AN INTRODUCTION TO LIE THEORY THROUGH MATRIX GROUPS

BRENDEN COLLINS

Abstract. In this paper we prove that matrix groups are manifolds and use them as a special case to introduce the concepts of Lie groups, Lie algebras, and the exponential map.

Contents

1. Introduction

The aim of this paper is to introduce the reader to the topic of Lie groups through the specific example of matrix groups. While matrix groups do not characterize Lie groups as a whole, many of the most studied and useful Lie groups arise as matrix groups, so matrix groups are an excellent concrete example to introduce a student to Lie theory. In the first section, we begin by defining matrix groups and giving a number of important examples of them. In the second section and third section we introduce the reader to the idea of a tangent space to a matrix group (specifically its Lie algebra, when considering the matrix group as a Lie group) and the exponential map, which maps elements from the Lie algebra of a matrix group into the group. We also derive the Lie algebras for the example matrix groups. The fourth and final section gives a general introduction to Lie groups and Lie algebras and shows that the previous examples of matrix groups are in fact Lie groups.

2. Introduction to Matrix Groups

In this section we introduce the concept of matrix groups, which are sets of matrices that satisfy the group axioms with matrix multiplication as the group operation.

Throughout this paper we shall use $\mathbb{K}$ to refer to both $\mathbb{R}$ and $\mathbb{C}$ when the definitions and propositions apply to both fields.

Date: 9/4/14.

Definition 2.1. $M_n(\mathbb{K})$ is the set of $n \times n$ square matrices with entries in $\mathbb{K}$.

$M_n(\mathbb{K})$ is not a group under matrix multiplication because some of its elements do not admit a multiplicative inverse. We thus consider an important subgroup of $M_n(\mathbb{K})$.

Definition 2.2. The general linear group over $\mathbb{K}$ is:

$$GL_n(\mathbb{K}) := \{A \in M_n(\mathbb{R}) \mid \exists B \in M_n(\mathbb{R}) \text{ with } AB = BA = I\},$$

where I is the standard identity matrix:

$$I_{ij} = \begin{cases} 1 \text{ if } i = j. \\ 0 \text{ if } i \neq j. \end{cases}$$

Proposition 2.3. *$GL_n(\mathbb{K})$ forms a group with matrix multiplication as the group operation.*

Proof. To show this we shall verify that it satisfies the four group axioms: closure, associativity, identity, and invertibility.

(i) Recall that the product of two $n \times n$ matrices is an $n \times n$ matrix, as well as the fact that a square matrix is invertible if and only if its determinant is non-zero. Then closure can be satisfied by observing that for some elements A, B in $GL_n(\mathbb{K})$, $\det(A)$ and $\det(B)$ are both non-zero. Thus

$$0 \neq \det(A)\det(B) = \det(AB),$$

so we have that AB is an element of $GL_n(\mathbb{K})$.

(ii) Matrix multiplication is associative.

(iii) Notice that $II = I$. This shows that I is its own inverse, which means $I \in GL_n(\mathbb{K})$. Additionally, for any $A \in GL_n(\mathbb{K}), AI = IA = A$, thus I is the identity element of the group.

(iv) Every element A of $GL_n(\mathbb{K})$ has some inverse $B = A^{-1}$ by definition. □

The general linear group of matrices is isomorphic to the set of invertible linear transformations. That is, $GL_n(\mathbb{K})$ acts on an n-dimensional vector space over $\mathbb{K}$ and represents the set of invertible linear transformations on such a vector space. This is stated without proof, but follows from the fact that any linear transformation of a vector space can be represented by a matrix chosen with respect to some basis.

Each element of $M_n(\mathbb{K})$ has n^2 elements, so $M_n(\mathbb{K})$ can be regarded as a vector space over $\mathbb{K}$ of dimension n^2. This allows us to consider topological properties of $M_n(\mathbb{K})$, which is important for the following definition of matrix groups.

Definition 2.4. A matrix group is a subgroup $G \subset GL_n(\mathbb{K})$ which is topologically closed in $GL_n(\mathbb{K})$.

This topological closure of every matrix group $G \subset GL_n(\mathbb{K})$ means that any convergent sequence of elements of G will convergent to some $g \in G$. This will become important later when we discuss matrix groups as manifolds.

We now introduce several examples of subsets of $GL_n(\mathbb{K})$ and prove that they are matrix groups by showing that they are closed in $GL_n(\mathbb{K})$.

Definition 2.5. We define the orthogonal group over $\mathbb{K}$ as:

$$\mathcal{O}_n(\mathbb{K}) := \{A \in GL_n(\mathbb{K}) \mid \langle XA, YA \rangle = \langle X, Y \rangle \text{ for all } X, Y \in \mathbb{K}^n\},$$

where $\langle X, Y \rangle$ represents the standard dot product of X and Y.

If $\mathbb{K} = \mathbb{R}$ it is denoted $O(n)$ and called the orthogonal group.

If $\mathbb{K} = \mathbb{C}$ it is denoted $U(n)$ and called the unitary group.

Equivalently, a real matrix is said to be orthogonal and a complex matrix is said to be unitary if

$$A \cdot A^* = I.$$

We next prove a simple property of the orthogonal group before defining a subgroup of particular interest called the special orthogonal group.

Proposition 2.6. *If* $A \in \mathcal{O}_n(\mathbb{K})$*, then* $|\det(A)| = 1$.

Proof. By definition, we have that $A \cdot A^* = I$.

Thus $1 = \det(A \cdot A^*) = \det(A)\det(A^*) = \det(A)\det(\overline{A}^T) = \det(A)\det(\overline{A}) = \det(A)\overline{\det(A)} = |\det(A)|^2$. □

We now define the special subgroups of the orthogonal group.

Definition 2.7. The special orthogonal group is:

$$SO(n) := \{A \in O(n) \mid \det(A) = 1\}.$$

The special unitary group is:

$$SU(n) := \{A \in U(n) \mid \det(A) = 1\}.$$

Both groups are subgroups of the general linear group as well as the special linear group:

$$SL_n(\mathbb{K}) := \{A \in GL_n(\mathbb{K}) \mid \det(A) = 1\}.$$

Proposition 2.8. $\mathcal{O}_n(\mathbb{K}), SL_n(\mathbb{K}), SO(n)$*, and* $SU(n)$ *are matrix groups.*

Proof. These examples are all subgroups of $GL_n(\mathbb{K})$, so it can be easily shown that they satisfy the group axioms through the properties of the determinant. This means that to show that they are matrix groups we simply have to prove that they are closed in $GL_n(\mathbb{K})$.

For $\mathcal{O}_n(\mathbb{K})$, we define a function $f : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ as $f(A) = A \cdot A^*$. This function is continuous, because for all i, j the component function

$$f(A)_{ij} = (A \cdot A^*)_{ij}$$

is continuous because it is a polynomial in the entries of A. We know that the image of f restricted on $\mathcal{O}_n(\mathbb{K})$ is $\{I\}$, which is a closed subset of $M_n(\mathbb{K})$. Thus $\mathcal{O}_n(\mathbb{K}) = f^{-1}(\{I\})$. The continuity of f tells us that $f^{-1}(\{I\})$ is closed in $M_n(\mathbb{K})$, thus $\mathcal{O}_n(\mathbb{K}) = f^{-1}(\{I\})$ is closed in $M_n(\mathbb{K})$ and is therefore closed in $GL_n(\mathbb{K})$.

For $SL_n(\mathbb{K})$, we first show that the determinant function $\det : M_n(\mathbb{K}) \to \mathbb{K}$ is continuous. Continuity follows from the fact that $\det(A)$ is a polynomial of degree n in the entries of A (this can be seen from the expansion of minors formula for calculating a determinant, the proof of which is not in the scope of this paper). The image of the determinant function restricted to $SL_n(\mathbb{K})$ is $\{1\}$, so we consider the set $\{1\}$ which is closed in $\mathbb{K}$. In this case we use $\{1\}$ instead of $\{I\}$ because the determinant function maps any object in $SL_n(\mathbb{K})$ to 1, whereas our function in the previous case mapped objects in $\mathcal{O}_n(\mathbb{K})$ to I. Then $SL_n(\mathbb{K}) = \det^{-1}(\{1\})$ is closed in $M_n(\mathbb{K})$ and therefore also in $GL_n(\mathbb{K})$ by the continuity of $\det(A)$.

For $SO(n)$ and $SU(n)$ recall that the intersection of two closed sets is closed and notice that $SO(n) = O(n) \cap SL_n(\mathbb{R})$ and $SU(n) = U(n) \cap SL_n(\mathbb{C})$.

□

3. Lie algebras and the exponential map

A matrix group $G \subset GL_n(\mathbb{K})$ is a subset of the Euclidean space $M_n(\mathbb{K})$, so we can discuss its tangent spaces.

Definition 3.1. We define a path through a point p in a matrix group G as a parametrization

$$\gamma(t) : (-\epsilon, \epsilon) \to G \text{ for } \epsilon > 0 \text{ and } \gamma(0) = p$$

We next define the tangent space to a matrix group at a point.

Definition 3.2. Let $G \subset \mathbb{R}^m$ be a matrix group and let $p \in G$. We shall call two differentiable paths γ_1 and γ_2 through p in G equivalent if $\gamma_1'(0) = \gamma_2'(0)$. This defines an equivalence relation on such curves, where the equivalence classes $\gamma'(0)$ are the tangent vectors of G at p. We then can define the tangent space to G at p, T_pG, as the set of equivalence classes of differentiable paths through p in G. That is:

$$T_pG := \{\gamma'(0) \mid \gamma_i(t) \text{ are equivalent differentiable paths through } p \text{ in } G\ \}.$$

In other words, if we think of γ as a differentiable path through p in $G \subset \mathbb{R}^m$, then T_pG is the set of initial velocity vectors.

We next define a vector space which we shall refer to as the Lie algebra of a matrix group. This definition is not the general definition of a Lie algebra, but we shall show that the two definitions coincide in the final section of this paper.

Definition 3.3. For some matrix group $G \subset GL_n(\mathbb{K})$, we shall refer to the tangent space to G at I as the Lie algebra of G. It is denoted $\mathfrak{g} := \mathfrak{g}(G) := T_IG$.

This definition is not the formal definition of a Lie algebra. However, we shall show later that this vector space satisfies the formal definition, so for now we ignore the Lie algebra structure and refer to it as a Lie algebra in anticipation of this later proof.

We shall now show that $\mathfrak{g}$ is a subspace of $M_n(\mathbb{K})$ which is important when using it to study its associated matrix group. Proving this will require the product rule for matrices.

Proposition 3.4. *If $\gamma, \beta : (-\epsilon, \epsilon) \to M_n(\mathbb{K})$ are differentiable, then so is the product path $(\gamma \cdot \beta)'(t) = \gamma(t) \cdot \beta'(t) + \gamma'(t) \cdot \beta(t)$.*

The proof of this proposition is simply verification through matrix multiplication, so we omit it here.

This result allows us to prove the main proposition:

Proposition 3.5. *The Lie algebra $\mathfrak{g}$ of matrix group $G \subset GL_n(\mathbb{K})$ is a real subspace of $M_n(\mathbb{K})$.*

Proof. We first note that a differentiable path in G can be represented by a matrix. The parametrization is simply composed of n^2 differentiable functions that map from $(-\epsilon, \epsilon)$ to $\mathbb{K}$ for each coordinate, yielding an $n \times n$ matrix over $\mathbb{K}$. The derivative of the path is simply the derivative of each coordinate function, which is thereby also an $n \times n$ matrix. Thus $\mathfrak{g}$ is an element of $M_n(\mathbb{K})$.

Now we must show that $\mathfrak{g}$ is closed under addition and scalar multiplication.

Consider $A, B \in \mathfrak{g}$. Then $A = \alpha'(0)$ and $B = \beta'(0)$ for some α, β in G with $\alpha(0) = \beta(0) = I$. The product path $\sigma(t) = \alpha(t) \cdot \beta(t)$ is differentiable and lies in G. By the product rule, $\sigma'(0) = \alpha'(0) \cdot \beta(0) + \alpha(0) \cdot \beta'(0) = A + B \in \mathfrak{g}$.

Next, let $\lambda \in \mathbb{R}$ and $A \in \mathfrak{g}$. Then $A = \gamma'(0)$ for some path γ such that $\gamma(0) = I$. The path $\sigma(t) := \gamma(\lambda t)$ is a path in G. $\sigma(0) = \gamma(\lambda \cdot 0) = I$, thus $\sigma'(0)$ in $\mathfrak{g}$. We have that $\sigma'(0) = \lambda \cdot A$, thus $\lambda \cdot A$ in $\mathfrak{g}$. □

To prove that a subspace $\mathfrak{g}$ of $M_n(\mathbb{K})$ is a Lie algebra of a matrix group G, we need to construct a path $\gamma : (-\epsilon, \epsilon) \to G$ for each element A of $\mathfrak{g}$ such that $\gamma(0) = I$ and $\gamma'(0) = A$. While it is possible to construct paths explicitly for individual examples of matrix groups such as $GL_n(\mathbb{K})$ or $\mathcal{O}_n(\mathbb{K})$, we shall instead introduce a path that can be contained in any matrix group G for which A is a tangent vector.

Since the elements of Lie algebras in this special case are simply matrices, we can associate a vector field with any element $A \in \mathfrak{g}$ for a matrix group G by simply multiplying A by every $g \in G$. When looking at paths in G, we want a simple path to associate with A, for which we use integral curves of A.

Definition 3.6. A path $\alpha : (-\epsilon, \epsilon) \to \mathbb{R}^m$ is called an <u>integral curve</u> of a vector field $F : \mathbb{R}^m \to \mathbb{R}^m$ if $\alpha'(t) = F(\alpha(t))$ for all $t \in (-\epsilon, \epsilon)$.

Intuitively, this means that the tangent vector to a path at a point t is equal to the value of the vector field at $\alpha(t)$. That is, the vector field gives the value of the tangent to the path at every point. It turns out that this natural path can be be given for any A with a simple function, given by <u>matrix exponentiation</u>.

Before introducing matrix exponentiation, we briefly discuss series in $M_n(\mathbb{K})$. A series
$$\sum A_l = A_0 + A_1 + A_2 + \dots$$
of elements $A_l \in M_n(\mathbb{K})$ converges if for all i, j the series $(A_0)_{ij} + (A_1)_{ij} + (A_2)_{ij} + \dots$ converges to some A_{ij} in $\mathbb{K}$. We write $\sum A_l = A$.

This allows us to consider power series of matrices. We introduce the following proposition:

Proposition 3.7. *Let $f(x) = c_0 + c_1 x + c_2 x^2 + \dots$ be a power series with coefficients $c_i \in \mathbb{K}$ with a radius of convergence R. If $A \in M_n(\mathbb{K})$ satisfies $|A| < R$, then $f(A) = c_0 I + c_1 A + c_2 A^2 + \dots$ converges absolutely.*

The proof of this proposition is not difficult, but we omit it here for the sake of brevity. The proposition, however, allows us to define the exponential function $f(x) = \exp(x) = e^x$ on matrices. Evaluating the power series of the exponential function on a matrix $A \in M_n(\mathbb{K})$ gives us <u>matrix exponentiation</u>:
$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots.$$
The radius of convergence of this power series is ∞, so e^A converges absolutely for all $A \in M_n(\mathbb{K})$ by Proposition 2.6. In order to prove the following propositions, we introduce some properties of the matrix exponential.

Proposition 3.8. *For a path $\gamma(t) = e^{tA}$ for $A \in M_n(\mathbb{K})$, the derivative γ' is:*

$$\gamma'(t) = A \cdot e^{tA}.$$

Proof. The result is obtained by termwise differentiating the power series of the matrix exponential.

$$\frac{d}{dt}(I + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \dots) = A + tA^2 + \frac{1}{2!}t^2A^3 + \dots .$$

□

Proposition 3.9. *If $A, B \in M_n(\mathbb{K})$ commute, then $e^{A+B} = e^A \cdot e^B$.*

Proof. We expand the function into its power series and perform a series of manipulations (one of which requires that A and B commute). To illustrate the method, we look at the first few terms:

$$\begin{aligned} e^A e^B &= (I + A + \frac{1}{2}A^2 + \dots)(I + B + \frac{1}{2}B^2 + \dots) \\ &= I + A + B + AB + \frac{1}{2}A^2 + \frac{1}{2}B^2 + \dots \\ &= I + (A + B) + \frac{1}{2}(A^2 + 2AB + B^2) + \dots \end{aligned}$$

Notice that if A and B commute, we have:

$$\begin{aligned} &= I + (A + B) + \frac{1}{2}(A^2 + AB + BA + B^2) + \dots \\ &= I + (A + B) + \frac{1}{2!}(A + B)^2 + \dots \\ &= e^{A+B}. \end{aligned}$$

To verify that this pattern continues for the whole series, we consider the full power series:

$$\begin{aligned} e^A e^B &= \left(\sum_{r \geq 0} \frac{1}{r!}A^r\right)\left(\sum_{s \geq 0} \frac{1}{s!}B^s\right) \\ &= \sum_{r \geq 0, s \geq 0} \frac{1}{r!s!}A^r B^s \\ &= \sum_{n \geq 0}\left(\sum_{r=0}^{n} \frac{1}{r!(n-r)!}A^r B^{n-r}\right) \\ &= \sum_{n \geq 0} \frac{1}{n!}\left(\sum_{r=0}^{n} \binom{n}{r} A^r B^{n-r}\right) \\ &= \sum_{n \geq 0} \frac{1}{n!}(A + B)^n \\ &= e^{A+B}. \end{aligned}$$

Notice that commutativity of A and B is required for the identity

$$\sum_{r=0}^{n} \binom{n}{r} A^r B^{n-r} = (A + B)^n.$$

□

This result is particularly important because it allows us to show that the exponential map is a group homomorphism. A group homomorphism is best described as a map between two groups that preserves the group structure. Explicitly, for G, H groups and $A, B \in G$, $\phi : G \to H$ is a homomorphism if $\phi(A + B) = \phi(A)\phi(B)$. From this it is easy to show that a group homomorphism will map the identity element of the pre-image group to the identity element of the image.

In fact, this proposition allows us to prove a number of other interesting propositions.

Definition 3.10. Let $gl_n(\mathbb{K}) = M_n(\mathbb{K})$.

Proposition 3.11. *For any $A \in M_n(\mathbb{K})$, $e^A \in GL_n(\mathbb{K})$. Therefore matrix exponentiation gives us a map* $\exp : gl_n(\mathbb{K}) \to GL_n(\mathbb{K})$.

Proof.
$$I = e^0 = e^{A+-A} = e^A \cdot e^{-A}$$
This tells us that e^A is invertible with inverse e^{-A}. Thus $e^A \in GL_n(\mathbb{K})$. □

Theorem 3.12. *$gl_n(\mathbb{K})$ is the Lie algebra of $GL_n(\mathbb{K})$.*

In other words, the Lie algebra of the set of invertible $n \times n$ matrices is simply the set of <u>all</u> $n \times n$ matrices.

Proof. To prove this, we show containment both ways. That is, we show that $gl_n(\mathbb{K}) \subset \mathfrak{g}(GL_n(\mathbb{K}))$ and that $\mathfrak{g}(GL_n(\mathbb{K})) \subset gl_n(\mathbb{K})$.

First we show that $gl_n(\mathbb{K}) \subset \mathfrak{g}(GL_n(\mathbb{K}))$. For any $A \in gl_n(\mathbb{K})$, Proposition 3.11 states that $e^A \in GL_n(\mathbb{K})$. We can parametrize the exponential function by defining $\gamma(t) = e^{tA} \in GL_n(\mathbb{K})$. Then we have $\gamma : (-\epsilon, \epsilon) \to GL_n(\mathbb{K})$ with $\gamma(0) = e^{t\cdot 0} = I$ and $\gamma'(0) = A$. This shows for any $A \in gl_n(\mathbb{K})$, a parametrized matrix exponential yields a path in $GL_n(\mathbb{K})$ with initial velocity equal to A, thus A is in $\mathfrak{g}(GL_n(\mathbb{K}))$, the Lie algebra of $GL_n(\mathbb{K})$.

Showing that $\mathfrak{g}(GL_n(\mathbb{K})) \subset gl_n(\mathbb{K})$ is trivial because the derivative at zero of any invertible $n \times n$ matrix $\gamma(t)$ is going to be an $n \times n$ matrix and is thus in $M_n(\mathbb{K}) = gl_n(\mathbb{K})$. □

We shall also consider the Lie algebra of the orthogonal group.

Definition 3.13. $o_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid A + A^* = 0\}$.

Proposition 3.14. *If $A \in o_n(\mathbb{K})$, then $e^A \in \mathcal{O}_n(\mathbb{K})$.*

Proof. Let $A \in o_n(\mathbb{K})$. It can be easily demonstrated through manipulation of terms in the series that $(e^A)^* = e^{A^*}$.
Then we have:
$$e^A(e^A)^* = e^A e^{A^*} = e^{A+A^*} = e^0 = I.$$
□

Theorem 3.15. *The Lie algebra of $\mathcal{O}_n(\mathbb{K})$ equals $o_n(\mathbb{K})$.*

Proof. We again aim to show containment in both directions.

We first show that $o_n(\mathbb{K}) \subset \mathfrak{g}(\mathcal{O}_n(\mathbb{K}))$. By Proposition 3.14, for any $A \in o_n(\mathbb{K})$, $\gamma(t) = e^{tA}$ is a differentiable path in $\mathcal{O}_n(\mathbb{K})$ with $\gamma(0) = I$ and $\gamma'(0) = A$. Thus there exists a path through the identity in $\mathcal{O}_n(\mathbb{K})$ with derivative equal to A at zero, proving that any A in $o_n(\mathbb{K})$ is in $\mathfrak{g}(\mathcal{O}_n(\mathbb{K}))$.

To show that $\mathfrak{g}(\mathcal{O}_n(\mathbb{K})) \subset o_n(\mathbb{K})$, consider a path $\gamma(t)$ in $\mathcal{O}_n(\mathbb{K})$. Then we have $\gamma(t) \cdot \gamma(t)^* = I$. Using the product rule to differentiate, we get:

$$\gamma'(t) \cdot \gamma(t)^* + \gamma(t) \cdot \gamma'(t)^* = 0$$

Evaluating at $t = 0$ gives

$$\gamma'(0) \cdot \gamma(0)^* + \gamma(0) \cdot \gamma'(0)^* = \gamma(0)' + \gamma'(0)^* = 0$$

which shows that $\gamma'(0) \in o_n(\mathbb{K})$. This demonstrates that $\mathfrak{g}(\mathcal{O}_n(\mathbb{K})) \subset o_n(\mathbb{K})$.

□

Definition 3.16. A <u>one parameter group</u> in a matrix group G is a differentiable group-homomorphism $\gamma : (\mathbb{R}, +) \to G$.

A one parameter group is both an algebraic object (as a homomorphism) and a geometric object (a differentiable path). This duality is one of the reasons that matrix groups and the exponential map are so interesting, as the following proposition helps illustrate:

Proposition 3.17.

(1) *For every $A \in gl_n(\mathbb{K})$, $\gamma(t) := e^{tA}$ is a one parameter group in $GL_n(\mathbb{K})$.*
(2) *Every one parameter group in $GL_n(\mathbb{K})$ has the description $\gamma(t) = e^{tA}$ for some $A \in gl_n(\mathbb{K})$.*

Proof. For part (1) we appeal to Proposition 3.9:

$$\gamma(t_1 + t_2) = e^{t_1 A + t_2 A} = e^{t_1 A} e^{t_2 A} = \gamma(t_1)\gamma(t_2).$$

Notice that $\gamma(t) \cdot \gamma(-t) = I$, which shows that $\gamma(t)$ has inverse $\gamma(-t)$.

For part (2), suppose $\gamma(t)$ is a one-parameter group in $GL_n(\mathbb{K})$. Let $A := \gamma'(0)$. Notice that for all $t \in \mathbb{R}$,

$$\begin{aligned}
\gamma'(t) &= \lim_{h \to 0} \frac{\gamma(t+h) - \gamma(t)}{h} \\
&= \lim_{h \to 0} \frac{\gamma(t)\gamma(h) - \gamma(t) \cdot I}{h} \quad \text{(by part (1))} \\
&= \gamma(t) \lim_{h \to 0} \frac{\gamma(h) - I}{h} \\
&= \gamma(t) \cdot \gamma'(0) \\
&= \gamma(t) \cdot A.
\end{aligned}$$

This tells us that $\gamma'(t) = \gamma(t)A$, so we suspect that $\gamma(t) = e^{tA}$. We verify this by applying the product rule:

$$\begin{aligned}
\frac{d}{dt}(\gamma(t)e^{-tA}) &= \gamma'(t)e^{-tA} + \gamma(t)\frac{d}{dt}(e^{-tA}) \\
&= \gamma(t)Ae^{-tA} - \gamma(t)Ae^{-tA} = 0.
\end{aligned}$$

Thus $\gamma(t)e^{-tA} = I$, which tells us that $\gamma(t) = e^{tA}$. □

In this section we introduced an incomplete concept of the Lie algebra, as well as the exponential map. Both of these concepts will be extremely important in the following section, where we shall prove that all matrix groups are manifolds.

Additionally, we defined and proved what the Lie algebras are for some examples of matrix groups and we showed that matrix exponentiation maps elements from a Lie algebra into their associated matrix group (actually a Lie group, but we will define that later). In the following section, we generalize this to show that the exponential function maps from any Lie algebra to its associated matrix (Lie) group.

4. Matrix groups are manifolds

In this section, we prove two important facts about how the exponential map acts on elements of a matrix group's Lie algebra. For $r > 0$, denote

$$B_r := \{A \in M_n(\mathbb{K}) \mid |A| < r\}$$

where $|A|$ is the Euclidean norm of A, that is the square root of the sum of the squares of each coordinate in A.

Theorem 4.1. *Let $G \subset GL_n(\mathbb{K})$ be a matrix group, with Lie algebra $\mathfrak{g} \in gl_n(\mathbb{K})$.*

(1) For all $X \in \mathfrak{g}, e^X \in G$.
(2) For sufficiently small $r > 0$, $V := \exp(B_r \cap \mathfrak{g})$ is a neighborhood of I in G, and the restriction $\exp : B_r \cap \mathfrak{g} \to V$ is a homeomorphism.

The proof of this theorem requires the use of a number of concepts in analysis. The proof of these are beyond the scope of this paper, but necessary definitions and theorems shall be stated for reference.

Definition 4.2. The directional derivative of f in the direction v at p is defined as:

$$df_p(v) := \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t},$$

if this limit exists.

We also present an alternative formulation:

Proposition 4.3. *$df_p(v)$ is the initial velocity vector of the image under f of any differentiable path $\gamma(t)$ in $\mathbb{R}^m$ with $\gamma(0) = p$ and $\gamma'(0) = v$.*

Theorem 4.4. *(Inverse function theorem). If $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^r on a neighborhood of $x \in \mathbb{R}^n (r \geq 1)$ and df_x is an invertible linear map, then there exists a (possibly smaller) neighborhood U of x such that $V := f(U)$ is a neighborhood of $f(x)$, and $f : U \to V$ is invertible with C^r inverse.*

These theorems, accompanied by standard analysis concepts such as differentiability and smoothness, allow us to prove Theorem 4.1.

Proof. We first prove part (1).
As a Lie algebra, $\mathfrak{g}$ has a subspace structure, as proven earlier. As such, let $\{X_1, \ldots, X_k\}$ be a basis of $\mathfrak{g}$. For each $i = 1, \ldots, k$ choose a differentiable path $\alpha_i : (-\epsilon, \epsilon) \to G$ with $\alpha_i(0) = I$ and $\alpha_i'(0) = X_i$. Define

$$F_{\mathfrak{g}} : (\text{neighborhood of } 0 \text{ in } \mathfrak{g}) \to G$$

as follows:

$$F_{\mathfrak{g}}(c_1 X_1 + \ldots + c_k X_k) = \alpha_1(c_1) \cdot \ldots \cdot \alpha_k(c_k).$$

Notice that

$$F_{\mathfrak{g}}(0) = F_{\mathfrak{g}}(0 \cdot X_1 + \ldots + 0 \cdot X_k) = \alpha_1(0) \cdot \ldots \cdot \alpha_k(0) = I.$$

Additionally, $d(F_{\mathfrak{g}})_0$ is the identity function. This can be easily verified on the basis elements as follows:

$$\begin{aligned}
d(F_{\mathfrak{g}})_0(X_i) &= \lim_{t\to 0} \frac{F_{\mathfrak{g}}(0+tX_i) - f(0)}{t} \\
&= \lim_{t\to 0} \frac{F_{\mathfrak{g}}(t(0\cdot X_1 + \ldots + X_i + \ldots + 0\cdot X_k)) - I}{t} \\
&= \lim_{t\to 0} \frac{\alpha_1(t\cdot 0)\cdot \ldots \cdot \alpha_i(t\cdot 1)\cdot \ldots \cdot \alpha_k(t\cdot 0) - I}{t} \\
&= \lim_{t\to 0} \frac{\alpha_i(0+t) - \alpha_i(0)}{t} \\
&= \alpha_i'(0) \\
&= X_i \text{ for all } i \text{ with } 1 \le i \le k.
\end{aligned}$$

Next, choose a subspace $\mathfrak{p} \subset M_n(\mathbb{K})$ which is complementary to $\mathfrak{g}$. We define $\mathfrak{p}$ complementary as follows: complete the set $\{X_1, \ldots, X_k\}$ to a basis of all of $M_n(\mathbb{K})$ and define $\mathfrak{p}$ to be the span of the added basis elements. Then $M_n(\mathbb{K}) = \mathfrak{g} \times \mathfrak{p}$.

Choose a function $F_{\mathfrak{p}} : \mathfrak{p} \to M_n(\mathbb{K})$ with $F_{\mathfrak{p}}(0) = I$ and $d(F_{\mathfrak{p}})_0(V) = V$ for all $V \in \mathfrak{p}$. As an example, $F_{\mathfrak{p}}(V) := I + V$ would satisfy these conditions.
Next define the function

$$F : (\text{neighborhood of } 0 \text{ in } \mathfrak{g}\times\mathfrak{p} = M_n(\mathbb{K})) \to M_n(\mathbb{K})$$

by the rule $F(X+Y) = F_{\mathfrak{g}}(X)\cdot F_{\mathfrak{p}}(Y)$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{p}$. Notice that $F(0) = I$ and dF_0 is the identity:

$$dF_0 = d(F_{\mathfrak{g}}\cdot F_{\mathfrak{p}})_0 = d(F_{\mathfrak{g}})_0 \cdot F_{\mathfrak{p}}(0) + F_{\mathfrak{g}}(0)\cdot d(F_{\mathfrak{p}})_0 = d(F_{\mathfrak{g}})_0 + d(F_{\mathfrak{p}})_0,$$

so

$$dF_0(X+Y) = d(F_{\mathfrak{g}})_0(X) + d(F_{\mathfrak{p}})_0(Y) = X + Y.$$

F is a product of differentiable functions and is thus differentiable and dF_0 is both linear and invertible. Therefore the inverse function theorem tells us that F has an inverse function defined on a neighborhood of I in $M_n(\mathbb{K})$. We can express the inverse as follows for matrices a in this neighborhood:

$$F^{-1}(a) = u(a) + v(a) \in \mathfrak{g}\times\mathfrak{p}.$$

By definition, $u(F(X+Y)) = X$ and $v(F(X+Y)) = Y$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{p}$ near 0. This means that v tests whether an element $a \in M_n(\mathbb{K})$ near I lies in G:

$$v(a) = 0 \implies F^{-1}(a) = u(a) \in \mathfrak{g} \implies F(u(a)) = a \in G.$$

Let $X \in \mathfrak{g}$ and define $a(t) = e^{tX}$. We wish to show that $a(t)$ lies in G for small t by showing that $v(a(t)) = 0$. It is clear that $v(a(0)) = 0$, because $a(0) = I \in G$, so we simply need to show that $\frac{d}{dt}v(a(t)) = 0$ for small t. Since

$$\frac{d}{dt}v(a(t)) = dv_{a(t)}(a'(t)) = dv_{a(t)}(X\cdot a(t)),$$

the result will come as a result of the following lemma:

Lemma 4.5. *For all $a \in M_n(\mathbb{K})$ near I and all $X \in \mathfrak{g}$, $dv_a(X\cdot a) = 0$.*

Proof. Express a as

$$a = F(Z+Y) = F_{\mathfrak{g}}(Z)\cdot F_{\mathfrak{p}}(Y),$$

where $Z \in \mathfrak{g}$ and $Y \in \mathfrak{p}$. Then for all $W \in \mathfrak{g}$, and for sufficiently small t

$$v(F_{\mathfrak{g}}(Z + tW) \cdot F_{\mathfrak{p}}(Y)) = Y,$$

which means that v is not changing at a in these directions:

$$\begin{aligned} 0 &= \frac{d}{dt}\bigg|_{t=0} v(F_{\mathfrak{g}}(Z + tW) \cdot F_{\mathfrak{p}}(Y)) \\ &= dv_a((d(F_{\mathfrak{g}})_Z(W)) \cdot F_{\mathfrak{p}}(Y)) \\ &= dv_a((d(F_{\mathfrak{g}})_Z(W)) \cdot F_{\mathfrak{g}}(Z)^{-1} \cdot a) \\ &= dv_a(X \cdot a), \end{aligned}$$

where $X := (d(F_{\mathfrak{g}})_Z(W)) \cdot F_{\mathfrak{g}}(Z)^{-1}$. It remains to be shown that X is an arbitrary element of $\mathfrak{g}$. From the definition of directional derivative, it is clear that X is the initial tangent vector of the path

$$t \mapsto F_{\mathfrak{g}}(Z + tW) \cdot F_{\mathfrak{g}}(Z)^{-1}$$

in G. Next, X is arbitrary because the linear map from $\mathfrak{g} \to \mathfrak{g}$ which sends

$$W \mapsto (d(F_{\mathfrak{g}})_Z(W)) \cdot F_{\mathfrak{g}}(Z)^{-1}$$

is the identity map when $Z = 0$. So by continuity it has determinant near 1 and is thus an isomorphism when Z is close to 0. In other words, W can be chosen such that X is any element of $\mathfrak{g}$. □

This lemma completes the proof that if $X \in \mathfrak{g}$, then $e^{tX} \in G$ for $t \in (-\epsilon, \epsilon)$. This result can be extended to all real t by observing that for all $t \in (-\epsilon, \epsilon)$ and $n \in \mathbb{Z}^+$,

$$e^{NtX} = e^{tX+tX+\ldots+tX} = e^{tX} \cdot e^{tX} \cdot \ldots \cdot e^{tX} \in G.$$

Therefore $e^{tX} \in G$ for all $t \in \mathbb{R}$ and $X \in \mathfrak{g}$, which completes the proof.

□

We now move on to the proof of part (2) of Theorem 4.1. We shall first prove part (2) in the special case where $G = GL_n(\mathbb{K})$. From there we shall introduce a lemma that will allow us to complete the full proof.

Lemma 4.6. *For sufficiently small $r > 0$, $V := \exp(B_r)$ is a neighborhood of I in $GL_n(\mathbb{K})$, and $\exp : B_r \to V$ is a homeomorphism (which is smooth and has smooth inverse).*

Proof. Notice that for all $X \in M_n(\mathbb{K})$, $d(\exp)_0(X)$ is the initial tangent vector to the path $t \mapsto e^{tX}$, that is $d(\exp)_0(X) = X$. So $d(\exp)_0$ is the identity map, which means that it is linear and invertible. Additionally, it can be shown that any power series gives a smooth function on the set of matrices with norm less than its radius of convergence, although a proof of this fact is not given here. In particular,

$$\exp : M_n(\mathbb{K}) \to M_n(\mathbb{K})$$

is smooth. Thus, by the inverse function theorem, there exists a function $\exp^{-1}$, which we denote as "log" which is smooth on a sufficiently small neighborhood $U \subset \exp(B_r)$ which contains the identity. This, combined with the observation that a sufficiently small neighborhood of I in $M_n(\mathbb{K})$ must lie in $GL_n(\mathbb{K})$ gives us our desired result. □

We introduce the following lemma to complete the proof of part (2):

Lemma 4.7. *Let $G \subset GL_n(\mathbb{K})$ be a matrix group with Lie algebra $\mathfrak{g}$. In Lemma 3.6, $r > 0$ can be chosen such that*

$$\exp(B_r \cap \mathfrak{g}) = \exp(B_r) \cap G.$$

Proof. We aim to prove containment in both directions.

For any r, $\exp(\mathfrak{g}) \subset G$ by part (1) of Theorem 4.1. Additionally, $\exp(B_r)$ is a subset of itself, thus $\exp(B_r \cap \mathfrak{g}) \subset \exp(B_r) \cap G$ for all r. Demonstrating containment in the other direction proves to be more challenging:

Similarly to the proof of part (1), we choose a subspace $\mathfrak{p}$ that is complementary to $\mathfrak{g}$. Define a function $\Phi : \mathfrak{g} \times \mathfrak{p} \to M_n(\mathbb{K})$ as $\Phi(X + Y) = e^X e^Y$ for $X \in \mathfrak{g}$ and $Y \in \mathfrak{p}$. Notice that $\Phi(X) = \exp(X)$ for all $X \in \mathfrak{g}$. Φ and exp are also similar in that the derivative of each at zero is the identity. This tells us that Φ is locally invertible by the inverse function theorem.

Assume for the sake of contradiction that the lemma is false. That is, assume that for all r

$$\exp(B_r) \cap G \not\subset \exp(B_r \cap \mathfrak{g}).$$

Then for all r, there exists a point $A \in \exp(B_r) \cap G \subset GL_n(\mathbb{K})$ such that $A \notin \exp(B_r \cap \mathfrak{g})$. Since we assume a lack of containment <u>for all</u> r, there must exist a sequence (as opposed to a single point) of non-zero vectors $\{A_1, A_2, \dots\}$ in $M_n(\mathbb{K})$ with $|A_i| \to 0$ such that $A_i \notin \mathfrak{g}$ and $\Phi(A_i) \in G$ for all i. Write $A_i = X_i + Y_i$ with $X_i \in \mathfrak{g}$ and $Y_i \in \mathfrak{p}$. For notational convenience, let $g_i := \Phi(A_i) = e^{X_i} e^{Y_i} \in G$ for all i. Notice that $e^{Y_i} = (e^{X_i})^{-1} g_i \in G$.

Consider the sphere of unit length vectors in $\mathfrak{p}$, which we will denote here as S. Since it is compact, any sequence of points $\{x_i\}$ with $x_i \in S$ will admit a subsequence that converges to some $x \in S$. This tells us that the sequence $\{\frac{Y_1}{|Y_1|}, \frac{Y_2}{|Y_2|}, \dots\}$ must have a convergent subsequence that converges to some unit-length vector $Y \in \mathfrak{p}$. Without loss of generality, re-choose the A_i's to be this subsequence so that we have a sequence that converges to Y, rather than one that sub-converges.

Let $t \in \mathbb{R}$. Since $|Y_i| \to 0$, it is possible to choose a sequence of positive integers n_i such that $n_i Y_i \to tY$. Since $e^{n_i Y_i} = (e^{Y_i})^{n_i}$ for $e^{Y_i} \in G$ and G is closed under multiplication, it is clear that $e^{n_i Y_i} \in G$. Since G is closed in $GL_n(\mathbb{K})$, we have that $e^{tY} \in G$, because it is the limit of a sequence in G. Thus $e^{tY} \in G$ for all t, which is a contradiction because $Y \notin \mathfrak{g}$.

□

We now turn to the proof of Theorem 3.1 (2).

Proof. Pick $r > 0$ as in Lemma 4.7. Then $V = \exp(B_r \cap \mathfrak{g})$ is a neighborhood of I in G because it is equal to the set $\exp(B_r) \cap G$, and $\exp(B_r)$ is open in $M_n(\mathbb{K})$ by Lemma 4.6. The restriction $\exp : B_r \cap \mathfrak{g} \to V$ is continuous. Additionally, its inverse function $\log : V \to B_r \cap \mathfrak{g}$ is continuous because it is a restriction of the continuous function $\log : \exp(B_r) \to B_r$. □

With Theorem 4.1, we are finally able to prove that matrix groups are manifolds.

Definition 4.8. A subset $X \subset \mathbb{R}^m$ is called a <u>manifold</u> of dimension n if for all $p \in X$ there exists a neighborhood V of p in X which is diffeomorphic to an open subset $U \subset \mathbb{R}^n$.

Theorem 4.9. *Any matrix group of dimension n is a manifold of dimension n.*

Proof. Let $G \subset GL_n(\mathbb{K})$ be a matrix group of dimension n with Lie algebra $\mathfrak{g}$. Choose $r > 0$ as in Theorem 4.1. Then $V := \exp(B_r \cap \mathfrak{g})$ is a neighborhood of I in G, and the restriction $\exp : B_r \cap \mathfrak{g} \to V$ is a parametrization at I. In doing this we are implicitly identifying $\mathfrak{g}$ with $\mathbb{R}^n$ by choosing a basis.

Let $g \in G$ be arbitrary. Define $\mathcal{L}_g : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ as

$$\mathcal{L}_g(A) := g \cdot A.$$

Notice that $\mathcal{L}_g$ restricts to a diffeomorphism from G to G. So $\mathcal{L}_g(V)$ is a neighborhood of $g \in G$, and $(\mathcal{L}_g \circ \exp) : B_r \cap \mathfrak{g} \to \mathcal{L}_g(V)$ is a parametrization at g. □

This proves that all matrix groups are also smooth manifolds.

5. Lie groups and Lie algebras

In the previous section, we proved that the exponential map always maps from a Lie algebra down into its associated matrix group. We also proved that all matrix groups are smooth manifolds. However, to complete the aims of this paper we are missing a link to general Lie theory. In this section, we define Lie groups and show that all matrix groups are Lie groups. Additionally, we introduce the general definition of a Lie algebra and a Lie bracket and show that what we previously called a Lie algebra in fact satisfies this definition.

Definition 5.1. A Lie group is a group that is also a smooth manifold in which the group multiplication map $\mu : G \times G \to G$ and the group inverse map are smooth maps.

Proposition 5.2. *Matrix groups are Lie groups.*

Proof. In the previous section we showed that all matrix groups are also smooth manifolds. Group multiplication on a matrix group is simply matrix multiplication, which is smooth as can be seen through repeated application of the product rule. The inverse map is obtained by matrix multiplication of a matrix and its inverse matrix, which smooth for the same reason as above. □

We now follow a chain of reasoning to motivate and define Lie brackets and Lie algebras. We first state an important theorem about Lie groups, however, the proof is omitted.

Theorem 5.3. *Let G and H be Lie groups, with G connected. A map $\rho : G \to H$ is uniquely determined by its differential $d\rho_e : T_eG \to T_eH$ at the identity.*

This theorem is a principle reason why Lie algebras are of interest: they are the tangent spaces to Lie groups at the identity, and they allow us to study the structure of a Lie group by simply looking at a linear map (the differential) of vector spaces (the Lie algebras). However, we must find out what maps between vector spaces actually arise as the differential of a group homomorphism.

Consider an arbitrary homomorphism $\rho : G \to H$. A homomorphism respects the action of a group on itself, that is if we define $m_g : G \to G$ as $m_g(X) = g \cdot X$ for $X \in G$, then $\rho(m_g(X)) = m_{\rho(g)}(\rho(X))$. However, we want a mapping from G to G that has a fixed point. A natural choice is the conjugation map $\Psi_g : G \to G$ defined by

$$\Psi_g(h) = g \cdot h \cdot g^{-1}$$

for any $g \in G$. This gives us a natural map $\Psi : G \to \mathrm{Aut}(G)$ (the set of automorphisms of G). The homomorphism ρ respects this mapping in that the diagram

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, "\Psi_g"'] & H \arrow[d, "\Psi_{\rho(g)}"] \\ G \arrow[r, "\rho"'] & H \end{tikzcd}$$

commutes. This mapping fixes the identity element in $e \in G$ which allows us to extract some of its structure by examining its differential at e. We define

$$\mathrm{Ad}_g = (d\Psi_g)_e : T_eG \to T_eG.$$

This is a representation $\mathrm{Ad} : G \to \mathrm{Aut}(T_eG)$ of the group G on its tangent space, called the adjoint representation of the group. A homomorphism respects the adjoint action of G on T_eG. In other words, for any $g \in G$, the action of Ad_g on T_eG and $\mathrm{Ad}_{\rho(g)}$ on T_eH must commute with the differential $(d\rho)_e : T_eG \to T_eH$, that is the diagram

$$\begin{tikzcd} T_eG \arrow[r, "(d\rho)_e"] \arrow[d, "\mathrm{Ad}_g"'] & T_eH \arrow[d, "\mathrm{Ad}_{\rho(g)}"] \\ T_eG \arrow[r, "(d\rho)_e"'] & T_eH \end{tikzcd}$$

commutes. An equivalent formulation is the following: for any $v \in T_eG$,

$$d\rho(\mathrm{Ad}_g(v)) = \mathrm{Ad}_{\rho(g)}(d\rho(v)).$$

This map is almost what we want, but it still uses ρ on the right hand side of the equation, and our goal is to obtain an expression purely in terms of $(d\rho)_e$. Thus we differentiate Ad to obtain $\mathrm{ad} : T_eG \to \mathrm{End}(T_eG)$ (the set of endomorphisms on T_eG). This operator maps from T_eG to a function from T_eG to itself, and thus can be regarded as a bilinear map

$$T_eG \times T_eG \to T_eG.$$

We introduce the bracket notation $[\ ,\ ]$ for this map, that is, for $X, Y \in T_eG$, we write

$$[X, Y] := \mathrm{ad}_X(Y).$$

This map ad involves only the tangent space to G at e and thus gives us our final characterization: the differential $(d\rho)_e$ of a homomorphism ρ on a Lie group G respects the adjoint action of the tangent space to G on itself. That is, the diagram

$$\begin{tikzcd} T_eG \arrow[r, "(d\rho)_e"] \arrow[d, "\mathrm{ad}_X"'] & T_eH \arrow[d, "\mathrm{ad}_{d\rho(X)}"] \\ T_eG \arrow[r, "(d\rho)_e"'] & T_eH \end{tikzcd}$$

commutes. In other words, for $X, Y \in T_eG$:

$$d\rho_e(\mathrm{ad}_X(Y)) = \mathrm{ad}_{d\rho_e(X)}(d\rho_e(Y)),$$

or, equivalently,

$$d\rho_e([X, Y]) = [d\rho_e(X), d\rho_e(Y)].$$

While this seems complicated, there is a very simple formula for this function when we restrict G to a subgroup of $GL_n(\mathbb{K})$. Let $X, Y \in gl_n(\mathbb{K})$ and $\gamma : (-\epsilon, \epsilon) \to G$ be a path in G with $\gamma(0) = I$ and $\gamma'(0) = X$. Note also that because $\gamma(0) = I$, we have that $\gamma(0)^{-1} = I$. Then $[X, Y]$ is defined as follows:

$$[X, Y] = \mathrm{ad}_X(Y) = \left.\frac{d}{dt}\right|_{t=0} (\mathrm{Ad}_{\gamma(t)}(Y)).$$

We apply the product rule to $\mathrm{Ad}_\gamma(t)(Y) = \gamma(t)Y\gamma(t)^{-1}$ to obtain

$$\begin{aligned} [X, Y] &= \gamma'(0) \cdot Y \cdot \gamma(t)^{-1} + \gamma(0) \cdot Y \cdot (\gamma(t)^{-1})' \\ &= \gamma'(0) \cdot Y \cdot \gamma(0)^{-1} + \gamma(0) \cdot Y \cdot (-\gamma(0)^{-1} \cdot \gamma'(0) \cdot \gamma(0)^{-1}) \\ &= X \cdot Y - Y \cdot X. \end{aligned}$$

Notice that the second line is obtained from applying the product rule to $\gamma(t)\gamma(t)^{-1} = I$, which yields

$$0 = \gamma'(0)\gamma(0)^{-1} + \gamma(0)(\gamma(0)^{-1})' \implies (\gamma(0)^{-1})' = -\gamma'(0) \cdot \gamma(0)^{-1} \cdot \gamma(0)^{-1}.$$

Therefore we see that the bracket operation is simply a commutator on matrix Lie groups. Another important property of the bracket operation is that it satisfies the Jacobi identity:

Proposition 5.4. *For $X, Y, Z \in T_eG$,*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

The proof of this is a simple but space consuming computation, so we omit it. This bracket operation is called the Lie bracket. Having it defined allows us to finally complete our general definition of a Lie algebra:

Definition 5.5. A Lie algebra $\mathfrak{g}$ is a vector space together with a skew-symmetric bilinear map

$$[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

that satisfies the Jacobi identity.

This definition completes our more general development of basic Lie theory and allows us to verify that the examples we have been working with are in fact Lie groups with Lie algebras. We proved above that matrix groups are Lie groups. We now consider the objects we earlier called "Lie algebras" and hope to verify that they satisfy the rigorous definition of a Lie algebra. From our prior definition, every element of $\mathfrak{g}$ is the initial tangent vector to some path in $GL_n(\mathbb{K})$. Additionally, $\mathfrak{g}$ is a subspace of $M_n(\mathbb{K})$, that is, it is closed under scalar multiplication and vector addition. Thus it is a vector space, as has been discussed before. Then, as described above, the Lie bracket operation for matrix Lie groups is simply a commutator operation and thus exists on elements of $\mathfrak{g}$. We last need to verify the Jacobi identity for the special case of matrix groups.

Proposition 5.6. *For matrix groups, the Jacobi identity is satisfied by the commutator formula for the Lie bracket.*

Proof.

$$\begin{aligned}[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]] = \\ &(X(YZ-ZY)-(YZ-ZY)X) \\ &+(Y(ZX-XZ)-(ZX-XZ)Y) \\ &+(Z(XY-YX)-(XY-YX)Z) \\ &=0.\end{aligned}$$

□

This completes the goal of this paper. We have described matrix groups and Lie groups generally and have shown that all matrix groups are Lie groups. Additionally, we have given examples of various matrix Lie groups and verified their Lie algebras. Lastly we have described a number of important properties of Lie groups, Lie algebras, and the exponential map, both in terms of matrix Lie groups and general Lie groups.

Acknowledgements. I'd like to thank my mentor, Weston Ungemach, for helping me work through an entirely new area of mathematics, making technical concepts intuitively simple, and giving me advice for writing my paper. I'd like to thank László Babai for teaching an excellent linear algebra course which made this topic more understandable. I'd also like to thank Peter May for organizing this REU and giving me the chance to learn a bunch of math I wouldn't otherwise see, as well as all the lecturers who made a diversity of topics both interesting and accessible.

References

[1] Kristopher Tapp. Matrix Groups for Undergraduates. American Mathematical Society. 2005.
[2] William Fulton and Joe Harris. Representation Theory: A First Course. Springer-Verlag New York Inc. 1991.
[3] Andrew Baker. An introduction to matrix groups and their applications. http://www.maths.gla.ac.uk/?ajb.