Elementary Matrix Lie Theory.

Samuel Harshe
Yale University
Professor Igor Frenkel
MATH 480: *Mathematical Topics*
April 14, 2024

Contents.

1 Introduction. 1

2 Smooth manifolds. 2
  2.1 Topological definitions. 2
  2.2 Topological manifolds. 3
  2.3 Smooth manifolds. 4

3 Matrix groups. 5
  3.1 Maximality of the general linear group. 5
  3.2 Classical matrix group definitions. 6
  3.3 Classical matrix groups are matrix groups. 6

4 Matrix Lie groups. 8
  4.1 Definitions and discussion. 8
  4.2 Direct proof: the general linear group is a Lie group. 9

5 Matrix exponential and logarithm functions. 10
  5.1 Matrix exponential function. 10
  5.2 Matrix logarithm function. 13
  5.3 One-parameter subgroups. 14

6 Lie algebras. 16
  6.1 Definitions. 16
  6.2 The Lie algebra of a matrix Lie group. 18

7 Examples. 24

1 Introduction.

Before we go nose to grindstone and deal carefully with the details of Lie groups and Lie algebras, it merits sketching the goal of this paper both to motivate the preliminary results and to allow readers to try to anticipate significant connections as they go. We therefore begin with the most important definition: a **Lie group** is a group with smooth composition and inversion maps that is also a smooth manifold. All of this structure makes Lie groups very nice to work with. We construct a tangent space at the identity of the group, endow it with a multiplication, and call it the **Lie algebra**. We can then define a map thats recovers much of the group structure from the algebra.

This means we can do much of our work on the algebra (which, conveniently, is linear) instead of on the group, partially reducing group operation to a much simpler linear algebra problem.

2 Smooth manifolds.

Perhaps surprisingly, much Lie theory can be covered without explicitly using manifolds. This is done through matrix Lie groups. Since there is is more than enough matrix Lie theory to fill two lectures, and I believe it is the quickest and most intuitive way to cover ground, matrices are the idiom of most of this paper. However, it would be wrong to leave manifolds completely out of the picture. We will therefore begin with smooth manifolds and prove a connection to matrix Lie groups before leaving them in the background for the rest of the paper. A significant result toward the end establishes the limit of this approach.

2.1 Topological definitions.

We collect several preliminary definitions pertinent to smooth manifolds (§5.1 in [2]).

- A set X equipped with a function $\mathcal{N}(x)$ that assigns to each $x \in X$ a nonempty collection of subsets $\{N\}$ (called **neighborhoods** of x) is a **topological space** if it satisfies the following four axioms.
  1. Each point $x \in X$ belongs to each of its neighborhoods.
  2. Each superset of a neighborhood of x is also a neighborhood of x.
  3. The intersection of any two neighborhoods of x is a neighborhood of x.
  4. Any neighborhood N contains a subneighborhood M such that N is a neighborhood of each point in M.
- A set is **open** if it contains a neighborhood of each point.
- The **basis** of a topological space M is some family $\mathcal{B}$ of open subsets such that every open set in M can be generated by taking the union of some some sub-family of $\mathcal{B}$.
- A topological space is **second countable** if it has a countable basis.
- A **Hausdorff space** is a topological space M where $\forall x, y \in M$, $\exists U_x, U_y$, neighborhoods of x and y, such that $U_x \cap U_y = \emptyset$.
- A **homemorphism** is a continuous bijection between topological spaces with a continuous inverse.
- A **open cover** of M is a collection of open sets in U_a whose union $\cup U_a = M$.
- A space is **locally Euclidean of dimension** n if every point has a neighborhood homeomorphic to an open subset of $\mathbb{R}^n$.
- A **chart** is a neighborhood U of a point p together with a homeomorphism $\phi : U \to V$ for an open subset V of $\mathbb{R}^n$. The chart is written as the pair (U, ϕ).

- A function is **smooth** if it is C^∞.

2.2 Topological manifolds.

§5.1 in [2].

Definition 2.1. We simply concatenate several definitions to define a **topological manifold of dimension *n***: a Hausdorff, second countable, locally Euclidean space of dimension n. Intuitively, this means that if you are standing on a topological manifold with terrible eyesight, as far as you can tell, you are standing in $\mathbb{R}^n$. We also require a few topological properties that make the manifold nicer to deal with.

Remark 2.2. Note that we require each neighborhood on a manifold to be homeomorphic to an open subset of $\mathbb{R}^n$. An object with some neighborhoods homeomorphic to open subsets of $\mathbb{H}^n$ but not $\mathbb{R}^n$ is called a **manifold with boundary**, which we will not consider a manifold.

Example 2.3. Two nonintersecting open line segments $M = \{(x, 0) : x \in (-1, 0) \cup (0, 1)\}$ form a 1-manifold in $\mathbb{R}^2$. Note that no part of the definition requires the manifold to be connected. Here, we have two components, each locally homeomorphic to an open neighborhood of $\mathbb{R}^1$.

We will prove an non-example of a manifold, but first, we need a lemma.

Lemma 2.4. Homeomorphisms preserve connected components.

Proof. Let X and Y be topological manifolds and $X = \cup X_i$ and $Y = \cup Y_j$ be their decompositions into their connected components. Let $f : X \to Y$ be a homeomorphism. Because f is a homeomorphism, f is continuous, so $f(X_i)$ is connected. Therefore, $f(X_i) \subseteq Y_j$ for some Y_j. Now because f is a homeomorphism, f^{-1} is continuous, so $f^{-1}(Y_j)$ is connected. Also, as long as X_i is nonempty, $f^{-1}(Y_j) \cap X_i \neq \emptyset$. Therefore, $f^{-1}(Y_j) \subseteq X_i$, since X_i is connected. Applying f to both sides yields $Y_j \subseteq f(X_i)$, completing the proof that $f(X_i) = Y_j$. □

Non-example 2.5. The unit cross in $\mathbb{R}^2$, $M = \{(x, y) : x \in (-1, 1)\ \&\ y = 0 \cup y \in (-1, 1)\ \&\ x = 0\}$, is not a topological manifold.

Proof. Assume toward contradiction that there exists a local homeomorphism $f : M \to \mathbb{R}^n$ such that $\forall x \in M$, $\exists$ a neighborhood $U \ni x$ such that $f(U)$ is open in $\mathbb{R}^n$ and $f\big|_U$ is a homeomorphism. Let $x = 0$. Then $\exists V$ a neighborhood containing x such that $f\big|_V$ is a homeomorphism. Put $f(V) = Y$. Define f' as the restriction of f to $V \setminus \{0\}$. This is a homeomorphism $V \setminus \{0\} \to Y \setminus \{f(0)\}$ because restrictions of homeomorphisms are still homeomorphisms. However, $V \setminus \{0\}$ has 4 connected components, while $Y \setminus \{f(0)\}$ has 2 components for $n = 1$ and 1 component for $f > 1$. Therefore, there exists no homeomorphism $f : M \to \mathbb{R}^n$, so M is not a topological manifold. □

2.3 Smooth manifolds.

§5.2 and §5.3 in [2].

Definition 2.6. Two charts (U, ϕ) and (U, ψ) are **compatible** if $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are both C^∞.

In other words, we can start in Euclidean space, go up to the manifold via the inverse of one map, and come back down via the other map, all without trouble.

Remark 2.7. Obvserve that if $U \cap V = \emptyset$, then the functions $\phi \circ \psi^{-1}$ and $\psi \circ \phi^{-1}$ are trivially C^∞, so this definition is only restrictive for charts on nondisjoint neighborhoods.

Definition 2.8. An **atlas** on a topological manifold M is a collection $\mathcal{U} = \{(U_a, \phi_a)\}$ of pairwise compatible charts such that $\cup\{U_a\}$ forms an open cover of M.

Definition 2.9. An atlas (U_a, ϕ_a) is **maximal** if it is not the proper subset of any atlas on M. (There is a fair amount of discussion about the philosophical status of this definition, but I will spare you the details and remark that I have never seen it play a crucial part in any interesting result. Any atlas admits extension to a unique maximal atlas, so we need not worry about which maximal atlas to proceed with, allowing this requirement to disappear into the background).

Finally, our main definition:

Definition 2.10. A **smooth manifold** is a topological manifold M together with a maximal atlas.

Example 2.11. Any open subset of the Euclidean space $S \subseteq \mathbb{R}^n$ is a topological manifold with chart (S, id).

Example 2.12. The graph of $y = |x|$ in $\mathbb{R}^2$ is also a smooth manifold of dimension 1 with the coordinate map $(x, |x|) \mapsto x$. The cusp at $x = 0$ does not prevent the graph from being a smooth 1-manifold: the projection down to the x-axis remains smooth.

Example 2.13. The sphere S^1 is a 1-manifold in $\mathbb{R}^2$. Define four open neighborhoods $U_+ = \{(x, y) \in S^1 : x > 0\}$, $U_- = \{(x, y) \in S^1 : x < 0\}$, $V_+ = \{(x, y) \in S^1 : y > 0\}$, and $V_- = \{(x, y) \in S^1 : y < 0\}$, with maps projecting onto the axis of the variable not restricted. Clearly, these neighborhoods cover S^1, and each is homeomorphic to an open subset of $\mathbb{R}^1$.

We define here two basic sets of matrices.

Definition 2.14. $M(n, \mathbb{K})$ is the set of all $n \times n$ matrices with entries in $\mathbb{K}$.

Remark 2.15. From now on, we let $\mathbb{K} = \mathbb{C}$ and elide the field in our definition of matrix groups: we write $M(n)$ instead of $M(n, \mathbb{C})$, $GL(n)$ instead of $GL(n, \mathbb{C})$, and so on, flagging each case when the field is not $\mathbb{C}$. Most but not all of the results below hold also for real matrices.

Definition 2.16. $GL(n)$ is the set of all $n \times n$ matrices with nonzero determinant.

To prove the next theorem, we need a lemma:

Lemma 2.17. Any open subset of a smooth manifold is a smooth manifold.

Proof. Let M' be an open subset of M and M a smooth manifold with atlas $\{(U_a, \phi_a)\}$. Then $\{(U_a \cap M', \phi_a)\}$ is an atlas of M', so M' is itself a smooth manifold. □

Theorem 2.18. *$GL(n)$ is a smooth $2n^2$ manifold.*

Proof. To begin, we identify $M(n)$ with $\mathbb{R}^{2n^2}$. $\mathbb{R}^{2n^2}$ is trivially a smooth manifold: take atlas $\{(\mathbb{R}^{2n^2}, id)\}$. Thus, $M(n)$ is a smooth manifold. The determinant $\det : GL(n) \to \mathbb{R}^2 \setminus \{0\}$ is polynomial in the entries of the matrix of which it is taken, so it is continuous. Its image is an open subset of $\mathbb{R}^2$. The preimage of any continuous map to an open set is open, so the preimage of det, $GL(n)$, is an open subset of $M(n)$. Therefore, by 2.17, $GL(n)$ is a smooth manifold of dimension $2n^2$. □

3 Matrix groups.

Before proceding to matrix Lie group theory, we prove some crucial results of a few classical matrix groups.

3.1 Maximality of the general linear group.

We first prove that $GL(n)$ is the maximal matrix group, then go on to prove that several other classical matrix groups are subgroups.

Theorem 3.1. *$GL(n)$ is the maximal matrix group.*

Proof. $GL(n)$ inherits associativity from the definition of matrix multiplication.
Recall from linear algebra that $\det(MN) = \det(M)\det(N)$. Let $M, N \in GL(n)$. By the definition of $GL(n)$, $\det(M) \neq 0$ and $\det(N) \neq 0$. Then $\det(MN) = \det(M)\det(N) \neq 0$, so $MN \in GL(n)$.
Also, $\det(I) \neq 0$ and $MI = IM = M, \forall AM \in GL(n)$, so $GL(n) \ni I$.
Finally, recall from linear algebra that a square matrix with nonzero determinant possesses an inverse. Then $\forall M \in GL(n), \exists M^{-1} : MM^{-1} = M^{-1}M = I$. This implies $\det(M^{-1}) = 1/\det(M)$, so $\det(M^{-1}) \neq 0$, so $M^{-1} \in GL(n)$. Therefore, $GL(n)$ is a matrix group.
Recall from linear algebra that nonsquare matrices and matrices with determinants of 0 lack inverses. Then any set with such a matrix fails to provide an inverse for each of its elements, so it is not a group.
$\therefore$ $GL(n)$ is a matrix group and any set of matrices with an element not in $GL(n)$ is not a matrix group, so $GL(n)$ is the maximal matrix group. □

3.2 Classical matrix group definitions.

First, we define these classical matrix groups.

Definition 3.2. $SL(n) = \{M \in GL(n) : \det(M) = 1\}$.

Definition 3.3. $SO(n, \mathbb{R}) = \{M \in GL(n, \mathbb{R}) : M^T M = MM^T = I \text{ and } \det(M) = 1\}$. This is a very important matrix group, so we define it and prove that it is a subgroup of $GL(n, \mathbb{R})$, even though the rest of the paper deals with complex matrices.

Definition 3.4. $SU(n) = \{M \in GL(n) : MM^* = M^*M = I \text{ and } \det(M) = 1 \text{ and } M_{ij} \in \mathbb{C}\}$.

Definition 3.5. There are several ways of defining $SP(2n)$, but we will take as definitional $SP(2n) = \{M \in GL(2n) : M^T \Omega M = \Omega\}$, with

$$\Omega = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

Note the $2n$: this definition breaks if we tried $SP(3)$, for example.

3.3 Classical matrix groups are matrix groups.

Proposition 3.6. $SL(n)$ is a matrix group.

Proof. $SL(n)$ inherits associativity from the definition of matrix multiplication.
For $M, N \in SL(n)$, $\det(MN) = \det(M)\det(N) = 1 \cdot 1 = 1$, so $MN \in SL(n)$, so $SL(n)$ is closed under composition.
For $M \in SL(n)$, $\det(M^{-1}) = 1/\det(M) = 1/1 = 1$, so $M^{-1} \in SL(n)$.
$\det(I) = 1$, so $SL(n) \ni I$.
$\therefore$ $SL(n)$ is associative, closed under inversion and composition, and contains the identity, so $SL(n)$ is a matrix group. $\square$

Proposition 3.7. $SO(n, \mathbb{R})$ is a matrix group.

Proof. $SO(n, \mathbb{R})$ inherits associativity from the definition of matrix multiplication.
For $M \in SO(n, \mathbb{R})$, $MM^T = I$ by definition, so $(MM^T)^{-1} = I^{-1} = I$. Also

$$\begin{aligned} (MM^T)^{-1} &= (M^T)^{-1} M^{-1} \\ &= (M^{-1})^T M^{-1} \\ &= M^{-1}(M^{-1})^T. \end{aligned}$$

Each equality is given by a straightforward property of matrices from linear algebra. This implies that $M^{-1}(M^{-1})^T = I$. $\det(M^{-1}) = 1$ is provided by $SO(n, \mathbb{R}) \subseteq SL(n)$. The $M_{ij} \in \mathbb{R}$ condition is

provided by the closure of $\mathbb{R}$ under negation and division, which are the operations on the entries that produce inversion. $M^{-1} \in SO(n,\mathbb{R})$. Therefore, $SO(n,\mathbb{R})$ is closed under inversion.
For $M, N \in SO(n,\mathbb{R})$:

$$\begin{aligned}(MN)(MN)^T &= MN(N^T M^T)\\ &= M(NN^T)M^T\\ &= MIM^T\\ &= MM^T\\ &= I.\end{aligned}$$

Again each equality is provided by a straightforward property of matrices from linear algebra. $\det(MN) = 1$ is provided by $SO(n,\mathbb{R}) \subseteq SL(n)$. The $M_{ij} \in \mathbb{R}$ condition is provided by the closure of $\mathbb{R}$ under addition and multiplication, the operations on the entries that produce composition.
$II^T = II = I$, so $SO(n,\mathbb{R}) \ni I$.
$\therefore$ $SO(n,\mathbb{R})$ is associative, closed under inversion and composition, and contains the identity, so $SO(n,\mathbb{R})$ is a matrix group. $\square$

Proposition 3.8. $SU(n)$ is a matrix group.

Proof. This proof is identical to the preceding one, *mutatis mutandis*. $\square$

Proposition 3.9. $SP(2n)$ is a matrix group.

Proof. $SP(2n)$ inherits associativity from the definition of matrix multiplication.
Let $M \in SP(2n)$, so $M^T \Omega M = \Omega$. Observe that $\Omega^{-1} = -\Omega$. Then

$$\begin{aligned}(M^T\Omega M)^{-1} &= \Omega^{-1}\\ &= -\Omega,\end{aligned}$$

so $-(M^T\Omega M)^{-1} = \Omega$, which allows finally

$$\begin{aligned}-(M^T\Omega M)^{-1} &= -M^{-1}\Omega^{-1}(M^T)^{-1}\\ &= M^{-1}\Omega(M^T)^{-1}\\ &= \Omega.\end{aligned}$$

Multiplying the penultimate and ultimate terms by M on the left and by M^T on the right gives

$$\Omega = M\Omega M^T,$$

showing that $M^T \in SP(2n)$. Then we may replace M with M^T in any of the above equations without jeopardizing the equality. Take $\Omega = M^{-1}\Omega(M^T)^{-1}$ and replace M with M^T:

$$\begin{aligned}(M^T)^{-1}\Omega((M^T)^T)^{-1} &= (M^T)^{-1}\Omega M^{-1}\\ &= (M^{-1})^T\Omega M^{-1}\\ &= \Omega,\end{aligned}$$

showing that $M^{-1} \in SP(2n)$, which is (finally) the desired result. Therefore, $SP(2n)$ is closed under inversion.

For $M, N \in SP(2n)$,

$$\begin{aligned}(MN)^T\Omega(MN) &= N^T M^T \Omega MN\\ &= N^T\Omega N\\ &= \Omega,\end{aligned}$$

so $MN \in SP(2n)$, so $SP(2n)$ is closed under composition.

Trivially, $I^T\Omega I = I\Omega I = \Omega$, so $SP(2n) \ni I$.

$\therefore$ $SP(2n)$ is associative, closed under inversion and composition, and contains the identity, so $SP(2n)$ is a matrix group. □

Before moving on to matrix Lie groups, we emphasize that this section has established that $SL(n)$, $SO(n, \mathbb{R})$, $SU(n)$, and $SP(n)$ are subgroups of $GL(n)$ (stipulating that n is even in the case of $SP(n)$, respecting our previous notation). With Cartan's Theorem, to be established later, this gets us further than we might expect.

4 Matrix Lie groups.

4.1 Definitions and discussion.

There are at least three definitions of "matrix Lie group" that could come into effect here.

Definition 4.1. A matrix Lie group is a matrix group that is also a Lie group.

In particular, this means that the group is realized as a set of matrices satisfying the group axioms, that the group has smooth inversion and composition maps, and also that the group has smooth manifold structure. This is the most honest and least insightful definition. To use this with no other machinery, for example, we might need to build a new atlas for each matrix Lie group by hand.

Definition 4.2. A matrix Lie group is a subgroup of a matrix group and a submanifold of a manifold.

This definition is slightly smarter. It does not require us to establish all the structure for our matrix Lie group *de novo*. Instead, all it asks is that we show that the set in question can inherit all the relevant properties from superset. Allowing a matrix group to be a subgroup of itself and a manifold to be a submanifold of itself, this definition is also perfectly broad.

Definition 4.3. A matrix Lie group is a subgroup of $GL(n)$ closed under nonsingular limits.

We have already established that $GL(n)$ is the maximal matrix group and that $GL(n)$ is a smooth manifold, so this definition should not come completely out of the blue. However, it should still be surprising. For any subgroup G of $GL(n)$, as long as any convergent sequence of matrices $M_n \in G$ converges to a matrix in G or leaves $GL(n)$ altogether, G is a matrix Lie group. No direct proof of its manifold structure is necessary. This is the route that allows one to study Lie theory in a surprising amount of depth without ever touching manifolds.

4.2 Direct proof: the general linear group is a Lie group.

It would be cheating to go through this whole paper without a single complete and direct proof that some object is a Lie group. It has already been shown that $GL(n)$ is a matrix group and a smooth manifold, so all that remains is to show that the inversion and composition maps are smooth.

Theorem 4.4. *$GL(n)$ is a Lie group.*

Proof. Let $\mu : GL(n) \times GL(n) \to GL(n)$ be the matrix multiplication ($\sim$ group composition) map. Then $\mu(A, B) = AB$, with entries calculated

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

This is a polynomial in the entries of A and B, so it is C^∞.
Let $\iota : GL(n) \to GL(n)$ be the matrix inversion ($\sim$ group inversion) map. Recall from linear algebra that the (i, j)-minor, denoted $M_{i,j}$, of a matrix is the determinant of the submatrix formed by deleting row i and column j of the matrix. The formula for entries of the inverse of a matrix A is

$$\begin{aligned}(\iota(A))_{ij} &= (A^{-1})_{ij} \\ &= \frac{1}{\det(A)}(-1)^{i+j} M_{j,i}\end{aligned}$$

by Cramer's rule. $M_{j,i}$ is a polynomial in the entries of A, so it is C^∞, and $1/\det(A)$ is also C^∞ as long as $\det(A) \neq 0$, which is assured by the fact that $A \in GL(n)$. Therefore this formula is C^∞ in the entries of A, so the version map is C^∞.
This completes the proof that $GL(n)$ is a Lie group. □

5 Matrix exponential and logarithm functions.

We now pivot to build the machinery that will be used to connect Lie groups to Lie algebras. (From chapter 2 in [1].)

5.1 Matrix exponential function.

Recall the power series of e^x for $x \in \mathbb{C}$:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This is a theorem due to Euler in the scalar case, but we will take it as a definition in the matrix case.

Definition 5.1.

$$\begin{aligned} e^X = \exp(X) &= \sum_{n=0}^{\infty} \frac{X^n}{n!} \\ &= I + \sum_{n=1}^{\infty} \frac{X^n}{n!}. \end{aligned}$$

Some texts use e^X rather than $\exp(X)$. The latter is preferred here because it can be confusing to conflate the scalar- and matrix-valued functions.

It is not yet clear that this is a sensible notion, for we do not know whether the power series converges.

Definition 5.2. To help prove the convergence of this power series, we define the **Hilbert-Schmidt norm** of a matrix:

$$\|X\| = \left[\sum_{i,j=1}^{n} (x_{ij})^2 \right]^{1/2}.$$

The following two inequalities follow quickly from properties of the metric on the field over which $M(n, \mathbb{K})$ is defined, and are not proven here:

$$\begin{aligned} \|X + Y\| &\leq \|X\| + \|Y\| \\ \|XY\| &\leq \|X\|\|Y\|. \end{aligned}$$

By induction on the second inequality, setting $X = Y$, we also have $\|X^n\| \leq \|X\|^n$.

Theorem 5.3. $\exp(X)$ *converges absolutely and is continuous* $\forall X \in M(n)$.

Proof. We say that a sequence of matrices $\{X_m\}$ converges to X if $(X_m)_{ij} \to (X)_{ij}$ as $m \to \infty$. It follows straightforwardly that X_m converges to X if and only if $\|X_m - X\| \to 0$ as $m \to \infty$. Then

$$\sum_{n=0}^{\infty} \frac{\|X^n\|}{n!} \le \sum_{n=0}^{\infty} \frac{\|X\|^n}{n!}.$$

The right-hand side is the power series for $e^{\|X\|}$ with $\|X\| \in \mathbb{R}$, which converges absolutely. Therefore, $\exp(X)$ converges absolutely.

Note also that each entry in X^n is a product of the entries of X, so X^n is a continuous function of X, $\forall n \in \mathbb{N}$, so the partial sums are continuous. By the Weierstrass M-test, $\exp(X)$ converges uniformly for matrices with norms in $(0, \|X\|)$. For any matrix $M \in M(n)$, we can choose $M + \varepsilon$ so that exp converges uniformly on an open set containing M. Therefore by the uniform convergence theorem, $\exp(X)$ is continuous on all of $M(n)$. □

Proposition 5.4. We now prove several important properties of the matrix exponential function.

1. $\exp(0) = I$.
2. If $XY = YX$, $\exp(X + Y) = \exp(X)\exp(Y) = \exp(Y)\exp(X)$.
3. $\exp(X)^{-1} = \exp(-X)$.
4. $\exp((\alpha + \beta)X) = \exp(\alpha X) \cdot \exp(\beta X)$ for $\alpha, \beta \in \mathbb{C}$.
5. $\forall C \in GL(n)$, $\exp(CXC^{-1}) = C\exp(X)C^{-1}$.

Proof. (1) follows straightforwardly from the definition of the power series beginning at $n = 1$.

To see (2), consider $\exp(X)\exp(Y)$, multiplying term-by-term in such a way that we collect all terms where powers add to m. Multiplying term-by-term is permitted because both series converge absolutely. This means:

$$\begin{aligned}
\exp(X)\exp(Y) &= \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{X^n}{n!} \frac{Y^{m-n}}{(m-n)!} \\
&= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} X^n Y^{m-n},
\end{aligned}$$

after multiplying by $m!/m!$ and rearranging. Now, note that exponentiation of matrices does not behave exactly like exponentiation of reals. For example

$$(X + Y)^2 = X^2 + XY + YX + Y^2,$$

which does not equal the familiar form

$$\sum_{n=0}^{2} \binom{m}{n} X^n Y^{m-n}$$

if X and Y fail to commute. In this case, however, since X and Y commute,

$$(X+Y)^m = \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} X^n Y^{m-n}.$$

Plugging in to the previous equation

$$\begin{aligned} \exp(X)\exp(Y) &= \sum_{m=0}^{\infty} \frac{(X+Y)^m}{m!} \\ &= \exp(X+Y). \end{aligned}$$

To prove (3), let $Y = -X$. We know that $-XX = X(-X)$, so $-X$ and X commute, so (2) applies. Then

$$\begin{aligned} \exp(-X+X) &= \exp(-X)\exp(X) \\ &= \exp(0) \\ &= I, \end{aligned}$$

so $\exp(-X) = \exp(X)^{-1}$. This implies that $\exp(X) \in GL(n)$, $\forall X \in M(n)$.
(4) is also a special case of (2), since αX is equivalent to $(\alpha I)X$, and αI commutes with $\forall X \in M(n)$.
To prove (5), note that $(CXC^{-1})^n = CX^nC^{-1}$ is a basic result in linear algebra. This implies that the power series of $\exp(CXC^{-1})$ and $C\exp(X)C^{-1}$ are term-by-term equivalent. □

Proposition 5.5. Let $X \in M(n)$. Then for $t \in \mathbb{R}$, $\exp(tX)$ is a smooth curve in $M(n)$, and

$$\begin{aligned} \frac{d}{dt}\exp(tX) &= X\exp(tX) \\ &= \exp(tX)X. \end{aligned}$$

This implies

$$\frac{d}{dt}\exp(tX)\Big|_{t=0} = X,$$

since $\exp(0) = I$.

Proof. This follows simply from differentiating the power series term by term. This is permitted because each entry $(\exp(tX))_{jk}$ of $\exp(tX)$ is given by a convergent power series in t, and one can differentiate a power series term by term within its radius of convergence. This holds for all entries of the matrix, so it holds for the matrix as a whole. □

Proposition 5.6. For all $X, Y \in M(n)$, we have

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

Proof. The proof is long and not very interesting and, moreover, requires a fairly substantial lemma that is also not very interesting, so it is omitted with a reference to ([1], 39–40). □

Proposition 5.7. For $\forall X \in M(n)$, we have

$$\det(\exp(X)) = e^{tr(X)}.$$

Proof. Suppose X is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $X = CDC^{-1}$ for D diagonal, so $\exp(X) = \exp(CDC^{-1}) = C\exp(D)C^{-1}$ by proprty (5). Clearly $\exp(D)$ is a diagonal matrix with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$. Then $\det(X) = \det(D) = e^{\lambda_1} \cdot \ldots \cdot e^{\lambda_n}$. On the other hand, $tr(X) = tr(D) = \lambda_1 + \ldots + \lambda_n$. Therefore $e^{tr(X)} = e^{tr(D)} = e^{\lambda_1+\ldots+\lambda_n} = e^{\lambda_1} \cdot \ldots \cdot e^{\lambda_n}$.
If $X \in M(n, \mathbb{C})$ is not diagonalizable, then there is a sequence $\{D_n\}$ with $\lim_{n\to\infty} D_n = D$ such that $X = CDC^{-1}$ for some invertible matrix C. Then also $tr(X) = tr(D)$ and $\det(X) = \det(D)$, so the proof goes through identically. (This theorem does not hold for $X \in M(n, \mathbb{R})$ because an arbitrary real matrix cannot be approximated by a diagonal matrix.) □

5.2 Matrix logarithm function.

It is also necessary to define the matrix logarithm function, which, as we will see, is the inverse of the matrix exponential in its radius of convergence.
For $z \in \mathbb{C}$, recall the power series

$$\log(z) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(z-1)^n}{n},$$

which we state but do not prove is defined and holomorphic in a circle of radius 1 about $z = 1$. (Proof in [1] 36–7.) This function has the following two crucial properties:

$$e^{\log(z)} = z,$$

for z with $|z-1| < 1$, and

$$\log(e^u) = u$$

for u with $|u| < \log 2$. (This condition means $|e^u - 1| < 1$).

Definition 5.8. Analogously, we define for $A \in M(n)$

$$\log(X) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(X-I)^n}{n}$$

whenever the series converges. Since the series has radius of convergence 1 for $z \in \mathbb{C}$ and $\|(X-I)^n\| \leq \|X-I\|^n$ for $n \geq 1$, if $\|X-I\| < 1$, the matrix-valued series converges and is continuous. Note that even outside that radius the series converges if $X - I$ is nilpotent, that is, if $\exists n : (X-I)^n = 0$. (We call such an X **unipotent**.)

Proposition 5.9. Within this radius of convergence,

$$\exp(\log(X)) = X,$$

and for all $X \in M(n)$ with $\|X\| < \log(2)$, $\|\exp(X)\| < 1$ and

$$\log(\exp(X)) = X.$$

We state but do not prove this theorem of which a proof exists in ([1], 38–9).

5.3 One-parameter subgroups.

With the matrix exponential and logarithm functions defined, we can now define one-parameter subgroups, which are used in generating the Lie algebra of a Lie group.

Definition 5.10. A one-parameter subgroup of $GL(n)$ is a group homomorphism $A : \mathbb{R}^+ \to GL(n)$. This implies the following:

- A is continuous.
- $A(t+s) = A(t)A(s)$.
- $A(0) = I$.

We state but do not prove the following lemma, of which a proof be found in ([1], 41–2).

Lemma 5.11. Fix some ε with $\varepsilon < \log 2$. Let $B_{\varepsilon/2}$ be the ball of radius $\varepsilon/2$ around the 0 in $M(n)$, and let $U = \exp(B_{\varepsilon/2})$. Then every $X \in U$ has a unique square root Y in U, given by $Y = \exp(\frac{1}{2}\log X)$.

Intuitively, this says that if we take some sufficiently small ball around the origin as the preimage of exp, every matrix X in the image has a unique square root Y also in the image, which is found simply by $Y = \exp(\frac{1}{2}\log X)$. It is easy to check that $Y^2 = X$, so the only real work is to prove uniqueness.

With this, we can prove the crucial result relating one-parameter subgroups to the matrix exponential function.

Theorem 5.12. *If A is a one-parameter subgroup of $GL(n)$, there exists a unique $n \times n$ matrix such that $A(t) = \exp(tX)$.*

Proof. Uniqueness is immediate: if $\exists X : \exp(tX) = A(t)$, then $X = \frac{d}{dt}A(t)\big|_{t=0}$. Now to prove existence. Let $U = \exp(B_{\varepsilon/2})$ as in the lemma, so U is an open set in $GL(n)$. By the continuity of A, $\exists t_0 > 0$ such that $A(t) \in U$ for all $t : |t| \leq t_0$. In other words, A is continuous, and its output is in U at $t = 0$, so if we keep t within some t_0 of 0, the output will stay within U. Now define

$$X = \frac{1}{t_0}\log(A(t_0)),$$

which means that

$$t_0 X = \log(A(t_0)).$$

Because $\log(A(t_0)) \in B_{\varepsilon/2}$, we also have $t_0 X \in B_{\varepsilon/2}$ and $\exp(t_0 X) = \exp(\log(A(t_0))) = A(t_0)$. Clearly, $A(t_0/2)$ is also in U, and $A(t_0/2)^2 = A(t_0)$ by a property of group homomorphisms (5.10). By 5.11, $A(t_0)$ has a unique square root in U, which is $\exp(t_0 X/2)$. This implies that $A(t_0/2) = \exp(t_0 X/2)$. By induction, $A(t_0/2^k) = \exp(t_0 X/2^k)$, $\forall k \in \mathbb{N}$. We have

$$\begin{aligned} A(mt_0/2k^2) &= A(t_0/2^k)^m \\ &= \exp(t_0 X/2^k)^m, \\ &= \exp(mt_0 X/2^k) \end{aligned}$$

with the first equality by a property of group homomorphisms (5.10), the second by $A(t_0/2^k) = \exp(t_0 X/2^k)$, and the third the property of the exponential function $\exp(M)^m = \exp(mM)$.

Therefore $A(t) = \exp(tX)$ for all $t = mt_0/2^k$. There exist m, t_0, k to recover any arbitrary t because the set of numbers of the form $t = mt_0/2^k$ is dense in $\mathbb{R}$. Moreover, $\exp(tX)$ and $A(t)$ are both continuous. Therefore, $A(t) = \exp(tX)$ for $t \in \mathbb{R}$. □

Proposition 5.13. The exponential map $\exp(X)$ is smooth (infinitely differentiable).

Proof. We have already proven that $\exp(tX)$ is smooth, but the present proposition is different: we are proving that we can take the derivative in the direction of an arbitrary matrix, which is stronger than taking the derivative with respect to the parameter t. Nonetheless, this proof proceeds similarly. Note that each entry $(X^m)_{jk}$ of X^m is a homogeneous polynomial of degree m in the entries of

X. Thus, the series for the function $(X^m)_{jk}$ has the form of a multivariable power series. Since the series converges on all of $M(n)$, it is permissible to differentiate the power series term-wise as many times as desired, which means that the function $(X^m)_{jk}$ is smooth. The smoothness of the exponential map follows immediately. □

We state but do not prove the following proposition, which is noteworthy but will not be used anywhere in this paper. The proof is sketched in ([1], 48)

Proposition 5.14. For $X \in GL(n)$, $\exists A \in GL(n) : \exp(A) = X$.

6 Lie algebras.

We now have the tools to make good use of the notion of a **Lie algebra**, which we will define then apply.

6.1 Definitions.

Definition 6.1. A Lie algebra $\mathfrak{g}$ over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is a $\mathbb{K}$-vector space with a bracket operation $[\cdot, \cdot]$ that satisfies the following properties:

- bilinearity: $[aX + bY, Z] = a[X, Z] + b[Y, Z]$ and $[Z, aX + bY] = a[Z, X] + b[Z, Y]$
- antisymmetry: $[X, Y] = -[Y, X]$
- Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

for $X, Y, Z \in \mathfrak{g}$ and $a \in \mathbb{K}$.

Remark 6.2. We use lowercase Gothic characters to denote Lie algebras, with the Lie algebra of a Lie group G as $\mathfrak{g}$.

Definition 6.3. The vector space of matrices in $M(n)$ with the bracket defined by the commutator

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

is denoted $\mathfrak{gl}(n)$, or $\mathfrak{gl}$.

Theorem 6.4. $\mathfrak{gl}$ *is a Lie algebra.*

Proof. We check that the commutator satisfies bilinearity, antisymmetry, and the Jacobi identity.

$[aX + bY, Z] = (aX + bY)Z - Z(aX + bY)$ by the definition of the commutator, which

$$
\begin{aligned}
&= aXZ + bYZ - ZaX - ZbY \\
&= aXZ + bYZ - aZX - bZY \\
&= aXZ - aZX + bYZ - bZY \\
&= a(XZ - ZX) + b(YZ - ZY) \\
&= a[X, Z] + b[Y, Z].
\end{aligned}
$$

The proof for linearity in the second coordinate is identical, *mutatis mutandis*.
Straightforwardly from the definition of the commutator, it follows that

$$
\begin{aligned}
[X, Y] &= XY - YX \\
&= -(YX - XY) \\
&= -[Y, X].
\end{aligned}
$$

Expanding by the definition of the commutator

$$
\begin{aligned}
&[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] \\
&= [X, YZ - ZY] + [Y, ZX - XZ] + [Z, XY - YX] \\
&= X(YZ - ZY) - (YZ - ZY)X + Y(ZX - XZ) - (ZX - XZ)Y + Z(XY - YX) - (XY - YX)Z \\
&= XYZ - XZY - YZX + ZYX + YZX - YXZ - ZXY + XZY + ZXY - ZYX - XYZ + YXZ \\
&= 0,
\end{aligned}
$$

cancelling like terms.
Therefore, the commutator is a valid bracket, so $\mathfrak{gl}(n)$ is a Lie algebra. □

The following theorem is significant but surprisigly difficult to prove, so we merely state it.

Theorem 6.5. *(**Ado's Theorem**.) Every finite-dimensional Lie algebra $\mathfrak{g}$ over a field $\mathbb{K}$ of characteristic zero is isomorphic to a Lie algebra of square matrices under the commutator bracket.*

This means that as long as we are working over non-pathological fields, $\mathfrak{gl}$ and its subalgebras are the only ones we need. This theorem is one reason that we can get further than we might expect using matrix Lie groups.

Example 6.6. The most familiar example of a Lie algebra is $\mathbb{R}^3$ equipped with the traditional cross-product. To prove that the cross product is a valid Lie bracket operation, it suffices to demonstrate that it is antisymmetric and follows the Jacobi identity on the basis vectors. Antisymmetry is definitional:

$$\hat{\mathbf{\imath}} \times \hat{\mathbf{\jmath}} = -\hat{\mathbf{\jmath}} \times \hat{\mathbf{\imath}},$$

$$\hat{\mathbf{\jmath}} \times \hat{\mathbf{k}} = -\hat{\mathbf{k}} \times \hat{\mathbf{\jmath}},$$

$$\hat{\mathbf{k}} \times \hat{\mathbf{\imath}} = -\hat{\mathbf{\imath}} \times \hat{\mathbf{k}}.$$

The Jacobi identity follows from computation:

$$\begin{aligned} \hat{\mathbf{\imath}} \times (\hat{\mathbf{\jmath}} \times \hat{\mathbf{k}}) + \hat{\mathbf{\jmath}} \times (\hat{\mathbf{k}} \times \hat{\mathbf{\imath}}) + \hat{\mathbf{k}} \times (\hat{\mathbf{\jmath}} \times \hat{\mathbf{\jmath}}) &= \hat{\mathbf{\imath}} \times \hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} \times \hat{\mathbf{\jmath}} + \hat{\mathbf{k}} \times \hat{\mathbf{k}} \\ &= 0. \end{aligned}$$

Remark 6.7. Any commutative algebra is also trivially a Lie algebra, where $[\cdot,\cdot] \equiv 0$ because $XY = YX$.

6.2 The Lie algebra of a matrix Lie group.

Definition 6.8. Let G be a matrix Lie group. The **"Lie algebra"** of G, denoted $\mathfrak{g}$, is the set of all matrices X such that $\exp(tX) \in G, \forall t \in \mathbb{R}$.

This definition states that the "Lie algebra" of a Lie group is the set of all matrices whose corresponding one-parameter subgroup lies entirely in G. We call it a "Lie algebra" for now because we have not shown that it is a Lie algebra according to the original definition.
Note that $\exp(X) \in G$ does not necessarily imply $X \in \mathfrak{g}$: our requirement is stronger.
We now show that $\mathfrak{g}$, our "Lie algebra," is indeed a Lie algebra.

Theorem 6.9. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then $\forall X, Y \in \mathfrak{g}$, the following results hold.*

- $\forall A \in G, AXA^{-1} \in \mathfrak{g}$.
- $\forall s \in \mathbb{R}, sX \in \mathfrak{g}$.
- $X + Y \in \mathfrak{g}$.
- $XY - YX \in \mathfrak{g}$.

It follows from (2) and (3) that $\mathfrak{g}$ is a vector space, and from (4) that $\mathfrak{g}$ is closed under the bracket $[X, Y] = XY - YX$, so it is a Lie algebra.

Proof. For (1), recall that

$$\exp(t(AXA^{-1})) = A\exp(tX)A^{-1}$$

by 5.4. This is in G because all three of its terms are in G.

For (2), note that $\exp(t(sX)) = \exp((ts)X)$, which is in G by the definition of $\mathfrak{g}$. This implies that $sX \in \mathfrak{g}$.
For (3), note that

$$\exp(t(X+Y)) = \lim_{t\to\infty} [\exp(tX/m)\exp(tY/m)]^m$$

by 5.6. Then $(\exp(tX/m))$ and $(\exp(tX/m))$ are in G. We know that G is closed under composition, so $\exp(tX/m)\exp(tY/m) \in G$. Exponentiation is repeated composition and, again, G is closed under composition, so $[\exp(tX/m)\exp(tY/m)]^m \in G$. By 5.4 $\exp(t(X+Y))$ in $GL(n)$. G is defined as a matrix Lie group, so by 4.3 it is closed in $GL(n)$. Therefore $\exp(t(X+Y)) \in G$. This finally shows that $\exp(t(X+Y)) \in G$, so $X+Y \in \mathfrak{g}$.
For (4), let $X, Y \in \mathfrak{g}$ and consider

$$\begin{aligned} \frac{d}{dt}(\exp(tX)Y\exp(-tX))\Big|_{t=0} &= (XY)\exp(0) + (\exp(0)Y)(-X) \\ &= XY - YX. \end{aligned}$$

Now by (1), $\exp(tX)Y\exp(-tX) \in G, \forall t \in \mathbb{R}$. Moreover, by (2) and (3), $\mathfrak{g}$ is a real subspace of $M(n)$, so it is topologically closed, so

$$\lim_{t\to 0} \frac{\exp(tX)Y\exp(-tX) - Y}{h}$$

remains in the subspace. This is the definition of the derivative of $\exp(tX)Y\exp(-tX)$ at $t = 0$, which has just been shown to equal $XY - YX$. Therefore $\mathfrak{g}$ is closed under the bracket $[X, Y] = XY - YX$.
This completes the proof that $\mathfrak{g}$ is a Lie algebra, so our notion of "Lie algebra" is indeed a Lie algebra, and we can remove the scare quotes. □

Remark 6.10. We do not have the space to do much with the bracket of a Lie algebra in this paper, so I would at least like to remark on it here. Our bracket, the commutator, is easily interpreted as a measure of non-abelian-ness: if X and Y commute, $[X, Y] = 0$, while if XY and YX differ significantly, then $[X, Y] = XY - YX$ is very large. The purpose of the bracket, then, is to recover some of the non-abelian structure of the group. If we did not endow our Lie algebra with a bracket, our only operation would be vector addition, which is commutative, so the algebra would obliterate all of the non-abelian structure of the group. For example, we saw above that $\exp(X)\exp(Y) = \exp(X+Y)$ if and only if X and Y commute. The full formula is the **Baker-Campbell-Hausdorff formula**:

$$\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \ldots),$$

continuing by adding higher-order compositions of the commutator. Thus, the bracket allows us to reprise non-commutative group operations on our vector space.

Two straightforward facts about the correspondence between Lie groups and Lie algebras follow. Then, we prove a theorem that gets us the rest of the significant results of this paper.

Definition 6.11. The identity component of G, denoted G_0, is the connected component of G containing the identity. In the context of matrix Lie groups, connectedness is equivalent to path-connectedness, so G_0 is also the path-connected component of identity.

Proposition 6.12. Let G be a matrix Lie group and $X \in \mathfrak{g}$ an element of its Lie algebra. Then $\exp(X) \in G_0$.

Proof. By the definition of $\mathfrak{g}$, $\exp(tX) \in G$, $\forall t \in \mathbb{R}$. Then for $t : 0 \to 1$, $\exp(tX) : I \to \exp(X)$, so I and $\exp(X)$ are path-connected, so $\exp(X) \in G_0$. □

Proposition 6.13. If G is commutative, then $\mathfrak{g}$ is commutative.

Proof. For $X, Y \in M(n)$, we can calculate

$$[X, Y] = \frac{d}{dt}\left(\frac{d}{ds} \exp(tX) \exp(sY) \exp(-tX)\Big|_{s=0}\right)\Big|_{t=0}.$$

If $X, Y \in \mathfrak{g}$ and G is commutative, then $\exp(tX)$ commutes with $\exp(sY)$, giving

$$\begin{aligned}[X, Y] &= \frac{d}{dt}\left(\frac{d}{ds} \exp(tX) \exp(-tX) \exp(sY)\Big|_{s=0}\right)\Big|_{t=0} \\ &= \frac{d}{dt}\left(\frac{d}{ds} \exp(sY)\Big|_{s=0}\right)\Big|_{t=0}.\end{aligned}$$

We are differentiating a function that is independent of t with respect to t, so $[X, Y] \equiv 0$, so $\mathfrak{g}$ is commutative. □

Remark 6.14. The reverse direction requires additionally that G be connected. It will be shown shortly.

Lemma 6.15. Let $\{B_m\}$ be a sequence of matrices in G such that $B_m \to I$ as $m \to \infty$. Define $Y_m = \log B_m$, which is defined for all sufficiently large m because log is defined around I. Suppose $Y_m \neq 0, \forall m$: this is equivalent to supposing $B_m \neq I, \forall m$. Define further that $Y_m / \|Y_m\| \to Y \in M(n)$ as $m \to \infty$. Then $Y \in \mathfrak{g}$.

Proof. For any $t \in \mathbb{R}$, $(t/\|Y_m\|)Y_m \to tY$ by construction. $B_m \to I$, so $\|Y_m\| \to 0$. Then we can construct a sequence k_m such that $k_m\|Y_m\| \to t$. Then

$$\exp(k_m Y_m) = \exp\left[(k_m\|Y_m\|)\frac{Y_m}{\|Y_m\|}\right] \to \exp(tY),$$

since the parentheses within the bracket approach t and the fraction within the bracket approaches Y. On the other hand,

$$\exp(k_m Y_m) = (\exp(Y_m))^{k_m}$$

by a property of the exponential map. This equals $(B_m)^{k_m}$ by the definition of B_m, which is in G. This implies that $\exp(tY) \in G$. Then by definition $Y \in \mathfrak{g}$. $\square$

Theorem 6.16. *For* $0 < \varepsilon < \log 2$, *let* $U_\varepsilon = \{X \in M(n) : \|X\| < \varepsilon\}$ *and let* $V_\varepsilon = \exp(U_\varepsilon)$. *Suppose* $G \subseteq GL(n)$ *is a martrix Lie group with Lie algebra* $\mathfrak{g}$. *Then* $\exists \varepsilon \in (0, \log 2)$ *such that* $\forall A \in V_\varepsilon$, $A \in G$ *if and only if* $\log A \in \mathfrak{g}$.

Proof. Begin by identifying $M(n)$ with $\mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$. Let $\mathfrak{g}^\perp$ denote the orthogonal complement of $\mathfrak{g}$ with respect to the usual inner product on $\mathbb{R}^{2n^2}$. Let $Z = X \oplus Y$ with $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}^\perp$. Consider $\Phi : M(n) \to M(n)$ given by $\Phi(Z) = \Phi(X, Y) = \exp(X)\exp(Y)$. The exponential is smooth, so Φ is also smooth. By $D_Z\Phi(0)$ denote the derivive of Φ at 0 in the direction Z. Then

$$\begin{aligned} D_{(X,0)}\Phi(0,0) &= \frac{d}{dt}\Phi(tX, 0)\Big|_{t=0} \\ &= (X, 0), \end{aligned}$$

and

$$\begin{aligned} D_{(0,Y)}\Phi(0,0) &= \frac{d}{dt}\Phi(0, tY)\Big|_{t=0} \\ &= (0, Y), \end{aligned}$$

both by direct calculation, employing the definition $\Phi(X, Y) = \exp(X)\exp(Y)$. Then

$$\begin{aligned} D_Z\Phi(0) &= D_{(X,Y)}\Phi(0,0) \\ &= D_{(X,0)+(0,Y)}\Phi(0,0), \end{aligned}$$

first by the definition of Z and second simply by addition. By the linearity of the derivative, this equals

$$\begin{aligned} D_{(X,0)}\Phi(0,0) + D_{(0,Y)}\Phi(0,0) &= (X, 0) + (0, Y) \\ &= Z. \end{aligned}$$

This is all to say that $D_Z\Phi(0) = Z$. Clearly, then, the derivative of Φ is nonsingular at 0. Thus, the inverse function theorem applies.

By the inverse function theorem, Φ has a continuous inverse in a neighborhood of $\Phi(0) = I$.

Let $A \in V_\varepsilon \cap G$. Assume toward contradiction that $\log A \notin \mathfrak{g}$. By the local inverse of Φ permitted by the inverse function theorem, $A_m = \exp(X_m)\exp(Y_m)$ for sufficiently large m, with $X_m, Y_m \to 0$

as $m \to \infty$. Then $Y_m \neq 0$, otherwise we would have

$$\begin{aligned}\log A_m &= \log[\exp(X_m)\exp(Y_m)]\\ &= \log\exp(X_m) \in \mathfrak{g},\end{aligned}$$

which violates our supposition that $\log A \notin \mathfrak{g}$. However, $\exp(X_m), A_m \in G$, so defining

$$\begin{aligned}B_m &= \exp(-X_m)A_m\\ &= \exp(Y_m) \in G.\end{aligned}$$

The unit sphere in $\mathfrak{g}^{\perp}$ is compact, so there exists a subsequence $\{Y_m\}$ such that $Y_m/\|Y_m\|$ converges to $Y \in \mathfrak{g}^{\perp}$, where $\|Y\| = 1$. But by the lemma this implies that $Y \in \mathfrak{g}$. Yet $\mathfrak{g}^{\perp}$ is the orthogonal complement of $\mathfrak{g}$, so the two are only trivially nondisjoint, so $Y \in \mathfrak{g}$ and $Y \in \mathfrak{g}^{\perp}$ is a contradiction. Therefore, there must be ε such that $\log A \in \mathfrak{g}$ for all $A \in V_{\varepsilon} \cap G$. □

Corollary 6.17. If G is a matrix Lie group with corresponding Lie algebra $\mathfrak{g}$, then $\exists U$ a neighborhood of 0 in $\mathfrak{g}$ and a neighborhood V of I in G such that the exponential map takes U homeomorphicallly onto V.

Proof. Let ε sufficiently small that 6.2 holds. Set $U = U_{\varepsilon} \cap \mathfrak{g}$ and $V = V_{\varepsilon} \cap G$. Then 6.2 implies that $\exp : U \to V$ is surjective. Moreover, exp is a homeomorphism, since there is a continuous inverse map $\log\big|_V$. □

Corollary 6.18. (**Cartan's Theorem**). Any closed subgroup H of a Lie group G is a Lie subgroup (and thus a submanifold) of G.

Proof. By the previous corollary, $\exp^{-1} : U \to V$ is a diffeomorphism from some neighborhood of $I \in G$ to some neighborhood of $0 \in V$. This implies that $\exp^{-1}\big|_H : U \cap H \to V \cap \mathfrak{h}$ is a diffeomorphism from some neighborhood of H at the identity to some neighborhood of $\mathfrak{h}$ at the identity. Then $(H \cap U, \exp^{-1}\big|_H)$ is a chart, and we can use left translation to get a chart for any other $h \in H$ [5]. □

Corollary 6.19. Let G be a matrix Lie group with Lie algbebra $\mathfrak{g}$ and let k be the dimension of $\mathfrak{g}$ as a real vector space. Then G is a smooth submanifold of $M(n)$ of dimension k and hence a Lie group according to 4.2.

Proof. In the interest of space, this proof, fairly tedious, is omitted, with a reference to ([1], 71). □

Corollary 6.20. Suppose $G \subseteq GL(n)$ is a matrix Lie group with corresponding Lie algebra $\mathfrak{g}$. Then $X \in \mathfrak{g}$ if and only if $\exists \gamma$ a smooth curve in $M(n)$ with $\gamma(t) \in G, \forall t$ and such that $\gamma(0) = I$ and $d\gamma/dt|_{t=0} = X$.

Proof. The forward direction is easy. Define $\gamma(t) = \exp(tX)$. Then $\gamma(0) = I$ and $d\gamma/dt|_{t=0} = X$ by properties of exp already established. In the other direction, let $\gamma(t)$ be smooth with $\gamma(0) = I$. For sufficiently small t, $\gamma(t) = \exp(\delta(t))$, where δ is a smooth curve in $\mathfrak{g}$. The derivative of $\delta(t)$ at $t = 0$ is the same as the derivative of $f(t) : t \mapsto t\delta'(0)$ at $t = 0$. This trick simplifies the algebra. Then by the chain rule

$$\begin{aligned}\gamma'(0) &= \frac{d}{dt}\exp(\delta(t))\Big|_{t=0} \\ &= \frac{d}{dt}\exp(f(t))\Big|_{t=0} \\ &= \frac{d}{dt}\exp(t\delta'(0))\Big|_{t=0} \\ &= \delta'(0).\end{aligned}$$

Now $\gamma(t) \in G$ by construction, and $\exp(\delta(t)) = \gamma(t)$ for sufficiently small t, so $\delta(t) \in \mathfrak{g}$ for sufficiently small t by the definition of $\mathfrak{g}$. Then also $\delta'(0) \in \mathfrak{g}$, so $\gamma'(0) \in \mathfrak{g}$. □

Remark 6.21. This means that $\mathfrak{g}$ is the tangent space at the identity to G. Many textbooks work the other way, defining the Lie algebra as the tangent space at the identity and recovering other properties we already have.

To prove the next corollary, we require the following lemma, which we state but do not prove.

Lemma 6.22. Suppose $A : [a, b] \to GL(n)$ is continuous. Then $\forall\varepsilon > 0, \exists\delta > 0$ such that for s, t where $|s - t| < \delta$, $\|A(s)A(t)^{-1} - I\| < \varepsilon$.

Corollary 6.23. If G is a connected matrix Lie group, then every element A of G can be written in the form $A = \exp(X_1)\dots\exp(X_m)$ for $X_1, \dots, X_m \in \mathfrak{g}$.

Proof. Let $V_\varepsilon = \exp(U_\varepsilon)$ for U_ε a neighborhood of 0, as in 6.2. For $A \in G$, define a continuous path $\gamma : [0, 1] \to G$ where $\gamma(0) = I$ and $\gamma(1) = A$. By 6.22, we can pick $\delta > 0$ such that $\gamma(s)\gamma(t)^{-1} \in V_\varepsilon$ for $|s - t| < \delta$.

Next, we partition $[0, 1]$ into m pieces of size $1/m$, choosing m so that $1/m < \delta$. Then for $j \in \{1, \dots, m\}$, $\gamma((j - 1)/m)^{-1}\gamma(j/m) \in V_\varepsilon$ because the two arguments are within δ of each other. This implies that

$$\gamma((j - 1)/m)^{-1}\gamma(j/m) = \exp(X_j)$$

for some element $X_j \in \mathfrak{g}$. Then

$$\begin{aligned} A &= \gamma(1) \\ &= \gamma(0)\gamma(0)^{-1}\gamma(1) \\ &= I \cdot \gamma(0)^{-1}\gamma(1) \\ &= \gamma(0)^{-1}\gamma(1) \\ &= \gamma(0)^{-1}\gamma(1/m)\gamma(1/m)^{-1} \dots \gamma((m-1)/m)\gamma((m-1)/m)^{-1}\gamma(1) \\ &= \exp(X_1) \dots \exp(X_m) \end{aligned}$$

for $X_1, \dots X_m$ as constructed earlier. □

Corollary 6.24. If G is a connected matrix Lie group and the Lie algebra $\mathfrak{g}$ of G is commutative, then G is commutative.

Proof. Since $\mathfrak{g}$ is commutative, any two elements of G, when written as in 6.23, will commute. □

Corollary 6.25. If G is a matrix Lie group, the identity component $G_0 \subseteq G$ is a closed subgroup of $GL(n)$ and thus a matrix Lie group. Moreover, $\mathfrak{g}_0 = \mathfrak{g}$.

Proof. Take $\{A_m\}$, a sequence in G_0 converging to some $A \in GL(n)$. If G is a matrix Lie group, then G is closed under nonsingular limits by 4.3, so $A \in G$. Moreover, $A_m A^{-1} \in G$ for all m because G is a group. Also, $A_m A^{-1} \to I$ as $m \to \infty$ because $A_m \to A$. By 6.2, $A_m A^{-1} = \exp(X)$ for $X \in \mathfrak{g}$ for m large enough. Left-multiplying by $\exp(-X)$ and right-multiplying by A gives $\exp(-X)A_m = A$. Because $A_m \in G_0$ by construction, there is a path joining I to G_0. Since $\exp(-X)A_m = A$, the path $\exp(-tX)A_m$ connects A_m to A, letting $t : 0 \to 1$. Combining this path with the path from I to A_m provides a path from I to A, so $A \in G_0$. Therefore, G_0 is a closed subgroup of $GL(n)$, so it is a matrix Lie group.
Now, since $G_0 \subseteq G$, it follows that $\mathfrak{g}_0 \subseteq \mathfrak{g}$. Now, pick an arbitrary element X from $\mathfrak{g}$. By the definition of $\mathfrak{g}$, we have $\exp(tX) \in G, \forall t \in \mathbb{R}$. Consider an arbitrary element $Y = \exp(t_0 X) \in G$. Then $\exp(tX)$ connects I to Y, letting $t : 0 \to t_0$. Then $\mathfrak{g} \subseteq \mathfrak{g}_0$. Therefore, $\mathfrak{g}_0 = \mathfrak{g}$. □

7 Examples.

I am not sure how to select the curriculum for this somewhat artificial crash course on Lie groups and Lie algebras. I hope that the preceding section will give the reader some familiarity with the basics of the correspondence between Lie groups and Lie algebras. The original intention of this section was to conclude by providing several interesting examples. Here, I have kept the proofs of the closure of the classical matrix groups but removed all the examples I had of Lie groups that are not matrix Lie groups.

I am more than happy to include these in my final paper if you reviewers believe that the exposition would be aided by examples that resist the matrix-based approach. I thought it was a nice and tidy to finish with proofs that the classical groups are in fact matrix Lie groups, but I defer to your preferences. The proofs that these are not matrix Lie groups are pretty slick, but I doubted that anyone would complain that this paper lacked length, so I have excluded them for now.

The following proofs all go through very similarly: we prove that each matrix group is the preimage of a continuous function to a closed set.

Theorem 7.1. *$SL(n)$ is closed.*

Proof. The determinant map $\det : GL(n) \to \mathbb{R}^2$ is a polynomial in the entries of the input. The set $\{1\}$ is closed in $\mathbb{C}$. By definition $SL(n) = \{\det^{-1}(\{1\})\}$.
$\therefore$ $SL(n)$ is the preimage of a continuous function to a closed set, so it is closed. □

Theorem 7.2. *$SO(n, \mathbb{R})$ is closed.*

Proof. The determinant map $\det : GL(n, \mathbb{R}) \to \mathbb{R}$ is a polynomial in the entries of the input. Define $L : M(n, \mathbb{R}) \to M(n, \mathbb{R})$ by $L(M) = M^T M$ and $R : M(n, \mathbb{R}) \to M(n, \mathbb{R})$ by $R(M) = MM^T$. The sets $\{1\}$ and $\{I\}$ are closed. By definition $SO(n, \mathbb{R}) = \{L^{-1}(\{I\}) \cap R^{-1}(\{I\}) \cap \det^{-1}(\{1\})\}$.
$\therefore$ $SO(n, \mathbb{R})$ is the finite intersection of the preimage of continuous functions to closed sets, so it is closed ([4]). □

Theorem 7.3. *$SU(n)$ is closed.*

Proof. This proof is identical to that for $SO(n, \mathbb{R})$, *mutatis mutandis*. □

Theorem 7.4. *$SP(n)$ is closed.*

Proof. Define $S : M(2n) \to M(2n)$ by $S(M) = A^T \Omega A - \Omega$. S is a polynomial in the entries of the input, so it is continuous. The set $\{0\}$ is closed in $M(2n)$. By definition $SP(2n) = S^{-1}(\{0\})$.
$\therefore$ $SP(2n)$ is the preimage of a continuous function to a closed set, so it is closed ([3]). □

Then, by 6.18, these classical matrix groups are Lie subgroups and therefore also submanifolds of $GL(n)$ (or $GL(n, \mathbb{R})$ in the case of $SO(n, \mathbb{R})$.)

The following is a consequence of Peter-Weyl.

Theorem 7.5. *All compact Lie groups are matrix groups.*

As this theorem suggests, there are non-compact Lie groups that do not have faithful representations as matrices. The most common example of such a Lie group is the universal cover of $SL(2)$. Another example, slightly easier to show, is the quotient of the Heisenberg group H, defined as

$$M \in M(n, \mathbb{R}) = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

by the discrete normal subgroup N

$$N \in H \cap M(n, \mathbb{Z}) = \left\{ \begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

is not a matrix Lie group.

In any event, thank you for taking the time to read this, and I look forward to improving with your feedback! ♥□

References

[1] Brian C. Hall (2015), *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer Graduate Texts in Mathematics.

[2] Loring C. Tu (2011), *An Introduction to Manifold*, Springer Universitexts.

[3] B.Heuber *Show that the symplectic group* $\mathrm{Sp}(2n, \mathbb{C})$ *is a closed sub group of* $\mathrm{Gl}(2n, \mathbb{C})$, URL (version 2020–11–08): https://bit.ly/3Jj2O7n, Mathematics Stack Exchange.

[4] Freakish, *Showing SO(n) is a compact topological group for every n.*, URL (version 2023–05–18): https://bit.ly/4cTQ6cV, Mathematics Stack Exchange.

[5] Zuoqin Wang, *Lecture 11: Cartan's Closed Subgroup Theorem*, URL https://bit.ly/3PX0AhF.