$GL_n(\mathbb{R})$ AS A LIE GROUP

MAXWELL LEVINE

ABSTRACT. This paper will provide some simple explanation about matrix Lie groups. I intend for the reader to have a background in calculus, point-set topology, group theory, and linear algebra, particularly the Spectral Theorem. I focus on $SO(3)$, the group of rotations in three-dimensional space.

CONTENTS

1. MATRIX GROUPS OVER $\mathbb{R}$, $\mathbb{C}$, AND $\mathbb{H}$

We focus on the matrix groups over the rings $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$, which are all Lie groups. For the most part, we will not discuss abstract Lie groups. However, it is a fact that every compact Lie group is isomorphic to a matrix group (a statement of the theorem can be found in [3]). Hence, the use of matrix groups simply provides a more direct approach to the material.

Definition 1.1. For a ring R, the set $M_n(R)$ denotes all n-by-n matrices over R.

$M_n(R)$ forms a ring under matrix addition and multiplication, but it has zero divisors (i.e., $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$). Furthermore, its additive group is uninteresting because it is simply the additive group of R^{n^2}. Therefore, we focus our interest on the *multiplicative* group of $M_n(R)$.

Definition 1.2. For a ring R, the set $GL_n(R)$ denotes the multiplicative group of $M_n(R)$, namely the group of invertible matrices over R, and is called the *general linear group*. If R is commutative, then the determinant function is well-defined. In this case, the set of matrices of determinant 1 is denoted $SL_n(R)$ and is called the *special linear group*. For this paper, we will focus on the case in which $R = \mathbb{R}, \mathbb{C}, \mathbb{H}$.

We know that $GL_n(R)$ is a group by virtue of the multiplicative property of determinants: If $A, B \in GL_n(R)$ and so $\det A$ and $\det B$ are nonzero, then

Date: August 21, 2009.

$\det(AB) = \det(A)\det(B) \neq 0$, so $AB \in GL_n(R)$. Similarly, we know that $\det(A^{-1}) = 1/\det(A) \neq 0$, so $GL_n(R)$ is closed under inverses. Analogous arguments can be made to show that $SL_n(R)$ is a subgroup of $GL_n(R)$.

The general linear group of n-by-n matrices is important first of all because it has the group structure that $M_n(R)$ lacks but also because it is all-encompassing in the sense that it contains many important groups.

Definition 1.3. For $n \times n$ matrices A, $O(n, \mathbb{R}) = \{A : A^T A = I\}$ (where A^T is the transpose of A) is the *orthogonal group*, and the subgroup $SO(n, \mathbb{R})$ of $O(n, \mathbb{R})$ of matrices of positive determinant is called the *special orthogonal group*.

Note that it is equivalent to define the orthogonal group as the subgroup of $GL_n(\mathbb{R})$ such that $A^T = A^{-1}$. $O(n, \mathbb{R})$ is a group because $A, B \in O(n, \mathbb{R})$ implies that $(AB)(AB)^T = A(BB^T)A^T = AA^T = I$ and $A^{-1}(A^{-1})^T = A^T A = I$. $SO(n, \mathbb{R})$ is a group because $A, B \in SO(n, \mathbb{R})$ implies that $\det(AB) = \det A \cdot \det B = 1 \cdot 1 = 1$.

Note that every matrix $n \times n$ can be interpreted as a linear transformation of n-dimensional space. The following proposition elaborates on this notion.

Proposition 1.4. *Let $A \in GL_n(\mathbb{R})$ and let $\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$ be the Euclidean inner product. The following are equivalent:*

1. *$AA^T = I$ (that is, $A \in O(n, \mathbb{R})$).*
2. *For entries a_{ij}, a_{ji}, $a_{ij}a_{ji} = \delta_{ij}$.*
3. *The columns of A are an orthonormal basis, and so are the rows of A.*
4. *A sends orthonormal bases to orthonormal bases.*
5. *A preserves the Euclidean norms of vectors under multiplication.*

The proof is, step by step, more or less immediate. Since the orthogonal group preserves the lengths of all vectors, we can thing of it is representing rotations and reflections in $\mathbb{R}^n$. In this case, the sign of the determinant of a matrix indicates whether it is orientation-preserving. Elements of the special orthogonal group all have positive determinant, which indicates that these transformations are rotations.

Definition 1.5. If $A = (a_{ij}) \in GL_n(\mathbb{C})$, then $\overline{A} = (\overline{a_{ij}})$ is the matrix of the complex conjugates of the entries of A. We define $A^* = \overline{A^T} = (\overline{a_{ji}})$ to be the *conjugate transpose* of A. The set $U(n) = \{A : A^* A = I\}$ is called the *unitary group* and the subgroup of $U(n)$ of matrices of determinant 1 is the *special unitary group*.

One of the interesting aspects of the general linear group is that $GL_n(\mathbb{C})$ is a subgroup of $GL_{2n}(\mathbb{R})$. This can be understood on a conceptual level because an element in $GL_n(\mathbb{C})$ is a transformation from $\mathbb{C}^n$ to itself. Since $\mathbb{C}$ is a vector space over $\mathbb{R}$ with the basis $\{1, i\}$, we can identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ using an injection that sends each entry $a + bi$ of $A \in GL_n(\mathbb{C})$ to a 2-by-2 section $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ of a $2n$-by-$2n$ matrix over $\mathbb{R}$. The result is that we reinterpret A as a transformation from $\mathbb{R}^{2n}$ to itself. This discussion can be summed up as the following:

Proposition 1.6. *$GL_n(\mathbb{C})$ is isomorphic to a subgroup of $GL_{2n}(\mathbb{R})$.*

Since we can think of these matrix groups as subgroups of $GL_n(\mathbb{R})$, we can conceive of intersections of certain matrix groups, even if they are over different fields. For example, the intersection of $O(2n, \mathbb{R})$ and $GL_n(\mathbb{C})$ is $U(n)$. Suppose A

is an element of this intersection. Since $GL_n(\mathbb{C})$ consists of all $\mathbb{C}$-linear invertible matrices, we consider A as an element of $GL_{2n}(\mathbb{R})$. Since $A \in O(2n, \mathbb{R})$, we know that $A^T A = I$. Given the isomorphism φ, we can think of the matrix $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)^T = \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ as the conjugate of the element $a + bi \in \mathbb{C}$. It follows that A^T corresponds to A^*, so that $A \in U(n)$.

Definition 1.7. In order to define operations such that $\mathbb{R}^4$ is a ring, we write it over the basis $\{1, i, j, k\}$ and use the relations

$$\begin{aligned} &i^2 = j^2 = k^2 = -1 \\ &ij = k \qquad ji = -k \\ &jk = i \qquad kj = -i \\ &ki = j \qquad ik = -j \end{aligned}$$

Under these operations, $\mathbb{R}^4$ is a non-commutative division ring which we call the *Hamiltonian numbers.* This ring contains $\mathbb{R}$ in its center and is denoted $\mathbb{H}$.

Despite the lack of commutativity, the Hamiltonians are similar to the complex numbers in that they have a conjugation operation $q \longrightarrow \overline{q}$ which sends $a+bi+cj+dk$ to $a-bi-cj-dk$, and a norm $|q| = \sqrt{q\overline{q}} = \sqrt{a^2 + b^2 + c^2 + d^2}$, which can be checked. Furthermore, just as multiplication by a complex number is an isometry of $\mathbb{C}$, multiplication by an element of $\mathbb{H}$ is an isometry is well, meaning that it preserves distances. This is because for $u, v \in \mathbb{H}$, $|uv| = |u||v|$, so $|u(v - w)| = |u||v - w|$.

Furthermore, we can define an operation on the matrices in $GL_n(\mathbb{H})$ that sends $A = (q_{ij})$ to $A^* = \overline{A^T} = (\overline{q_{ji}})$, which is an analog of the conjugate transpose operation. (The question of which is being used should be clear from context).

Definition 1.8. The $*$ operation defined is the *Hamiltonian conjugate transpose.* The group of matrices $Sp(n) = \{A \in GL_n(\mathbb{H}) : AA^* = I\}$ is the *symplectic group.*

In fact, there exists an injection of $\mathbb{H}$ into $GL_4(\mathbb{R})$ given by the action of each of $a + bi + cj + dk$ on the basis elements of $\mathbb{H}$ which takes the form

$$a + bi + cj + dk \mapsto \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix} \tag{1.9}$$

Thus, we find that $GL_n(\mathbb{H}) \leq GL_{4n}(\mathbb{R})$.

The property that makes $\mathbb{H}$ important in the context of this paper is that the Hamiltonian numbers represent rotations in three-dimensional space. Of particular interest are the Hamiltonians of norm 1, which is denoted $\mathbb{S}^3$. $\mathbb{S}^3$ is a group under multiplication, and it is isomorphic to $SU(2)$. Furthermore, $\mathbb{S}^1$ acts on the space $i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$ of imaginary Hamiltonians, which we denote $\mathbb{R}^3$, by rotation.

Proposition 1.10. *The conjugation operation $t \mapsto qtq^{-1}$ for fixed $q \in \mathbb{S}^3$ is a rotation of $i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$, the three-dimensional hyperplane in $\mathbb{H}$. In particular, if we write $q = \cos\theta + u \sin\theta$, then conjugation rotates $i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$ by the angle 2θ around the axis through u.*

Proof. First of all, note that the multiplication of elements in $\mathbb{R}^3$ is $uv = -u \cdot v + u \times v$, where $\cdot$ and $\times$ are the respective dot and cross products in our standard conception of $\mathbb{R}^3$. So if $u \in i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$, then $u^2 = -1$.

Since $\mathbb{R}$ commutes with all elements of $\mathbb{H}$, $t \in \mathbb{R}$ means that $qtq^{-1} = t$ for $q \neq 0$, so this operation maps $\mathbb{R}$ to itself. Furthermore, it is a ring homomorphism and it has the inverse $t \mapsto q^{-1}tq$, so it is a bijection. Now suppose that $q(bi+cj+dk)q^{-1} = a + b'i + c'j + d'k$. But then $q^{-1}(a + b'i + c'j + d'k)q = a + q^{-1}(b'i + c'j + d'k)q$, which contradicts the fact that the conjugation operation has an inverse. Hence, facts together imply that conjugation by elements of the unit hypersphere maps $i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$ to itself.

Now, we can write $q = \cos\theta + u\sin\theta$ for $\theta \in \mathbb{R}$ since $u^2 = -1$ (think in terms of the fact that $|\cos\theta + u\sin\theta| = \cos^2\theta + \sin^2\theta = 1$). One can check that conjugation fixes u, and therefore real multiples of u is well. Thus, since conjugation is an isometry that preserves a line, this function is a rotation of $i\mathbb{R} + j\mathbb{R} + k\mathbb{R}$.

Now choose a vector v such that $|v| = 1$ and v is orthogonal to u and use $w = v \times u$. The set $\{u, v, w\}$ is thus an orthonormal basis for $\mathbb{R}^3$. The angle of the rotation can then be found by the action of conjugation on this basis, from which we get the following matrix:

$$\begin{pmatrix} quq^{-1} & qvq^{-1} & qwq^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{pmatrix}$$

This matrix is a rotation of the plane spanned by v and w by the angle 2θ.

□

One more important consideration with regard to the action of $\mathbb{H}$ is that every three-dimensional rotation corresponds to *two* Hamiltonian numbers. Consider the fact that a three-dimensional rotation can be described in terms of finding a vector v in the three-dimensional sphere $\mathbb{S}^2$ and an angle θ for the rotation (v, θ). Then this rotation is the same as that for $(-v, -\theta)$. With this consideration, keep in mind the following proposition, a proof of which can be found in [2].

Proposition 1.11. *The map that sends the conjugation action by an element $q \in \mathbb{S}^3$ to a rotation in $\mathbb{R}^3$ is surjective, with the kernel $\{1, -1\}$.*

2. Topology and Matrix Groups

Now we can introduce topological concepts. We frame them in terms of the general linear group, since we have shown that all of the matrix groups that strike our interest are subgroups of the general linear group. I am assuming familiarity with point-set topology.

Definition 2.1. Suppose $A \in GL_n(\mathbb{R})$. Let $|A| = \sqrt{\sum a_{ij}^2}$. This is the Euclidean norm on n space, and we can define a Euclidean metric with $d(A, B) = |A - B|$. The set of balls $B(x, \epsilon)$ in this metric form a basis for the standard topology on $\mathbb{R}^{n^2}$.

This notion allows us to speak of open and closed subsets of $GL_n(\mathbb{R})$.

Example 2.2. The determinant function $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$ is continuous because it is a polynomial function. The set $\mathbb{R} \setminus \{0\}$ is an open set, and therefore so is its preimage, $GL_n(\mathbb{R})$.

Since there exist matrices with negative determinants, and since $\mathbb{R}^\times = (-\infty, 0) \cup (0, \infty)$ is disconnected, is follows that $GL_n\mathbb{R}$ is disconnected as well. In particular, this implies that $O(n, \mathbb{R})$ is disconnected. However, $\{A \in M_n(\mathbb{R}) : \det A > 1\}$ is connected because $\mathbb{R}^+$ is connected, and so is its coset of matrices with negative determinants.

Definition 2.3. A group G is a *topological group* if both $G \times G \to G$ is continuous and $g \mapsto g^{-1}$ is continuous as well.

It follows from the definition that $x \mapsto g^{-1}x$ is continuous as well, so the map $x \mapsto gx$ is a homeomorphism. It maps open sets to open sets and the preimages of open sets are open.

Proposition 2.4. *$GL_n(\mathbb{R})$ ands its subgroups are topological groups.*

Proof. Fix $A \in GL_n(\mathbb{R})$ and define $f : B \mapsto AB$. Then the entries in AB are polynomials in the entries of A and B. If the entries in B are variables, then each entry of AB changes continuously. Thus, we find that f is a continuous function. Furthermore, the map $A \mapsto A^{-1}$ is continuous for the same reason. □

Note that *right* multiplication in $GL_n(\mathbb{R})$ is continuous as well. Therefore, if H is a closed subgroup of a matrix group G, then the cosets gH and the conjugates gHg^{-1} are closed in G. We will use these facts to prove more topological properties of matrix groups later on.

Example 2.5. $O(n, \mathbb{R})$ is closed in $M_n(\mathbb{R})$. Consider its compliment $\{A : AA^T \neq I\}$. As in the above proposition, $f : X \mapsto XX^T$ is continuous because the entries are polynomials in the entries of X. Therefore, $O(n)$ is closed because it is the preimage of the point I under the continuous function f.

Furthermore, since $SL_n(\mathbb{R})$ is closed (since it is the preimage of $\{1\}$ under det) it follows that $SO(n, \mathbb{R}) = O(n, \mathbb{R}) \cap SL_n(\mathbb{R})$ is closed. This means that the coset $\{A \in O(n, \mathbb{R}) : \det A = -1\}$ is closed as well by the fact that $GL_n(\mathbb{R})$ is a topological groups.

Although we may have been emphasizing the properties of matrix groups as sets of linear functions, we can think of matrix groups as spatial objects as well. $M_n(\mathbb{R})$ is an n^2-dimensional vector space (and $M_n(\mathbb{C})$ a $2n^2$-dimensional vector space and so on), so we can conceive of elements of $M_n(\mathbb{R})$ as being points in n^2-dimensional Euclidean space.

Definition 2.6. A *path* through $M_n(\mathbb{R})$ is a function $\gamma : \mathbb{R} \to M_n(\mathbb{R})$ which sends a variable t to a matrix $A(t) = (a_{ij}(t))$.

Proposition 2.7. *In $GL_n(\mathbb{R})$, connectedness is equivalent to path-connectedness.*

Proof. Topologically, path-connectedness always implies connectedness. To prove the converse, we must prove that $GL_n(\mathbb{R})$ is locally path-connected.

Given $A \in GL_n(\mathbb{R})$, we know by definition that $\det A \neq 0$, so $\det A \in \mathbb{R}^\times$, which is open. Let U be an open neighborhood of $\det A$. Then the preimage of U is an open neighborhood of A. Without loss of generality, we can assume that U is an open ball centered at A. Thus, given $B \in U$, we can find a path $\varphi : [0, 1] \to U$ from A to B defined by $t \mapsto A + t(B - A)$.

Local path-connectedness and connectedness together imply path-connectedness because of the following: Let U be a connected set of matrices, let $A \in U$, and let

$X, Y \subset U$ be defined respectively as the matrices which can be connected by a path to A and those which cannot. Clearly, $X \cup Y = U$ and $X \cap Y = \emptyset$.

Then we can apply basic point-set topology: Using the concatenation of paths and local connectedness, we can prove that X and Y are both open because all of their points are interior points. But since U is connected, it must follows that $X = U$, since $X \neq \emptyset$.

□

The topological properties we have discussed here are not mere curiosities. They are vital to our discussion of $GL_n(\mathbb{R})$, as we will see in the forthcoming sections.

3. The Exponential Function

The exponential function for matrices, an analog of the exponential function for real numbers, is the one of the most important tools for the discussion of matrix Lie groups. It takes the following form:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots = \sum_{k=0}^{\infty} \frac{A^k}{k!} \tag{3.1}$$

Note that in some contexts we will denote the exponential function of A as $\exp(A)$. Remember that the groups we discuss are subgroups of the general linear group, so it suffices to prove the convergence of $\exp A$ for $A \in GL_n(\mathbb{R})$. Now let us define the norm $\|A\| = n \cdot |A|$ where $|A| = \max |a_{ij}|$ over $GL_n(\mathbb{C})$ so that $|a_{ij}| = |c_{ij} + id_{ij}| = \sqrt{c_{ij}^2 + d_{ij}^2}$. Considered in this way,. the norm $\|\cdot\|$ over $GL_n\mathbb{R}$ is just the restriction of the norm over $GL_n(\mathbb{C})$, so it suffices to prove that exp converges over $GL_n(\mathbb{C})$. Note that this norm defines the same standard topology as the Euclidean norm.

Proposition 3.2. *The following are properties of* $\|\cdot\|$*:*

(1) $\|A + B\| \leq \|A\| + \|B\|$
(2) $\|\lambda A\| = |\lambda| \|A\|$
(3) $\|AB\| \leq \|A\| \|B\|$
(4) exp *converges under this norm.*

Proof. *(1)* and *(2)* are fairly simple to prove. We get *(3)* from

$$|(AB)_{ij}| \leq \sum_{i,j} |a_{ik}||b_{jk}| \leq \frac{n\|A\|\|B\|}{n^2} = \frac{\|A\|\|B\|}{n}$$

Then we see that the series converges entrywise, because the $\|\frac{A^k}{k!}\| \leq \frac{|a_{ij}|_k^k}{k!}$ and applying the ratio test shows us that $\|\frac{A^k}{k!}\| / \|\frac{A^{k-1}}{(k-1)!}\| \leq \frac{|a_{ij}|_k^k}{k} \to 0$.

□

Now we can detail some of the important properties of the exponential function.

Proposition 3.3. *The exponential function for matrices has the following properties for $n \times n$ matrices A and B:*

(1) $e^0 = I$
(2) If A and B commute, then $e^{A+B} = e^A e^B$
(3) e maps $M_n(\mathbb{R})$ onto nonsingular matrices.

(4) *If B is invertible, then* $e^{BAB^{-1}} = Be^A B^{-1}$.
(5) $\det e^A = e^{Tr(A)}$

Proof. The first two points follow from calculations on the infinite series.

To find (3), note that A and $-A$ commute, so $\det e^A \det e^{-A} = \det e^A e^{-A} = \det e^{A-A} = \det e^0 = \det I = 1$. Thus, we must have $\det A \neq 0$.

(4) follows from the fact that

$$(BAB^{-1})^n = BA(B^{-1}B)AB^{-1}\ldots = BA^nB^{-1}$$

To find the last fact, we apply the Jordan Theorem for linear operators and the multiplicative property of the determinant function. Since A is conjugate to an upper-triangular matrix, we can write $BAB^{-1} = C$, where C is upper-triangular. Now, if $C = (a_{ii})$ is a diagonal matrix, then $C^k = (a_{ii}^k)$, so $e^D = (e^{a_{ii}})$. It follows that

$$\begin{aligned} e^{Tr(A)} =& e^{a_{11}+\ldots+a_{nn}} = e^{a_{11}}\ldots e^{a_{nn}} = \det e^D = \det(e^{BAB^{-1}}) = \\ =& \det(Be^AB^{-1}) = \det(BB^1)\det(e^A) = \det e^A \end{aligned}$$

Note that this works over $M_n(\mathbb{R})$ as well as $M_n(\mathbb{C})$, because even if the diagonal conjugate is in $M_n(\mathbb{C})$, the calculation returns to $M_n(\mathbb{R})$.

□

The exponential function has a local inverse around the identity matrix in the form of the logarithmic function. Again, we define it in terms of a well-known series:

$$\log A = (A-I) - \frac{1}{2}(A-I)^2 + \frac{1}{3}(A-I)^3 - \ldots = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}(A-I)^k \tag{3.4}$$

The logarithm converges because if $\|A\| < \epsilon$, then $\|\frac{A^k}{k}\| \leq \frac{n^{k-1}\epsilon^k}{k}$, so the ratio test shows us that $\|\frac{A^k}{k!}\|/\|\frac{A^{k-1}}{(k-1)!}\| = \frac{k}{k+1}\epsilon \to \epsilon$. Thus, the series converges as long as $\|A - I\| < 1$.

Proposition 3.5. *The matrix logarithm is analogous to the logarithm for real numbers. That is, if* $\log A$ *and* $\log B$ *are defined, and if* A *and* B *commute, then we get the following:*

(1) $\log(e^A) = A$
(2) $\log(AB) = \log(A) + \log(B)$

Proof. We find one (1) by computing the series for $\log e^A$ and collecting the terms for $(A-I)^k$ and finding that the coefficients sum to zero. Using the fact that e and log are inverses (meaning that the exponential is injective near 0), we find that

$$e^{\log AB} = AB = e^{\log A}e^{\log B} = e^{\log A + \log B}$$

□

4. Tangent Spaces

The importance of the exponential function is that it maps the tangent space of a matrix group to the group itself. The tangent spaces are essentially the same concept as tangent lines in calculus and tangent planes in geometry. Consider the one-dimensional example of the circle group $\mathbb{S}^1 \subset \mathbb{C}$. It is well-known that, using the Euler identity, we map the line $i\theta$ onto $= \{e^{i\theta} : 0 \leq \theta < 2\pi\} = \mathbb{S}^1$. We can imagine that the line $i\theta$ is a tangent space to the identity (actually, the line $1 + i\theta$, but we have the same geometry) on S^1. Essentially, the exponential function wraps the tangent space around the group.

This example does not reveal much about the relation between matrix groups and their tangent spaces because it is one-dimensional and commutative. In this section, we will discuss how this concept is generalized to multiple dimensions and yields more complicated results. Since tangents in $\mathbb{C}$ are paths, we must define paths in $M_n(\mathbb{R})$. First, let us prove an important fact about paths.

Definition 4.1. A path is *smooth* if each $a_{ij}(t)$ is differentiable. In this case, we have $A'(t) = (a'_{ij}(t))$, so that each entry is differentiated.

Example 4.2. Consider the path $A(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ This is a path through four-dimensional space, but also through $SO(2)$. It is smooth and it has the derivative $A'(t) = \begin{pmatrix} -\sin t & -\cos t \\ \cos t & -\sin t \end{pmatrix}$.

If we have two paths $A(t)$ and $B(t)$, we can define a new path $A(t)B(t)$, which is the product of the matrices. In this case, the standard differentiation rules apply. In particular $\frac{d}{dt}A(t)B(t) = A'(t)B(t) + A(t)B'(t)$.

Definition 4.3. Consider a path $A(t)$ through a matrix group G such that $A(0) = I$. A *tangent vector* at the identity is a matrix $A'(0)$. Furthermore, the set of tangent vectors at the identity is denoted $T(G)$.

Many texts denote the tangent space at the identity as $T(G)_1$ or something similar. One can consider the possibility of tangent spaces at other points defined analogously, but since we will not make use of this notion, we will continue with this notation.

Proposition 4.4. . *$T(G)$ is an $\mathbb{R}$-linear vector space.*

Proof. Let $X \in T(G)$, and suppose that $A(t)$ is the path through G such that $X = A'(0)$. Then if $r \in \mathbb{R}$, $A(rt)$ is also a path through G, and $\frac{d}{dt}A(rt) = rA'(t)$. Now $A(r \cdot 0) = A(0) = I$, so $rA'(0) = rX \in T(G)$.

Now suppose that $X, Y \in T(G)$ and $A(t)$ and $B(t)$ are their corresponding paths through G (so that $A(t) = B(t) = I$). Then the path $A(t)B(t)$ is equal to I at 0. We differentiate $A(t)B(t)$ to get a path $A'(t)B(t) + A(t)B'(t)$ which at 0 is $A'(0) + B'(0) = X + Y \in T(G)$. □

Definition 4.5. A matrix A is *skew-symmetric* if $A + A^T = 0$, meaning that $a_{ij} = -a_{ji}$, and in particular that $a_{ii} = 0$. The set of n-by-n skew-symmetric matrices is denoted $so(n)$.

Skew-symmetric matrices are important because of their relationship with $SO(n)$.

Proposition 4.6. $\exp$ *maps $so(n)$, the set of skew-symmetric matrices, into $O(n)$*

Proof. If $X \in T(G)$, then $X + X^T = 0$, and X and X^T commute because $X(X^T) = X(-X) = (-X)X = X^T X$. Thus, $e^X(e^X)^T = e^X e^{X^T} = e^{X+X^T} = e^0 = I$. □

Proposition 4.7. *The tangents at the identity of $SO(n)$ are skew-symmetric matrices.*

Proof. Let $A(t)$ be such that $X = A(0) = I$ and $A(t)^T A(t) = I$ (i.e. $A(t)$ is a path in $SO(n)$). Then we differentiate both sides of $A(t)^T A(t) = I$ to get $A'(t)^T A(t) + A(t)^T A'(t) = 0$. If we evaluate this equation at 0, we get $A'(0)^T I + IA'(0) = A'(0)^T + A'(0) = X + X^T = 0$. □

It can be proved similarly that the tangents of $U(n)$ are the matrices X such that $X + X^* = 0$ and that the tangents of $Sp(n)$ are the matrices X such that $X + X^* = 0$.

Proposition 4.8. *All n-by-n skew-symmetric matrices can be realized as tangent vectors of $SO(n)$.*

Proof. Let X be skew-symmetric, and let $\gamma(t) = e^{tX}$, so γ is a path through $O(n)$. Since $\gamma(0) = I$, and since $SO(n)$ and $O(n) - SO(n)$ are disconnected, γ is a path through $SO(n)$ in particular. We utilize the fact that $\gamma'(t) = Xe^{tX}$, which applies for any X and can be justified by differentiating the terms of the series. Since $\gamma'(0) = X$, we find that X is a tangent vector of $SO(n)$. □

We can sum all of this up with the following:

Theorem 4.9. *The tangent space of $SO(n)$ is $so(n)$.*

Example 4.10. A skew-symmetric matrix $A \in so(2)$ takes the form $A = \left(\begin{smallmatrix} 0 & x \\ -x & 0 \end{smallmatrix}\right)$.

$$\begin{aligned} e^A &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}^2 + \frac{1}{3!}\begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}^3 + \ldots = \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix} + \ldots = \\ &= \begin{pmatrix} 1 - \frac{1}{2!}x^2 + \ldots & x - \frac{1}{3!}x^3 + \ldots \\ -(x - \frac{1}{3!}x^3 + \ldots) & 1 - \frac{1}{2!}x^2 + \ldots \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \end{aligned}$$

The tangent space allows us to study the group locally. Again, consider $\mathbb{S}^1$, which exists in $\mathbb{R}^2$, but has a one-dimensional surface. Although we represent it in two dimensions, it has a one-dimensional form.

Definition 4.11. The *dimension* of G is the dimension of $T(G)$ as a vector space.

Example 4.12. The dimension of $GL_n(\mathbb{R})$ is n^2 because its tangent space is $M_n(\mathbb{R})$. Suppose that $A \in M_n(\mathbb{R})$, and define a path $\gamma : t \mapsto I + tA$. Since $\det I \neq 0$, there is a neighborhood of I that is contained entirely by $GL_n(\mathbb{R})$. hence, for a small enough t, γ is contained in $GL_n(\mathbb{R})$. Then, since $\gamma'(0) = A$, A is a tangent vector of $GL_n(\mathbb{R})$.

Example 4.13. The dimension of $SO(n)$ as a lie group is defined by the dimension of $so(n)$ as a linear space. Skew symmetric matrices are defined by the entries strictly above the diagonal, since the entries below the diagonal are the additive inverses of those above. Hence, the dimension of $SO(n)$ is $n + (n-1) + \ldots + 1 = \frac{n(n-1)}{2}$.

Tangent spaces can be explored in great detail, but first it is worth discussing another type of important structure in $GL_n(\mathbb{R})$.

5. Maximal Tori

Definition 5.1. A *circle group* is a group of matrices isomorphic to $\mathbb{S}^1$. A *torus* is a direct sum of circle groups. A *maximal torus* is a torus in a matrix group that is not contained in any other torus.

Example 5.2. The subgroup of matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

in $SO(4)$ is a torus, but not a maximal torus, because it sits inside the subgroup of matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\phi & -\sin\phi \\ 0 & 0 & \sin\phi & \cos\phi \end{pmatrix}$$

On the other hand, in $SO(3)$, the subgroup of matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is, in fact, a maximal torus. This makes intuitive sense because we cannot "fit" any more rotation blocks into the matrix.

Definition 5.3. The *rank* of a maximal torus T is the number k such that

$$T = \underbrace{\mathbb{S}^1 \oplus \ldots \oplus \mathbb{S}^1}_{k\text{-times}}$$

Since maximal tori are conjugate, as we will prove at the end of this section, and since the conjugate operation is a group automorphism, the rank of a group is well-defined.The rank is an isomorphism-invariant property of matrix groups separate from dimension. For example, $SO(2)$ has a dimension of 1, while $SO(3)$ has a dimension of 3. However, both have a rank of 1.

Proposition 5.4. *An example of a maximal torus in $SO(2n)$ is*

$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cos\theta_k & -\sin\theta_k \\ 0 & 0 & 0 & \sin\theta_k & \cos\theta_k \end{pmatrix}$$

Furthermore, all maximal tori in $U(n)$ take the form

$$\begin{pmatrix} e^{i\theta_1} & 0 & 0 & 0 \\ 0 & e^{i\theta_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & e^{i\theta_n} \end{pmatrix}$$

Proof. Let A be an element of any torus containing the torus T depicted in the above example. It is enough to show that if A commutes with every element of T, then $A \in T$. Let R_k be the matrix with just the $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$-block in place of $\left(\begin{smallmatrix} a_{2k-1,2k-1} & a_{2k-1,2k} \\ a_{2k,2k-1} & a_{2k,2k} \end{smallmatrix}\right)$ with ones in the rest of the diagonal entries and zeroes everywhere else. Then R_k is an element of the maximal torus T. Now let $v_1, \ldots, v_{2n}$ be the orthonormal basis represented by the identity matrix. Then

$$\begin{aligned} R_k A(v_{2k-1}) =& R_k(a_1 v_1 + \ldots + a_{2n} v_{2n}) = \\ =& a_1 v_1 + \ldots + (-a_{2k-1} v_{2k-1} - a_{2k} v_{2k}) + \ldots + a_{2n} v_{2n} \\ A R_k(v_{2k-1}) =& A(-v_{2k-1}) = -a_1 v_1 - \ldots - a_{2n} v_{2n} \end{aligned}$$

Since $R_k A(v_{2k-1}) = A R_k(v_{2k-1})$, it follows that $a_i = -a_i$, and hence $a_i = 0$ for all $i \neq 2k, 2k-1$. This means that $A(v_{2k-1})$ is in the $2k-1, 2k$-plane and therefore that A represents a rotation through that plane. This means that $A \in T$.

The proof for $SO(n)$ can be generalized to $U(n)$ following the same methods. □

From this proof, one could generalize to the case of $SO(2n+1)$, in which there are k blocks and a leftover 1 in the diagonal. We will proceed to prove that tori are entirely characterized by being compact, connected, abelian matrix groups.

Lemma 5.5. *A matrix group G is connected if and only if $\langle \exp(x) : x \in T(G) \rangle = G$.*

Proof. Suppose first that G is connected. Let U be a neighborhood of the identity of G, and H be the subgroup of G generated by U. Given any $h \in H$, hU is an open neighborhood containing h. Therefore, H as a whole is open, and so are its cosets in G. Furthermore, $G = \bigcup_{g \in G} gH$, meaning that the cosets of H disjointly partition G. If $H \neq G$, then the cosets constitute a partition of G in open sets, which is contrary to our hypothesis. Therefore, it must be the case that $H = G$, meaning that G is generated by any open neighborhood of the identity. Since $\exp(T(G))$ is a neighborhood of the identity, this implies that $G = \exp(T(G))$.

Now suppose that G is generated by $G = \exp(T(G))$, and recall that connectedness in G is equivalent to path-connectedness. If g and h are two points of G, we can write $g^{-1}h = \exp(X_1)\exp(X_2)\ldots\exp(X_k)$ where $X_1, \ldots, X_k \in T(G)$ and define a path $\gamma(t) = g \cdot \exp(tX_1)\exp(tX_2)\ldots\exp(tX_k)$ through G. Then $\gamma(0) = g$ and $\gamma(1) = h$. □

Lemma 5.6. *The kernel of* exp *is discrete.*

Proof. The fact that the exponential function is bijective on the neighborhood U or 0 implies that 0 is the only element of $K = ker(\exp)$ is contained in this neighborhood. Therefore, there are no elements $x \neq y \in ker(\exp)$ such that $x - y \in U$. □

Lemma 5.7. $ker(\exp) \cong \mathbb{Z}^k$ *for some k.*

A detailed proof can be found in [3], but we can provide a sketch: We can choose a vector v_1 in $ker(\exp)$ such that $v_1 = \min\{v : v = \lambda v_1\}$, meaning that v_1 is minimal in a particular direction. Then we choose v_2 in the orthogonal compliment of v_1 the same way and so on until we exhaust the dimension of the kernel and get a set of vectors $v_1, \ldots, v_k$ which form a linear basis for $ker(\exp)$. This shows that $ker(\exp) = \mathbb{Z}v_1 \oplus \ldots \oplus \mathbb{Z}v_k \cong \mathbb{Z}^k$.

Theorem 5.8. *Every compact, connected, abelian matrix group G is isomorphic to a torus.*

Proof. $G = \exp(T(G))$ by Lemma 5.4, so if $T(G)$ is of dimension $n \geq k$, then by the fundamental theorem of group isomorphisms, $G \cong \exp T(G)/\mathbb{Z}^k \cong \mathbb{R}^{n-k} \oplus \mathbb{R}^k/\mathbb{Z}^k$. If $n \neq k$, then this space is unbounded and therefore not compact. Thus, $\dim K = n$ and so $G \cong \mathbb{R}^k/\mathbb{Z}^k = \mathbb{T}^k$.

□

Note that the converse of this theorem is also true: All maximal tori are compact, connected, and abelian. It is left to show how some connected groups G are covered by the conjugates of their maximal tori.

For some matrix groups, the conjugates of the maximal tori cover the group. However, for this condition to hold, the group must be connected. This is because for a maximal torus T, $\cup_{g \in G} gTg^{-1}$ is connected because T, and hence gTg^{-1}, contains the identity and the union of connected sets sharing an element is connected. Hence, it is *not* the case that the conjugates of the maximal tori in $O(n, \mathbb{R})$ cover the group.

Proposition 5.9. *The conjugates of the maximal torus in $U(n)$ cover the group.*

Proof. This is essentially an application of the Spectral Theorem for linear operators using the fact that the maximal torus is the group of diagonal unitary matrices. If $A \in U(n)$, then we know that $A^* = A^{-1}$, and hence A and A^* commute. Thus, we can apply the spectral theorem to find that A is conjugate by a unitary matrix B to a diagonal matrix. Since the maximal torus T consists of diagonal matrices, this means that $BAB^{-1} \in T$. □

Proposition 5.10. *The conjugates of the maximal torus in $SO(n)$ cover the group.*

Proof. There is a more specific statement that can be deduced from the spectral theorem, namely that every real symmetric matrix is conjugate via $B \in SO(n)$ to a diagonal matrix. From this we can deduce that for $A \in SO(n)$, $\mathbb{R}^n$ can be broken up into subspaces of dimension 1 or 2 that are invariant under A.

It suffices to demonstrate that $\mathbb{R}^n$ has at least one invariant subspace of desired dimension. Define $S = A + A^T$, which is symmetric and therefore has an eigenvector v. If Av and v are linearly dependent, then our A-invariant subspace is the space spanned by v, so suppose that Av and v are linearly independent and define $V = span(Av, v)$. Then an element of V can be written $\alpha v + \beta Av$. Since v is an eigenvector, $Sv = (A + A^T)v = (A + A^{-1})v = Av + A^{-1}v = \lambda v$ for some $\lambda \in \mathbb{R}$, which can be written as $A^2v + v = \lambda Av$. Hence, we can calculate $A(\alpha v + \beta Av) = \alpha Av + \beta A^2 v = \alpha Av + \beta(\lambda Av - v) = -\beta v + (\alpha + \beta\lambda)Av$, which is an element of V.

We now choose an orthonormal basis $v_1, \ldots, v_n$ of $\mathbb{R}$ such that each A-invariant subspace is generated by precisely one or two vectors in the basis. To be more specific, let us say that $v_1, \ldots v_k$ are the one-dimensional A-invariant spaces and

$w_{k+1}, w_{k+2}, \ldots, w_{n-1}, w_n$ are the two-dimensional A invariant spaces. Then let B be the matrix that maps e_i from the standard basis to v_i in our new basis for all $1 \leq i \leq n$. If $\det B = 1$, then it follows that $BAB^{-1} \in T$. If $\det B = -1$, then we can reorder two of the "w" vectors and relabel the new matrix as B, and again we can say that $BAB^{-1} \in T$.

□

Explanations of the usage of the spectral theorem can be found in [1].

Lemma 5.11. *Let $r \in \mathbb{S}^1$ be a rotation by an irrational angle α. Then $\mathcal{O} = \{r^k(1)\}_{k=0}^{\infty}$ is dense in $\mathbb{S}^1$.*

Proof. First of all, $\mathcal{O}$ is infinite. If it were finite, then it would the case that $k\alpha = z \in \mathbb{Z}$ and thus $\alpha = z/k \in \mathbb{Q}$, which is a contradiction. Now, since $\mathbb{S}^1$ is compact, we can cover $\mathbb{S}^1$ with a finite set $I = \{I_1, \ldots, I_n\}$ of open intervals of length (by which I mean $|x-y|$ is the angle of their difference) less than ϵ. By the pigeon-hole principle, there exist $n, m \in \mathbb{Z}$ such that $|r^m(1) - r^n(1)| < \epsilon$. Now find x such that $r^m(x) = r^n(1)$, so $r^{(n-m)}(1) = x$. Then $|x-1| < \epsilon$. Since it is possible to reach a point arbitrarily close to 1 in $m - n = k$ steps, $\mathcal{O}$ comes within ϵ of any point on $\mathbb{S}^1$ through iterations of r^k.

□

Theorem 5.12. *In a connected matrix group G, all of the maximal tori are conjugate.*

Proof. Let G be of rank n, and let T and T' be two maximal tori in G. Now let $\{r_1, \ldots, r_n\}$ be a set of irrational rotations of $\mathbb{S}^1$ that are linearly independent over $\mathbb{Q}$. In this context, this means that $q_1 r_1 + \ldots + q_n r_n = 0$ implies that the rational q_i's are all zero. Now define $r = r_1 \oplus \ldots \oplus r_n$ to be a rotation of G. By the lemma, the orbit $\mathcal{O}$ of r is dense in each component. Given $x = (x_1, \ldots, x_n) \in T$, we can independently choose elements in the orbits $\{r_1^k(1)\}, \ldots, \{r_n^k(1)\}$ to be arbitrarily close to $x_1, \ldots, x_n$. Therefore, the orbit of r is dense in T'.

Since every element in G is in a conjugate of T, it follows that $r \in xTx^{-1}$ for some $x \in G$. But xTx^{-1} is a group, so $\mathcal{O} \subset xTx^{-1}$. But we already know that $\mathcal{O} \subset T'$. Since $T' \cap xTx^{-1}$ is closed, this means that $T' \cap xTx^{-1} = T'$, meaning $T' \subset xTx^{-1}$. But since T' is maximal, it follows that $T' = xTx^{-1}$.

□

6. Lie Algebras

In this section, we will examine the relationship between conjugation in a matrix group G and operations in the tangent space $T(G)$.

Proposition 6.1. *If X and Y are tangent vectors of G, then $XY - YX \in T(G)$.*

Proof. Define $\gamma(s,t) = e^{sX}e^{tY}(e^{sX})^{-1} = Xe^{tY}X^{-1}$, which is a smooth path such that $\gamma(0) = 1$, so $\frac{d}{dt}\gamma|_{t=0} = X(Ye^0)X^{-1} \in T(G)$. Furthermore, $\eta(s) = \frac{d}{dt}\gamma|_{t=0} = e^{sX}e^{tY}(e^{sX})^{-1} = e^{sX}Y(e^{sX})^{-1}$ is a smooth function of s, so its tangent is also in $T(G)$.

$$\frac{d}{ds}\eta(s) = \frac{d}{ds}(e^{sX}Ye^{-sX}) = (Xe^{sX})(Ye^{-sX}) + (e^{sX}Y)(-e^{sX})$$

Therefore, $\eta'(0) = XY - YX \in T(G)$.

□

Hence, the matrix $XY - YX$ in a sense measures the extent to which X and Y are non-commutative, and the extent to which the paths that they are tangent to twist one another.

Definition 6.2. The operation $[A, B] = AB - BA$ which we derived above is called the *bracket operation.*

Definition 6.3. The *Lie algebra* defined on the tangent space of a matrix group is the tangent space with linear operations as well as the bracket operation.

The Lie algebra is also closed under the bracket operation by Proposition 6.1, and the bracket operation characterizes the tangent space as an algebraic object.

Proposition 6.4. *The Lie bracket has the following properties:*

1. *Linearity:* $[A + B, C] = [A, C] + [B, C]$
2. *Anti-commutativity:* $[A, B] = -[B, A]$
3. *The Jacobi Identity:* $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$

Example 6.5. If G is abelian, then $T(G)$ is abelian, and $[X, Y] = XY - XY = 0$ for all $X, Y \in T(G)$. This is why the tangent space of $\mathbb{S}^1$ does not give us a clear idea of what tangent spaces are like in other matrix groups.

Definition 6.6. An *ideal* I in $T(G)$ is a subspace such that $[X, Y] \in T(G)$ for all $X \in T(G)$ and $Y \in I$.

Proposition 6.7. *The tangent space $T(H)$ of a normal subgroup H of a matrix group G is an ideal in $T(G)$.*

Proof. It should be clear that $T(H)$ a Lie algebra, so we need so show that it is closed under the bracket operation with elements of $T(G)$. Let $X \in T(G)$ and $Y \in T(H)$. Since H is normal, it is closed under conjugation This means that we can define a function η like the one we used in Proposition 6.1, and its tangent vector is $XY - YX$. □

Proposition 6.8. *A differentiable group homomorphism $\psi : G \to H$ induces a homomorphism of Lie algebras between $T(G)$ and $T(H)$.*

Proof. Let $\varphi(X) = \frac{d}{dt}\psi(\exp(tX))|_{t=0}$. φ will be the homomorphism of Lie algebras in question. First, we can find

$$\begin{aligned}\frac{d}{ds}\psi(\exp(sX)) =& \frac{d}{dt}\psi(\exp((s+t)X)|_{t=0} = \\ =& \frac{d}{dt}\psi(\exp(sX)\exp(tX))|_{t=0} = \\ =& \psi(\exp(sX)\frac{d}{dt}\psi(\exp(tX))|_{t=0} = \\ =& \psi(\exp(sX))\varphi(X)\end{aligned}$$

This is the familiar differential equation $\frac{d}{dt}f(at) = af(at)$, which has the solution $f(at) = e^{at}$. Hence, this shows that $\psi(\exp X) = \exp(\varphi(X))$. From this, we get

$$\psi(\exp(tX)\exp(sY)\exp(-tX)) = \exp(t\varphi(X))\exp(s\varphi(Y))\exp(-t\varphi(X))$$

We differentiate both sides with respect to s as $s = 0$ to get

$$\varphi(\exp(tX)Y\exp(-tX) = \exp(t\varphi(X))\varphi(Y)\exp(-t\varphi(X))$$

Finally, we differentiate both sides with respect to t at $t = 0$. The computation is the same as that for the proof of proposition 6.1. It follows that $\varphi[X, Y] = [\varphi(X), \varphi(Y)]$.

□

This proof can be found in [3].

Example 6.9. The skew-symmetric three-by-three matrices all take the form

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

So they can be thought of as a vector space with the basis of the matrices,

$$E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

One can calculate that $[E_1, E_2] = E_3$, $[E_2, E_3] = E_1$, and so on. This space is isomorphic as an algebra to $\mathbb{R}^3$ under the cross product. If φ sends E_1 to $\hat{\imath}$, E_2 to $\hat{\jmath}$, and E_3 to $\hat{k}$, then $\varphi[A, B] = [\varphi(A), \varphi(B)]$.

Suppose we tried to solve the equation $e^X e^Y = e^Z$ for Z. Then we could write $Z = \log(e^X e^Y)$ and expand to get

$$Z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left\{ \left(\sum_{i=0}^{\infty} \frac{X^i}{i!} \right) \left(\sum_{j=0}^{\infty} \frac{Y^j}{j!} \right) - 1 \right\}^k = \tag{6.10}$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left\{ \sum_{i,j \geq 0, i+j \geq 1} \frac{X^i Y^j}{i! j!} \right\}^k = \tag{6.11}$$

$$= \sum \frac{(-1)^{k-1}}{k} \frac{X^{i_1} Y^{j_1} \dots X^{i_k} Y^{j_k}}{i_1! j_1! \dots i_k! j_k!} \tag{6.12}$$

This is indeed a very cumbersome expression, but it turns out that the first few terms are

$$Z = X + Y + \frac{1}{2}[X, Y] + \dots \tag{6.13}$$

In fact, all of the terms can be expressed in nested Lie brackets. This statement is known as the Campbell-Baker-Hausdorff Theorem:

Theorem 6.14. *The terms in the series* (6.12) *can all be expressed in terms of rational coefficients multiplied by Lie brackets.*

Proofs of entirely different natures can be found in [3] and [4]. Note that if X and Y commute, then all of the terms are 0 and $Z = X + Y$. But this is the result of Proposition 3.5. The Campbell-Baker-Hausdorff Theorem defines Lie theory on a qualitative level, because it demonstrates the extent of the interplay between conjugation in the Lie group and the bracket operation in the Lie algebra. It produces many interesting results.

Proposition 6.15. *Suppose that G and H are connected matrix groups. Then a homomorphism between their respective tangent spaces $T(G)$ and $T(H)$ induces a homomorphism between G and H.*

Proof. Let $\varphi : T(G) \to T(H)$ be a homomorphism. Then for $X \in T(G)$, define a function $\Phi(X) = e^{\varphi(X)}$. Note that any element of $g \in G$ can be written as e^X for $X \in T(G)$ and every element of $h \in H$ can be written as e^Y for $Y \in T(H)$. Then we use the fact that $\varphi[X, Y] = [\varphi(X), \varphi(Y)]$ in the following:

$$\Phi(gh) = \Phi(de^X e^Y) = \Phi(e^{(X+Y+\frac{1}{2}[X,Y]+\dots)}) = e^{\varphi(X+Y+\frac{1}{2}[X,Y]+\dots)} =$$
$$= e^{(\varphi(X)+\varphi(Y)+\frac{1}{2}[\varphi(X),\varphi(Y)]+\dots)} = e^{\varphi(X)}e^{\varphi(Y)} = \Phi(g)\Phi(h)$$

□

In this proof, we used the facts that the solution of $e^X e^Y = e^Z$ can be written in terms of Lie brackets. In fact, this proof can be generalized to matrix groups that are not necessarily connected, but that proof involves more topology. It can be found in [3] and [4].

Acknowledgments. I would like to thank my mentor, Vipul Naik, for putting so much time and effort into helping me with my REU project, and also for paving the way for my understanding of the material with his work as the teaching assistant for my abstract algebra course last year.

References

[1] Michael Artin. Algebra. Prentice Hall. 1991.
[2] Morton L. Curtis. Matrix Groups. Springer-Verlag. 1979.
[3] Wulf Rossmann. Lie Groups: An Introduction Through Linear Groups. Oxford Science Publications. 2002.
[4] John Stillwell. Naive Lie Theory. Springer. 2008.