CONTROL OF SYSTEMS ON LIE GROUPS

by

Claire Tomlin

Memorandum No. UCB/ERL M97/50

21 July 1997

CONTROL OF SYSTEMS ON LIE GROUPS

by

Claire Tomlin

Memorandum No. UCB/ERL M97/50

21 July 1997

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Control of Systems on Lie Groups

Claire Tomlin
Notes for the Course EE290B, taught by Professor S. Sastry
Department of Electrical Engineering
University of California, Berkeley, 94720
Fall, 1994

1 Introduction

These notes were developed from the second part of an Advanced Topics in Control Theory course taught at U.C. Berkeley in the fall of 1994.

The first chapter describes some of the mathematics of matrix Lie groups in a self-contained manner. The second chapter introduces control systems with left-invariant vector fields on matrix Lie groups. The examples are restricted to $SO(3)$ and $SE(3)$, although a section about the Wei-Norman formula discusses how one may deal with the higher dimension case. Some recent work by Walsh, Sarti, and Sastry [8] about steering algorithms on SO(3) is also described.

2 Mathematical Preliminaries

This chapter describes some of the main topics in the mathematics of matrix Lie groups. The coverage is by no means exhaustive; its purpose is to provide a good base for the applications in the next chapter.

2.1 Groups, Fields, and Algebras

We begin with a set of definitions.

Definition 1 (Group) *A group G is a set with a binary operation $(\cdot)$: $G \times G \to G$, such that, $\forall a, b, c$ in G, the following properties are satisfied:*

1. *associativity:* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
2. $\exists$ *an identity* $e \ni a \cdot e = e \cdot a = a$
3. $\exists$ *an inverse* $a^{-1} \ni a \cdot a^{-1} = a^{-1} \cdot a = e$.

A group G is called *abelian* if $a \cdot b = b \cdot a$, $\forall a, b$ in G.

Definition 2 (Homomorphism) *A homomorphism between groups, ϕ : $G \to H$, is a map which preserves the group operation:*

$$\phi(a \cdot b) = \phi(a) \cdot \phi(b).$$

Definition 3 (Isomorphism) *An isomorphism is a homomorphism which is bijective.*

Definition 4 (Field) *A field K is a set with two binary operations: addition $(+)$, and multiplication $(\cdot)$, such that:*

1. *K is an abelian group under $(+)$, with identity* 0
2. *$K - \{0\}$ is an (abelian) group under $(\cdot)$, with identity* 1
3. *$(\cdot)$ distributes over $(+)$* $\ni a \cdot (b + c) = a \cdot b + a \cdot c$.

Some examples of fields are presented below.

$\mathbb{R}$ is a field with addition and multiplication defined in the usual way.

$\mathbb{R}^2$, with addition defined in the usual way and with multiplication defined as:

$$(x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, x_2 y_2)$$

for $\{x_1, x_2, y_1, y_2\}$ in $\mathbb{R}$, is not a field. Why not? If it were, we would have:

$$\begin{aligned}(1,0)\cdot(0,1) &= (0,0)\\ (1,0)^{-1}\cdot(1,0)\cdot(0,1) &= (1,0)^{-1}\cdot(0,0)\\ (0,1) &= (0,0).\end{aligned}$$

This is clearly a contradiction. $\mathbb{R}^2$ can be made into a field if we define $(\cdot)$ as $(x_1,x_2)\cdot(y_1,y_2) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1)$. We denote this field as $\mathbb{C}$, the set of complex numbers, where $(x_1,x_2) = x_1 + \imath x_2$.

If we relax the requirement that $K - \{0\}$ be an abelian group under multiplication, we may define the *quaternions* as a field. This field is denoted $\mathbf{H}$, for *Hamiltonian field.*

The quaternions $\mathbf{H}$ are the set of 4-tuples $(x_1,x_2,x_3,x_4) = (x_1+\imath x_2+\jmath x_3+kx_4)$ with addition defined in the usual way, and multiplication defined according to the following table:

$(\cdot)$	1	$\imath$	$\jmath$	k
1	1	$\imath$	$\jmath$	k
$\imath$	$\imath$	-1	k	$-\jmath$
$\jmath$	$\jmath$	$-k$	-1	$\imath$
k	k	$\jmath$	$-\imath$	-1

In the following we will be defining similar constructions for each of the fields $\mathbb{R}$, $\mathbb{C}$, and $\mathbf{H}$. For ease of notation, we denote the set as $\mathbf{K} \in \{\mathbb{R}, \mathbb{C}, \mathbf{H}\}$.

We write $\mathbf{K}^n$ as the set of all n-tuples whose elements are in $\mathbf{K}$. If we denote $\psi : \mathbf{K}^n \to \mathbf{K}^n$ as a linear map, then ψ has matrix representation $M_n(\mathbf{K}) \in \mathbf{K}^{n\times n}$. $\mathbf{K}^n$ and $M_n(\mathbf{K})$ are both vector spaces over $\mathbf{K}$.

Definition 5 (Algebra) *An algebra is a vector space with a multiplication operation which distributes over addition.*

$M_n(\mathbf{K})$ is an algebra with multiplication defined as the usual multiplication of matrices: for $A, B, C \in M_n(\mathbf{K})$,

$$\begin{aligned}A(B+C) &= AB + AC\\ (B+C)A &= BA + CA.\end{aligned}$$

Definition 6 (Unit) *If $\mathcal{A}$ is an algebra, $x \in \mathcal{A}$ is a unit if there exists $y \in \mathcal{A}$ such that $xy = yx = 1$.*

If $\mathcal{A}$ is an algebra with an associative multiplication operation, and $U \in \mathcal{A}$ is the set of units in $\mathcal{A}$, then U is a group with respect to this multiplication operation.

2.2 Matrix Groups

The class of groups whose elements are $n \times n$ matrices is introduced in this section.

General and Special Linear Groups

- The group of units of $M_n(\mathbf{K})$ is the set of matrices M for which $\det(M) \neq 0$, where 0 is the additive identity of $\mathbf{K}$. This group is called the *general linear group* and denoted by $GL(n, \mathbf{K})$.
- $SL(n, \mathbf{K}) \subset GL(n, \mathbf{K})$ is the subgroup of $GL(n, \mathbf{K})$ whose elements have determinant 1. $SL(n, \mathbf{K})$ is called the *special linear group.*

Orthogonal Matrix Groups

- $O(n, \mathbf{K}) \subset GL(n, \mathbf{K})$ is the subgroup of $GL(n, \mathbf{K})$ whose elements matrices A satisfy the orthogonality condition: $\bar{A}^T = A^{-1}$, where $\bar{A}^T$ is the complex conjugate transpose of A.

Examples of orthogonal matrix groups are:

$O(n) \equiv O(n, \mathbf{R})$ is called the *orthogonal group.*

$U(n) \equiv U(n, \mathbf{C})$ is called the *unitary group.*

$Sp(n) \equiv Sp(n, \mathbf{H})$ is called the *symplectic group.*

Note that for $A \in GL(n, \mathbf{H})$, $\bar{A}$ denotes the complex conjugate of the quaternion, defined by conjugating each element, using

$$\overline{x + iy + jz + kw} = x - iy - jz - kw.$$

An equivalent way of defining the symplectic group is as a subset of $GL(2n,\mathbb{C})$, such that

$$Sp(n) = \{B \in GL(2n,\mathbb{C}) : B^T J B = J; \bar{B}^T = B^{-1}\},$$

where the matrix J is called the *infinitesimal symplectic matrix*, and is written as:

$$J = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & 0 & \cdots & 0 \\ 0 & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{bmatrix}.$$

Special Orthogonal Matrix Groups

- $SO(n) = O(n) \cap SL(n,\mathbb{R})$ is the set of all orthogonal matrices of determinant 1. It is called the *special orthogonal group.*
- $SU(n) = U(n) \cap SL(n,\mathbb{C})$, the set of all unitary matrices of determinant 1, is called the *special unitary group.*

Euclidean Matrix Groups

- The *Euclidean group* is the set of matrices $E(n)$ such that

$$E(n) = \{A \in \mathbb{R}^{(n+1)\times(n+1)} : A = \begin{bmatrix} R & p \\ 0^{1\times n} & 1 \end{bmatrix}, R \in GL(n), p \in \mathbb{R}^n\}.$$

- The *special Euclidean group* is the set of matrices $SE(n)$ such that

$$SE(n) = \{A \in \mathbb{R}^{(n+1)\times(n+1)} : A = \begin{bmatrix} R & p \\ 0^{1\times n} & 1 \end{bmatrix}, R \in SO(n), p \in \mathbb{R}^n\}.$$

Proposition 1 *Let $G \subset M_n(\mathbb{K})$ be a matrix group. Let $\gamma : [a,b] \to G$ be a curve with $0 \in (a,b)$ and $\gamma(0) = I$. Let T be the set of all tangent vectors $\gamma'(0)$ to curves γ. Then T is a real subspace of $M_n(\mathbb{K})$.*

Proof: If γ and σ are two curves in G, then $\gamma'(0)$ and $\sigma'(0)$ are in T. Also, $\gamma\sigma$ is a curve in G with $(\gamma\sigma)(0) = \gamma(0)\sigma(0) = I$.

$$\begin{aligned} \frac{d}{du}(\gamma(u)\sigma(u)) &= \gamma'(u)\sigma(u) + \gamma(u)\sigma'(u) \\ (\gamma\sigma)'(0) &= \gamma'(0)\sigma(0) + \gamma(0)\sigma'(0) \\ &= \gamma'(0) + \sigma'(0). \end{aligned}$$

Since $\gamma\sigma$ is in G, $(\gamma\sigma)'(0)$ is in T. Therefore, $\gamma'(0)\sigma(0) + \gamma(0)\sigma'(0)$ is in T, and T is closed under vector addition.

Also, if $\gamma'(0) \in T$ and $r \in \mathbf{R}$, if we let $\sigma(u) = \gamma(ru)$, then $\sigma(0) = \gamma(0) = I$ and $\sigma'(0) = r\sigma'(0)$. Therefore, $r\sigma'(0) \in T$, and T is closed under scalar multiplication. □

Definition 7 (Dimension of a Matrix Group) *The dimension of the matrix group G is the dimension of the vector space T of tangent vectors to G at I.*

We now introduce a family of matrices which we will use to determine the dimensions of our matrix groups. Let $so(n)$ denote the set of all skew-symmetric matrices in $M_n(\mathbf{R})$,

$$so(n) = \{A \in M_n(\mathbf{R}) : A^T + A = 0\}.$$

Similarly, the set

$$su(n) = \{A \in M_n(\mathbf{C}) : \bar{A}^T + A = 0\}$$

denotes the skew-hermitian matrices, and the set

$$sp(n) = \{A \in M_n(\mathbf{H}) : \bar{A}^T + A = 0\}$$

denotes the skew-symplectic matrices. We also define

$$sl(n) = \{A \in M_n(\mathbf{R}) : \text{trace}(A) = 0\},$$

and

$$se(n) = \{A \in \mathbf{R}^{(n+1)\times(n+1)} : A = \begin{bmatrix} \hat{w} & p \\ 0 & 0 \end{bmatrix}, \hat{w} \in SO(n), p \in \mathbf{R}^n\}.$$

Now consider the orthogonal matrix group. Let $\gamma : [a, b] \to O(n)$, such that $\gamma(u) = A(u)$, where $A(u) \in O(n)$. Therefore, $A^T(u)A(u) = I$. Taking the derivative of this identity with respect to u, we have:

$$A'^T(u)A(u) + A^T(u)A'(u) = 0.$$

Since $A(0) = I$,

$$A'^T(0) + A'(0) = 0.$$

Thus, the vector space $T_I O(n)$ of tangent vectors to $O(n)$ at I is a subset of the set of skew-symmetric matrices, $so(n)$:

$$T_I O(n) \subset so(n).$$

Similarly, we can derive

$$\begin{aligned} T_I U(n) &\subset su(n) \\ T_I Sp(n) &\subset sp(n) \end{aligned}$$

Armed with our definition of the dimension of a matrix group, we conclude that:

$$\begin{aligned} \dim O(n) &\leq \dim so(n) \\ \dim U(n) &\leq \dim su(n) \\ \dim Sp(n) &\leq \dim sp(n) \end{aligned}$$

We will now show that these inequalities are actually equalities.

Definition 8 (Exponential and Logarithm) *The matrix exponential function, $exp : M_n(\mathbf{K}) \to M_n(\mathbf{K})$, is defined in terms of the Taylor series expansion of the exponential:*

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots$$

The matrix logarithm $log : M_n(\mathbf{K}) \to M_n(\mathbf{K})$ is defined only for matrices near the identity matrix I:

$$\log X = (X - I) - \frac{(X-I)^2}{2} + \frac{(X-I)^3}{3} - \ldots$$

Proposition 2 $A \in so(n) \Longrightarrow e^A \in SO(n)$.

Proof: $(e^A)^T = e^{A^T} = e^{A^{-1}} = (e^A)^{-1}$, therefore, $e^A \in O(n)$. Using $\det(e^A) = e^{\mathrm{trace}(A)}$, we have $\det(e^A) = e^0 = 1$. □

Similarly,

$$\begin{aligned} A \in su(n) &\implies e^A \in U(n); \\ A \in sp(n) &\implies e^A \in Sp(n); \\ A \in sl(n) &\implies e^A \in SL(n); \\ A \in se(n) &\implies e^A \in SE(n). \end{aligned}$$

Proposition 3 $X \in SO(n) \implies \log(X) \in so(n)$.

Proof: Noting that $\log(XY) = \log(X) + \log(Y)$ iff $XY = YX$, we take the logarithm on both sides of the equation: $XX^T = X^TX = I$. Thus, $\log(X) + \log(X^T) = 0$, so $\log(X) \in so(n)$. □

Similarly,

$$\begin{aligned} X \in U(n) &\implies \log(X) \in su(n); \\ X \in Sp(n) &\implies \log(X) \in sp(n); \\ X \in SL(n) &\implies \log(X) \in sl(n); \\ X \in SE(n) &\implies \log(X) \in se(n). \end{aligned}$$

The logarithm and exponential thus define maps which send a matrix group G to its tangent space T, and vice versa.

Definition 9 (One Parameter Subgroup) *A one parameter subgroup γ of a matrix group G is a smooth homomorphism $\gamma : \mathbf{R} \to G$.*

The group operation in $\mathbf{R}$ is addition, thus, $\gamma(u+v) = \gamma(u) \cdot \gamma(v)$. Since $\mathbf{R}$ is an abelian group under addition, we have that

$\gamma(u+v) = \gamma(v+u) = \gamma(u) \cdot \gamma(v) = \gamma(v) \cdot \gamma(u)$.

Note that by defining γ on some small neighbourhood U of $0 \in \mathbf{R}$, γ is defined over all $\mathbf{R}$, since for any $x \in \mathbf{R}$, some $\frac{1}{n}x \in U$ and $\gamma(x) = (\gamma(\frac{1}{n}x))^n$.

Proposition 4 *If $A \in M_n(\mathbf{K})$, then e^{Au} is a one parameter subgroup.*

Proof: Noting that $e^{X+Y} = e^X e^Y$ iff $XY = YX$, we have

$$e^{A(u+v)} = e^{Au+Av} = e^{Au}e^{Av},$$

since A commutes with itself. $\square$

Proposition 5 *Let γ be a one parameter subgroup of $M_n(\mathbf{K})$. Then there exists $A \in M_n(\mathbf{K})$ such that $\gamma(u) = e^{Au}$.*

Proof: Define $A = \sigma'(0)$, where $\sigma(u) = \log \gamma(u)$, (ie. $\gamma(u) = e^{\sigma(u)}$). We need to show that $\sigma(u) = Au$, a line through 0 in $M_n(\mathbf{K})$.

$$\begin{aligned} \sigma'(u) &= \lim_{v \to 0} \frac{\sigma(u+v) - \sigma(u)}{v} \\ &= \lim_{v \to 0} \frac{\log \gamma(u+v) - \log \gamma(u)}{v} \\ &= \lim_{v \to 0} \frac{\log \gamma(u)\gamma(v) - \log \gamma(u)}{v} \\ &= \lim_{v \to 0} \frac{\log \gamma(v)}{v} \\ &= \sigma'(0) \\ &= A. \end{aligned}$$

Therefore, $\sigma(u) = Au$. $\square$

So, given any element in the tangent space of G at I, its exponential belongs to G.

Proposition 6 *Let $A \in T_I O(n, \mathbf{K})$, the tangent space at I to $O(n, \mathbf{K})$. Then there exists a unique one parameter subgroup γ in $O(n, \mathbf{K})$ with $\gamma'(0) = A$.*

Proof: $\gamma(u) = e^{Au}$ is a one parameter subgroup of $GL(n, \mathbf{K})$, and γ lies in $O(n, \mathbf{K})$ since $\gamma(u)^T \gamma(u) = (e^{Au})^T e^{Au} = I$. $\square$

Thus,

$$\dim O(n, \mathbf{K}) \geq \dim so(n, \mathbf{K}).$$

But we have shown using our definition of the dimension of a matrix group that

$$\dim O(n, \mathbf{K}) \leq \dim so(n, \mathbf{K}).$$

Therefore,

$$\dim O(n, \mathbf{K}) = \dim so(n, \mathbf{K}),$$

and the tangent space, at I, to $O(n, \mathbf{K})$ is exactly the set of skew-symmetric matrices.

Now the dimension of $so(n, \mathbf{R})$ is easily computable: we simply find a basis. Let E_{ij} be the matrix whose entries are all zero except the ij^{th} entry, which is 1, and the ji^{th} entry, which is -1. Then E_{ij}, for $i < j$, form a basis for $so(n)$. There are $\frac{n(n-1)}{2}$ of these basis elements. Therefore, $\dim O(n) = \frac{n(n-1)}{2}$.

Similarly,

$$\begin{aligned}
\dim SO(n) &= \frac{n(n-1)}{2} \\
\dim U(n) &= n^2 \\
\dim SU(n) &= n^2 - 1 \\
\dim Sp(n) &= n(2n+1).
\end{aligned}$$

2.3 Matrix Lie Groups and their Lie Algebras

We start our discussion of matrix Lie groups with some definitions from differential geometry.

Definition 10 (Topological Space) *A topological space is a set M with a collection of subsets $\mathcal{T}$ of M having the properties:*

1. *$\emptyset$ and M are in $\mathcal{T}$;*
2. *the union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$;*
3. *the intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.*

$\mathcal{T}$ *is called a* topology *on* M.

Definition 11 (Homeomorphism) *A homeomorphism f between two topological spaces M and N is a bijective, continuous map $f : M \to N$ with a continuous inverse $f^{-1} : N \to M$.*

Definition 12 (Manifold) *An n-manifold is a topological space M with the property that, if $x \in M$, then there is some neighbourhood U of x such that U is homeomorphic to $\mathbb{R}^n$.*

Definition 13 (Chart, Atlas, Maximal Atlas) *A chart (φ, U) on an n-manifold M is an open set U of M and a homeomorphism $\varphi : U \to \mathbb{R}^n$. Two charts, (φ, U) and (ϕ, V), are said to have smooth overlap if the maps $\phi^{-1} \circ \varphi : \mathbb{R}^n \to \mathbb{R}^n$ and $\varphi^{-1} \circ \phi : \mathbb{R}^n \to \mathbb{R}^n$ are smooth. A family of charts which covers M and whose members have smooth overlap is called an atlas. A maximal atlas for M is an atlas which contains the maximum number of charts.*

Definition 14 (Differentiable Manifold) *A differentiable manifold is a manifold with an associated maximal atlas.*

The atlas allows us to perform calculus on the manifold: the charts in the atlas provide explicit homeomorphisms which refer the manifold to $\mathbb{R}^n$, a space in which we know how to integrate and differentiate. The requirement that the charts have smooth overlap guarantees that these operations are well-defined over the whole manifold.

We may characterize a smooth function $f : M \to N$, where M is an m-manifold and N is an n-manifold, according to the corresponding atlases on M and N. Let (φ, U) be a chart on M and (ϕ, V) be a chart on N. Let $p \in M$ such that U is an open neighbourhood of p and V is an open neighbourhood of $f(p)$. Then f is said to be smooth at p if

$$\phi \circ f \circ \varphi^{-1} : \mathbb{R}^m \to \mathbb{R}^n$$

is smooth at $\varphi(p)$.

We now introduce the concept of tangent vectors to manifolds. In $\mathbf{R}^n$, tangent vectors to smooth surfaces are easy to picture, however smooth surfaces and tangent vectors in an arbitrary manifold are not as intuitive to visualize.

Let M be a differentiable m-manifold with $p \in M$. Let

$$A(p) = \{(W, f) : p \in W, W \text{ open in } M, f : W \to \mathbf{R} \text{ is smooth}\}.$$

Define vector addition and scalar multiplication on $A(p)$ as, for $(W_1, f_1), (W_2, f_2) \in A(p)$ and $r \in \mathbf{R}$,

$$\begin{aligned} (W_1, f_1) + (W_2, f_2) &= (W_1 \cap W_2, f_1 + f_2) \\ r(W_1, f_1) &= (W_1, r f_1). \end{aligned}$$

Under these operations, $A(p)$ is a real vector space. We make $A(p)$ into an algebra by defining vector multiplication as

$$(W_1, f_1)(W_2, f_2) = (W_1 \cap W_2, f_1 f_2).$$

Definition 15 (Tangent Vectors) *A <u>tangent vector</u> ξ to M at p is a linear map $\xi : A(p) \to \mathbf{R}$ satisfying, for $f, g \in A(p)$:*

1. *if $f = g$ in a neighbourhood of p, then $\xi(f) = \xi(g)$;*
2. *$\xi(fg) = f(p)\xi(g) + \xi(f)g(p)$.*

The second condition above is called the *derivation law.*

Consequences of this definition:

- If f is a constant function ($f(q) \equiv r, \forall q \in U$), then $\forall$ tangent vectors ξ, $\xi(f) = 0$.

  Proof:

$$\begin{aligned} \xi(f \cdot g) &= f(p)\xi(g) + \xi(f)g(p) \\ \xi(r \cdot g) &= r\xi(g) + \xi(f)g(p) \end{aligned}$$

  Therefore, $\xi(f)g(p) = 0$ since $\xi(r \cdot g) = r\xi(g)$. Since this holds $\forall g$, we have $\xi(f) = 0$.

- If $f(p) = g(p) = 0$ then $\xi(fg) = 0$.

Tangent vectors are operators which act on functions: if γ is a smooth curve in a manifold M, then γ gives rise to a linear function

$$\xi \equiv \gamma_*(t) : A(p) \to \mathbf{R}$$

defined by

$$\xi(t)(f) = \gamma_*(t)(f) = (f \circ \gamma)'(t),$$

which may be described as the directional derivative of f at p in the direction of γ. The *tangent space* of M at a point p, denoted T_pM, is the set of all tangent vectors to M at p.

Proposition 7 *T_pM is a real vector space of dimension m, the dimension of M.*

Proof: With the definition of vector addition and scalar multiplication on tangent vectors as follows, for $\xi, \eta \in T_p(M), r \in \mathbf{R}$,

$$\begin{aligned} (\xi + \eta)(f) &= \xi(f) + \eta(f) \\ (r\xi)(f) &= r\xi(f), \end{aligned}$$

it is easy to verify that T_pM is a real vector space.

We now prove that the dimension of T_pM is m.

Let (φ, U) be a chart on M, where $p \in \varphi(U) \subset M$ and $U \in \mathbf{R}^m$. Suppose that M sits in the ambient space $\mathbf{R}^N$, and assume that $\varphi(0) = p$. The best approximation to $\varphi : U \to M$ at 0 is the map:

$$\varphi(u) = \varphi(0) + d\varphi_0(u) = x + d\varphi_0(u).$$

Recall that φ^{-1} is a smooth map from M to $\mathbf{R}^m$ Choose an open set W in $\mathbf{R}^N$ and a smooth map $\Phi' : \mathbf{R}^N \to \mathbf{R}^m$ that extends φ^{-1}. Thus $\Phi' \circ \varphi$ is the identity map of U, so, by the chain rule,

$$\mathbf{R}^m \overset{d\varphi_0}{\to} T_p(M) \overset{d\Phi'_p}{\to} \mathbf{R}^m$$

is the identity map of $\mathbf{R}^m$. Therefore, $d\varphi_0 : \mathbf{R}^m \to T_p(M)$ is an isomorphism, and the dimension of $T_p(M)$ is m. □

Definition 16 (Smooth Vector Fields) *A smooth vector field X on a manifold M is an assignment of $X_p \in T_pM$ for each $p \in M$, such that, if $f : M \to \mathbf{R}$ is a smooth function, then*

$$(Xf)_p \equiv X_p(f) : M \to \mathbf{R}$$

is smooth over p.

The smooth vector fields on M form a real vector space. Indeed, it is easy to check that if X and Y are smooth vector fields, then $X+Y$ is a smooth vector field, since $(X+Y)(f) = X(f) + Y(f)$. If X is a smooth vector field and $r \in \mathbf{R}$ then rX is smooth, where $rX(f) = r(X(f))$.

Definition 17 (Integral Curve) *Let $c : [0,1] \to M$ be a curve on the differential manifold M, and let X be a smooth vector field on M. The curve c is said to be an integral curve of the vector field X if*

$$\dot{c} = X(c(t)).$$

Vector fields thus represent differential equations on manifolds.

The space of smooth vector fields becomes an algebra under the appropriate multiplication operation. If we have two smooth vector fields X and Y, let us define

$$(X \circ Y)(p)(f) \equiv X_p((Yf)_p).$$

Now $(Yf)_p$ is a smooth function from M to $\mathbf{R}$, thus

$$X_p(Yf)_p : A(p) \to \mathbf{R}.$$

However, $X_p(Yf)_p$ may not necessarily be a tangent vector. Consider an example in which

$$\begin{aligned} M &= \mathbf{R}^n, \\ X &= \frac{\partial}{\partial x_1}, \\ Y &= \frac{\partial}{\partial x_2}. \end{aligned}$$

Thus,

$$X_p(Yf)_p = \frac{\partial^2 f}{\partial x_1 \partial x_2}|_p;$$

and, for this example, the derivation law is not satisfied. Therefore, $(X \circ Y)$ is not a tangent vector, so the vector space of smooth vector fields is not closed under this operation.

The candidate multiplication operation under which the vector space of smooth vector fields becomes an algebra must therefore somehow cancel these mixed partial derivatives.

Proposition 8 *For smooth vector fields X, Y, the operator*

$$f \longmapsto X_p(Yf) - Y_p(Xf)$$

is a tangent vector.

Proof: We prove only that the operator defined above satisfies the derivation law:

$$\begin{aligned} X_p(Y(fg)) &= X_p(f(Y(g)) + Y(f)g) \\ &= X_p(f)Y_p(g) + f(p)X_p(Y(g)) + X_p(Y(f))g(p) + Y_p(f)X_p(g). \end{aligned}$$

There is a symmetric formula for $Y_p(X(fg))$. Thus

$$X_p(Y(fg)) - Y_p(X(fg)) = (X_pY - Y_pX)(f)g(p) + f(p)(X_pY - Y_pX)(g). \quad \square$$

We now have a multiplication operation which makes smooth vector fields on M into algebras. Let $[X, Y] = XY - YX$ denote the vector field defined by $[X, Y]_p = X_pY - Y_pX$.

Definition 18 (Lie Algebra) *A Lie algebra is a real vector space, V, with a multiplication operation $[\ ,\ ]$ which satisfies, for $A, B \in V$,*

1. $[A, B] = -[B, A]$;
2. $[A, B + C] = [A, B] + [A, C]$,
   $[A + B, C] = [A, C] + [B, C]$;

3. *for* $r \in \mathbf{R}$, $r[A,B] = [rA,B] = [A,rB]$;
4. $[A,[B,C]] + [B,[A,C]] + [C,[A,B]] = 0$.

The fourth condition is called the *Jacobi Identity*.

Proposition 9 *The set $\mathcal{L}(M)$ of smooth vector fields on a differentiable manifold M forms a Lie Algebra under $[\ ,\ ]$.*

Proof: We have shown that $\mathcal{L}(\mathcal{M})$ is a vector space, and it is a matter of substitution to show that $[X,Y]_p = X_pY - Y_pX$ satisfies the four properties listed above. □

The multiplication operation $[X,Y]$ is called the *Lie bracket* of X and Y.

Definition 19 (Lie Group) *A Lie group is a group G which is also a differentiable manifold such that, for $a, b \in G$,*

1. $(a,b) \longmapsto ab$
2. $a \longmapsto a^{-1}$

are smooth functions.

All finite dimensional Lie groups may be represented as matrix groups. For example, since the function *det* : $\mathbf{R}^{n^2} \to \mathbf{R}$ is continuous, the matrix group $GL(n,\mathbf{R}) = \det^{-1}(\mathbf{R} - \{0\})$ is open. It can be given a differentiable structure which makes it an open submanifold of $\mathbf{R}^{n^2}$. Multiplication of matrices in $GL(n,\mathbf{R})$ is continuous, and smoothness of the inverse map follows from Cramer's Rule. Thus, $GL(n,\mathbf{R})$ is a Lie group. Similarly, $O(n), SO(n), E(n)$, and $SE(n)$ are Lie groups.

In order to study the algebras associated with matrix Lie groups, the concepts of differential maps and left translations are first introduced.

Let M, N be differentiable manifolds and let $M \xrightarrow{\psi} N$ be a smooth map. Then ψ induces a linear map $T_p(M) \xrightarrow{d\psi} T_{\psi(p)}N$:

$$(d\psi \circ \xi)(f) = \xi(f \circ \psi),$$

for $\xi \in T_pM, f \in A(\psi(p))$. The map $d\psi$ is called the *differential* of ψ.

Let G be a Lie group with identity I, and let X_I be a tangent vector to G at I. We may construct a vector field defined on all of G in the following way. For any $g \in G$, define the *left translation* by g to be a map $L_g : G \rightarrow G$ such that $L_g(x) = gx$, where $x \in G$. Since G is a Lie group, L_g is a diffeomorphism of G for each g. Taking the differential of L_g at e results in a map from the tangent space of G at e to the tangent space of G at g:

$$dL_g : T_eG \rightarrow T_gG$$

such that

$$X_g = dL_g(X_e).$$

The vector field formed by assigning $X_g \in T_gG$ for each $g \in G$ is called a *left invariant* vector field.

Proposition 10 *If X and Y are left invariant vector fields on G, then so is $[X, Y]$.*

Proof: Let $g \in G$ and $f \in A(g)$.

$$\begin{aligned} dL_g[X,Y]_e(f) &= [X,Y]_e(f \circ L_g) \\ &= X_e(Y(f \circ L_g)) - Y_e(X(f \circ L_g)) \\ &= dL_gX_e(Yf) - dL_gY_e(Xf)) \\ &= X_g(Yf) - Y_g(Xf) \\ &= [X,Y]_g(f). \quad \square \end{aligned}$$

Also, if X and Y are left invariant vector fields, then $X + Y$ and $rX, r \in \mathbb{R}$ are also left invariant vector fields on G. Thus, the left invariant vector fields of G form an algebra under $[\ ,\ ]$, which is called the *Lie algebra* of G and denoted $\mathcal{L}(G)$. $\mathcal{L}(G)$ is actually a subalgebra of the Lie algebra of all smooth vector fields on G.

With this notion of a Lie group's associated Lie algebra, we can now look at the Lie algebras associated with some of our matrix Lie groups. We first look at three examples, and then, in the next section, study the general map from a Lie algebra to its associated Lie group.

Examples:

- The Lie algebra of $GL(n, \mathbf{R})$ is denoted $gl(n, \mathbf{R})$, the set of all $n \times n$ real matrices. The tangent space of $GL(n, \mathbf{R})$ at the identity can be identified with $\mathbf{R}^{n^2}$ since $GL(n, \mathbf{R})$ is an open submanifold of $\mathbf{R}^{n^2}$. The Lie bracket operation is simply $[A, B] = AB - BA$, matrix multiplication.

- The special orthogonal group $SO(n)$ is a submanifold of $GL(n, \mathbf{R})$, so $SO(n)_I$ is a subspace of $GL(n, \mathbf{R})_I$. The Lie algebra of $SO(n)$, denoted $so(n)$, may thus be identified with a certain subspace of $\mathbf{R}^{n^2}$. We have shown in the previous section that the tangent space at I to $SO(n)$ is the set of skew- symmetric matrices; it turns out that we may identify $so(n)$ with this set. For example, for $SO(3)$, the Lie algebra is:

$$so(3) = \left\{ \hat{w} \equiv \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}, w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \right\}.$$

  The Lie bracket on $so(n)$ is defined as $[\hat{w}_a, \hat{w}_b] = \widehat{(w_a \times w_b)}$, the skew-symmetric matrix form of the vector cross product.

- The Lie algebra of $SE(3)$, called $se(3)$, is defined as follows:

$$se(3) = \left\{ \hat{\xi} = \begin{bmatrix} \hat{w} & v \\ 0 & 0 \end{bmatrix} \quad w, v \in \mathbf{R}^3 \right\}.$$

  The Lie bracket on $se(3)$ is defined as

$$[\hat{\xi}_1, \hat{\xi}_2] = \hat{\xi}_1\hat{\xi}_2 - \hat{\xi}_2\hat{\xi}_1 = \begin{bmatrix} \widehat{(w_1 \times w_2)} & w_1 \times v_2 - w_2 \times v_1 \\ 0 & 0 \end{bmatrix}.$$

2.4 The Exponential Map

In computing the dimension of $O(n, \mathbf{K})$ in Section 2.2, we showed that for each matrix A in $O(n, \mathbf{K})_I$, there is a unique one parameter subgroup γ in $O(n, \mathbf{K})$, with $\gamma(u) = e^{Au}$, such that $\gamma'(0) = A$. In this section we introduce a function

$$\exp : T_eG \to G,$$

for a general Lie group G. This map is called the *exponential map* of the Lie algebra $\mathcal{L}(G)$ into G. We then apply this exponential map to the Lie algebras of the matrix Lie groups discussed in the previous section.

Consider a general Lie group G with identity e. For every $\xi \in T_eG$, let $\phi_\xi : \mathbf{R} \rightarrow G$ denote the integral curve of the left invariant vector field X_ξ passing through e at $t = 0$. Thus,

$$\phi_\xi(0) = e$$

and

$$\frac{d}{dt}\phi_\xi(t) = X_\xi(\phi_\xi(t)).$$

One can show that $\phi_\xi(t)$ is a one parameter subgroup of G. Now the *exponential map* of the Lie algebra $\mathcal{L}(G)$ into G is defined as $\exp : T_eG \rightarrow G$ such that for $s \in \mathbf{R}$,

$$\begin{aligned} \exp(\xi s) &= \phi_\xi(s) \\ \exp(\xi) &= \phi_\xi(1). \end{aligned}$$

Thus, a line ξs in $\mathcal{L}(G)$ is mapped to a one parameter subgroup $\phi_\xi(s)$ of G.

We differentiate the map $\exp(\xi s) = \phi_\xi(s)$ with respect to s at $s = 0$ to obtain $d(\exp) : T_eG \rightarrow T_eG$ such that:

$$d(\exp)(\xi) = \phi_\xi'(0) = \xi,$$

thus, $d(\exp)$ is the identity map on T_eG.

By the inverse function theorem,

$$\exp : \mathcal{L}(G) \rightarrow G$$

is a local diffeomorphism from a neighbourhood of zero in $\mathcal{L}(G)$ onto a neighbourhood of e in G, which is denoted as G_0, the *identity component* of G.

We now discuss the conditions under which the exponential map is surjective onto the Lie group.

Definition 20 (Path Connected) *For any two points x and y of a topological space X, a path in X from x to y is a continuous map $f : [0,1] \rightarrow X$ such that $f(0) = x$ and $f(1) = y$. X is said to be path connected if every pair of points of X can be joined by a path in X.*

G_0 is path connected by construction: the one parameter subgroup $\exp(\xi s) = \phi_\xi(s)$ defines a path between any two elements in G_0.

Proposition 11 *If G is a path connected Lie group and H is a subgroup which contains an open neighbourhood U of e in G, then $H = G$.*

Proof: See Curtis [1].

We may thus conclude that if G is a path connected Lie group, then $\exp : \mathcal{L}(G) \to G$ is surjective. If G is not path connected, $\exp(\mathcal{L}(G))$ is the identity component G_0 of G.

For matrix Lie groups, the exponential map is just the matrix exponential function, e^A, where A is a matrix in the associated Lie algebra.

- For $G = SO(3)$, the exponential map $\exp \hat{w}$, $\hat{w} \in so(3)$, is given by

  $$e^{\hat{w}} = I + \hat{w} + \frac{\hat{w}^2}{2!} + \frac{\hat{w}^3}{3!} + \dots,$$

  which can be written in closed form solution as:

  $$e^{\hat{w}} = I + \frac{\hat{w}}{\| w \|} \sin \| w \| + \frac{\hat{w}^2}{\| w \|^2}(1 - \cos \| w \|).$$

  This is known as *Rodrigues' formula.*

- For $G = SE(3)$, the exponential map $\exp \hat{\xi}$, $\hat{\xi} \in se(3)$ is given by

  $$e^{\hat{\xi}} = \begin{bmatrix} I & v \\ 0 & 1 \end{bmatrix},$$

  for $w = 0$, and

  $$e^{\hat{\xi}} = \begin{bmatrix} e^{\hat{w}} & Av \\ 0 & 1 \end{bmatrix},$$

  for $w \neq 0$, where

  $$A = I + \frac{\hat{w}}{\| w \|^2}(1 - \cos \| w \|) + \frac{\hat{w}^2}{\| w \|^3}(\| w \| - \sin \| w \|).$$

2.5 Canonical Coordinates on Matrix Lie Groups

Let $\{X_1, X_2, \ldots, X_n\}$ be a basis for the Lie algebra $\mathcal{L}(G)$. Since

$$\exp : \mathcal{L}(G) \to G$$

is a local diffeomorphism, the mapping $\sigma : \mathbb{R}^n \to G$ defined by

$$g = \exp\{\sigma_1 X_1 + \ldots + \sigma_n X_n\}$$

is a local diffeomorphism between $\sigma \in \mathbb{R}^n$ and $g \in G$ for g in a neighbourhood of the identity e of G. Therefore, $\sigma : U \to \mathbb{R}^n$, where $U \subset G$ is a neighbourhood of e, may be considered a coordinate mapping with coordinate chart (σ, U). Using the left translation L_g, we can construct an atlas for the Lie group G from this single coordinate chart. The functions σ_i are called the *Lie-Cartan coordinates of the first kind* relative to the basis $\{X_1, X_2, \ldots, X_n\}$.

A different way of writing coordinates on a Lie group using the same basis is to define $\theta : \mathbb{R}^n \to G$ by:

$$g = \exp X_1\theta_1 \exp X_1\theta_2 \ldots \exp X_n\theta_n$$

for g in a neighbourhood of e. The functions $(\theta_1, \theta_2, \ldots \theta_n)$ are called the *Lie-Cartan coordinates of the second kind.*

An example of a parameterization of $SO(3)$ using the Lie-Cartan coordinates of the second kind is just the product of exponentials formula:

$$\begin{aligned} R &= e^{\hat{z}\theta_1} e^{\hat{y}\theta_2} e^{\hat{x}\theta_3} \\ &= \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) \\ 0 & 1 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_3) & -\sin(\theta_3) \\ 0 & \sin(\theta_3) & \cos(\theta_3) \end{bmatrix}, \end{aligned}$$

where $R \in SO(3)$ and

$$\hat{x} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \hat{y} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \hat{z} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This is known as the ZYX Euler angle parameterization. Similar parameterizations are the YZX Euler angles, and the ZYZ Euler angles.

A *singular configuration* of a parameterization is one in which there does not exist a solution to the problem of calculating the Lie-Cartan coordinates from the matrix element of the Lie group. For example, the ZYX Euler angle parameterization for $SO(3)$ is singular when $\theta_2 = -\pi/2$. The ZYZ Euler angle parameterization is singular when $\theta_1 = -\theta_3$ and $\theta_2 = 0$, in which case $R = I$, illustrating that there are infinitely many representations of the identity rotation in this parameterization.

2.6 The Campbell-Baker-Hausdorff Formula

The exponential map may be used to relate the algebraic structure of the Lie algebra $\mathcal{L}(G)$ of a Lie group G with the group structure of G. The relationship is described through the Campbell-Baker-Hausdorff (CBH) formula which is introduced in this section. The notions of conjugation and adjoint maps are first described. The structure of a Lie algebra in terms of its structure constants is also presented.

If M is a differentiable manifold and G is a Lie group, we define a *left action* of G on M as a smooth map $\Phi : G \times M \to M$ such that

1. $\Phi(e, x) = x$ for all $x \in M$
2. for every $g, h \in G$ and $x \in M$, $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$.

The left action of G on itself defined by $C_g : G \to G$:

$$C_g(h) = ghg^{-1} = R_{g^{-1}} L_g h$$

is called the *conjugation map* associated with g.

The derivative of the conjugation map at e is called the *adjoint map*, defined as $Ad_g : \mathcal{L}(G) \to \mathcal{L}(G)$ such that, for $\xi \in \mathcal{L}(G)$, $g \in G$,

$$Ad_g(\xi) = (T_e(C_g))(\xi) = T_e(R_{g^{-1}} L_g)(\xi).$$

If $G \subset GL(n, \mathbb{C})$, then $Ad_g(\xi) = g\xi g^{-1}$.

The lower-case adjoint map $ad_\xi : \mathcal{L}(G) \to \mathcal{L}(G)$ is defined as

$$ad_\xi(\eta) = [\xi, \eta].$$

Lemma 1 (Campbell-Baker-Hausdorff Formula) *If $x, y \in \mathcal{L}(G)$, then*

$$\begin{aligned} Ad_{e^x} y &= e^x y e^{-x} = y + [x,y] + \frac{1}{2!}[x,[x,y]] + \frac{1}{3!}[x[x,[x,y]]] + \ldots \\ &= y + ad_x y + \frac{1}{2!} ad_x^2 y + \frac{1}{3!} ad_x^3 y + \ldots \end{aligned}$$

The CBH formula is a measure of how much x and y fail to commute over the exponential: if $[x, y] = 0$, then $Ad_{e^x} y = y$.

If $\{X_1, X_2, \ldots, X_n\}$ is a basis for the Lie algebra $\mathcal{L}(G)$, the *structure constants* of $\mathcal{L}(G)$ with respect to $\{X_1, X_2, \ldots, X_n\}$ are the values $c_{ij}^k \in \mathbb{R}$ defined by:

$$[X_i, X_j] = \sum_k c_{ij}^k X_k.$$

Lemma 2 *Consider $\mathcal{L}(G)$ with basis $\{X_1, X_2, \ldots, X_n\}$ and structure constants c_{ij}^k with respect to this basis. Then*

$$\prod_{j=1}^{r} \exp(p_j X_j) X_i \prod_{j=r}^{1} \exp(-p_j X_j) = \sum_{k=1}^{n} \xi_{ki} X_k,$$

where $p_j \in \mathbb{R}$ and $\xi_{ki} \in \mathbb{R}$.

Proof of 2: We first prove the lemma for $r = 1$. Using the CBH formula, write:

$$\exp(p_1 X_1) X_i \exp(-p_1 X_1) = X_i + \sum_{k=1}^{\infty} \frac{ad_{X_1}^k X_i}{k!} p_1^k.$$

The terms $ad_{X_1}^k X_i$ are calculated using the structure constants:

$$\begin{aligned} ad_{X_1} X_i &= \sum_{n_1=1}^{n} c_{1i}^{n_1} X_{n_1} \\ ad_{X_1}^2 X_i &= \sum_{n_1=1}^{n} \sum_{n_2=1}^{n} c_{1i}^{n_1} c_{1n_1}^{n_2} X_{n_2} \\ &\cdot \\ &\cdot \\ &\cdot \\ ad_{X_1}^k X_i &= \sum_{n_1=1}^{n} \sum_{n_2=1}^{n} \cdots \sum_{n_k=1}^{n} c_{1i}^{n_1} c_{1n_1}^{n_2} c_{1n_2}^{n_3} \cdots c_{1n_{k-1}}^{n_k} X_{n_k}. \end{aligned}$$

Substitute the above formula for $ad^k_{X_1}X_i$ into $X_i + \sum_{k=1}^{\infty} \frac{ad^k_{X_1}X_i}{k!} p_1^k$, and note that since each of the c_{ij}^k is finite, the infinite sum is bounded. The ξ_{ki} are consequently bounded and are functions of c_{ij}^k, $k!$, and p_1^k. The proof is similar for $r > 1$. □

3 Left Invariant Control Systems on Matrix Lie Groups

This chapter uses the mathematics developed in the previous chapter to describe control systems with left-invariant vector fields on matrix Lie groups. For an n-dimensional Lie group G, the type of system described in this section has state which can be represented as an element $g \in G$. The time differential equation which describes the evolution of g can be written as:

$$\dot{g} = g(\sum_{i=1}^{n} X_i u_i),$$

where the u_i are the inputs, and the X_i are a basis for the Lie algebra $\mathcal{L}(g)$. In the above equation, gX_i is the notation for the left invariant vector field associated with X_i. The equation represents a *driftless* system, since if $u_i = 0$ for all i, $\dot{g} = 0$.

In the first section, the state equation describing the motion of a rigid body on $SE(3)$ is developed. The second section develops a relationship, called the Wei-Norman formula, between the inputs u_i and the Lie-Cartan coordinates of the group. In the third section, the problem of steering a control system on $SO(3)$ is studied through a specific example.

3.1 Frenet-Serret Equations: A Control System on $SE(3)$

In this section, arc-length parameterization of a curve describing the path of a rigid body in $\mathbb{R}^3$ is used to derive the state equation of the motion of this left invariant system.

Consider a curve

$$\alpha(s) : [0,1] \rightarrow \mathbf{R}^3,$$

representing the motion of a rigid body in 3-space. Represent the tangent to the curve as

$$t(s) = \alpha'(s).$$

Constrain the tangent to have unity norm, $\| t(s) \| = 1$, so that

$$< t(s), t(s) > = 1.$$

Now taking the derivative of the above with respect to s, we have

$$< t'(s), t(s) > + < t(s), t'(s) > = 0,$$

so that $t'(s) \perp t(s)$. Denote the norm of $t'(s)$ as

$$\| t'(s) \| = \kappa(s),$$

where $\kappa(s)$ is called the *curvature* of the motion: it measures how quickly the curve is pulling away from the tangent. Let us assume $\kappa > 0$. Denoting the unit normal vector to the curve $\alpha(s)$ as $n(s)$, we have that

$$t'(s) = \kappa(s) n(s),$$

and also

$$< n(s), n(s) > = 1$$

so that $n'(s) \perp n(s)$.

The *binormal* to the curve at s is denoted as $b(s)$, where

$$b(s) = t(s) \times n(s),$$

or equivalently,

$$n(s) = b(s) \times t(s).$$

Let

$$n'(s) = \tau(s) b(s),$$

where $\tau(s)$, called the *torsion* of the motion, measures how quickly the curve is pulling out of the plane defined by $n(s)$ and $b(s)$. Thus

$$\begin{aligned}
b'(s) &= t'(s) \times n(s) + t(s) \times n'(s) \\
&= \kappa(s)n(s) \times n(s) + t(s)\tau(s)b(s) \\
&= \tau(s)t(s) \times b(s) \\
&= \tau(s)n(s),
\end{aligned}$$

where $n(s) \times n(s) = 0$.

Similarly,

$$\begin{aligned}
n'(s) &= b'(s) \times t(s) + b(s) \times t'(s) \\
&= \tau(s)n(s) \times t(s) + \kappa(s)b(s) \times n(s) \\
&= -\tau(s)b(s) - \kappa(s)t(s).
\end{aligned}$$

We thus have:

$$\begin{aligned}
\alpha'(s) &= t(s) \\
t'(s) &= \kappa(s)n(s) \\
n'(s) &= -\tau(s)b(s) - \kappa(s)t(s) \\
b'(s) &= \tau(s)n(s).
\end{aligned}$$

Since $t(s), n(s)$, and $b(s)$ are all orthogonal to each other, the matrix with these vectors as its columns is an element of $SO(3)$:

$$[t(s),\ \ n(s),\ \ b(s)] \in SO(3).$$

Thus,

$$\left[\begin{array}{ccc|c} t(s) & n(s) & b(s) & \alpha(s) \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \in SE(3),$$

and

$$\frac{d}{ds}\left[\begin{array}{ccc|c} t(s) & n(s) & b(s) & \alpha(s) \\ \hline 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{ccc|c} t(s) & n(s) & b(s) & \alpha(s) \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{ccc|c} 0 & -\kappa(s) & 0 & 1 \\ \kappa(s) & 0 & \tau(s) & 0 \\ 0 & -\tau(s) & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right].$$

These are known at the *Frenet-Serret* equations of a curve. The evolution of the Frenet-Serret frame in $\mathbb{R}^3$ is given by

$$\dot{g} = gX,$$

where $g \in SE(3)$ and X is an element of the Lie algebra $se(3)$. We may regard the curvature $\kappa(s)$ and the torsion $\tau(s)$ as inputs to the system, so that if

$$\begin{aligned} u_1 &= \kappa(s) \\ u_2 &= -\tau(s), \end{aligned}$$

then

$$\dot{g} = g \left[\begin{array}{ccc|c} 0 & -u_1 & 0 & 1 \\ u_1 & 0 & -u_2 & 0 \\ 0 & u_2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right],$$

which is a special case of the general form describing the state evolution of a left invariant control system in $SE(3)$.

An example of the general form of a left invariant control system in $SE(3)$ is given by an aircraft flying in $\mathbb{R}^3$:

$$\dot{g} = g \left[\begin{array}{ccc|c} 0 & -u_3 & u_2 & u_4 \\ u_3 & 0 & -u_1 & 0 \\ -u_2 & u_1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right].$$

The inputs u_1, u_2, and u_3 control the *roll*, *pitch*, and *yaw* of the aircraft, and the input u_4 controls the velocity in the forward direction.

Specializing the above to $SE(2)$, we have the example of the unicycle rolling on the plane:

$$\dot{g} = g \left[\begin{array}{cc|c} 0 & -u_2 & 1 \\ u_2 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right].$$

In this case, the input u_2 controls the angle of the wheel.

The previous formulation describes *kinematic* steering problems since it is assumed that we have direct control of the velocities of the rigid bodies. In the control of physical systems, though, we generally only have access to the forces and torques which drive the motion. A more realistic approach would therefore be to formulate the steering problem with a ***dynamic*** model of the rigid body, which uses these forces and torques as inputs. Dynamic models are more complex than their kinematic counterparts, and the control problem is harder to solve.

3.2 The Wei-Norman Formula

In this section we derive the Wei-Norman formula, which describes a relationship between the open loop inputs to a system and the Lie-Cartan coordinates used to parameterize the system.

Consider the state equation of a left-invariant control system on a Lie group G with state $g \in G$:

$$\dot{g} = g(\sum_{i=1}^{n} X_i u_i),$$

where the u_i are inputs and the X_i are a basis of the Lie algebra $\mathcal{L}(g)$.

We may express g in terms of its Lie-Cartan coordinates of the 2^{nd} kind:

$$g(t) = exp(\gamma_1(t)X_1)exp(\gamma_2(t)X_2)\dots exp(\gamma_n(t)X_n).$$

Thus,

$$\begin{aligned} \dot{g} &= \sum_{i=1}^{n} \gamma_i'(t) \prod_{j=1}^{i-1} \exp(\gamma_j X_j) X_i \prod_{j=i}^{n} \exp(\gamma_j X_j) \\ &= g \sum_{i=1}^{n} \gamma_i'(t) (\prod_{j=1}^{n} \exp(\gamma_j X_j))^{-1} X_i (\prod_{j=1}^{n} \exp(\gamma_j X_j)) \\ &= g \sum_{i=1}^{n} \gamma_i'(t) \sum_{k=1}^{n} \xi_{ki}(\gamma) X_k, \end{aligned}$$

Where the last equation results from Lemma 2. If we compare this equation with the state equation, we may generate a formula for the inputs to the system in terms of the Lie-Cartan coordinates:

$$\begin{bmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ \cdot \\ u_n(t) \end{bmatrix} = \begin{bmatrix} & & \\ & \xi_{ij}(\gamma) & \\ & & \end{bmatrix} \begin{bmatrix} \gamma_1'(t) \\ \gamma_2'(t) \\ \cdot \\ \cdot \\ \gamma_n'(t) \end{bmatrix}$$

so that

$$\begin{bmatrix} \gamma_1'(t) \\ \gamma_2'(t) \\ \cdot \\ \cdot \\ \gamma_n'(t) \end{bmatrix} = \begin{bmatrix} & & \\ & \xi_{ij}(\gamma) & \\ & & \end{bmatrix}^{-1} \begin{bmatrix} u_1(t) \\ u_2(t) \\ \cdot \\ \cdot \\ u_n(t) \end{bmatrix}.$$

The above is known as the *Wei-Norman formula.* It transforms the problem from a differential equation on Lie groups to one in $\mathbb{R}^n$: steering from an initial configuration g_i to a final configuration g_f is converted into steering from $\gamma(0)$ to $\gamma(1)$, both vectors in $\mathbb{R}^n$.

3.3 Steering a Satellite on $SO(3)$

This section introduces a set of algorithms for steering on $SO(3)$. The simple structure of the group is exploited to solve for the control inputs directly, instead of using the Wei-Norman formula of the previous section to transform the problem. The algorithms and results are described in detail in the paper by Walsh, Sarti, and Sastry [8].

A satellite is a rigid body floating in space. There are *rotors* or *momentum wheels* attached to its body which create linearly independent momentum fields which rotate the satellite to any configuration in $SO(3)$. The satellite may be modelled as a drift free system on $SO(3)$:

$$\dot{g} = g\hat{b}_1 u_1 + g\hat{b}_2 u_2 + g\hat{b}_3 u_3,$$

where $g \in SO(3)$, $\hat{b}_i \in so(3)$, and the $u_i \in \mathbb{R}$ are scalars. The vector $b_i \in \mathbb{R}^3$ describes the direction and magnitude of the momentum field created by the i^{th} momentum wheel on the satellite. Given an initial state g_i and a desired final state g_f, we wish to find control inputs $u_1(t), u_2(t), u_3(t)$ which will steer the system from g_i to g_f in finite time T.

First, consider the case in which $u_i \neq 0$ for $i \in \{1,2,3\}$. Since we have assumed that the momentum fields are linearly independent, the input vector fields $\hat{b}_1 u_1$, $\hat{b}_2 u_2$, and $\hat{b}_3 u_3$ span the tangent space $so(3)$ at every point of $SO(3)$. If we assume that the inputs are constant (u_1, u_2, u_3) and applied over one second, the solution to the state equation is

$$g_f = g_i \exp(\hat{b}_1 u_1 + \hat{b}_2 u_2 + \hat{b}_3 u_3).$$

Since the exponential map is surjective onto $SO(3)$, for any initial configuration g_i and final configuration g_f, inputs (u_1, u_2, u_3) may be calculated which steer the system from g_i to g_f, hence the system is completely controllable. This is the easy case.

The second case that we consider is the two input system in which $u_1 = 0$. We claim that even without the first vector field, the system is completely controllable. An intuitive way to see why this claim is valid is to consider the ZYZ Euler angle parameterization of $SO(3)$. If rotation about the y axis corresponds to *pitch*, and rotation about the z axis corresponds to *roll*, then rotation about the x axis (*yaw*) may be generated from the two vector fields corresponding to pitch and roll. To see this, let $R \in SO(3)$ be defined as $R = g_i^{-1} g_f$. We may parameterize R as

$$R = e^{\hat{z}\theta_1} e^{\hat{y}\theta_2} e^{\hat{z}\theta_3}.$$

Recalling that the ZYZ Euler angle parameterization is singular at $R = I$, we perturb the representation about $R = I$ with respect to the angles θ_2 and θ_3 to obtain:

$$\begin{aligned}
\left.\frac{dR}{d\theta_3}\right|_{R=I} &= \hat{z} \\
\left.\frac{dR}{d\theta_2}\right|_{R=I} &= (e^{-\hat{z}\theta_3}\hat{y}e^{\hat{z}\theta_3}) \\
&= (Ad_{e^{-\hat{z}\theta_3}}\hat{y}) \\
&= \hat{y} + ad_{-\hat{z}\theta_3}\hat{y} + \frac{1}{2!}ad^2_{-\hat{z}\theta_3}\hat{y} + \ldots \\
&= \hat{y} + -\theta_3[\hat{z}, \hat{y}] + \ldots
\end{aligned}$$

Where the last equation results from the CBH formula. Since

$$[\hat{z}, \hat{y}] = (\widehat{z \times y}) = -\hat{x},$$

perturbations of R with respect to θ_2 produce motion in the direction of x.

Considering the same control problem as in the three input case, we proceed to calculate the required inputs $(u_1(t), u_2(t))$ through a series of steps.

If the momentum vectors b_2 and b_3 are not orthogonal, the first step is to orthogonalize their actions using the Gram-Schmidt algorithm. If v_2 corresponds to pitch and v_3 corresponds to roll, then Gram-Schmidt produces:

$$\begin{bmatrix} u_3 \\ u_2 \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} v_3 \\ v_2 \end{bmatrix},$$

where

$$\begin{aligned} \beta_{11} &= (\| b_3 \|)^{-1}, \\ \beta_{22} &= (\| b_2 - b_2^T b_3 \beta_{11}^2 b_3 \|)^{-1}, \\ \beta_{12} &= -b_2^T b_3 \beta_{11}^2 \beta_{22}. \end{aligned}$$

Thus, defining a_1, a_2, and a_3 as the lengths of time that v_3 (roll) is applied, v_2 (pitch) is applied, and then v_3 applied again, the equation to solve is:

$$g_i^{-1} g_f = \exp(\beta_{11} \hat{b}_3 a_1) \exp(\beta_{12} \hat{b}_3 a_2 + \beta_{22} \hat{b}_2 a_2) \exp(\beta_{11} \hat{b}_3 a_3).$$

The second step is to transform the system so that the input vector fields are the canonical ones for $\mathbb{R}^3$. We construct the matrix $K \in SO(3)$:

$$K = [\beta_{11} b_3 \quad (\beta_{12} b_3 + \beta_{22} b_2) \quad (\beta_{11} b_3 \times (\beta_{12} b_3 + \beta_{22} b_2))].$$

Now $K^{-1} \beta_{11} b_3 = e_1$ and $K^{-1}(\beta_{12} b_3 + \beta_{22} b_2) = e_2$, where e_1 and e_2 are the standard first two basis vectors for $\mathbb{R}^3$. Defining the similarity transform as

$$\tilde{g}(t) = K^{-1} g_i^{-1} g(t) K,$$

and taking the derivative of the above, results in

$$\dot{\tilde{g}}(t) = \tilde{g}(t)(\hat{e}_1 v_1 + \hat{e}_2 v_2),$$

which is in the desired canonical representation.

The third step is to solve the general form of the roll-pitch-roll equation for the coordinates (a_1, a_2, a_3). Denoting

$$\begin{aligned} \sin a_1 &\equiv s_{a_1} \\ \cos a_2 &\equiv c_{a_2} \end{aligned}$$

etc., we obtain

$$\begin{aligned} \tilde{g}_f \equiv (g_i K)^{-1} g_f K &= \exp(\hat{e}_1 a_1) \exp(\hat{e}_2 a_2) \exp(\hat{e}_1 a_3) \\ &= \begin{bmatrix} c_{a_2} & s_{a_2} s_{a_3} & s_{a_2} c_{a_3} \\ s_{a_1} s_{a_2} & c_{a_1} c_{a_3} - s_{a_1} c_{a_2} s_{a_3} & -c_{a_1} s_{a_3} - s_{a_1} c_{a_2} c_{a_3} \\ -c_{a_1} s_{a_2} & s_{a_1} c_{a_3} + c_{a_1} c_{a_2} s_{a_3} & -s_{a_1} s_{a_3} + c_{a_1} c_{a_2} c_{a_3} \end{bmatrix}. \end{aligned}$$

Denoting the elements of $\tilde{g}_f$ as $\tilde{g}_{ij}$, we may solve for (a_1, a_2, a_3):

$$\begin{aligned} a_1 &= \operatorname{atan2}(\tilde{g}_{21}, -\tilde{g}_{31}) \text{ if } \tilde{g}_{31} \neq 0 \\ &= \operatorname{acot2}(-\tilde{g}_{31}, \tilde{g}_{21}) \text{ else} \\ a_2 &= \operatorname{atan2}(\tilde{g}_{11} \sin(a_1), \tilde{g}_{21}) \text{ if } \tilde{g}_{21} \neq 0 \\ &= \operatorname{atan2}(\tilde{g}_{11} \cos(a_1), -\tilde{g}_{31}) \text{ else} \\ a_3 &= \operatorname{atan2}(\tilde{g}_{12}, \tilde{g}_{13}) \text{ if } \tilde{g}_{13} \neq 0 \\ &= \operatorname{acot2}(\tilde{g}_{13}, \tilde{g}_{12}) \text{ else} \end{aligned}$$

Finally, in the fourth step, the a_i calculated in the previous step are used to compute the actual controls. Assuming the system is steered from g_i to g_f in the time duration T, the controls below are applied each for a duration of $\frac{T}{3}$:

$$\begin{aligned} (u_1, u_2)_1 &= (3\beta_{11} \frac{a_1}{T}, 0) \\ (u_1, u_2)_2 &= (3\beta_{12} \frac{a_2}{T}, 3\beta_{22} \frac{a_2}{T}) \\ (u_1, u_2)_3 &= (3\beta_{11} \frac{a_3}{T}, 0). \end{aligned}$$

3.4 Concluding Remarks

The representation of a system as one with left-invariant vector fields on matrix Lie groups:

$$\dot{g} = g(\sum_{i=1}^{n} X_i u_i),$$

is a natural one for systems describing the motion of a rigid body with respect to a coordinate frame attached to the body. Aircraft, underwater vehicles, and satellites such as the one modelled in the previous section are all important examples of systems which may be modelled and controlled using matrix Lie groups.

The appeal of this theory is that it is mathematically simple: once the system has been modelled using matrix Lie groups, the computation of the controls required to place the system in a desired final configuration is no harder than the corresponding problem in linear system theory.

The theory of Lie groups and Lie algebras is of current interest in optimal control theory, in which a control solution is sought to minimize a prespecified cost function. The optimal control problem reduces to solving two differential equations - a problem which is theoretically simple on $\mathbf{R}^n$ but becomes very complicated on a differentiable manifold, since the differential equations are only defined locally on each coordinate chart. If the manifold is a Lie group, however, the optimal control problem may be simplified. Using the exponential map from $\mathcal{L}(G)$ to G, the problem may be formulated on the Lie algebra $\mathcal{L}(G)$ of the group. For groups such as $SO(3)$ and $SE(3)$, the Lie algebras are isomorphic to R^n, allowing for a global definition of the differential equations. Optimal control on Lie groups therefore not very much more complicated than optimal control on R^n.

References

[1] M.L. Curtis. *Matrix Groups.* Springer-Verlag, 1979.

[2] V. Guillemin, A. Pollack. *Differential Topology.* Prentice-Hall, 1974.

[3] J.R. Munkres. *Topology, a first course.* Prentice-Hall, 1975.

[4] R.M. Murray, Z.X. Li, S.S. Sastry. *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1994.

[5] A. Sarti, G. Walsh, S.S. Sastry. *Steering left-invariant control systems on matrix Lie groups* . Proceedings of the 1993 IEEE Conference on Decision and Control, 1993.

[6] M. Spivak. *A Comprehensive Introduction to Differential Geometry, Vol 1.* Publish or Perish, 1979.

[7] V.S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations.* Prentice-Hall, 1974.

[8] G. Walsh, A. Sarti, S.S. Sastry. *Algorithms for steering on the group of rotations* . Proceedings of the 1993 IEEE Automatic Control Conference, 1993.