GEOMETRIC NONLINEAR DYNAMICAL CONTROL

JOHN A. GEMMER[1]

ABSTRACT. This paper presents a brief introduction to the controllability of nonlinear systems through geometrical techniques. More specifically, Lie brackets are used to steer the system in directions not initially apparent from the linearization of the system. This allows us to prove an analogous result to the rank Kalman condition for controllability of nonlinear systems known as the accessibility rank condition.

1. Introduction

The goal of this paper is to present and prove the accessibility rank theorem and illustrate how it can be used to determine the controllability of nonlinear, affine systems of differential equations. The mathematics presented in this paper is geometrical and the reader who is unfamiliar with these concepts can think of n-dimensional manifolds as subsets of $\mathbb{R}^n$ and vector fields as mappings that assign tangent vectors to every point on a manifold.

The paper is organized as follows: In section 2 we begin by presenting the essential definitions from nonlinear dynamical control theory that we will need throughout the rest of the paper. In section 3 we then state the rank Kalman theorem which provides an easily verifiable criterion for affine linear systems to be controllable. We then show how the rank Kalman theorem can be extended to the linearization of a nonlinear system to provide a sufficient condition for nonlinear systems to be controllable. We conclude this section with a physical example illustrating the inadequacy of studying the linearization of nonlinear systems. In section 4 we give a brief introduction to Lie brackets and motivate how they can be used to solve the controllability problem for nonlinear systems. Finally, in section 5 we use Lie brackets to state and prove an analogous result to the Kalman rank condition for nonlinear systems known as the accessibility rank theorem. Also, by using Frobenius' theorem we describe the structure of the reachable set if the accessibility rank theorem is satisfied on a suitable integral submanifold.

2. Definitions

In this paper we will consider $m \times n$ *affine control systems* [4] of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sum_{i=1}^{m} \mathbf{g_j}(\mathbf{x}) u_j(t), \tag{2.1}$$

where $\mathbf{x} = \{x_1, \dots, x_n\}$ are local coordinates on a smooth n dimensional manifold M, $\mathbf{u}(t) = (u_1(t), \dots, u_m(t)) \in U \subset \mathbb{R}^m$, and $\mathbf{f}, \mathbf{g_1} \dots \mathbf{g_m}$ are smooth vectorfields on M. $\mathbf{f}$ is commonly called the *drift vector field* [4] and the vector fields $\mathbf{g_i}$ are called the *input vectorfields* [4]. The function $\mathbf{u}(t)$ is called the *control* or the *input*

Date: May 3rd, 2007.

function [4]. Additionally, we will assume that the *input space* U [4] is such that the the family of vectorfields

$$\mathcal{F} = \left\{ \mathbf{f} + \sum_i^m \mathbf{g_i} u_i : \mathbf{u} \in U \right\} \tag{2.2}$$

contains the vector fields $\mathbf{f}$ and $\mathbf{f} + \mathbf{g_i}$ for $i \in \{1, \dots, m\}$. Furthermore, we will assume that $\mathbf{u} \in \mathcal{U}$, where $\mathcal{U}$ is the set of all piecewise constant functions which are continuous from the right. We call $\mathcal{U}$ the *space of admissible controls* [4] and if $\mathbf{u} \in \mathcal{U}$ we call $\mathbf{u}$ an *admissible control* [4].

Definition 2.1. *The system 2.1 is controllable if for any two points $\mathbf{x}_0$ and $\mathbf{x}_1$ on M there exists a finite T and an admissible control $\mathbf{u} : [0, T] \to U$ such that for $\mathbf{x}$ satisfying $\mathbf{x}(0) = \mathbf{x}_0$ we have that $\mathbf{x}(T) = \mathbf{x_1}$.*

3. Linearization and Controllability

Before we analyze the controllability of (2.1) lets investigate the controllability of the affine linear system

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u},\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{u} \in \mathbb{R}^m, \tag{3.1}$$

where A is an $n \times n$ matrix and B is an $n \times m$ matrix. It is easy to verify that the solution to (3.1) satisfying $\mathbf{x}(0) = \mathbf{x}_0$ is given by

$$\mathbf{x}(t) = e^{tA}\mathbf{x}_0 + e^{tA} \int_0^t e^{-sA} B\mathbf{u}(s)\, ds.$$

Now, define the set $\mathcal{R}$ by

$$\mathcal{R} = \{\mathbf{x} \in M : \exists \mathbf{u} \in \mathcal{U} \text{ and } 0 \le t < \infty \text{ such that } \mathbf{x}(t) = 0\}.$$

So, if $\mathbf{x}_0 \in \mathcal{R}$ we have that

$$\mathbf{x}_0 = -\int_0^t e^{-sA} B\mathbf{u}(s)\, ds = -\int_0^t \sum_{k=0}^{\infty} \frac{(-s)^k A^k}{k!} B\mathbf{u}(s)\, ds.$$

From this result one can prove the following theorem concerning the controllability of (3.1). For the details of the proof see [1].

Theorem 3.1. ***Kalman rank condition*** *(3.1) is controllable if and only if the matrix $G = [B|AB|A^2B|\cdots|A^{n-1}B]$ satisfies rank$(G) = n$.*

One approach to studying the controllability of (2.1) is to investigate its *linearization* about a point $\mathbf{x}_0$, where $\mathbf{f}(\mathbf{x}_0) = 0$ and use the Kalman rank condition.

Proposition 3.2. *Consider the system (2.1), and let $\mathbf{x}_0 \in M$ satisfy $\mathbf{f}(\mathbf{x}_0) = 0$. Let U contain a neighborhood V of $\mathbf{u} = 0$ and suppose that the linearization of (2.1) about $\mathbf{x}_0$ given by*

$$\dot{\mathbf{z}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_0)\mathbf{z} + \sum_{j=1}^m \mathbf{g}_j(\mathbf{x}_0) v_j, \quad \mathbf{z} \in \mathbb{R}^n,\ \mathbf{v} \in \mathbb{R}^m$$

is controllable. Then, for any $T > 0$ the set of points that can be reached from $\mathbf{x}_0$ contains a neighborhood of $\mathbf{x}_0$.

Proof.
Since, the linearization is completely controllable there exists admissible controls $\mathbf{v}^1(t), \ldots, \mathbf{v}^n(t)$ defined on $[0, T]$ that steer the origin $\mathbf{z} = 0$ in time T to linearly independent vectors $\mathbf{z}^1, \ldots, \mathbf{z}^n \in \mathbb{R}^n$. Now, consider the input function $\mathbf{u}(t)$ defined for $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ by

$$\mathbf{u}(t, \xi) = \xi_1 \mathbf{v}^1(t) + \ldots + \xi_n \mathbf{v}^n(t).$$

Now, $\forall \epsilon > 0$ it follows that for ξ_i small enough $\|\mathbf{u}(t, \xi)\|_2 < \epsilon$. Now, for all possible ξ satisfying this criterion apply $\mathbf{u}(t, \xi)$ to (2.1) and obtain the trajectory $\mathbf{x}(t, \xi))$ on M. Therefore $\mathbf{x}(t, \xi)$ satisfies

$$\dot{\mathbf{x}}(t, \xi) = \mathbf{f}(\mathbf{x}(t, \xi)) + \sum_{i=1}^{m} \mathbf{g_i}(\mathbf{x}(t, \xi)) u_i(t, \xi).$$

Differentiating with respect to ξ and evaluating at $\xi = 0$ yields the matrix differential equation

$$Z(t) = AZ(t) + B[\mathbf{v}^1(t)| \cdots |\mathbf{v}^n(t)], \;\; Z(0) = 0,$$

where $Z(t) = \frac{\partial \mathbf{x}}{\partial \xi}(t, \xi)|_{\xi = \mathbf{0}}$, $A = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_0)$ and $B = [\mathbf{g}_1(\mathbf{x}_0)| \cdots |\mathbf{g}_n(\mathbf{x}_0]$. It follows from the definition of the controls $\mathbf{v}^1, \ldots, \mathbf{v}^n$ that $Z(T)$ is full rank. Therefore, by the implicit function theorem applied to the map $\xi \mapsto \mathbf{x}(t, \xi)$ it follows that the set of points reachable from $\mathbf{x}_0$ contains a neighborhood of $\mathbf{x}_0$. $\square$

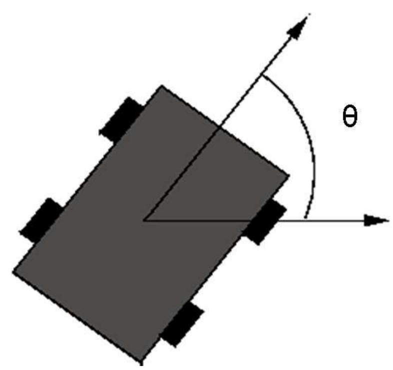

FIGURE 1

Although the above proposition gives us a simple condition to check for whether or not (2.1) is locally controllable, it is not a necessary criterion for the controllability of a system. For example, consider the following simple model of a car. Let $(x_1, x_2) \in \mathbb{R}^2$ be the position of the center of the car and $\theta \in S^1$ be the angle the front of the car makes with the x_1 axis (see figure 1). The motion of the car can be modeled by the following system:

$$\begin{cases} \dot{x_1} = u_1(t) \cos\theta, \\ \dot{x_2} = u_1(t) \sin\theta, \\ \dot{\theta} = u_2(t). \end{cases}$$

The functions $u_1(t)$ and $u_2(t)$ correspond to how the driver controls the speed of the car with the gas pedal and the orientation of the car with the steering wheel.

Clearly, the linearization about the origin given by

$$\begin{cases} \dot{x_1} = v_1(t), \\ \dot{x_2} = 0, \\ \dot{\theta} = v_2(t), \end{cases}$$

is not controllable. But, our everyday experience tells us that this model for a car should be controllable. So, perhaps there is an indirect way we can steer the car so that the system is controllable.

4. Lie Brackets

Recall that a *smooth derivation* δ on M is a mapping $\delta : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ such that $\forall f, g \in \mathcal{C}^\infty, \lambda \in \mathbb{R}$ the following properties hold:

(1) $\delta(f + \lambda g) = \delta(f) + \lambda\delta(g)$,
(2) $\delta(fg) = f\delta g + g\delta(f)$,

(see [6]). It follows that there is a natural correspondence between vector fields and smooth derivations (see [5]). Namely, if we are given a smooth derivation δ, then the vector field corresponding to δ is given in local coordinates by

$$\mathbf{v}_\delta = (\delta(x_1), \ldots, \delta(x_m)), \tag{4.1}$$

while if we are given a vector field $\mathbf{v}$ the corresponding smooth derivation is given in local coordinates by

$$\delta_{\mathbf{v}}(\mathbf{f}) = \sum_{i=1}^{m} v_i \frac{\partial \mathbf{f}}{\partial x_i}. \tag{4.2}$$

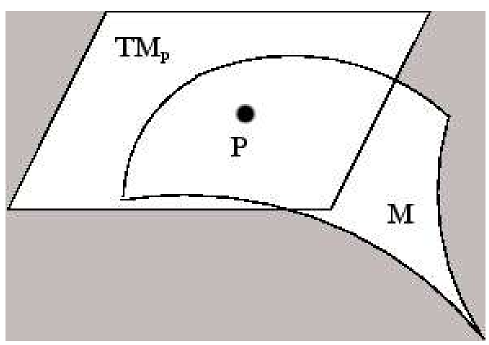

Figure 2. Illustration of the tangent plane TM_p corresponding to a point p in M.

Recall that the *tangent space* $TM_{\mathbf{x}}$ to M at $\mathbf{x}$ is the linear space of smooth derivations $\delta(\mathbf{x}) : \mathcal{C}^\infty \to \mathbb{R}$ (see [6]). Equivalently, using the correspondence between smooth derivations and vector fields the tangent space can also be thought of as the collection of all tangent vectors to all possible curves passing through $\mathbf{x} \in M$ (see figure 2). Furthermore, given that derivations act as first order differential operators we will denote the basis vectors of $TM_{\mathbf{x}}$, which clearly has the same dimension of M, by $\frac{\partial}{\partial x_i}$. So, if $\delta \in TM_{\mathbf{x}}$, then $\exists \alpha_1 + \ldots + \alpha_n \in \mathbb{R}$ such that

$$\delta = \alpha_1 \frac{\partial}{\partial x_1}, \ldots, \alpha_n \frac{\partial}{\partial x_n}.$$

Now, let δ_1 and δ_2 be two smooth derivations and consider their commutator, i.e. the map $f \mapsto \delta_1(\delta_2(f)) - \delta_2(\delta_1(f))$. It follows by direct calculation that this map is a smooth derivation as well. Therefore, for two vector fields $\mathbf{u}$ and $\mathbf{v}$ on M we have that $\delta_{[\mathbf{u},\mathbf{v}]} = \delta_{\mathbf{u}}(\delta_{\mathbf{v}}(f)) - \delta_{\mathbf{v}}(\delta_{\mathbf{u}}(f))$ will give us another vector field denoted by $[\mathbf{u}, \mathbf{v}]$. This operation is called the *Lie Bracket* [6] of $\mathbf{u}$ and $\mathbf{v}$.

Furthermore, using the correspondence between vector fields and derivations, we have that since

$$\delta_{[\mathbf{u},\mathbf{v}]} = \delta_{\mathbf{u}}(\delta_{\mathbf{v}}(f)) - \delta_{\mathbf{v}}(\delta_{\mathbf{u}}(f)), \tag{4.3}$$

it follows by (4.2) that

$$\begin{aligned} \delta_{[\mathbf{u},\mathbf{v}]} &= \delta_{\mathbf{u}} \left(\sum_{i=1}^{m} v_i \frac{\partial f}{\partial x_i} \right) - \delta_{\mathbf{v}} \left(\sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i} \right) \\ &= \sum_{j=1}^{m} u_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} v_i \frac{\partial f}{\partial x_i} \right) - \sum_{j=1}^{m} v_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i} \right). \end{aligned}$$

Consequently, by (4.1), the k-th component of the vector field corresponding to the commutator is given by

$$\begin{aligned} [u, v]_k &= \sum_{j=1}^{m} u_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} v_i \frac{\partial x_k}{\partial x_i} \right) - \sum_{j=1}^{m} v_j \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{m} u_i \frac{\partial x_k}{\partial x_i} \right) \\ &= \sum_{j=1}^{m} u_j \frac{\partial v_k}{\partial x_j} - \sum_{j=1}^{m} v_j \frac{\partial u_k}{\partial x_j}. \end{aligned}$$

Therefore,

$$[\mathbf{u}, \mathbf{v}] = J(\mathbf{v})\mathbf{u} - J(\mathbf{u})\mathbf{v}, \tag{4.4}$$

where $J(\mathbf{w})$ is the Jacobian matrix corresponding to a vector field $\mathbf{w}$.

Let M and N be manifolds and let $F : M \to N$ be a smooth map between them. The *tangent map* is the map is the map defined by $F_* : TM_{\mathbf{x}} = TN_{F(\mathbf{x})}$ (see [5]). That is, if we are given a smooth curve $\gamma(t)$ on M such that $\gamma(0) = \mathbf{x}$ then $F_*(\gamma(t))$ maps the tangent vector to $\gamma(t)$ at $t = 0$ to the tangent vector of the curve $F(\gamma(t))$ at $t = 0$. Furthermore, if we think of elements in the tangent plane as smooth derivations we have that for $\delta \in TM_{\mathbf{x}}$ and $f \in \mathcal{C}^\infty$

$$F_*(\delta(f)) = (\delta(f \circ F)) \circ F^{-1}, \tag{4.5}$$

(see [4]). If two vectorfields, $\mathbf{v}$ on M and $\mathbf{u}$ on N, satisfy $F_*\mathbf{v} = \mathbf{u}$ we say that $\mathbf{v}$ and $\mathbf{u}$ are *F-related* [4].

Proposition 4.1. *Let $F : M \to N$ and suppose that $F_*\mathbf{v}_i = \mathbf{u}_i$, for $i = 1, 2$ for vector fields $\mathbf{v}_1, \mathbf{v}_2$ on M and $\mathbf{u}_1$ and $\mathbf{u}_2$ on N. Then, $F_*[\mathbf{v}_1, \mathbf{v}_2] = [\mathbf{u}_1, \mathbf{u}_2]$.*

Proof.
Thinking of $\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2$ as derivations we have that for $g : N \to \mathbb{R}$, $g \in \mathcal{C}^\infty$,

$$\{[\mathbf{u}_1, \mathbf{u}_2](g)\} \circ F = \{\mathbf{u}_1(\mathbf{u}_2(g))\} \circ F - \{\mathbf{u}_2(\mathbf{u}_1(g)\} \circ F.$$

Now, let $f : N \to \mathbb{R}$, such that $f \in \mathcal{C}^\infty$. By, (4.5) we have that $\mathbf{u}_i(f) \circ F = F_*(\mathbf{v}_i(f)) \circ F = \mathbf{v}_i(f \circ F)$. Therefore,

$$\begin{aligned} \{[\mathbf{u}_1, \mathbf{u}_2](g)\} \circ F &= \{\mathbf{v}_1(\mathbf{u}_2(g) \circ F)\} - \{\mathbf{v}_2(\mathbf{u}_1(g) \circ F)\} \\ &= \{\mathbf{v}_1(\mathbf{v}_2(g \circ F)\} - \{\mathbf{v}_2(\mathbf{v}_1(g \circ F)\} \\ &= [\mathbf{v}_1, \mathbf{v}_2](g \circ F). \end{aligned}$$

Therefore, $F_*[\mathbf{v}_1, \mathbf{v}_2] = [\mathbf{u}_1, \mathbf{u}_2]$. □

We will also need the following important proposition but the proof is long and not particularly important for our purposes.

Proposition 4.2. *Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be linearly independent vectorfields in a neighborhood of $\mathbf{p}$ satisfying $[\mathbf{v}_i, \mathbf{v}_j] = 0$ for $i, j \in \{1, \dots, k\}$. Then, there is a coordinate chart U with local coordinates $x_1, \dots, x_n$ around p such that*

$$\mathbf{v}_i = \frac{\partial}{\partial x_i}$$

Recall, that the *local flow* [2] of a vector field $\mathbf{v}$ on M denoted by $\mathbf{v}^t(\mathbf{x}_0)$ is the solution to the differential equation $\frac{d\phi(t)}{dt} = \mathbf{v}$ for time t and with initial condition $\phi(0) = \mathbf{x}_0$. Therefore, it can be shown that the Lie Bracket of the vector fields $\mathbf{u}$ and $\mathbf{v}$ tells us that $\mathbf{v}^t(\mathbf{u}^s(\mathbf{x}_0)) = \mathbf{u}^s(\mathbf{v}^t(\mathbf{x}_0))$ if and only if $[\mathbf{u}, \mathbf{v}] = 0$. In other words, the Lie Bracket corresponds to how the local flows generated by $\mathbf{u}$ and $\mathbf{v}$ commute. Furthermore, this failure to commmute is entirely a non-linear phenomenon for if we are given two constant vector fields their Lie Bracket is identically 0. The next proposition show us how this failure to commute gives us, at least approximately, the indirect method we need to steer the car.

Proposition 4.3. *Let $\mathbf{u}, \mathbf{v}$ be vector fields on M. For each $\mathbf{x} \in M$ define the curve $\Psi(\mathbf{x}, t)$ on M by*

$$\Psi(\mathbf{x}, t) = \mathbf{u}^{-t}(\mathbf{v}^{-t}(\mathbf{u}^t(\mathbf{v}^t(\mathbf{x})))). \tag{4.6}$$

The Lie Bracket $[\mathbf{v}, \mathbf{u}](\mathbf{x})$ is the tangent vector to this curve at the endpoint $\Psi(\mathbf{x}, 0)$, i.e,

$$[\mathbf{v}, \mathbf{u}](\mathbf{x}) = \frac{d}{dt}\Psi(\mathbf{x}, t)|_{t=0}. \tag{4.7}$$

Proof.
First, to simplify notation all functions will be evaluated at $\mathbf{x}$ but the evaluation will not be explicitly written down, i.e, $\mathbf{v}(\mathbf{x}) = \mathbf{v}$ and $\mathbf{v}^t = \mathbf{v}^t(\mathbf{x})$. Also, to reduce the length of the proof many intermediate computations have been removed. To prove the result we will use repeated applications of Taylor's Theorem. Computing gives us

$$\begin{gathered} \mathbf{v}^t = \mathbf{x} + t\mathbf{v} + \frac{1}{2}t^2 J(\mathbf{v})\mathbf{v} + \mathcal{O}(t^3) \\ \Rightarrow \mathbf{u}^t(\mathbf{v}^t) = \mathbf{x} + t(\mathbf{v} + \mathbf{u}) + t^2\left(\frac{1}{2}J(\mathbf{v})\mathbf{v} + J(\mathbf{u})\mathbf{v} + \frac{1}{2}J(\mathbf{u})\mathbf{u}\right) + \mathcal{O}(t^3) \\ \Rightarrow \mathbf{v}^{-t}(\mathbf{u}^t(\mathbf{v}^t)) = \mathbf{x} + t\mathbf{u} + t^2\left(J(\mathbf{u})\mathbf{v} - J(\mathbf{v})\mathbf{u} + \frac{1}{2}J(\mathbf{u})\mathbf{u}\right) + \mathcal{O}(t^3) \\ \Rightarrow \Psi(\mathbf{x}, t) = \mathbf{u}^{-t}(\mathbf{v}^{-t}(\mathbf{u}^t(\mathbf{v}^t))) = \mathbf{x} + t^2(J(\mathbf{u})\mathbf{v} - J(\mathbf{v})\mathbf{u}) + \mathcal{O}(t^3). \end{gathered}$$

Therefore,

$$\lim_{t\to 0} \frac{\Psi(\mathbf{x},t)-\Psi(\mathbf{x},0)}{t^2} = \lim_{t\to 0} \frac{\Psi(\mathbf{x},t)-\mathbf{x}}{t^2} = [\mathbf{v},\mathbf{u}](\mathbf{x})$$

□

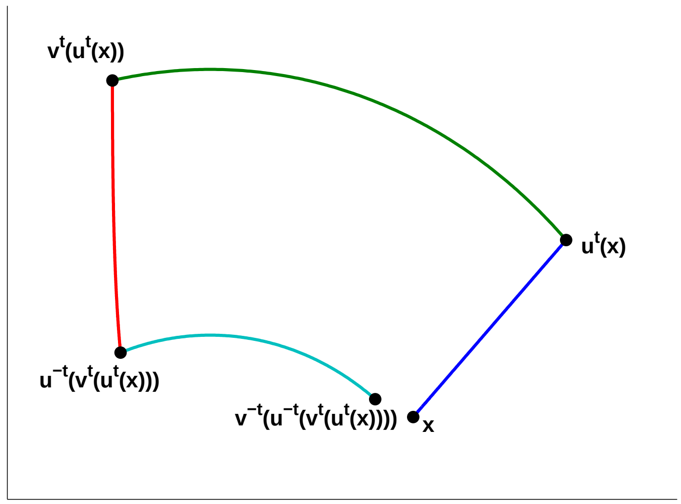

FIGURE 3. Graphical illustration of the above proof. The Lie bracket is the tangent vector to the curve $\Psi(t,\mathbf{x})$ at $t=0$ and gives us another direction to steer our system in.

By picking more complicated input function it follows by the above proposition that we can approximately steer the car in the direction of higher order brackets, i.e. for vector fields $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ terms of the form

$$[[\mathbf{v}_1,\mathbf{v}_2],\mathbf{v}_3],\ \ [\mathbf{v}_k,[\mathbf{v}_{k-1},[\cdots,[\mathbf{v}_2,\mathbf{v}_1]\cdots]]]$$

etc.

Definition 4.4. *A vector space V over $\mathbb{R}$ is said to be a Lie algebra if there exists a binary operation $V\times V\to V$ denoted by $[\,,\,]$ such that $\forall \mathbf{u},\mathbf{v},\mathbf{w}\in V$ and $\forall\alpha,\beta\in\mathbb{R}$ the following properties are satisfied*

(1) $[\alpha\mathbf{u},\beta\mathbf{v},\mathbf{w}] = \alpha[\mathbf{u},\mathbf{w}]+\beta[\mathbf{v},\mathbf{w}]$ *(bilinearity),*
(2) $[\mathbf{u},\mathbf{v}] = -[\mathbf{v},\mathbf{u}]$ *(anti-symmetry),*
(3) $[[\mathbf{u},[\mathbf{v},\mathbf{w}]]+[\mathbf{v},[\mathbf{w},\mathbf{u}]]+[\mathbf{w},[\mathbf{v},\mathbf{v}]]=0$ *(Jacobi-identity).*

Let $V^\infty(M)$ be the linear space of smooth vector fields on M. Clearly, for the Lie bracket it follows by (4.4) that the bilinearity and skew symmetry properties are satisfied for elements in $V^\infty(M)$. Also, it can be shown, with a a little bit of work, that the Jacobi identity holds as well. So, $V^\infty(M)$ with the Lie bracket is an infinite dimensional Lie algebra. A *subalgebra* of $V^\infty(M)$ is a linear subspace $U^\infty\subset V^\infty(M)$ satisfying $\forall\mathbf{u},\mathbf{v}\in U^\infty$, $[\mathbf{u},\mathbf{v}]\in U^\infty$. For our purposes, we will be interested in the following subalgebra.

Definition 4.5. *For the affine system (2.1) the accessibility algebra $\mathfrak{C}$ is the smallest subalgebra of $V^\infty(M)$ containing the vector fields $\mathbf{f}, \mathbf{g}_1, \ldots, \mathbf{g}_m$.*

Proposition 4.6. *Every element of $\mathfrak{C}$ is a linear combination of repeated Lie brackets of the from*

$$[\mathbf{v}_k, [\mathbf{v}_{k-1}, [\cdots, [\mathbf{v}_2, \mathbf{v}_1]\cdots]]] \tag{4.8}$$

where $\mathbf{v}_i$, $i \in 1, \ldots, k$ is in the set $\{\mathbf{f}, \mathbf{g}_1, \ldots, \mathbf{g}_m\}$ and $k = 0, 1, 2, \ldots$.

Proof.
Let $\mathcal{L}$ be the linear subspace of V^∞ spanned by (4.8). $\forall \mathbf{w} \in \mathcal{L}$ let $L(\mathbf{w})$ denote the number of Lie brackets contained in $\mathbf{w}$. I.e. if $\mathbf{w}$ is given by (4.8) then $L(\mathbf{w}) = k$. Now, $\forall n \in \mathbb{N}$ let $P(n)$ denote the logical statement that for an arbitrary $\mathbf{u} \in \mathcal{L}$ we have that $\forall \mathbf{v} \in \mathcal{L}$ satisfying $L(\mathbf{v}) \leq n$, $[\mathbf{v}, \mathbf{u}] \in \mathcal{L}$. $P(1)$ is trivially true. Suppose $\exists k \in \mathbb{N}$ such that $P(k)$ is true. Let $\mathbf{v} = [\mathbf{v}_{k+1}, [\mathbf{v}_k, [\cdots, [\mathbf{v}_2, \mathbf{v}_1]\cdots]]]$ and let $\mathbf{z} = [\mathbf{v}_k, [\cdots, [\mathbf{v}_2, \mathbf{v}_1]\cdots]]$. Then, by the Jacobi identity we have that

$$\begin{aligned}
[\mathbf{v}, \mathbf{u}] &= [[\mathbf{v}_{k+1}, \mathbf{z}], \mathbf{u}] \\
&= -[\mathbf{u}, [\mathbf{v}_{k+1}, \mathbf{z}]] \\
&= [\mathbf{v}_{k+1}, [\mathbf{z}, \mathbf{u}]] + [\mathbf{z}, [\mathbf{u}, \mathbf{v}_{k+1}]].
\end{aligned}$$

Since $L(\mathbf{z}) = k$ it follows by $P(k)$ that $[\mathbf{v}_{k+1}, [\mathbf{z}, \mathbf{u}]] \in \mathcal{L}$ and $[\mathbf{z}, [\mathbf{u}, \mathbf{v}_{k+1}]] \in \mathcal{L}$. By the principle of mathematical induction it follows that $\mathcal{L}$ is a subalgebra containing $\{\mathbf{f}, \mathbf{g}_1, \ldots, \mathbf{g}_m\}$ so $\mathcal{L} = \mathfrak{C}$. □

5. Local Accessibility and Controllability

Now, to answer the question of controllability of a system we need to examine the space spanned by the elements in the accessibility algebra.

Definition 5.1. *A distribution D on M is a map which assigns to each $\mathbf{x} \in M$ a linear subspace $D(\mathbf{x})$ of the tangent space TM_x. D is called a smooth distribution if around any point $\mathbf{x}$, $D(\mathbf{x})$ is spanned by a set of smooth vector fields, i.e, for $\mathbf{x} \in M$ there exists a neighborhood U of $\mathbf{x}$ and a set of smooth vectorfields $\mathbf{v}_i$, $i \in I$ with I some index set such that*

$$\forall \mathbf{y} \in U,\, D(\mathbf{y}) = span\{\mathbf{v}_i(y);\, i \in I\}.$$

The *dimension* [4] of a distribution at $\mathbf{x} \in M$ is the dimension of the subspace $D(\mathbf{x})$. A distribution D is called *constant dimensional* [4] if $\forall \mathbf{x} \in M$ the dimension of $D(\mathbf{x})$ does not depend on x. Furthermore, a distribution D is called called *involutive* [4] if for all $\mathbf{u}, \mathbf{v} \in D$, $[\mathbf{u}, \mathbf{v}] \in D$. Define the *accessibility distribution* [4] to be the involutive distribution generated by $\mathfrak{C}$:

$$C(\mathbf{x}) = span\{\mathbf{v}(\mathbf{x}) : \mathbf{v} \in \mathfrak{C}\}. \tag{5.1}$$

We say that a submanifold N of M is an *integral manifold* [4] of a distribution D on M if

$$\forall \mathbf{x} \in N,\, TN_x = D(\mathbf{x}).$$

Proposition 5.2. *Let D be a distribution on M such that through each point of M there passes an integral manifold of D. Then, D is an involution.*

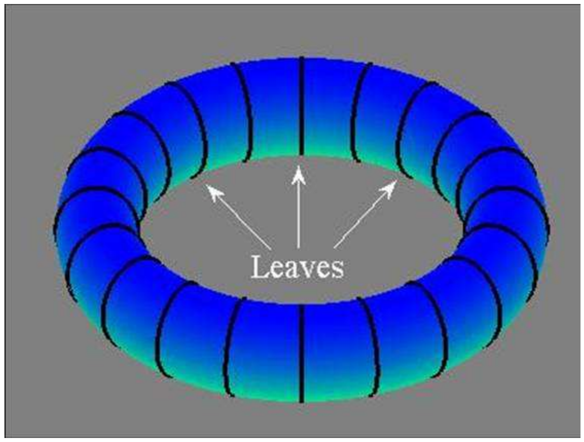

FIGURE 4. This figure illustrates one-dimensional integral submanifolds of a torus. The black curves labeled as leaves are integral submanifolds corresponding to a distribution that is spanned by a single vector field on the torus.

Proof.
Let $\mathbf{u}, \mathbf{v} \in D$ and $\mathbf{x} \in M$. Let N be an integral manifold of D with dimension $n < m$. Then, for every $\mathbf{x} \in N$,

$$\mathbf{u}(\mathbf{x}) \in D(\mathbf{x}) = TN_x, \text{ and } \mathbf{v}(\mathbf{x}) \in D(\mathbf{x}) = TN_x.$$

Since N is a submanifold of M, it follows that there exists a coordinate chart U with local coordinates $x_1, \ldots, x_m$, such that

$$U \cap N = \{\mathbf{y} \in U : x_i(\mathbf{y}) = x_i(\mathbf{x}),\, i = n+1, \ldots, m\}.$$

Therefore, the last $m-n$ components of $\mathbf{u}$ and $\mathbf{v}$ in the basis $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m}\right\}$ are precisely zero. Consequently, by (4.4) it follows that the last $m-n$ components of $[\mathbf{u}, \mathbf{v}]$ are precisely zero as well. Therefore, $[\mathbf{u}, \mathbf{v}] \in TM_x = D(\mathbf{x})$. $\square$

Now we are in a position to give our first major result concerning the controllability of (2.1). Related to the notion of controllability is the *reachability set* [4] $R^V(\mathbf{x}_0, T)$ defined by

$$\begin{aligned} R^V(\mathbf{x}_0, T) \quad = \quad & \{\mathbf{x} \in M : \exists \mathbf{u} \in \mathcal{U},\, \mathbf{u} : [0, T] \to U \text{ such that the evolution of} \\ & (2.1) \text{ for } \mathbf{x}(0) = \mathbf{x}_0 \text{ satisfies } \mathbf{x}(t) \in V,\, 0 \le t \le T \text{ and } \mathbf{x}(T) = \mathbf{x}\}. \end{aligned}$$

Let

$$R_T^V(\mathbf{x}_0) = \bigcup_{\tau \le T} R^V(\mathbf{x}_0, \tau). \tag{5.2}$$

Theorem 5.3. ***Accessibility Rank Condition***. *Consider the system (2.1). Assume that $dim(C(\mathbf{x}_0)) = n$. Then, for any open neighborhood V of $\mathbf{x}_0$ and $T > 0$ the set $R_T^V(\mathbf{x}_0)$ contains a non-empty open set of M.*

Proof.
By continuity there exists a neighborhood $W \subset V$ of $\mathbf{x}_0$ such that $dim(C(\mathbf{x}_0)) = n$ for any $\mathbf{x} \in W$. It follows that since $dim(C(\mathbf{x}_0)) \neq 0$ we can pick $\mathbf{v}_1 \in \mathcal{F}$, where

$\mathcal{F}$ is defined in (2.2), such that $\mathbf{v}_1(\mathbf{x}_0) \neq 0$. Therefore, for sufficiently small $\epsilon_1 > 0$ the set

$$N_1 = \{\mathbf{v}_1^{t_1}(\mathbf{x}_0) : 0 < t_1 < \epsilon_1\}$$

is a submanifold of dimension 1, contained in W. Now, assume we have constructed a submanifold $N_{j-1} \subset W$ of dimension $j-1 < n$ defined as

$$N_{j-1} = \{\mathbf{v}_{j-1}^{t_{j-1}} \circ \mathbf{v}_{j-2}^{t-2} \circ \cdots \mathbf{v}_1^{t_1}(\mathbf{x}_0) : 0 < t_i < \epsilon_i,\, i \in \{1, \ldots, j-1\}\},$$

where $\mathbf{v}_i$, $i \in \{1, \ldots, j-1\}$ are vectorfields in $\mathcal{F}$. Now, we can find $\mathbf{v}_j \in \mathcal{F}$ and $\mathbf{q} \in N_{j-1}$ such that $\mathbf{v}_j(\mathbf{q}) \notin TN_{j-1,\mathbf{q}}$. For if this was not possible, then $\forall \mathbf{v} \in \mathcal{F}, \mathbf{q} \in N_{j-1}$ we have have that $\mathbf{v}(\mathbf{q}) \in TN_{j-1,\mathbf{q}}$. By the proof of proposition 5.2 it follows that all linear combinations of Lie brackets of elements in $\mathcal{F}$ would also lie in $TN_{j-1,\mathbf{q}}$ which contradicts the fact that $dim(C(\mathbf{q})) = n$. It also follows that we may take $\mathbf{q}$ arbitrarily close to $\mathbf{x}_0$. Therefore, the map

$$(t_j, \ldots, t_1) \to \mathbf{v}_j^{t_j} \circ \mathbf{v}_{j-1}^{t_{j-1}} \circ \cdots \mathbf{v}_1^{t_1}(\mathbf{x}_0)$$

has rank equal to j on some set $0 < t_i < \epsilon_i$ for $i \in \{1, \ldots, j\}$. Hence, $N_j \subset W$ is a submanifold of dimension j. By the principle of mathematical induction it follows that N_n is the desired open set contained in $R_T^V(\mathbf{x}_0)$. □

Motivated by this result we say that the system (2.1) is *locally accessible* [4] from $\mathbf{x}_0$ if $R_T^V(\mathbf{x}_0)$ contains a non-empty open set of M for all neighborhoods V of $\mathbf{x}_0$ and all $T > 0$. If this is true for all $\mathbf{x}_0 \in M$ we say that the system is *locally accessible* [4]. It immediately follows from theorem 5.3 that if $dim(C(\mathbf{x})) = n$ for all $\mathbf{x} \in M$ then the system is locally accessible.

But, we cannot say that the system is controllable even if the system is locally accessible. For example, consider the system

$$\begin{cases} \dot{x}_1 = x_2^2 \\ \dot{x}_2 = u \end{cases}$$

In this case the drift vector field is given by $(x_2^2, 0)$ and the only input vector field is $(1, 0)$. Computing the lie brackets gives us

$$C(\mathbf{x}) = span\{(-2x_2, 0), (2, 0), (x_2^2, 0), (0, 1)\}.$$

Clearly, $dim(C(\mathbf{x})) = 2$ for all $\mathbf{x} \in M$ so the system is locally accessible. But, since $x_2^2 \geq 0$ the x_1 coordinate is always increasing so the system is not controllable.

Now, a natural question to ask is what can we say about the reachable set if $C(\mathbf{x})$ is constant dimensional but $dim(C(\mathbf{x})) < n$? This question is essentially answered by the celebrated Frobenius Theorem which says that a constant dimensional distribution D is involutive if and only if it is integrable. Moreover, Frobenius' theorem gives us a method for computing the integral manifold corresponding to the distribution (see [3]). Then, if we restrict the system (2.1) to the integral submanifold the system will be locally accessible.

Theorem 5.4. ***Frobenius' Theorem*** *Let D be an involutive constant dimensional distribution on M. Then, around any $\mathbf{p} \in M$ there exists a coordinate chart $(U, x_1, \ldots, x_m)$ such that*

$$D(\mathbf{q}) = span\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}\right\}. \tag{5.3}$$

Proof.
Let D be a involutive, constant dimensional distribution on M. Let (U, φ) be a coordinate chart about $\mathbf{p} \in M$ with $\varphi(\mathbf{p}) = 0$. Since U locally looks like a subset of $\mathbb{R}^m$ we may as well work in $\mathbb{R}^m$ with $\mathbf{p} = 0$. So, we can assume that there exists k such that $D(0) \subset T\mathbb{R}^m$ is spanned by

$$\left\{ \frac{\partial}{\partial r_1}, \ldots, \frac{\partial}{\partial r_k} \right\}.$$

Define the function $\pi : \mathbb{R}^m \to \mathbb{R}^k$ to be the projection map onto the first k components. It follows that $\pi_{*0} : T\mathbb{R}^m_0 \to T\mathbb{R}^k_{\pi(0)}$ when restricted to to $D(0)$ is one-to-one, onto and since it is linear a homomorphism. Therefore, π_{*0} is an isomorphism. By continuity it follows that for $\mathbf{q}$ close to 0 $\pi_{*\mathbf{q}} : T\mathbb{R}^m_{\mathbf{q}} \to T\mathbb{R}^k_{\pi_{\mathbf{q}}}$ is an isomorphism when restricted to $D(\mathbf{q})$ as well. So, for $\mathbf{q}$ sufficiently close to 0 we have vectors $\mathbf{v}_1(\mathbf{q}), \ldots, \mathbf{v}_k(\mathbf{q}) \in D(\mathbf{q})$ such that

$$\pi_{*\mathbf{q}}(\mathbf{v}_i(\mathbf{q})) = \left. \frac{\partial}{\partial r_i} \right|_{\pi(\mathbf{q})}.$$

By proposition 4.1 it follows that for $i, j \in 1 \ldots, k$,

$$\pi_{*\mathbf{q}}\left([\mathbf{v_i}, \mathbf{v_j}]\right) = \left. \left[\frac{\partial}{\partial r_i}, \frac{\partial}{\partial r_j} \right] \right|_{\pi(\mathbf{q})} = 0.$$

Now, since D is involutive we have that $[\mathbf{v}_i, \mathbf{v}_j] \in D$. Therefore, since $\pi_{*\mathbf{q}}$ is one-to-one when restricted to $D(\mathbf{q})$ and $\pi_{*\mathbf{q}}(0) = 0$ it follows that $[\mathbf{v}_i, \mathbf{v}_j] = 0$. Finally, it follows by proposition 4.2 that for all $\mathbf{q} \in U$

$$\mathbf{v_i} = \frac{\partial}{\partial x_i}.$$

□

Corollary 5.5. *Let D be an involutive distribution of constant dimension k on M. Then, for any $\mathbf{p} \in M$ there is a coordinate chart $(U, x_1, \ldots, x_m)$ with $\varphi(\mathbf{p}) = 0$ and*

$$\varphi(U) = (-\epsilon, \epsilon), \ldots, (-\epsilon, \epsilon),$$

such that for each $a_{k+1}, \ldots, a_m$, smaller in absolute value than ϵ, the submanifold

$$\{\mathbf{q} \in U : x_{k+1}(\mathbf{q}) = a_{k+1}, \ldots, x_m(\mathbf{q}) = a_m\} \tag{5.4}$$

is an integral manifold of D. Moreover, every integral manifold is of this form.

The union over all possible submanifolds defined by (5.4) is called a *foliation* [6] of the coordinate chart U. For fixed values of a_i the submanifold given in (5.4) is called a *leaf* [6] of the foliation. Therefore, Frobenius' theorem tells us that every involutive constant dimensional distribution generates a foliation of M whose leaves are integral manifolds of the distribution.

Proposition 5.6. *Suppose that for all $\mathbf{x} \in M$, $C(\mathbf{x})$ has constant dimension $k < n$. By Frobenius's theorem we can find a neighborhood W of $\mathbf{x}$ and local coordinates $x_1, \ldots, x_n$, such that the submanifold*

$$S_{\mathbf{x}} = \{\mathbf{q} \in W : x_i(\mathbf{q}) = x_i(\mathbf{x}),\ i = k+1, \ldots, n\}$$

is an integral manifold of C. Then, the system restricted to $S_{\mathbf{x}}$ is locally accessible.

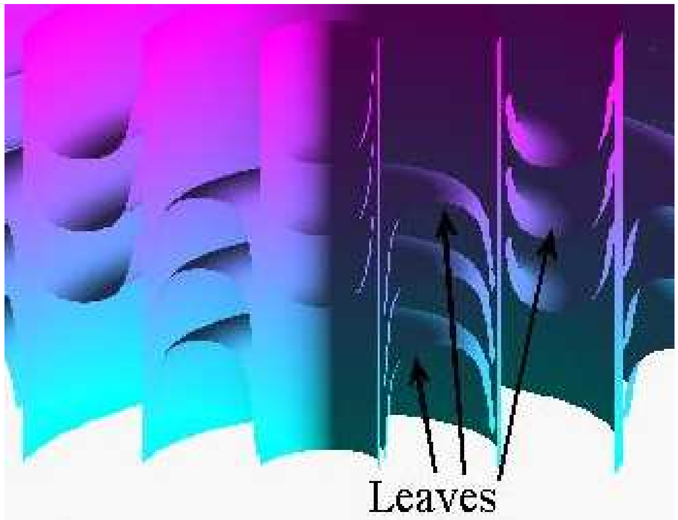

FIGURE 5. This figure illustrates a foliation of $\mathbb{R}^3$. One can think of the leaves of the foliation as slices of $\mathbb{R}^3$ that correspond to submanifolds that are locally accessible from a system starting on the submanifold. Physically, the system we are trying to control can only be steered to other points on the leaf the system starts on.

If we now return to our previous example where we illustrated why local accessibility does not imply controllability, it is clear that the problem was with the input vector fields inability to steer the system in the opposite direction of the drift vector field. So, to guarantee controllability we will need the system to remain locally accessible if $\mathbf{f}$ is removed from $\mathcal{F}$ and in addition we will need the input vector fields and their associated Lie brackets to be able to steer the system in all directions.

Theorem 5.7. ***Controllability*** *Suppose the drift vector field* $\mathbf{f}$ *satisfies* $\mathbf{f} = 0$ *or* $\mathbf{f} \in span\{g_i(\mathbf{x}),\, i \in \{1, \ldots, m\}\}$ *and suppose also that for any* $\mathbf{v} \in \mathcal{F}$ *we have that* $-\mathbf{v} \in \mathcal{F}$*. Then, if* $dim(\mathcal{C}(\mathbf{x}) = n$ *for all* $\mathbf{x} \in M$ *and* M *is connected then (2.1) is controllable.*

Proof.
Recall the proof of theorem 5.3 and consider the n-dimensional submanifold $N_n(\mathbf{x})$ which is the image of the map

$$(t_n, \ldots, t_1) \to \mathbf{v}_n^{t_n} \circ \mathbf{v}_{n-1}^{t_{n-1}} \circ \cdots \circ \mathbf{v_1}^{t_1}(\mathbf{x}),\, 0 < t_i < \epsilon_i$$

for some $\mathbf{v}_i \in \mathcal{F}$, $i \in \{1, \ldots, n\}$. Now, let $(s_1, \ldots, s_n)$ satisfy $0 < s_i < \epsilon_i$ and for $s_i \in \{1, \ldots, n\}$ fixed consider the map

$$(t_n, \ldots, t_1) \to (-\mathbf{v}_1)^{s_1} \circ (-\mathbf{v}_2)^{s_2} \circ \cdots (\mathbf{v}_n)^{s_n} \circ \mathbf{v}_n^{t_n} \circ \mathbf{v}_{n-1}^{t_{n-1}} \circ \cdots \mathbf{v}_1^{t_1}(\mathbf{x}),\, t_i < \epsilon_i.$$

Now, since $(-\mathbf{v}_i)^{s_i} = \mathbf{v}_i^{-s_i}$ it follows that the image of this map is an open set of M containing $\mathbf{x}$. Therefore, by the symmetry of $\mathcal{F}$ we have proven that for all $\mathbf{x} \in M$, $R_T^V(\mathbf{x})$ contains an open neighborhood of $\mathbf{x}$ for all neighborhoods V of $\mathbf{x}$ and $T > 0$.

Now, let

$$R(\mathbf{x}) = \bigcup_{\tau>0} R^M(\mathbf{x}, \tau).$$

and B be the boundary of $R(\mathbf{x})$. Suppose there exists $\mathbf{z} \in M$ such that $\mathbf{z} \in B$. By the above argument $R(\mathbf{z})$ contains a neighborhood of $\mathbf{z}$. Consequently, $\exists \mathbf{z}' \in R(\mathbf{z})$ such that $\mathbf{z}' \notin R(\mathbf{x})$. Consequently $\mathbf{z}'$ can be reached by first steering $\mathbf{x}$ to $\mathbf{z}$ and then steering $\mathbf{z}$ to $\mathbf{z}'$. But, this contradicts the fact that $\mathbf{z}$ is in the boundary of $R(\mathbf{x})$. Therefore, $B = \emptyset$ which proves that $R(\mathbf{x}) = M$. □

References

[1] Evans, Lawrence. An Introduction to Mathematical Optimal Control Theory. 5/3/07, ¡http://math.berkeley.edu/ evans/control.course.pdf¿

[2] Frankel, Theodore. The Geometry of Physics. New York: Cambridge University Press, 2004

[3] Lee, John. Introduction to Smooth Manifolds. New York: Springer-Verlag, 2002

[4] Nijmeijer, Henk and Van der Shaft, Arjan. Nonlinear Dynamical Control Systems. New York: Springer-Verlag, 1990

[5] Olver, Peter. Applications of Lie Groups to Differential Equations. New York: Springer-Verlag, 2000

[6] Warner, Frank. Foundations of Differentiable Manifolds and Lie Groups. Glenview, Illinois: Scott, Foresman and Company, 1971

Interdisciplinary Department of Applied Mathematics, University of Arizona, Tucson, AZ 85721

E-mail address: jgemmer@math.arizona.edu