Summer 2016

Geometric Control Theory: Nonlinear Dynamics and Applications

Geoffrey A. Zoehfeld
San Jose State University

GEOMETRIC CONTROL THEORY: NONLINEAR DYNAMICS AND APPLICATIONS

A Thesis

Presented to

The Faculty of the Department of Mathematics and Statistics

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Geoffrey A. Zoehfeld

August 2016

The Designated Thesis Committee Approves the Thesis Titled

GEOMETRIC CONTROL THEORY: NONLINEAR DYNAMICS AND APPLICATIONS

by

Geoffrey A. Zoehfeld

APPROVED FOR THE DEPARTMENT OF MATHEMATICS AND STATISTICS

SAN JOSÉ STATE UNIVERSITY

August 2016

Dr. Slobodan N. Simić	Department of Mathematics and Statistics
Dr. Matthew D. Johnston	Department of Mathematics and Statistics
Dr. Samih Obaid	Department of Mathematics and Statistics

ABSTRACT

GEOMETRIC CONTROL THEORY: NONLINEAR DYNAMICS AND APPLICATIONS

by Geoffrey A. Zoehfeld

We survey the basic theory, results, and applications of geometric control theory. A control system is a dynamical system with parameters called controls or inputs. A control trajectory is a trajectory of the control system for a particular choice of the inputs. A control system is called controllable if every two points of the underlying space can be connected by a control trajectory. Two fundamental problems of control theory include: 1) Is the control system controllable? 2) If it is controllable, how can we construct an input to obtain a particular control trajectory? We shall investigate the first problem exclusively for affine drift free systems. A control system is affine if it is of the form:
$\dot{x} = X_0(x) + u_1 X_1(x) + ... + u_k X_k(x)$ where X_0 is the drift vector field, $X_1(x), ..., X_k(x)$ are the control vector fields, and $u_1, ..., u_k$ are the inputs. An affine system is called drift-free if $X_0 = 0$. The fundamental theorem of control theory (known as Chow-Rashevsky theorem) states that an affine drift-free control system is controllable if the control vector fields together with their iterated Lie brackets span the entire tangent bundle of the underlying space. We prove this result in the simplest case when the space is 3-dimensional and $k = 2$.

DEDICATION

I dedicate this thesis to my cat Cupcake, the cutest, most adorable cat you'll ever meet. Actually, it'd be more appropriate to dedicate this to my parents since they raised us! Thanks Mom and Dad!

ACKNOWLEDGEMENTS

I would like to thank the faculty of Berkeley High School, Columbia University, and San José State University who have inspired and challenged me throughout the culmination of my academic career. In particular, I wish to express my utmost gratitude to my thesis advisor, Dr. Simić for his patience and time. I can only hope to surpass him one day, but today is not that day.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION AND PRELIMINARIES

In order to understand control theory, we shall first elucidate the underlying basic topological and geometric concepts. We shall use definitions and examples from a variety of sources such as [Boo86, Mun00, Sas99, Jur97, Sim11, Kap14].

1.1 Basic Topology

We begin with the open set definition of a topological space.

Definition 1.1.1. Let X be a set and T a family of subsets of X. Then the pair (X, T) is a *topological space* and T is a *topology* on X if the following axioms are satisfied.

(1) The empty set $\emptyset$ and X are members of T. (If only $\emptyset$ and X belong to T, this is the *trivial topology.*)

(2) The union of any members of T is still a member of T.

(3) The intersection of finitely many members of T is still a member of T.

Definition 1.1.2. If X is a topological space with topology T, then we call a subset U of X an *open set* of X if U belongs to T.

With this definition of open sets, we can view a topological space as a set X with a collection of subsets of X that are open sets (including $\emptyset$ and X), and that unions and finite intersections of open sets are open.

Example 1.1.3. Let $X = \{1, 2, 3, 4\}$ and

$$T = \{\emptyset, \{2\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}, \{1, 2, 3, 4\}\}.$$

The pair (X, T) clearly satisfies the axioms of a topological space.

Example 1.1.4. Let $X = \mathbb{R}^n$ and T be the Euclidean topology, which is a topology induced by the Euclidean metric $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. This metric is a function that assigns two vectors in Euclidean n-space $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ the number $d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + ... + (x_n + y_n)^2}$, which is the distance between two vectors in $\mathbb{R}^n$. With this metric we may define an open n-ball $B^n_\epsilon(\mathbf{x})$ for $\epsilon > 0$ as follow:

$$B^n_\epsilon(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n | d(\mathbf{x}, \mathbf{y}) < \epsilon\}.$$

For our given set X, the Euclidean topology is the empty set $\emptyset$, X, and the unions of open n-balls. If $X = \mathbb{R}$, then the Euclidean topology would have unions of open intervals instead of open n-balls.

Definition 1.1.5. Given a topological space (X, T) and a subset S of X, the *subspace topology* on S is defined by $T_S = \{S \cap U | U \in T\}$, where U is an open set in X. The topological space (S, T_S) is called a *subspace* of (X, T).

Sometimes authors refer to the set X as a topological space or a subspace without explicitly declaring the topology T. We shall occasionally make use of this convention in this paper.

Definition 1.1.6. Let (X, T) be a topological space. A *separation* of X is a pair U, V of disjoint nonempty open subsets of X whose union is X. The space X is said to be *connected* if there does not exist a separation of X.

Roughly speaking, a space can be separated if it can be broken up into two parts that are disjoint open sets; otherwise it is connected. We may also define connectedness for a subspace Y of a topological space X via the following lemma. Refer to [Mun00] for a proof of the lemma. For the sake of completeness, we shall also define a limit point first.

Definition 1.1.7. If A is a subset of the topological space X and if x is a point of X, we say that x is a *limit point* of A if every neighborhood of x intersects A in some point other than x itself.

Lemma 1.1.8. *If Y is a subspace of X, a separation of Y is a pair of disjoint nonempty open sets A and B whose union is Y, neither of which contains a limit point of the other. The space Y is connected if there exists no separation of Y.*

Example 1.1.9. Let X be the subspace $[-1, 1]$ of the real line. The sets $[-1, 0]$ and $(0, 1]$ are disjoint and nonempty, but they do not form a separation of X since $[-1, 0]$ is not open in X. It can be shown that all other pairs of sets whose union is $[-1, 1]$ do not constitute a separation of $[-1, 1]$. Hence X is connected. However, if the point 0 was removed from X, then X would be disconnected.

Example 1.1.10. The set of rational numbers $\mathbb{Q}$ is not connected. The only connected subspaces of $\mathbb{Q}$ are one point sets.

Before we define a manifold, we shall need to define the notions of a homeomorphism, chart, atlas, countable set, and a basis.

Definition 1.1.11. A map $f : X \to Y$ is called a *homeomorphism* if f is continuous, one-to-one, onto, and its inverse f^{-1} is continuous. The sets X and Y are called *homeomorphic* if there exists a homeomorphism between them.

One can imagine two homeomorphic sets as being equivalent to each other via continuous deformations. A classic example illustrating this concept is the deformation of a donut to a coffee mug and vice-versa.

Example 1.1.12. Open intervals of $\mathbb{R}$ are homeomorphic to other open intervals of $\mathbb{R}$. Let $X = (-1, 1)$ and $Y = (a, b)$, where $a < b < \infty$ and let $f : X \to Y$ be defined by

$$f(x) = \frac{b-a}{2}(x + \frac{b+a}{b-a}).$$

We see that f is bijective and continuous since it is linear. Also, the inverse

$$f^{-1}(x) = \frac{2}{b-a}(x + \frac{a+b}{2})$$

exists (since f is injective) and is continuous since it is also linear.

Definition 1.1.13. A set S is *countable* if there exists an injective function f from S to the set of natural numbers $\mathbb{N}$.

A countable set is a set with the same cardinality (number of elements) as the natural numbers.

Definition 1.1.14. A *basis* for a topology X is a collection $\mathcal{B}$ of subsets of X (called *basis elements*) satisfying the following properties.

(1) For each $x \in X$, there is at least one basis element B containing x.

(2) If x belongs to the intersection of two basis elements B_1 and B_2, there is a basis element B_3 containing x such that $B_3 \subset B_1 \cap B_2$.

We may now define a manifold [Boo86].

Definition 1.1.15. A *manifold* of dimension n or n-manifold, is a topological space with the following properties.

(1) M is *Hausdorff*, that is for each pair of points p and q on the manifold M, there exist open sets U and V such that $p \in U, q \in V$ and $U \cap V = \emptyset$.

(2) M is locally Euclidean of dimension n. That is, each point p in M has a neighborhood U which is homeomorphic to an open subset U' of $\mathbb{R}^n$.

(3) M has a countable basis of open sets.

Definition 1.1.16. A pair (U, ϕ), where U is an open set of a manifold M and ϕ is a homeomorphism of U to an open subset of $\mathbb{R}^n$, is called a *chart* or a *coordinate neighborhood*.

See Figure 1.1 for an illustration of a chart on a manifold.

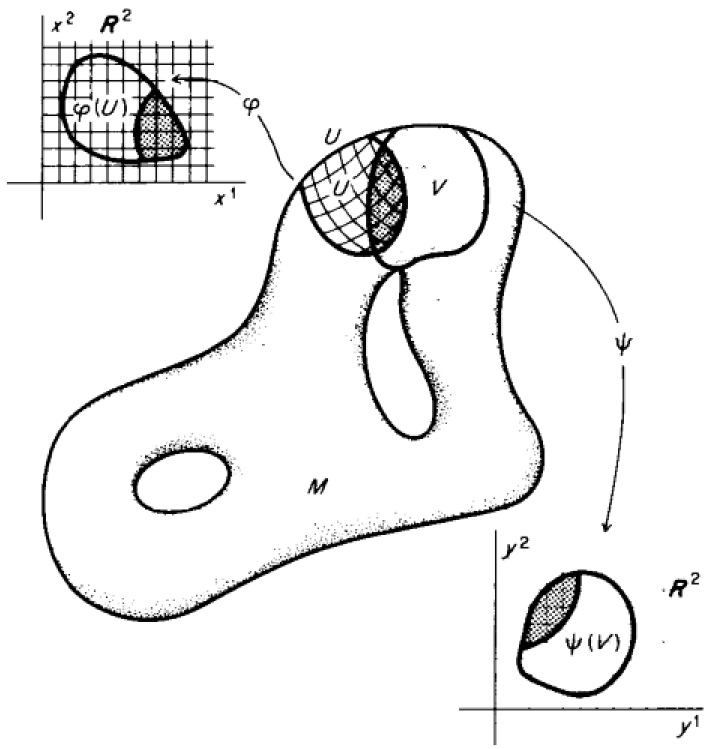

Figure 1.1: Charts on a manifold.[1]

[1] This figure was published in An Introduction to Differentiable Manifolds and Riemannian Geometry, William M. Boothby, Copyright Elsevier (1986). Reprinted with permission by Elsevier.

Definition 1.1.17. An *atlas* for a topological space M is a collection $\{(U_\alpha, \phi_\alpha)\}$ of charts on M such that $\bigcup U_\alpha = M$.

An intuitive real life example that ties together the aforementioned ideas is the planet Earth. We can imagine Earth as a manifold that is round, but locally flat (two dimensional Euclidean space). A single chart cannot adequately describe the entire structure of the manifold (Earth), hence an atlas containing a collection of several of these charts is needed to capture the full structure of the manifold.

Finally, we need to define the notion of smoothness, and a few extra definitions in order to define smooth manifolds.

Definition 1.1.18. Let $U \subset \mathbb{R}^k$ and $V \subset \mathbb{R}^l$ be open sets, where $k, l \in \mathbb{N}$. A mapping $f : U \to V$ is called *smooth* if all the partial derivatives $\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}$ exist and are continuous. If $X \subset \mathbb{R}^k$ and $Y \subset \mathbb{R}^l$ are subsets of Euclidean spaces (not necessarily open!), then $f : X \to Y$ is *smooth* if there exists an open set $U \subset \mathbb{R}^k$ containing X and a smooth mapping $F : U \to \mathbb{R}^l$ that coincides with f in $U \cap X$.

Definition 1.1.19. Suppose X and Y are subsets of Euclidean spaces. Then a map $f : X \to Y$ is called a *diffeomorphism* if f is a *homeomorphism* (continuous bijection with a continuous inverse) and both f and f^{-1} are smooth. The sets X and Y are called *diffeomorphic* if there exists a diffeomorphism between them.

Example 1.1.20. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^3$. Clearly, f is C^∞ and a homeomorphism, but it not a diffeomorphism since $f^{-1}(x) = x^{1/3}$, which is not differentiable at $x = 0$.

Since it is not always possible to explicitly compute inverses, we would like to be able to show that an inverse exists. The following theorem tells us when a function is invertible in a neighborhood of a point in its domain, and gives us a formula for the derivative of the inverse.

Theorem 1.1.21 (Inverse Function Theorem)**.** *Let W be an open subset of $\mathbb{R}^n$ and $F : W \to \mathbb{R}^n$ be a C^r mapping, where $r \in \mathbb{N}$ or ∞. If $a \in W$ and the Jacobian $DF(a)$ is nonsingular, then there exists an open neighborhood U of a in W such that $V = F(U)$ is open and $F : U \to V$ is a C^r diffeomorphism. If $x \in U$ and $y = F(x)$, the derivatives of F^{-1} at y is given by*

$$DF^{-1}(y) = (DF(x))^{-1},$$

where the term on the right denotes the inverse matrix of $DF(x)$.

Refer to [Boo86] for a proof of the Inverse Function Theorem that makes use of the Contraction Mapping Theorem. The Inverse Function Theorem will be useful for proving the Chow-Rashevsky Theorem later on.

Example 1.1.22. Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$F \begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}.$$

The Jacobian DF is given by

$$DF = \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix},$$

thus the determinant of the Jacobian at $a = (r, \theta)$ is

$$r\cos^2(\theta) + r\sin^2(\theta) = r.$$

As long as $r \neq 0$, the Inverse Function Theorem tells us that the inverse F^{-1} exists in the neighborhood of a. In this example, we may explicitly compute $(DF)^{-1}$ via the inversion formula for 2×2 matrices:

$$(DF)^{-1} = \frac{1}{r} \begin{bmatrix} r\cos(\theta) & r\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Converting from polar coordinates to Cartesian coordinates via the usual formulae $x = r\cos(\theta), y = r\sin(\theta)$, we obtain

$$(DF)^{-1} = \begin{bmatrix} x/\sqrt{x^2+y^2} & y/\sqrt{x^2+y^2} \\ -y/(x^2+y^2) & x/(x^2+y^2) \end{bmatrix}.$$

Definition 1.1.23. Let (U, ϕ) and (V, ψ) be coordinate neighborhoods on a manifold M. We say that (U, ϕ) and (V, ψ) are C^∞*-compatible* if $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are diffeomorphisms of the open subsets $\phi(U \cap V)$ and $\psi(U \cap V)$ of $\mathbb{R}^n$, where n=dim M.

Definition 1.1.24. A *differentiable or* C^∞ *(smooth) structure* on a manifold M is a family $\mathbb{U} = \{(U_\alpha, \phi_\alpha)\}$ of coordinate neighborhoods such that:

(1) the U_α cover M,

(2) for any α, β the neighborhoods (U_α, ϕ_α) and (U_β, ϕ_β) are C^∞-compatible,

(3) any coordinate neighborhood (V, ψ) compatible with every $(U_\alpha, \phi_\alpha) \in \mathbb{U}$ is itself in $\mathbb{U}$.

Definition 1.1.25. A C^∞ *(smooth) manifold* is a manifold with a C^∞-differentiable structure.

Definition 1.1.26. Let $C^\infty(p)$ denote the algebra of all smooth functions defined in a neighborhood of p. A *tangent space* to M at p, T_pM, is the vector space of mappings $X_p : C^\infty(p) \to \mathbb{R}$ that satisfies two properties for all functions $f, g \in C^\infty(p)$.

(1) (*Linearity*) For $\alpha, \beta \in \mathbb{R}$, $X_p(\alpha f + \beta g) = \alpha X_p(f) + \beta X_p(g)$.

(2) (*Product/Leibniz Rule*) $X_p(fg) = X_p(f)g(p) + f(p)X_p(g)$.

Then we can define addition and scalar multiplication of vector space operations as follows.

(1) $(X_p + Y_p)(f) = X_p(f) + Y_p(f)$.

(2) $(\alpha X_p)(f) = \alpha X_p(f), \alpha \in \mathbb{R}$.

Elements of the tangent space T_pM are called tangent vectors.

Figure 1.2 portrays the tangent space T_pM and one tangent vector.

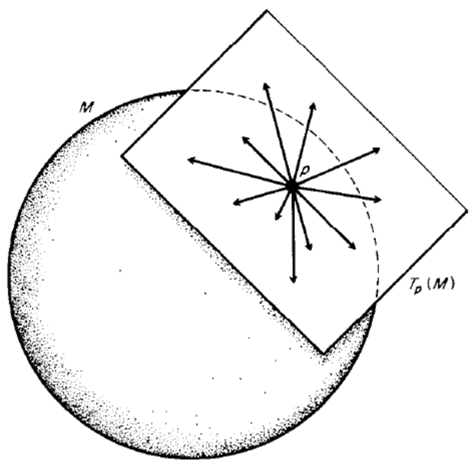

Figure 1.2: The tangent space T_pM at $p \in M$.[2]

Definition 1.1.27. Given a smooth map $F : M \to N$ and $p \in M$ we define the *differential* (or *push forward*) of F at p by $F_*(X_p)(g) = X_p(g \circ F)$, where $X_p \in T_pM$ and $g \in C^\infty(F(p))$. Thus F_* is a map $T_pM \to T_{F(p)}N$. (Sometimes the notation DF_p will be used.) The differential of F at a point p can be thought of as the best linear approximation of F near p, and it pushes tangent vectors on M forward to tangent vectors on N.

[2] This figure was published in An Introduction to Differentiable Manifolds and Riemannian Geometry, William M. Boothby, Copyright Elsevier (1986). Reprinted with permission by Elsevier.

Definition 1.1.28. A differentiable mapping $F : N \to M$ between two smooth manifolds is an *immersion* if the rank of F at each point is equal to the dimension of the domain manifold N. That is, rank $F = n =$ dim N.

Example 1.1.29. Let $F : \mathbb{R} \to \mathbb{R}^3$ be given by $F(t) = (\cos 2\pi t, \sin 2\pi t, t)$. To verify that F is an immersion, we check that the Jacobian has rank 1 at every point. This means one of the derivatives with respect to t differs from zero for every value of t for which the mapping F is defined. We only need to differentiate each component of $F(t)$ with respect to t to obtain the Jacobian of F:

$$\frac{dF}{dt} = (-2\pi \sin 2\pi t, 2\pi \cos 2\pi t, 1).$$

Since the third component of $\frac{dF}{dt}$ is nonzero for all $t \in \mathbb{R}$, the rank of F is 1 for all $t \in \mathbb{R}$, which is equal to $\dim \mathbb{R}$. Hence, F is indeed an immersion.

Definition 1.1.30. A subset $\tilde{N}$ of the differentiable manifold M is an *immersed submanifold* of M if $\tilde{N}$ is the image of a one-to-one immersion $F : N \to M$. That is, $\tilde{N} = F(N)$, is endowed with a topology and C^∞ structure from the correspondence $F : N \to \tilde{N}$, which is a diffeomorphism.

Definition 1.1.31. An *embedding* is a one-to-one immersion $F : N \to M$, which is a homeomorphism of N into M. This means that F is a homeomorphism of N onto its image, $\tilde{N} = F(N)$ endowed with subspace topology. The image of an embedding is called an *embedded submanifold.*

Example 1.1.32. Let $F : \mathbb{R} \to \mathbb{R}^3$ be given by $F(t) = (\cos 2\pi t, \sin 2\pi t, t)$. Here, $N = \mathbb{R}, M = \mathbb{R}^3$, and $N \subset M$. The image $F(N) = F(\mathbb{R})$ is a helix lying on the unit cylinder, whose axis is the x_3 axis in $\mathbb{R}^3$. The differential mapping F is one-to-one immersion and F is a homeomorphism of $\mathbb{R}$ onto $F(\mathbb{R})$, so it is an embedding. This helix is an embedded submanifold of $\mathbb{R}^3$.

Example 1.1.33. Let $F : \mathbb{R} \to \mathbb{R}^2$ be given by $F(t) = (\cos 2\pi t, \sin 2\pi t)$. The image $F(\mathbb{R})$ is the unit circle $S^1 = \{(x_1, x_2) | x_1^2 + x_2^2 = 1\}$ in $\mathbb{R}^2$. Once again, we verify that F is an immersion by checking that the Jacobian has rank 1 at every point.

$$\frac{dF}{dt} = (-2\pi \sin 2\pi t, 2\pi \cos 2\pi t).$$

Since both components of $\frac{dF}{dt}$ are not both zero for all time t, F is an immersion. However, F is not an embedding since it is not a homeomorphism of $\mathbb{R}$ onto its image $F(\mathbb{R})$: the mapping is not one-to-one. It can be shown that S^1 is an embedded submanifold of $\mathbb{R}^2$.

The following figure compares the submanifolds of the two previous examples. In Figure 1.3 part (a) we have the embedded submanifold of $\mathbb{R}^3$, the helix, and in part (b) the embedded submanifold of $\mathbb{R}^2$ is the unit circle.

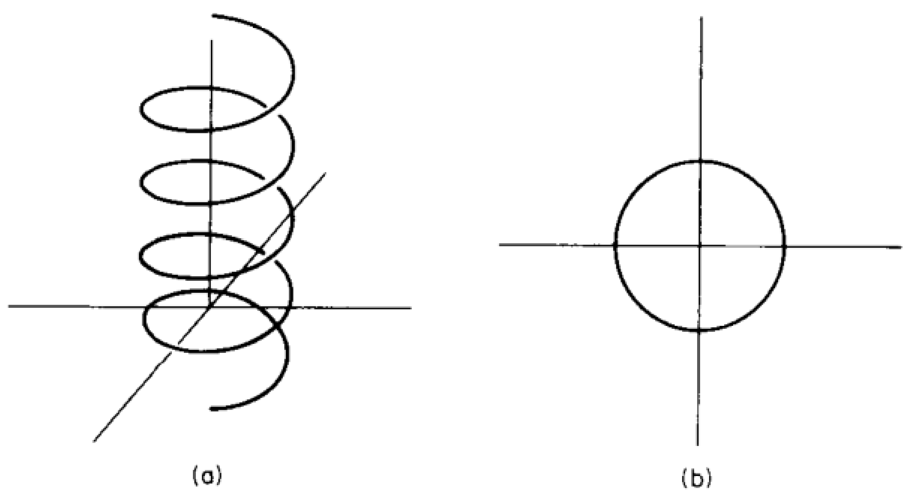

Figure 1.3: Embedded submanifold of $\mathbb{R}^3$ and embedded submanifold of $\mathbb{R}^2$.[3]

[3] This figure was published in An Introduction to Differentiable Manifolds and Riemannian Geometry, William M. Boothby, Copyright Elsevier (1986). Reprinted with permission by Elsevier.

Definition 1.1.34. The *tangent bundle* of a smooth manifold is the disjoint union of all tangent spaces $TM = \bigcup_{p \in M} T_pM$.

It can be shown that the tangent bundle TM is a smooth $2m$-dimensional manifold, where m is the dimension of the manifold M.

1.2 Vector Fields

Definition 1.2.1. A smooth *vector field* X on a smooth manifold M is a smooth map $X : M \to TM$ such that $X_p \in T_pM$, for all $p \in M$.

In coordinates, the formula for a vector field $X(p) \in M$ is given by

$$X_p = \sum_{i=1}^{m} \alpha_i(p) \frac{\partial}{\partial x_i},$$

where $\alpha_i(p)$ are smooth functions of $p \in M$, $\frac{\partial}{\partial x_i}$ denotes the coordinate basis of T_pM, and m is the dimension of the manifold M. Here, we define the coordinate basis at a point p by

$$\frac{\partial}{\partial x_i}\Big|_p = \phi_*^{-1}\left(\frac{\partial}{\partial x_i}\Big|_{\phi(p)}\right)$$

where (U, ϕ) is a chart. The coordinate basis at $\phi(p)$ is just the standard basis e_i:

$$\frac{\partial}{\partial x_i}\Big|_{\phi(p)} = e_i\Big|_{\phi(p)}.$$

A vector field X assigns a tangent vector at p to any point p on M. A weather chart that displays wind velocity and direction is a real life example of a vector field: each point on the chart is assigned a vector that describes the wind's velocity and direction.

Example 1.2.2. Suppose $M = \mathbb{R}^3 \setminus \{0\}$. Then the gravitational field of an object of unit mass at 0 is a smooth vector field with components $\alpha_i, \alpha_2, \alpha_3$ relative to the basis $\frac{\partial}{\partial x_1} = E_1$, $\frac{\partial}{\partial x_2} = E_2$, $\frac{\partial}{\partial x_3} = E_3$ given by

$$\alpha_i = \frac{x_i}{r^3}, \quad i = 1, 2, 3 \quad \text{with } r = ((x_1)^2 + (x_2)^2 + (x_3)^2))^{1/2}.$$

In other words,

$$X = \frac{x_1}{r^3}\frac{\partial}{\partial x_1} + \frac{x_2}{r^3}\frac{\partial}{\partial x_2} + \frac{x_3}{r^3}\frac{\partial}{\partial x_3}.$$

Vector fields can also be viewed as a set of ordinary differential equations, or a dynamical system on M. To see this, we shall first need to define integral curves.

Definition 1.2.3. A curve γ which maps from $t \mapsto \gamma(t)$ defined on an open interval $J \subset \mathbb{R}$ on a manifold M is an *integral curve* of the vector field X if $\dot{\gamma}(t) = X(\gamma(t))$ on J. By definition, an integral curve is connected.

Figure 1.4 shows an integral curve of a vector field.

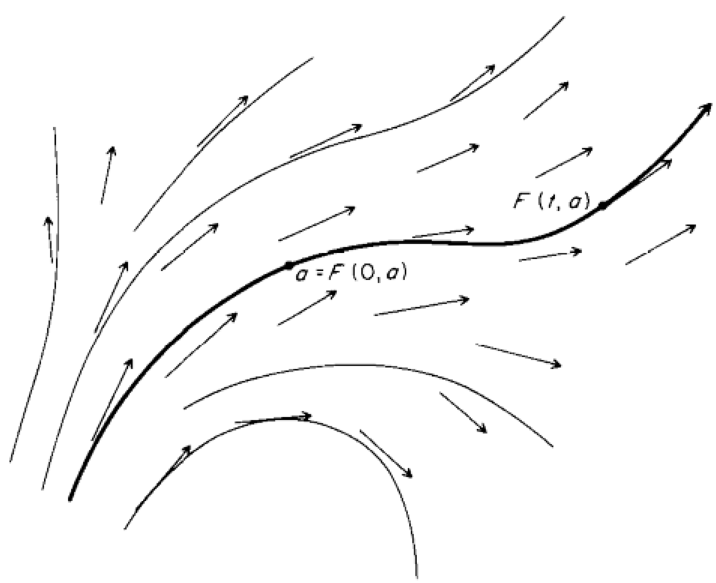

Figure 1.4: Integral curves of a vector field.[4]

The curve γ is the solution to the differential equation $\dot{\gamma}(t) = X(\gamma(t))$. In local coordinates, if $X = \sum_{i=1}^{m} \alpha_i \frac{\partial}{\partial x_i}$, this system can be written as a set of m ordinary differential equations:

$$\dot{\gamma}_1(t) = \alpha_1(\gamma_1(t), \ldots, \gamma_m(t))$$
$$\vdots$$
$$\dot{\gamma}_m(t) = \alpha_m(\gamma_1(t), \ldots, \gamma_m(t)).$$

By the Existence Theorem for ODEs, the existence of integral curves for a smooth vector field is guaranteed.

Theorem 1.2.4 (Existence Theorem for ODEs)**.** *Let $U \subset \mathbb{R}^n$ be an open set and $I_\epsilon, \epsilon > 0$, denote the interval $-\epsilon < t < \epsilon, t \in \mathbb{R}^n$. Suppose $f_i(t, x_1, ..., x_n), i = 1, ..., n$ be C^r functions, $r \geq 1$, on $I_\epsilon \times U$. Then for each $x \in U$ there exists $\delta > 0$ and a neighborhood V of $x, V \subset U$, such that:*

(1) For each $a = (a_1, ..., a_n) \in V$, there exists an n-tuple of C^{r+1} functions $x(t) = (x_1(t), ..., x_n(t))$, defined on I_δ and mapping I_δ into U, which satisfy the system of first-order differential equations

$$\frac{dx_i}{dt} = f_i(t, x), \quad i = 1, ..., n, \tag{1.1}$$

and the initial conditions

$$x_i(0) = a_i, \quad i = 1, ..., n. \tag{1.2}$$

For each a the functions $x(t) = (x_1(t), ..., x_n(t))$ are uniquely determined in the sense that any other functions $\bar{x}_1(t), ..., \bar{x}_n(t)$ satisfying (1.1) and (1.2) must agree with $x(t)$ on an open interval around $t = 0$.

[4] This figure was published in An Introduction to Differentiable Manifolds and Riemannian Geometry, William M. Boothby, Copyright Elsevier (1986). Reprinted with permission by Elsevier.

(2) The functions are uniquely determined by $a = (a_1, ..., a_n)$ for every $a \in V$. We can write them as $x_i(t, a_1, ..., a_n), i = 1, ..., n$, hence they are of class C^r in all variables and thus determine a C^r map of $I_\delta \times V \to U$.

A proof of the Existence Theorem for ODEs may be found in any standard textbook for ODEs such as [Cod89]. If the right hand side of (1.1) is independent of t, then the system of ODEs is called *autonomous*. For our purposes, we shall henceforth assume that the system in (1.1) is autonomous, and that each f_i is C^∞. Let us define a C^∞-vector field X on $U \subset \mathbb{R}^n$ by

$$X = f_1(x)\frac{\partial}{\partial x^1} + ... + f_n(x)\frac{\partial}{\partial x^n}.$$

From our definition of an integral curve, if we write γ in terms of its coordinate functions $\gamma(t) = (x_1(t), ..., x_n(t))$, then the vector equation $\dot{\gamma}(t) = X(\gamma(t))$ is satisfied if and only if $\frac{dx_i}{dt} = f_i(x_1(t), ..., x_n(t)), i = 1, ..., n$, so $x(t) = (x_1(t), ..., x_n(t))$ is a solution of (1.1). For any given $x \in U$, the Existence Theorem says that for each a in a neighborhood V of x there exists a unique integral curve $\gamma(t)$ satisfying $\gamma(0) = a$. Given an arbitrary manifold M instead of $U \subset \mathbb{R}^n$, the existence of an integral curve is also guaranteed.

Theorem 1.2.5. *Let X be a C^∞-vector field on a manifold M. Then for each $p \in M$ there exists a neighborhood V and real number $\delta > 0$ such that there corresponds a C^∞ mapping $\theta_V : I_\delta \times V \to M$, satisfying*

$$\dot{\theta}_V(t, q) = X(\theta_V(t, q)) \tag{1.3}$$

and

$$\theta_V(0, q) = q, \forall q \in V. \tag{1.4}$$

If $\gamma(t)$ is an integral curve of X with $\gamma(0) = q \in V$, then $\gamma(t) = \theta_V(t, q)$ for $|t| < \delta$.

This mapping is unique in the sense that if V_1, δ_1 is another pair for $p \in M$, then $\theta_V = \theta_{V_1}$ on the common part of their domains.

Proof. This theorem is a restatement of the Existence Theorem for ODEs on an arbitrary manifold M, instead of some open set $U \subset \mathbb{R}^n$. For some $p \in M$, we choose a coordinate neighborhood (U, ϕ) and define another vector field $\tilde{X} = \phi_*(X)$ on $\tilde{U} = \phi(U) \subset \mathbb{R}^n$. Now we can apply the local Existence Theorem to obtain $F : I_\delta \times \tilde{V} \to \tilde{U}$ defined by $F(t, a) = (x_t(t, a), \ldots, x_n(t, a))$ on a neighborhood $\tilde{V} \subset \tilde{U}$ of $\phi(p)$. We set $V = \phi^{-1}(\tilde{V})$ and define $\theta_V : I_\delta \times V \to U$ by $\theta_V(t, q) = \phi^{-1}(F(t, \phi(q)))$. Since ϕ and ϕ^{-1} are diffeomorphisms, we see that θ_V satisfies (1.3) and (1.4). This mapping is unique due to the uniqueness of solutions in the Existence Theorem. □

We would like our integral curves to be defined for all time t, a property known as completeness.

Definition 1.2.6. A vector field X is called *complete* if the integral curves through each point p in M are defined for all values of $t \in \mathbb{R}$.

Example 1.2.7. Let X be a vector field on $\mathbb{R}$ be given by $X(x) = x^2$. The ODE $\frac{dx}{dt} = x^2$ with initial condition $x(0) = x_0$ has a unique solution $x(t) = x_0/(1 - x_0 t)$ if $x_0 \neq 0$. For $x_0 \neq 0, x(t)$ is undefined at $t = 1/x_0$, so X is incomplete.

If a manifold M is compact, then every vector field X on M is complete. This fact is a corollary of a lemma regarding integral curves on a manifold. For details and a proof, refer to [Boo86].

Definition 1.2.8. The *flow* of a vector field X is given by a one parameter family of diffeomorphisms $\gamma_t : M \to M$, where M is a compact manifold and $t \mapsto \gamma_t(p)$ is the unique integral curve of X starting at p. Hence, the vector field X is complete on M and γ_t is defined by the flow $\gamma_t(p) = x_p(t)$ where x_p is the unique solution of the ODE $\dot{x} = X(x)$ satisfying the initial condition $x(0) = p$. Otherwise, if M is not compact, we have a *local flow*.

At $t = 0, \gamma_0$ is simply the identity mapping since $\gamma_0(p) = x_p(0) = p$. The Existence Theorem for ODEs says that solutions of $\dot{x} = X(x)$ are unique, which implies that $\gamma_s \circ \gamma_t = \gamma_{s+t}$ for all $s, t \in \mathbb{R}$.

1.3 Lie Brackets

Lie brackets are indispensible tools for understanding and analyzing control systems. We shall first define the Lie bracket, then proceed to present a few useful properties of Lie brackets and examples.

Definition 1.3.1. Let X and Y be smooth vector fields on a manifold M. Then we can define the *Lie bracket* in the following manner.

$$[X, Y](f) = X(Yf) - Y(Xf) \tag{1.5}$$

For any $f \in C^\infty(M)$, the function $Yf : M \to \mathbb{R}$ is defined by $(Yf)(p) = Y_p f$, where Y_p is the value of the vector field Y at $p \in M$.

Intuitively, the Lie Bracket $[X, Y]$ calculates the change of Y along the integral curve of X.

There is another way of defining a Lie bracket from a geometric point of view. Suppose we fix $p \in M$ and we consider the vector $Y_{\phi_t(p)}$, where ϕ_t is the flow of X. We wish to define $[X, Y]_p$ as the rate of change of Y along the flow lines $t \mapsto \phi_t(p)$. We cannot differentiate $Y_{\phi_t(p)}$ with respect to t since the vectors $Y_{\phi_t(p)}$ lie in different tangent spaces for different values of t. We can make use of the time-t map ϕ_t of the flow, which is a diffeomorphism $\phi_t : M \to M$, to compare $Y_{\phi_t(p)}$ with Y_p. Since ϕ_t

takes p to ϕ_t, then the differential map $D\phi_t(p)$ maps T_pM isomorphically onto $T_{\phi_t(p)}M$. (Note that $(D\phi_t(p))^{-1} = D\phi_{-t}(\phi_t(p))$ from the Inverse Function Theorem.) This allows us to compare Y_p with $(D\phi_t(p))^{-1}(Y_{\phi_t(p)})$ to obtain an expression for $[X, Y]_p$:

$$[X, Y]_p = \frac{d}{dt}\Big|_{t=0} (D\phi_t(p))^{-1}(Y_{\phi_t(p)}) = \lim_{t\to 0} \frac{(D\phi_t(p))^{-1}(Y_{\phi_t(p)}) - Y_p}{t}. \tag{1.6}$$

Figure 1.5 demonstrates how the vector field Y varies along the flow of X.

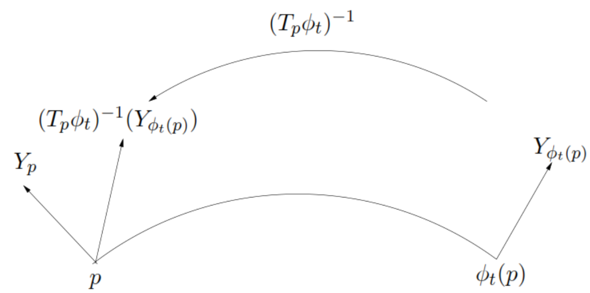

Figure 1.5: How Y varies along the flow of X.[5]

Note that the Lie bracket itself is another vector field on M.

Proposition 1. Let X and Y be two smooth vector fields on a manifold M. Then the Lie bracket $[X, Y]$ is also a vector field.

Proof. Linearity can be easily checked for the Lie bracket, so we just need to verify that it satisfies the Leibniz Rule. Let $f, g \in C^\infty(M)$.

$$\begin{aligned}
[X, Y](fg) = X(Y(fg)) - Y(X(fg)) &= X(Y(f)g + fY(g)) - Y(X(f)g + fX(g)) \\
&= X(Y(f))g + Y(f)X(g) + X(f)Y(g) + fX(Y(g)) \\
&- Y(X(f)) - X(f)Y(g) - Y(f)X(g) - Y(X(g)f) \\
&= [X, Y](f)g + f[X, Y](g).
\end{aligned}$$

□

[5] This figure was reprinted with permission from [Sim11].

If we restrict our vector fields X and Y to be defined on an open subset $U \subset \mathbb{R}^n$, then we can compute Lie brackets via the following formula.

$$[X, Y](x) = DY(x)X(x) - DX(x)Y(x) \tag{1.7}$$

The terms $DX(x)$ and $DY(x)$ denote the Jacobian matrices of X and Y.

Example 1.3.2. Let

$$X = {x_1}^2 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} = \begin{bmatrix} {x_1}^2 \\ x_2 \end{bmatrix}$$

and

$$Y = (x_1 + x_2) \frac{\partial}{\partial x_1} + {x_2}^2 \frac{\partial}{\partial x_2} = \begin{bmatrix} x_1 + x_2 \\ {x_2}^2 \end{bmatrix}$$

be vector fields on $\mathbb{R}^2$. Then we can compute the Jacobian matrices

$$\frac{\partial X}{\partial x} = \begin{bmatrix} 2x_1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\frac{\partial Y}{\partial x} = \begin{bmatrix} 1 & 1 \\ 0 & 2x_2 \end{bmatrix}.$$

Using the definition of a Lie bracket, we find that

$$[X, Y] = \begin{bmatrix} -{x_1}^2 - 2x_1x_2 + x_2 \\ {x_2}^2 \end{bmatrix}.$$

When computing Lie brackets, it is useful to take advantage of the following properties to facilitate computations.

(1) $[X, Y] = -[Y, X]$ (antisymmetry)

(2) $[X_1 + X_2, Y] = [X_1, Y] + [X_2, Y]$

(3) For any smooth functions $a, b : M \to \mathbb{R}$

$[aX, bY] = ab[X, Y] + a(Xb)Y - b(Ya)X$

(4) $[X, \alpha Y + \beta Z] = \alpha[X, Y] + \beta[X, Z]$ where $\alpha, \beta \in \mathbb{R}$ (bilinearity)

(5) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (Jacobi Identity)

Remark 1.3.3. The Lie bracket of two constant vector fields is zero. That is, if $X = a$ and $Y = b$ for $a, b \in \mathbb{R}$, then $[X, Y] = 0$.

Theorem 1.3.4. *Let X and Y be vector fields on a manifold M and let $F : M \to N$ be a diffeomorphism. Then*

$$F_*([X, Y]) = [F_*(X), F_*(Y)]_p, \tag{1.8}$$

for all $p \in M$.

Proof. Let $g \in C^\infty(N)$ be a smooth function. From the definition of the push forward, we have

$$F_*(X_p)g = X_p(g \circ F),$$

so we can write

$$F_*(X)(g) \circ F = X(g \circ F).$$

We can apply this push foward to the Lie bracket $[X, Y]_p$ to obtain

$$F_*[X, Y]_p(g) = [X, Y]_p(g \circ F).$$

Using the definition of the Lie bracket to expand this, we get

$$X_p(Y(g \circ F)) - Y_p(X(g \circ F)).$$

This can be rewritten using the definition of the push forward again as

$$X_p(F_*(Y)g \circ F) - Y_p(F_*(X)g \circ F).$$

Finally, rewriting this equation as

$$F_*(X_p)F_*(Y)(g) - F_*(Y_p)F_*(X)(g),$$

we obtain our desired expression

$$[F_*(X), F_*(Y)]_p(g).$$

□

Theorem 1.3.5. *If the Lie bracket of two vector fields X and Y on a manifold M is 0, the corresponding flows ϕ_t and ψ_t of the vector fields X and Y are said to commute:*

$$[X, Y] \equiv 0 \iff \phi_s \circ \psi_t(p) = \psi_t \circ \phi_s(p) \qquad \forall s, t \in \mathbb{R}, \forall p \in M \tag{1.9}$$

Thus we can view the Lie bracket as a measure of the noncommutativity of two vector fields.

Proof. We first prove $\Rightarrow$. We claim that $[X, Y] = 0$ implies that $D\phi_s(Y_q) = Y_{\phi_s(q)}$ via the following lemma.

Lemma 1.3.6. *If $[X, Y] \equiv 0$, then $\frac{d}{ds}\Big|_s D\phi_{-s}(Y_{\phi_s(p)}) = 0$ for all s such that ϕ_s is well-defined. In other words, $s \mapsto D\phi_{-s}Y_{\phi_s}$ is a constant: $D\phi_s(Y_p) = Y_{\phi_s(p)}$.*

Proof. Let

$$[X, Y]_p = \frac{d}{dr}\Big|_{r=0} D\phi_{-r}(Y_{\phi_r(p)}) \tag{1.10}$$

for all $p \in M$. We apply the pushforward $(\phi_{-s})_* = D\phi_{-s}$ to the Lie bracket $[X, Y]_p$ to obtain

$$D\phi_{-s}([X, Y]_p) = [D\phi_{-s}(X), D\phi_{-s}(Y)]_p = [X, D\phi_{-s}(Y)]_p = 0,$$

where the first equality follows from (1.8). But this last Lie bracket can be rewritten from (1.10) using our geometric definition of the Lie bracket (1.6)

$$\frac{d}{dr}\Big|_{r=0} D\phi_{-r}(D\phi_{-s}(Y_{\phi_{r+s}(p)}) = \frac{d}{dr}\Big|_{r=0} D\phi_{-(r+s)}(Y_{\phi_{r+s}(p)}).$$

Simplifying and applying the chain rule, we get

$$\frac{d}{ds}\Big|_{s=0} D\phi_{-s}(Y_{\phi_s(p)}) = 0.$$

□

Fix a point $p \in M$ and define

$$\gamma(t) = \phi_s(\psi_t(p)), \tag{1.11}$$

where s is fixed and t varies. We want to show that $\gamma(t) = \psi_t(\phi_s(p))$. We take the derivative of both sides of (1.11) with respect to t,

$$\dot{\gamma}(t) = \frac{d}{dt}\phi_s(\psi_t(p)).$$

We apply the chain rule to obtain

$$\dot{\gamma}(t) = D\phi_s(\frac{d}{dt}\psi_t(p)).$$

By definition of the flow, this can be expressed as

$$\dot{\gamma}(t) = D\phi_s(Y_{\psi_t(p)}). \tag{1.12}$$

Let

$$\gamma(0) = \phi_s(p) = q. \tag{1.13}$$

This implies that $\dot{\gamma}(t) = Y_{\phi_s(\psi_t(p))}$ for all t and $\gamma(0) = \phi_s(p)$. Then by the Uniqueness Theorem, $\gamma(t) = \psi_t(\phi_s(p))$. From the result of the lemma, it follows that

$$D\phi_{-s}(Y_{\phi_s(p)}) = Y_p$$

Applying $D\phi_s$ to both sides of the previous equation yields

$$Y_{\phi_s(p)} = D\phi_s(Y_p). \tag{1.14}$$

Combining this result with (1.12) and using uniqueness implies that

$$\gamma(t) = \psi_t(\phi_s(p)) = \phi_s(\psi_t(p)).$$

Now we prove $\Leftarrow$. Assume that $\phi_s(\psi_t(p)) \equiv \psi_t(\phi_s(p))$ $\forall s, t$ and for all $p \in M$. We evaluate the Lie bracket $[X, Y]_p$:

$$[X, Y]_p = \frac{d}{ds}\Big|_{s=0} D\phi_{-s}(Y_{\phi_s(p)}).$$

However, it follows from our assumption that $D\phi_{-s}(Y_{\phi_s(p)}) \equiv Y_p$, which is a constant, so $[X, Y]_p = 0$. This holds for all p, so $[X, Y] = 0$ as desired. □

Example 1.3.7. In a coordinate system, we can obtain commuting vector fields from the basis vectors of the coordinate system. Let us consider the elliptic coordinate system whose elliptic coordinates (u, v) in terms of Cartesian coordinates are given by

$$x = a\cosh u \cos v$$

$$y = a\sinh u \sin v,$$

where u is a nonnegative real number, $v \in [0, 2\pi]$, and a is the distance of either of the foci of the ellipse from the origin on the x-axis. The unit basis vectors of the elliptic coordinate system are given by

$$\frac{\partial}{\partial u} = \sinh u \cos v \frac{\partial}{\partial x} + \cosh u \sin v \frac{\partial}{\partial y} = \begin{bmatrix} \sinh u \cos v \\ \cosh u \sin v \end{bmatrix}$$

$$\frac{\partial}{\partial v} = -\cosh u \sin v \frac{\partial}{\partial x} + \sinh u \cos v \frac{\partial}{\partial y} = \begin{bmatrix} -\cosh u \sin v \\ \sinh u \cos v \end{bmatrix},$$

It can be shown with a bit of computation that

$$[\frac{\partial}{\partial u}, \frac{\partial}{\partial v}] = 0,$$

thus these two vector fields commute.

Using local coordinates to denote the vector field X with its respective differential operator, we may write $L_X = X = \sum_{i=1}^{n} X_i \frac{\partial}{\partial x_i}$. Then the commutator of the operators L_X, L_Y is defined as $[L_X, L_Y] = L_X L_Y - L_Y L_X$, which corresponds to the Lie bracket $[X, Y]$.

Lie brackets may be expressed in coordinates via the following equation.

$$[X, Y] = \sum_{i=1}^{n} \sum_{j=1}^{n} (\frac{\partial X_i}{\partial x_j} Y_j - \frac{\partial Y_i}{\partial x_j} X_j) \frac{\partial}{\partial x_i}, \qquad i = 1, 2, ..., n \tag{1.15}$$

In practice it is easier to use the properties of the Lie bracket for computational purposes.

Finally, we would like to introduce the concept of a Lie algebra, which will be convenient when we discuss the Chow Rashevsky theorem later on.

Definition 1.3.8. Let $F^\infty(M)$ denote the space of all smooth vector fields on M. This is a real vector space under pointwise addition of vectors:

$$(\alpha X + \beta Y)(x) = \alpha X(x) + \beta Y(x) \qquad \forall x \in M \tag{1.16}$$

where α and β are real numbers and X and Y are vector fields. We shall regard $F^\infty(M)$ as an *algebra* with addition given by (1.16) and multiplication given by the Lie bracket. Any algebra equipped with a Lie bracket that satisfies antisymmetry, bilinearity, and the Jacobi identity is called a *Lie algebra*. For any family of vector fields $\mathcal{F}$ we shall denote the Lie algebra of vector fields generated by $\mathcal{F}$ with Lie($\mathcal{F}$).

Example 1.3.9. Let $\mathcal{M}_n(\mathbb{R})$ denote the algebra of $n \times n$ matrices over $\mathbb{R}$ and let XY represent the usual matrix product of X and Y. Then the commutator $[X, Y] = XY - YX$ defines a Lie algebra on $\mathcal{M}_n(\mathbb{R})$.

1.4 Distributions

Definition 1.4.1. A k-dimensional *distribution* Δ on a smooth manifold M is map that assigns each point $p \in M$ a k-dimensional linear subspace of the tangent space at p, which we write as $p \mapsto \Delta(p) \subset T_pM$. If these subspaces are locally spanned by smooth vector fields, then Δ is a smooth distribution.

Given a set of smooth (locally defined) vector fields $X_1, X_2, ..., X_m$, we may define an m-dimensional distribution $\Delta = \text{span}\{X_1, X_2, ..., X_m\}$. To be more specific, Δ is the set of finite linear combinations of the vector fields X_i over the ring of smooth functions $\alpha_i(x)$, i.e. of the form

$$\alpha_1(x)X_1(x) + \alpha_2(x)X_2(x) + ... + \alpha_m(x)X_m(x).$$

Example 1.4.2. Let $M = \mathbb{R}^3 \setminus \{0\}$ and $\Delta(x) = \{v \in \mathbb{R}^3 | v^T x = 0\}$. The distribution $\Delta(x)$ is the tangent space at x to the sphere centered at the origin passing through x.

Definition 1.4.3. Let Δ be a k-dimensional distribution on M. An immersed submanifold $N \subset M$ is called an *integral manifold* for Δ if $T_pN = \Delta(p)$ for every $p \in N$.

Definition 1.4.4. A distribution Δ is *integrable* if for every $p \in M$ there exists an integral manifold of Δ passing through p.

Example 1.4.5 ([Wan13])**.** Consider the smooth distribution Δ on $\mathbb{R}^3$ spanned by two vector fields

$$X_1 = \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_3}, X_2 = \frac{\partial}{\partial x_2}.$$

This distribution is not integrable because there is no integral manifold through the origin. To see this, assume that Δ is integrable. Then the integrable manifold N of

Δ containing the origin must also contain the integral curve of X_1 passing through the origin. The integral curve of X_1 is a part of the x_1-axis, so N contains all points of the form $(t, 0, 0)$ for $|t| < \epsilon$. Likewise, N must contain the integral curves of the vector field X_2 passing through all points $(t, 0, 0)$. Hence, for each $|t| < \epsilon, N$ contains all points of the form $(t, s, 0), |s| < \delta_t$. This means that N contains a part of the x_1x_2-plane that contains the origin. But this is a contradiction because the vector $\frac{\partial}{\partial x_1}$ is a tangent vector to this part of the plane, but is not in $\Delta(p)$ for any $p \neq 0$.

It is natural to ask: under what conditions is a distribution integrable? To answer this question, we define the following notion.

Definition 1.4.6. A distribution Δ is *involutive* if the Lie bracket of two vector fields X, Y in Δ (i.e. $X_p, Y_p \in \Delta(p)$, for all p) is also an element of Δ, i.e. $[X, Y]_p \in \Delta(p)$ for all p.

Definition 1.4.7. A distribution Δ is *nonsingular* if there exists an integer d such that dim $\Delta(p) = d$ for all $p \in M$.

Theorem 1.4.8 (Frobenius Theorem)**.** *A nonsingular distribution is completely integrable if and only if it is involutive.*

For a proof, see [Boo86].

Example 1.4.9. Let $\Delta = \text{span}\{X_1, X_2\}$, where

$$X_1(x) = x_1\frac{\partial}{\partial x_1} + x_2\frac{\partial}{\partial x_2} + x_3\frac{\partial}{\partial x_3}, X_2(x) = \frac{\partial}{\partial x_3}.$$

We will show that Δ is integrable.

Using the properties of Lie brackets, we can write

$$[X_1, X_2] = [x_1\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_3}] + [x_2\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}] + [x_3\frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_3}].$$

Using the formula for coordinates, we find that

$$[x_1\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_3}] = 0, [x_2\frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}] = 0, [x_3\frac{\partial}{\partial x_3}, \frac{\partial}{\partial x_3}] = \frac{\partial}{\partial x_3}.$$

Clearly, $[X_1, X_2] \in \text{span}\{X_1, X_2\}$ since $[X_1, X_2]$ is just a multiple of X_2. Hence the distribution Δ is involutive, and by the Frobenius Theorem, it is also integrable.

CHAPTER 2

CONTROL SYSTEMS

We shall first define a control system in the nonlinear case and various related terminology, and then consider simpler linear control systems. A few examples illustrating linear control systems will be discussed in depth. The sources for this chapter include [Sim01, Lei10, Sas99].

2.1 Definitions and Examples

Definition 2.1.1. A *control system* on a manifold M is a family of vector fields $X(x, u)$ on M parameterized by the controls u. Typically control systems can be characterized as a system of ODEs (dynamical system) by the following equation.

$$\dot{x}(t) = X(x(t), u(t)), \qquad x(t) \in M, \qquad u(t) \in U \subset \mathbb{R}^m \tag{2.1}$$

We describe the *state* of the system by the variable x, the *input* or *control* of the system by u, the *input set* by $U(x)$ and the *state space* (usually Euclidean space or a smooth manifold) by M. If we take the union of all input sets for each $x \in M$, the set $\mathbb{U} = \bigcup_{x \in M} U(x)$ is called the *control bundle.*

For each control system we denote the set of *admissible control functions*, $\mathcal{U}$, that consists of functions $u : [0, T] \to \mathbb{U}$, for some $T > 0$. These control functions are usually piecewise continuous or smooth.

Definition 2.1.2. A curve $x : [0, T] \to M$ is called a *control trajectory* if there exists an admissible control function $u : [0, T] \to \mathbb{U}$ such that for all $t \in [0, T]$, we have $\dot{x}(t) = X(x(t), u(t))$.

Figure 2.1 is an illustration of a control trajectory.

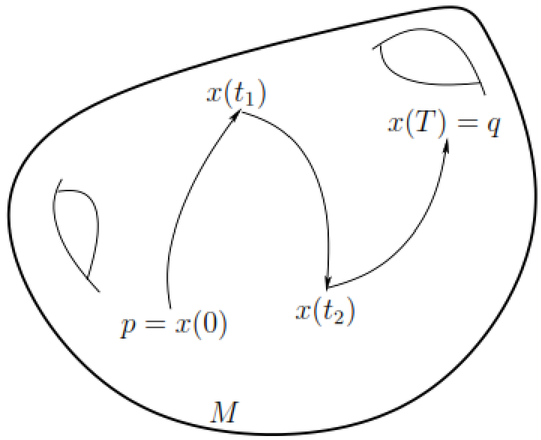

Figure 2.1: A control trajectory connecting p and q.[1]

Example 2.1.3 (Linear Systems)**.** Suppose we have the control system from (2.1). If we consider the case where $M = \mathbb{R}^n$ and X is a linear function of x and u, then we have a *linear control system.*

$$\dot{x} = Ax + \sum_{t=1}^{m} u_i b_i = Ax + Bu \tag{2.2}$$

Here B is the $n \times m$ matrix whose columns are the n-dimensional vectors $b_1, ..., b_m \in \mathbb{R}^n$.

Example 2.1.4 (Affine Systems)**.** Once again we consider the control system from (2.1), but we restrict X to be affine in u. This restriction gives us an *affine control system.*

$$\dot{x} = X_0(x) + \sum_{t=1}^{m} u_i X_i(x) \tag{2.3}$$

[1] This image was reprinted with permission from [Sim01].

The vector field g_0 is called the *drift* and the vector fields $X_1, ..., X_m$ on M are called *control vector fields.* The system is called *drift free* if $X_0 = 0$.

We are interested in the question of *controllability* for control systems. That is, can we steer a system from a given initial state $x(0)$ to some final state $x(T)$ in finite time using the available controls u? We do not consider the optimal ways to achieve this goal; we only care if controllability is feasible. Optimal control uses methods from calculus of variations to consider optimal ways to control the system, but that discussion is beyond the scope of this paper.

Definition 2.1.5. If $x : [0, T] \to M$ is a control trajectory from $x(0) = p$ to $x(T) = q$, then q is called *reachable* or *accessible* from p. We shall denote the set of points reachable from p by $\mathcal{R}(p)$.

Definition 2.1.6. If the interior of $\mathcal{R}$ in M is not empty, the system is *locally accessible at p*. If the system is locally accessible for all p, it is *locally accessible.*

Definition 2.1.7. If $\mathcal{R}(p) = M$ for all p, the system is *controllable.*

Example 2.1.8. The system $\dot{x_1} = u_1, \dot{x_2} = u_2$ where $(u_1, u_2) \in \mathbb{R}^2$ (unconstrained inputs) is (trivially) controllable. Solving the system of ODEs yields $x_1 = tu_1$, and $x_2 = tu_2$. The system is clearly locally accessible at every point in $\mathbb{R}^2$, which coincides with our two dimensional Euclidean state space; hence the system is controllable.

Example 2.1.9. Suppose we have the control system $\dot{x} = \alpha x + u$ where $\alpha \neq 0$ and $u \in [-1, 1]$ (constrained inputs). We shall show that this system is uncontrollable for any $\mathcal{U}$. Consider the case $\alpha > 0$. Let $u \in \mathcal{U}$ be arbitrary and let $x(t)$ be the control trajectory. Since this system is a first order linear ODE, we may use an integrating factor $\mu(t) = e^{\int -\alpha dt} = e^{-\alpha t}$ to obtain an expression for $x(t)$.

$$\begin{aligned}
[x(t)e^{-\alpha t}]' &= ue^{-\alpha t} \\
x(t)e^{-\alpha t} &= x(0) + \int_0^t e^{-\alpha s}u(s)ds \\
x(t) &= e^{\alpha t}x(0) + e^{\alpha t}\int_0^t e^{-\alpha s}u(s)ds \\
&\geq x(0) - e^{\alpha t}\int_0^t e^{-\alpha s}ds \qquad \text{(since } u(s) \geq -1\text{)} \\
&= x(0) + \frac{e^{\alpha t} - 1}{\alpha} \\
&> x(0) \qquad \forall t > 0
\end{aligned}$$

Since points less than $x(0)$ cannot be reached from $x(0)$, the system is uncontrollable. The case for $\alpha < 0$ is similar, hence we claim that the control system is uncontrollable.

CHAPTER 3

MAIN RESULTS

One of the fundamental questions of control theory is the issue of controllability: under what conditions is a control system controllable? The Chow-Rashevsky Theorem gives us the answer for drift free affine systems:

$$\dot{x} = \sum_{t=1}^{m} u_i X_i(x), \tag{3.1}$$

where u_i denotes the control functions and $X_1, ..., X_m$ are vector fields.

3.1 Chow-Rashevsky Theorem

Definition 3.1.1. A distribution $\Delta = \text{span}\{X_1, ..., X_m\}$ on M is said to be *bracket generating* if the iterated Lie brackets $X_i, [X_i, X_j], [X_i, [X_j, X_k]], ..., 1 \leq i, j, k \leq m$ span the tangent space of M at every point.

Theorem 3.1.2 (Chow-Rashevsky)**.** *If M is connected and the control distribution Δ is bracket generating, then the drift free affine system* (3.1) *is controllable.*

Proof. We shall refer to the proof of the general case in [Jea14], but only prove it for the dim M=3, m=2 case. That is, we shall consider the following affine drift free system:

$$\dot{x} = u_1 X_1(x) + u_2 X_2(x), \tag{3.2}$$

where X_1, X_2 are vector fields on $\mathbb{R}^3$. First, we shall utilize the following lemma.

Lemma 3.1.3. *If* (3.2) *satisfies Lie$(X_1, X_2)(q) = T_q M$ for all $q \in M$, then for every $p \in M$ the reachable set $\mathcal{R}_p$ is a neighborhood of p.*

Proof. Let $U \subset M$ be a neighborhood of p in $\mathbb{R}^3$. Let ϕ_t^i denote the flow of the vector field X_i for $i = 1, 2$. Every curve that $t \mapsto \phi_t^i(p)$ is a trajectory of (3.2), so we have

$$\phi_t^i(p) = p + tX_i(p) + o(t).$$

Now we define the local diffeomorphism Φ_t on U by

$$\Phi_t = [\phi_t^1, \phi_t^2] := \phi_{-t}^2 \circ \phi_{-t}^1 \circ \phi_t^2 \circ \phi_t^1.$$

Hence by construction, Φ_t may be expanded as a composition of flows of the vector field $X_i, i = 1, 2$. Thus $\Phi_t(q)$ is the endpoint of a trajectory of (3.2) issued from q.

Lemma 3.1.4. *On a neighborhood of p the following holds*

$$\Phi_t(p) = p + t^2[X_1, X_2](p) + o(t^2).$$

This formula is proven in the appendix of [Jea14]. In order to obtain a diffeomorphism whose derivative with respect to time is exactly $[X_1, X_2]$, we set

$$\Psi_t = \begin{cases} \Phi_{t^{1/2}} & t \geq 0, \\ [\phi_{|t|^{1/2}}^2, \phi_{|t|^{1/2}}^1] & t < 0. \end{cases}$$

Then we have

$$\Psi_t(p) = p + t[X_1, X_2](p) + o(t) \tag{3.3}$$

and $\Psi_t(q)$ is the endpoint of a trajectory of (3.2) issued from q. We now choose commutators $X_1, X_2, [X_1, X_2]$ whose values at p span T_pM. We introduce the map φ defined on a small neighborhood Ω of 0 in $\mathbb{R}^3$ by

$$\varphi(t_1, t_2, t_3) = \Psi_{t_3} \circ \phi_{t_2}^2 \circ \phi_{t_1}^1(p) \in M.$$

Next, we check the derivatives of this map at 0.

$$D_0\varphi(e_1) = \frac{\partial \varphi}{\partial t_1}(0,0,0) = X_1(p)$$
$$D_0\varphi(e_2) = \frac{\partial \varphi}{\partial t_2}(0,0,0) = X_2(p)$$
$$D_0\varphi(e_3) = \frac{\partial \varphi}{\partial t_3}(0,0,0) = [X_1, X_2](p).$$

Note that these three preceding expressions correspond to the differential map of φ defined at the basis vectors e_1, e_2, e_3 for $\mathbb{R}^3$ at the origin. That is, $D_0\varphi(e_i) = \frac{\partial \varphi}{\partial t_i}(0,0,0)$ for $i = 1,2,3$. Thus $D_0\varphi$ is a linear isomorphism that maps the standard basis (e_1, e_2, e_3) of the tangent space $T_0\mathbb{R}^3$ onto $(X_1(p), X_2(p), [X_1, X_2](p))$, the basis of the tangent space T_pM

$$D_0\varphi : (e_1, e_2, e_3) \mapsto (X_1(p), X_2(p), [X_1, X_2](p)).$$

We conclude from (3.3) and by the Inverse Function Theorem that the map φ is a local C^1-diffeomorphism near 0 and has an invertible derivative at 0 since the Jacobian is nonzero at 0. That is, there exists a neighborhood Ω of the origin in $\mathbb{R}^3$ and a neighborhood U of p in M such that $\varphi : \Omega \mapsto U$ is a diffeomorphism. Therefore $\varphi(\Omega)$ contains a neighborhood of p. For every $t \in \Omega, \varphi(t)$ is the endpoint of a concatenation of trajectories of (3.2), the first one being issued from p. It is then the endpoint of a trajectory starting from p. Therefore $\varphi(\Omega) \subset \mathcal{R}_p$, which implies that $\mathcal{R}_p$ is a neighborhood of p. □

Let $p \in M$. If $q \in \mathcal{R}_p$, then $p \in \mathcal{R}_q$. This means $\mathcal{R}_p = \mathcal{R}_q$ for any $q \in M$ and the lemma implies $\mathcal{R}_p$ is an open set. Thus the manifold M is covered by the union of sets $\mathcal{R}_p$ that are pairwise disjointed. Since M is connected, there is only one such open set. This completes the proof of the theorem when dim M=3 and m (number of inputs)=2. □

Example 3.1.5 (Heisenberg group)**.** Let us consider the distribution $\Delta = \text{span}\{X_1, X_2\}$ on $\mathbb{R}^3$. Let

$$X_1(x,y,z) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, X_2(x,y,z) = \begin{bmatrix} 0 \\ 1 \\ x \end{bmatrix}.$$

We have dim Δ=2 and

$$X_3 = [X_1, X_2] = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus X_1, X_2, X_3 span $T\mathbb{R}^3$ at every point and Δ is bracket generating. The triple $(\mathbb{R}^3, \Delta, \langle \cdot, \cdot \rangle)$, where for $v, w \in \Delta$, $\langle v, w \rangle = v_1 w_1 + v_2 w_2$ is an inner product on Δ, called the *Heisenberg group.*

Figure 3.1 is a visual representation of the distribution for the Heisenberg group.

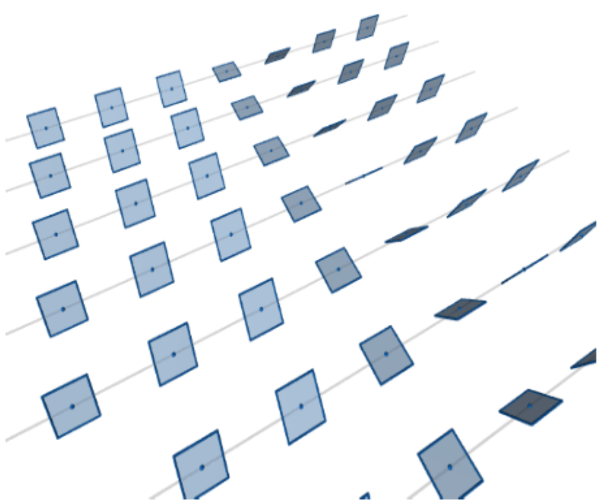

Figure 3.1: The distribution for the Heisenberg group on $\mathbb{R}^3$.[1]

[1] This image was taken from [dM09]. Original image created by Patrick Massot and reprinted with permission.

3.2 Car Example

Example 3.2.1. We can model the motion of a car parked parallel parked between two other cars on a street by the following control system:

$$\dot{x} = u_1 X_1(x) + u_2 X_2(x), \tag{3.4}$$

where $x = (x_1, x_2, x_3)$. Note that $x \in \mathbb{R}^2 \times S^1$: x_1 and x_2 denote the Euclidean coordinates of the center of the rear axle of the car, and x_3 denotes the angle between the rear axle of the car and the x_1 axis, which runs parallel to the street. Let

$$X_1(x) = \begin{bmatrix} \sin x_3 \\ \cos x_3 \\ 0 \end{bmatrix} = (\sin x_3)\frac{\partial}{\partial x_1} + (\cos x_3)\frac{\partial}{\partial x_2} \quad \text{(Rolling)}$$

$$X_2(x) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{\partial}{\partial x_3} \quad \text{(Rotation)}$$

We will show that (3.4) is controllable.

First we use the following property of a Lie bracket to make computations easier: $[X_1 + X_2, Y] = [X_1, Y] + [X_2, Y]$. Applying this to our problem, instead of computing $[X_1, X_2]$, we compute $[\sin x_3 \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_3}] + [\cos x_3 \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}]$. To compute Lie brackets in coordinates, we make use of formula (1.15). It is easy to compute and verify that $[\sin x_3 \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_3}] = -\cos x_3 \frac{\partial}{\partial x_1}$ and $[\cos x_3 \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}] = \sin x_3 \frac{\partial}{\partial x_2}$. Hence

$$[X_1, X_2] = -\cos x_3 \frac{\partial}{\partial x_1} + \sin x_3 \frac{\partial}{\partial x_2} = \begin{bmatrix} -\cos x_3 \\ \sin x_3 \\ 0 \end{bmatrix}.$$

From this it is clear that $[X_1, X_2] \notin \text{span}\{X_1, X_2\}$. However, the vectors $X_1, X_2, [X_1, X_2]$ span the tangent space of $\mathbb{R}^2 \times S^1$. Thus, by Chow's Theorem, the system (3.4) is controllable. This means that we can successful unpark the car regardless of how close the other cars are. If the other cars are too close, we cannot simply turn the car, but we can use a zigzag motion that roughly follows the Lie bracket of $[X_1, X_2]$. If turning is not possible, then setting $x_3 = 0$ yields

$$[X_1, X_2] = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

This result tells us we can unpark the car by driving it parallel in a direction perpendicular (the x_2 axis) to the street.

3.3 Orbit Theorem

The notion of an immersed submanifold, defined in the basic topology section, is crucial for understanding the Orbit Theorem. Recall that $\mathcal{R}_p$ denoted the set of points reachable from p. The Frobenius Theorem gives us the case when $\mathcal{R}_p$ is the smallest, whereas the Chow-Rashevsky Theorem describes the case when $\mathcal{R}_p$ is the biggest, i.e. the manifold M. The Orbit Theorem characterizes the intermediate case when $\mathcal{R}_p$ is between its minimum and maximum size and simply states that these reachable sets are immersed submanifolds of M.

Definition 3.3.1. Let ϕ_t^k denote the flow of the vector field f_k at time t. The *orbit* of a family of vector fields $\mathcal{F}$ on a manifold M through a point $x_0 \in M$ can be defined as $\mathcal{O}(x_0) = \{\phi_{t_k}^k \circ \cdots \circ \phi_{t_1}^1 : t_i \in \mathbb{R}, k \in \mathbb{N}, f_1, \ldots, f_k \in \mathcal{F}\}$.

Points in the orbit of x_0 are points that can be reached by starting at x_0 and traveling along integral curves of vector fields in $\mathcal{F}$ for any time. Note that orbits differ from reachable sets since reachable sets restrict time to forwards time only. See Figure 3.2 to compare reachable sets and orbits.

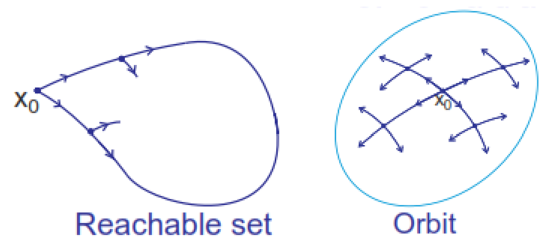

Figure 3.2: Reachable sets versus orbits.[2]

We now state the orbit theorem without proof. For details of the proof, refer to [Jur97].

Theorem 3.3.2 (Orbit Theorem)**.** *Each orbit $\mathcal{O}(x)$ is an immersed submanifold of M.*

Example 3.3.3. If the family $\mathcal{F}$ of vector fields consists of only one smooth vector field X on a smooth manifold M, then the orbit $\mathcal{O}(p)$ of $\mathcal{F}$ through a point p is just the orbit (trajectory) of X through p. If M is the 2-torus T^2 endowed with the quotient space $\mathbb{R}^2/\mathbb{Z}^2$ and

$$X = \frac{\partial}{\partial x} + \alpha\frac{\partial}{\partial y},$$

where $\alpha \notin \mathbb{Q}$, then every orbit of X is a dense immersed submanifold of T^2.

Example 3.3.4. Let $M = \mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R}$, with coordinates $x = (x_1, x_2, x_3, x_4)$, and let $\mathcal{F} = \{X_1, X_2\}$, where

$$X_1(x) = \frac{\partial}{\partial x_1}, \qquad X_2(x) = \frac{\partial}{\partial x_2} + x_1\frac{\partial}{\partial x_3}.$$

[2] This image was reprinted with permission from [Lei10].

Note that X_1, X_2 are the vector fields that define the Heisenberg group on $\mathbb{R}^3$ extended to $\mathbb{R}^4$ in the trivial way (they don't depend on x_4). Then the orbit of $\mathcal{F}$ through (x_1, x_2, x_3, x_4) is the plane $\mathbb{R}^3 \times \{x_4\}$, which is an immersed submanifold of $\mathbb{R}^3 \times \mathbb{R}$.

CHAPTER 4

CONCLUSION

Starting from the foundational concepts of topology, we defined several important notions such as manifolds, tangent spaces, submanifolds, Lie brackets, and distributions. We restricted our attention to affine control systems. In particular, we were interested in answering the question of controllability for affine drift free systems. The Chow-Rashevsky Theorem told us that affine drift free systems were controllable if the manifold M was connected and the control distribution Δ was bracket generating. Although the proof can be generalized for higher dimensional distributions on any connected manifold, we proved the Chow-Rashevsky Theorem in the case of $\dim \Delta = 2$ on $\mathbb{R}^3$ for clarity. Naturally, control systems more complex than the affine drift free control system require more sophisticated conditions that are beyond the scope of this paper. For example, the linear control system (2.2) must fulfill the Kalman rank condition in order to be controllable. Additionally, nonlinear systems are more difficult to deal with. One of the main ways to deal with nonlinear systems is to check if a linearized form of it is controllable, but this only answers questions locally. The question of controllability, along with construction and optimization of specific control trajectories would be worthy of further investigation.

BIBLIOGRAPHY

[Boo86] William M. Boothby, *An introduction to differentiable manifolds and riemannian geometry*, Academic, 1986.

[Cod89] Earl A. Coddington, *An introduction to ordinary differential equations*, Dover Publications, 1989.

[dM09] Arnaud de Mesmay, *The heisenberg group and pansu's theorem*, 2009, http://www.gipsa-lab.fr/~arnaud.demesmay/mesmay.pdf.

[Jea14] Frédéric Jean, *Control of nonholonomic systems: from sub-riemannian geometry to motion planning*, Unpublished paper, 2014.

[Jur97] Velimir Jurdjevic, *Geometric control theory*, Cambridge UP, 1997.

[Kap14] Ilya Kapovich, *Vector and tensor analysis introduction to differential geometry suitable for scientists and engineers*, January 2014, http://www.math.uiuc.edu/~kapovich/481-14/commutator.pdf.

[Lei10] Fátima Silva Leite, *Geometric nonlinear control -an introduction-*, Unpublished paper, 2010.

[Mun00] James Munkres, *Topology*, Pearson, 2000.

[Sas99] Shankar Sastry, *Nonlinear system: Analysis, stability, and control*, Springer, 1999.

[Sim01] Slobodan N. Simić, *A tutorial on geometric control theory*, Unpublished paper, 2001.

[Sim11] ______, *Three faces of the lie bracket*, http://www.math.sjsu.edu/~simic/Spring11/Math213B/lie.bracket.pdf, 2011.

[Wan13] Zuoqin Wang, *Lecture 14: Distributions*, February 2013, http://www-personal.umich.edu/~wangzuoq/437W13/Notes/Lec%2014.pdf.