$, $g(x_0+2k\pi) = g(x)$ for all *positive* integers k. In that case we can conclude both that $O_{2\pi}(x_0)$ has interior and that $A_{2\pi}(x_0) = \{2k\pi, k\geq 1\}$. An example of a g like this is

$$g(x):=(\sin x)/x .$$

This example can be modified to obtain one where even the orbit under Σ_δ equals **M** *for all* (x,δ) but such that still $A_\delta(x_0)$ has empty interior for some x_0. For this, take the above g and introduce a $g_1(x):= \sum 2^{-n}g^2(x+2\pi n)$, the sum over $n\geq 0$. Again $x_0=\delta=2\pi$ serves as a counterexample.n

10. References

[BA] Baillieul,J., "Controllability and observability of polynomial dynamical systems," *Nonl.Anal.,TMA* **5**(1981):543-552.
[BL] Bar-Ness,Y. and G.Langholz, "Preservation of controllability under sampling," *Int.J.Control* **22**(1975):39-47.
[BO] Bonnard,B., "Controle de l'attitude d'un satellite," Report #8019, Univ.Bordeaux, Oct.1980.
[BC] Brickell,F. and R.S.Clark, *Differentiable Manifolds*, Van Nostrand, New York, 1970.
[FN] Fliess,M. and D.Normand-Cyrot, "A group-theoretic approach to discrete-time nonlinear controllability," *Proc.IEEE Conf.Dec.Control*, 1981.
[GH] Gibson,J.A. and T.T.Ha, "Further to the preservation of controllability under sampling," *Int.J.Control* **31**(1980):1013-1026.
[GO] Goodman,R., "Lifting vector fields to nilpotent Lie groups," *J.Math. Pures et Appl.* **57**(1978): 77-86.
[HK] Hermann,R. and A.J.Krener, "Nonlinear controllability and observability," *IEEE Trans.Autom.Ctr.* **22**:728-740.
[IS] Isidori, A., *Nonlinear Control Systems: An Introduction*, Springer, Berlin, 1985.
[JA] Jackubczyk,B., "Invertible realizations of nonlinear discrete time systems," *Proc.Princeton Conf.Inf.Sc.and Syts.* (1980):235-239.
[JN] Jackubczyk,B., and D.Normand-Cyrot, "Orbites de pseudo-groupes de diffeomorphismes et commandabilitB des systemes non linearires en temps discret," *C.R.Acad. Sc. Paris*, **298-I**(1984): 257-260.
[JC] Jacobson,N., *Lie Algebras*, Dover, N.Y., 1979.

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

[KL] Kalouptsidis,N., *Accessibility and Stability Theory of Nonlinear Control Systems*, Dr.Sci. Thesis, Washington University, 1977.

[KS] Kupka,I., and G.Sallet, "A sufficient condition for the transitivity of pseudo-groups: Application to system theory," *J. Diff. Eqs.* **47**(1973): 462-470.

[KR] Krener,A., "(Adf,g), (adf,g) and locally (adf,g) Invariant and Controllability Distributions," preprint, UC-Davis, 1984.

[LO] Lobry,C., "Bases mathematiques de la theorie de systemes asservis non lineaires," Report #7505, Univ.Bordeaux, 1976.

[MO] Mohler,R.M., *Bilinear Control Processes*, Academic Press, NY, 1973

[NC] Normand-Cyrot, Dorothee, *Theorie et Pratique des Systemes Non Lineaires en Temps Discret*, These de Docteur d'Etat, Univ. Paris-Sud, March 1983.

[ST] Stefan, P., "Attainable sets are manifolds," preprint, Univ. of Wales, 1973(?).

[SO] Sontag,E.D., *Polynomial Response Maps*, Springer, Berlin-NY, 1979.

[SO1] Sontag,E.D., "Remarks on the preservation of various controllability properties under sampling," in *Developpement et Utilisation d'Outils et ModeAles MatheBmatiques en Automatique, Analyse de SysteAmes et Traitement de Signal*, Coll. CNRS, RCP 567, Belle-Ile, 1983, pp.623-637.

[SO2] Sontag,E.D., "A concept of local observability," *Systems and Control Letters* **5**(1984): 41-47.

[SO3] Sontag,E.D., "An eigenvalue condition for sampled weak controllability of bilinear systems," to appear.

[SS] Sontag,E.D. and H.J.Sussmann, "Accessibility under sampling," *Proc. IEEE Conf. Dec. and Control, Orlando, Dec. 1982*.

[SU1] Sussmann,H.J., "Orbits of families of vector fields and integrability of distributions," *Trans.AMS* **180**(1973):171-188.

[SU2] Sussmann,H.J., "Lie brackets, real analiticity, and geometric control," in *Differential Geometric Control theory* (R.W.Brockett, R.S.Millman, and H.J.Sussmann, eds.), Birkhauser, Boston, 1983.

[SJ] Sussmann,H.J. and V.Jurdjevic, "Controllability of nonlinear systems," *J.Diff.Eqs.* **12**(1972):95-116.

Table of Contents