LIE GROUPS AND LIE ALGEBRAS

NEIL KRISHNAN

Abstract. We discuss the definitions of Lie groups and Lie algebras and their relationship.

1. Overview

This paper assumes some familiarity with group theory, linear algebra, and differential geometry. *Lie groups* have two main characteristics: they are groups and they are *smooth* in some way. For example, the group $\mathbb{Z}_3$ of group of integers modulo 3 under addition has a finite number of elements and as a result the operations do not seem smooth. Although $\mathbb{Z}^+$ (the group of integers under addition) is a group with an infinite number of elements, somehow the group operation seems discrete as there is no action in between adding 0 and adding 1. On the other hand, the group $\mathbb{R}^+$ of reals under addition seems smooth.

Smoothness allows us to connect Lie groups to vector spaces called Lie algebras through the exponential map. Because algebras are easier objects to understand, we can understand many properties of the Lie group more easily through the Lie algebra. This relationship, in a few of its aspects, will be the culmination of this paper. For more information on linear Lie groups and Lie algebras, see [2], and for a more gentle introduction and more detail on more general lie groups, see [1].

2. Lie Groups

We must first properly define what we mean by smooth. Consider a sphere. A sphere is smooth because there are no cusps or pointy bumps. In other words, if we zoom in on any point of the sphere, it locally looks like a plane. We can formalize this notion of a section A of the sphere looking like a plane through a diffeomorphism from the plane to A.

Definition 2.1. Let $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$. A *diffeomorphism* $f : X \to Y$ is a smooth (infinitely differentiable) bijection with smooth inverse.

Every small section on the sphere must locally look like a plane, so for every point $\mathbf{p}$, there must be a subset of the sphere V containing $\mathbf{p}$ such that there is a diffeomorphism from a subset of the plane U to V. More generally, we a have a manifold.

Definition 2.2. A *m-dimensional manifold* M is a subset of $\mathbb{R}^n$ such that for every point $\mathbf{p} \in M$ there is a subset $V \subseteq M$ and a subset $U \subseteq \mathbb{R}^m$ such that there is a diffeomorphism from U to V.

Date: August 15, 2025.

Example 2.3. For a unit sphere, we can use the spherical parametrization

$$\boldsymbol{\sigma}_1(u, v) = (\cos u \sin v, \sin u \sin v, \cos v),$$

which maps from $(0, 2\pi) \times (0, \pi)$ to the entire sphere except the great arc from $(0, 0, 1)$ to $(0, 0, -1)$ intersecting $(1, 0, 0)$. For points on this arc, we must find another parametrization. Consider

$$\boldsymbol{\sigma}_2(u, v) = (-\cos u \sin v, \cos v, \sin u \sin v),$$

which maps from $(0, 2\pi) \times (0, \pi)$ to the entire sphere except the arc from $(0, 1, 0)$ to $(0, -1, 0)$ intersection $(-1, 0, 0)$. This is just a rotation of our previous parametrization, but it covers the arc from $(0, 0, 1)$ to $(0, 0, -1)$.

Both $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are diffeomorphisms, so the sphere is a 2-dimensional manifold.

We can now define Lie groups.

Definition 2.4. A *Lie group* is a manifold M combined with a smooth operation $\cdot : M \times M \to M$ and smooth inverse ${}^{-1} : M \to M$ such that the set of points in M along with the operation and inverse forms a group.

Let $\mathbf{GL}(n; \mathbb{R})$ or the *general linear group* be the group of real invertible $n \times n$ matrices. The main class of Lie groups we will talk about are *linear Lie groups* which are Lie groups that are subgroups of $\mathbf{GL}(n; \mathbb{R})$. We can determine whether a subgroup of $\mathbf{GL}(n; \mathbb{R})$ is a Lie group very easily as a result of the following theorem.

Theorem 2.5 (Von Neumann and Cartan 1927)**.** *Let G be a closed subgroup of $\mathbf{GL}(n; \mathbb{R})$, i.e., for every convergent sequence of elements of G, the resulting limiting matrix is either in G or not in $\mathbf{GL}(n; \mathbb{R})$. Then, G is a linear Lie group.*

Example 2.6. The following sets are linear Lie groups.

(1) $\mathbf{GL}(n; \mathbb{R})$ is a subgroup of $\mathbf{GL}(n; \mathbb{R})$. Furthermore, it is closed as all matrices are either invertible or not invertible. Thus the limit of a sequence of matrices is either in $\mathbf{GL}(n; \mathbb{R})$ or not in $\mathbf{GL}(n; \mathbb{R})$.
(2) $\mathbf{SL}(n; \mathbb{R})$ or the *special linear group* is the group of real $n \times n$ matrices of determinant 1. Because $\det(AB) = \det(A)\det(B)$ for all matrices, we have $\mathbf{SL}(n; \mathbb{R})$ is a subgroup of $\mathbf{GL}(n; \mathbb{R})$ and it is closed as the determinant is a smooth function, so the limit of matrices with determinant 1 must also have determinant 1 and be in $\mathbf{SL}(n; \mathbb{R})$.
(3) $\mathbf{O}(n; \mathbb{R})$ or the *orthogonal group* is the group of orthogonal matrices, i.e., matrices A such that $A^T A = I$. This is indeed a subgroup as if A and B are orthogonal matrices, then $(AB)^T(AB) = B^T A^T AB = I$. Note that matrix multiplication and transposes are smooth functions, so the relation $A^T A = I$ is preserved under limits.
(4) $\mathbf{SO}(n)$ or the *special orthogonal group* is a group of orthogonal matrices with determinant 1 (orthogonal matrices can have determinant 1 or -1). This is a closed subgroup as it is the intersection of two closed subgroups, $\mathbf{SL}(n; \mathbb{R})$ and $\mathbf{O}(n)$.

Remark 2.7. Consider the subgroup H of $\mathbf{GL}(1; \mathbb{R})$ of just the integer powers of 2. This is a subgroup and it is closed (the limit of $1, 1/2, 1/4, \ldots$ is 0 which is not in $\mathbf{GL}(1; \mathbb{R})$). Yet H seems discrete. In actuality, H is a union of disconnected 0-dimensional manifolds. This is similar to how $\mathbf{O}(n)$ has two connected components: one where matrices have determinant 1 and another where matrices have determinant -1.

3. Exponential Map

We restrict ourselves to linear Lie groups unless otherwise specified. To make this section more rigorous, we define the norm of a matrix.

Definition 3.1. The *norm* of an $n \times n$ matrix A is

$$\|A\| = \max_{\mathbf{x} \in \mathbb{R}^n} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

This definition of the norm satisfies several properties. For example, we have the triangle inequality as

$$\|A+B\| = \max_{x \in \mathbb{R}^n} \frac{\|(A+B)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{x \in \mathbb{R}^n} \frac{\|A\mathbf{x}\| + \|B\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{\mathbf{x} \in \mathbb{R}^n} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} + \max_{\mathbf{y} \in \mathbb{R}^n} \frac{\|B\mathbf{y}\|}{\|\mathbf{y}\|} = \|A\| + \|B\|.$$

In addition

$$\|AB\| = \max_{x \in \mathbb{R}^n} \frac{\|AB\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{x \in \mathbb{R}^n} \frac{\|AB\mathbf{x}\|}{\|B\mathbf{x}\|} \cdot \frac{\|Bx\|}{\|\mathbf{x}\|} \leq \left(\max_{\mathbf{x} \in \mathbb{R}^n} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \right) \left(\max_{\mathbf{y} \in \mathbb{R}^n} \frac{\|B\mathbf{y}\|}{\|\mathbf{y}\|} \right) = \|A\|\|B\|.$$

This norm allows us to define limits and convergence of sequences in the usual way.

Definition 3.2. The *exponential map* is a function from matrices to matrices given by

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

for a matrix A.

Notice that this sum must converge for all A as

$$\|e^A\| \leq \|I\| + \|A\| + \left\| \frac{A^2}{2!} \right\| + \cdots \leq \|I\| + \|A\| + \frac{\|A\|^2}{2!} + \cdots \leq e^{\|A\|}.$$

Example 3.3. Let

$$A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}.$$

Let

$$I_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad I_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad I_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_0 \quad I_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -I_1,$$

and define I_k for $k \geq 4$ to be I_ℓ where $k \equiv \ell \bmod 4$ and $0 \leq \ell \leq 3$. Through matrix multiplication, we can see that $A^k = \theta^k I_k$, so

$$e^A = \sum_{k=0}^{\infty} \frac{\theta^k I_k}{k!} = \left(I_0 + \frac{\theta^2 I_2}{2!} + \cdots \right) + \left(\theta I_1 + \frac{\theta^3 I_3}{3!} + \cdots \right) = I_0 \cos\theta - I_1 \sin\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

As a result, the exponential map takes 2×2 skew-symmetric matrices to matrices in $\mathbf{SO}(2)$. This map is actually surjective as all matrices in $\mathbf{SO}(2)$ are rotation matrices.

The exponential map also has the following properties which are simple to prove.

Proposition 3.4. *For matrices X and Y we have*

- $e^0 = I$

- $(e^X)^T = e^{X^T}$.
- $(e^X)^{-1} = e^{-X}$.
- $e^{C^{-1}XC} = C^{-1}e^X C$.
- *If X and Y commute, then $e^{X+Y} = e^X e^Y$.*

As a side note on calculating the exponential map, notice that we can easily find the exponential map in two cases. If a matrix A is diagonalizable, it can be expressed as $S^{-1}\Lambda S$ where Λ is diagonal, so by Proposition 3.4, we have

$$e^A = S^{-1}e^\Lambda S = S^{-1}\exp\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix} S = S^{-1}\begin{pmatrix}e^{\lambda_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & e^{\lambda_n}\end{pmatrix} S.$$

If N is nilpotent, or in other words there is some m such that $A^m = 0$, then there are only a finite number of terms in e^A, so it can be easily calculated.

Because of the Jordan normal form, every matrix can be written as

$$S^{-1}BS = S^{-1}\begin{pmatrix}D_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & D_k\end{pmatrix} S,$$

where the D_i are upper triangular blocks where every element on the diagonal is the same and every element on the superdiagonal (smaller diagonal above main diagonal) is 0 or 1. Therefore splitting B into a diagonal and superdiagonal portions, we see that A is the sum of a diagonalizable matrix D and a nilpotent matrix N. It can be shown that D and N commute, so $e^A = e^D e^N$. Thus we can find e^A more easily by multiplying e^D and e^N.

Recall that the determinant of a matrix is the product of the eigenvalues and the trace is the sum of the eigenvalues. Furthermore, the determinant of e^N where N is nilpotent is one as e^N is the sum of upper triangular matrices and only the first term of I contributes to the terms on the diagonal. From the previous discussion, we then have

$$\det(e^A) = \det(e^D)\det(e^N) = \det(S^{-1}e^\Lambda S)\cdot 1 = e^{\operatorname{tr}(\Lambda)} = e^{\operatorname{tr}(D)} = e^{\operatorname{tr}(D)+\operatorname{tr}(N)} = e^{\operatorname{tr}(A)}.$$

Note that the fifth property of Proposition 3.4 is not true for noncommutative matrices X and Y. For example, in the case of

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} = \exp\left(\begin{pmatrix}0 & 0\\ \theta & 0\end{pmatrix} + \begin{pmatrix}0 & -\theta\\ 0 & 0\end{pmatrix}\right) \neq \exp\begin{pmatrix}0 & \theta\\ 0 & 0\end{pmatrix}\exp\begin{pmatrix}0 & -\theta\\ 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\ \theta & 1\end{pmatrix}\begin{pmatrix}1 & -\theta\\ 0 & 1\end{pmatrix}.$$

Still there is a relationship between e^{X+Y}, e^X, and e^Y as stated in the following theorem.

Theorem 3.5. *For matrices X and Y, we have*

$$e^{X+Y} = \lim_{m\to\infty}\left(e^{X/m}e^{Y/m}\right)^m.$$

The proof of this theorem is omitted, but it boils down to using Taylor approximations and the fact that as $m \to \infty$, higher order terms tend to 0 faster to simplify the right hand side into the left hand side.

Consider e^{tX} where $t \in \mathbb{R}$ and $X \in \mathbf{GL}(n;\mathbb{R})$. This corresponds to some curve in $\mathbb{R}^{n^2}$, so we can consider the derivative of this curve. We have

$$\frac{d}{dt}e^{tX} = \frac{d}{dt}\left(I + \sum_{k=1}^{\infty} t^k \frac{X^k}{k!}\right) = 0 + \sum_{k=1}^{\infty} t^{k-1}\frac{X^k}{(k-1)!} = Xe^{tX}.$$

4. Lie Algebras

To make Lie algebras more concrete, let us first examine the Lie algebra of a Lie group.

Definition 4.1. The *Lie algebra* of a linear Lie group G of $n \times n$ matrices is the subset $\mathfrak{g} \subseteq \mathbf{GL}(n;\mathbb{R})$ of all $A \in \mathfrak{g}$ where $e^{tA} \in G$ for all $t \in \mathbb{R}$.

Example 4.2. The following are the corresponding Lie algebras for the examples of Lie groups in Example 2.6.

(1) The Lie algebra $\mathfrak{gl}(n;\mathbb{R})$ of $\mathbf{GL}(n;\mathbb{R})$ is the set of all $n \times n$ matrices as for every matrix A, e^{tA} has inverse e^{-tA}, so e^{tA} is always invertible.
(2) The Lie algebra $\mathfrak{sl}(n;\mathbb{R})$ of $\mathbf{SL}(n;\mathbb{R})$ is the set of all matrices with null trace as for any matrix A with $\operatorname{tr}(A) = 0$, we have $\det(e^{tA}) = e^{\operatorname{tr}(tA)} = 1$.
(3) The Lie algebra $\mathfrak{o}(n)$ of $\mathbf{O}(n)$ is the set of all skew-symmetric matrices, i.e., matrices A with $A^T = -A$. For an example, see Example 3.3. If e^{tA} is orthogonal, then $(e^{tA})^T(e^{tA}) = I$, so $e^{tA^T}e^{tA} = I$ which implies $e^{tA^T} = e^{-tA}$. Taking the derivative with respect to t and evaluating at $t = 0$, we then have $A^T = -A$.
(4) The Lie algebra $\mathfrak{so}(n)$ of $\mathbf{SO}(n)$ is also the set of all skew-symmetric matrices as only skew-symmetric matrices A satisfy e^A is orthogonal, and for such A, we see $\det(e^A) = e^{\operatorname{tr}(A)} = 1$.

More generally, Lie algebras $\mathfrak{g}$ satisfy a few properties. First, if $X \in \mathfrak{g}$ then $sX \in \mathfrak{g}$ as $e^{tsX} = e^{(ts)X} \in \mathfrak{g}$. We also know that if $X, Y \in \mathfrak{g}$, then $X+Y \in \mathfrak{g}$ as X/m and Y/m are in $\mathfrak{g}$ so $\left(e^{X/m}e^{Y/m}\right)^m$ is in G. Because G is closed, the limit as $m \to \infty$ of these elements is in G, so $X + Y \in \mathfrak{g}$. Thus, Lie algebras are actually vector spaces with scalar multiplication by reals and vector addition through matrix addition. In fact, Lie algebras are the tangent spaces of their Lie groups G at the identity element. Recall that the tangent space of a manifold M at a point $\mathbf{p}$ is

$$T_{\mathbf{p}}M = \{\dot{\boldsymbol{\gamma}}(0) : \boldsymbol{\gamma} \text{ is a smooth curves on } M \text{ with } \boldsymbol{\gamma}(0) = \mathbf{p}\}.$$

Theorem 4.3 (Von Neumann and Cartan 1927)**.** *Let $\mathfrak{g}$ be the Lie algebra of a linear Lie group G. Then $X \in \mathfrak{g}$ if and only if $X \in T_I G$.*

Proof. The forward direction is immediate. If $X \in \mathfrak{g}$, then $e^{tX} \in G$ for all t and e^{tX} is a curve on G intersecting the identity and the derivative of this curve at 0 is X.

For the reverse direction, consider a curve $\boldsymbol{\gamma}$ on G with $\boldsymbol{\gamma}(0) = I$. Let $X = \dot{\boldsymbol{\gamma}}(0)$. Then

$$\boldsymbol{\gamma}(t) = I + tX + \text{higher order terms}.$$

Consider $\boldsymbol{\gamma}(t/k)$ for a positive integer k. As $k \to \infty$, we see that the higher order terms decay on the order of k^{-2} while the main terms decay only at the rate of k^{-1}. Thus

$$\lim_{k\to\infty} \boldsymbol{\gamma}\left(\frac{t}{k}\right)^k = \lim_{k\to\infty}\left(I + \frac{tX}{k}\right)^k = e^{tX}.$$

Because $\gamma(t/k)^k$ is in G for all k, we must have that e^{tX} is in G as well. ■

So far we have been discussing the Lie algebra, but we have only discussed addition. Multiplication comes in the form of the Lie bracket.

Definition 4.4. For matrices $X, Y \in \mathfrak{g}$ the *Lie bracket* is $[X, Y] = XY - YX$.

Proposition 4.5. *For all $X, Y \in \mathfrak{g}$, we have $[X, Y] \in \mathfrak{g}$.*

Proof. Consider $e^{tX}Ye^{-tX} \in \mathfrak{g}$. Notice that this is in $\mathfrak{g}$ as

$$\exp(te^{tX}Ye^{-tX}) = e^{tX}e^{tY}e^{-tX} \in G.$$

Because this curve is contained in $\mathfrak{g}$, its derivative at $t = 0$ must also be contained in $\mathfrak{g}$. Therefore, using the product rule,

$$\frac{d}{dt}e^{tX}Ye^{-tX}|_{t=0} = Xe^{tX}Ye^{-tX}|_{t=0} + e^{tX}Y(-X)e^{-tX}|_{t=0} = XY - YX = [X, Y].$$

■

The Lie bracket has a nice interpretation in terms of G as the tangent space as well through the Baker-Campbell-Hausdorff formula. Let $X, Y \in \mathfrak{g}$. Then X and Y are the derivatives of e^{tX} and e^{sY} evaluated at $t = 0$ and $s = 0$, respectively. Now consider $e^{tX}e^{sY}$ for fixed t and s which lies along a different curve from the identity element I. If X and Y commuted, then $e^{tX+sY} = e^{tX}e^{sY}$ which implies that $e^{tX}e^{sY}$ lies on a curve of the form e^{tA} with derivative a multiple of $tX + sY$. Thus, we get a sort of linearity between multiplication in the group and addition in the algebra. If X and Y do not commute, we can use the following theorem.

Theorem 4.6 (Baker-Campbell-Hausdorff)**.** *For X and Y in $\mathfrak{g}$ and $t, s \in \mathbb{R}$ we have*

$$e^{tX}e^{sY} = \exp\left(tX + sY + \frac{1}{2}ts[X, Y] + \cdots\right),$$

where the dots are higher order terms.

Therefore $[X, Y]$ gives information about the relationship between addition in the Lie algebra and multiplication in the Lie group.

References

[1] J. Gallier and J. Quaintance, *Differential geometry and Lie groups. A second course*, Geom. Comput., vol. 13, Cham: Springer, 2020 (English). doi:10.1007/978-3-030-46047-1

[2] B. C. Hall, *An Elementary Introduction to Groups and Representations*, 2000.