Matrices Lie: An introduction to matrix Lie groups and matrix Lie algebras

By

Max Lloyd

A Journal submitted in partial fulfillment of the requirements
for graduation in Mathematics.

Abstract:

This paper is an introduction to Lie theory and matrix Lie groups. In working with familiar transformations on real, complex and quaternion vector spaces this paper will define many well studied matrix Lie groups and their associated Lie algebras. In doing so it will introduce the types of vectors being transformed, types of transformations, what groups of these transformations look like, tangent spaces of specific groups and the structure of their Lie algebras.

Whitman College
2015

Contents

1 Acknowledgments

A special thank you to Barry Balof, Moira Gresham and Patrick Keef for providing a guiding hand throughout the project, Alex Porter for editing my papers and John Stillwell, whose book *Naive Lie Theory* is the basis of this paper. Without your support and instruction this would not have been possible.

2 Introduction

Lie theory was developed by mathematician Sophus Lie in the late 19^{th} century. In essence it rests on the theory of *continuous groups* or groups with a continuous operation. In this paper we specifically discuss groups of matrices corresponding to linear transformations on different spaces. In order to do this we will utilize the conventions and structure presented in [1]. In general *Lie theory* is strongly rooted in topology and the group structure of *smooth differentiable manifolds*. This, however, is beyond the scope of this paper. Since we only aim to introduce the fundamental concepts of Lie theory we choose to narrow our focus onto *matrix Lie groups* or Lie groups whose elements are all matrices. What makes a group "Lie" is that it has an associated vector algebra or *Lie algebra*. This algebra can be found by exploiting the continuous nature of a Lie group and bestowing upon it the structure of a *Lie Bracket*. This algebra can be far easier to work in so operations can be made here before moving back to the group through a process called exponentiation. Though we do not talk explicitly about exponentiation the purpose of this paper is to get the reader to become antiquated with what a matrix Lie group is and how it generates an algebra.

To help the reader understand this process the paper will focus specifically on the *general geometric groups* which are groups of matrices corresponding to rotations and isometries of real, complex and quaternion vector spaces. This is done by slowly constructing the definitions of each group before finding their Lie algebras. Specifically, in section 3 we will introduce what real, complex and quaternion elements look like and learn how we represent them. In section 4 we will explicitly define the types of transformations we are interested in and the conditions for matrices in each geometric group. In section 5 we will define the actual elements in each general geometric group. In section 6 we will use the explicit elements in each geometric group to define their Lie algebra and in section 7 we will dive into the structure of each Lie algebra.

3 Types of Numbers and Their Representations

To begin we will discuss and use of specific numbers and their representations. Specifically we will examine the reals ($\mathbb{R}$), complexes ($\mathbb{C}$) and quaternions ($\mathbb{H}$). Usually when one thinks about something real or complex they imagine a *number* like 2 or $3 + 4i$. This, however, is just a *representation* of an element in either set. For the purposes of this paper we will define each element of our three sets using a point in a real vector field and choose to express it as either a number or a matrix. In this way we say that a real number characterizes a one dimensional point, a complex number characterizes an ordered pair, and a quaternion characterizes an ordered quadruple. The explicit form of for each is outlined in the table in Figure 1. Each number will now be discussed in greater detail.

3.1 Real ($\mathbb{R}$)

We should be familiar with the real numbers. Let's start with some real number a. This characterizes a point that is at the location a on a number line and can be represented as either the number a or the 1×1 matrix $[a]$. The difference between these two representations is trivial. Clearly all algebraic operations behave exactly the same between the numerical and matrix representations. The only other operation worth note is that $\det([a]) = a$ which the distance from a to the origin.

3.2 Complex ($\mathbb{C}$)

Now consider some complex number $z = a + ib$. This characterizes the ordered pair $(a, b) \in \mathbb{R}^2$ and can be represented as the 2×2 real matrix

$$Z = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

We want to explore this matrix a bit farther. Notice that we can decompose z into

$$Z = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

From now on we will call any $n \times n$ identity matrix *unity* and denote it $\mathbf{1}$. Now let us explore the second matrix in the term a bit further. Notice that

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -\mathbf{1}.$$

This means that

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \sqrt{\mathbf{-1}}$$

Type	Characterizes	As a Number	As a Matrix
"Real" ($\mathbb{R}$)	$(a) \in \mathbb{R}$	a	$[a]$
"Complex" ($\mathbb{C}$)	$(a, b) \in \mathbb{R}^2$	$a + bi$	$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$
"Quaternion" ($\mathbb{H}$)	$(a, b, c, d) \in \mathbb{R}^4$	$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$	$\begin{bmatrix} a + bi & -c - di \\ c - di & a - bi \end{bmatrix}$

Figure 1: Numerical and matrix representations of reals, complexes and quaternions

and therefore behaves just like the imaginary number i. Thus, we can rewrite our matrix form as $Z = a\mathbf{1} + b\sqrt{\mathbf{-1}}$ which looks strikingly similar to our numerical form for a complex number. This offers a linear algebra explanation as a rationale behind the behavior of complex numbers.

Now we want to explore the matrix form a bit further. First notice that

$$\det(Z) = \begin{vmatrix} a & -b \\ b & a \end{vmatrix} = a^2 + b^2$$

which is the distance between (a, b) and the origin. Now lets explore some more properties of the matrix form.

Let $Z_1 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and $Z_2 = \begin{bmatrix} c & -d \\ c & d \end{bmatrix}$ be two complex number in matrix from. First notice that

$$Z_1 Z_2 = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac - bd & -(ad + bc) \\ ad + dc & ac - bd \end{bmatrix} = Z_2 Z_1$$

which is the matrix form of the complex number

$$ad - dc + i(ac - bd) = (a + bi)(c + di).$$

This is an expected result and shows that multiplication between representations is isomorphic which includes the abelian nature of complex multiplication. For the remainder of the paper we will define the *conjugate* of any complex number $z = a + ib$ as $\overline{z} = a - ib$ and $\overline{Z} = a\mathbf{1} - b\sqrt{\mathbf{-1}}$ for any complex $Z = a\mathbf{1} + b\sqrt{\mathbf{-1}}$ in matrix form.

3.3 Quaternion ($\mathbb{H}$)

To begin our discussion of quaternions first examine some complex number z. Notice that we associate it with the ordered pair $(a, b) \in \mathbb{R}^2$ and number $a + ib$ under the condition that $i^2 = -1$. We can do the same for ordered quadruples. Start with the ordered quadruple $(a, b, c, d) \in \mathbb{R}^4$ we can now associate it with the quaternion

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}.$$

under the condition that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{-1}$, $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$, $\mathbf{ki} = -\mathbf{ik} = \mathbf{j}$ and $\mathbf{jk} = -\mathbf{kj} = \mathbf{i}$. Or more succinctly

$$\mathbf{ijk} = \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1.$$

This is the numerical representation of a quaternion. We will call the ladder equation the *Hamilton Relation*. A note on quaternions is that they are *non-commutative*; that is for any two quaternions q_1 and q_2 it is not usually the case that $q_1 q_2 = q_2 q_1$. This is a critical feature of quaternions and is what sets them apart from most other types of numbers.

Now let us discuss the matrix form of a quaternion. Any quaternion $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ has the matrix form

$$q = \begin{bmatrix} a + id & -b - ic \\ b - ic & a - id \end{bmatrix} = \begin{bmatrix} \alpha & -\beta \\ \beta^* & \alpha^* \end{bmatrix}$$

if we let $\alpha = a + id$ and $\beta = b + ic$. We can treat any quaternion as either a 2×2 complex matrix or a 4×4 real matrix (if we treat each complex entry as a real 2×2 block). Just like the matrix form of a complex number notice that we can break down the matrix form of a quaternion into

$$q = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + c \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

and define

$$\mathbf{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \mathbf{j} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \mathbf{k} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

Thus, as a matrix, $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Now we want to verify that the matrices **i**, **j** and **k** verify the Hamilton relation. Notice that

$$\mathbf{i}^2 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}^2 = \mathbf{-1},$$

$$\mathbf{j}^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \mathbf{-1},$$

and

$$\mathbf{k}^2 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}^2 = \mathbf{-1}.$$

Now also notice that

$$\mathbf{ijk} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = \mathbf{-1}.$$

Thus, we have satisfied the Hamilton relation. Now notice that the matrices **i**, **j** and **k** behave the same as the numerical **i**, **j** and **k** so if we drop the **1** from the matrix form it is indistinguishable from its numerical form. Thus, we can talk about quaternions as numbers and matrices interchangeably.

Now notice in the matrix form for any quaternion q,

$$\det(q) = (a + ib)(a - ib) + (-c - id)(c - id) = a^2 + b^2 + c^2 + d^2$$

which is the distance of (a, b, c, d) from the origin. This means that as long as $q \neq 0$ than any q has an inverse.

Now we want to define some properties of quaternions:

Definition 1. *The **real** part of any quaternion $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ is a.*

Definition 2. *The* ***imaginary*** *part of any quaternion* $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ *is the sum* $b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$*.*

Definition 3. *A quaternion* p *is* ***pure imaginary*** *if it has no real part. The space of all pure imaginary quaternions is all quaternions of the form* $p = b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$*.*

Definition 4. *Given any quaternion* $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ *we define the* ***quaternion conjugate*** *of* q*, denoted* $\overline{q}$*, as* $\overline{q} = a - b\boldsymbol{i} - c\boldsymbol{j} - d\boldsymbol{k}$*.*

Now we want to use the matrix form of any quaternion q to find q^{-1}. Again let

$$q = \begin{bmatrix} a+id & -b-ic \\ b-ic & a-id \end{bmatrix}$$

we can use the inverse formula for any two 2×2 matrix to find

$$q^{-1} = \frac{1}{\det(q)} \begin{bmatrix} a-id & b+ic \\ -b+ic & a+id \end{bmatrix}$$

or using our basis vectors,

$$q^{-1} = \frac{1}{\det(q)}(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) = \frac{1}{\det(q)}\overline{q}.$$

In this way for any quaternion q, $q\overline{q} = \overline{q}q = \det(q) = a^2 + b^2 + c^2 + d^2$. This is the same behavior as complex numbers under conjugation.

Now we want to outline a theorem involving quaternions and rotations. Though it will not be proven in this paper a complete proof may be found in [1]. We mention this now because it is used later in section 7.

Theorem 1. *Any rotation of a quaternion* q *in* $\mathbb{H}$ *can be written as*

$$q \to pq\overline{w}$$

where p *and* w *are both quaternions.*

Definition 5. *A unit quaternion is any quaternion* q *such that* $|q| = 1$ *or* $\det(q) = 1$*.*

Definition 6. *A* ***pure imaginary unit quaternion*** *is any unit quaternion that happens to be pure imaginary.*

Let us discuss pure imaginary quaternions a bit more. We know that any pure imaginary quaternion u has the form $u = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. This is strikingly similar to a vector in $\mathbb{R}^3$. For this, just as we can associate a quaternion to vector in $\mathbb{R}^4$, we can associate a pure imaginary quaternion to a vector in $\mathbb{R}^3$! The space of pure imaginary unit quaternions can therefore be thought of as points on the unit sphere in $\mathbb{R}^3$. Furthermore, any axis in $\mathbb{R}^3$ through the origin can be characterized by a point on the unit sphere and therefore also as a pure imaginary unit quaternion.

Theorem 2. *Any **unit quaternion**, q can be written in the form*

$$q = \cos(\theta) + u\sin(\theta)$$

for some pure imaginary unit quaternion u and real θ. Moreover, any quaternion of this form is unit.

Theorem 3. *Given any vector, $\boldsymbol{v}$ and axis in the direction of unit vector $\boldsymbol{u}$ in $\mathbb{R}^3$ the rotation of $\boldsymbol{v}$ by some angle θ around $\boldsymbol{u}$ is the vector represented by the pure imaginary quaternion p where*

$$p = (\cos(\theta) + u\sin(\theta))v(\cos(\theta) - u\sin(\theta))$$

and u and v are the pure imaginary quaternion counterparts to $\boldsymbol{u}$ and $\boldsymbol{v}$.

In this way we can characterize any rotation of $\mathbb{R}^3$ as unit quaternions. Notice however that if we are given a rotation of a pure imaginary quaternion u characterized by the unit quaternion q then the resulting pure imaginary quaternion p is given by

$$p = quq^{-1} = (-q)u(-q^{-1}).$$

So for any one rotation in $\mathbb{R}^3$ there are two corresponding anti-podial unit quaternions. This will come up again in section 7.

The algebra of quaternions was developed by Hamilton the mid 1800's. The space of quaternions is therefore denoted $\mathbb{H}$ in his honor.

4 Transformations and General Geometric Groups

4.1 Linear Transformations

Recall that any *linear transformation*, T on some n–dimensional field $\mathbb{F}^n$ can be written as

$$T(\overline{x}) = A\overline{x}$$

where A is a $n \times n$ matrix whose entries are elements of $\mathbb{F}$ and $\overline{x}$ is some vector in $\mathbb{F}^n$. In this way every linear transformation only maps sets to themselves so $T(\overline{x})$ is a vector in $\mathbb{F}^n$. This paper is primarily concerned with linear transformations over the fields $\mathbb{R}^n$, $\mathbb{C}^n$ and $\mathbb{H}^n$.

Definition 1. *Any linear transformation T is **length preserving** if T preserves the inner product, that is, a transformation is length preserving if*

$$T(\boldsymbol{u}) \cdot T(\boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{v}.$$

Definition 2. *A Matrix Linear Transformation, $T(\boldsymbol{x}) = A\boldsymbol{x}$ is **orientation preserving** if $\det A = 1$ and **non-orientation preserving** if $\det A \neq 1$.*

For the context of this paper we are concerned with witch length preserving transformations are rotations. Formally we will give use the definition given by [1] which says the following:

Definition 3. *A **rotation** is a linear transformation that preserves length and orientation.*

Our goal in examining transformations is to classify how objects like rotations are structured. For the remainder of this paper, a group of transformations over a field $\mathbb{F}^n$ with a given property will refer to the set of $n \times n$ matrices, A, such that the linear transformation $T(\overline{x}) = A\overline{x}$ has that property.

4.2 Geometric Matrix Groups

This paper will be examine the following sets:

Name	Symbol	Description
Orthogonal Group	$O(n)$	The set of all isometrics of $\mathbb{R}^n$ leaving the origin fixed
Special Orthogonal Group	$SO(n)$	The set of rotations of $\mathbb{R}^n$
Unitary Group	$O(n)$	The set of all isometrics of $\mathbb{C}^n$ leaving the origin fixed
Special Unitary Group	$SO(n)$	The set of rotations of $\mathbb{C}^n$
Symplectic Group	$Sp(n)$	The Set of all rotations of $\mathbb{H}^n$

We will call this collection of sets the *general geometric groups* and note that each is a *matrix Lie group* (In the next section we will prove that each is actually a group under matrix multiplication). Now let us examine the two conditions on the orthogonal and unitary groups. In general an isometry is not necessarily a linear transformation. For our purposes however, we only want to classify linear transformations as isometries. Since any transformation $T(\overline{x}) = A\overline{x}$ maps $\mathbf{0}$, the zero vector, to itself we get for free that elements in $O(n)$ and $U(n)$ leave the origin fixed.

Now we want to note two "obvious" isomorphisms between some of the geometric groups. The first is $SO(2) \approx SU(1)$. This makes sense because $\mathbb{R}^2$ and $\mathbb{C}$ are both a two dimensional Cartesian plots with differently tittled axes. The second is $SU(2) \approx Sp(1)$. This one is clear from our definition of a quaternion. We defined the matrix form of any quaternion q as the 2×2 complex matrix

$$q = \begin{bmatrix} \alpha & -\beta \\ \beta^* & \alpha^* \end{bmatrix}.$$

Since each matrix can be thought of as the ordered pair of complex numbers (α, β) we expect the rotations of $\mathbb{H}$ to behave like $\mathbb{C}^2$.

4.3 Defining $SO(2)$

To become more familiar with the special unitary group let us find an explicit definition of $SO(2)$. Remember that $SO(2)$ is the set of all rotations on $\mathbb{R}^2$. If

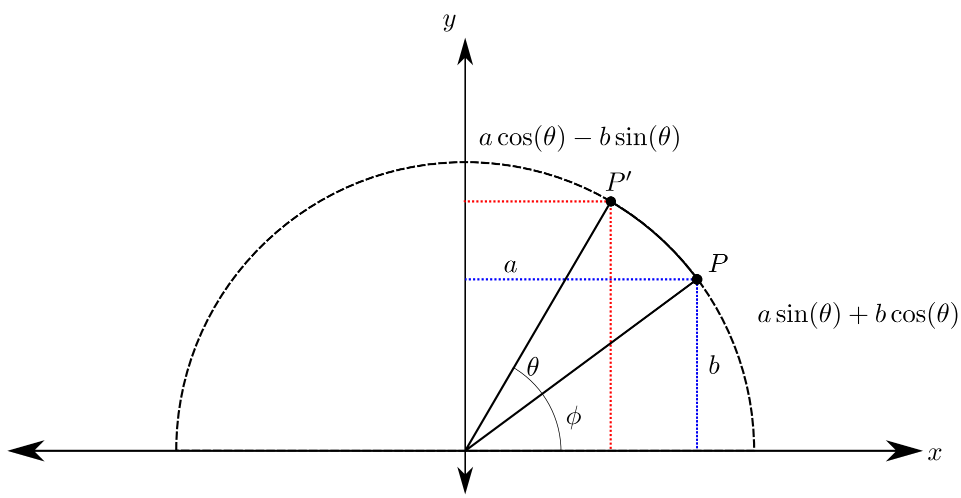

Figure 2: A point P before and after a rotation. We know that P' is at the point $(\cos(\theta+\phi), \sin(\theta+\phi))$. Using a trigonometric identity we know that P' is at the point $(\cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi), \sin(\theta)\cos(\phi) + \sin(\phi)\cos(\theta))$ but $\cos(\phi) = a$ and $\sin(\phi) = b$. Thus, P' is at the point $(a\cos(\theta) - b\sin(\phi), a\sin(\theta) + b\sin(\phi))$.

we begin with some point P at (a, b) in $\mathbb{R}^2$ and we rotate it through an angle θ then our new point P' is at the position

$$(a', b') = (a\cos(\theta) - b\sin(\theta), a\sin(\theta) + b\cos(\theta)).$$

A short proof of this is illustrated in the figure above. We can therefore define the linear transformation of this rotation as

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = A \begin{bmatrix} a \\ b \end{bmatrix}$$

where

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Thus,

$$SO(2) = \left\{ \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : 0 \leq \theta \leq 2\pi \right\}.$$

Now notice that each matrix element of $SO(2)$ can be viewed as the matrix representation of the complex number $\cos(\theta) + i\sin(\theta)$. This provides an avenue for us to visualize $SO(2)$. If we treat each element as the ordered pair $(\cos(\theta), \sin(\theta))$ we can plot the result in $\mathbb{R}^2$ (Figure 3). Thus, we can view the structure of $SO(2)$ in $\mathbb{R}^2$ as the unit circle!

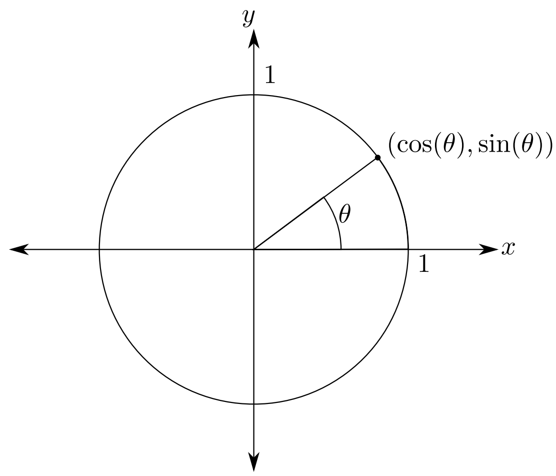

Figure 3: Graphical representation of $SO(2)$

5 Conditions for Matrix Elements of General Geometric Groups

5.1 $SO(n)$ and $O(n)$

Definition 1. *Given two real vectors* $\boldsymbol{u} = (u_1, u_2, ..., u_n)$ *and* $\boldsymbol{v} = (v_1, v_2, ..., v_n)$ *in* $\mathbb{R}^n$ *the* ***inner product*** $\boldsymbol{u} \cdot \boldsymbol{v}$ *is defined by*

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

Theorem 1. *Let* A *be some* $n \times n$ *real matrix corresponding to the linear transformation* $T(\boldsymbol{x}) = A\boldsymbol{x}$. *We know that* $A \in O(n)$ *if only if* $AA^T = \boldsymbol{1}$ *and* $A \in SO(n)$ *if and only if* $\det(A) = 1$ *in addition to the first condition.*

Proof. Let's assume first that $AA^T = \mathbf{1}$ We want to show that T is distance preserving. We first need to notice that for each column and row,

$$AA^T = \mathbf{1} \rightarrow (i^{th}\text{row of } A)(j^{th}\text{column of } A^T) = \delta_{ij}.$$

Hence, the rows of A form an orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ in $\mathbb{R}^n$. We can therefore rewrite A as

$$A = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}$$

Let's examine two arbitrary basis vectors $\mathbf{b}_i = (b_{i_1}, b_{i_2}, ..., b_{i_n})$ and $\mathbf{b}_j = (b_{j_1}, b_{j_2}, ..., b_{j_n})$.

If we transform $\mathbf{b}_i$ and $\mathbf{b}_j$ using our transformation, notice that

$$A\mathbf{b}_i \cdot A\mathbf{b}_j = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix} \mathbf{b}_i \cdot \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{bmatrix} \mathbf{b}_j \tag{1}$$

$$= \begin{bmatrix} \mathbf{b}_1 \cdot \mathbf{b}_i \\ \mathbf{b}_2 \cdot \mathbf{b}_i \\ \vdots \\ \mathbf{b}_n \cdot \mathbf{b}_i \end{bmatrix} \cdot \begin{bmatrix} \mathbf{b}_1 \cdot \mathbf{b}_j \\ \mathbf{b}_2 \cdot \mathbf{b}_j \\ \vdots \\ \mathbf{b}_n \cdot \mathbf{b}_j \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ \mathbf{b}_i \cdot \mathbf{b}_i \\ \vdots \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ \mathbf{b}_j \cdot \mathbf{b}_j \\ \vdots \\ 0 \end{bmatrix} \tag{3}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1_j \\ \vdots \\ 0 \end{bmatrix} = \delta_{ij} = \mathbf{b}_i \cdot \mathbf{b}_j \tag{4}$$

where steps (3) and (4) take advantage of the orthonormal properties of $\mathbf{b}_i$ and $\mathbf{b}_j$. Now since any vector in $\mathbb{R}^n$ can be written as a linear combination of basis vectors, the resulting finding proves that, under the matrix transformation $T(\mathbf{x}) = A\mathbf{x}$, the inner product is preserved. Hence A is a transformation that preserves length and $A \in O(n)$. If we further assume that $\det(A) = 1$ than we get for free that $A \in SO(n)$.

Now let's assume that $A \in O(n)$: We know that A must preserve inner product so for any two vectors in $\mathbb{R}^n$,

$$A(\mathbf{u}) \cdot A(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}.$$

This means, the image of any orthonormal basis in $\mathbb{R}^n$ under a transformation by A is also an orthonormal basis of $\mathbb{R}^n$. Hence, the columns of A form an orthonormal basis which means that the rows of A^T are orthonormal. Thus $A^T A = \mathbf{1}$ so $A^T = A^{-1}$ and $(AA^T) = \mathbf{1}$.

Now notice that by properties of the determinate $\det(AA^T) = \det(A)\det(A^T) = \det(A)^2 = 1$. So $\det(A) = \pm 1$. If T is orientation preserving than $\det A = 1$. □

Now we want to make sure that our previous definition of $SO(2)$ matches with our new generalized definition:

To begin Let A be an element of $SO(2)$. We know, by our definition of general geometric groups, that A must be a 2×2 real matrix so we will begin by letting

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Now we apply our first condition; we know A is length preserving if $AA^T = \mathbf{1}$. Thus $A^{-1} = A^T$ and by the well known formula for the inverse of any 2×2 matrix we know

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} = A^T.$$

Since A is also orientation preserving we know $\det(A) = 1$ and therefore that:

1. $a = d$
2. $-b = c$
3. $ad - cb = a^2 + b^2 = 1$.

Using the third criteria, if we let $a = \cos(\theta)$ for some arbitrary θ than $b = \sin(\theta)$. Thus, every matrix A in $SO(2)$ has the form

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

This matches our previous definition.

Definition 2. *Let G be some Matrix Lie Group and I some connected interval in $\mathbb{R}$. We say a **path** through G is any function $f : I \to G$ such that $f(t) = A(t)$ where $A(t)$ is a matrix whose entries are all continuous functions of t.*

Definition 3. *A matrix group G is **path connected** if there is a path connecting any two elements of G.*

Theorem 2. *For any n, $SO(n)$ is path connected.*

Proof. It suffices to show that $SO(n)$ is path connected if there is a path between $\mathbf{1}$ and any element of $SO(n)$. This is because if there is a path between $\mathbf{1}$ and A and $\mathbf{1}$ and B than there is also a path between A and B.

The next part will be a proof by mathematical induction.

First examine the case where $n = 2$. Notice that $SO(2)$ is the unit circle and clearly path connected. Notably for any

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

the function

$$A(t) = \begin{bmatrix} \cos(\theta t) & -\sin(\theta t) \\ \sin(\theta t) & \cos(\theta t) \end{bmatrix}$$

is a path that connects $\mathbf{1}$ to A as t is pushed from 0 to 1.

Now suppose that $SO(n-1)$ is path connected and let $A \in SO(n)$. We want to find a continuous motion that takes the basis vectors of $\mathbb{R}^n$, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ to $\{A\mathbf{e}_1, A\mathbf{e}_2, \ldots, A\mathbf{e}_n\}$ which are the columns of A. This corresponds to a path connecting $\mathbf{1}$ to A in $SO(n)$.

If $\mathbf{e}_1 = A\mathbf{e}_1$ then the overall rotation can take place in $\mathbb{R}^{n-1}$ and thus is path connected by our induction hypothesis.

If $\mathbf{e}_1$ and $A\mathbf{e}_1$ are distinct than they define some plane P. By the path connectedness of $SO(2)$ we can create a path between $\mathbf{e}_1$ and $A\mathbf{e}_1$ within the plane. Now, much like the ladder argument, the rest of the rotation can take place in $\mathbb{R}^{n-1}$ and is path connected due to our induction hypothesis. Hence there is a path between $\mathbf{1}$ and A in $SO(n)$. Thus $SO(n)$ is path connected. □

5.2 $U(n)$ and $SU(n)$

Definition 4. *Given two vectors $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{C}^n$ we define the **inner product** denoted $\boldsymbol{u} \cdot \boldsymbol{v}$ as*

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1\overline{v_1} + u_2\overline{v_2} + \cdots + u_n\overline{v_n}$$

where $\overline{v_i}$ denotes the complex conjugate of v_i.

Definition 5. *Given a complex square matrix A the **conjugate transpose** of A, denoted $A^\dagger$, is the transpose of $\overline{A}$ where $\overline{A}$ is A with every entry replaced by its conjugate.*

Theorem 3. *Let A be some $n \times n$ complex matrix corresponding to the linear transformation $T(\boldsymbol{x}) = A\boldsymbol{x}$ of $\mathbb{C}^n$. We give that the criteria that $A \in U(n)$ if and only if*

$$AA^\dagger = \mathbf{1}.$$

Proof. Let $AA^\dagger = \mathbf{1}$ much like the last section we know that the row of A form an orthonormal basis. Hence if $\mathbf{a}_i$ and $\mathbf{a}_j$ are row vectors of A then $A\mathbf{a}_i \cdot A\mathbf{a}_j = 0$ if $i \neq j$ and 1 if $i = j$ in an argument similar to theorem 1. Hence, $T(\mathbf{x})$ preserves the inner product between vectors in $\mathbb{C}^n$ and A preserves length. Thus $A \in U(n)$.

Now assume $A \in U(n)$. We know that $T(\mathbf{x})$ preserves inner product. This means that if $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthonormal basis of $\mathbb{C}^n$ the set $\{A\mathbf{e}_1, A\mathbf{e}_2, \ldots, A\mathbf{e}_n\}$ is also an orthonormal basis of $\mathbb{C}^n$. Hence, the columns of A are orthonormal and

$$A^\dagger A = \mathbf{1}$$

Hence

$$A^\dagger = A^{-1}$$

and

$$A^\dagger A = \mathbf{1} = AA^\dagger.$$

□

Definition 6. *We say any $n \times n$ complex matrix $A \in U(n)$ is an element of $SU(n)$ if and only if* $\det(A) = 1$

Lemma 1. *The elements of $SU(2)$ are all complex matrices, A of the form:*

$$A = \begin{bmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix}.$$

Proof. To prove this we will consider the following arbitrary 2×2 complex matrix:

$$A = \begin{bmatrix} z_1 & z_2 \\ z_3 & z_4 \end{bmatrix}$$

where each z_i is an arbitrary complex number.

Now, if we force $AA^\dagger = \mathbf{1}$ and $\det A = 1$ we see that $A^{-1} = A^\dagger$. Also, by the inverse formula for any 2×2 matrix,

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} z_4 & -z_2 \\ -z_3 & z_1 \end{bmatrix}$$

but $\det A = 1$ so we know

$$\begin{bmatrix} z_4 & -z_2 \\ -z_3 & z_1 \end{bmatrix} = A^{-1} = A^\dagger = \begin{bmatrix} \overline{z_1} & \overline{z_3} \\ \overline{z_2} & \overline{z_4} \end{bmatrix}.$$

Thus, we know $z_1 = \overline{z_4}$ and $-z_2 = \overline{z_3}$.

Now if we let $z_1 = \alpha$ and $z_2 = \beta$ then we see that our matrix A has the general form

$$A = \begin{bmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix}.$$

Where $|\alpha|^2 + |\beta|^2 = 1$. Thus any element of $SO(2)$ is also the matrix form of a unit quaternion! This completes the proof.

□

Theorem 4. *For any n, $SU(n)$ is path connected.*

Proof. We will once again prove this by induction beginning with the case where $n = 2$.

Again we want to show that there is a path between $\mathbf{1}$ and any Matrix in $SO(2)$.

Let $A \in SU(2)$. We know $\det(A) = 1$ and A is a complex 2×2 matrix of the form

$$A = \begin{bmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{bmatrix}.$$

Furthermore, since $\det A = 1$ we know that $|\alpha|^2 + |\beta|^2 = 1$ so we can rewrite α and β as $\alpha = u \cos\theta$ and $\beta = v \sin\theta$ where $\theta \in \mathbb{R}$ and $u, v \in \mathbb{C}$ such that $|u| = |v| = 1$.

Now notice that u and v are elements of $SU(1) \approx SO(2)$. Thus u and v can be connected to $\mathbf{1}$ with a path. Let's call these paths $u(t)$ and $v(t)$ such that $u(0) = v(0) = \mathbf{1}$, $u(1) = u$ and $v(1) = v$.

Now let's define a the path

$$A(t) = \begin{bmatrix} u(t)\cos(\theta t) & -v(t)\sin(\theta t) \\ \overline{v(t)}\sin(\theta t) & \overline{u(t)}\cos(\theta t) \end{bmatrix}.$$

Notice that $A(0) = \mathbf{1}$ and $A(1) = A$. Now, we want to show that $A(t) \in SU(2)$ for each t. Based on its construction it will suffice to show that $\det(A(t)) = 1$ for all t.

Since $u(t)$ and $v(t)$ are paths in $SU(1)$ we know that $|u(t)| = |v(t)| = 1$ for each t. Hence,

$$\det(A(t)) = |u(t)|^2 \cos^2(\theta t) + |v(t)|^2 \sin^2(\theta t) = 1.$$

Thus $SU(2)$ is path connected.

Now assume that $SU(n-1)$ is path connected. We can now ue an argument that exactly mirrors Theorem 2 that proves $SU(n)$ is path connected. □

5.3 $Sp(n)$

Definition 7. *On the space $\mathbb{H}^n$ the **inner product of two quaternion vectors**, $\boldsymbol{q} = \{q_1, q_2, \ldots, q_n\}$ and $\boldsymbol{p} = \{p_1, p_2, \ldots, q_n\}$, denoted $\boldsymbol{q} \cdot \boldsymbol{v}$ is defined as,*

$$\boldsymbol{q} \cdot \boldsymbol{v} = q_1\overline{p_1} + q_2\overline{p_2} + \cdots + q_n\overline{p_n}.$$

Note that in general $\boldsymbol{q} \cdot \boldsymbol{v} \neq \boldsymbol{p} \cdot \boldsymbol{q}$.

Definition 8. *If A is any $n \times n$ quaternion matrix, $A^\dagger$ represents the transpose of the matrix whose ij^{th} component is the quaternion conjugate of the of ij^{th} component of A.*

Definition 9. *The **complex form** of any matrix in $Sp(n)$ is the $2n \times 2n$ complex matrix in which each quaternion entry is replaced by its 2×2 complex counterpart. If $A \in Sp(n)$ its complex form is denoted $C(A)$. This helps avoid potential problems that arise from the non–commutative nature of quaternions.*

Definition 10. *Let A be some $n \times n$ quaternion matrix. We define **the determinant of** A as*

$$\det(A) = \det(C(A)).$$

Theorem 5. *Let A be some $n \times n$ quaternion matrix corresponding to the linear transformation $T(\boldsymbol{x}) = A\boldsymbol{x}$ of $\mathbb{H}^n$. We give that the criteria that $A \in Sp(n)$ if and only if*

$$AA^\dagger = \mathbf{1}.$$

and

$$\det(C(A)) = 1.$$

Proof. It is trivial that $T(\mathbf{x})$ is orientation preserving if and only if $\det(C(A)) = 1$ so we must prove that $T(\mathbf{x})$ is length preserving if and only if $AA^\dagger = \mathbf{1}$. In order to prove this we want to use the same argument from Theorem 3. We must first, however, take some care to show that A and $A^\dagger$ commute.

Notice that for each column and for each row

$$AA^\dagger = \mathbf{1} \to \mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij}$$

where $\mathbf{b}_i$ denotes the i^{th} row of A. Now, since $\delta_{ij} = \delta_{ji}$ we know that $\mathbf{b}_i \cdot \mathbf{b}_j = \mathbf{b}_j \cdot \mathbf{b}_j$. Thus, $AA^\dagger = A^\dagger A = \mathbf{1}$ and the columns of A form an orthonormal basis of $\mathbb{H}^n$.

Hence by the same argument as Theorem 3, we prove that $T(\mathbf{x})$ preserves length. □

Theorem 6. *The complex form of any $A \in Sp(n)$ is an element of $U(2n)$.*

Proof. Let A be some $n \times n$ quaternion matrix in $Sp(n)$. It follows by block multiplication and properties of $Sp(n)$ that

$$C(A)C(A)^\dagger = C(AA^\dagger) = C(\mathbf{1}) = \mathbf{1'}$$

where $\mathbf{1}$ is the $n \times n$ identity and $\mathbf{1'}$ is the $2n \times 2n$ identity.

Thus $C(A) \in U(2n)$. □

Theorem 7. *The matrix group $Sp(n)$ is path connected.*

Proof. Again we will prove this by induction on n. Since $Sp(1) \approx SU(2)$ we know that $Sp(1)$ is path connected.

Now assume $Sp(n-1)$ is path connected. We can use an argument that exactly mirrors Theorem 2 that shows $Sp(n)$ is path connected. □

Theorem 8. *Each geometric groups is a group under matrix multiplication.*

Proof. Let A^* denote either A^T or $A^\dagger$.

Let G be one of the geometric groups. To begin we know that $\mathbf{1}(\mathbf{1})^T = \mathbf{1}$ so $\mathbf{1} \in G$. Now notice that if $A \in G$ than $AA^* = \mathbf{1}$ and $A^* = A^{-1}$. Thus $A^*(A^*)^* = A^*A = \mathbf{1}$ so $A^{-1} \in G$. Now let A and B be arbitrary elements in G. Notice that by properties of G

$$AB(AB)^* = ABB^*A^* = A\mathbf{1}A^* = \mathbf{1}.$$

Furthermore, $\det(AB) = \det(A)\det(B) = 1$ if $\det(A) = 1 = \det(B)$.

Thus G is closed under matrix multiplication and we know it is a group. □

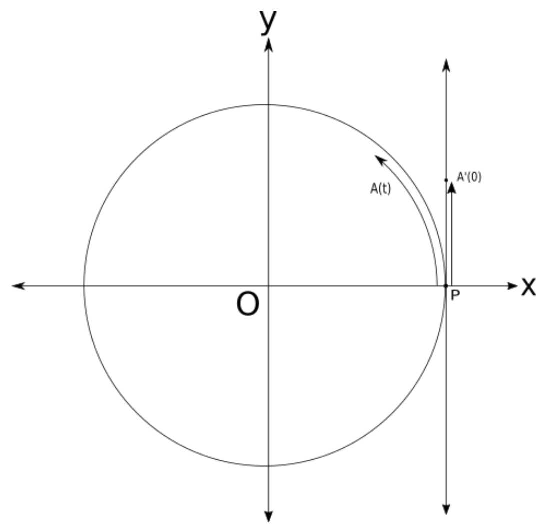

Figure 4: A path through $SO(2)$ and its tangent vector at the identity.

6 Tangent Spaces and Lie Algebras

6.1 Introductions

Definition 1. *Let G be some Matrix Lie Group and $A(t)$ a path through G. we say $A(t)$ is **smooth** if each entry in $A(t)$ is differentiable. The derivative of $A(t)$ at t, denoted $A'(t)$, is the matrix whose ij^{th} element is the time derivative of the ij^{th} element of $A(t)$.*

6.1.1 Tangent Space of $SO(2)$

Let's consider some smooth path, $A(t)$ in $SO(2)$ that goes through $\mathbf{1}$ at $t = 0$ and a point P moving along it. We want to consider the possible velocities of P as it moves through $t = 0$. Easily visualized in Figure 4, we see that P can move at any speed either clockwise or counter–clockwise around the unit circle so the set of all velocities of P is just the line tangent to the circle at $\mathbf{1}$. This line is the *tangent space at unity* of $SO(2)$.

6.1.2 Formal Definition of the Tangent Space

Now we want to introduce a formal definition of the tangent space. To do this let's consider two smooth paths $A(t)$ and $B(t)$ through some matrix Lie group G such that both pass a point P at $t = 0$. Now, we will impart an equivalence relation on any two such paths and say that $A(t)$ and $B(t)$ are in the same *equivalence class* if $A'(0) = B'(0)$. Now, we define **the tangent space of G at P** to be the set of all possible equivalence classes.

This may sound complicated but in essence it is a simple concept. The tangent space of a group G at a point P is just the set of all possible velocities a point could have while moving through P. This becomes a little convoluted, however, because instead of talking about velocity vectors in the traditional sense we are referring to specific matrices. In this way, for any smooth path $A(t)$, we will call $A'(t)$ the *velocity vector* of $A(t)$ at t.

Fundamentally the tangent space of different points in the same group are isomorphic. That is why the rest of this section will focus on **the tangent space at unity** since it is the easiest to describe mathematically. The tangent space at unity is for any matrix Lie group G will be denoted $T_{\mathbf{1}}(G)$

Definition 2. *A set V is a real* ***vector space*** *if for each $v_1, v_2 \in V$*

- $v_1 + v_2 \in V$
- $cv_1 \in V$ *for each* $c \in \mathbb{R}$

Theorem 1. *For any general geometric group G, $T_{\mathbf{1}}(G)$ is a vector space.*

Proof. Let G be a general geometric group and let $X, Y \in T_{\mathbf{1}}(G)$. We know that $X = A'(0)$ and $Y = B'(0)$ for two smooth paths $A(t)$ and $B(t)$ in G through unity at $t = 0$.

For any matrix A let A^* denote either $A^\dagger$ or A^T

We first want to prove that $X + Y \in T_{\mathbf{1}}(G)$. To do this we want to show that $C(T) = A(t)B(t)$ is a smooth path through G. Since G is one of our general linear groups we know that $A(t)(A(t))^* = B(t)(B(t))^* = \mathbf{1}$. Hence by group properties of the geometric groups $C(t) \in G$. Now notice that

$$C'(t) = \frac{d}{dt}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t)$$

and

$$C'(0) = A'(0)B(0) + A(0)B'(0) = X + Y$$

since $A(t)$ and $B(t)$ both pass through $\mathbf{1}$ at $t = 0$. Hence $X + Y \in T_{\mathbf{1}}(G)$.

Now let $c \in \mathbb{R}$ and let $C(t) = A(ct)$. Clearly, $C(t) \in G$ and $C'(t) = \frac{d}{dt}(A(ct)) = cA'(ct)$ so, $C'(0) = cA'(0) = cX \in T_{\mathbf{1}}(G)$.

Thus $T_{\mathbf{1}}(G)$ is a vector space. □

6.1.3 Tangent space of $Sp(1)$ and introduction to Lie Algebras

Now we want to find the tangent space of $Sp(1)$ at unity. First, however we must recall the following outlined earlier in the paper:

1. Any quaternion has the form $a + bi + c\mathbf{j} + d\mathbf{k}$.
2. Any quaternion, p is pure imaginary if it has the form $p = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$,

3. If $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ than its quaternion conjugate, $\overline{q}$, is equal to $a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$ and

4. $q\overline{q} = a^2 + b^2 + c^2 + d^2$.

Now, we have the following theorem:

Theorem 2. *The tangent space at unity of $Sp(1)$ is the space of all pure imaginary quaternions.*

Proof. Let $q(t)$ be an arbitrary path passing through $\mathbf{1}$ at $t = 0$. Notice that since $q(t) \in Sp(1)$ for all t, $q(t)\overline{q(t)} = 1$. Hence by differentiation,

$$\begin{aligned} q(t)\bar{q}(t) = 1 &\to 0 = q'(t)\bar{q}(t) + q(t)\bar{q}'(t) \\ &\to 0 = q'(0) + \bar{q}'(0) && \text{at } t = 0 \\ &\to 0 = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} \\ &\to 0 = 2a \end{aligned}$$

Thus any tangent vector at $\mathbf{1}$ is purely imaginary. □

Now, we know that $T_{\mathbf{1}}(Sp(1))$ is a vector space, however we would like to add some extra structure by bestowing an algebra upon it.

Definition 3. *An **algebra** is a vector space with a bilinear operator. That is, a function that takes any two elements in the vector space and maps them to a third.*

Usually algebras use multiplication as their bilinear operators. This is a problem, however, because the space of pure imaginary quaternions is *not* closed under multiplication. This is easily demonstrated by $i^2 = -\mathbf{1}$. To solve this problem, Sophus Lie introduced the "Lie Bracket" defined as follows:

Definition 4. *Let G be a matrix Lie group. For any two elements $X, Y \in T\mathbf{1}(G)$ we define the **Lie bracket**, denoted $[X, Y]$ to be the following binary operation:*

$$[X, Y] = XY - YX$$

In other words the Lie bracket measures how much two elements commute. We also note that for any three vectors X, Y and Z and any real number c:

- $[X + Y, Z] = XZ + YZ - ZY - ZX = [X, Z] + [Y, Z]$
- $[X, Y + Z] = XZ + XY - ZX - YX = [X, Y] + [X, Z]$
- $[cX, Y] = c[X, Y]$
- $[X, cY] = c[X, Y]$

Now, $T_{\mathbf{1}}(Sp(1))$ has the obvious basis vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$. Now we notice from section 3 that the under the lie bracket operation,

- $[\mathbf{i}, \mathbf{j}] = 2\mathbf{k}$
- $[\mathbf{k}, \mathbf{i}] = 2\mathbf{j}$
- $[\mathbf{j}, \mathbf{k}] = 2\mathbf{i}$.

Now, let $\mathbf{i}' = \mathbf{i}/2$, $\mathbf{j}' = \mathbf{j}/2$ and $\mathbf{k}' = \mathbf{k}/2$ we see that:

- $[\mathbf{i'}, \mathbf{j'}] = \mathbf{k'}$
- $[\mathbf{k'}, \mathbf{i'}] = \mathbf{j'}$
- $[\mathbf{j'}, \mathbf{k'}] = \mathbf{i'}$

Paired with the ladder algebraic properties this demonstrates that the space of pure imaginary quaternions is closed under the Lie bracket operation. We call the $T_{\mathbf{1}}(Sp(1))$ paired with the Lie bracket operation, the *Lie algebra* of $Sp(1)$ and denote it $\mathfrak{sp}(1)$.

Definition 5. *Given some matrix Lie group G, the Lie algebra of G, denoted $\mathfrak{g}$, is $T_{\mathbf{1}}(G)$ under the Lie bracket operation.*

Now we will dive a bit deeper into $\mathfrak{sp}(1)$. Notice that if we treat each of $\mathbf{i}', \mathbf{j}'$ and $\mathbf{k}'$ as unit vectors in the x, y and z directions in $\mathbb{R}^3$, than we see that:

- $[\mathbf{i'}, \mathbf{j'}] = \mathbf{k'} = \mathbf{i'} \times \mathbf{j'}$
- $[\mathbf{k'}, \mathbf{i'}] = \mathbf{j'} = \mathbf{k'} \times \mathbf{i'}$
- $[\mathbf{j'}, \mathbf{k'}] = \mathbf{i'} = \mathbf{j'} \times \mathbf{k'}$

Hence, $\mathfrak{sp}(1)$ is isomorphic to the cross product algebra over $\mathbb{R}^3$.

Now we want to classify the different tangent spaces and Lie algebras of each general geometric groups.

6.2 Tangent Vectors of $O(n)$, $U(n)$ and $Sp(n)$

Theorem 3. *The tangent vectors, X, of $O(n)$, $U(n)$ and $Sp(n)$ at $\mathbf{1}$ are the following:*

1. *For $O(n)$, X is any $n \times n$ real matrix such that $X + X^T = \mathbf{0}$*
2. *For $U(n)$, X is any $n \times n$ complex matrix such that $X + X^\dagger = \mathbf{0}$*
3. *For $Sp(n)$, X is any $n \times n$ quaternion matrix such that $X + X^\dagger = \mathbf{0}$*

Proof. Let us first find the tangent vectors of $O(n)$.

Recall that any matrix $A \in O(n)$ satisfies $AA^T = \mathbf{1}$. Now consider some smooth path $A(t)$ such that $A(0) = \mathbf{1}$. We know that

$$A(t)A(t)^T = \mathbf{1}$$

for all t. Now, if we differentiate this equation we find that because $\frac{d}{dt}(A(t)^T) = (\frac{d}{dt}A(t))^T$

$$A'(t)A(t)^T + A(t)A'(t)^T = \mathbf{0}$$

and at $t = 0$ we know

$$A'(0) + A'(0)^T = \mathbf{0}.$$

Thus any tangent vector $X = A'(0)$ satisfies $X + X^T = \mathbf{0}$

Now consider a smooth path $A(t) \in U(n)$.

First note that if we look at the components of $A(t)$ we can easily see that $\frac{d}{dt}(A(t)^\dagger) = (\frac{d}{dt}A(t))^\dagger$. Since all matrices $A \in U(n)$ satisfy $AA^\dagger = \mathbf{1}$ we can use an argument like the last to show any tangent vector $X = A'(0)$ satisfies $X + X^\dagger = \mathbf{0}$.

Now consider smooth paths $A(t) \in Sp(n)$. We can use a similar proof to the last to show that any tangent vector $X = A'(0)$ satisfies $X + X^\dagger = \mathbf{0}$

□

6.3 Tangent Space and Lie algebra of $SO(n)$

Definition 6. *Given any $n \times n$ matrix, X we define e^X through by polynomial expansion of the function $f(x) = e^x$. That is,*

$$e^X = \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

Theorem 4. *Let A and B be $n \times n$ matrices. If $AB = BA$ then*

$$e^{A+B} = e^A e^B$$

Proof. We will prove this theorem by examining the Taylor expansions of e^{A+B} and $e^A e^B$ about $\mathbf{0}$. Notice that

$$e^{A+B} = \left(\mathbf{1} + \frac{A+B}{1!} + \cdots + \frac{(A+B)^n}{n!} + \dots\right)$$

and

$$e^A e^B = \left(\mathbf{1} + \frac{A}{1!} + \cdots + \frac{(A)^n}{n!} + \dots\right)\left(\mathbf{1} + \frac{B}{1!} + \cdots + \frac{(B)^n}{n!} + \dots\right).$$

We will now prove that the coefficients for $A^l B^m$ is the when both expressions are expanded.

In the first expression we know $A^l B^m$ is found in the larger term

$$\frac{(A+B)^{l+m}}{(l+m)!}.$$

Hence when this is expanded, because A and B commute, the coefficient, $c_{l,m}$, of $A^l B^m$ is

$$c_{l,m} = \frac{{}_{(l+m)}C_l}{(l+m)!} = \frac{(l+m)!}{(l!)(m!)(l+m)!} = \frac{1}{(l!)(m!)}.$$

Now in the second expression we know $c_{l,m}$ is equal to the product of the coefficient of A^l in the expansion of e^A and the coefficient of B^m in the expansion of e^B.
Hence,

$$c_{l,m} = \frac{1}{l!}\frac{1}{m!} = \frac{1}{(l!)(m!)}.$$

Hence the coefficient's for every term $A^l B^m$ are equal between the two expressions and $e^{A+B} = e^A e^B$. □

Theorem 5. *Given some $n \times n$ matrix X that satisfies $X + X^T = \mathbf{0}$, the function $A(t) = e^{tX} \in SO(n)$ for each t.*

Proof. Notice that by looking at the elements of the expansion, it is clear that for each t

$$(e^X)^T = e^{X^T}$$

since $(X^n + X^m)^T = (X^T)^n + (X^T)^m$ for each m and n. Furthermore, we know that $X^T = -X$ which implies

$$XX^T = X(-X) = -XX = X^T X.$$

Therefore X and X^T commute so

$$e^X (e^X)^T = e^X e^{X^T} = e^{X+X^T} = e^{\mathbf{0}} = \mathbf{1}.$$

Hence $e^X \in O(n)$. Now define the function $A(t) = e^{tX}$. Notice that each tX satisfies $(tX) + (tX)^T = t(X + X^T) = \mathbf{0}$. Thus, $A(t) \in O(n)$ for each t.

Now notice that $A(t)$ is continuous and therefore $\det(A(t))$ is also continuous. Now, we know that at $A(0) = e^{(0)X} = \mathbf{1}$. Hence $\det(A(0)) = 1$. Now, because $\det(A(t)) = 1$ or -1 for every t it follows that $\det(A(t)) = 1$ for all t. Hence for each t, $A(t) \in SO(n)$.

In addition to this, notice that we have also proven that $T_{\mathbf{1}}(O(n)) = T_{\mathbf{1}}(SO(n))$ for all n. □

Theorem 6. *The tangent space of $SO(n)$ consists of precisely the $n \times n$ real vectors X that satisfy $X + X^T = \mathbf{0}$.*

Proof. We have already shown that any smooth path through unity has a tangent vector, X at $\mathbf{1}$ that satisfies $X + X^T = \mathbf{0}$.

Now consider a matrix, X that satisfies $X + X^T = \mathbf{0}$. We want to prove that X is a tangent vector to some smooth path.

We conjecture that this path could be $A(t) = e^{tX}$. We already know that $A(t) \in SO(n)$. Now notice that if we, again, look at each component in the expansion we know

$$\frac{d}{dt}e^{tX} = Xe^{tX}.$$

Thus $A'(0) = Xe^{(0)X} = Xe^0 = X\mathbf{1}$ so $A(t) = e^{tX}$ is a path whose tangent vector at $\mathbf{1}$ is X. This concludes the proof. □

Theorem 7. *The tangent space of $SO(n)$, denoted $\mathfrak{so}(n)$, is a Lie Algebra.*

Proof. To prove this we must only prove that $\mathfrak{so}(n)$ is closed under the Lie Bracket operation. Let A and B be two vectors in $\mathfrak{so}(n)$. Recall that $A^T = -A$ and $B^T = -B$ Notice that

$$([A, B])^T = (AB - BA)^T = B^T A^T - A^T B^T = BA - AB = [B, A].$$

Hence,

$$[A, B] + ([A, B])^T = [A, B] + [B, A] = \mathbf{0}$$

and $[A, B] \in \mathfrak{so}(n)$! □

6.4 Tangent Space and Lie algebras of $U(n)$, $SU(n)$ and $Sp(n)$

Theorem 8. *The Tangent sapce of $U(n)$ consists of all the $n \times n$ complex matrices satisfying $X + X^\dagger = \mathbf{0}$. The Tangent sapce of $Sp(n)$ consists of all the $n \times n$ quaternion matrices satisfying $X + X^\dagger = \mathbf{0}$.*

Proof. From the first section we know that any smooth path through unity in $U(n)$ ($Sp(n)$) at has a tangent vector, X, at $\mathbf{1}$ that satisfies $X + X^\dagger = \mathbf{0}$.

Now let X be a $n \times n$ complex (quaternion) matrix that satisfies $X + X^\dagger = \mathbf{0}$ we want to find a path through $U(n)$ ($Sp(n)$) whose tangent vector at unity is X. We conjecture that this path is $A(t) = e^{tX}$. Again we notice that $X^\dagger = -X$ so

$$XX^\dagger = X(-X) = -XX = X^\dagger X.$$

Therefore, because $(e^X)^\dagger = e^{X^\dagger}$ we know for each value of t

$$e^{tX}(e^{tX})^\dagger = e^{tX}e^{tX^\dagger} = e^{t(X+X^\dagger)} = e^{\mathbf{0}} = \mathbf{1}$$

Hence $A(t) \in U(n)$ ($Sp(n)$). Now we know that

$$A'(t) = \frac{d}{dt}e^{tX} = Xe^{tX}$$

so $A'(0) = X$. Since $A(0) = \mathbf{1}$, this means that X is the tangent vector of $A(t)$ at unity! Hence, the tangent space of $U(n)$ ($Sp(n)$) is all $n \times n$ matrices, X that satisfy $X + X^T = \mathbf{0}$.

□

We can now use the exact same proof as theorem 7 to show that the tangent spaces of $O(n)$ and $Sp(n)$ are closed under the Lie bracket operation just replacing by transpose operation with the transpose conjugate. Thus, $T_{\mathbf{1}}(U(n))$ and $T_{\mathbf{1}}(Sp(n))$ are the Lie algebras $\mathfrak{u}(n)$ and $\mathfrak{sp}(n)$ respectively.

Theorem 9. *For any upper triangular $n \times n$ complex matrix T,*

$$\det(e^T) = e^{Tr(T)}$$

where $Tr(T)$ is the trace function of T.

Proof. Let T be some upper triangular complex $n \times n$ matrix. We know that T has the following form:

$$\begin{pmatrix} t_{11} & * & * & \dots & * \\ 0 & t_{22} & * & \dots & * \\ 0 & 0 & t_{33} & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t_{nn} \end{pmatrix}$$

From this we can see that T^2 is upper triangular with i^{th} diagonal entry t_{ii}^2 and similarly and T^m is upper triangular with i^{th} diagonal entry t_{ii}^m. It therefore follows that, e^T is upper diagonal with i^{th} diagonal entry $e^{t_{ii}}$.

It now follows that

$$\det(e^T) = e^{t_{11}} e^{t_{22}} \dots e^{t_{nn}} = e^{Tr(T)}$$

since the determinant of any triangular matrix is the product of its diagonal. This concludes the proof. □

Theorem 10. *For any square complex matrix A,*

$$\det(e^A) = e^{Tr(A)}$$

Proof. Recall that any complex matrix, A there is an invertible matrix B and an upper triangular matrix T such that $A = BTB^{-1}$. Hence,

$$A^n = (BTB^{-1})^n = BT^nB^{-1}.$$

Now,

$$e^A = \sum_{m \geq 0} \frac{A^m}{m!} = B \left(\sum_{m \geq 0} \frac{T^m}{m!} \right) B^{-1} =$$

and by properties of the determinant ,

$$\begin{aligned}
\det(e^A) &= \det\left[B\left(\Sigma_{m\geq 0}\frac{T^m}{m!}\right)B^{-1}\right] \\
&= \det(B)\det\left(\Sigma_{m\geq 0}\frac{T^m}{m!}\right)\det(B^{-1}) \\
&= \det(B)\det(B^{-1})\left[\Sigma_{m\geq 0}\frac{\det(T^m)}{m!}\right] \\
&= \det(BB^{-1})\det\left(\Sigma_{m\geq 0}\frac{T^m}{m!}\right) \\
&= \det(e^T) \\
&= e^{Tr(T)}
\end{aligned}$$

But by properties of the trace we know

$$Tr(A) = Tr(BTB^{-1}) = Tr(B^{-1}BT) = Tr(T).$$

Hence $\det(e^A) = e^{Tr(A)}$. This concludes the proof. □

Theorem 11. *Let G be some matrix Lie group. If $X \in T_{\mathbf{1}}(G)$ than $e^X \in G$.*

Unfortunately this theorem is beyond the scope of this paper so we will leave it without proof. A complete proof of it, however can be found in [4].

Theorem 12. *The tangent space os $SU(n)$ consists of all complex $n \times n$ matrices, X such that $X + X^T = \mathbf{0}$ and $Tr(X) = 0$.*

Proof. Let X be a $n \times n$ complex matrix such that $X + X^T = \mathbf{0}$ and $Tr(X) = 0$. Now let $A(t) = e^{Xt}$. We know that X and X^T commute so

$$A(t)A(t)^T = e^{Xt}e^{X^T t} = e^{(X+X^T)t} = \mathbf{1}.$$

Furthermore under our trace condition,

$$\det(A(t)) = \det(e^{Xt}) = e^{Tr(Xt)} = e^{Tr(X)t} = 1.$$

Thus, $A(t) \in SU(n)$. Now notice that $A'(t) = Xe^{Xt}$ so, $A'(0) = X$ and $X \in T_{\mathbf{1}}(SU(n))$.

Now assume $X \in T_{\mathbf{1}}(SU(n))$. We know that $SU(n) \subset U(n)$ so $X + X^T$. Furthermore we know that $e^X \in SU(n)$. Thus, $\det(e^X) = e^{Tr(X)} = 1$ and $Tr(X) = 0$. This concludes the proof. □

6.5 Dimensions of Tangent Spaces

Since each of $\mathfrak{o}(n), \mathfrak{so}(n), \mathfrak{u}(n), \mathfrak{su}(n)$ and $\mathfrak{sp}(n)$ are vector spaces it makes sense to talk about the dimensionality of each Lie algebra; that is the amount of independent free parameters we can plug into a matrix in each tangent space. Now, because we have shown that $\mathfrak{o}(n) = \mathfrak{so}(n)$ we have the following theorem:

Theorem 13. *Let $D(\mathfrak{g})$ denote the dimension of the Lie algebra for some matrix Lie group G. The Dimension of the Lie algebras associated with each general geometric group is as follows:*

1. $D(\mathfrak{so}(n)) = \frac{1}{2}n(n-1)$
2. $D(\mathfrak{u}(n)) = n^2$
3. $D(\mathfrak{su}(n)) = n^2 - 1$
4. $D(\mathfrak{sp}(n)) = n(2n+1)$

Proof. 1. We know that each element in $\mathfrak{so}(n)$ is an $n \times n$ real matrix, X that satisfies $X + X^T = \mathbf{0}$. Now let a_{ij} denote the entry in the i^{th} row and j^{th} column of X. From the ladder relation we know that each diagonal entry of X is 0 and each entry above the diagonal are the negatives of those below. That is, for each i and j, $a_{ii} = 0$ and $a_{ij} = -a_{ji}$. Thus we are only free to choose entries above the diagonal. In an $n \times n$ matrix there are

$$1 + 2 + \cdots + n - 1 = \frac{1}{2}n(n-1)$$

entries above the diagonal. Hence $D(\mathfrak{so}(n)) = \frac{1}{2}n(n-1)$.

2. We know that $\mathfrak{u}(n)$ is comprised of all complex $n \times n$ matrices X such that $X + X^\dagger = \mathbf{0}$. Thus, like the last arguement we are free to choose the $n(n-1)/2$ entries above the diagonal. In addition to that, however, we are also free to put pure imaginary entries along the diagonal. Thus, because each free complex entry gives two degrees of freedom,

$$D(\mathfrak{u}(n)) = 2(n(n-1))/2 + n = n^2.$$

3. We know that any element, X of $\mathfrak{su}(n)$ is also an element of $\mathfrak{u}(n)$ under the condition that $Tr(X) = 0$. Thus, we lose a degree of freedom between $\mathfrak{u}(n)$ because we do not have the freedom to pick the last element along the diagonal of X. Hence, $D(\mathfrak{so}(n)) = D(\mathfrak{u}(n)) - 1 = n^2 - 1$.

4. We know that $\mathfrak{sp}(n)$ is comprised of all quaternion matrices X such that $X + X^\dagger = \mathbf{0}$. Thus, like 2., we know that we are free to choose any quaternion we like for the $n(n-1)/2$ entries above the diagonal. Furthermore we know the entries along the diagonal are pure imaginary so we find that

$$D(\mathfrak{sp}(n)) = 4(\frac{1}{2}n(n-1)) + 3(n) = n(2n+1).$$

This completes the proof. □

It is natural to suspect that the dimension of a Matrix Lie Group is the same as its tangent space. To prove this, however, requires mathematical machinery that is beyond the scope of this paper. For parties interested, one can prove it by constructing a *homeomorphism* between the tangent space near $\mathbf{0}$ and the group near $\mathbf{1}$. Most Lie theory texts, however, circumvent this issue by simply defining the dimension of a Lie group to be the dimension of its Lie algebra.

6.6 Non-Geometric Lie Goups and Their Lie Algebras

Now that we are familiar with the Lie algebra of each general geometric group, it is time to introduce , $GL(n, \mathbb{C})$. and $SL(n, \mathbb{C})$.

6.6.1 $GL(n, \mathbb{C})$ and $\mathfrak{gl}(n, \mathbb{C})$

The **general linear group** $GL(n, \mathbb{C})$ is the group of all invertible complex $n \times n$ matrices or, more succinctly,

$$GL(n, \mathbb{C}) = \{X \in M_n(\mathbb{C}) | \det(X) \neq 0\}$$

where $M_n(\mathbb{C})$ is the set of all $n \times n$ complex matrices.

Theorem 14. *The tangent space of $GL(n, \mathbb{C})$ at unity is $M_n(\mathbb{C})$*

Proof. We know for free that the tangent space of $GL(n, \mathbb{C})$ is a subset of $M_n(\mathbb{C})$.

Now let $X \in M_n(\mathbb{C})$ and let $A(t) = e^{Xt}$ be some smooth path. Notice that since X and $-X$ commute we know

$$e^{Xt}e^{-Xt} = e^{(X-X)t} = \mathbf{1}$$

so $A(t)$ has an inverse for each t. Thus $A(t)$ is a path through $GL(n, \mathbb{C})$. Now from previous proofs we know that $A'(t) = Xe^{Xt}$ and $A'(0) = X$ hence $X \in T_{\mathbf{1}}(GL(n, \mathbb{C}))$ and $T_{\mathbf{1}}(GL(n, \mathbb{C})) = M_n(\mathbb{C})$. ☐

Now, by properties of square matrices we know automatically that $M_n(\mathbb{C})$ is closed under Lie brackets. Thus, $T_{\mathbf{1}}(GL(n, \mathbb{C}))$ is a Lie algebra that we will denote $\mathfrak{gl}(n, \mathbb{C})$.

6.6.2 $SL(n, \mathbb{C})$ and $\mathfrak{sl}(n, \mathbb{C})$

The *special linear group*, $SL(n, \mathbb{C})$ is the group of invertible $n \times n$ matrices with determinant 1.

Theorem 15. *The tangent space of $SL(n, \mathbb{C})$ at unity is the set*

$$\{X \in M_n(\mathbb{C}) | Tr(X) = 0\}$$

Proof. Let X be a trace zero complex $n \times n$ matrix. Notice that, similar to the last proof, function $A(t) = e^{Xt} \in GL(n, \mathbb{C})$ but

$$\det(A(t)) = \det(e^{Xt}) = e^{Tr(Xt)} = e^{Tr(X)t} = 1$$

so $A(t) \in SL(n, \mathbb{C})$. Furthermore, we know from the last proof that $A'(0) = X$ so $X \in T_{\mathbf{1}}(SL(n, \mathbb{C}))$.

Now let $X \in T_{\mathbf{1}}(SL(n, \mathbb{C}))$. We know that $e^X \in SL(n, \mathbb{C})$ by a previous theorem. Using this result, however, we know that $\det(e^X) = e^{Tr(X)} = 1$ so $Tr(X) = 0$. This concludes the proof. ☐

The Lie algebras generated by the general and special linear groups are relatively easy to deal with and provide a good example of complicated Lie algebras that can be constructed from $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$. This will be discussed in the next chapter.

7 Lie Algebra Structure

This chapter is dedicated to diving deeper into the actual structure of a Lie algebra and the Lie Bracket to focus in on what is meant when an algebra is considered *simple*. To begin we will recall some Algebra definitions and then dive into the important theorems behind some Lie algebras.

7.1 Algebra Topics

The idea behind a *simple* Algebras are deeply rooted in the concepts of *normal subgroups* and *ideals*.

Very briefly recall the following definitions:

Definition 1. *Given a group G a normal subgroup of G is any subgroup H of G such that for all $h \in H$ and all $g \in G$, $ghg^{-1} \in H$.*

A special note that this is usually presented as a theorem, however for the purposes of this paper we will treat it as a definition.

Definition 2. *Let R be any ring that forms a group under the operation $\times$. An ideal of R is any subgroup I such that for each $i \in I$ and $r \in R$ we know $i \times r \in I$.*

7.2 Normal Subgroups, Ideals and Simplicity

Theorem 1. *If $A(t)$ is a smooth path through a matrix group G such that $A(0) = \mathbf{1}$, then $\frac{d}{dt}(A(0)^{-1}) = -A'(0)$.*

Proof. First notice that if $A(0) = \mathbf{1}$ then $A(0)^{-1} = \mathbf{1}$. Now, let $A(t)^{-1} = M(t)$. Notice that $\mathbf{1} = M(t)A(t)$. By differentiating this equation and using the product rule, we find that $\mathbf{0} = M'(t)A(t) + M(t)A'(t)$. Thus, $-M'(t)A(t) = A'(t)M(t)$ and $-M'(t)M(t)^{-1} = A'(t)A(t)^{-1}$. So, $-M'(0)M(0)^{-1} = A'(0)A(0)^{-1}$ and $M'(0) = \frac{d}{dt}(A(0)^{-1}) = -A'(0)$. □

Definition 3. *An ideal $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a subspace of $\mathfrak{g}$ closed under Lie brackets with arbitrary members of $\mathfrak{g}$. That is, if $Y \in \mathfrak{h}$ and $X \in \mathfrak{g}$ then $[X, Y] \in \mathfrak{h}$.*

Theorem 2. *If H is a normal subgroup of a matrix Lie group G, then $T_{\mathbf{1}}(H)$ is an ideal of the Lie algebra $T_{\mathbf{1}}(G)$.*

Proof. To start we know that $T_{\mathbf{1}}(H)$, like all tangent spaces, is a vector space. Furthermore, because any vector tangent to H at $\mathbf{1}$ is tangent to G at $\mathbf{1}$ we know $T_{\mathbf{1}}(H)$ is a subset of $T_{\mathbf{1}}(G)$. So it remains to show that for any $X \in T_{\mathbf{1}}(G)$ and $Y \in T_{\mathbf{1}}(H)$, $[X, Y] \in T_{\mathbf{1}}(H)$.

Now let $A(s)$ and $B(t)$ be arbitrary smooth paths in G and H respectively such that $A(0) = \mathbf{1} = B(0)$. So, $A'(0) = X \in T_{\mathbf{1}}(G)$ and $B'(0) = Y \in T_{\mathbf{1}}(H)$ where X and Y can be arbitrarily picked based on the choice of $A(s)$ and $B(t)$.

Now, for fixed s let's define the path

$$C_s(t) = A(s)B(t)A(s)^{-1}$$

By properties of an ideal subgroup we know that for any value of t or s, $C_s(t)$ is a smooth path in H. Now let's define a new function

$$D(s) = C_s(0) = A(s)YA(s)^{-1}.$$

It follows from the previous comment that $D(s)$ is a smooth path in H. Thus, by a result from theorem 1,

$$\begin{aligned} D'(s) &= A'(s)YA(s)^{-1} + A(s)Y(A)\frac{d}{ds}(A(t)^{-1}) \\ &= A'(s)YA(s)^{-1} - A(s)Y(A)A'(s) \end{aligned}$$

and at $s = 0$, $D'(0) = XY - YX = [X, Y] \in T_{\mathbf{1}}(H)$. Hence $T_{\mathbf{1}}(H)$ is an ideal of $T_{\mathbf{1}}(G)$.

□

Definition 4. *A Lie algebra is* ***simple*** *if it contains no ideals other than itself and* $\{\mathbf{0}\}$

This is all the machinery we need to begin our discussions of Lie algebras. We will spend the remainder of the section showing the *simplicity* or *non–simplicity* of specific groups.

Theorem 3. *The Lie algebra* $\mathfrak{so}(3)$ *is simple.*

Recall from section 3 that $SO(3)$, $\mathbb{R}^3$, is also the set of antipodial points of $Sp(1)$ (the set of unit quaternions). We know that the tangent space at unity for $SO(3)$ is the same as $Sp(1)$. Thus, we know the Lie algebra is isomorphic to the cross product algebra on $\mathbb{R}^3$. Therefore to prove that $\mathfrak{so}(3)$ is simple it will suffice to prove the following theorem.

Theorem 4. *The cross product algebra is simple.*

Proof. For the purposes of this proof, $\mathfrak{R}^3$ will denote the cross product algebra over $\mathbb{R}^3$

Let $\mathfrak{J}$ be a non-zero ideal of the cross product algebra on $\mathfrak{R}^3$. We will prove that $\mathfrak{J} = \mathfrak{R}^3$.

We know that there is some non-zero $u = x\mathbf{i} + y\mathbf{y} + z\mathbf{j} \in \mathfrak{J}$. Without loss of generality we will assume that $x \neq 0$. Now notice,

$$u \times \mathbf{j} = x\mathbf{k} - z\mathbf{i} \in \mathfrak{J}$$

so

$$(x\mathbf{k} - z\mathbf{i}) \times (\frac{1}{x})\mathbf{i} = \mathbf{j} \in \mathfrak{J}.$$

Hence,

$$\mathbf{j} \times -\mathbf{i} = \mathbf{k} \in \mathfrak{J}$$

and

$$\mathbf{j} \times \mathbf{k} = \mathbf{i} \in \mathfrak{J}.$$

Now, because $\mathbf{i}, \mathbf{j}, \mathbf{k} \in \mathfrak{J}$ we know that $\mathfrak{J} = \mathfrak{R}^3$. □

7.3 Homomorphisms

Now we can discover another way to generate ideals within a Lie algebra using the kernel of a homomorphism.

Definition 5. *Let $\phi : \mathfrak{g} \to \mathfrak{g}'$ be a function. We say that ϕ is a **Lie algebra homomorphism** if ϕ preserves sums, scalar multiplication and Lie Brackets. That is, for all $c \in \mathbb{R}$ and $X, Y \in \mathfrak{g}$,*

$$\phi(cX) = c\phi(X),$$

$$\phi(X + Y) = \phi(X) + \phi(Y)$$

and

$$\phi([X, Y]) = [\phi(X), \phi(Y)]$$

Definition 6. *Let $\phi : \mathfrak{g} \to \mathfrak{g}'$ be a Lie algebra homomorphism. The **Kernel** of ϕ is the set*

$$\mathfrak{h} = \{X \in \mathfrak{g} : \phi(X) = \boldsymbol{0}\}.$$

Theorem 5. *The kernel of any given Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{g}'$ is an ideal of $\mathfrak{g}$.*

Proof. Let $\phi : \mathfrak{g} \to \mathfrak{g}'$ be a Lie algebra homomorphism and let $\mathfrak{h}$ be its kernel.

We will first prove that $\mathfrak{h} \leq \mathfrak{g}$. Let X, Y be any two elements in $\mathfrak{h}$. Notice that

$$\begin{aligned} X, Y \in \mathfrak{h} &\to \phi(X) = \mathbf{0} = \phi(Y) \\ &\to \phi(X + Y) = \phi(X) + \phi(Y) = \mathbf{0} \\ &\to X + Y \in \mathfrak{h} \end{aligned}$$

and for any $c \in \mathbb{R}$

$$\begin{aligned} X \in \mathfrak{h} &\rightarrow c\phi(X) = \mathbf{0} \\ &\rightarrow \phi(cX) = \mathbf{0} \\ &\rightarrow cX \in \mathfrak{h}. \end{aligned}$$

Thus $\mathfrak{h}$ is a subspace of $\mathfrak{g}$.

Now let $G \in \mathfrak{g}$. Notice that for any $X \in \mathfrak{h}$,

$$\phi([X, G]) = [\phi(X), \phi(G)] = [\mathbf{0}, \phi(G)] = \mathbf{0}.$$

Thus, $[X, G] \in \mathfrak{h}$ and we know that $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. □

In the next section we will use this theorem to take a closer look at non-simple Lie algebras.

7.4 Non-simple Lie Algebras

7.4.1 $\mathfrak{gl}(n, \mathbb{C})$

We first motivate that $\mathfrak{gl}(n, \mathbb{C})$ is not simple because $SL(n, \mathbb{C})$ is the kernel of the function $\phi : GL(n, \mathbb{C}) \rightarrow \mathbb{C}^{\times}$ in which for any $M \in GL(n, \mathbb{C})$

$$\phi(M) = \det(M).$$

Now, recall that $\mathfrak{gl}(n, \mathbb{C}) = M_n$ (the space of all $n \times n$ matrices) and $\mathfrak{sl}(n, \mathbb{C})$ is the set of all $n \times n$ matrices with trace 0.

Now, we want find a rigorous proof that $\mathfrak{sl}(n, \mathbb{C})$ is an ideal of $\mathfrak{gl}(n, \mathbb{C})$ using the previous theorem.

Proof. Let $\phi : M_n \rightarrow \mathbb{C}$ be the trace function. Clearly $\mathfrak{sl}(n, \mathbb{C})$ is the kernel of ϕ so if we prove this function is a homomorphism, we know for free that $\mathfrak{sl}(n, \mathbb{C})$ is an ideal of $\mathfrak{gl}(n, \mathbb{C})$.

Consider two arbitrary elements X, Y of M_n. By the properties of the trace function we know that

$$Tr(X + Y) = Tr(X) + Tr(Y)$$

and for any $z \in \mathbb{C}$

$$Tr(zX) = zTr(X).$$

Furthermore, we know that in general $Tr(XY) = Tr(YX)$ so

$$Tr([X, Y]) = Tr(XY) - Tr(YX) = \mathbf{0} = [Tr(X), Tr(Y)].$$

Hence, ϕ is a Lie algebra homomorphism and $\mathfrak{sl}(n, \mathbb{C})$ is an ideal of $\mathfrak{gl}(n, \mathbb{C})$. □

7.4.2 $\mathfrak{so}(4)$

The goal of this section is to prove that there is an ideal in $\mathfrak{so}(4)$.

In section 3 we have a theorem that states that we can define the rotation of any vector $q \in \mathbb{R}^4$ as $q \to v^{-1}qw$ where $v, w \in Sp(1)$. Thus, the function $\Phi : Sp(1) \oplus Sp(1) \to SO(4)$ defined as,

$$\Phi(v, w) = v^{-1}qw.$$

for any $v, w \in Sp(1)$ gives all possible rotations of q in $\mathbb{R}^4$ in a 2–1 homomorphism.

In starting our discussion of Lie groups we first expect $\mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$ to be the Lie algebra of $Sp(1) \oplus Sp(1)$. Indeed this is the case because any smooth path in $Sp(1) \oplus Sp(1)$ has the form $u(t) = (v(t), w(t))$ and

$$u'(0) = (v'(0), w'(0)) \in \mathfrak{sp}(1) \oplus \mathfrak{sp}(1).$$

We expect that the function $\phi : \mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \to \mathfrak{so}(4)$ in which

$$\phi(u'(0), v'(0)) = u'(0)qv'(0)$$

for arbitrary paths is also a homomorphism. In fact ϕ ends up being a 1–1. This is because we know that there are only two pairs $(u(t), v(t))$ and $(-u(t), -v(t))$ map to the same rotation $q \to u(t)qv(t)$. We also know that $v'(0)$ is in $\mathfrak{sp}(1)$ so $-v'(0)$ not. The same goes for $u'(0)$. Now define a new function $\psi : \mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \to \mathfrak{sp}(1) \oplus \mathfrak{sp}(1)$ such that

$$\psi(v'(0), w'(0)) = (0, w'(0)).$$

Clearly ψ has the nontrivial kernel $(\mathfrak{sp}(1), 0) \approx \mathfrak{sp}(1) \approx \mathfrak{so}(3)$. Since all are isomorphic to the cross product algebra over $\mathbb{R}^3$. Thus there is an ideal that lives inside $\mathfrak{so}(4)$ isomorphic to $\mathfrak{so}(3)$.

7.4.3 Constructing $\mathfrak{gl}(n, \mathbb{C})$ From Simple Algebras

It is possible to construct $\mathfrak{gl}(n, \mathbb{C})$ from $\mathfrak{sp}(1)$ through a process called *complexification.* What complexification does is it takes any "real" vector space (in the sense that it is closed under scalar multiplication by real numbers) and extends it to be closed under scalar multiplication by *complex* numbers. Thus for any matrix group G we can complexify its tangent space $\mathfrak{g}$ to create the new set:

$$\mathfrak{g} + i\mathfrak{g} = \{X + iY : X, Y \in \mathfrak{g}\}.$$

Now we have the following theorem discussing some properties of a complexified Lie algebra.

Theorem 6. *A complexified Lie algebra,* $\mathfrak{g} + i\mathfrak{g}$ *is closed under scalar multiplication by complex numbers, vector addition and Lie brackets.*

Proof. Scalar multiplication by complex numbers: Let $X + Yi \in \mathfrak{g} + i\mathfrak{g}$ and let $a + bi \in \mathbb{C}$. Notice that by properties of the real vector space $\mathfrak{g}$:

$$(a + bi)(X + Yi) = aX - bY + i(aY + bX) \in \mathfrak{g} + i\mathfrak{g}.$$

Vector Addition: This directly follows from properties of the real vector space $\mathfrak{g}$.

Lie Bracket: Let $X_1 + iY_1$ and $X_2 + iY_2$ be elements of $\mathfrak{g} + i\mathfrak{g}$. Notice that:

$$[X_1 + iY_1, X_2 + iY_2] = [X_1, X_2] - [Y_1, Y_2] + i([X_1, Y_2] + [Y_1, X_2]).$$

This is clearly an element of $\mathfrak{g} + i\mathfrak{g}$ because $\mathfrak{g}$ is closed under Lie brackets. $\square$

Theorem 7. *The Lie algebra* $\mathfrak{gl}(n, \mathbb{C}) = \mathfrak{u}(n) + i\mathfrak{u}(n)$.

Proof. We know for free that any matrix in $\mathfrak{u}(n) + i\mathfrak{u}(n)$ will be a complex $n \times n$ matrix. Thus $\mathfrak{u}(n) + i\mathfrak{u}(n) \subset \mathfrak{gl}(n, \mathbb{C})$.

Now let $X \in \mathfrak{gl}(n, \mathbb{C})$, or rather let X be any complex $n \times n$ matrix. Notice that

$$X = \frac{X - X^\dagger}{2} + i\frac{X + X^\dagger}{2i}.$$

Now notice that

$$\frac{X - X^\dagger}{2} + \left(\frac{X - X^\dagger}{2}\right)^\dagger = \frac{X - X^\dagger}{2} + \frac{X^\dagger - X}{2} = \mathbf{0}.$$

Thus, $\frac{X - X^\dagger}{2} \in \mathfrak{u}(n)$. Now notice that

$$\frac{X + X^\dagger}{2i} + \left(\frac{X + X^\dagger}{2i}\right)^\dagger = \frac{X + X^\dagger}{2i} - \frac{X^\dagger + X}{2i} = \mathbf{0}.$$

Thus $i\frac{X + X^\dagger}{2i} \in \mathfrak{u}(n)$ and $X \in \mathfrak{u}(n) + i\mathfrak{u}(n)$. This completes the proof. $\square$

This just gives a taste of how one can begin to construct more complex algebras using simple algebras as building blocks.

References

[1] John Stillwell. *Naive Lie Theory.* University of San Francisco, San Francisco, CA, 2010.

[2] Pertti Lounesto. *Clifford Algebras and Spinors.* Cambridge University Press, Cambridge, England, 2002.

[3] David J. Griffiths. *Introduction to Elementary Particle Physics 2nd edition.* Wiley-VCH, 2008.

[4] Jean Gallier. *Basics of Classical Lie Groups: The Exponential Map, Lie Groups, and Lie Algebras* University of Pennsylvania