This file contains the original version of the following report:

Integrability of certain distributions associated to actions on manifolds and an introduction to Lie-algebraic control

Rutgers Center for Systems and Control
(SYCON) Report 88-04, Eduardo D. Sontag, July 88.

This report consists of two parts: pages 1-41 contain a paper published as *Nonlinear Controllability and Optimal Control* (H.J. Sussmann, ed.), pp. 81-131, Marcel Dekker, NY 1990. (Proceedings of the Conference on Nonlinear Control at Rutgers, May 87.) Pages 42-50 apply the results to the control of continuous time systems; this is an exposition of some of the basic results of the Lie algebraic accessibility theory.

Erratum: there is (at least) one error in the report, namely, the hypotheses of Theorem 5 should include the assumption that the vector fields are everywhere defined.

INTEGRABILITY OF CERTAIN DISTRIBUTIONS ASSOCIATED TO ACTIONS ON MANIFOLDS AND APPLICATIONS TO CONTROL PROBLEMS

Eduardo D. Sontag [1]
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
(201)932-3072 – *sontag@fermat.rutgers.edu*

ABSTRACT

Results are given on the integrability of certain distributions which arise from smoothly parametrized families of diffeomorphisms acting on manifolds. Applications to control problems and in particular to the problem of sampling are discussed.

1 Introduction

The first objective of this paper is to provide a tutorial introduction to integrability results for distributions –or "singular vector bundles",– on manifolds. These distributions arise from actions of smoothly parametrized families of diffeomorphisms. Such results generalize Frobenious' Theorem in two ways: they deal with diffeomorphisms not necessarily associated to flows, and they do not require the distribution to be nonsingular. Results along these lines are important in various areas of control theory, and they originated in the work of Hermann ([7]) and subsequent research by Sussmann ([19]) and Stephan ([14]) in the early 70's, who removed the nonsingularity assumption and showed that a form of the theorem still holds in the singular case . We shall present an abstract version which summarizes all that is needed for various applications. Our result is more abstract in that it deals with rather general classes of diffeomorphisms, not just those arising from flows as in [19] and [14], but the main ideas of the proof are very similar to the ones in the former reference. Following [19], we also show how more special results due to Nagano, Lobry, and others can be obtained as consequences of the general theorem.

Our interest in actions different from those arising from flows is due to the possibility of applying the obtained results in the study of *discrete time invertible systems.* The general theory regarding controllability questions for discrete time systems remained until recently much weaker than that possible in the more classical continuous time case. In principle, noninvertibility of transition maps in discrete time implies that semigroups appear where groups would appear in the continuous case, so less algebraic structure is available. Another difficulty is that no analogue of the infinitesimal information obtained by taking derivatives with respect to time is available for difference equations. One avoids the first of these difficulties by restricting attention to invertible systems, for which by assumption transition maps are invertible. The lack of infinitesimal information is dealt with by substituting derivations with respect to control values, assuming that, as is often the case, there is a differentiable structure in the control value

[1]Research supported in part by US Air Force Grant 0247.

set. This gives rise to an action in the sense of this paper, and the results can then be applied to these systems.

Although invertibility is in principle a strong assumption in the context of general discrete time systems, it is the case that for systems that result from *sampling*, this assumption is always satisfied. Recall that sampling is the process under which the state of a continuous time system is measured at discrete instants, and control actions are taken also at discrete instants. Under such a process, the obtained transition maps as observed at the sampling times give rise to an invertible model. This is analogous to the situation in classical dynamical systems, where one studies time-one diffeomorphisms and Poincaré maps associated to differential equations. The paper [8] introduced the idea of studying invertible discrete nonlinear control systems. For other work in this area, and more in the spirit of the present paper, see for instance [5], [13], [9], [18], [10], and related papers. Applications to the sampling problem are presented in the last section.

2 The Orbit Theorem

2.1 Vector fields and diffeomorphisms

For differential geometric definitions and elementary results, our main reference is [2]. Other excellent references are for instance [3] and [1]. We deviate from standard terminology mainly in allowing non-second countable manifolds and singular distributions, as described later. Much of the terminology about singular distributions is borrowed from [19]; in fact, many of the proofs on integrability are either taken from that reference or are easy generalizations of the proofs there.

Throughout this paper, *manifold* will mean smooth, i.e. C^∞, and paracompact manifold. Thus manifolds are Hausdorff, not necessarily connected, but each connected component is second countable. (Recall that a paracompact space is one for which every open covering has a locally finite refinement. Thus a space is paracompact iff each connected component is. Every locally compact and second countable Hausdorff space, in particular thus every connected second countable manifold, is paracompact; see for instance [2], section V.4.) An *analytic* manifold means a real-analytic paracompact manifold. The tangent space to the manifold $\mathcal{M}$ at the point ξ will be denoted by $T_\xi\mathcal{M}$, and $T\mathcal{M}$ is the tangent bundle to $\mathcal{M}$.

By a *submanifold* $\mathcal{N}$ of a manifold $\mathcal{M}$ we mean an immersed submanifold. We do not require $\mathcal{N}$ to be an embedded (or regular) submanifold, nor to be connected or even second countable, even if $\mathcal{M}$ is. It will turn out that this generality is needed in order to establish a number of the results. Thus, for instance, if $\mathcal{M} = \mathbb{R}$ with the usual topology, we may consider the submanifold $\mathcal{N} = \mathbb{R}_{\text{discr}}$ which equals $\mathbb{R}$ as a set but which is endowed with the discrete topology. This is a manifold of dimension 0, and it has uncountably many components since each real number is a component. It is pathological in that it equals $\mathcal{M}$ as a set even though it has lower dimension.

A central issue in controllability is that of determining when certain submanifolds of reachable sets fill an open subset of the ambient space, and to be able to determine this based only on algebraic computations involving the relative dimensions of $\mathcal{M}$ and $\mathcal{N}$. Thus one would like to have the property that the submanifold $\mathcal{N}$ has a nonempty interior with respect to the topology of $\mathcal{M}$ precisely when the dimensions of the two coincide. When the dimensions do coincide then this does indeed hold, but as the example $\mathcal{M} = \mathbb{R}, \mathcal{N} = \mathbb{R}_{\text{discr}}$ illustrates, the converse is false in general. However, *if* $\mathcal{N}$ is know to be a second countable manifold, then indeed it cannot have lower dimension than $\mathcal{M}$ unless it has measure zero in $\mathcal{M}$: see for instance a proof in [3], proposition 8.5.6. Thus a central objective of our study will be to give general results that insure that certain submanifolds are second countable. Equivalently, because of the paracompactness assumption and the assumption of second countability that we shall make on the state spaces, we will be interested in determining when these submanifolds have only countably many components (in the submanifold topology).

Given two manifolds $\mathcal{N}_1, \mathcal{N}_2$, the product manifold is denoted $\mathcal{N}_1 \times \mathcal{N}_2$. This is the cartesian product of the two sets, endowed with the usual differentiable structure: typical coordinate functions are $(\varphi_1(\xi_1), \varphi_2(\xi_2))$ for each set of local coordinates φ_1, φ_2 for $\mathcal{N}_1$ and $\mathcal{N}_2$ respectively. Note that there is a natural identification $T_{(\xi_1,\xi_2)}(\mathcal{N}_1 \times \mathcal{N}_2) \simeq T_{\xi_1}(\mathcal{N}_1) \times T_{\xi_2}(\mathcal{N}_2)$. When $\mathcal{N}_2$ is a discrete manifold, that is to say a manifold of dimension zero, we can identify $\mathcal{N}_1 \times \mathcal{N}_2$

with a disjoint union of copies of $\mathcal{N}_1$ (one for each element of $\mathcal{N}_2$), and we may think of each tangent space $T_{(\xi_1,\xi_2)}(\mathcal{N}_1 \times \mathcal{N}_2) = T_{\xi_1}(\mathcal{N}_1) \times 0$ as just $T_{\xi_1}(\mathcal{N}_1)$. Such products of one manifold by another one which is discrete will appear in some of the examples to be considered.

We shall allow most concepts, such as vector fields and diffeomorphisms, to be partially defined. This is necessary because many of these objects will be typically derived from flows of vector fields, and solutions of differential equations are in general only locally defined. More precisely, by a *vector field X on $\mathcal{M}$* we shall mean a smooth vector field (smooth section of the tangent bundle) defined on some open subset $\mathcal{V}_X$ of $\mathcal{M}$; we denote by $\Xi(\mathcal{M})$ the set of all such X. For a vector field X and each $\xi \in \mathcal{V}_X$, we denote by $X(\xi)$ the value of X at $\xi \in \mathcal{M}$; this is a vector in the tangent space $T_\xi\mathcal{M}$ of $\mathcal{M}$ at the point ξ. (The notation X_ξ is more standard.)

Given X and Y in $\Xi(\mathcal{M})$, we let $\mathcal{V}_X \cap \mathcal{V}_Y = \mathcal{V}$ and we define the Lie bracket $[X, Y]$ as the Lie bracket of X restricted to $\mathcal{V}$ and Y restricted to $\mathcal{V}$. If $\mathcal{V}$ is empty, the bracket is undefined. Similarly for the sum of X and Y, and products by constants. Thus, the subset Φ of $\Xi(\mathcal{M})$ will be said to be *involutive* if $[X, Y]$ is in Φ whenever the product is defined and X, Y are in Φ, and we say that Φ is a *subspace* of $\Xi(\mathcal{M})$ if for all X, Y in Φ and all $r \in \mathbb{R}$, rX and $X + Y$ are in Φ whenever the operations are defined. The smallest subspace containing Φ we shall denote by $\operatorname{span}\Phi$, and the smallest involutive subspace containing Φ by Φ_{LA}. With the above operations, $\Xi(\mathcal{M})$ is a pseudo-Lie algebra; for simplicity we shall take in this paper the term *Lie algebra* to imply only partially defined operations. The set $\Xi(\mathcal{M})$ can be also seen as a module over the ring of smooth functions on $\mathcal{M}$, again in the sense of partially defined operations; thus $\Phi \subseteq \Xi(\mathcal{M})$ is a submodule of $\Xi(\mathcal{M})$ if all defined linear combinations

$$\alpha_1 X_1 + \ldots \alpha_k X_k$$

of vector fields in Φ, whose coefficients $\alpha_i(\xi)$ are smooth functions on $\mathcal{M}$, are again in Φ. When it is clear from the context that we are dealing with analytic objects, vector field will mean analytic vector field.

If $\pi : \mathcal{N} \to \mathcal{M}$ is a smooth map and $\xi \in \mathcal{M}$, we let

$$\pi_*[\xi] : T_\xi\mathcal{N} \to T_{\pi(\xi)}\mathcal{M}$$

denote the differential of π at ξ. We write simply π_* when ξ is clear from the context. Note that for each tangent vector ν at ξ, $\pi_*[\xi](\nu)$ is a vector at $\pi(\xi)$, When $\pi : \mathcal{N}_1 \times \mathcal{N}_2 \to \mathcal{M}$ is defined on a product of manifolds, we may consider the differentials of each of the partial maps $\pi(\xi, \cdot)$ and $\pi(\cdot, \zeta)$, for fixed ξ and ζ respectively. In that case, we use the alternative notation

$$\left.\frac{\partial}{\partial z}\right|_{z=\zeta} \pi(\xi, z) \;:=\; \pi(\xi, \cdot)_*[\zeta] \tag{1}$$

and similarly for the other partial and for products of more factors. A *smooth partial map* $\pi : \mathcal{N} \to \mathcal{M}$ will be by definition a smooth mapping $\pi : D \to \mathcal{M}$, where D is an open subset of $\mathcal{N}$.

By a *partial diffeomorphism* γ of $\mathcal{M}$ we mean a diffeomorphism (analytic if clear from the context) from an open subset D of $\mathcal{M}$ onto another open subset of $\mathcal{M}$. The inverse γ^{-1} of γ is defined on the image of γ. The composition $\gamma_2 \circ \gamma_1$ of γ_1 and γ_2 is only defined if the image of γ_1 intersects the domain of γ_2. We let let $\mathrm{Diff}(M)$ denote the set of such partial diffeomorphisms on $\mathcal{M}$. As a general rule, if the Lie bracket of two vector fields, or the composition of two partial diffeomorphisms, appears in a statement, that statement should be taken to be read 'if this composition is defined, then...'. We often leave this implicit.

If γ is a partial diffeomorphism of $\mathcal{M}$ and X is a vector field, we consider the pull-back $\mathrm{Ad}_\gamma X$ of X under γ. This is a new vector field, and is defined by the formula

$$\mathrm{Ad}_\gamma X(\xi) = (\gamma^{-1})_*[\gamma(\xi)]X(\gamma(\xi)) = (\gamma_*[\xi]^{-1})X(\gamma(\xi)). \tag{2}$$

This is the same as what is denoted sometimes by $(\gamma^{-1})_*X$ in the differential geometric literature. Its open (possibly empty) domain is $\gamma^{-1}\mathcal{V}_X$. The pull-back is natural with respect to Lie brackets, that is,

$$Ad_\gamma[X,Y] = [Ad_\gamma X, Ad_\gamma Y] \tag{3}$$

for all vector fields X, Y. Further, for any two partial diffeomorphisms γ_1, γ_2, and any X,

$$\mathrm{Ad}_{\gamma_2}\mathrm{Ad}_{\gamma_1}X = \mathrm{Ad}_{\gamma_1\gamma_2}X$$

For any vector field $X \in \Xi(\mathcal{M})$ and each $\xi \in \mathcal{V}_X$, elementary existence theorems for differential equations insure that the initial value problem

$$\dot{x}(\tau) = X(x(\tau)), \ \ x(0) = \xi \tag{4}$$

has a unique solution $x(\tau)$, defined for an open set of pairs (τ, ξ) (which depends on X), and that this solution is smooth as a function of (τ, ξ), analytic if X is analytic. We denote by $\exp(tX)(\xi)$ the value of this solution at time t, if it exists. Here t may be positive or negative. By definition of solution, if $\exp(tX)(\xi)$ is defined then also $\exp(\tau X)(\xi)$ is defined for each τ between 0 and t. For each fixed t, $\exp(tX)$ is a partial diffeomorphism of $\mathcal{M}$. Using the time reparameterization $x(s\tau)$ one proves that $\exp(t(sX)) = \exp((ts)X)$ for each t, s. In particular, there is no ambiguity in denoting $\exp(1X)$ just as $\exp(X)$. We also use sometimes the notation e^{tX} instead of $\exp(tX)$. When $\exp(tX)(\xi)$ is defined for all $t \in \mathbb{R}$ and ξ, one says that X is *complete*.

If $\mathcal{N}$ is a submanifold of $\mathcal{M}$ such that X is tangent to $\mathcal{N}$, that is, if $X(\xi) \in T_\xi\mathcal{N}$ all those $\xi \in \mathcal{N}$ for which X is defined, then (4) can be solved in $\mathcal{N}$, and by uniqueness we conclude that $\exp(tX)(\xi)$ stays in $\mathcal{N}$ for small t. [Note that for large t it may hold that $\exp(tX)(\xi)$ is undefined with respect to $\mathcal{N}$ but not with respect to $\mathcal{M}$. For instance, take $\mathcal{M} = \mathbb{R}, \mathcal{N} = (-1,1), \dot{x}(t) = 1, x(0) = 0$. Then the solution is defined globally on $\mathcal{M}$ but only for $-1 < t < 1$ on $\mathcal{N}$.]

2.2 Smooth actions on manifolds

We now introduce a notion of smooth action on a manifold. The definition generalizes that of action of a Lie group. One of the main results to be proved is a theorem that describes the orbits under such actions as submanifolds associated to certain vector fields. This theorem will generalize the fundamental orbit result of [19], who essentially established it for actions associated to flows of vector fields. The idea of the proof given here is essentially the same, however.

Definition 2.1 Let $\mathcal{M}$ be a manifold. An *action* $\Sigma = (\mathcal{U}, \gamma)$ *on* $\mathcal{M}$ is given by a manifold $\mathcal{U}$ and a smooth partial map

$$\gamma : \mathcal{M} \times \mathcal{U} \to \mathcal{M}$$

such that $\gamma_u := \gamma(\cdot, u)$ is a partial diffeomorphism for each $u \in \mathcal{U}$. ∎

The domain of γ is denoted by D, and we let

$$D_u := \{\xi | (\xi, u) \in D\} \tag{5}$$

be the (possibly empty) domain of γ_u. The notations $\gamma_u(\xi)$ and $\gamma(\xi, u)$ mean the same, and can be used interchangeably, but we will tend to use the first when we wish to use the fact that γ_u is a diffeomorphism, while the second will be used when it is relevant that γ depends smoothly on u. For each u, we may also consider the inverse γ_u^{-1} of γ_u; this is again a partial diffeomorphism, whose domain is $\gamma_u(D_u)$. From now on, we fix an action Σ.

Definition 2.2 The action Σ is *analytic* if $\mathcal{M}$ and $\mathcal{U}$ are analytic manifolds and γ is analytic. It is *complete* if the mapping γ_u is an (everywhere defined) diffeomorphism of $\mathcal{M}$ onto itself, for each fixed $u \in \mathcal{U}$. (Thus, $D = \mathcal{M} \times \mathcal{U}$.) ∎

Given any family $\{\gamma_\lambda, \lambda \in \Lambda\} \subseteq \text{Diff}(M)$, we may always consider Λ as a zero dimensional manifold, and consider the action given by $\gamma(\xi, \lambda) := \gamma_\lambda(\xi)$, with domain D equal to the set of all (ξ, λ) such that γ_λ is defined at ξ. Typically however we have a continuous component to the parameter set, and the interaction between the topologies of this parameter set and the manifold will be the interesting part of the study, as will be clear from the examples given later. The orbit theorems to be proved are all trivial in the case in which $\mathcal{U}$ is discrete.

Remark 2.3 A somewhat different definition of action was given in [17]; the present one will allow us to distinguish between forward and backward motions, which was not possible there. On the other hand, the definition in [17] was more general than the one given here, mainly in that more than one γ is allowed to act on $\mathcal{M}$ at the same time. The present definition appears to be general enough to cover most applications of interest, however. A related but more restrictive concept is studied in [11]. ∎

Remark 2.4 Another more general definition would replace the partial diffeomorphism assumption on γ by a local invertibility assumption, which may be more natural in modeling some applications:

$$\text{for each fixed } (\xi, u) \in D, \ \text{rank}\,(\gamma_u)_*[\xi] = \dim \mathcal{M}.$$

(Note that a map γ_u is a partial diffeomorphism if and only if this property holds and in addition γ_u is one-to-one. Thus the generalization is in dropping the global one-to-one requirement.) Arguing via the implicit mapping theorem, we can conclude that there is in this case an open covering $\{V_\lambda, \lambda \in \Lambda\}$ of D so that for each u, $\gamma(\cdot, u)$ is a diffeomorphism whose domain is the set of ξ for which $(\xi, u) \in V_\lambda$. If we now view Λ as a zero-dimensional manifold and introduce the product manifold $\mathcal{U} \times \Lambda$, we may consider the action with $\tilde{\gamma}(\xi, u, \lambda) := \gamma(\xi, u)$, defined on the open set

$$\{(\xi, u, \lambda) \mid (\xi, u) \in V_\lambda\}.$$

The results to be given later can then be applied to this action. ∎

We now associate to Σ several types of reachable sets involving positive, negative, and mixed positive and negative motions. The latter are the least interesting from a practical point of view, since they typically do not correspond to physically realizable trajectories of systems, but they are easier to study and they provide much geometric insight into the structure of the corresponding system.

Definition 2.5 We shall say that the state $\xi \in \mathcal{M}$ *can be reached from itself in zero steps*, and for each positive integer k we define by induction that the state ζ *can be reached from the state* ξ *in* k *steps*, if there exists a state ζ' which can be reached from ξ in $k-1$ steps and some $u \in \mathcal{U}$ such that $\zeta' \in D_u$ and $\gamma_u(\zeta') = \zeta$. Equivalently, we say that case that ξ *can be controlled to* (or *steered to*) ζ *in* k *steps*. Finally, we define inductively that ξ *is accessible from* ζ *in* k *steps* iff either $k = 0$ and $\xi = \zeta$ or there exists a state ζ' which is accessible from ξ in $k-1$ steps and such that either ζ can be reached from ζ' in one step or ζ can be controlled to ζ' in one step.∎

Note that accessibility is symmetric:

ξ *is accessible from* ζ *in* k *steps* iff ζ *is accessible from* ξ *in* k *steps*,

but that reachability and controllability are not. If ζ is accessible from ξ in k steps for some k, we say simply that ζ is accessible from ξ, and similarly for the other notions. Accessibility is an equivalence relation, and the state space $\mathcal{M}$ is partitioned into equivalence classes, the *orbits*

$$O(\xi) = \{\zeta \mid \zeta \text{ is accessible from } \xi\}.$$

In general, we will use script letters such as $\mathcal{N}$ to denote manifolds, with the corresponding roman letter N denoting the underlying sets. Thus, when we later impose a manifold structure on $O(\xi)$, we shall denote such an orbit as $\mathcal{O}(\xi)$.

Definition 2.6 The action is *transitive* if $O(\xi) = \mathcal{M}$ for all ξ. The action has the *accessibility property from* $\xi \in \mathcal{M}$ if $O(\xi)$ is open in $\mathcal{M}$. ∎

Since the γ_u's are partial diffeomorphisms, the accessibility property is equivalent to the requirement that $O(\xi)$ contain a neighborhood of ξ in $\mathcal{M}$, or even that just that $O(\xi)$ have a nonempty interior. (Sometimes the term accessibility property is used for the stronger concept that the reachable set from ξ should have a nonempty interior. Here we shall use the term in the sense of the above definition. In the context of continuous time analytic systems both possible concepts coincide.)

Definition 2.7 A subset $N \subseteq \mathcal{M}$ is *stable* for the action Σ if for each $\xi \in N$ and each ζ accessible from ξ, ζ is again in N. ∎

Note that stable sets are precisely the same as unions of orbits. When N is stable, we denote by $\gamma \mid N \times \mathcal{U}$ the restriction of γ to the set $D \cap (N \times \mathcal{U})$, seen as a map into N. When $\mathcal{N}$ is a manifold, as below, saying that this restriction is smooth means smooth as a map into $\mathcal{N}$, with the submanifold structure in $\mathcal{N}$.

Definition 2.8 Let $\mathcal{N}$ be a stable submanifold of $\mathcal{M}$. If $(\mathcal{U}, \gamma \mid \mathcal{N} \times \mathcal{U})$ is an action on $\mathcal{N}$, we call this induced action a *subaction* of Σ, and denote it by $\Sigma \mid \mathcal{N}$. When $\mathcal{N}$ is an analytic submanifold and the induced action is analytic, we say that $\Sigma \mid \mathcal{N}$ is an analytic subaction. ∎

Remark 2.9 By the implicit mapping theorem, it follows that a stable submanifold $\mathcal{N}$ of $\mathcal{M}$ induces a subaction if and only if $\gamma \mid \mathcal{N} \times \mathcal{U}$ is smooth. ∎

The following is one of the fundamental results about actions. It will be a consequence of Theorem 2 in section 3.2. Its main interest is in the case when $\mathcal{N}$ is a single orbit.

Theorem 1 *Let Σ be an action and let N be any stable set. Then there exists a submanifold structure $\mathcal{N}$ on N which makes the restriction $\Sigma \mid \mathcal{N}$ a Σ-subaction. When Σ is an analytic action, this is an analytic subaction.* ∎

Remark 2.10 There may be more than one submanifold structure $\mathcal{O}(\xi)$ on an orbit $O(\xi)$, or more generally on a stable set N, for which $\mathcal{O}(\xi)$ induces a Σ-subaction. As an illustration, consider the case of the action $\gamma(\xi, u) := \xi + u$, with $\mathcal{M} = \mathbb{R}$ (with the usual 1-dimensional structure), and $\mathcal{U} = \mathbb{R}_{\text{disc}}$. For this action, $O(0)$ is as a set all of $\mathbb{R}$, but it is a Σ-subaction under both the submanifold structure $\mathbb{R}_{\text{disc}}$ and the usual structure $\mathbb{R}$. However, we prove later that, in general, *there is a unique submanifold structure on $\mathcal{O}(\xi)$ of minimal possible dimension.* Thus the 0-dimensional $\mathbb{R}_{\text{disc}}$ is more natural than $\mathbb{R}$ for this example. Another uniqueness statement can be given in terms of integrability of distributions, and this is done in section 3.2. ∎

A natural topology can be imposed on the set $O(\xi)$, as follows. We start by introducing the set

$$D^- := \{(\zeta, u) \mid \zeta \in \gamma_u(D_u)\} \tag{6}$$

and defining γ^- by

$$\gamma^-(\zeta, u) = \gamma_u^{-1}(\zeta). \tag{7}$$

The implicit function theorem can be applied to the equation for ξ

$$\gamma(\xi, u) = \zeta$$

about any ξ_0, u_0, ζ_0 such that $\gamma(\xi_0, u_0) = \zeta_0$, because the partial differential of γ with respect to ξ is invertible. It follows that D^- is open and that γ^- is smooth. It is convenient to denote γ also as γ^+ and D as D^+.

With these notations, we can say that $O(\xi)$ is the union of the images of the (partial) mappings

$$\theta_\xi^b : \mathcal{U}^k \to \mathcal{M} : (u_1, \ldots, u_k) \mapsto \gamma^b(\xi, u_1, \ldots, u_k), \tag{8}$$

where

$$\gamma^b(\xi, u_1, \ldots, u_k) := \gamma^{a_k}(\gamma^{a_{k-1}}(\ldots(\gamma^{a_2}(\gamma^{a_1}(\xi, u_1), u_2)\ldots), u_{k-1}), u_k) \tag{9}$$

one such mapping for each possible finite sequence $b = (a_1, \ldots, a_k)$ of $+$'s and $-$'s. These mappings are smooth as mappings into $\mathcal{M}$, since they are obtained as compositions of smooth maps. Note that each γ^b is defined on an open subset D^b of $\mathcal{M} \times \mathcal{U}^k$ (see below), and θ^b_ξ on an open subset of $\mathcal{U}^k$. Moreover, if $O(\xi)$ has a submanifold structure $\mathcal{O}(\xi)$ which induces a subaction, then they must also be continuous as maps into $\mathcal{O}(\xi)$.

Definition 2.11 The *orbit topology* on $O(\xi)$ is the finest topology for which all the maps θ^b_ξ are continuous. ∎

Some more terminology will help later when working with the maps γ^b. We let $\mathcal{A} = \{+, -\}$, and for the element $a = \pm$ we let $-a := \mp$ respectively. The free monoid on $\mathcal{A}$ is $\mathcal{A}^*$, the set of all possible sequences b of $+$'s and $-$'s. The subset $\mathcal{A}^*_+$ is the set of all sequences of $+$'s alone, and we think of $\mathcal{A}$ as a subset of $\mathcal{A}^*$ consisting of sequences of length 1. For each $b = (a_1, ..., a_r) \in \mathcal{A}^*$, $-b$ is the sequence $(-a_r, ..., -a_1)$, (note the reversed order) and we use the notation $\mathcal{U}^b$ instead of $\mathcal{U}^r$, for the product

$$\mathcal{U}^b = \underbrace{\mathcal{U} \times \ldots \times \mathcal{U}}_{r \text{ copies}}$$

So $\mathcal{U}^+ = \mathcal{U}^- = \mathcal{U}$. The sets $\mathcal{A}^*$ and $\mathcal{A}^*_+$ include the empty sequence ϕ for which $r = 0$ and $\mathcal{U}^\phi$ is a one-point set. Each of above maps γ^b is defined on a subset D^b of $\mathcal{M} \times \mathcal{U}^b$; inductively on the length of b we have $\gamma^\phi(\xi, \phi) = \xi$ and for $a \in \mathcal{A}$,

$$(x, u\omega) \in D^{ab} \quad \text{iff} \quad (x, u) \in D^a \text{ and } (\gamma^a(x, u), \omega) \in D^b,$$

for u in $\mathcal{U}^a$ and ω in $\mathcal{U}^b$, and then

$$\gamma^{ab}(x, u\omega) := \gamma^b(\gamma^a(x, u), \omega).$$

Each set D^b is open, and the maps γ^b are smooth. If $\mathcal{N}$ induces a subaction, we denote by $\gamma^b \mid \mathcal{N} \times \mathcal{U}^b$ the restriction of γ^b to $D^b \cap (\mathcal{N} \times \mathcal{U}^b)$, seen as a map into $\mathcal{N}$. A concatenation notation is alternatively used to exhibit sequences in $\mathcal{U}^b$, as in $u\omega$ above, and similarly for words in $\mathcal{A}^*$. If $\omega = (u_1, \ldots, u_r)$ is in $\mathcal{U}^b$, we let $\tilde{\omega} := (u_r, \ldots, u_1)$, an element of $\mathcal{U}^{-b}$ (note the reversed order). Even though $\mathcal{U}^b = \mathcal{U}^{-b}$, we write $\mathcal{U}^{-b}$ in order to emphasize that γ^{-b} is being used. Then $(\gamma^b(x, \omega), \tilde{\omega})$ is in D^{-b} whenever (x, ω) is in D^b and

$$\gamma^{-b}(\gamma^b(x, \omega), \tilde{\omega}) = x.$$

Consistently with the case when $b \in \mathcal{A} = \{\pm\}$, we denote

$$\gamma^b_\omega : D^b_\omega \to \mathcal{M}, \gamma^b_\omega(\xi) := \gamma^b(\xi, \omega), \tag{10}$$

where $D^b_\omega := \{\xi | (\xi, \omega) \in D^b\}$.

2.3 Main examples

We now provide the main examples of actions that we shall be concerned with. The above definition was formulated so that these become particular cases. We start with the most classical one. Here $\mathbb{R}_{>0}$ denotes the set of positive reals.

Definition 2.12 Assume that $\mathcal{M}$ is a manifold and that Φ is a set of vector fields on $\mathcal{M}$. The *action associated to* Φ, denoted $\Sigma(\Phi)$, is the action on $\mathcal{M}$ which has $\mathcal{U} = \mathbb{R}_{>0} \times \Phi$, where Φ is thought of as a discrete manifold, and

$$\gamma(\xi, t, X) := \exp(tX)(\xi).\blacksquare$$

This a well-defined action because of the smooth dependence of solutions of differential equations on time and initial conditions. It is analytic (respectively, complete,) iff each $X \in \Phi$ is an analytic (respectively, complete,) vector field. Note that here

$$\gamma_{t,X}^{-1} = \exp(tX)^{-1} = \exp(-tX) = \gamma_{-t,X}$$

and if $-X \in \Phi$ this inverse also equals $\gamma_{t,-X}$.

Later we shall show how, conversely, certain sets of vector fields arise naturally when studying actions. We now turn to continuous and discrete time systems. We define the former in such a way that existence and uniqueness theorems for differential equations apply.

By a (time-invariant) *continuous time system* we shall mean a controlled set of differential equations

$$\dot{x}(t) = P(x(t), u(t)), \quad t \in \mathbb{R}, \tag{11}$$

where the *state* $x(t)$ belongs to a second countable manifold $\mathcal{M}$, controls $u(t)$ take values in a metric space $\mathcal{K}$, and

$$P : \mathcal{M} \times \mathcal{K} \to T\mathcal{M} \tag{12}$$

is a continuous mapping defined on an open subset of $\mathcal{M} \times \mathcal{K}$ such that $X_u := P(\cdot, u)$ is a smooth vector field for each fixed $u \in \mathcal{K}$ and such that $[X_u, Y]$ is again continuous in (ξ, u) for all vector fields Y on $\mathcal{M}$. An *analytic* system is one for which $\mathcal{M}$ as well as each X_u is analytic.

We say that the system is *smooth in controls* if $\mathcal{K}$ is a second countable manifold and P is smooth (jointly on $\mathcal{M}$ and $\mathcal{K}$). For systems smooth in controls, analyticity is taken to mean that (12) is jointly analytic, and $\mathcal{K}$ is an analytic manifold. When P is jointly smooth, the above continuity requirement on each bracket $[X_u, Y]$ is automatically satisfied. We will be especially interested in such systems, but the more general definition is needed in order to prove some of the intermediate results.

Any set of vector fields Φ can be seen as a continuous time system, simply by introducing a discrete metric on Φ, but the interest here will be on time functions $u(\cdot)$ as controls, so the structure of $\mathcal{K}$ will be relevant. For a discrete metric, only controls taking a finite set of values will be admissible in the sense to be defined below.

A particular class of continuous time systems which appears in modeling a large number of physical systems is that of *systems affine in the control.* These are systems for which $\mathcal{K} \subseteq \mathbb{R}^m$

and for which there exist smooth vector fields $f, g_1, \ldots, g_m$ on $\mathcal{M}$ such that $P(x,u) = f + \sum_i u_i g_i$, that is, the equations are

$$\dot{x} = f(x) + \sum_{i=1}^{m} u_i g_i(x). \tag{13}$$

Note that any set of smooth vector fields $\{f, g_1, \ldots, g_m\}$ provides such a system. The equations are always smooth on $(x,u) \in \mathcal{M} \times \mathbb{R}^m$. The system is analytic when these vector fields are. When dealing with systems affine in the control, we shall always assume that $\mathcal{K}$ has a nonempty interior in $\mathbb{R}^m$.

In typical examples of continuous time systems (11), $\mathcal{K}$ is a subset of an Euclidean space $\mathbb{R}^m$ that may include magnitude or energy constraints. For instance, $\mathcal{K}$ may consist of the set of all vectors u whose components satisfy $|u_i| \leq 1$, or the unit ball $u_1^2 + \ldots + u_m \leq 1$. If the dynamics P restricts to a smooth mapping on $\mathcal{M} \times \operatorname{int} \mathcal{K}$, we may restrict controls to the interior of $\mathcal{K}$ and consider this as a smooth system. We shall see later (see remark 4.8) that many controllability properties are not changed by restricting to this interior, as long as $\mathcal{K}$ is, as for these examples, included in the closure of its interior.

To say that a map $u : [0,T] \to \mathcal{K}$ is measurable means that $u^{-1}(V)$ is measurable for each open subset $V \subseteq K$. To say it is essentially bounded means that there is some compact subset $K \subseteq \mathcal{K}$ such that $u(t) \in K$ for almost all t. With this terminology, we recall a basic fact about differential equations. For each $\xi \in \mathcal{M}$ and each measurable and essentially bounded $u : [0,T] \to \mathcal{K}$, the initial value problem

$$\dot{x}(t) = P(x(t), u(t)), \quad x(0) = \xi \tag{14}$$

admits a unique solution for small t. (By solution one means an absolutely continuous curve. For manifolds, absolute continuity is defined as absolute continuity of each restriction to an open subinterval for which the image is entirely in a chart.) This follows from standard existence theorems as follows. Since the statement is local, we may work in $\mathcal{M} = \mathbb{R}^n$. Taking above for Y the possible derivatives $\partial/\partial x_i$, we have that, with $Z(\xi,t) := P(\xi, u(t))$, Z satisfies the classical Carathéodory conditions for existence and uniqueness (see e.g. [4], chapter II).

We shall say that $\xi \in \mathcal{M}$ can be *controlled to* $\zeta \in \mathcal{M}$ with respect to the system (11) iff there exists some interval $[0,T]$, $T \geq 0$, an essentially bounded measurable map $u(\cdot) : [0,T] \to \mathcal{K}$, and a solution of the differential equation (11) with this $u(\cdot)$, defined on the entire interval $[0,T]$, such that $x(0) = \xi, x(T) = \zeta$. We say equivalently that ζ can be *reached from* ξ, and that $u(\cdot)$ *steers* ξ to ζ. If there exists a finite sequence of states $\xi_1 = \xi, \xi_2, \ldots, \xi_k = \zeta$ such that for each $i = 2, \ldots, k$, ξ_i is either reachable from or controllable to ξ_{i-1}, we say that ζ is *accessible* from ξ. The terminology *weakly reachable* is often used in the literature to refer to this last concept. We say that the *accessibility property* is satisfied from a state ξ with respect to the system (11) if the set of states accessible from ξ is an open subset of $\mathcal{M}$. A *completely controllable system* will be one for which ξ can be controlled to ζ for every pair ξ, ζ.

Remark 2.13 Given any essentially bounded measurable map $u(\cdot) : [0,T] \to \mathcal{K}$, we may always find a sequence of piecewise constant controls $u_l(\cdot) : [0,T] \to \mathcal{K}$ with the property that, if the solution of (11) with $x(0) = \xi$ and control $u(\cdot)$ is defined on the interval $[0,T]$, then the solutions are also defined for each of the controls $u_l(\cdot)$ and same initial state, and the corresponding final states $x_l(T)$ converge to $x(T)$. This is a consequence of general theorems on continuous dependence of solutions of (11). Moreover, the same argument allows taking the approximating piecewise constant controls to have values in any dense subset of $\mathcal{K}$. ∎

There are two very different types of actions that can be associated to a given continuous time system. The first is is the one implicitly used in the literature when dealing with continuous time systems.

Definition 2.14 The *time-topology action associated* to the system (11) is the action $\Sigma(\Phi)$, where Φ is the set of all vector fields of the type $X_u = P(\cdot, u), u \in \mathcal{K}$. ∎

We also write $\gamma(\xi, t, u)$, or $\gamma_{t,u}(\xi)$, instead of $\gamma(\xi, t, X_u)$ in this case, and think of $\mathcal{U}$ as $\mathbb{R}_{>0} \times \mathcal{K}$ (second factor with the *discrete* topology). In this definition, the structure of $\mathcal{K}$ is irrelevant. The name time-topology is due to the fact that the orbit topology (definition 2.11) is induced by the dependency of $\gamma(\xi, t, u)$ on t. A different type of action is associated in the next definition, for which the orbit topology will be induced by the topology on control values.

Definition 2.15 Let (11) be a system smooth in controls. The *input-topology action associated* to it is the action having also $\mathcal{U} = \mathbb{R}_{>0} \times \mathcal{K}$ but now with the second factor having its differentiable structure and the factor $\mathbb{R}_{>0}$ as a discrete manifold, and again

$$\gamma_{t,u}(\xi) := \exp(tX_u)(\xi). \blacksquare \tag{15}$$

This second type of action is not usually studied in the theory of continuous time systems (11). It turns out however to provide the right framework for understanding sampling results. Note that one could just as well introduce a "joint" action of time and inputs, where $\mathcal{U} = \mathbb{R}_{>0} \times \mathcal{K}$ and both factors are given their natural differentiable structures. This would give yet another natural type of action associated to such systems, but seems of less interest for applications. Whenever the input-topology action is mentioned, we assume implicitly that we are dealing with systems that are smooth in controls.

We now relate the controllability definitions for continuous time systems to those for the corresponding actions. Under either type of action, time-topology or input-topology, the k-step notions correspond to states reachable in positive, negative, or mixed positive and negative time using *piecewise constant* controls with precisely k switches. As discussed in remark 2.13, elements in $O(\xi)$ can in principle only be expected to be dense in the set of states accessible using the differential equation and arbitrary measurable controls. However it will turn out that the two concepts of accessibility are equivalent. This result is proved in section 4.5:

Proposition 2.16 ζ is accessible from ξ with respect to the system (11) iff $\zeta \in O(\xi)$. ∎

Finally, we may consider (time-invariant) *invertible discrete time* systems. These are a natural class of discrete time control systems, and where studied explicitely first by [8]. Their control properties are studied in detail in [10]. They are described by controlled difference equations

$$x(t+1) = P(x(t), u(t)), \quad t \in \mathbb{Z}, \tag{16}$$

where the *state* $x(t)$ belongs to a second countable manifold $\mathcal{M}$ and controls $u(t)$ take values in a second countable manifold $\mathcal{K}$. The map $P : D \to \mathcal{M}$ is smooth on an open subset D of $\mathcal{M} \times \mathcal{K}$, and for each u in $\mathcal{K}$, $P(\cdot, u)$ is a partial diffeomorphism ("invertibility"). An *analytic* system is one for which all data is analytic. A *complete* system is one for which $P(\cdot, u)$ is a (everywhere defined) diffeomorphism of $\mathcal{M}$ onto $\mathcal{M}$, for each $u \in \mathcal{K}$.

Such systems arise for instance under *sampling* of a (smooth in controls) continuous time system (11). For any such system and each fixed real number $\delta > 0$, one introduces the δ-*sampled* system associated to (11); this has the same $\mathcal{M}$ and $\mathcal{K}$, and equations as in (16) with

$$P(\xi, u) := \exp(\delta X_u)(\xi). \tag{17}$$

Even if the vector fields X_u are everywhere defined, that is P in (11) is defined on all of $\mathcal{M} \times \mathcal{K}$, this is in general not a complete discrete time system, since the vector fields need not be complete.

We identify discrete time systems with the associated actions $(\mathcal{K}, P)$, and also write γ for P and $\mathcal{U}$ for $\mathcal{K}$. Thus discrete time systems are the same as actions for which $\mathcal{M}$ and $\mathcal{U}$ are second countable. For discrete time systems, we define controllability and related notions in terms of the associated action.

When the system (16) happens to be the δ-sampled system associated to (11), for some δ, the discrete time action is related to the input-topology action associated to (11), in the sense that γ_u in this definition is the particular element $\gamma_{\delta,u}$ in (15). For small enough δ the elements $\gamma_{k\delta,u}$, for integer k, approximate arbitrary $\gamma_{t,u}$. What will turn out to be true, but is far from obvious at this stage, is that the actions corresponding to sampled systems will be in a certain sense *equal* to, not just approximations of, the (input-topology) action associated to the original system, for every small enough δ. This fact will be made precise and proved later, after more machinery is in place, and it is one of the main results about sampling.

Often one might want to use control value sets which are not manifolds. Consider the following definition. A subset $\mathcal{C}$ of a manifold $\mathcal{K}$ has *nice boundary* if the following property holds: for each $u \in \mathcal{C}$ there is a smooth curve

$$\rho : (-\varepsilon, 1 + \varepsilon) \to \mathcal{K}$$

for some $\varepsilon > 0$, such that $\rho(0) = u$ and $\rho(t)$ is in $\operatorname{int} \mathcal{C}$, the interior of $\mathcal{C}$ with respect to $\mathcal{K}$, for each $t \in (0, 1]$. If $\mathcal{C}$ is a subset with nice boundary of $\mathcal{K}$, we may consider the action with $\mathcal{U} := \operatorname{int} \mathcal{C}$. It turns out that the orbits with respect to $\operatorname{int} \mathcal{C}$ are the same as the states that can be obtained by allowing controls in $\mathcal{C}$ and defining accessibility in the same manner; see proposition 4.9, stated in section 4.5. There is no a priori reason for this equality to hold; all that one may say in principle is that each state ζ accessible from ξ using arbitrary controls is a limit of states which are accessible with respect to controls in the interior. Since orbits are not necessarily closed (consider as an illustration a rotation by an angle not commensurable with π on $\mathcal{M}$= unit circle), this result is somewhat surprising.

3 Integrability of Distributions

The manifold structure in Theorem 1 can be described naturally in terms of the integrability of distributions.

3.1 Distributions

We start by defining (possibly singular) distributions of tangent vectors. This concept is defined in many different and not always equivalent ways in the literature. We shall take it to mean the choice of a subspace of the tangent space at each ξ in $\mathcal{M}$. That is, a *distribution* $\mathcal{D}$ on $\mathcal{M}$ is a subset of the tangent bundle $T\mathcal{M}$ with the property that

$$\mathcal{D}(\xi) := \{v \in T_\xi\mathcal{M} | (\xi, v) \in \mathcal{D}\}$$

is a subspace for each ξ. Its *rank at* ξ is the dimension of $\mathcal{D}(\xi)$. The distribution is *nonsingular* if this rank is independent of ξ. (In differential geometry the term distribution is often taken to already imply nonsingularity.) We shall say that $\mathcal{D}$ *has full rank at* ξ if $\mathcal{D}(\xi) = T_\xi\mathcal{M}$, and simply that $\mathcal{D}$ *has full rank* if this holds for each $\xi \in \mathcal{M}$.

If $\mathcal{N}$ is a submanifold of $\mathcal{M}$ such that $\mathcal{D}(\xi) \subseteq T_\xi\mathcal{N}$ for each $\xi \in \mathcal{N}$, there is a well-defined *restriction* of $\mathcal{D}$ to a distribution on $\mathcal{N}$, which we denote by $\mathcal{D} \mid \mathcal{N}$. We use the inclusion notation $\mathcal{D} \subseteq \mathcal{D}'$ to mean that

$$\mathcal{D}(\xi) \subseteq \mathcal{D}(\xi')$$

for each $\xi \in \mathcal{M}$. Similarly, we define $\mathcal{D} + \mathcal{D}'$ and intersection $\mathcal{D} \cap \mathcal{D}'$ pointwise:

$$\begin{aligned} (\mathcal{D} + \mathcal{D}')(\xi) &:= \mathcal{D}(\xi) + \mathcal{D}'(\xi) \\ (\mathcal{D} \cap \mathcal{D}')(\xi) &:= \mathcal{D}(\xi) \cap \mathcal{D}'(\xi). \end{aligned}$$

We now show how to associate in a natural way a distribution to each set of vector fields, and conversely, a set of vector fields to each distribution. Given $\mathcal{D}$, a vector field X *pointwise belongs to* $\mathcal{D}$, or *takes values in* $\mathcal{D}$, if $X(\xi) \in \mathcal{D}(\xi)$ for each $\xi \in \mathcal{V}_X$. The set of vector fields that pointwise belong to $\mathcal{D}$ is denoted by $\mathrm{vf}(\mathcal{D})$. Note that $\mathrm{vf}(\mathcal{D})$ is always a submodule of vector fields.

Conversely, starting with a family of vector fields Φ, the *distribution* $\mathcal{D}(\Phi)$ *determined by* Φ is the smallest distribution $\mathcal{D}$ for which all $X \in \Phi$ pointwise belong to $\mathcal{D}$. Thus for each $\xi \in \mathcal{M}$

$$\mathcal{D}(\Phi)(\xi) = \mathrm{span}\,\{X(\xi) | X \in \Phi \text{ and } X \text{ is defined at } \xi\}.$$

(If no such X are defined at ξ, $\mathcal{D}(\Phi)(\xi) = 0$.) The *rank* of the family Φ at a point ξ is by definition the rank of the associated distribution $\mathcal{D}(\Phi)$ at that point.

We shall say that the distribution $\mathcal{D}$ is *smooth* if it equals $\mathcal{D}(\Phi)$ for some set of vector fields Φ. Note that always $\mathcal{D}(\mathrm{vf}(\mathcal{D})) \subseteq \mathcal{D}$, with equality iff $\mathcal{D}$ is smooth.

An *integral (sub)manifold* $\mathcal{N}$ *of* $\mathcal{D}$ is a submanifold of $\mathcal{M}$ with the property that

$$T_\xi\mathcal{N} = \mathcal{D}(\xi)$$

for every $\xi \in \mathcal{N}$. This generalizes the notion of integral curve for a vector field X (or rather, for its associated distribution $\mathcal{D}(\{X\})$). We shall say that the distribution $\mathcal{D}$ is *integrable* if there is a partition of $\mathcal{M}$ into integral manifolds of $\mathcal{D}$.

A related definition is as follows. An integral manifold $\mathcal{N}$ of $\mathcal{D}$ is a *maximal integral manifold of* $\mathcal{D}$ if it is connected and if for every other connected integral manifold $\mathcal{N}'$ of $\mathcal{D}$ intersecting $\mathcal{N}$, $\mathcal{N}'$ is an open submanifold of N. A distribution satisfies the *maximal integral manifolds property* if it induces a (singular) foliation, that is, there is a partition of $\mathcal{M}$ into maximal integral manifolds of $\mathcal{D}$, the *leaves* of $\mathcal{D}$. We remark later that if $\mathcal{D}$ is integrable then it also satisfies this property; the connected components of the integral manifolds in the definition of integrability provide the leaves of the foliation.

We shall say that $\mathcal{D}$ is *involutive* if $\mathrm{vf}(\mathcal{D})$ is an involutive set of vector fields. Integrable distributions are necessarily involutive. Indeed, if X, Y are in $\mathrm{vf}(\mathcal{D})$ and $\xi \in \mathcal{M}$, let $\mathcal{N}$ be the integral manifold of $\mathcal{D}$ passing through ξ. Then the vector fields X, Y are tangent to $\mathcal{N}$ and therefore their Lie bracket at ξ is again tangent to $\mathcal{N}$. It is in general false that (smooth) involutive distributions are necessarily integrable. The classical theorem of Frobenius states that this implication does hold provided that $\mathcal{D}$ be nonsingular, and the Hermann-Nagano theorem states that the same conclusion is true if $\mathcal{D}$ is generated by analytic vector fields. Following [19], we shall review later how these results are easy consequences of the orbit theorem to be proved below.

3.2 The integrability result

If $\Sigma \mid \mathcal{N}$ is a subaction, then from the fact that each γ_u is a partial diffeomorphism it follows that, for each u and each $\xi \in \mathcal{N} \cap D_u$, the differential

$$(\gamma_u)_*[\xi] : T_\xi \mathcal{M} \to T_{\gamma_u(\xi)}\mathcal{M}$$

maps $T_\xi \mathcal{N}$ onto $T_{\gamma_u(\xi)}\mathcal{N}$. If an orbit has a manifold structure $\mathcal{O}(\xi)$ inducing a subaction, this says that in particular

$$(\gamma_u)_*[\xi_u]\, T_\zeta \mathcal{O}(\xi) \;\; = \;\; T_{\gamma_u(\zeta)}\mathcal{O}(\xi)$$

for each $\zeta \in \mathcal{O}(\xi)$ in the domain of γ_u. In particular, if Theorem 1 where to hold, then with respect to any such manifold structure the distribution $\mathcal{D}$ defined at each ξ by the formula

$$\mathcal{D}(\xi) := T_\xi \mathcal{O}(\xi) \tag{18}$$

would be invariant in the sense of the following definition.

Definition 3.1 Let Σ be an action and $\mathcal{D}$ a distribution on $\mathcal{M}$. We shall say that $\mathcal{D}$ is Σ-*invariant* if for each $(\xi, u) \in D$, $(\gamma_u)_*[\xi]\mathcal{D}(\xi) = \mathcal{D}(\gamma_u(\xi))$. ∎

Note that it follows directly from the definition that a Σ-invariant distribution must have constant rank along orbits of Σ.

An equivalent way of defining invariance is in terms of pull-backs. For any distribution $\mathcal{D}$ and any partial diffeomorphism $\gamma : \mathcal{M} \to \mathcal{M}$ we let $\mathrm{Ad}_\gamma \mathcal{D}$ be the distribution defined as follows. If ξ is not in the domain of γ then $\mathrm{Ad}_\gamma \mathcal{D}(\xi) := 0$, otherwise

$$\mathrm{Ad}_\gamma \mathcal{D}(\xi) := (\gamma_*[\xi])^{-1}\mathcal{D}(\gamma(\xi)) = \{(\gamma_*[\xi])^{-1}(\nu), \nu \in \mathcal{D}(\gamma(\xi))\}.$$

With this definition,

$$\mathrm{Ad}_\alpha \mathrm{Ad}_\beta \mathcal{D} = \mathrm{Ad}_{\beta\circ\alpha}\mathcal{D} \tag{19}$$

where we interpret $\beta \circ \alpha(\xi)$ as undefined if $\alpha(\xi)$ is. In particular,

$$\mathrm{Ad}_{\gamma^{-1}} \mathrm{Ad}_\gamma \mathcal{D}(\xi) = \mathcal{D}(\xi) \tag{20}$$

whenever $\gamma(\xi)$ is defined. Since γ is invertible, we have that whenever $\gamma(\xi)$ is defined,

$$\mathrm{Ad}_\gamma \mathcal{D}(\xi) = \mathcal{D}(\xi) \;\text{ if and only if }\; \gamma_*[\xi]\mathcal{D}(\xi) = \mathcal{D}(\gamma(\xi)).$$

Therefore,

Lemma 3.2 $\mathcal{D}$ is Σ-invariant iff $\mathrm{Ad}_{\gamma_u}\mathcal{D}(\xi) = \mathcal{D}(\xi)$ for all $(\xi, u) \in D$. ∎

When $\mathcal{D}$ is smooth, $\mathcal{D} = \mathcal{D}(\Phi)$, it follows from the definition (2) that

$$\mathrm{Ad}_\gamma \mathcal{D}(\xi) = \mathrm{span}\,\{\mathrm{Ad}_\gamma X(\xi) \mid X \in \Phi, X \;\text{ defined at }\; \gamma(\xi)\}. \tag{21}$$

(We make the convention that $\mathrm{span}\,\emptyset = \{0\}$.)

There is always a (unique) smallest Σ-invariant distribution containing any given $\mathcal{D}$, which we denote $\mathrm{Ad}_\Sigma\mathcal{D}$. This is because there is at least one such containing distribution, namely the one with $\mathcal{D}(\xi) = T_\xi\mathcal{M}$ for all ξ, and given any family of such distributions, their intersection is again invariant. In fact, $\mathrm{Ad}_\Sigma\mathcal{D}$ can be obtained by starting with the vectors at each $\mathcal{D}(\xi)$ and taking all possible iterated images and preimages under the differentials of the γ_u's. Because of formula (21), $\mathrm{Ad}_\Sigma\mathcal{D}$ is smooth whenever $\mathcal{D}$ is. When $\Sigma \mid \mathcal{N}$ is a subaction for which the restriction $\mathcal{D} \mid \mathcal{N}$ is defined, all these iterates remain in tangent spaces to $\mathcal{N}$, and we conclude the following fact.

Lemma 3.3 If $\Sigma \mid \mathcal{N}$ is a subaction and $\mathcal{D}$ is a distribution such that $\mathcal{D} \mid \mathcal{N}$ is defined, then $\mathrm{Ad}_\Sigma\mathcal{D} \mid \mathcal{N}$ is also defined and it equals $\mathrm{Ad}_\Sigma(\mathcal{D} \mid \mathcal{N})$. ∎

We shall now reverse the reasoning and define directly the distribution in equation (18), not assuming known that a subaction can be induced on orbits. For each $(\xi, u) \in D^-$, consider the mapping $\gamma^-(\gamma(\xi, \cdot), u)$ defined in a neighborhood of u. We may consider its differential at u, applied to each tangent vector $\nu \in T_u\mathcal{U}$:

$$\mathcal{X}_{u,\nu}(\xi) := \left.\frac{\partial}{\partial v}\right|_{v=u} \gamma^-(\gamma(\xi, v), u)(\nu). \tag{22}$$

It is clear from its expression in local coordinates as a product of Jacobians multiplied by a vector that this expression is smooth in (ξ, u). In particular, (22) defines a smooth vector field on $\mathcal{M}$, which is analytic if the action is analytic. These vector fields play an important role in characterizing reachability and accessibility properties of actions.

Definition 3.4 The set of *vector fields associated to the action* Σ is the set

$$\mathrm{vf}(\Sigma) := \{\mathcal{X}_{u,\nu} \;\mid\; u \in \mathcal{U}, \nu \in T_u\mathcal{U}\}.$$

The *distribution associated to the action* Σ is $\mathcal{D}_\Sigma := \mathrm{Ad}_\Sigma\mathcal{D}(\mathrm{vf}(\Sigma))$, the smallest Σ-invariant distribution which contains the distribution determined by the vector fields associated to Σ. ∎

Alternatively, from the formula

$$(\gamma_u)_*[\xi](\mathcal{X}_{u,\nu}(\xi)) = \frac{\partial}{\partial v}\Big|_{v=u} \gamma(\gamma^-(\gamma(\xi,v),u),u)(\nu) = \frac{\partial}{\partial v}\Big|_{v=u} \gamma(\xi,v)(\nu) \tag{23}$$

we also know that $\mathcal{D}_\Sigma$ is the same as the smallest Σ-invariant distribution containing the distribution

$$\mathcal{E}(\zeta) := \text{span}\,\{ \frac{\partial}{\partial v}\Big|_{v=u} \gamma(\xi,v)(\nu) \quad | \quad (\xi,u) \in D, \nu \in T_u\mathcal{U}, \gamma_u(\xi) = \zeta\}. \tag{24}$$

Lemma 3.5 If $\mathcal{N}$ is a submanifold structure on $O(\xi)$ inducing a subaction, then $\mathcal{D}_\Sigma(\zeta) \subseteq T_\zeta\mathcal{N}$ for each $\zeta \in \mathcal{N}$.

Proof. Since each of the mappings γ_u^{-1}, γ_v leaves $\mathcal{N}$ invariant, all the vector fields (22) are tangent to $\mathcal{N}$. Thus $\mathcal{D}(\text{vf}(\Sigma)) \mid \mathcal{N}$ is well-defined. The conclusion follows from lemma 3.3:

$$\mathcal{D}_\Sigma \mid \mathcal{N} = \text{Ad}_\Sigma(\mathcal{D}(\text{vf}(\Sigma)) \mid \mathcal{N}) = \mathcal{D}_{\Sigma|\mathcal{N}}. \blacksquare$$

Because $\mathcal{D}_\Sigma$ can be generated by iteratively applying the differentials of the maps γ_u and their inverses, we have with the notation in (10) that $\mathcal{D}_\Sigma$ can be also defined as

$$\mathcal{D}_\Sigma = \mathcal{D}(\{\text{Ad}_{\gamma_\omega^b} \mathcal{X}_{u,\nu} \mid \omega \in \mathcal{U}^b, b \in \mathcal{A}^*, u \in \mathcal{U}, \nu \in T_u\mathcal{U}\}). \tag{25}$$

Furthermore, from the chain rule for derivatives it follows that each partial derivative of an expression such as (9) with respect to a fixed u_i is one of the vector fields appearing in the generating set in (25). It follows that

Lemma 3.6 For each $b \in \mathcal{A}^*$ and each $(\xi,\omega) \in D^b$, the image of $(\theta_\xi^b)_*[\omega]$ is contained in $\mathcal{D}_\Sigma(g^b(\xi,\omega))$. $\blacksquare$

We can state a theorem which summarizes the main facts about integrability of distributions associated to actions. Note this implies Theorem 1, since the set N there can be given the submanifold structure which consists of the disjoint union of the manifold structures on the orbits contained in it. For part (4) recall the definition 2.11 of the orbit topology, and for part (2) the notations in equation (9).

Theorem 2 *For any action Σ, the distribution $\mathcal{D}_\Sigma$ is integrable. Furthermore, pick any $\xi \in \mathcal{M}$ and let s be the rank of $\mathcal{D}_\Sigma$ along the orbit of ξ. Then, $O(\xi)$ admits a unique s-dimensional submanifold structure $\mathcal{O}(\xi)$ which induces a subaction, and the following properties hold for $\mathcal{O}(\xi)$:*

1. *When Σ is an analytic action, $\Sigma \mid \mathcal{O}(\xi)$ is an analytic subaction.*
2. *Given any $\zeta \in O(\xi)$, there is some sequence b and some $\omega \in \mathcal{U}^k$ such that $\gamma^b(\xi,\omega) = \zeta$ and $(\theta_\xi^b)_*$ has rank s at ω.*

3. *$\mathcal{O}(\xi)$ is an integral manifold of $\mathcal{D}_\Sigma$.*

4. *The topology of $\mathcal{O}(\xi)$ is the orbit topology.*

The following fact, to be proved in section 3.4, will provide the needed technical construction:

Lemma 3.7 Let Σ be any action on $\mathcal{M}$, and pick any $\xi \in \mathcal{M}$. Then, $O(\xi)$ has a unique structure $\mathcal{O}(\xi)$ of submanifold of $\mathcal{M}$ such that

a. For each $b \in \mathcal{A}^*$, $\gamma^b \mid (\mathcal{O}(\xi) \times \mathcal{U}^b)$ is smooth.

b. For any ζ in $\mathcal{O}(\xi)$, the dimension r of $\mathcal{O}(\xi)$ is the largest possible value $r(\xi, \zeta)$ of the rank of $(\theta^b_\xi)_*[\omega]$ among all b and ω such that $\gamma^b(\xi, \omega) = \zeta$.

When Σ is analytic, $\mathcal{O}(\xi)$ and the map $\gamma \mid \mathcal{O}(\xi) \times \mathcal{U}$ are analytic. ∎

Proof of Theorem 2. We shall show here that all the conclusions of the Theorem follow from the lemma. We start by imposing on $O(\xi)$ the submanifold structure $\mathcal{O}(\xi)$ given by the lemma. Statement (1) then holds. From lemma 3.5, we know that $\mathcal{D}_\Sigma(\zeta) \subseteq T_\zeta\mathcal{O}(\zeta)$ for each ζ in the orbit. In particular, $s \leq r$. By property (b) in lemma 3.7, applied with $\zeta = \xi$, there is some ω so that $(\theta^b_\xi)_*[\omega]$ has rank r; since by lemma 3.6 the image of this differential is included in $\mathcal{D}_\Sigma(\xi)$, it follows that

$$s = r. \tag{26}$$

Thus $\mathcal{O}(\xi)$ has dimension s, as wanted, and this also establishes that $\mathcal{O}(\xi)$ is an integral manifold, property (3), as well as (2). Statement (a) in the lemma together with remark 2.9 show that there is an induced subaction on $\mathcal{O}(\xi)$. If there were any other submanifold structure $\mathcal{N}$ on $O(\xi)$ inducing a subaction, then (a) in the lemma is satisfied for $\mathcal{N}$; if $\mathcal{N}$ has dimension s then the equality (26) gives that (b) holds for $\mathcal{N}$ too, hence $\mathcal{N} = \mathcal{O}(\xi)$ by the uniqueness assertion in the lemma.

Applying part (3) of the Theorem at each $\xi \in \mathcal{M}$, we know that the orbits $\mathcal{O}(\xi)$ constitute a partition of $\mathcal{M}$ into integral manifolds of $\mathcal{M}$; this provides the integrability of $\mathcal{D}_\Sigma$.

Finally, we prove statement (4). Assume that Υ is a topology on $O(\xi)$ for which the mappings θ^b_ξ are all continuous. Pick any $\zeta \in O(\xi)$; we need to show that for each open neighborhood V of ζ relative to Υ there is a $V_1 \subseteq V$ which is a neighborhood of ζ relative to $\mathcal{O}(\xi)$ (with the topology given by the differentiable structure in the Theorem). By part (2) and the implicit function theorem applied to θ^b_ξ as a mapping into $\mathcal{O}(\xi)$, there is a sequence b of $+$'s and $-$'s and an open subset W of $\mathcal{U}^k$ such that θ^b_ξ gives a diffeomorphism between W and its image in the topology of $\mathcal{O}(\xi)$, and ζ is in this image. Then

$$V_1 := \theta^b_\xi((\theta^b_\xi)^{-1}(V) \cap W)$$

is as desired. ∎

Note that the claim made in remark 2.10 regarding the uniqueness of the manifold structure of minimal dimension is a consequence of the Theorem and of lemma 3.5, since the latter implies that any such structure must have dimension at least s.

3.3 Using the distribution to determine accessibility

One of the main applications of the integrability theorem is in verifying the various controllability properties by checking appropriate rank conditions on vector fields.

The integrability statement says that in particular the dimension of $\mathcal{O}(\xi)$ is the same as the rank of $\mathcal{D}_\Sigma(\xi)$. Note that if $\mathcal{D}_\Sigma$ has full rank at a given ξ this means that $\mathcal{O}(\xi)$ has the same dimension as $\mathcal{M}$, and hence that it is open in $\mathcal{M}$, that is, the accessibility property holds from ξ. Conversely, if $\mathcal{O}(\xi)$ contains an open subset of $\mathcal{M}$, and if in addition it is second countable in its own topology, it must have the same dimension as $\mathcal{M}$. From the discussion in section 2.1, we conclude the following.

Proposition 3.8 If $\mathcal{D}_\Sigma$ is full rank at ξ then the accessibility property holds from ξ. Conversely, if the accessibility property holds from ξ and $\mathcal{O}(\xi)$ is a second countable manifold, then $\mathcal{D}_\Sigma$ is full rank at ξ. ∎

The converse statement does not hold in general. Take for example the case of continuous time systems with input topology, example 2.15. In general, orbits will not be second countable here. For a trivial counterexample, take the equation

$$\dot{x}(t) = 1$$

with $\mathcal{M} = \mathbb{R}$ and any $\mathcal{K}$. The accessibility property holds from every point, since the system is transitive. But $\mathcal{D}_\Sigma$ is identically zero, so not full rank. In fact, in this example, the (unique) orbit is the zero dimensional submanifold $\mathbb{R}_{\mathrm{disc}}$ mentioned in section 2.1.

On the other hand, when $\mathcal{U}$ is second countable, as with discrete time systems (16), $\mathcal{O}(\xi)$ is necessarily second countable in its topology. This is because $\mathcal{O}(\xi)$ can be described as the union of the (countably many) images of the continuous mappings (9), and each of these is defined on an open, hence second countable, subset.

Proposition 3.9 For discrete time systems, accessibility from ξ is equivalent to full rank of $\mathcal{D}_\Sigma$ at ξ. ∎

The same result holds for continuous time systems with the time-topology. More generally, we shall use the following concept.

Definition 3.10 The action Σ is *connected* if the following property holds. For each $(\xi, u) \in D$ there exists a $\zeta \in \mathcal{M}$, some sequence $b \in \mathcal{A}^+$, and a pair of elements ω_1, ω_2 in the same connected component of D^b, such that, with the notations in equation (9), $\gamma^b(\zeta, \omega_1) = \xi$ and $\gamma^b(\zeta, \omega_2) = \gamma_u(\xi)$. ∎

If an action is connected, then continuity of γ^b on its second factor implies that ξ and $\gamma_u(\xi)$ are in the same component of $\mathcal{O}(\xi)$, and iterating this we conclude that orbits are connected submanifolds of the paracompact manifold $\mathcal{M}$, hence second countable. So:

Proposition 3.11 For connected actions, accessibility from ξ is equivalent to full rank of $\mathcal{D}_\Sigma$ at ξ. ∎

Time-topology actions associated to continuous time systems are always connected. Indeed, if $\exp(tX)(\xi)$ is defined then

$$\exp(-\varepsilon X)\exp((s+\varepsilon)X)(\xi)$$

is defined for each $s \in [0,t]$ for any fixed small enough $\varepsilon > 0$. We can then use $b = (-,+)$, $\omega_1 = (-\varepsilon, X)(\varepsilon, X)$, and $\omega_2 = (-\varepsilon, X)(t+\varepsilon, X)$.

Corollary 3.12 For time-topology actions, accessibility from ξ is equivalent to full rank of $\mathcal{D}_\Sigma$ at ξ. ∎

3.4 Proof of the integrability lemma

In the proof, the letters χ, ω, possibly primed, will always denote elements of sets of the form $\mathcal{U}^b$ (that is, sequences of elements of $\mathcal{U}$), while b, c stand for words in $\mathcal{A}^*$. Fix an $\xi \in \mathcal{M}$ as in the statement of the lemma, and let $O = O(\xi)$. For each b, we shall use θ^b instead of θ^b_ξ to denote the map $\gamma^b(\xi, \cdot)$; its domain is

$$L^b := \{\omega | (\xi, \omega) \in D^b\}.$$

We first prove that $r(\xi, \zeta) = r(\xi, \zeta')$ for any ζ, ζ' in O. Pick b, c in $\mathcal{A}^*$ and ω, ω' such that

$$\theta^b(\omega) = \zeta \;\text{ and }\; \theta^c(\omega') = \zeta'$$

with full rank:

$$\operatorname{rank} \theta^b_*[\omega] = r(\xi, \zeta) \quad \text{and} \quad \operatorname{rank} \theta^c_*[\omega'] = r(\xi, \zeta').$$

Introduce

$$e := (b, -b, c), \quad \chi := \omega\widetilde{\omega}\omega'.$$

Since $\theta^e(\chi) = \zeta'$,

$$\operatorname{rank} \theta^e_*[\chi] \le r(\xi, \zeta'). \tag{27}$$

Let $F := \gamma^{(-c,b)}(\cdot, \widetilde{\omega}'\omega)$. Then, for each $v \in \mathcal{U}^b$,

$$\gamma^b(\xi, v) = \gamma^{(b,-b,c,-c,b)}(\xi, v\widetilde{\omega}\omega'\widetilde{\omega}'\omega) = F(\gamma^e(\xi, v\widetilde{\omega}\omega')).$$

It follows that

$$r(\xi, \zeta) = \operatorname{rank} \left.\frac{\partial}{\partial v}\right|_{v=\omega} \gamma^b(\xi, v) = \operatorname{rank} F_*[\zeta'] \circ \left.\frac{\partial}{\partial v}\right|_{v=\omega} \gamma^e(\xi, v\widetilde{\omega}\omega')$$

and therefore by (27),

$$r(\xi, \zeta) \le \operatorname{rank} \theta^e_*[\chi] \le r(\xi, \zeta').$$

Interchanging the roles of ζ and ζ' establishes the other inequality, as desired.

Let r be the common value of the $r(\xi, \zeta)$, for this fixed ξ, and consider the set S of all triples (b, Q, h), where b is in $\mathcal{A}^*$ and:

$$Q \text{ is an } r- \text{ dimensional embedded submanifold of } L^b \tag{28}$$

$$\theta^b \mid Q : Q \to \mathcal{M} \text{ is injective and has differential of constant rank } r \tag{29}$$

$$h : Q \to \mathbb{R}^r \text{ is a diffeomorphism with an open subset } h(Q). \tag{30}$$

In the analytic case, we restrict to triples such that Q is an analytic manifold and h is an analytic diffeomorphism.

Fix one such triple, and consider the set $\theta^b(Q) \subseteq O$. The bijection $\theta^b \mid Q$ induces a canonical manifold structure on this set for which both $\theta^b \mid Q$ and $\varphi := h \circ (\theta^b \mid Q)^{-1}$ are diffeomorphisms and φ is a chart. We now prove that with respect to this structure,

(i) the inclusion $i : \theta^b(Q) \to \mathcal{M}$ has injective differential at every point, and

(ii) for any smooth structure $\mathcal{O}$ on O for which the lemma holds, the subset $\theta^b(Q)$ is open relative to $\mathcal{O}$ and the identity map provides a diffeomorphism between the two structures.

The inclusion i factors as

$$\theta^b \circ j \circ (\theta^b \mid Q)^{-1},$$

where j is the embedding of Q in L^b. Therefore property (i) follows from the corresponding properties for its factors (for θ^b, the properties hold on the restriction to Q, which is sufficient). We now prove (ii). Consider θ^b as a map from L^b into $\mathcal{O}$; this map is smooth because of the assumed property (a) in the lemma, since θ^b is a restriction of γ^b. So $\theta^b \mid Q$ is also smooth into $\mathcal{O}$. Since the latter is a submanifold of $\mathcal{M}$ and $\theta^b_* \mid Q$ has constant rank r as a map into $\mathcal{M}$, this rank is also r as a map into $\mathcal{O}$. But this submanifold has dimension r, by part (b) of the lemma. Thus $\theta^b(Q)$ is indeed open relative to $\mathcal{O}$, and $\theta^b \mid Q$ is a diffeomorphism between $\theta^b(Q)$ as a subset of $\mathcal{O}$ and Q. We have then proved that both (i) and (ii) hold.

We now establish that the family of all such charts $(\theta^b(Q), \varphi)$ defines a smooth r-dimensional manifold structure on O, analytic in the case of analytic actions. It will then follow from (i) above that this structure makes O into a submanifold of $\mathcal{M}$, and the uniqueness statement follows from (ii). We start by showing that the sets $\theta^b(Q)$ cover O. Indeed, pick any ζ in O and let b, ω be such that

$$\gamma^b(\xi, \omega) = \zeta, \ \ \operatorname{rank} \theta^b_*[\omega] = r.$$

Since θ^b_* has maximal rank at ω, there is an r-dimensional embedded submanifold Q of L^b containing ω such that both equations (28) and (29) are satisfied. Replacing if necessary Q by an open subset, a suitable h can be found so that equation (30) holds too. Thus these sets form a covering. It is only left to prove compatibility of different charts. For this, pick any two charts $(\theta^b(Q), \varphi)$ and $(\theta^c(P), \beta)$ corresponding to (b, Q, h) and (c, P, k) respectively.Let

$$V := \theta^b(Q) \cap \theta^c(P).$$

We need to establish that:

(I) $\varphi(V)$ is open in $\varphi(\theta^b(Q))$.

(II) $\beta \circ \varphi^{-1}$ is smooth on V.

Pick an arbitrary ζ in V. By definition of the set V, there are ω, ω' in Q and P respectively so that

$$\zeta = \theta^b(\omega) = \theta^c(\omega')$$

with full rank r. Let $e := (b, -c, c)$ in $\mathcal{A}^*$, and take $\chi := \omega\widetilde{\omega}'\omega'$, so that

$$\theta^e(\mu\widetilde{\omega}'\omega') = \theta^b(\mu) \ \text{ for all } \ \mu \in L^b \tag{31}$$

and

$$\theta^e(\omega\widetilde{\omega}'\mu) = \theta^c(\mu) \ \text{ for all } \ \mu \in L^c. \tag{32}$$

Therefore

$$r \geq \operatorname{rank} \theta^e_*[\chi] \geq \operatorname{rank} \theta^b_*[\omega] = r,$$

and $\theta^e_*[\chi]$ must have maximal rank too. So there is an open subset Z of L^e which contains χ and such that $\theta^e(Z)$ is an r-dimensional embedded submanifold of $\mathcal{M}$. Introduce the open sets

$$W := \{\mu \in L^b \,|\, \mu\widetilde{\omega}'\omega' \in Z\}$$

and

$$W' := \{\mu \in L^c \,|\, \omega\widetilde{\omega}'\mu \in Z\}.$$

Since $\chi \in Z$ also $\omega \in W$ and $\omega' \in W'$. Consider the sets $P' := P \cap W'$ and $Q' := Q \cap W$. Since Q is an embedded submanifold of L^b and W is open in L^b, it follows that also Q' is open in Q, and similarly for P, P'. By (31) and the definition of Z it follows that $\theta^b \mid Q'$ maps into $\theta^e(Z)$, and it is injective with differential of constant rank r. Thus θ^b establishes a diffeomorphism between Q' and an open subset $\mathcal{E}$ of $\theta^e(Z)$. Similarly, using (32), for $\theta^c \mid P'$ and an open $\mathcal{F}$ in $\theta^e(Z)$. Note that

$$\mathcal{E} \cap \mathcal{F} \subseteq V.$$

Since ω' and ω are in P', Q' respectively, $\zeta \in \mathcal{E} \cap \mathcal{F}$. By injectivity of $\theta^b \mid Q$, $(\theta^b \mid Q)^{-1}(\mathcal{E} \cap \mathcal{F})$ equals $(\theta^b \mid Q')^{-1}(\mathcal{E} \cap \mathcal{F})$. Continuity of the restrictions and openness of $\mathcal{E} \cap \mathcal{F}$ in $\mathcal{E}$ imply that this is open, so also

$$\varphi(\mathcal{E} \cap \mathcal{F})$$

is an open subset of

$$h(Q) = \varphi(\theta^b(Q)).$$

Thus $\varphi(\zeta)$ has a neighborhood included in $\varphi(\theta^b(Q))$, and (I) follows. To prove (II), note that φ maps $\mathcal{E} \cap \mathcal{F}$, seen as an embedded submanifold of $\theta^e(Z)$, diffeomorphically onto $\varphi(\mathcal{E} \cap \mathcal{F})$. This latter set is open in $h(Q)$ and contains $\varphi(\zeta)$. A similar statement holds for β. So

$$\beta \circ \varphi^{-1} : \varphi(\mathcal{E} \cap \mathcal{F}) \simeq \beta(\mathcal{E} \cap \mathcal{F}),$$

and the second statement (II) follows too. We then have a manifold structure as desired. By construction, it is an analytic structure for analytic actions.

Finally, we must prove the joint smoothness of γ^b on u and ξ, property (a) of the lemma, with respect to the above manifold structure. We first establish that each of the maps θ^b is smooth. Pick $\omega \in L^b$, $\zeta = \theta^b(\omega)$. Since $r(\xi, \zeta) = r$, there are c, ω' so that $\theta^c(\omega') = \zeta$ with full rank r. Let

$$e := (b, -c, c), \quad \chi := \omega\widetilde{\omega}'\omega'.$$

It will suffice to prove that θ^e is smooth on some neighborhood of χ, because

$$\theta^b(\mu) = \theta^e(\mu\widetilde{\omega}'\omega')$$

for each μ in a neighborhood of ω. Since $\theta^{(b,-c)}(\omega\widetilde{\omega}') = \xi$, also $\theta^c(\mu) = \theta^e(\omega\widetilde{\omega}\mu)$, so

$$r \geq \operatorname{rank} \theta^e_*[\chi] \geq \operatorname{rank} \theta^c_*[\omega'] = r.$$

So θ^e achieves maximal rank at χ. There is then a chart C of L^e, centered at χ, and diffeomorphic to a cube in $\mathbb{R}^s \times \mathbb{R}^r$, such that, if Q is the embedded submanifold corresponding to the factor $\mathbb{R}^r$, then $\operatorname{rank} \theta^e_*$ is constantly r on Q and θ^e is injective on Q. Let h give the corresponding diffeomorphism of Q with $\mathbb{R}^r$. Then (e, Q, h) gives rise to a chart, say $(\theta^e(N), \varphi)$. So $\theta^e \mid C$ is the composition of the projection onto Q and of $\theta^e \mid Q$, and is therefore smooth. Thus θ^b is smooth.

To prove now that each γ^b is smooth as a map into $\mathcal{O}$, pick any (ξ, ω) in D^b, ξ in $\mathcal{O}$. Let (c, Q, h) give a chart around ξ. For (ζ, χ) in a neighborhood of (ξ, ω) relative to

$$(\mathcal{O} \times \mathcal{U}^b) \cap D^b,$$

$\gamma^b(\zeta, \chi)$ equals the composition

$$\theta^{(c,b)}_\xi((\theta^c_\xi \mid Q)^{-1}(\zeta), \chi),$$

and it is therefore smooth because of the smoothness of $\theta^{(c,b)}_\xi$ and of θ^c_ξ. This gives property (a) of the lemma, as wanted. ∎

4 Lie algebraic conditions and invariance

In this section we study the relationships between various properties of distributions and their implications for actions.

4.1 Some Lie algebraic criteria

We introduce a Lie algebra associated to each action.

Definition 4.1 The *Lie algebra* $\mathcal{L}_\Sigma$ *associated to* Σ is $\mathrm{vf}(\Sigma)_{LA}$. ∎

The distribution $\mathcal{D}_\Sigma$ is in general very difficult to compute for actual examples, since it involves the arbitrary compositions γ^b of the basic mappings γ defining the action. In examples like those associated to continuous time systems, this would involve having explicit solutions for the differential equations defining the system. On the other hand, $\mathcal{L}_\Sigma$ can be typically readily computed using symbolic manipulation systems, by iterating the application of the Lie bracket operation to the set of vector fields in $\mathrm{vf}(\Sigma)$, and the latter can be often obtained easily, as illustrated in the next section. Thus the following fact is of interest.

Proposition 4.2 The distribution $\mathcal{D}(\mathcal{L}_\Sigma)$ is included in $\mathcal{D}_\Sigma$.

Proof. By Theorem 2, $\mathcal{D}_\Sigma$ is integrable and therefore involutive, that is, $\mathrm{vf}(\mathcal{D}_\Sigma)$ is an involutive set of vector fields. With $\Phi := \mathrm{vf}(\Sigma)$, we then have that

$$\Phi \subseteq \mathrm{vf}(\mathcal{D}_\Sigma) \Rightarrow \Phi_{LA} \subseteq \mathrm{vf}(\mathcal{D}_\Sigma) \Rightarrow \mathcal{D}(\Phi_{LA}) \subseteq \mathcal{D}(\mathrm{vf}(\mathcal{D}_\Sigma)) = \mathcal{D}_\Sigma,$$

the last equality because $\mathcal{D}_\Sigma$ is smooth. ∎

It follows that checking if $\mathcal{D}(\mathcal{L}_\Sigma)$ has full rank provides a sufficient condition for transitivity. This condition is however far from necessary, even for continuous time systems for which the rank of $\mathcal{D}_\Sigma$ *does* provide a necessary criterion. This is because the converse of proposition 4.2 is not true in general. For time-topology actions corresponding to continuous time systems, however, such a result will be given in section 4.6 for the analytic and other important special cases. Roughly speaking, $\mathcal{D}(\mathcal{L}_\Sigma)$ provides information based only on local data, while $\mathcal{D}_\Sigma$ uses all global information.

4.2 Explicit form of the vector fields

The vector fields in $\mathrm{vf}(\Sigma)$ can be often computed in closed form if a closed form expression is available for γ. This is typically the case for actions corresponding to discrete time systems. Furthermore, when $\mathcal{K}$ is an open subset of $\mathbb{R}^m$ for some m, it is only necessary for each u to compute the vectors $\mathcal{X}_{u,\nu}$ for each of the m canonical tangent vectors $\nu = e_i = (0, \ldots, 0, 1, 0, \ldots, 0)'$. Any other element of $\mathrm{vf}(\Sigma)$ is a linear combination of these. Also, because of formula (23), there is no need to compute the inverses of the mappings γ_u, just the inverses of their Jacobians. This is useful when there is no easy manner to obtain these inverses. For example, take the discrete time system with $\mathcal{M} = \mathbb{R}, \mathcal{K} = [-1, 1]$ and

$$P(\xi, u) := \xi^3 + 2\xi + u \sin \xi.$$

This is strictly increasing for each fixed value of u, so it does give rise to an invertible system, but there is no closed form expression for the inverses of the maps $P(\cdot, u)$. Since $\mathcal{K} = \mathbb{R}$, it is enough to compute just one vector, for $\nu = 1$:

$$X_u(\xi) = \frac{\partial P}{\partial u} / \frac{\partial P}{\partial \xi} = \frac{\sin \xi}{3\xi^2 + 2 + u \cos \xi}.$$

We now compute $\mathrm{vf}(\Sigma)$ explicitely for the case of continuous time systems with time topology in definition 2.15, or more generally for example 2.12. It is trivial here to find $\mathrm{vf}(\Sigma)$: if Φ is a family of vector fields and $X \in \Phi$ is defined at ξ then for each $u = (t, X)$ so that $\exp(tX)(\xi)$ is defined and using $\nu = (1, 0)$, we obtain that (22) equals

$$\mathcal{X}_{t,X}(\xi) = \left.\frac{\partial}{\partial s}\right|_{s=t} \exp(-tX)\exp(sX)(\xi) = \left.\frac{\partial}{\partial s}\right|_{s=t} \exp((s-t)X)(\xi) \tag{33}$$

$$= \left.\frac{\partial}{\partial v}\right|_{s=0} \exp(sX)(\xi) = X(\xi). \tag{34}$$

(Independent of t.) The vectors in the tangent space to $\mathcal{U}$ are all of the form $(r, 0), r \in \mathbb{R}$, so all the possible vectors in $\mathrm{vf}(\Sigma)$ are multiples of this. We conclude that for time-topology actions associated to continuous time systems, $\mathrm{vf}(\Sigma) = \{X_u, u \in \mathcal{K}\}$ and more generally for actions as in definition 2.12:

Proposition 4.3 For $\Sigma = \Sigma(\Phi)$, $\mathrm{span}\,\mathrm{vf}(\Sigma) = \mathrm{span}\,\Phi$, $\mathcal{D}(\Sigma(\Phi)) = \mathcal{D}(\Phi)$, and $\mathcal{L}_\Sigma = \Phi_{LA}$. ∎

Proposition 4.4 For systems affine in control as in equation (13),

$$\mathcal{L}_\Sigma = \{f, g_1, \ldots, g_m\}_{LA}.$$

(Recall that $\mathcal{K} \subseteq \mathbb{R}^m$ must have a nonempty interior.) ∎

For the case of continuous time systems with input topology, for simplicity we restrict attention to analytic systems (13). For each $i = 1, \ldots, m$, each positive t, and each $u \in \mathbb{R}^m$, we must consider

$$\mathcal{X}_{t,u,i}(\xi) := \left.\frac{\partial}{\partial v_i}\right|_{v=u} \exp(-t(f + u_1 g_1 + \ldots + u_m g_m))\exp(t(f + v_1 g_1 + \ldots + v_m g_m))(\xi). \tag{35}$$

For small t, we may use the expansion ([6]), with $X_u = f + u_1 g_1 + \ldots + u_m g_m$:

$$\mathcal{X}_{t,u,i}(\xi) = \sum_{i=1}^{\infty} \frac{t^i}{i!} \mathrm{ad}^{i-1}(X_u)(g_i)(\xi). \tag{36}$$

Here $\mathrm{ad}^k(X)(Y)$ denotes the iterated Lie bracket

$$[X, [X, [\ldots [X, Y] \ldots]]]$$

(k times). The explicit form (36) for the vector fields associated to the input-topology action forms the basis basis in [15] and [16] for an eigenvalue criterion for sampling.

4.3 Criteria involving Lie algebras of vector fields

We know from propositions 3.8 that if $\mathcal{D}_\Sigma$ is full rank at ξ then the accessibility property holds from ξ. With the stronger statement that $\mathcal{L}_\Sigma$ has full rank, it is possible to obtain results on reachability. Given any $(\zeta, u) \in D$, consider the differential of γ at (ζ, u). For any vector $\sigma \in T_\zeta\mathcal{M}$,

$$\left.\frac{\partial}{\partial x}\right|_{x=\zeta} \gamma(x, u)(\sigma) = (\gamma_u)_*[\zeta](\sigma) \tag{37}$$

by definition, while for each $\nu \in T_u\mathcal{U}$,

$$\left.\frac{\partial}{\partial v}\right|_{v=u} \gamma(\zeta, v)(\nu) = \left.\frac{\partial}{\partial v}\right|_{v=u} \gamma(\gamma^-(\gamma(\zeta, v), u), u)(\nu) = (\gamma_u)_*[\zeta](\mathcal{X}_{u,\nu}(\zeta)). \tag{38}$$

Thus, if the vectors $\sigma_1, \ldots, \sigma_r \in T_\zeta\mathcal{M}$ are linearly independent and if $\mathcal{X}_{u,\nu}(\zeta)$ is not in their span, it follows that $\gamma_*[\zeta, u]$ maps

$$\sigma_1, \ldots, \sigma_r, \mathcal{X}_{u,\nu}(\zeta)$$

(seen as tangent vectors to $\mathcal{M} \times \mathcal{U}$ at (ζ, u)) into a set of $r + 1$ linearly independent vectors. From this it follows that when S is an r-dimensional submanifold containing ζ and the σ_i are a basis of $T_\zeta S$, if $\mathcal{X}_{u,\nu}(\zeta)$ is not in $T_\zeta S$ then $\gamma(S \times \mathcal{U})$ must contain a submanifold of $\mathcal{M}$ of dimension at least $r + 1$ (implicit mapping theorem). We therefore have the following result.

Theorem 3 *If $\mathcal{L}_\Sigma$ has full rank then Σ has the accessibility property from every $\xi \in \mathcal{M}$ and, furthermore, the set of states that can be reached from each ξ contains an open subset of $\mathcal{M}$.*

Proof. Let s be the maximal possible dimension of a submanifold S of $\mathcal{M}$ which is included in the set of states reachable from ξ. (There is always such a submanifold, for instance $S = \{\xi\}$.) If some vector field $\mathcal{X}_{u,\nu}(\zeta)$ were not to be defined but not tangent to S at some $\zeta \in S$, then the above argument shows that $\gamma(S \times \mathcal{U})$ contains a submanifold of dimension $r + 1$, all whose points are reachable from ξ, contradicting maximality of r. Thus each vector field in $\mathrm{vf}(\Sigma)$ is tangent to S at every point of S. So also every vector field in $\mathcal{L}_\Sigma$ must be tangent, since S is a submanifold. It follows from the full rank assumption that S must have the dimension as $\mathcal{M}$, and therefore it is open as wanted. ∎

The accessibility statement also follows as a consequence of proposition 4.2, since full rank of the Lie distribution implies full rank of $\mathcal{D}_\Sigma$, and hence $\mathcal{O}(\xi)$, an integral manifold of this distribution, must have full dimension.

Remark 4.5 The above theorem is a particular case of a stronger result established in [10]; the result there says that the same conclusion is true if one considers instead the Lie algebra generated by the larger set consisting of all vector fields of the form (25) with $b \in \mathcal{A}_+^*$. ∎

4.4 Foliations

For completeness, we prove here that integrability coincides with the existence of singular foliations. Together with Theorem 2, this will mean that the connected components of the possible orbits are the maximal integral manifolds of $\mathcal{D}_\Sigma$. In general, if $\mathcal{D}$ is an integrable distribution and a partition into integral manifolds is given, we may refine this partition by taking all connected components of each element. This exhibits $\mathcal{M}$ as the disjoint union of connected integral submanifolds of $\mathcal{D}$; the next result shows that these are then maximal integral manifolds.

Proposition 4.6 If $\{\mathcal{N}_\lambda, \lambda \in \Lambda\}$ is a partition of $\mathcal{M}$ into connected integral submanifolds of a distribution $\mathcal{D}$, then each $\mathcal{N}_\lambda$ is a maximal integral manifold of $\mathcal{D}$.

Proof. We first prove that if $\mathcal{N}$ and $\mathcal{N}'$ are two integral manifolds of a distribution $\mathcal{D}$ then their intersection is open in each of them. For this, we pick any $\xi \in \mathcal{N} \cap \mathcal{N}'$ and claim that there is a set V which is a neighborhood of ξ in $\mathcal{N}'$ such that V is entirely contained in $\mathcal{N}$. Thus ξ is in the interior of $\mathcal{N} \cap \mathcal{N}'$ with respect to $\mathcal{N}'$, and the same argument with the roles of $\mathcal{N}$ and $\mathcal{N}'$ reversed gives that ξ is in the interior of this intersection with respect to $\mathcal{N}$.

Without loss of generality, we may assume for the local statement that $\mathcal{N}'$ is an embedded submanifold of $\mathcal{M}$, say an open subset of the slice $x_{k+1} = \ldots = x_n = 0$ in a coordinate chart about ξ, and that ξ is 0 under these coordinates. In particular, the vector fields $\partial/\partial x_i$, $i = 1, \ldots, k$ are pointwise in $\mathcal{D}$, and hence are tangent to both $\mathcal{N}$ and $\mathcal{N}'$. Consider the control system

$$\dot{x} = \sum_{i=1}^{k} u_i \frac{\partial}{\partial x_i}(x)$$

with controls in $\mathbb{R}^k$, defined on this coordinate chart. This differential equation has a solution in $\mathcal{N}$ for small t and $x(0) = 0$, for each constant control $u(t) \equiv (a_1, \ldots, a_k)$, $a_i \in \mathbb{R}$. Moreover, by continuity of solutions on parameters, there are $T, \varepsilon > 0$ such that the solution exists on $[0, T]$ for every constant control with $\sum_i a_i^2 = \varepsilon$. Thus the unit ball V of radius $T\varepsilon$ in the slice is included in $\mathcal{N}$. This is an open subset of $\mathcal{N}'$ if $T\varepsilon$ is small enough, because the slice is an open subset of $\mathcal{N}'$. This establishes the claim.

Assume now that $\mathcal{N}'$ is any connected integral manifold of $\mathcal{D}$. Then, it is contained in precisely one of the $\mathcal{N}_\lambda$. Indeed, $\mathcal{N}' = \cup(\mathcal{N}' \cap \mathcal{N}_\lambda)$, and the above argument gives that each of the elements in this union is open in $\mathcal{N}'$, so by connectedness there must be exactly one nonempty intersection. Finally, $\mathcal{N}' = \mathcal{N}' \cap \mathcal{N}_\lambda$ is open in $\mathcal{N}_\lambda$, again by the same argument. ∎

4.5 Accessible sets for systems are included in orbits

In this section, we establish proposition 2.16 and a related result for discrete time systems. We start with this lemma.

Lemma 4.7 Consider a continuous time system (11) and assume that $\mathcal{D}$ is a distribution with the property that $P(\xi, u) \in \mathcal{D}(\xi)$ for each (ξ, u). Suppose that there is a partition $\{\mathcal{N}_\lambda, \lambda \in \Lambda\}$ of $\mathcal{M}$ into integral submanifolds of $\mathcal{D}$, and that ξ is controllable to ζ. Then, ξ and ζ are in the same element $\mathcal{N}_\lambda$ of the partition.

Proof. Assume that $u : [0,T] \to \mathcal{K}$ is the control steering ξ into ζ, and let $x(t)$ be the corresponding solution with $x(0) = \xi, x(T) = \zeta$. Pick any fixed $t_0 \in [0,T]$, denote $\xi' := x(t_0)$, and let $\mathcal{N}$ the element $\mathcal{N}_\lambda$ of the partition which contains ξ'. We claim that there is a neighborhood V of t_0 in $[0,T]$ such that $u(V) \subseteq \mathcal{N}$. Since the vector fields $X_u = P(\cdot, u)$ are tangent to (the integral manifold) $\mathcal{N}$, we may also consider the controlled differential equation (11) as an equation on $\mathcal{N}$. The continuity requirements on X_u and its derivatives are still satisfied for this equation, from which it follows that there is a solution $x'(t)$ of the equation for t near t_0 with initial condition $x'(t_0) = \xi'$ and so that $x'(t) \in \mathcal{N}$ on its domain. (Solving the equation backwards gives a solution for $t < t_0$; if t_0 is an endpoint, we only have a solution in one direction.) By uniqueness of solutions over $\mathcal{M}$, $x'(t) = x(t)$ where defined, and so the claim is established. We conclude that for each λ the set $\{t | u(t) \in \mathcal{N}_\lambda\}$ is open. Connectedness of $[0,T]$ then implies the lemma. ∎

Proof of proposition 2.16. We apply the above with $\mathcal{D} := \mathcal{D}_\Sigma$, using the time-topology action Σ. The possible orbits will be the elements of the partition $\mathcal{N}_\lambda$; Theorem 2 insures that these are integral manifolds. It is only necessary to see that $X_u(\xi) \in \mathcal{D}(\xi)$ for all $u \in \mathcal{K}$, so that the lemma can be applied and one may conclude that ξ and ζ are in the same orbit whenever one state is accessible from the other. But this follows from proposition 4.3. ∎

Remark 4.8 If $\mathcal{K}'$ is a dense subset of $\mathcal{K}$, then controls with values in $\mathcal{K}'$ give the same orbit. This is because the above proof can be applied with the distribution $\mathcal{D}'$ corresponding to using $\mathcal{K}'$, but the distributions are the same. Indeed, because $X_u(\xi)$ depends continuously on u and $\mathcal{D}'(\xi)$ is a finite dimensional vector space, we have that also those $X_u(\xi)$ with $u \in \mathcal{K}$ are in this space, as desired. ∎

For discrete time systems, one could use sets with nice boundary as control sets, and accessibility would not change:

Proposition 4.9 If $\mathcal{C}$ is an open subset of $\mathcal{K}$ with nice boundary and $P(\zeta, u) = \zeta'$ for some u in $\mathcal{C}$, then ζ' is in the orbit of ζ with respect to the discrete time system that uses $\operatorname{int}\mathcal{C}$ as control value set.

Proof. We apply the lemma now with $\mathcal{D} = \mathcal{D}_\Sigma$, where Σ is the action associated to the discrete time system having $\mathcal{U} := \operatorname{int}\mathcal{C}$. The $\mathcal{N}_\lambda$'s are the orbits for this action. Assume that $P(\zeta, u) = \zeta'$, $u \in \mathcal{C}$. By the nice boundary assumption, there is some smooth curve ρ with $\rho(0) = u$ and $\rho(t) \in \mathcal{U}$ for $t \in (0,1]$. Since the domain of P is open, there is some $T \leq 1$ such that for each $t \in (0,T]$ the element

$$x(t) := P(\zeta, \rho(t)) = \gamma(\zeta, \rho(t))$$

is defined and belongs to the orbit $\mathcal{O}(\zeta)$. We claim that $\zeta' = x(0)$ is in the same orbit as $\zeta'' := x(T)$. By transitivity it will then follow that ζ and ζ' are in the same orbit, which is what we want to establish. For each fixed t in this interval, we compute the derivative of the curve $x(t)$. This is

$$\dot{x}(t) = \frac{\partial}{\partial v}\Big|_{v=\rho(t)} \gamma(\zeta, v)(\dot{\rho}(t)),$$

which is in $\mathcal{D}(x(t))$ for each t, by the fact that $\mathcal{D}$ includes (24). If we introduce the vector fields

$$\mathcal{X}'_{u,\nu}(\xi) := \frac{\partial}{\partial v}\bigg|_{v=u} \gamma(\gamma^{-}(\xi,u),v)(\nu),$$

we may think of $\dot{x} = \mathcal{X}'_{\rho(t),\dot{\rho}(t)}(x)$ as a continuous time system whose control value space is the tangent bundle to $\mathcal{U}$. Applying the lemma, the whole trajectory must remain in one orbit, as wanted. ∎

4.6 Invariance and integrability

We now concentrate on actions corresponding to time topologies for continuous time systems, or more generally actions of the type $\Sigma(\Phi)$. Let Φ be a set of vector fields, $\mathcal{D} := \mathcal{D}(\Phi)$, and $\Sigma := \Sigma(\Phi)$. "Invariance" will mean in this section $\Sigma(\Phi)$-invariance. By lemma 3.2, we know that $\mathcal{D}$ is invariant if and only if

$$\mathrm{Ad}_{\exp(tX)}\mathcal{D}(\xi) = \mathcal{D}(\xi)$$

whenever $\xi \in \mathcal{M}$, $X \in \Phi$, $t > 0$, and $\exp(tX)(\xi)$ is defined. By proposition 4.3, $\mathcal{D}(\mathrm{vf}(\Sigma)) = \mathcal{D}$. Since $\mathcal{D}_\Sigma$ is the smallest invariant distribution containing this, it follows that:

Lemma 4.10 $\mathcal{D}$ is invariant if and only if $\mathcal{D} = \mathcal{D}_\Sigma$. ∎

Since by Theorem 2 $\mathcal{D}_\Sigma$ is integrable, it follows that an invariant distribution must be integrable. Conversely, if $\mathcal{D}$ is integrable and $\{\mathcal{N}_\lambda, \lambda \in \Lambda\}$ is a partition into maximal integral manifolds, it follows from lemma 4.7, applied to the equations $\dot{x} = X(x)$, for $X \in \Phi$, that each $\mathcal{N}_\lambda$ is stable and therefore that

$$\mathrm{Ad}_{\exp(tX)} T_\xi \mathcal{N}_\lambda(\xi) \subseteq T_\xi \mathcal{N}_\lambda(\xi) \tag{39}$$

for each $\xi \in \mathcal{N}_\lambda$ and each t, X so that the flow is defined. Because $T_\xi\mathcal{N}_\lambda(\xi)$ has the same dimension as $T_{\exp(tX)(\xi)}\mathcal{N}_\lambda(\exp(tX)(\xi))$, equation (39) is an equality. Since these are integral manifolds, $\mathcal{D}$ must be invariant. Thus we established the following result.

Theorem 4 *A smooth distribution $\mathcal{D}$ is invariant if and only if it is integrable.* ∎

Recall that here invariance is under $\Sigma(\Phi)$, for any given set of vector fields such that $\mathcal{D} = \mathcal{D}(\Phi)$. The following local version of invariance is useful.

Lemma 4.11 $\mathcal{D}$ is invariant if and only if for each $\xi \in \mathcal{M}$ and each $X \in \Phi$ defined at ξ there exists an $\varepsilon > 0$ such that

$$\mathrm{Ad}_{\exp(tX)}\mathcal{D}(\xi) \subseteq \mathcal{D}(\xi) \tag{40}$$

for each $t \in (-\varepsilon, \varepsilon)$.

Proof. Note first that it is equivalent to assume equality in equation (40). Indeed, when this is applied for small enough $-t$ it implies, because of equation (20), that $\mathcal{D}(\xi) = \mathrm{Ad}_{\exp(tX)}\mathrm{Ad}_{\exp(-tX)}\mathcal{D}(\xi) \subseteq \mathrm{Ad}_{\exp(tX)}\mathcal{D}(\xi)$.

If the distribution is invariant, then $\mathrm{Ad}_{\exp(tX)}\mathcal{D}(\xi) = \mathcal{D}(\xi)$ for all positive t. But then equality also holds for small negative t, since $\exp(-tX)(\xi)$ is defined for small t, again using equation (20),

Assume now that the local statement in the lemma holds. Pick any X and ξ, and let (t_0, t_1) be the interval of definition of $\exp(tX)(\xi)$. We will prove that (40) holds as an equality for all t in this interval. Let S be the subset of (t_0, t_1) where the equality holds. It will be enough to prove that S is both open and closed. (Note that S is nonempty, since $0 \in S$.) We pick an ε as in the statement. For each $\tau \in (t_0, t_1)$ and each small enough $t \in (-\varepsilon, \varepsilon)$,

$$\mathrm{Ad}_{\exp((t+\tau)X)}\mathcal{D}(\xi) = \mathrm{Ad}_{\exp(\tau X)}\mathrm{Ad}_{\exp(tX)}\mathcal{D}(\xi) = \mathrm{Ad}_{\exp(\tau X)}\mathcal{D}(\xi).$$

Thus if $\tau \in S$ the last term equals $\mathcal{D}(\xi)$, so also $t + \tau \in S$, and S is open. If τ is not in S, the last term is different from $\mathcal{D}(\xi)$, and therefore $t + \tau$ is also not in S, so the complement of S is open as well. ∎

Since $\mathcal{D} = \mathcal{D}(\Phi)$, we see that $\mathcal{D}$ is invariant iff for each $X, Y \in \Phi$, each ξ, and each small t, $\mathrm{Ad}_{\exp(tX)}Y(\xi) \in \mathcal{D}(\xi)$.

We now prove that integrability is equivalent to involutivity provided that $\mathcal{D}$ is determined by analytic vector fields or that certain other sufficient conditions hold. For this, we need the *Baker-Campbell-Hausdorff* formula:

$$\frac{\mathrm{d}^l}{\mathrm{d}t^l}\mathrm{Ad}_{\exp(tX)}Y(\xi) = \mathrm{Ad}_{\exp(tX)}(\mathrm{ad}^l(X)(Y))(\xi) \tag{41}$$

valid for all vector fields X, Y defined at ξ, all t such that the flow is defined, and all nonnegative integers l. Assume that Φ is involutive, and pick X, Y, ξ. Consider

$$\alpha(t) := \mathrm{Ad}_{\exp(tX)}Y(\xi)$$

as a function taking values in the finite dimensional space $T_\xi\mathcal{M}$, which we identify with Euclidean space $\mathbb{R}^n$. Because of involutivity and formula (41), all derivatives of α at $t = 0$ are in the subspace $\mathcal{D}(\xi)$. When both X and Y are analytic, this implies that $\alpha(t) \in \mathcal{D}(\xi)$ for all t, so the distribution is invariant, or equivalently, integrable.

We next show that the same conclusion is true if Φ is an involutive *locally finitely generated* set of vector fields. For involutive Φ, we define this latter property to mean: for each $\xi \in \mathcal{M}$ there are vector fields $Y_1, \ldots, Y_k \in \Phi$ such that any $Y \in \Phi$ can be expressed in some neighborhood V of ξ as

$$Y(\zeta) = \sum_{i=1}^{k} \rho_i(\zeta) Y_i(\zeta) \tag{42}$$

for a set of smooth functions $\rho_1, \ldots, \rho_k$ defined on V. In other words, Φ is locally finitely generated as a module over the ring of smooth functions. Examples of locally finitely generated involutive Φ are *finite dimensional* Lie algebras of vector fields (just take for the Y_i's a basis of this Lie algebra), as well as involutive sets of vector fields for which the distribution $\mathcal{D}(\Phi)$

has constant rank (if $\{Y_1(\xi), \dots, Y_k(\xi)\}$ is a basis of $\mathcal{D}(\xi)$, then $\{Y_1(\zeta), \dots, Y_k(\zeta)\}$ must also be a basis in a neighborhood of ξ; the ρ_i's are then obtained from Cramer's rule). Also, it is possible to prove that distributions generated by analytic vector fields are always locally finitely generated, so this case in fact contains the previous one.

One could define the concept of local finite generation for noninvolutive Φ, but a useful definition, though equivalent to the above under the added assumption of involutivity, in the general case would have to be somewhat more complicated.

Assume now that Φ is involutive locally finitely generated. Pick any ξ and $X, Y \in \Phi$, and a set $\{Y_1, \dots, Y_k\}$ as above. Let ν be any element of $T_\xi\mathcal{M} = \mathbb{R}^n$ perpendicular to all elements of $\mathcal{D}$, $\langle \nu, Y_i(\xi)\rangle = 0$ for all i, and introduce the functions

$$\beta_i(t) := \langle \nu, \mathrm{Ad}_{\exp(tX)} Y_i(\xi)\rangle$$

and the vector $\alpha(t) := (\beta_1(t), \dots, \beta_k(t))'$. If we prove that α is identically zero in a neighborhood of $t = 0$, it will follow that for each $Y \in \Phi$ also $\langle \nu, \mathrm{Ad}_{\exp(tX)} Y(\xi)\rangle \equiv 0$, because Y is a linear combination of the Y_i's and $\mathrm{Ad}_{\exp(tX)}$ is linear. Repeating with a basis $\nu_1, \dots,$ of the annihilator of $\mathcal{D}(\xi)$, we conclude that $\mathrm{Ad}_{\exp(tX)} Y(\xi) \in \mathcal{D}(\xi)$, as desired. Note that $\alpha(0) = 0$. By the local finite generation property, there exist smooth functions ρ_{ij} such that

$$[X, Y_i] = \sum_{j=1}^{k} \rho_{ij} Y_j .$$

Applying formula (41) with $l = 1$, we have that

$$\dot{\alpha}(t) = R(t)\alpha(t),$$

where R is the matrix of the ρ_{ij}'s. Thus $\alpha(t)$ satisfies a homogeneous linear differential equation, so $\alpha(0) = 0$ implies that $\alpha(t) \equiv 0$ as desired. We summarize the conclusions in the next theorem.

Theorem 5 *Assume that Φ is an involutive set of vector fields, $\mathcal{D} = \mathcal{D}(\Phi)$, and one of the following properties holds:*

- *All the vector fields in Φ are analytic.*
- *The span of Φ is finite dimensional.*
- *$\mathcal{D}$ is nonsingular.*

Then $\mathcal{D}$ is integrable. ∎

Note that if $\mathcal{D}$ is involutive, then it can be written as $\mathcal{D}(\mathrm{vf}(\mathcal{D}))$, and thus as $\mathcal{D}(\Phi')$ for some involutive set Φ'. The constant rank part of the Theorem is basically Frobenius Theorem. The analytic version is due to Nagano and Hermann, and the method of proof is due to Lobry and Sussmann.

We now establish a converse of proposition 4.2.

Proposition 4.12 Assume that Σ is an action of the type $\Sigma(\Phi)$. If either $\mathcal{L}_\Sigma$ is analytic, or it spans a finite dimensional space, or the distribution $\mathcal{D}(\mathcal{L}_\Sigma)$ is nonsingular, then $\mathcal{D}(\mathcal{L}_\Sigma) = \mathcal{D}_\Sigma$.

Proof. Recall that by definition $\mathcal{D}_\Sigma$ is the smallest Σ-invariant distribution containing $\mathcal{D}(\mathrm{vf}(\Sigma))$, and by proposition 4.2 it contains $\mathcal{D}(\mathcal{L}_\Sigma)$. By definition the latter includes $\mathcal{D}(\mathrm{vf}(\Sigma))$; thus it will be enough to establish that it is invariant. We apply Theorem 5 with "Φ" there being $\mathcal{L}_\Sigma$. It follows that $\mathcal{D}(\mathcal{L}_\Sigma)$ is invariant under $\mathcal{L}_\Sigma$, and therefore in particular under $\mathrm{vf}(\Sigma)$, which by proposition 4.3 equals Φ. ∎

5 Zero-time orbits and sampling

In this section we wish to show how the previously obtained results can be applied to the problem of preservation of accessibility under sampling. The main fact to be proved is that, for continuous time smooth systems with connected control value set, the "fixed time accessibility property" from a state ξ is a sufficient (as well as a necessary) condition to insure that the sampled systems in (17) satisfy the accessibility property from this ξ as discrete time systems, for each δ small enough. Fixed time accessibility is the requirement that the set of states accessible from ξ in total time zero be a neighborhood of ξ. The idea of the proof is to first show that these orbits are second countable in the input topology, and to then approximate the distribution associated to the input topology by distributions corresponding to the sampled systems. Second countability will insure that distributions have full rank precisely when this strong accessibility condition holds. The sampling result can be alternatively proved using a fixed-point argument, as done in [18]; however the present approach is much more natural, and the second countability proof is of considerable interest in its own right.

For this entire section, Σ is the input-topology action associated to a fixed smooth-in-controls system.

5.1 Zero-time orbits

The *zero-time orbit* $O_0(\xi)$ of $\xi \in \mathcal{M}$ with respect to the system (11) is the set consisting of ξ as well as all states of the form

$$\exp(t_1 X_{u_1}) \dots \exp(t_k X_{u_k})(\xi) \tag{43}$$

obtained for all possible sequences of nonzero numbers $(t_1, \dots, t_k)$ such that

$$\sum_{i=1}^{k} t_i = 0$$

and all positive integers k and all sequences of elements $u_1, \dots, u_k \in \mathcal{K}$. The state space $\mathcal{M}$ can be partitioned into zero-time orbits. Our first claim is that these are open subsets of the orbits $\mathcal{O}(\xi)$, when the latter are given the input topology.

Indeed, assume given any $\zeta \in O_0(\xi)$. In particular, $\zeta \in \mathcal{O}(\xi)$, so by part (2) of Theorem 2 there are positive numbers

$$s_1, \dots, s_l,$$

a sequence $b = (a_1, \dots, a_l)$ of $+$'s and $-$'s, and control values

$$u_1, \dots, u_l \in \mathcal{K},$$

such that

$$\zeta = \exp(a_1 s_1 X_{u_1}) \dots \exp(a_l s_k X_{u_l})(\xi)$$

and so that

$$(v_1, \dots, v_l) \mapsto \exp(a_1 s_1 X_{v_1}) \dots \exp(a_l s_k X_{v_l})(\xi) \tag{44}$$

has rank at $(u_1, \ldots, u_l)$ equal to $\dim \mathcal{M}$. Note that the image of (44) is included in $O_0(\zeta)$, since one may go first from ζ to ξ in time $-\sum s_i$ using the controls u_i. But $O_0(\zeta) = O_0(\xi)$, because of the assumption that $\zeta \in O_0(\xi)$. These maps are smooth with respect to the input topology structure on $\mathcal{O}(\zeta)$. By the implicit mapping theorem, there is then a neighborhood of ζ in $\mathcal{O}(\zeta) = \mathcal{O}(\xi)$ included in the image of (44), and hence in $O_0(\xi)$, as desired.

From now on, we shall write $\mathcal{O}_0(\xi)$ to denote the zero time orbit $O_0(\xi)$ endowed with the (input) topology from $\mathcal{O}(\xi)$.

The use of the input topology is essential here. Zero-time orbits are not open with respect to the time topology, as illustrated by the example $\dot{x} = 1$. Here $O(0) = \mathbb{R}$ and $O_0(0) = \{0\}$; the former is $\mathbb{R}_{disc}$ in the case of the input topology case but has the usual structure in the time-topology case.

In general, $\mathcal{O}(\xi)$ is not second countable in the topology being considered, as illustrated again by the above example. However, we shall establish in the next subsection that its open subset $\mathcal{O}_0(\xi)$ *is* second countable, *provided* that the control value set $\mathcal{K}$ be connected. Perhaps surprisingly, the assumption of connectedness of $\mathcal{K}$ is essential. Even a $\mathcal{K}$ with just two components may result in a non second countable zero-time orbit. As an example of this latter phenomenon, consider the discrete manifold $\mathcal{K}$ consisting of two points, say $\{0, 1\}$, and the equation on $\mathcal{M} = \mathbb{R}$

$$\dot{x} = u, \quad u \in \mathcal{K} = \{0, 1\}. \tag{45}$$

For this system, $\mathcal{O}_0(0) = \mathcal{O}(0) = \mathbb{R}_{disc}$, not second countable. In contrast, for the time-topology, connectedness of $\mathcal{K}$ is irrelevant; the same proof to be given below provides second countability for that topology when the roles of controls and times are interchanged.

5.2 Second countability of zero-time orbits

Pick an integer r and a permutation π on $\{1 \ldots r\}$ of order 2 ($\pi^2 =$ identity), and any fixed $\xi \in \mathcal{M}$. For any such π, we let

$$\mathbb{R}_\pi := \{t = (t_1, \ldots, t_r) \in \mathbb{R} \mid t_i = -t_{\pi i} \text{ for all } i\}.$$

(If $\pi i = i$ this forces t_i to vanish.) Consider also the function

$$\alpha_{r\pi}(t, \omega) := \exp(t_r X_{u_r}) \ldots \exp(t_1 X_{u_1})(\xi) \tag{46}$$

where $t = (t_1, ..., t_r)$ and $\omega = (u_1, ..., u_r)$, thought of as defined on the open subset

$$\mathcal{E}_{r\pi} \subseteq \mathbb{R}_\pi \times \mathcal{K}^r$$

where the expression in (46) is defined. We write just $\alpha(t, \omega)$ if r, π are clear from the context.

Because of the restriction to $t \in \mathbb{R}_\pi$, the image of each map (46) is included in $\mathcal{O}_0(\xi)$. This map is continuous on ω for each fixed t, by definition of the input topology; note that α is independent of those u_i for which t_i vanishes. We shall prove that the union of the images of the maps (46) cover $\mathcal{O}_0(\xi)$, and that each such image is second countable. Because there is a countable number of possible maps like this, second countability of $\mathcal{O}_0(\xi)$ will follow. Note that the "proof" that $\mathcal{E}_{r\pi}$ is second countable and therefore the image of it under α is also second countable is fallacious, because for the input topology on orbits α *is not continuous on* t.

The first step is then to establish that every element of $\mathcal{O}_0(\xi)$ can be written as $\alpha(t,\omega)$ for suitable r, π, ω, and t. We shall call such an expression, where for each time t_i there is also a corresponding $-t_i$, a *balanced* expression. The naive approach to proving this would be as follows. Assume that

$$\zeta = \exp(aX)\exp(bY)\exp(cZ)(\xi), \tag{47}$$

where X, Y, Z are of the form X_u and $a+b+c=0$. Then, since $a=-b-c$, we may also formally write the above as

$$\zeta = \exp(-cX)\exp(-bX)\exp(bY)\exp(cZ)(\xi), \tag{48}$$

or as

$$\zeta = \exp(-bX)\exp(-cX)\exp(bY)\exp(cZ)(\xi), \tag{49}$$

both of which are balanced expressions. Unfortunately, neither (48) nor (49) may be well-defined. This difficulty is illustrated by the example

$$\dot{x} = (1-u)x^2,\ \dot{y} = u,$$

with $\mathcal{M} = \mathbb{R}^2$ and $\mathcal{K} = \mathbb{R}$. Now take $a=-0.5, b=1.5, c=-1$ and for X, Y, Z the vector fields corresponding respectively to $u = 0,1,1$. Let $\xi := (1,0)'$. Then (47) is well defined, but neither (48) nor (49) are, because the solution of

$$\dot{x} = x^2,\ x(0) = 1$$

is only defined for $t \in (-1,1)$. The argument is somewhat more involved.

Lemma 5.1 If $\zeta = \exp(\tau_0 X_0)\exp(\tau_1 X_1)\dots\exp(\tau_k X_k)(\xi)$, $\sum_{h=0}^{k}\tau_h = 0$, then ζ can also be obtained from a balanced expression.

Proof. We assume without loss that $\tau_0 > 0$, otherwise the argument is the same interchanging signs. We show that

$$\zeta = \exp((\sum_{i=1}^{l}\lambda_i)X_0)\exp(\tau_1 X_1)\dots\exp(\tau_k X_k)(\xi), \tag{50}$$

where $l = km$ for some positive integer m, and where the elements of the sequence

$$-\lambda_1, \dots, -\lambda_l$$

are a permutation of

$$\underbrace{\frac{\tau_1}{m},\dots,\frac{\tau_1}{m}}_{m \text{ times}},\underbrace{\frac{\tau_2}{m},\dots,\frac{\tau_2}{m}}_{m \text{ times}},\dots,\underbrace{\frac{\tau_k}{m},\dots,\frac{\tau_k}{m}}_{m \text{ times}}, \tag{51}$$

such that

$$0 \le \sum_{i=1}^{j}\lambda_i \le \tau_0 \tag{52}$$

for each $j = 1, \ldots, l$. Because each $\exp(\tau_h X_h)$ can be rewritten as the composition of m copies of $\exp((t_h/m)X_h)$, the expression (50) is balanced; by (52) it is well-defined. It equals ζ because necessarily $\sum_{i=1}^{l} \lambda_i = \tau_0$.

Let m be any positive integer such that

$$|\tau_h/m| < \tau_0/2 \tag{53}$$

for all $h = 1, \ldots, k$. We now construct the sequence of the λ_i's as a permutation of the negatives of the elements (51). Since $-\sum \tau_h = \tau_0 > 0$, there must be some h so that $-\tau_h > 0$; let λ_1 be this element $-\tau_h/m$. We keep defining $\lambda_2, \ldots, \lambda_j$ each equal to some $-\tau_h/m$, for negative elements of the sequence (51), as long as the constraint (52) is satisfied and there are elements left in the sequence. Because of (53), at the end of this process the sum in (52) will be at least $\tau_0/2$. If we didn't exhaust the sequence (51) (in which case we are done), we stopped because the sum would become larger than τ_0 when adding some element $-\tau_h/m$, which means (because the sum of all the elements in the sequence is exactly τ_0) that there is some $\tau_h/m > 0$ in the sequence. We now start adding the negatives of such elements to the sequence of the λ_i's, again as long as the constraint (52) is satisfied. Since the sum of the λ_i up to this point is at least $\tau_0/2$, this process can be done for at least one step. When we stop, either all the sequence has been reordered or the sum is less than $\tau_0/2$. Now we repeat the ascending process. It is clear that this algorithm stops after a finite number of iterations, and results in the desired reordering. ∎

Theorem 6 *If $\mathcal{K}$ is connected, then for every ξ the input-topology zero-time orbit $\mathcal{O}_0(\xi)$ is second countable.*

Proof. We will prove that for each fixed permutation π as above, and for each (t^0, ω^0) in $\mathcal{E}_{r\pi}$ *there exists a neighborhood N of (t^0, ω^0) in $\mathcal{E}_{r\pi}$ such that $\alpha_{r\pi} N \subseteq \mathcal{O}_0(\xi)$ is connected.* Since $\mathcal{O}_0(\xi)$ is a countable union of sets of the form

$$\alpha_{r\pi}(\mathcal{E}_{r\pi}),$$

it will then be enough to show that each of these latter sets intersects at most countably many components of $\mathcal{O}_0(\xi)$. We cover $\mathcal{E}_{r\pi}$ by open sets N each of which maps into a connected set. Since $\mathcal{E}_{r\pi}$ is an open subset of the second countable manifold $\mathbb{R}_\pi \times \mathcal{K}^r$ (recall that $\mathcal{K}$ is assumed to be connected, and hence second countable), it is itself second countable. It follows that there is a countable subcover by these sets N (Lindeloff property), and the theorem will follow. Thus the construction of N is the critical part of the proof.

This is done as follows. First we introduce some auxiliary mappings. For each pair of integers $1 \le i < j \le r$, we let $\varphi_{ij}(t, \tau)$ be defined as

$$(t_1, 0, t_2, 0, \ldots, 0, t_i, \tau, t_{i+1}, 0, \ldots, 0, t_j, -\tau, t_{j+1}, 0, \ldots, 0, t_r)$$

and let $\chi_{ij}(\omega, u, v)$ be

$$(u_1, u_1, u_2, u_2, \ldots, u_{i-1}, u_i, u, u_{i+1}, u_{i+1}, \ldots, u_{j-1}, u_j, v, u_{j+1}, u_{j+1}, \ldots, u_r, u_r).$$

We also let π_{ij} be the permutation of $\{1, \ldots, 2r\}$ with

$$\pi_{ij}(2i) = 2j,\ \pi_{ij}(2j) = 2i, \text{ and } \pi_{ij}(2l-1) = 2\pi l - 1 \text{ for all } l.$$

and $\pi_{ij}k = k$ for all other k. We denote by Z the set consisting of all

$$(t, \tau, \omega, u, v) \in \mathbb{R}_\pi \times \mathbb{R} \times \mathcal{K}^r \times \mathcal{K} \times \mathcal{K}$$

for which

$$(\varphi_{ij}(t,\tau), \chi(\omega, u, v)) \in \mathcal{E}_{2r,\pi_{ij}} \text{ for all } i < j.$$

The set Z is open, by continuity on controls and time of solutions of differential equations. For the given ω^0, let C be any compact subset of $\mathcal{K}$ such that $\operatorname{int} C$ is connected and all components u_0^i of ω^0 are in $\operatorname{int} C$. The existence of such a set C follows from the assumption that $\mathcal{K}$ is connected. (This is the only place where the assumption is used.) Let

$$K := \{(t^0, 0, \omega^0, u, v) | u, v \in C\}.$$

This set is compact and it is included in Z because

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

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 that

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

where B is an open set containing C, and for each i, A_i is a connected subset of $\operatorname{int} C$ which contains the corresponding u_0^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 $\mathbb{R}_\pi$.) Finally, we let

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

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 ρ with $\rho(0) = u_i, \rho(1) = u_i'$, and $\rho(\lambda)$ in A_i for all λ. Composing with α (as a function of u_i) we get the desired path in $\mathcal{O}_0(\xi)$.

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 as straightforward as above. 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

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

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 $\tau \in (-\varepsilon, \varepsilon)$, and note that

$$\alpha(t_1 \ldots t_r, \omega) = \alpha_{2r,\pi_{ij}}(\varphi_{ij}(t, 0), \chi(\omega, u_i, u_j)) \tag{54}$$

and

$$\alpha(s_1 \ldots s_r, \omega) = \alpha_{2r,\pi_{ij}}(\varphi_{ij}(t, \tau), \chi(\omega, u_i, u_j)). \tag{55}$$

We now claim that $\alpha(s_1 \ldots s_r, \omega)$ is in the same component as

$$\alpha_{2r,\pi_{kj}}(\varphi_{kj}(t,\tau), \chi_{kj}(\omega, u_k, u_j)) \tag{56}$$

if $i \leq k < j-1$ 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 (55). Assume now that the claim has been proved for k. Since ω is in $A_1 \times \ldots \times A_r$, both u_k and u_{k+1} are in $\operatorname{int} C$. Thus there is a path ρ connecting u_k and u_{k+1}, with the image of ρ contained in $\operatorname{int} C$, and hence in B. Consider the path

$$\rho'(\lambda) := \alpha_{2r,\pi_{kj}}(\varphi_{kj}(t,\tau), \chi_{kj}(\omega, \rho(\lambda), u_j)). \tag{57}$$

This is continuous into $\mathcal{O}_0(\xi)$ with the input topology. It is well defined because of the choice of the neighborhood V, and it connects the element in equation (56) with the the corresponding element having u_{k+1} instead of u_k. If $k+1 < j$, this equals

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

(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 that N is as desired, and therefore the proof of the theorem. ∎

5.3 A sampling result

Assume given a continuous time system (11) smooth in controls, and let Σ be its associated input-topology action (definition 2.15). For each $\delta > 0$, the *associated δ-sampled action* is the one for the discrete time system (17); we denote this action, which has $\mathcal{U} = \mathcal{K}$ and the same state space $\mathcal{M}$, by Σ_δ. For an introduction to the topic of sampling and its relevance in digital control, the reader is referred for instance to [17].

Let $\mathcal{D}_\Sigma^\delta$ be the distribution associated to the sampled action Σ_δ. It follows from the definition of this sampled action and from the explicit form (25) for the generating vector fields that

$$\mathcal{D}_\Sigma^\delta \subseteq \mathcal{D}_\Sigma. \tag{58}$$

Furthermore, we claim that in fact this is locally an equality for δ small enough. Indeed, pick any $\xi \in \mathcal{M}$. There is a finite set of vector fields as in (25) whose values at ξ are a basis of $\mathcal{D}_\Sigma(\xi)$. Seen as functions of t, we may arrange these vector fields (evaluated at ξ) into a real matrix

$$A(t_1, \ldots, t_k). \tag{59}$$

We view this as a matrix of functions on $\mathbb{R}^k$, where the t_i are all the times appearing in "ω" and "u" for each generator, and for a particular value $(\tau_1, \ldots, \tau_k)$ we have the above vector fields. The columns of A are in $\mathcal{D}_\Sigma(\xi)$ for each sequence of t_i's where well-defined, and they form a basis whenever $|t_i - \tau_i| < \Delta$ for all i, for some $\Delta > 0$. For each $0 < \delta < \Delta$ and for each i, then, we may pick an integer s_i with $|s_i - \tau_i/\delta| < 1$. It follows that $A(s_1\delta, \ldots, s_k\delta)$ is a basis of $\mathcal{D}_\Sigma(\xi)$ for such δ and s_i's. Since the columns of this matrix are evaluations at ξ of generators of Σ_δ, we conclude as follows:

Lemma 5.2 For each $\xi \in \mathcal{M}$ and each small enough $\delta > 0$, $\mathcal{D}_\Sigma^\delta(\xi) = \mathcal{D}_\Sigma(\xi)$. ∎

Definition 5.3 The system (11) satisfies the *strong* or *fixed-time accessibility property* from $\xi \in \mathcal{M}$ iff $\xi \in \operatorname{int} O_0(\xi)$. ∎

For analytic systems (13) affine in controls, it is a "classical" fact ([20]) that the strong accessibility property is equivalent to the possibility of reaching in fixed (positive) time an open subset of the state space, and is also equivalent to the rank condition

$$\operatorname{rank} \mathcal{L}_0(\xi) = \dim \mathcal{M}$$

where $\mathcal{L}_0$ is the strong accessibility Lie algebra defined as the smallest subspace of vector fields on $\mathcal{M}$ which contains $g_1, \dots, g_m$ and is closed under Lie brackets by f as well as all the g_i's.

Definition 5.4 The system (11) satisfies the *sampled accessibility property* from $\xi \in \mathcal{M}$ iff $\xi \in \operatorname{int} O_\delta(\xi)$ for some $\delta > 0$, where $O_\delta(\xi)$ is the orbit of ξ under the action Σ_δ. ∎

We may now prove the main result of this section.

Theorem 7 *If $\mathcal{K}$ is connected, sampled accessibility is equivalent to strong accessibility.*

Proof. Sampled accessibility at ξ is equivalent to the full rank of $\mathcal{D}_\Sigma^\delta$ at ξ, by proposition 3.9. Strong accessibility at ξ is equivalent to the full rank of $\mathcal{D}_\Sigma$ (input-topology action) at ξ, because of proposition 3.8 and Theorem 6, plus and the fact that $O_0(\xi)$ is open in $\mathcal{O}(\xi)$ and therefore is also an integral manifold of Σ. Thus (58) gives one implication and lemma 5.2 gives the other. ∎

The Theorem is false even if $\mathcal{K}$ has just two components, as again illustrated by example (45): with any fixed δ, the sampled orbit $O_\delta(0)$ is the set of all integer multiples of δ, but the zero-time orbit is as a set all of $\mathbb{R}$.

More precise estimates for what are "good" δ can be given, based on the expansions (36); see for instance [16]. These estimates generalize known results for linear systems and, in its dual form for controllability, the Nyquist-Shannon Sampling Theorem.

6 References

1. Abraham, R. and J. E. Marsden, *Foundations of Mechanics, second edition*, Benjamin Cummings, Reading, 1978.

2. Boothby, W. M., *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, NY, 1975.

3. Brickell, F. and R.S.Clark, *Differentiable Manifolds*, Van Nostrand, New York, 1970.

4. Coddington, E.A., and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, NY, 1955.

5. Fliess, M. and D.Normand-Cyrot, "A group-theoretic approach to discrete-time nonlinear controllability," *Proc.IEEE Conf.Dec.Control*, 1981.

6. Goodman, R., "Lifting vector fields to nilpotent Lie groups," *J.Math. Pures et Appl.* **57**(1978): 77-86.

7. Hermann,R., "The differential geometry of foliations, Part II," *J.Math and Mechanics* **11**(1962): 302-316.

8. Jakubczyk, B., "Invertible realizations of nonlinear discrete time systems," *Proc.Princeton Conf.Inf.Sc.and Syts.* (1980):235-239.

9. Jakubczyk, B., and D.Normand-Cyrot, "Orbites de pseudo-groupes de diffeomorphismes et commandabilité des systemes non linearires en temps discret," *C.R.Acad. Sc. Paris*, **298-I**(1984): 257-260.

10. Jakubczyk, B. and E. D. Sontag, "Controllability of nonlinear discrete time systems: a Lie-algebraic approach," *SIAM J Control and Optimization*, submitted.

11. 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.

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

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

14. Stefan, P., "Attainable sets are manifolds," preprint, Univ. of Wales, 1973.

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

16. Sontag, E.D., "A Chow property for sampled bilinear systems," in *Proceedings of the 8th International Symposium on Mathematical Theory of Networks and Systems*, (C.I. Byrnes, C.F. Martin, and R. Saeks, eds.,) North Holland, Amsterdam, 1988.

17. Sontag,E.D., "Orbit theorems and sampling," in *Algebraic And Geometric Methods in Nonlinear Control Theory*, M.Fliess and M.Hazewinkel (Eds.), Reidel, Dordrecht, 1986, pp. 441-486.

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

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

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

7 Continuous time systems

Form now on we concentrate exclusively on applications to continuous time systems.

7.1 The accessibility rank condition

We say that *the accessibility rank condition (ARC) holds at* ξ if $\mathcal{L}_\Sigma$ has full rank at ξ; if this happens at *all* ξ, we just say that the ARC holds. The *controllability Lie algebra* of (11) is the Lie algebra

$$\mathcal{L} := \{X_u, u \in \mathcal{K}\}_{LA}$$

which for systems of the special form (13) is the same as

$$\{f, g_1, \ldots, g_m\}_{LA}.$$

This is just the Lie algebra $\mathcal{L}_\Sigma$ obtained from the time-topology action, and the ARC is simply the statement that the vectors $X(\xi)$, $X \in \mathcal{L}$ must span a space of dimension $\dim \mathcal{M}$.

The sufficiency part of the following result is often called the (positive form of) *Chow's lemma.*

Theorem 8 *Assume that for the continuous time system (11) the ARC holds at* ξ*. Then, for each neighborhood* V *of* ξ*:*

- *The set of states in* V *that are accessible from* ξ *is open.*
- *The set of states that can be reached from* ξ *contains an open subset of* V*.*
- *The set of states that can be controlled to* ξ *contains an open subset of* V*.*

All these statements hold even if controls are restricted to be piecewise constant. Conversely, if the system is analytic, and if either of the sets of states accessible from, reachable from, or controllable to ξ *contains an open subset of* $\mathcal{M}$*, then the ARC must hold.*

Proof. The value at any state ξ of the Lie bracket of two vector fields depends continuously on ξ. Thus the ARC must hold in an open neighborhood $W \subseteq V$ of ξ. We now consider the system (11) restricted to the manifold W as its state space. For this system, the ARC holds everywhere, so we may apply Theorem 3, using the time-topology action associated to the continuous time system. Note that controls are piecewise constant and take values in the interior of $\mathcal{K}$. For the third statement, note that ζ can be controlled to ξ relative to the system (11) if and only if ξ can be controlled to ζ for the time-reversed system

$$\dot{x}(t) = -P(x(t), u(t)), \quad t \in \mathbb{R}. \tag{60}$$

To apply Theorem 3 to this reversed system, we compute its Lie algebra. This turns out to be the same as the Lie algebra of the original system, since the vector fields generating $\mathcal{L}$ are the negatives of the original ones. Thus the third assertion holds too. Each of the reachability and controllability conclusions implies accessibility, so the converse statement is a consequence of corollary 3.12 and proposition 4.12. ∎

Less than analiticity is needed for the converse to hold. The same conclusion is true if the algebra $\mathcal{L}_\Sigma$ is finite dimensional or if $\mathcal{D}_\Sigma$ is of constant rank

The one-dimensional example $\dot{x} = 1$ shows that the ARC does not imply that the reachable set from ξ must be open, or even that it must be a neighborhood of ξ, as is true for the accessible set.

If the ARC holds (at every point), every orbit must be open, so $\mathcal{M}$ is partitioned into disjoint open subsets. So we conclude as follows.

Corollary 7.1 If $\mathcal{M}$ is connected and the ARC holds, the continuous time system (11) is transitive. Conversely, transitivity implies the ARC provided that the system is analytic. ∎

For linear systems, transitivity is equivalent to (complete) controllability. This is because these are analytic systems and the reachable set from the origin, being a subspace, can only contain an open subset if it equals the entire space.

Exercise 7.2 By computing the controllability Lie algebra for linear continuous time systems, relate the ARC to the usual controllability rank condition. ∎

Exercise 7.3 Given a time-varying continuous time linear system $\dot{x} = A(t)x + B(t)u$, with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^n$, for which all the entries of A and B are smooth on t, we may consider the associated nonlinear system

$$\dot{x} = A(\tau)x + B(\tau)u \tag{61}$$

$$\dot{\tau} = 1 \tag{62}$$

where τ is a new state variable. The state space is now $\mathbb{R}^{n+1}$. *Prove* that if the ARC holds for (61) then for each $\xi \in \mathbb{R}^n$ there is a $T > 0$ such that the set of states reached at time T if starting from ξ at time 0 is all of $\mathbb{R}^n$. Relate the ARC to a known controllability condition for time-varying linear systems. ∎

A continuous time system is sometimes said to be *symmetric* if for each $u \in \mathcal{K}$ there is some $u' \in \mathcal{K}$ such that $-X_u = X_{u'}$. The typical example is that of systems (13) with $f \equiv 0$ and a symmetric $\mathcal{K}$, in which case $u' = -u$ satisfies gives symmetry. For such systems, being able to reach ζ from ξ with a constant control u is equivalent to being able to control ζ to ξ using u'. Since transitivity is equivalent to transitivity with piecewise constant controls, it follows that in this case transitivity implies reachability. So the above can be rephrased, for simplicity in the analytic case, as follows.

Corollary 7.4 If (11) is symmetric and analytic, and $\mathcal{M}$ is connected, complete controllability is equivalent to the ARC. ∎

Exercise 7.5 In the nonanalytic case, a system may be controllable but the ARC may not hold. Consider for instance the system on $\mathbb{R}^2$, with $\mathcal{K} = \mathbb{R}^2$,

$$\dot{x} = u_1 g_1 + u_2 g_2$$

where $g_1(x, y) = (1, 0)'$, $g_2(x, y) = (0, \alpha(x))'$, and α is the function with $\alpha(x) = e^{-1/x^2}$ for $x > 0$ and $\alpha(x) \equiv 0$ for negative x. Show that this is completely controllable but that the ARC does not hold. ∎

The following is a well-known example illustrating the use of the ARC. Assume that we model an automobile in the following way, as an object in the plane. The position of the center of the front axle has coordinates (x, y), its orientation is specified by the angle φ, and θ is the angle its wheels make relative to the orientation of the car.

FIGURE OF CAR INDICATING COORDINATES x, y, φ, θ

We assume that the angle θ can take values on an interval $(-\theta_0, \theta_0)$, corresponding to the maximum allowed displacement of the steering wheel, and that φ can take arbitrary values. As controls we take the steering wheel moves (u_1) and the engine speed (u_2). Using elementary trigonometry, the following (symmetric) model results:

$$\dot{z} = u_1 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + u_2 \begin{pmatrix} \cos(\varphi + \theta) \\ \sin(\varphi + \theta) \\ \sin\theta \\ 0 \end{pmatrix}, \tag{63}$$

where $z = (x, y, \varphi, \theta)'$ can be thought of as belonging to the state space

$$\mathcal{M} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times (-\theta_0, \theta_0) \subseteq \mathbb{R}^4.$$

(We could instead identify φ and $\varphi + 2\pi$ and take as state space the manifold $\mathbb{R} \times \mathbb{R} \times \mathbb{S}^1 \times (-\theta_0, \theta_0)$; the results to follow would be the same in that case.) We take the controls for instance as belonging to the set

$$\mathcal{K} = [-1, 1] \times [-1, 1].$$

A control with $u_2 \equiv 0$ corresponds to a pure steering move, while one with $u_1 \equiv 0$ models a pure driving move in which the steering wheel is fixed in one position. We let $g_1 = \textit{steer}$ be the vector field $(0, 0, 0, 1)'$ and $g_2 = \textit{drive}$ the vector field $(\cos(\varphi + \theta), \sin(\varphi + \theta), \sin\theta, 0)'$. It is intuitively clear that the system is completely controllable, but it is worth proving it mathematically as follows.

Exercise 7.6 Apply corollary 7.4 to show that the system is controllable. Do this by computing the vector fields *wriggle*:= [steer,drive] and *slide*:= [wriggle,drive], and showing that at each point of $\mathcal{M}$, the determinant of the matrix consisting of the columns (steer, drive, wriggle, slide) is nonzero. ∎

What is perhaps less obvious is that one can reach any neighborhood of a given state without large excursions. More precisely, given any open subset V of $\mathcal{M}$, the system could be thought of as a system with state space V, and as such the same result gives that the restricted system is completely controllable. This is particularly useful if we need to get out of a tight parking space.

FIGURE OF CAR IN TIGHT PARKING SPACE

Exercise 7.7 Show that for $\varphi = \theta = 0$ and any (x, y), *wriggle* is the vector $(0, 1, 1, 0)$, a mix of sliding in the y direction and a rotation, and that *slide* is the vector $(0, 1, 0, 0)$ corresponding to sliding in the y direction. ∎

The trajectories corresponding to following the vector *wriggle* thus represents the "wriggling" motion

steer - *drive* - *reverse steer* - *reverse drive,* *repeat*

because of the following basic fact about Lie brackets. (As it is often said, this is a computation that everyone should do at least once in their life.)

Exercise 7.8 Show that for any two vector fields X, Y and any ξ,

$$\exp(-tY)\exp(-tX)\exp(tY)\exp(tX)(\xi) = \exp(t^2[X, Y](\xi) + o(t^2))(\xi) \tag{64}$$

as $t \to 0$. You will need to use the facts that (a) for any vector field Z in $\mathbb{R}^n$

$$\exp(tZ)(\zeta) = \zeta + tZ(\zeta) + \frac{t^2}{2} Z_*(\zeta)Z(\zeta) + o(t^2),$$

which follows from the definition of $\exp(tZ)$ as a solution of a differential equation (Z_* denotes the Jacobian of Z), and (b) Taylor expansions to first order for each of $X(x)$ and $Y(x)$. ∎

Note that (64) characterizes the Lie bracket of X and Y as measuring how far they are from commuting with each other. The term t^2 explains why many small-time wriggling motions are needed in order to obtain a displacement in the wriggling direction: the order of magnitude t^2 of a displacement in time t is much smaller than t. The exercise also suggests how to implement the pure sliding motion: wriggle, drive, reverse wriggle, reverse drive, repeat (many times).

7.2 Some further controllability results

The symmetric case can be easily characterized, at least in the analytic case. The study of controllability for *non*symmetric systems is the subject of much current research, and is one of the main areas of theoretical work in nonlinear control. We will not pursue this difficult topic here, except for a few very elementary remarks.

Lemma 7.9 Assume that the ARC holds, and that the set of states reachable from a given ξ is dense in $\mathcal{M}$. Then every state is reachable from ξ.

Proof. Pick any $\zeta \in \mathcal{M}$. By the second assertion in Theorem 8, applied at the point ζ and taking $V = \mathcal{M}$, there exists an open set W such that each state in W can be controlled to ζ. By the density assumption, there must be some state in W which is reachable from ξ. This gives the result. ∎

A similar argument is used in the next result. Recall that proposition 2.16 says that accessibility for a continuous time system is equivalent to accessibility using piecewise constant controls. When the ARC holds, a similar statement can be made regarding controllability.

Corollary 7.10 Consider a continuous time system (11). Let ξ, ζ be any two states in $\mathcal{M}$, and assume that ζ is in the interior of the set $\mathcal{R}(\xi)$ of states reachable from x and also that the ARC holds at ζ. Then, ξ can be controlled to ζ using piecewise constant controls. In particular, if the system is analytic and completely controllable, then it is possible to reach any ζ from any other ξ using piecewise constant controls.

Proof. By the second assertion in Theorem 8, applied at the point ζ and taking for V the interior $\operatorname{int} \mathcal{R}(\xi)$ of the set of states reachable from ξ, we know that there exists an open set $W \subseteq \mathcal{R}(\xi)$ such that each state in W can be steered to ζ using piecewise constant controls. Pick any state $\zeta' \in W$. By the discussion in remark 2.13, there is a sequence of states approaching ζ' each of which is reachable from ξ with piecewise constant controls. In particular, there is one such state $\zeta'' \in W$. Concatenating the piecewise constant control sending ξ to ζ'' with one sending ζ'' to ζ, we obtain the desired conclusion. For the last statement, we use the converse part of Theorem 8 in order to conclude that the ARC must hold at every point. ∎

When control values are unbounded, one may be able to cancel in some sense the effect of the "drift" term f and reduce to the symmetric situation. The following result illustrates this fact.

Proposition 7.11 Assume given an affine-in-controls system (13) for which $\mathcal{K} = \mathbb{R}^m$ and $\mathcal{M}$ is connected. A sufficient condition for the system to be completely controllable is that $\{g_1, \ldots, g_m\}_{LA}$ have full rank at each point.

This condition is far from being necessary. For linear systems, for example, it would require the B matrix to have rank n, and in particular, that there be more controls than the dimension of the state space. However, it is a useful condition sometimes. In its proof we require the following fact.

Lemma 7.12 Assume given an affine-in-controls system (13) for which $\mathcal{K} = \mathbb{R}^m$ and the ARC holds. Then this system is completely controllable if and only if the system

$$\dot{x} = u_0 f(x) + \sum_{i=1}^{m} u_i g_i(x), \quad u_0 \geq 0 \tag{65}$$

is completely controllable. (This is a system whose controls take values in $\mathcal{K} = [0, +\infty) \times \mathbb{R}^m \subseteq \mathbb{R}^{m+1}$.)

In the analytic case, the hypothesis that the ARC holds is redundant, since it is implied by controllability.

Proof. Any trajectory of the original system is also one for the new system, (simply let $u_0 \equiv 1$,) so only one implication needs to be proved. Assume then that (65) is controllable. We shall show that, for the system (13), the reachable set $\mathcal{R}(\xi)$ from each $\xi \in \mathcal{M}$ is dense in $\mathcal{M}$. The ARC holds for this system, because its controllability Lie algebra coincides with that for (65), since both equal $\{f, g_1, \ldots, g_m\}_{LA}$. Then the result will follow from lemma 7.9.

By corollary 7.10, (65) is controllable using piecewise constant controls. Pick any two $\xi, \xi' \in \mathcal{M}$. Thus there exists a sequence of states $\xi_1 = \xi, \ldots, \xi_k = \xi'$ such that each for each $i = 2, \ldots, k$ there are real numbers $T, u_0, \ldots, u_m$ such that

$$\xi_i = \exp(T(u_0 f + \sum_{j=1}^{m} u_j g_j))(\xi_{i-1}) \tag{66}$$

and T, u_0 are nonnegative. We prove by induction on i that each ξ_i is in the closure of $\mathcal{R}(\xi)$. Assume that this is true for ξ_{i-1}, and let $\zeta_l \to \xi_{i-1}$ as $l \to \infty$, each $\zeta_l \in \mathcal{R}(\xi)$. By continuous dependence of solutions on controls and initial states, for λ large enough

$$\eta_l := \exp(T([u_0 + \frac{1}{l}]f + \sum_{j=1}^{m} u_j g_j))(\zeta_l)$$

is defined and $\eta_l \to \xi_i$. But each η_l is reachable from ξ_{i-1}, because it equals

$$\exp((Tu_0 + \frac{T}{l})(f + \sum_{i=1}^{m} \frac{u_j}{u_0 + 1/l} g_j))(\zeta_l).$$

(Note that the fact that $\mathcal{K} = \mathbb{R}^m$ is used here, since the quotients $u_j/(u_0 + 1/l)$ may be very large even if the u_j's where restricted to be small.) Thus the induction step is completed. ∎

Proof of proposition 7.11. By the above, it is only necessary to show that (65) is controllable. For this it is enough to establish that

$$\dot{x} = \sum_{i=1}^{m} u_i g_i(x)$$

is controllable. But this follows from the Lie algebra assumption, which is the same as the ARC for this (symmetric) system, together with corollary 7.4. ∎

The gap between controllability and accessibility, at least in the situation of the above lemma, is due to the negative motions $\exp(tf), t < 0$:

Corollary 7.13 Assume the affine-in-controls system (13) has $\mathcal{K} = \mathbb{R}^m$ and satisfies the ARC. Then, it is completely controllable if and only if $\exp(-tf)(\xi) \in \mathcal{R}(\xi)$ for each $\xi \in \mathcal{M}$ and each $t > 0$ where defined.

Proof. The necessity of the condition is clear, since controllability means that $\mathcal{R}(\xi) = \mathcal{M}$ for each ξ. Conversely, assume that the the ARC holds for the system (13). Pick any pair $\xi, \zeta \in \mathcal{M}$. We need to establish that $\zeta \in \mathcal{R}(\xi)$. There is an integer k and a sequence of real numbers $t_1, \ldots, t_k$ such that

$$\xi = \exp(t_1 X_1) \ldots \exp(t_k X_k)(\zeta), \tag{67}$$

where each X_i is one of the vector fields $f, g_1, \ldots, g_m$. This is because the system has the same controllability Lie algebra as

$$\dot{x} = X(x),\ X \in \{f, g_1, \ldots, g_m\}$$

thought of as a continuous time system having a discrete control space $\mathcal{K}$ with $m+1$ elements, and this latter system is transitive due to the ARC. We will prove by induction on k, that $\exp(tX)(\xi) \in \mathcal{R}(\xi)$ for each such X, with respect to the system (65) and therefore by lemma 7.12 also for the original system. But each $\exp(tg_i)$ is a motion of system (66) (just let $u_i \equiv 0$ for each $j \neq i$,) and similarly for $\exp(tf)$ for positive f. Finally, for negative t, the case $\exp(tf)$ follows by the hypothesis. ∎

Exercise 7.14 (a) Many physical systems can be described by equations such as (13) in which the flow of f is *periodic*, that is, for each ξ there is some T such that $\exp(Tf)(\xi) = \xi$. *Prove* that if the ARC holds, $\mathcal{M}$ is connected, $\mathcal{K} = \mathbb{R}^m$, and the flow of f is periodic, then the system is completely controllable.

(b) As an application, consider the control of the angular velocity of a satellite using a pair of opposing jets. The corresponding equations can be obtained from the classical Euler equations for rigid body motion. When the satellite is symmetric about an axis, and choosing simple values for the parameters defining the moments of inertia and the axis along which the control acts, these become

$$\begin{aligned}\dot{\omega}_1 &= \omega_2\omega_3 + u \\ \dot{\omega}_2 &= u \\ \dot{\omega}_3 &= -\omega_2\omega_3 + u\end{aligned}$$

with state space $\mathcal{M} = \mathbb{R}^3$ and control set $\mathcal{K} = \mathbb{R}$. Show that this system is completely controllable. ∎

Bilinear continuous time systems are systems affine in controls for which $\mathcal{M}$ is a submanifold of $\mathbb{R}^n$ and the vector fields $f, g_1, \ldots, g_m$ are linear. That is, the equations are

$$\dot{x} = (F_0 + \sum_{i=1}^{m} u_i F_i)x$$

where the F_i are matrices such that $F_i\xi \in T_\xi\mathcal{M}$ for each $\xi \in \mathcal{M}$ under the identification $T_\xi\mathbb{R}^n = \mathbb{R}^n$. For such systems, the computation of the ARC is particularly easy, since the Lie bracket of any two linear vector fields $f(x) = Ax$ and $g(x) = Bx$ is obtained from the Lie bracket of the corresponding matrices, $[f, g](x) = (BA - AB)x$. Since the space of all $n \times n$ matrices has dimension n^2, one must check at most Lie brackets involving n^2 of the matrices

F_j. (In fact, the computations can be arranged very efficiently by recursively computing a basis for the span of the set of all brackets formed out of k elements, $k = 1, \ldots, n^2$.) In particular, one may consider the case of *systems on Lie groups*, whose state space is a (connected) Lie subgroup G of $GL(n)$ and the equations are in matrix form

$$\dot{X} = (A_0 + \sum_{i=1}^{m} u_i A_i)X,$$

with the A_i in the Lie algebra of G. A system like this can be rewritten as a bilinear system on a submanifold of $\mathbb{R}^{n^2}$. The vector spaces $\mathcal{L}(\xi)$ are for such systems all isomorphic to each other, and equal the Lie algebra of matrices generated by $\{A_0, \ldots, A_m\}$. In other words, the ARC is equivalent to the requirement that the set of matrices obtained by taking arbitrary Lie brackets of the A_i's should span the Lie algebra of G. For such systems, the results given about transitivity reduce to well-known and elementary facts in the theory of Lie groups.

Exercise 7.15 Use corollary 7.4 to show that any 3×3 orthogonal matrix with determinant 1 is the product of rotations about the x-axis and rotations about the y-axis. You may use that the space $SO(3)$ of such matrices is a connected Lie group of dimension 3 whose Lie algebra is the set of all skew-symmetric 3×3 matrices. Note that the two types of rotations correspond to solutions of the differential equation

$$\dot{X} = AX, \quad X(0) = I,$$

where A is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

respectively. ∎

An important variation on the idea of controllability is that of *fixed time control*. This concept is central in the study of many optimal control problems. Let $\mathcal{R}^T(\xi)$ denote the set of states reachable from ξ in time exactly T for the continuous time system (11). It is often of interest to know if $\mathcal{R}^T(\xi)$ has a nonempty interior, at least for some positive T. (If this interior is nonempty for a given T, it is also nonempty for $T + \varepsilon$, for each small $\varepsilon > 0$, since for each ζ in the interior there is some vector field X_u and some ε such that the local diffeomorphism $\exp(\varepsilon X)$ is defined at ζ.) The ARC, though clearly necessary in the analytic case, is not sufficient for this: again the trivial example $\dot{x} = 1$ provides a counterexample. For simplicity, we restrict attention now to systems affine in control (13). Consider the *strong accessibility Lie algebra* $\mathcal{L}_0$ associated to any such system as follows: $\mathcal{L}_0$ is the smallest subset of $\mathcal{L}$ which contains the vector fields $g_1, \ldots, g_m$ and which is closed under Lie brackets by f as well as by all the g_i's. (In other words, $\mathcal{L}_0$ is the *ideal* of $\mathcal{L}$ generated by the g_i's.)

Thus $\mathcal{L}_0$ is generated in the same manner as $\mathcal{L}$, by taking all possible brackets of $f, g_1, \ldots, g_m$, except that f itself is not included in the generating set. The only way for f to be in $\mathcal{L}_0$ is if it happens to be a linear combination of the g_i's and of any brackets of at least two of the vector fields. Note that for each ξ, either

$$\dim \mathcal{L}_0(\xi) = \dim \mathcal{L}(\xi)$$

or

$$\dim \mathcal{L}_0(\xi) = \dim \mathcal{L}(\xi) - 1.$$

Intuitively, the effect of time (through the autonomous vector field f) is left out of the generating set. At an equilibrium state ($f(\xi) = 0$), both algebras give the same vectors. The system satisfies the *strong accessibility rank condition* at ξ if $\mathcal{L}_0(\xi)$ has full dimension at ξ.

Exercise 7.16 *Prove* that, for an analytic system (13) and any state ξ, the interior of $\mathcal{R}^T(\xi)$ is nonempty for some $T > 0$ if and only if the strong ARC holds at ξ. *Hint.* Consider the system

$$\begin{aligned} \dot{x} &= f + \sum u_i g_i \\ \dot{z} &= 1 \end{aligned}$$

where z is a new variable. Use that the ARC for this system is the same as the strong ARC for the original system. ∎