MATRIX GROUPS AND THEIR LIE ALGEBRAS

ELIJAH SORIA

Faculty Advisor: Dr. Keely Machmer-Wessels
Saint Mary's College Mathematics, 2016

Abstract. The purpose of this paper is to provide an introduction to Lie Theory through the use of matrix groups and examples of Lie groups and Lie algebras that pertain to matrix groups. We begin by giving background information in the mathematical areas that are used in the study of Lie groups and Lie algebras, which are mainly abstract algebra, topology, and linear algebra. Second, we introduce the reader to a formal definition of a matrix group, as well as give three examples; the general linear group, the special linear group, and the orthogonal group. Next, we define the meaning of a Lie algebra and how Lie algebras are developed through the use of matrix exponentiation, and then use this method in finding the Lie algebras of the three examples of matrix groups previously mentioned. Last, we prove the result that all matrix groups are Lie groups, and then conclude the essay with a small discussion on the importance of Lie groups and Lie algebras.

Date: May 15, 2016.

1. Introduction

The origins of Lie theory stem from the work of Felix Klein (1849-1925), who envisioned that the geometry of space is determined by the group of its symmetries. The developments in modern physics in the 19th century required an expansion in our understanding of geometry in space, and thus the notions of Lie groups and their representations expanded as well. Sophus Lie (1842-1899) began by investigating local group actions on manifolds, and his idea of looking at the group action infinitesimally began the study that would later be referred to as Lie theory [7]. During the late 19th and early 20th centuries, the importance of Lie theory became paramount due to the development of special and general relativity, and has hardly slowed down since. This paper looks to understand the structure of Lie groups and Lie algebras, giving the reader insight as to why the study of Lie theory has such strong ramifications in both the mathematical and physical worlds.

Lie theory is a field of mathematics that takes elements from topology, linear algebra, geometry, and other areas of mathematics. The study of Lie groups and Lie algebras can be convoluted and difficult to conceptualize due to the highly abstract nature of the objects themselves. Matrices, for the most part, are much more reasonable to deal with conceptually and provide a window into the world of Lie theory due to the structure of certain types of matrices under the operation of matrix multiplication. This paper looks to simplify the entry into the world of Lie theory by using matrix groups as a way to understand the structure of Lie groups and Lie algebras whilst dealing with objects that are familiar to most mathematicians. Drawing from methodologies of previous studies, such as the works of Tapp [6], this paper seeks to compile, review, and establish various introductory techniques in the study of Lie theory and matrix groups in an effort to establish and solidify preexisting works.

2. Matrix Groups

We begin by introducing some elementary definitions concerning groups. For relevant information concerning linear algebra, consult the appendix at the end of the paper.

Definition 2.1. [3] A **binary operation** $*$ on a set S is a function mapping $S \times S$ into S. For each $(a, b) \in S \times S$, we will denote the element $*((a, b))$ of S by $a * b$.

Definition 2.2. Let $*$ be a binary operation on S and let H be a subset of S. The subset H is **closed under** $*$ if for all $a, b \in H$ we also have $a * b \in H$. In this case, the binary operation on H given by restricting $*$ to H is the **induced operation** of $*$ on H.

Definition 2.3. A **group** $\langle G, * \rangle$ is a set G, closed under a binary operation $*$, such that the following properties are satisfied:

(1) (*Associativity*) For all $a, b, c \in G$,

$$(a * b) * c = a * (b * c).$$

(2) (*Identity*) There is a unique element e in G such that for all $x \in G$,

$$e * x = x * e = x.$$

(3) (*Inverse*) Corresponding to each $a \in G$, there is a unique element a' in G such that

$$a * a' = a' * a = e.$$

Definition 2.4. If a subset H of a group G is closed under the binary operation of G and if H with the induced operation from G is a group, then H is a **subgroup** of G.

Notation. Let $M_n(\mathbb{R})$ denote the set of all $n \times n$ square matrices with entries in $\mathbb{R}$.

For the purposes of this paper, we will only be studying matrices with entries in $\mathbb{R}$. The study of matrices over different fields and skew-fields, such as the complex numbers $\mathbb{C}$ and the quaternions $\mathbb{H}$, is a widely important area of study and gives rise to different matrix groups, but is beyond the scope of this paper.

Definition 2.5. [6] The **general linear group** over $\mathbb{R}$ is:

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

where I_n is the $n \times n$ identity matrix, i.e. $GL_n(\mathbb{R})$ is the collection of all invertible $n \times n$ matrices.

Theorem 2.6. *$GL_n(\mathbb{R})$ is a group with the operation being matrix multiplication.*

Proof. Let $n \in \mathbb{N}$ be arbitrary and consider $GL_n(\mathbb{R})$.
Recall that for square matrices $A, B \in M_n(\mathbb{R})$, $\det(A)\det(B) = \det(AB)$. Since for all $X, Y \in GL_n(\mathbb{R})$, $\det(X) \neq 0 \neq \det(Y)$, it follows that $\det(XY) = \det(X)\det(Y) \neq 0$. Thus, $XY \in GL_n(\mathbb{R})$, so $GL_n(\mathbb{R})$ is closed under matrix multiplication.
(*Associativity*) Let $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}) \in M_n(\mathbb{R})$. Then we have

$$\begin{aligned}
(A * B) * C &= ((a_{ij}) * (b_{ij})) * (c_{ij}) \\
&= \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) * (c_{ij}) \\
&= \left(\sum_{l=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{kl}\right) * (c_{lj})\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left(\sum_{l=1}^{n} a_{il} * \left(\sum_{k=1}^{n} b_{lk} * c_{kj} \right) \right) \\
&= (a_{ij}) * \left(\sum_{k=1}^{n} b_{lk} * c_{kj} \right) \\
&= (a_{ij}) * ((b_{ij}) * (c_{ij})) \\
&= A * (B * C).
\end{aligned}
$$

Thus, matrix multiplication is associative, so associativity holds for $GL_n(\mathbb{R})$ in particular.
(*Identity*) Let $A \in GL_n(\mathbb{R})$ where

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}
$$

Let I_n be the matrix defined as

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

It immediately follows that $AI_n = I_nA = A$, and since $\det(I_n) = 1$, $I_n \in GL_n(\mathbb{R})$.
(*Inverses*) Since for all $A \in GL_n(\mathbb{R})$, A is invertible by the definition of $GL_n(\mathbb{R})$, so A^{-1} exists. And since A^{-1} is also invertible ($(A^{-1})^{-1} = A$), $A^{-1} \in GL_n(\mathbb{R})$.
Therefore, $GL_n(\mathbb{R})$ is a group under matrix multiplication. $\square$

As the definition of matrix groups uses the definition of convergence, we define convergence explicitly here.

Definition 2.7. A sequence of real numbers (a_n) **converges** to a real number a if, for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that whenever $n \geq N$ it follows that $|a_n - a| < \epsilon$.

Definition 2.8. [4] Let A_m be a sequence of matrices in $M_n(\mathbb{R})$. We say that A_m **converges** to a matrix A if each entry of A_m converges (as $m \to \infty$) to the corresponding entry of A (i.e. if $(A_m)_{ij}$ converges to $(A)_{ij}$ for all $1 \leq i, j \leq n$).

Definition 2.9. [4] A **matrix group** is any subgroup $G \subset GL_n(\mathbb{R})$ with the following property: If A_m is any sequence of matrices in G, and A_m converges to some matrix A, then either $A \in G$, or A is not invertible.

Note that this definition is equivalent to stating that a matrix group is closed in the topological sense of the word, which is defined in Section 3, due to the fact that a matrix group contains its limit points. The group $GL_n(\mathbb{R})$ itself is a matrix group because any sequence of matrices (A_n) which converges to a matrix A has the property that A is invertible or not invertible. Two other important examples of matrix groups are the special linear group $SL_n(\mathbb{R})$ and the orthogonal group $O_n(\mathbb{R})$, which are the topics of the following paragraphs.

Definition 2.10. The **special linear group** over $\mathbb{R}$, denoted $SL_n(\mathbb{R})$, is the set of all $n \times n$ matrices with a determinant of 1, that is

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

We prove that $SL_n(\mathbb{R})$ is in fact a matrix group in Theorem 2.15; the following definitions and theorems will be useful in doing so.

Definition 2.11. A function $f : A \to \mathbb{R}$ is **continuous** at a point $c \in A$ if, for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - c| < \delta$ (and $x \in A$) it follows that $|f(x) - f(c)| < \epsilon$.

Definition 2.12. Let $f : A \to \mathbb{R}$, and let c be a limit point of the domain A. We say that $\lim_{x \to c} f(x) = L$ provided that, for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - c| < \delta$ it follows that $|f(x) - L| < \epsilon$.

Theorem 2.13. [1] *Let $f : A \to \mathbb{R}$, and let $c \in A$ be such that there exists some sequence (x_n) where $x_n \in A$ for all $n \in \mathbb{N}$ and $x_n \to c$. The function f is continuous at c if and only if any one of the following holds true:*

(1) $f(x_n) \to f(c)$;
(2) $\lim_{x \to c} f(x) = f(c)$.

Proof. (1) Let $f : A \to \mathbb{R}$, and let $c \in A$ be such that the sequence (x_n) where $x_n \in A$ for all $n \in \mathbb{N}$ has the property that $x_n \to c$. Suppose that f is continuous at c and let $\epsilon > 0$ be arbitrary. Since f is continuous at c, there exists some $\delta > 0$ such that whenever $x \in A$ and $|x - c| < \delta$, we are guaranteed that $|f(x) - f(c)| < \epsilon$. Towards contradiction, assume that $\lim_{n\to\infty} f(x_n) \neq f(c)$. Thus there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - c| < \delta$ and $|f(x_n) - f(c)| \geq \epsilon$, which is a contradiction to our assumption that $|f(x) - f(c)| < \epsilon$ for all $x \in A$. Thus $f(x_n) \to f(c)$ by contradiction.

Now suppose that f is not continuous at c. This implies that there exists some $\epsilon_0 > 0$ such that for all $\delta > 0$, there exists some $x_0 \in A$ such that $|x_0 - c| < \delta$ and $|f(x_0) - f(c)| \geq \epsilon_0$. For each $n \in \mathbb{N}$, let $\delta_n = 1/n$. This implies that there exists some $x_n \in A$ such that $|x_n - c| < \delta_n$ and $|f(x_n) - f(c)| \geq \epsilon_0$. Clearly, the sequence (x_n) has the property that $x_n \to c$, as for all $\epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n \geq N$, it follows that

$|x_n - c| < \delta_n < \epsilon$. Thus, the sequence (x_n) has the property that $x_n \to c$ and for all $N' \in \mathbb{N}$ there exists some $n_0 \geq N'$ such that $|f(x_{n_0}) - f(c)| \geq \epsilon_0$. This proves that if $x_n \to c$ (with $x_n \in A$), then $f(x_n) \to f(c)$. Thus, f is continuous at c by the contrapositive. Therefore, statement (1) of Theorem 2.13 holds if and only if f is continuous at c.
(2) We show that statement (1) is equivalent to statement (2). Using Definition 2.12, $\lim_{x \to c} f(x) = f(c)$ states that for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - c| < \delta$ it follows that $|f(x) - f(c)| < \epsilon$. This is equivalent to the statement "if $x_n \to c$ (with $x_n \in A$), then $f(x_n) \to f(c)$" using Definition 2.11. Therefore, statement (1) is equivalent to statement (2), proving Theorem 2.13 in its entirety. □

Theorem 2.14. *The determinant function* $\det : M_n(\mathbb{R}) \to \mathbb{R}$ *is continuous.*

Proof. The proof will proceed by induction on n. First, let $n = 1$. Since the determinant of a 1×1 real matrix is simply the entry itself, the determinant function is continuous as it just outputs the entry itself. Thus, the determinant function from $M_1(\mathbb{R})$ to $\mathbb{R}$ is continuous. Now, assume that the determinant function from $M_n(\mathbb{R})$ to $\mathbb{R}$ is continuous, with the goal of proving that the determinant function from $M_{n+1}(\mathbb{R})$ to $\mathbb{R}$ is continuous. Let $A \in M_{n+1}(\mathbb{R})$ where

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & a_{1,n+1} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & a_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n+1,1} & a_{n+1,2} & \cdots & a_{n+1,n} & a_{n+1,n+1} \end{bmatrix}.$$

By the definition of the determinant, $\det A = \sum_{i=1}^{n+1} (-1)^{i+j} a_{i,j} M_{i,j}$ where $M_{i,j}$ is the minor of the i–j^{th} entry. Since $M_{i,j}$ is the determinant of a $n \times n$ matrix for each $i, j \in \{1, 2, \cdots, n+1\}$, $\det A$ is simply a sum of continuous functions multiplied by a real number, so $\det A$ is continuous. Thus, the determinant function from $M_{n+1}(\mathbb{R})$ to $\mathbb{R}$ is continuous, proving the original statement by induction. □

Theorem 2.15. $SL_n(\mathbb{R})$ *is a matrix group.*

Proof. Let $n \in \mathbb{N}$ be arbitrary. We first prove that $SL_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$. Let $A, B \in SL_n(\mathbb{R})$. Since $\det(AB) = \det(A)\det(B)$ and $\det(A) = 1 = \det(B)$ since $A, B \in SL_n(\mathbb{R})$, it follows that $\det(AB) = \det(A)\det(B) = 1(1) = 1$. Thus $AB \in SL_n(\mathbb{R})$, so $SL_n(\mathbb{R})$ is closed under matrix multiplication. Also, since $\det(I_n) = 1$, $I_n \in SL_n(\mathbb{R})$. Lastly, Since $\det(AA^{-1}) = \det(I_n) = 1 = \det(A)\det(A^{-1})$ and $\det(A) = 1$, it follows that $\det(A^{-1}) = 1$, so $A^{-1} \in SL_n(\mathbb{R})$. Thus $SL_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.
Let (A_m) be a sequence of matrices where $A_m \in SL_n(\mathbb{R})$ for each $m \in \mathbb{N}$ and $A_m \to A$. Since $\det A_m = 1$ for all $m \in \mathbb{N}$ and since the determinant is a continuous function by Theorem

2.14, it follows by Theorem 2.13 that $\det A = 1$ as well. Therefore, $A \in SL_n(\mathbb{R})$, so $SL_n(\mathbb{R})$ is a matrix group. $\square$

To understand the orthogonal group $O_n(\mathbb{R})$, we will first cover what it means to be orthogonal.

Definition 2.16. [6] The **standard inner product on** $\mathbb{R}^n$ is the function from $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by:

$$\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle_{\mathbb{R}} := x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n.$$

Definition 2.17. [6] The **standard norm on** $\mathbb{R}^n$ is the function from $\mathbb{R}^n \to \mathbb{R}^+$ defined by

$$|x|_{\mathbb{R}} = \sqrt{\langle x, x \rangle_{\mathbb{R}}}.$$

Definition 2.18. [6] Vectors $x, y \in \mathbb{R}^n$ are called **orthogonal** if $\langle x, y \rangle = 0$.

Definition 2.19. A vector $x \in \mathbb{R}^n$ is called a **unit vector** $|x| = 1$.

Definition 2.20. A matrix $A \in M_n(\mathbb{R})$ is said to be **orthogonal** if the column vectors of A are orthogonal unit vectors.

Note that this definition is equivalent to stating that $\langle xA, yA \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{R}$. This condition is known as an isometry condition, meaning that an orthogonal matrix is a distance preserving linear transformation. It follows from the above definition alone that for all orthogonal matrices $A \in M_n(\mathbb{R})$, $A^T A = I_n = AA^T$ where A^T is the transpose of matrix A, that is, if $a_{i,j}$ is the entry of A in the i^{th} row and j^{th} column of A, then $a_{i,j}$ is the entry in the j^{th} row and i^{th} column of A^T.

The following definition generalizes orthogonality over different fields.

Definition 2.21. [6] The **orthogonal group over** $\mathbb{R}$ is defined as

$$O_n(\mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid \langle xA, yA \rangle = \langle x, y \rangle \text{ for all } x, y \in \mathbb{R}^n\}.$$

We reserve the proof that $O_n(\mathbb{R})$ is a matrix group until the end of Section 3. The following definitions and theorems will be useful throughout the paper in understanding $O_n(\mathbb{R})$.

Definition 2.22. [6] A set $\{x_1, x_2, \ldots, x_n\}$ of $\mathbb{R}^n$ is called **orthonormal** if $\langle x_i, x_j \rangle = 1$ when $i = j$ and $\langle x_i, x_j \rangle = 0$ when $i \neq j$.

As an example, an orthonormal set of $\mathbb{R}^n$, is the set

$$\mathcal{B} = \{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)\}.$$

The set $\mathcal{B}$ is called the *standard orthonormal basis* for $\mathbb{R}^n$.

Lemma 2.23. *If $A, B \in M_n(\mathbb{R})$, then $(AB)^T = B^T A^T$.*

Proof. Let $A, B \in M_n(\mathbb{R})$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

The following equalities hold.

$$\begin{aligned}
(AB)^T &= \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \right)^T \\
&= \left(\begin{bmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots a_{1n}b_{n2} & \cdots & a_{11}b_{1n} + \cdots + a_{1n}b_{nn} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots a_{2n}b_{n2} & \cdots & a_{21}b_{1n} + \cdots + a_{2n}b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}b_{11} + \cdots + a_{nn}b_{n1} & a_{n1}b_{12} + \cdots a_{nn}b_{n2} & \cdots & a_{n1}b_{1n} + \cdots + a_{nn}b_{nn} \end{bmatrix} \right)^T \\
&= \begin{bmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{21}b_{11} + \cdots + a_{2n}b_{n1} & \cdots & a_{n1}b_{11} + \cdots + a_{nn}b_{n1} \\ a_{11}b_{12} + \cdots a_{1n}b_{n2} & a_{21}b_{12} + \cdots a_{2n}b_{n2} & \cdots & a_{n1}b_{12} + \cdots a_{nn}b_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{1n} + \cdots + a_{1n}b_{nn} & a_{21}b_{1n} + \cdots + a_{2n}b_{nn} & \cdots & a_{n1}b_{1n} + \cdots + a_{nn}b_{nn} \end{bmatrix} \\
&= \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{n1} \\ b_{12} & b_{22} & \cdots & b_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1n} & b_{2n} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \\
&= B^T A^T
\end{aligned}$$

□

Theorem 2.24. *If $A \in M_n(\mathbb{R})$, then $(A^n)^T = (A^T)^n$.*

Proof. Let $A \in M_n(\mathbb{R})$. We proceed by induction on n. First, let $n = 1$. Clearly, $(A^1)^T = A^T = (A^T)^1$, so this case holds. Now, suppose that $(A^n)^T = (A^T)^n$ holds for some $n \in \mathbb{N}$. By Lemma 2.23, it follows that $(A^T)^{n+1} = (A^T)^n A^T = (A^n)^T A^T = (AA^n)^T = (A^{n+1})^T$. Therefore, $(A^n)^T = (A^T)^n$ is true by the principle of mathematical induction. □

Definition 2.25. [6] If $A \in M_n(\mathbb{R})$, define $R_A : \mathbb{R}^n \to \mathbb{R}^n$ and $L_A : \mathbb{R}^n \to \mathbb{R}^n$ such that for all $x \in \mathbb{R}^n$,

$$R_A(x) = x \cdot A \text{ and } L_A(x) = (A \cdot x^T)^T.$$

Theorem 2.26. *For all $A \in GL_n(\mathbb{R})$, $A \in O_n(\mathbb{R})$ if and only if $A \cdot A^T = I_n$.*

Proof. Let $A = [a_{ij}]_n \in GL_n(\mathbb{R})$ be arbitrary.
($\Longrightarrow$) Suppose that $A \in O_n(\mathbb{R})$. Since $\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)\}$ is an orthonormal basis for $\mathbb{R}^n$ and $\langle x \cdot A, y \cdot A\rangle = \langle x, y\rangle$, it follows that

$$\{R_A(e_1), R_A(e_2), \ldots, R_A(e_n)\}$$

is an orthonormal set of vectors. $\{R_A(e_1), R_A(e_2), \ldots, R_A(e_n)\}$ is precisely the set of row vectors of A, where $R_A(e_i)$ is the i^{th} row of A. Notice that

$$\begin{aligned}
(A \cdot A^T)_{ij} &= (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } A^T) \\
&= (\text{row } i \text{ of } A) \cdot (\text{row } j \text{ of } A) \\
&= \langle(\text{row } i \text{ of } A), (\text{row } j \text{ of } A)\rangle.
\end{aligned}$$

Thus, $(A \cdot A^T)_{ij} = 1$ when $i = j$ as $\langle(\text{row } i \text{ of } A), (\text{row } i \text{ of } A)\rangle = \langle R_A(e_i), R_A(e_i)\rangle = 1$ and $(A \cdot A^T)_{ij} = 0$ when $i \neq j$ as $\langle(\text{row } i \text{ of } A), (\text{row } j \text{ of } A)\rangle = \langle R_A(e_i), R_A(e_j)\rangle = 0$. Thus, $A \cdot A^T = I_n$.
($\Longleftarrow$) Suppose that $A \cdot A^T = I_n$. This implies that $\langle R_A(e_i), R_A(e_i)\rangle = 1$ and $\langle R_A(e_i), R_A(e_j)\rangle = 0$ when $i \neq j$. Let $x, y \in \mathbb{R}^n$ be arbitrary where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$. We see that

$$\begin{aligned}
\langle x \cdot A, y \cdot A\rangle &= \langle R_A(x), R_A(y)\rangle \\
&= \left\langle \sum_{i=1}^{n} x_i(\text{row } i \text{ of } A), \sum_{j=1}^{n} y_j(\text{row } j \text{ of } A) \right\rangle \\
&= \sum_{i=1}^{n} x_i\langle(\text{row } i \text{ of } A), (\text{row } j \text{ of } A)\rangle y_i \\
&= x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n \\
&= \langle x, y\rangle.
\end{aligned}$$

Therefore, $A \in O_n(\mathbb{R})$, proving the statement. $\square$

3. Topology of Matrix Groups

The goal of this section is to relate matrix groups with topologies, specifically with the Euclidean topology $\mathbb{R}^{n^2}$. To do so, a background in general topology and metric spaces is needed.

Definition 3.1. [5] A **topology** on a non-empty set X is a collection $\mathcal{T}$ of subsets of X having the following properties:

(1) $\emptyset$ and X are in $\mathcal{T}$.
(2) The union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$.
(3) The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

A set X for which a topology $\mathcal{T}$ has been specified is called a **topological space**, denoted $(X, \mathcal{T})$.

Definition 3.2. [5] If $(X, \mathcal{T})$ is a topological space, we say that a subset U of X is an **open set** of X if U belongs to the collection $\mathcal{T}$. Similarly, if U is an open set containing some point $x \in X$, then we say that U is a **neighborhood** of x.

Definition 3.3. A subset A of a topological space $(X, \mathcal{T})$ is said to be **closed** if the set $X - A$ is open in $\mathcal{T}$.

We now turn our attention to a basis for a topology, which will help us create a better understanding of building topologies on sets.

Definition 3.4. [5] If X is a non-empty set, a **basis** for a topology on X is a collection $\mathcal{B}$ of subsets of X such that

(1) For each $x \in X$, there exists some $B \in \mathcal{B}$ such that $x \in B$.
(2) If $x \in X$ such that $x \in B_1 \cap B_2$ for some $B_1, B_2 \in \mathcal{B}$, there exists a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

If we have some topology $\mathcal{T}$ and $\mathcal{B}$ is a basis for $\mathcal{T}$, then $\mathcal{T}$ is the collection of all arbitrary unions of elements of $\mathcal{B}$. A simple example of a topology and a basis is the real numbers, which is explained in the following example.

Example 3.5. *The collection $\mathcal{B}$ of open intervals in $\mathbb{R}$, precisely defined as*

$$\mathcal{B} = \{(a, b) \mid \textit{where } a < b \textit{ and } a, b \in \mathbb{R}\},$$

is a basis for a topology on $\mathbb{R}$.

Proof. Let $\mathcal{B}$ be the collection of all open subsets of $\mathbb{R}$, that is

$$\mathcal{B} = \{(a, b) \mid \text{where } a < b\}.$$

To satisfy condition 1 of a basis, it is easy to see that for any $x \in \mathbb{R}$, the open set $(x-1, x+1) \in \mathcal{B}$ contains x. For condition 2, let $x \in X$ be such that $x \in (a_1, b_1) \cap (a_2, b_2)$ for some $(a_1, b_1), (a_2, b_2) \in \mathcal{B}$. Without loss of generality, assume that $a_1 < a_2$ and $b_1 < b_2$. Thus $x \in (a_1, b_1) \cap (a_2, b_2) = (a_2, b_1) \in \mathcal{B}$, so $\mathcal{B}$ is in fact a basis. $\square$

The union of the elements of $\mathcal{B}$ gives us the standard topology on $\mathbb{R}$. The standard topology on $\mathbb{R}$ is one of the most fundamental examples of a topology, and will be used to associate matrix groups with topologies in the upcoming sections. Now we look to classify distance within topologies, specifically $\mathbb{R}^n$, through the use of a function called a metric.

Definition 3.6. [5] A **metric** on a non-empty set X is a function

$$d : X \times X \to R$$

having the following properties:

(1) $d(x, y) \geq 0$ for all $x, y \in X$; $d(x, y) = 0$ if and only if $x = y$.
(2) $d(x, y) = d(y, x)$ for all $x, y \in X$.
(3) (Triangle inequality) $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$.

The ultimate goal is to construct a topology using a metric, which will require generating a basis using the following definition.

Definition 3.7. Let d be a metric on a set X and let $x \in X$. Given $\epsilon > 0$, the set

$$B_d(x, \epsilon) = \{y \mid d(x, y) < \epsilon \text{ and } y \in X\}$$

is called the **ϵ-ball centered at** x.

Definition 3.8. [5] If d is a metric on the set X, then the collection of all ϵ-balls $B_d(x, \epsilon)$, for $x \in X$ and $\epsilon > 0$, is a basis for a topology on X, called the **metric topology** induced by d.

One important example of a metric is called the Euclidean metric on $\mathbb{R}^n$, which will be useful when relating real matrix groups of size n with the space $\mathbb{R}^{n^2}$.

Example 3.9. *Let* $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. *Let* $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *be the function defined as*

$$d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

is a metric on $\mathbb{R}^n$ *called the Euclidean metric.*

The proof of the example will require the use of the following lemma.

Lemma 3.10. [5] *For all $x, y \in \mathbb{R}^n$, $|x \cdot y| \leq \|x\|\|y\|$.*

Proof. First, suppose that $x = \bar{0}$ or $y = \bar{0}$. We see that $|x \cdot y| = 0 \leq 0 = \|x\|\|y\|$, so our claim holds in this case. Now, suppose that $x \neq \bar{0}$ and $y \neq \bar{0}$. Let $a_o = \frac{1}{\|x\|}$ and $b_o = \frac{1}{\|y\|}$. First note that $0 \leq \|ax \pm by\|$ for all $a, b \in \mathbb{R}$. Through the use of this inequality after squaring both sides, the following inequalities hold.

$$\begin{aligned}
0 &\leq \left\| \frac{1}{\|x\|}x \pm \frac{1}{\|y\|}y \right\|^2 \\
&= \left(\sqrt{\left(\frac{1}{\|x\|}x_1 \pm \frac{1}{\|y\|}y_1 \right)^2 + \cdots + \left(\frac{1}{\|x\|}x_n \pm \frac{1}{\|y\|}y_n \right)^2} \right)^2 \\
&= \frac{1}{\|x\|^2}x_1^2 \pm \frac{2}{\|x\|\|y\|}x_1y_1 + \frac{1}{\|y\|^2}y_1^2 + \cdots + \frac{1}{\|x\|^2}x_n^2 \pm \frac{2}{\|x\|\|y\|}x_ny_n + \frac{1}{\|y\|^2}y_n^2 \\
&= \frac{1}{\|x\|^2}(x_1^2 + \cdots + x_n^2) \pm \frac{2}{\|x\|\|y\|}(x_1y_1 + \cdots + x_ny_n) + \frac{1}{\|y\|^2}(y_1^2 + \cdots + y_n^2) \\
&= \frac{1}{\|x\|^2}\|x\|^2 \pm \frac{2}{\|x\|\|y\|}(x \cdot y) + \frac{1}{\|y\|^2}\|y\|^2 \\
&= 2 \pm \frac{2}{\|x\|\|y\|}(x \cdot y).
\end{aligned}$$

This implies that $\mp\frac{1}{\|x\|\|y\|}(x \cdot y) \leq 1$, which implies that $|x \cdot y| \leq \|x\|\|y\|$, proving the statement. $\square$

We will now use this lemma in Example 3.9, shown below. Example 3.9 is stated again for convenience.

Example 3.9. *Let $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Let $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the function defined as*

$$d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

is a metric on $\mathbb{R}^n$ called the Euclidean metric.

Proof. Let $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the function defined as

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

We will show that d satisfies the three conditions given in Definition 3.6

(1) Let $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Since $(x_i - y_i)^2 \geq 0$ for all $i \in \{1, 2, \ldots, n\}$, it immediately follows that $d(x, y) \geq 0$. If $x = y$, then

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

$$= \sqrt{(x_1 - x_1)^2 + (x_2 - x_2)^2 + \cdots + (x_n - x_n)^2}$$
$$= 0.$$

If $d(x, y) = 0$, then $(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2 = 0$, so $(x_i - y_i) = 0$ for all $1 \leq i \leq n$, implying that $x_i = y_i$ for all $1 \leq n \leq n$. Thus $x = y$.

(2) We have

$$\begin{aligned} d(x, y) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \\ &= \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2} \\ &= d(y, x). \end{aligned}$$

(3) Let us consider $\|x + y\|^2$. Recall the definition of a standard inner product from Definition 2.16; the following equalities hold.

$$\begin{aligned} \|x + y\|^2 &= (x + y) \cdot (x + y) \\ &= (x_1 + y_1)^2 + \cdots + (x_n + y_n)^2 \\ &= x_1^2 + 2x_1y_1 + y_1^2 + \cdots + x_n^2 + 2x_ny_n + y_n^2 \\ &= (x_1^2 + \cdots + x_n^2) + 2(x_1y_1 + \cdots + x_ny_n) + (y_1^2 + \cdots + y_n^2) \\ &= (x \cdot x) + 2(x \cdot y) + (y \cdot y) \\ &= \|x\|^2 + 2(x \cdot y) + \|y\|^2. \end{aligned}$$

Through our knowledge of absolute values and though the use of Lemma 3.10, we see that this implies

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2(x \cdot y) + \|y\|^2 \\ &\leq \|x\|^2 + 2|x \cdot y| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Taking the square root of both sides of the inequality above, we get $\|x + y\| \leq \|x\| + \|y\|$.

Now let $z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$. We see that $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.

Therefore, d is a metric on $\mathbb{R}^n$ □

The metric space induced by the Euclidean metric on $\mathbb{R}^n$ is known as the Euclidean topology on $\mathbb{R}^n$. The Euclidean topology on $\mathbb{R}^n$ is the topology that we need to relate $\mathbb{R}^n$ and $M_m(\mathbb{R})$ with each other.

To relate $\mathbb{R}^n$ and $M_m(\mathbb{R})$ to each other, we can create a one-to-one correspondence between $\mathbb{R}^{n^2}$ and $M_n(\mathbb{R})$ by creating the bijective function $\phi : \mathbb{R}^{n^2} \to M_n(\mathbb{R})$ defined as

$$\phi(x) = \phi((x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{nn}) = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix}.$$

Thus, we can actually talk about the Euclidean space $\mathbb{R}^{n^2}$ and still work with matrices, which allows us to study the geometry and topologies of matrix groups through the use of the Euclidean metric and the subspace topology applied to the Euclidean topology.

The definition of the subspace topology is given below.

Definition 3.11. Let $(X, \mathcal{T})$ be a topological space. If Y is a subset of X, the collection

$$\mathcal{T}_Y = \{Y \cap U \mid U \in \mathcal{T}\}$$

is a topology on Y, called the **subspace topology**. With this topology, Y is called a **subspace** of X.

Thus, the topologies of matrix groups are structurally equivalent to subspace topologies of the Euclidean topology.

With an understanding of the topology of matrix groups, we are poised to understand the proof that $O_n(\mathbb{R})$ is a matrix group. The following definition and theorems will be used in the proof that $O_n(\mathbb{R})$ is a matrix group.

Theorem 3.12. [5] *Let $(Y, \mathcal{T}_Y)$ be a subspace of $(X, \mathcal{T})$. Then a set A is closed in Y if and only if it equals the intersection of a closed set of X with Y.*

Proof. Let $(Y, \mathcal{T}_Y)$ be a subspace of $(X, \mathcal{T})$. Let A be closed in Y. Thus, $Y - A \in \mathcal{T}_Y$, so $Y - A = U \cap Y$ where $U \in \mathcal{T}$. Since $X - U$ is closed in X and $A = Y \cap (X - U)$, A is the intersection of Y with a closed set of X.

Now let $A \subset Y$ be such that $A = C \cap Y$ where C is closed in X. Then $X - C \in \mathcal{T}$, so $Y \cap (X - C) \in \mathcal{T}_Y$. Since $(X - C) \cap Y = Y - A$, $Y - A \in \mathcal{T}_Y$, so A is closed in Y, as desired. □

Definition 3.13. [5] Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces. A function $f : X \to Y$ is said to be **continuous** if for each open subset V of Y, the set $f^{-1}(V)$ is an open subset of X.

Theorem 3.14. [5] *Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces; let $f : X \to Y$. If f is continuous, then for every closed subset B of Y, the set $f^{-1}(B)$ is closed in X.*

Proof. Let $(X, \mathcal{T})$ and $(Y, \mathcal{T}')$ be topological spaces and let $f : X \to Y$. Suppose that f is continuous and let B be a closed set of Y. Since $Y - B \in \mathcal{T}'$, and f is continuous, $f^{-1}(Y - B)$ is open in $\mathcal{T}$. Since $f^{-1}(Y - B) = f^{-1}(Y) - f^{-1}(B) = X - f^{-1}(B)$, it follows that $f^{-1}(B)$ is closed in X, as desired. □

Theorem 3.15. *$O_n(\mathbb{R})$ is a matrix group.*

Proof. Let $n \in \mathbb{N}$ be fixed; we will first show that $O_n(\mathbb{R})$ is a group. First, note that the identity matrix I_n is in $O_n(\mathbb{R})$ since for any $x, y \in \mathbb{R}^n$, $\langle xI_n, yI_n \rangle = \langle x, y \rangle$. Second, note that $O_n(\mathbb{R})$ inherits inverses from $GL_n(\mathbb{R})$ since for any $M \in O_n(\mathbb{R})$, $M^{-1} = M^T$, so M^{-1} is orthogonal since $M^{-1}(M^{-1})^T = M^T(M^T)^T = I_n$ and $(M^{-1})^T M^{-1} = (M^T)^T M^T = I_n$. Lastly, orthogonal matrices are closed under multiplication since for any $A, B \in O_n(\mathbb{R})$, $(AB)^T = B^T A^T$ and thus $AB(AB)^T = ABB^T A^T = AA^T = I_n$ and $(AB)^T AB = B^T A^T AB = B^T B = I_n$. Thus, $O_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.

Since for all matrices $N, M \in M_n(\mathbb{R})$, $\det(N) = \det(N^T)$ and $\det(N)\det(M) = \det(NM)$, it follows that if $A \in O_n(\mathbb{R})$, then $\det(A)^2 = \det(AA) = \det(AA^T) = \det(I_n) = 1$, so $\det(A) = \pm 1$. Now, define $T : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ by $T(X) = XX^T$ for all $X \in GL_n(\mathbb{R})$. It is clear that T is continuous since, for all $X \in GL_n(X)$ where $X = [x_{ij}]$, the $i - j^{\text{th}}$ entry of $T(X)$ is simply $\sum_{k=1}^{n} x_{ik}x_{jk}$, which is a polynomial function in $\mathbb{R}$. Thus, since $T^{-1}(\{I_n\}) = O_n(\mathbb{R})$ and one-point sets are closed in $\mathbb{R}^{n^2}$, $\{I_n\}$ is closed in $GL_n(\mathbb{R})$ by Theorem 3.12, so it follows by Theorem 3.14 that $O_n(\mathbb{R})$ is closed in $GL_n(\mathbb{R})$. Therefore, $O_n(\mathbb{R})$ is a matrix group. □

4. Lie Algebras

Definition 4.1. [6] Let $M \subset \mathbb{R}^m$ and let $x \in M$. The **tangent space** to M at x is defined as

$$T_x M := \{\gamma'(0) \mid \gamma : (-\epsilon, \epsilon) \to M \text{ is differentiable with } \gamma(0) = x\}.$$

The function $\gamma : (-\epsilon, \epsilon) \to M$ in the previous definition is referred to commonly as a *path* through the point x. Thus, the tangent space to $M \subset \mathbb{R}^m$ at x is the collection of slopes of all paths such that each component function of γ is differentiable from $(-\epsilon, \epsilon)$ to $\mathbb{R}$.

Due to the correlation stated in Section 3 between $M_n(\mathbb{R})$ and $\mathbb{R}^{n^2}$, we are able to consider matrix groups as subsets of the Euclidean space. This gives us the ability to talk about tangent spaces of matrix groups, which gives us the definition of a Lie algebra, given below.

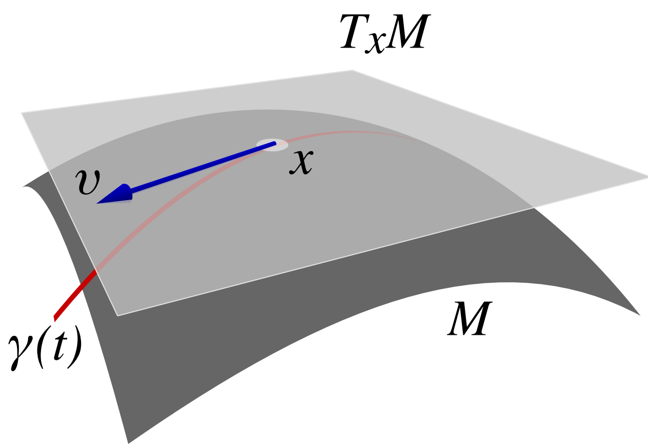

FIGURE 4.1. A visualization of the tangent space T_xM.

Definition 4.2. [6] The **Lie algebra** of a matrix group $G \subset GL_n(\mathbb{R})$ is the tangent space to G at the identity matrix I_n. We denote the Lie Algebra of G as $\mathfrak{g} := \mathfrak{g}(G) := T_{I_n}G$.

In Theorem 4.4, we prove that the Lie algebras of matrix groups are subspaces of $M_n(\mathbb{R})$. To do so, we will use the product rule for paths in $M_n(\mathbb{R})$, which is the subject of the following theorem.

Theorem 4.3. [6] *If $\gamma, \beta : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ are differentiable, then the product path $(\gamma \cdot \beta)(t) := \gamma(t) \cdot \beta(t)$ is differentiable. Furthermore,*

$$(\gamma \cdot \beta)'(t) = \gamma(t) \cdot \beta'(t) + \gamma'(t) \cdot \beta(t).$$

Proof. Let $\gamma, \beta : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ be differentiable. When $n = 1$, then we have the product rule from calculus. Since

$$((\gamma \cdot \beta)(t))_{ij} = \sum_{l=1}^{n} \gamma(t)_{il} \cdot \beta(t)_{lj}$$

and $\gamma(t)_{il} \cdot \beta(t)_{lj}$ is a product of functions from $(-\epsilon, \epsilon)$ to $\mathbb{R}$, it follows that

$$\begin{aligned} ((\gamma \cdot \beta)'(t))_{ij} &= \sum_{l=1}^{n} \gamma(t)_{il} \cdot \beta'(t)_{lj} + \gamma'(t)_{il} \cdot \beta(t)_{lj} \\ &= (\gamma(t) \cdot \beta'(t))_{ij} + (\gamma'(t) \cdot \beta(t))_{ij}. \end{aligned}$$

□

Theorem 4.4. [6] *The Lie algebra $\mathfrak{g}$ of a matrix group $G \subset GL_n(\mathbb{R})$ is a real subspace of $M_n(\mathbb{R})$.*

Proof. Let $G \subset GL_n(\mathbb{R})$ be an arbitrary matrix group. To prove that $\mathfrak{g}$ is a subspace of $M_n(\mathbb{R})$, we need to prove that $\mathfrak{g}$ is closed under scalar multiplication and matrix addition. Thus, let $\lambda \in \mathbb{R}$ and and let $A \in \mathfrak{g}$, so $A = \gamma'(0)$ where $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^n$ is a differentiable path such that $\gamma(0) = I_n$. Let $\sigma : (-\lambda\epsilon, \lambda\epsilon) \to \mathbb{R}^n$ be the path defined as $\sigma(t) := \gamma(\lambda \cdot t)$ for all $t \in (-\lambda\epsilon, \lambda\epsilon)$. Since $\sigma'(t) = \lambda \cdot \gamma'(\lambda \cdot t)$, it follows that $\sigma'(0) = \lambda \cdot A$. Thus, since $\sigma(0) = \gamma(\lambda \cdot 0) = I_n$, we can conclude that $\lambda \cdot A \in \mathfrak{g}$.

Next, let $A, B \in \mathfrak{g}$. Thus, $A = \gamma'(0)$ and $B = \beta'(0)$ where $\gamma : (-\epsilon_1, \epsilon_1) \to \mathbb{R}^n$ and $\beta : (-\epsilon_2, \epsilon_2) \to \mathbb{R}^n$ are differentiable paths such that $\gamma(0) = \beta(0) = I_n$. Let $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. Let $\pi : (-\epsilon, \epsilon) \to \mathbb{R}^n$ be the product path defined as $\pi(t) := \gamma(t) \cdot \beta(t)$ for all $t \in (-\epsilon, \epsilon)$. By Theorem 4.3, we know that π is a differentiable path that lies in G with

$$\pi'(0) = \gamma(0) \cdot \beta'(0) + \gamma'(0) \cdot \beta(0) = I_n \cdot B + A \cdot I_n = A + B.$$

Therefore $A + B \in \mathfrak{g}$, proving that $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{R})$. □

Since Lie algebras are vector spaces over $\mathbb{R}$, we are able to classify matrix groups and their Lie algebras according to their basis.

Definition 4.5. The **dimension** of a matrix group G is the dimension of its Lie algebra.

In order to give examples of Lie algebras $\mathfrak{g}$ of matrix group $G \subset GL_n(\mathbb{R})$, we must construct paths $\gamma_A : (-\epsilon, \epsilon) \to G$ for each $A \in G$ such that $\gamma(0) = I_n$ and $\gamma'(0) = A$. The simplest way to accomplish this is to use a function called *matrix exponentiation*, which requires a few definitions to understand the beautiful simplicity of the concept.

Definition 4.6. [6] A **vector field** is a continuous function $F : \mathbb{R}^m \to \mathbb{R}^m$.

Definition 4.7. [6] An **integral curve** of a vector field $F : \mathbb{R}^m \to \mathbb{R}^m$ is a path $\alpha : (-\epsilon, \epsilon) \to \mathbb{R}^m$ such that $\alpha'(t) = F(\alpha(t))$ for all $t \in (-\epsilon, \epsilon)$.

Intuitively, the vector field $F : \mathbb{R}^m \to \mathbb{R}^m$ gives the value of the tangent vector to every point on the path $\alpha : (-\epsilon, \epsilon) \to \mathbb{R}^m$. Surprisingly, matrix exponentiation gives us an integral curve for every element in the Lie algebra of a matrix group.

As matrix exponentiation is defined by power series of matrices, we will introduce terms and results that refer to series in $M_n(\mathbb{R})$.

Definition 4.8. Let $A \in M_n(\mathbb{R})$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

The **Euclidean norm** of A, denoted $|A|$, is defined as

$$|A| = \sqrt{(a_{11})^2 + \cdots + (a_{1n})^2 + (a_{21})^2 + \cdots + (a_{2n})^2 + \cdots + (a_{n1})^2 + \cdots + (a_{nn})^2}.$$

The Euclidean norm of a matrix is better understood as the square root of the sum of squares of the entries of the matrix.

Definition 4.9. Let $A_i \in M_n(\mathbb{R})$ for all $i \in \mathbb{Z}^*$. We say that the series $\Sigma_{i=0}^{\infty} A_i = A_0 + A_1 + A_2 + \cdots$ **converges (absolutely)** if, for all $i, j \in \mathbb{Z}^*$, $(A_0)_{ij} + (A_1)_{ij} + (A_2)_{ij} + \cdots$ converges (absolutely) to some $(A)_{ij} \in \mathbb{R}$. This is denoted as $\Sigma_{i=0}^{\infty} A_i = A$.

To prove a result concerning absolute convergence of a series, the following lemma will be used.

Lemma 4.10. *For all $X, Y \in M_n(\mathbb{R})$, $|XY| \leq |X| \cdot |Y|$.*

Proof. Let $X, Y \in M_n(\mathbb{R})$ be arbitrary. Recall that for all $x, y \in \mathbb{R}^n$, $|\langle x, y\rangle| \leq |x| \cdot |y|$ (the Schwarz inequality). Using the Schwarz inequality, it follows that for all indices i, j,

$$\begin{aligned}
|(XY)_{ij}|^2 &= \left|\sum_{l=1}^{n} X_{il} Y_{lj}\right|^2 \\
&= \left|\langle (\text{row } i \text{ of } X), (\text{column } j \text{ of } Y)^T \rangle\right|^2 \\
&\leq \left|(\text{row } i \text{ of } X)\right|^2 \cdot \left|(\text{column } j \text{ of } Y)^T\right|^2 \\
&= \left(\sum_{l=1}^{n} |X_{il}|^2\right) \cdot \left(\sum_{l=1}^{n} |Y_{lj}|^2\right).
\end{aligned}$$

Thus it follows that

$$\begin{aligned}
|XY|^2 &= \sum_{i,j=1}^{n} |(XY)_{ij}|^2 \\
&\leq \sum_{i,j=1}^{n} \left(\left(\sum_{l=1}^{n} |X_{il}|^2\right) \cdot \left(\sum_{l=1}^{n} |Y_{lj}|^2\right)\right) \\
&= \left(\sum_{i,j=1}^{n} |X_{ij}|^2\right) \cdot \left(\sum_{i,j=1}^{n} |Y_{ij}|^2\right) \\
&= |X|^2 |Y|^2 .
\end{aligned}$$

Taking the square root of this equation, we get $|XY| \leq |X| \cdot |Y|$, as desired. □

Theorem 4.11. [6] *Let $f(x) = c_0 + c_1 x + c_2 x^2 + \cdots = \sum_{i=1}^{\infty} c_i x^i$ be a power series with coefficients $c_i \in \mathbb{R}$ and a radius of convergence R. If $A \in M_n(\mathbb{R})$ satisfies $|A| < R$, then $f(A) = c_0 I_n + c_1 A + c_2 A^2 + \cdots = \sum_{i=1}^{\infty} c_i A^i$ converges absolutely.*

Proof. Let $f(x) = c_0 + c_1x + c_2x^2 + \cdots$ be a power series with coefficients $c_i \in \mathbb{R}$ with a radius of convergence R. Let $A \in M_n(\mathbb{R})$ be such that $|A| < R$. For any indices i, j, we must show that $|(c_0I_n)_{ij}| + |(c_1A)_{ij}| + |(c_2A^2)_{ij}| + \cdots$ converges. For any $k \in \mathbb{N}$, it follows by Lemma 4.10 that

$$\left|(c_kA^k)_{ij}\right| \leq \left|c_kA^k\right| = |c_k| \cdot \left|A^k\right| \leq |c_k| \cdot |A|^k.$$

Since $|A| < R$, it follows that $|(c_0I_n)_{ij}| + |(c_1A)_{ij}| + |(c_2A^2)_{ij}| + \cdots$ converges, so $f(A) = c_0I_n + c_1A + c_2A^2 + \cdots$ converges absolutely. $\square$

Through the use of Theorem 4.11, we are able to rigorously define matrix exponentiation.

Definition 4.12. Let $A \in M_n(\mathbb{R})$. The **matrix exponentiation** of A is the function

$$e^A = \exp(A) = I_n + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{i=1}^{\infty} \frac{1}{i!}A^i.$$

Those with sufficient calculus knowledge will recall that the radius of convergence for the power series of e^x is infinite, so e^A converges absolutely for all $A \in M_n(\mathbb{R})$ by Theorem 4.11.

Also, considering the function $\gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ defined as $\gamma(t) = e^{tA} = I_n + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \cdots$, it follows that $\gamma(0) = e^{0A} = I_n + 0A + \frac{1}{2!}(0A)^2 + \frac{1}{3!}(0A)^3 + \cdots = I_n$, so $\gamma(t) = e^{tA}$ is indeed a path. In fact, $\gamma(t) = e^{tA}$ is one of the most useful paths when trying to define Lie algebras of matrix groups. The following theorems will help us understand the power of matrix exponentiation.

Theorem 4.13. [6] *The path $\gamma(t) = e^{tA} = I_n + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \cdots$, where $A \in M_n(\mathbb{R})$, is differentiable with derivative $\gamma'(t) = A \cdot e^{tA}$.*

Proof. Let $A \in M_n(\mathbb{R})$ and let the function $\gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ be defined as $\gamma(t) = e^{tA} = I_n + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \cdots$ for all $t \in (-\epsilon, \epsilon)$. By Theorem 4.11, we know that $\gamma(t)$ is absolutely convergent, so we can take the derivative of $\gamma(t)$. Thus, through term-by-term differentiation, it follows that for all $t \in (-\epsilon, \epsilon)$,

$$\gamma'(t) = \frac{d}{dt}\left(I_n + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \cdots\right) = A + tA^2 + \frac{1}{2!}t^2A^3 + \cdots = A \cdot e^{tA}.$$

$\square$

Theorem 4.14. *Let $A, B \in M_n(\mathbb{R})$. If $AB = BA$, then $e^{A+B} = e^Ae^B$.*

Proof. Let $A, B \in M_n(\mathbb{R})$ be such that $AB = BA$. Due to the commutativity of A and B,

$$(A + B)^k = (A + B)(A + B)(A + B) \cdots (A + B)$$

$$
\begin{aligned}
&= (A^2 + AB + BA + B^2)(A + B) \cdots (A + B) \\
&= (A^2 + 2AB + B^2)(A + B) \cdots (A + B) \\
&= (A^3 + A^2B + 2ABA + 2AB^2 + B^2A + B^3) \cdots (A + B) \\
&= (A^3 + 3A^2B + 3AB^2 + B^3) \cdots (A + B) \\
&\ \ \vdots \\
&= A^k + kA^{k-1}B + \binom{k}{2}A^{k-2}B^2 + \cdots + \binom{k}{k-1}AB^{k-1} \\
&= \sum_{r=0}^{k} \binom{k}{r} A^{k-r}B^r.
\end{aligned}
$$

The following equalities hold.

$$
\begin{aligned}
e^{A+B} &= I_n + A + B + \frac{1}{2!}(A+B)^2 + \frac{1}{3!}(A+B)^3 + \cdots \\
&= \sum_{i=0}^{\infty} \frac{1}{i!}(A+B)^i \\
&= \sum_{i=0}^{\infty} \frac{1}{i!}\left(\sum_{j=0}^{i}\binom{i}{j}A^{i-j}B^j\right) \\
&= \sum_{i=0}^{\infty} \frac{1}{i!}\left(\sum_{j=0}^{i}\frac{i!}{(i-j)!j!}A^{i-j}B^j\right) \\
&= \sum_{i=0}^{\infty} \left(\sum_{j=0}^{i}\frac{1}{(i-j)!j!}A^{i-j}B^j\right) \\
&= I_n + A + B + \frac{1}{2!}A^2 + AB + \frac{1}{2!}B^2 + \frac{1}{3!}A^3 + \frac{1}{2!}A^2B + \frac{1}{2!}AB^2 + \frac{1}{3!}B^3 + \cdots \\
&= \left(\sum_{k=0}^{\infty}\frac{1}{k!}A^k\right)\left(\sum_{k=0}^{\infty}\frac{1}{k!}B^k\right) \\
&= e^A e^B.
\end{aligned}
$$

□

Now, using Theorems 4.13 and 4.14, we will find the Lie algebras for $GL_n(\mathbb{R})$, $O_n(\mathbb{R})$, and $SL_n(\mathbb{R})$. When referring to the Lie algebra of a matrix group, we write the Lie algebra in lower case letters. For example, the Lie algebra of $GL_n(\mathbb{R})$ is typically denoted $gl_n(\mathbb{R})$.

Theorem 4.15. *$M_n(\mathbb{R})$ is the Lie algebra of $GL_n(\mathbb{R})$.*

Proof. Let $A \in M_n(\mathbb{R})$. By Theorem 4.14, $e^A \cdot e^{-A} = e^{A-A} = e^0 = I_n$, so e^A is invertible and thus $e^A \in GL_n(\mathbb{R})$. Let $\gamma : (-\epsilon, \epsilon) \to GL_n(\mathbb{R})$ be defined as $\gamma(t) = e^{tA}$ for all $t \in (-\epsilon, \epsilon)$. By

Theorem 4.14, $e^{tA} \cdot e^{-tA} = e^{tA-tA} = e^0 = I_n$, so $e^{tA} \in GL_n(\mathbb{R})$ as well. Since $\gamma(0) = I_n$ and $\gamma'(0) = A$, it follows that $A \in gl_n(\mathbb{R})$ and thus $M_n(\mathbb{R}) \subset gl_n(\mathbb{R})$.

For the other direction, since the paths $\gamma(t)$ are all $n \times n$ matrices, their derivatives at 0 are $n \times n$ matrices as well, so $\mathfrak{g}(GL_n(\mathbb{R})) \subset M_n(\mathbb{R})$. Therefore, by double inclusion, $M_n(\mathbb{R}) = gl_n(\mathbb{R})$. $\square$

Notation. $o_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid A + A^T = 0\}$

Lemma 4.16. *If $A \in o_n(\mathbb{R})$, then $e^A \in O_n(\mathbb{R})$.*

Proof. Let $A \in o_n(\mathbb{R})$. By Theorem 2.24, we see that

$$(e^A)^T = \left(\sum_{n=0}^{\infty} \frac{A^n}{n!}\right)^T = \sum_{n=0}^{\infty} \frac{(A^n)^T}{n!} = \sum_{n=0}^{\infty} \frac{(A^T)^n}{n!} = e^{A^T}.$$

Since $A \in o_n(\mathbb{R})$, $A^T = -A$. Thus,

$$e^A(e^A)^T = e^A e^{A^T} = e^A e^{-A} = e^{A-A} = e^0 = I_n.$$

By Theorem 2.26, it follows that $e^A \in O_n(\mathbb{R})$. $\square$

Theorem 4.17. *$o_n(\mathbb{R})$ is the Lie algebra of $O(\mathbb{R})$.*

Proof. First, let $A \in o_n(\mathbb{R})$. By Lemma 4.16, it follows that the path $\gamma(t) = e^{tA} \in O_n(\mathbb{R})$. Since $\gamma(0) = I_n$ and $\gamma'(0) = A$, it follows that $A \in \mathfrak{g}(O_n(\mathbb{R}))$, so $o_n(\mathbb{R}) \subset \mathfrak{g}(O_n(\mathbb{R}))$.

Next, let $B \in \mathfrak{g}(O_n(\mathbb{R}))$. Thus, there exists some path $\sigma : (-\epsilon, \epsilon) \to O_n(\mathbb{R})$ such that $\sigma(t) \in O_n(\mathbb{R})$ for all $t \in (-\epsilon, \epsilon)$, $\sigma(0) = I_n$ and $\sigma'(0) = B$. Since $\sigma(t) \in O_n(\mathbb{R})$ for all $t \in (-\epsilon, \epsilon)$, $\sigma(t) \cdot \sigma(t)^T = I_n$ by Theorem 2.26. Using the product rule for differentiation, it follows that

$$\frac{d}{dt}\left(\sigma(t) \cdot \sigma(t)^T\right) = \sigma'(t) \cdot \sigma(t)^T + \sigma(t) \cdot \sigma'(t)^T,$$

and since $\sigma(t) \cdot \sigma(t)^T = I_n$, we get that

$$\frac{d}{dt}\left(\sigma(t) \cdot \sigma(t)^T\right) = \frac{d}{dt}(I_n) = 0.$$

When $t = 0$, we get

$$\begin{aligned} 0 &= \frac{d}{d0}\left(\sigma(0) \cdot \sigma(0)^T\right) \\ &= \sigma'(0) \cdot \sigma(0)^T + \sigma(0) \cdot \sigma'(0)^T \\ &= B \cdot I_n + I_n \cdot B^T \\ &= B + B^T \end{aligned}$$

Thus, $B \in o_n(\mathbb{R})$, which demonstrates that $\mathfrak{g}(O_n(\mathbb{R})) \subset o_n(\mathbb{R})$. Therefore, $o_n(\mathbb{R})$ is the Lie algebra of $O(\mathbb{R})$. $\square$

Lemma 4.18 will help in finding the Lie algebra of $SL_n(\mathbb{R})$. First, we introduce notation that will be used in the proof of Lemma 4.18.

Notation. Let $A \in M_n(\mathbb{R})$. We denote $A[i,j] \in M_n(\mathbb{R})$ as the matrix obtained by crossing out row i and column j of A. For example,

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} [1,1] = \begin{bmatrix} e & f \\ h & i \end{bmatrix}.$$

Lemma 4.18. [6] *If $\gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ is differentiable and $\gamma(0) = I_n$, then*

$$\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) = trace(\gamma'(0)),$$

where $trace(\gamma'(0))$ is the sum of the entries of the main diagonal of $\gamma'(0)$.

Proof. Let $\gamma : (-\epsilon, \epsilon) \to M_n(\mathbb{R})$ be differentiable with $\gamma(0) = I_n$

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) &= \left.\frac{d}{dt}\right|_{t=0} \sum_{j=1}^{n} (-1)^{j+1} \cdot \gamma(t)_{1j} \cdot \det(\gamma(t)[1,j]) \\ &= \sum_{j=1}^{n} (-1)^{j+1} \cdot \left(\gamma'(0)_{1j} \cdot \det(\gamma(0)[1,j]) + \gamma(0)_{1j} \cdot \left.\frac{d}{dt}\right|_{t=0} \det(\gamma(0)[1,j]) \right) \\ &= \gamma'(0)_{11} + \left.\frac{d}{dt}\right|_{t=0} \det(\gamma(0)[1,1]). \end{aligned}$$

Computing $\left.\frac{d}{dt}\right|_{t=0} \det(\gamma(0)[1,1])$ through the same argument n times, we get

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} \det(\gamma(t)) &= \gamma'(0)_{11} + \gamma'(0)_{22} + \cdots + \gamma'(0)_{nn} \\ &= trace(\gamma'(0)). \end{aligned}$$

□

Theorem 4.19. [6] *The Lie algebra of $SL_n(\mathbb{R})$ is*

$$sl_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid trace(A) = 0\}.$$

Proof. Let $A \in \mathfrak{g}(SL_n(\mathbb{R}))$. Thus, there exists some path $\gamma : (-\epsilon, \epsilon) \to SL_n(\mathbb{R})$ such that γ is differentiable, $\gamma(0) = I_n$, and $\gamma'(0) = A$. By Lemma 4.18, it follows that $trace(\gamma'(0)) = trace(A) = 0$. This shows that $A \in sl_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \mid trace(A) = 0\}$, so $\mathfrak{g}(SL_n(\mathbb{R})) \subset sl_n(\mathbb{R})$.

On the other hand, let $B \in M_n(\mathbb{R})$ be such that $trace(B) = 0$. Let $\sigma : (-\epsilon, \epsilon) \to SL_n(\mathbb{R})$ be defined as

$$\sigma(t) = \begin{bmatrix} \frac{ta_{11}+1}{\det(I_n+tB)} & \frac{ta_{12}}{\det(I_n+tB)} & \cdots & \frac{ta_{1n}}{\det(I_n+tB)} \\ ta_{21} & ta_{22}+1 & \cdots & ta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ta_{n1} & ta_{n2} & \cdots & ta_{nn}+1 \end{bmatrix}$$

Note that $\sigma(0) = I_n$ and

$$\sigma'(t) = \begin{bmatrix} \frac{a_{11}(\det(I_n+tB))-(ta_{11}+1)(\frac{d}{dt}\det(I_n+tB))}{\det(I_n+tB)^2} & \cdots & \frac{a_{1n}(\det(I_n+tB))-(ta_{1n}+1)(\frac{d}{dt}\det(I_n+tB))}{\det(I_n+tB)^2} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

Thus, by Lemma 4.18,

$$\begin{aligned} \sigma'(0) &= \begin{bmatrix} \frac{a_{11}(\det(I_n+(0)B))-((0)a_{11}+1)(trace(B))}{\det(I_n+(0)B)^2} & \cdots & \frac{a_{1n}(\det(I_n+(0)B))-((0)a_{1n}+1)(trace(B))}{\det(I_n+(0)B)^2} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \frac{a_{11}(1)-((0)a_{11}+1)(0)}{1^2} & \cdots & \frac{a_{1n}(1)-((0)a_{1n}+1)(0)}{1^2} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \\ &= A \end{aligned}$$

Since

$$\det(I_n + tB) = \sum_{j=1}^{n} (-1)^{j+1} \cdot (I_n + tB)_{1j} \cdot \det((I_n + tB)[1, j]),$$

it follows that

$$\begin{aligned} \det(\sigma(t)) &= \sum_{j=1}^{n} (-1)^{j+1} \cdot (I_n + tB)_{1j} \cdot \frac{1}{\det(I_n + tB)} \cdot \det((I_n + tB)[1, j]) \\ &= \frac{1}{\det(I_n + tB)} \cdot \left(\sum_{j=1}^{n} (-1)^{j+1} \cdot (I_n + tB)_{1j} \cdot \det((I_n + tB)[1, j]) \right) \\ &= \frac{1}{\det(I_n + tB)} \cdot (\det(I_n + tB)) \\ &= 1. \end{aligned}$$

Thus, $\sigma(t) \in SL_n(\mathbb{R})$ for all $t \in (-\epsilon, \epsilon)$, and since $\sigma'(0) = A$, $A \in \mathfrak{g}(SL_n(\mathbb{R})$. Therefore, $sl_n(\mathbb{R}) \subset \mathfrak{g}(SL_n(\mathbb{R})$, so $sl_n(\mathbb{R}) = \mathfrak{g}(SL_n(\mathbb{R})$. □

5. Manifolds and Lie Groups

Topologically, manifolds are abstract spaces, but around every point there is a neighborhood that looks and acts like Euclidean space. Due to this criterion, manifolds are important in the study of space as they provide a nice way to characterize and work with abstract structures since working with Euclidean space is simpler. Also, Lie groups are manifolds with more criteria attached, which is why we must first discuss manifolds before our discussion of Lie groups.

We must first go over a few fundamental theorems that will prove to be useful in later proofs.

Notation. Let $B_r := \{W \in M_n(\mathbb{R}) \mid |W| < r\}$.

Theorem 5.1. [6] *Let $G \subset GL_n(\mathbb{R})$ be a matrix group, with Lie algebra $\mathfrak{g} \subset gl_n(\mathbb{R})$.*

(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_n in G, and the restriction $\exp : B_r \cap \mathfrak{g} \to V$ is a diffeomorphism (see Definition 5.7).*

The proof of Theorem 5.1 requires delving into the world of analysis and is rather lengthy, and is therefore beyond the scope of this paper. The proof of Theorem 5.1 can be found in Tapp's *Matrix Groups for Undergraduates* [6], and is worth studying to understand the inner workings of matrix groups.

Going forwards, we look to define manifolds and prove that all matrix groups are manifolds, the proof of which relies heavily on Theorem 5.1. First we will add some more definitions to our stockpile, specifically those that pertain to functions in a topological space.

Let $U \subset \mathbb{R}^n$ be an open set in the Euclidean topology on $\mathbb{R}^n$. Any function $f : U \to \mathbb{R}^m$ can be thought of as m separate functions, that is, $f = (f_1, f_2, \ldots, f_m)$ where $f_i : U \to \mathbb{R}$ for each $i \in \{1, 2, \ldots, m\}$. An example of such a function would be the function $h : \mathbb{R}^2 \to \mathbb{R}^3$ defined as $h(x, y) = (xy, x^2 - y^2, x^3 + y)$ for all $x, y \in \mathbb{R}^2$, which is defined by the separate functions $h_1, h_2, h_3 : \mathbb{R}^2 \to \mathbb{R}^3$ where $h_1(x, y) = xy$, $h_2(x, y) = x^2 - y^2$, and $h_3(x, y) = x^3 + y$.

Definition 5.2. Let $U \subset \mathbb{R}^n$ be an open set in the standard topology on $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a function. Let $p \in U$ and let $v \in \mathbb{R}^n$. 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.

Consider the straight line $g(t) = p + tv$ in $\mathbb{R}^n$. Visually, $df_p(v)$ is the initial velocity vector of $(f \circ g)(t) = f(p + tv)$, if this velocity vector exists. In this sense, $df_p(v)$ can be thought of as an approximation of where the function f sends points near p in the direction of v.

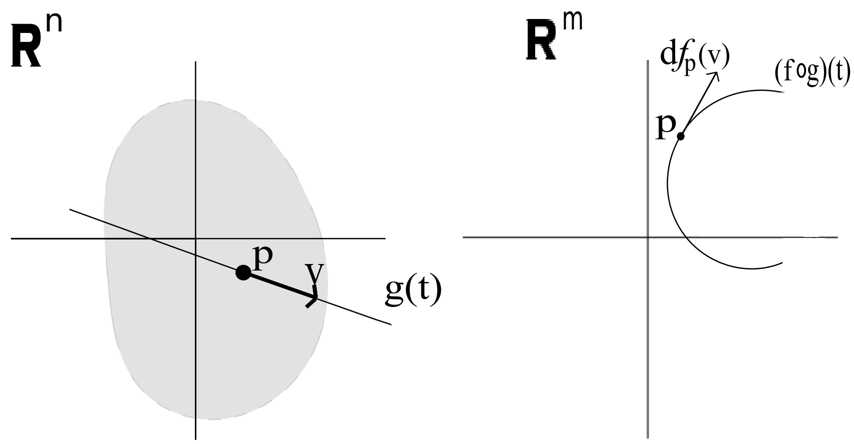

FIGURE 5.1. $df_p(v)$ is the initial velocity vector of $(f \circ g)(t)$ in the direction of v.

The following proposition is a more intuitive understanding of Definition 5.2.

Proposition 5.3. *$df_p(v)$ 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 $\gamma'(0) = v$.*

Definition 5.4. [6] Let $U \subset \mathbb{R}^n$ be an open set in the standard topology on $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a function. The directional derivatives of the component functions $\{f_1, f_2, \ldots, f_m\}$ in the directions of the standard orthonormal basis vectors $\{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ are called **partial derivatives** of f and are denoted as

$$\frac{\partial f_i}{\partial x_j}(p) := d(f_i)_p(e_j).$$

Directional derivatives of a function f measure the rates of change of each of the component functions of f. If we fix $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$ and if $\frac{\partial f_i}{\partial x_j}(p)$ exists for all $p \in U$, then the function $g : U \to \mathbb{R}^m$ defined as $g(p) = \frac{\partial f_i}{\partial x_j}(p)$ is a well-defined function from U to $\mathbb{R}^m$, so we can take the partial derivatives of g. If the partial derivatives of g exist, they are called the *second order* partial derivatives of f. Following in this matter, if we take r partial derivatives of a function f and the partial derivatives exits, then we say that they are the r^{th} order partial derivatives of f. This gives rise to the following definition.

Definition 5.5. [6] Let $U \subset \mathbb{R}^n$ be an open set in the standard topology on $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a function. The function f is called C^r **on** U if all r^{th} order partial derivatives exist and are continuous on U, and f is called **smooth on** U if f is C^r on U for all positive integers r.

Similarly, we can define smoothness for any set $X \subset \mathbb{R}^n$, not just open sets.

Definition 5.6. [6] If $X \subset \mathbb{R}^n$, then $f : X \to \mathbb{R}^n$ is called **smooth** if for all $p \in X$, there exists an open neighborhood U of p in $\mathbb{R}^m$ and a smooth function $\tilde{f} : U \to \mathbb{R}^n$ which agrees with f on $X \cap U$.

Using this more general definition of smoothness, we can create a type of similarity between subsets of $\mathbb{R}^n$, which will allow us to define a manifold in $\mathbb{R}^n$.

Definition 5.7. [6] $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ are called **diffeomorphic** if there exists a smooth bijective function $f : X \to Y$ whose inverse is also smooth. In this case, f is called a **diffeomorphism**.

Definition 5.8. [6] A subset $M \subset \mathbb{R}^m$ is called a **manifold** of dimension n if for all $p \in M$ there exists a neighborhood U of p in M which is diffeomorphic to an open set $V \subset \mathbb{R}^n$.

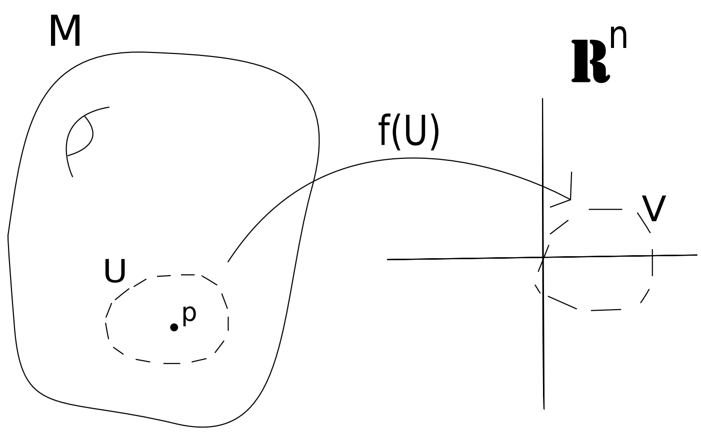

FIGURE 5.2. A visualization of a manifold of dimension n.

To prove that a set M is a manifold, we need to construct a *parametrization* at every point $p \in M$, which is a diffeomorphism ϕ from an open set $V \subset \mathbb{R}^n$ to a neighborhood U of $p \in M$. We use this method to prove that any matrix group is a manifold.

Theorem 5.9. *Any matrix group is a manifold.*

Proof. Let $G \subset GL_n(\mathbb{R})$ be a matrix group with Lie algebra $\mathfrak{g}$. Choose a sufficiently small $r > 0$ which is guaranteed by Theorem 5.1. Thus, $V := \exp(B_r \cap \mathfrak{g})$ is a neighborhood of I_n in G, and the restriction map $\exp : B_r \cap \mathfrak{g} \to V$ is a diffeomorphism, so $\exp : B_r \cap \mathfrak{g} \to V$ is a parametrization at I_n.

Next, let $g \in G$ be arbitrary. Define the function $\mathcal{L}_g : G \to G$ as $\mathcal{L}_g(A) = g \cdot A$ for all $A \in G$. $\mathcal{L}_g$ is injective because if $g \cdot A = g \cdot B$ for some $A, B \in G$, then $A = B$ through left multiplication by g^{-1}. Also, $\mathcal{L}_g$ is surjective because, for all $C \in G$, $\mathcal{L}_g(g^{-1} \cdot C) = C$, so $\mathcal{L}_g$ is bijective. Since matrix multiplication from $M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$ can be thought of as a function with n^2 component functions, it follows that $\mathcal{L}_g$ is smooth as each component function is a polynomial over $\mathbb{R}$, so all r^{th} order partial derivatives exist and are continuous on G. Also, since G is a group, g^{-1} exists. Thus, the inverse function of $\mathcal{L}_g$ is $\mathcal{L}_g^{-1}(B) = g^{-1} \cdot B$, which is also smooth through the same reasoning. Thus, $\mathcal{L}_g$ is a diffeomorphism from G to G, so $\mathcal{L}_g(V)$ in particular is a neighborhood of g in G as $\mathcal{L}_g$ maps open neighborhoods to open neighborhoods being a diffeomorphism. Therefore, $(\mathcal{L}_g \circ \exp) : B_r \cap \mathfrak{g} \to \mathcal{L}_g(V)$ is a parametrization at g as the composition of diffeomorphisms is diffeomorphic, proving that G is a manifold. $\square$

We are now able to move on to the high point of this section.

Definition 5.10. A **Lie group** is a manifold, G, with a smooth group operation $G \times G \to G$ and a smooth inverse map.

Theorem 5.11. *All matrix groups are Lie groups.*

Proof. Let $G \subset GL_n(\mathbb{R})$ be a matrix group. From Theorem 5.9, we know that G is a manifold. Also, from the proof of Theorem 5.9, we know that matrix multiplication over matrix groups is smooth, so the group operation of G is smooth. Further, the inverse map of G is the function $\iota : G \to G$ defined as $\iota(A) = \frac{1}{\det(A)}\text{adj}(A)$, which is smooth as this is also just a calculation of polynomials in $\mathbb{R}$ (this is a standard result from linear algebra). This shows that G is a Lie group, proving the statement. $\square$

6. Conclusion

Momentarily diverting from explaining the ramifications of Theorem 5.11, a direct example of matrix groups in practice is demonstrated through an application to computer animation. Suppose that we are given some smiley face graphic, properly named "Smiley". Our goal is to transform Smiley into an upside-down smiley face and then return them back to their original position. If we consider Smiley as a set of points in $\mathbb{R}^2$, then we can multiply Smiley as a whole by a matrix that corresponds to rotations in $\mathbb{R}^2$. If we don't like Smiley in

[h]

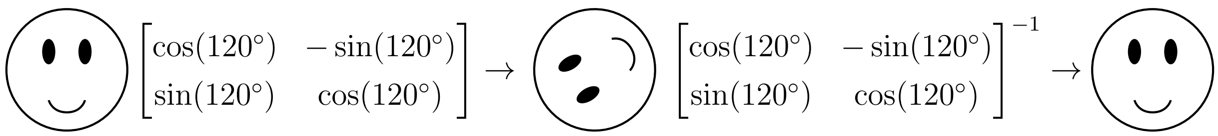

FIGURE 6.1. Smiley being rotated through matrix multiplication.

their new orientation, we can then multiply Smiley by the inverse of that matrix to return them back to their original position.

Since the matrix given in Figure 6.1 is a member of the matrix group $SO_2(\mathbb{R})$, the matrix group of all 2×2 orthogonal matrices with determinant equal to 1, we know that the matrix has an inverse that is also in $SO_2(\mathbb{R})$. Therefore, we can rotate Smiley around and be sure that we can return Smiley back to their original position through another multiplication operation. If $SO_2(\mathbb{R})$ was not a group, it is not certain that such an inverse would exist.

The inferences of the importance of all matrix groups being Lie groups in the study of Lie theory in general can be understood from the following two theorems.

Theorem 6.1. *Every compact Lie group is smoothly isomorphic to a matrix group.*

Theorem 6.2. *The Lie algebra of any Lie group is isomorphic to the Lie algebra of a matrix group.*

The proof of these two theorems are beyond the scope of this paper, but the significance of them can be felt none the less. The physical world is represented by real numbers, and so are matrix groups (or at least some of them). Thus, matrix groups are the medium to studying Lie theory as it pertains to the physical world. Modeling physical phenomenon as matrices and studying the Lie structure of those matrices allows us to use results from Lie theory and apply them to our physical phenomenon to understand the structure of these events. Matrix groups are the vehicle by which developments in understanding the physical world travel, and by first understanding matrix groups we are one step closer to a better understanding of the physical and mathematical realms which Lie theory pertains.

Appendix

Linear Algebra.

Definition 1. [2] Let V be an arbitrary nonempty set of elements on which two operations are defined: addition, and multiplication by numbers called **scalars**. By **addition** we mean a rule for associating with each pair of objects u and v in V an object $u+v$, called the sum of u and v; by **scalar multiplication** we mean a rule for associating with each scalar k and each object u in V an object ku, called the **scalar multiple** of u by k. If the following axioms are satisfied by all objects u, v, w in V and all scalars k and m, then we call V a **vector space** and we call the objects in V **vectors**.

(1) If u and v are objects in V, then $u+v$ is in V.
(2) $u+v=v+u$ for all $u, v \in V$.
(3) $u+(v+w)=(u+v)+w$ for all $u, v, w \in V$.
(4) There is an object 0 in V, called a **zero vector** for V, such that $0+u=u+0=u$ for all $u \in V$.
(5) For each $u \in V$, there is an object $-u \in V$ such that $u+(-u)=(-u)+u=0$.
(6) If k is any scalar and u is any object in V, then $ku \in V$.
(7) $k(u+v)=ku+kv$ for all $u, v \in V$ and for all $k \in \mathbb{R}$.
(8) $(k+m)u=ku+mu$ for all $u \in V$ and for all $k, m \in \mathbb{R}$.
(9) $k(mu)=(km)u$ for all $u \in V$ and for all $k, m \in \mathbb{R}$.
(10) $1u=u$ for all $u \in V$.

Definition 2. [2] A subset W of a vector space V is called a **subspace** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

Definition 3. [2] If $S=\{w_1, w_2, \ldots, w_r\}$ is a nonempty set of vectors in a vector space V, then the subspace W of V that consists of all possible linear combinations of the vectors in S is called the subspace of V **generated** by S, and we say that the vectors $w_1, w_2, ..., w_r$ **span** W. We denote this subspace as $W = \text{span}(S)$.

Definition 4. [2] If $S=\{v_1, v_2, \ldots, v_n\}$ is a set of two or more vectors in a vector space V, then S is said to be a **linearly independent set** if no vector in S can be expressed as a linear combination of the others. A set that is not linearly independent is said to be **linearly dependent**.

Definition 5. [2] A vector space V is said to be a **finite-dimensional** vector space if there is a finite set of vectors in V that span V and is said to be **infinite-dimensional** if no such set exists.

Definition 6. [2] If $S = \{v_1, v_2, \ldots, v_n\}$ is a set of vectors in a finite-dimensional vector space V, then S is called a **basis** for V if:
(a) S spans V.
(b) S is linearly independent.

Definition 7. If $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is a basis for a vector space V, then the **dimension** of V is the cardinality of $\mathcal{B}$.

Definition 8. [2] If $T : V \to W$ is a mapping form a vector space V to a vector space W, then T is called a **linear transformation** from V to W if the following two properties hold for all vectors u and v in V and for all scalars k:
(i) $T(ku) = kT(u)$
(ii) $T(u + v) = T(u) + T(v)$.

Acknowledgements

The author would like to acknowledge Professor Keely Machmer-Wessels for her help and guidance through this project. The author would also like to acknowledge Professor Kathy Porter for her facilitation of the Senior Seminar program. Lastly, the author would like to thank the Saint Mary's College of California Math Department for providing the space and opportunity to conduct this research.

References

[1] S. Abbott, *Understanding analysis*, Second, Undergraduate Texts in Mathematics, Springer, New York, 2015. MR3331079

[2] H. Anton, *Elementary linear algebra*, John Wiley & Sons, 2010.

[3] J. B. Fraleigh and V. J. Katz, *A first course in abstract algebra*, Addison-Wesley world student series, Addison-Wesley, 2003.

[4] B. Hall, *Lie groups, Lie algebras, and representations*, Second, Graduate Texts in Mathematics, vol. 222, Springer, Cham, 2015. An elementary introduction. MR3331229

[5] J. R. Munkres, *Topology, 2nd ed.*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 2000. MR0464128

[6] K. Tapp, *Matrix groups for undergraduates*, Student Mathematical Library, vol. 29, American Mathematical Society, Providence, RI, 2005. MR2141985 (2006d:20001)

[7] V. X. Varadarajan, *Historical review of lie theory*, 2007.