REU Notes

Katalin Berlow

June 24, 2019

Contents

Index

1 Linear Algebra

1.1 Bilinear forms and Orthogonality

Here we will redefine some things that we've learned in Matrices and Linear Transformations, but in a more general way.

Definition. Let U, V, W be vector spaces. A **bilinear form** B is a function from $U \times V \to W$ such that when one input is fixed, the function is linear.

Definition. Let V, W be vector spaces. A **symmetric bilinear form** B on V is a bilinear form from $V \times V \to W$ in which $B(a, b) = B(b, a)$ for all $a, b \in V$.

An example of a symmetric bilinear form is the dot product.

Definition. Given a symmetric bilinear form B, two vectors u and v are **orthogonal** if $B(u, v) = 0$. A set is said to be orthogonal if any two matrices in that set are orthogonal.

Any set of basis vectors must be orthogonal. This is a more general version of being perpendicular.

Definition. Two vectors are **orthonormal** if they are orthogonal and unit vectors.

Definition. An $n \times n$ *complex* matrix A is said to be **unitary** if the column vectors of A are orthonormal. An $n \times n$ *real* matrix A is said to be **orthogonal** if the column vectors of A are orthonormal.

Proposition 1.1.1. *A matrix U is unitary if and only if $U^* = U^{-1}$.*

Proposition 1.1.2. *A matrix U is orthogonal if and only if $U^T = U^{-1}$.*

1.2 Eigenvalues and Eigenvectors

We will start with a basic review of eigenvectors.

Definition. If A is a matrix in $M_n(\mathbb{C})$, then a nonzero vector v is called an **eigenvector** for A if there exists some complex number λ such that

$$Av = \lambda v$$

.

Definition. Similarly, a complex number λ is called an **eigenvalue** if there exists a nonzero vector v such that

$$Av = \lambda v \quad \text{or equivalently} \quad (A - \lambda I)v = 0$$

This condition is equivalent to the following

$$\det(A - \lambda I) = 0$$

Definition. For any $A \in M_n(\mathbb{C})$, we define the **characteristic polynomial** p of A to be

$$p(\lambda) = \det(A\lambda I) \quad \text{for } \in \mathbb{C}$$

This is a polynomial of degree n.

The eigenvalues of a matrix are equal to the zeroes of its characteristic polynomial.

1.3 SN Decomposition

Not all matrices are diagonalizable. However, the following convenient fact is true.

Theorem 1.3.1 (SN Decomposition). *Let $A \in M_n(\mathbb{C})$. There exists a unique pair (S, N) of matrices also in $M_n(\mathbb{C})$ with the following properties:*

1. *S is diagonalizable*
2. *N is nilpotent*
3. *$SN = NS$*
4. *$A = S + N$*

1.4 Miscellaneous Linear Algebra Facts

Proposition 1.4.1. *The derivative commutes with a linear transformation.*

Definition. A bilinear form $[\cdot, \cdot]$ on a vector space V is **skew-symmetric** if for all $X, Y \in V$, we have $[X, Y] =$ -$[Y, X]$.

1.5 Vector Spaces

Definition. A **vector space** is a set of vectors V and a designated scalar field $\mathbb{F}$ along with two operations: vector addition and scalar multiplication, for which the following properties are true. Let $X, Y, Z \in V$ and $r, s \in \mathbb{F}$.

1. vector addition is associative
2. vector addition commutative

3. vector addition has an identity
4. vector addition has inverses
5. scalar multiplication is associative
6. scalar mulriplication has an identity
7. $(r+s)X = rX + sX$
8. $r(X+Y) = rX + rY$

Definition. Let V be a vector space equipped with a bilinear form. Let W be a subspace of V. The **orthogonal compliment** of W is the set of all vectors in V which are orthogonal to every vector in W.

Definition. A set of vectors B in a vector space V is called a **basis** if every element of V can be written uniquely as a linear combination of elements of B.

Equivalently, A set of vectors B in a vector space V is a basis if its elements are linearly independent and every element of V can be written as a linear combination of elements of B.

Definition. Let U and V be vector spaces. The **direct sum** of U and V, denoted $U \oplus V$, is a the set of (u, v) such that $u \in U$ and $v \in V$ with addition and scalar multiplication defined entrywise.

2 Analysis

2.1 Topology

Definition. Let X be a set and let τ be a subset of $\mathcal{P}(X)$. τ is a **topology** on X if:

i. $\varnothing, X \in \tau$

ii. Any union of elements of τ is an element of τ.

iii. Any *finite* intersection of elements of τ is an element of τ.

Definition. Let X be a topological space. For any sest $A \subseteq X$, an **open cover** $\mathcal{C}$ of A is any collection of open sets $U \subseteq X$ such that:

$$A \subseteq \bigcup_{U \in \mathcal{C}} U$$

If you have two open covers $\mathcal{C}$ and $\mathcal{C}'$ of the same set A, then $\mathcal{C}'$ is a **subcover** of $\mathcal{C}$ if $\mathcal{C}' \subseteq \mathcal{C}$.

Definition. The set A is called **compact** if every open cover of A has a finite subcover.

Theorem 2.1.1 (Heine-Borel)**.** *A set $A \subseteq \mathbb{R}^n$ (or $\mathbb{C}^n$) is compact if and only if it is bounded and closed.*

Thus in order to check if a matrix Lie group is compact, one only needs to check that it is bounded and closed.

Definition. A continuous bijection whose inverse is also continuous is called a **homeomorphism**. This is equivalent to an isomorphism between topological spaces.;

Definition. Two continuous functions are **homotopic** if one can be continuously deformed into the other. More formally, two functions are homotopic if there exists a **homotopy** between them. A homotopy is defined as follows. Let f and g be two continuous functions from topological spaces X to Y. A homotopy is a continuous function $H : X \times [0, 1] \to Y$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$ for all $x \in X$.
We can think of the second parameter as time and of H as a continuous deformation of f into g.

2.2 Continuity

In order to understand more properties about Lie groups we first need to know a little bit of Analysis.

Definition. Given an interval I and a function $f : A \to \mathbb{R}^{(n?)}$ f is **continuous** on I if it has the following property:

$$\forall \varepsilon > 0\ \exists \delta > 0 s : \forall x \in I\ \forall c \in \mathbb{R} : |x - c| < \delta \Rightarrow |f(x) - f(c)| < \varepsilon$$

In words this property can be restated as: Fir any positive amount ε, there is a small enough δ such that for two points x and c, if they are less than δ distance apart, then their images differ by less than ε.

Continuous functions have no jumps or breaks.

Definition. A **smooth** function is a function with derivatives of all orders everywhere in its domain.

2.3 Weierstrass M-test

Recall back to Calculus 2, there was a method of determining convergence of infinite series called the comparison test. The comparison test states that if you have two infinite series, $\sum a_n$ and $\sum b_n$ and $0 \leq a_n \leq b_n$, then if $\sum b_n$ converges, so must $\sum a_n$. The analogous test for complex numbers is called the Weierstrass M-test.

Theorem 2.3.1 (Weierstrass M-test)**.** *Suppose that $\{f_n\}$ is a sequence of complex-valued functions defined on a set A, and that there is a sequence of positive real numbers $\{M_n\}$ satisfying the following:*

$$\sum_{n=1}^{\infty} M_n < \infty$$

$$(\forall x \in A)\ |f_n(x)| \leq M_n$$

then $\displaystyle\sum_{n=1}^{\infty} f_n(x)$ *converges uniformly and absolutely.*

In other words, if the norm of a sequence of functions forms a convergent series, then the series of functions converge.

2.4 Convergence

Definition. A series $\sum a_n$ is **absolutely convergent** if $\sum |a_n|$ is convergent.

Absolute convergence implies convergence. This is equivalent to the condition that the series converges to the same thing regardless of the order in which it is summed up.

Definition. Let $\{f_n(x) : \text{-} \in \omega\}$ be a sequence of functions defined on the interval $\mathcal{I}$.We say that $f_n(x)$ **converges pointwise** to the function $f(x)$ on the interval $\mathcal{I}$ if

$$f_n(x) \to f(x) \text{ as } n \to \infty \text{ for each } x \in \mathcal{I}$$

The limit function f may fail to be continuous even if all f_n are continuous.

We next want to define a stronger version of convergence: uniform convergence. In order to do this, we need to define something called the *supremum norm*.

Definition. We define a type of norm called the **supremum norm** for a real valued function f on an interval $\mathcal{I}$ as follows:

$$||f||_{\mathcal{I}} = \sup_{x \in I} |f(x)|$$

This is the biggest value of $f(x)$ on the interval $\mathcal{I}$.

Note that the supremum norm always exists when f is continuous and $\mathcal{I}$ is closed and bounded (or equivalently, when $\mathcal{I}$ is compact, see Heine-Borel Theorem in section 1.3).

We can now define uniform convergence.

Definition. A sequence of functions $f_n(x)$ defined on $\mathcal{I}$ is said to **converge uniformly** to $f(x)$ on $\mathcal{I}$ if

$$||f_n - f||_{\mathcal{I}} \to 0 \text{ as } n \to \infty$$

We write this as

$$f_n \to f \text{ uniformly on } \mathcal{I} \text{ as } n \to \infty$$

Uniform convergence implies pointwise convergence.

2.5 Power Series

Definition. A power series on the real line $\mathbb{R}$ is a formal expression

$$\sum_{n=0}^{\infty} a_n (x - \alpha)^n$$

Where a_n are real constants and $\alpha \in \mathbb{R}$.

Proposition 2.5.1. *Assume the power series*

$$\sum_{n=0}^{\infty} a_n(x-\alpha)^n$$

converges at $x = c$. Let $r = |c-\alpha|$. then the series converges uniformly and absolutely on compact subsets of $\mathcal{I} = \{x : |x-\alpha| < r\}$.

Note that because of the Heine-Borel Theorem, the *compact* subsets of $\mathcal{I}$ are just the closed and bounded subsets of $\mathcal{I}$.

Proof. Let K be a compact subset of $\mathcal{I}$ such that $K = [\alpha - s, \alpha + s]$ for some $0 < s < r$. In other words, K is an interval of radius s around α which is contained in I. The following holds:

$$\sum_{n=0}^{\infty} |a_n(x-\alpha)^n| = \sum_{n=0}^{\infty} |a_n(c-\alpha)^n| \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^n$$

The first term of the right hand side, $\sum_n |a_n(c-\alpha)^n|$ converges by the condition of the proposition. Thus the terms $|a_n(c-\alpha)^n|$ must be bounded by some constant C. Thus the following is true:

$$\sum_{n=0}^{\infty} |a_n(c-\alpha)^n| \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^n \le \sum_{n=0}^{\infty} C \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^n$$

The second term in this sum, $\left|\frac{x-\alpha}{c-\alpha}\right|^n$ is bounded by $\frac{s}{r}$ since $|x-\alpha| < s$ and $|c-\alpha| = r$. Since $s < r$, $\frac{s}{r} < 1$. Thus $\left|\frac{x-\alpha}{c-\alpha}\right|^n < 1$. Therefore $\sum C \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^n$ converges. By the comparison test, $\sum |a_n(c-\alpha)^n| \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^n$ must converge as well. Thus $\sum_{n=0}^{\infty} |a_n(x-\alpha)^n|$ converges. Therefore, by the Weierstrass-M Test, $\sum_{n=0}^{\infty} a_n(x-\alpha)^n$ converges absolutely and uniformly on K. ■

A consequence of this proposition is that the set in which $\sum |a_n(x-\alpha)^n|$ converges is an interval around α. This interval is called the **interval of convergence**. The **radius of convergence** is half the length of this interval.

The series doesn't necessarily converge at the endpoints of this interval. Thus we will use $\mathcal{C}$ to denote the *open* interval of convergence.

If the series converges at an endpoint of $\mathcal{C}$, then the convergence is uniform up to that point.

Definition. A function with domain an open set $U \subseteq \mathbb{R}$ and range in the real more complex numbers is said to be **real analytic** or **holomorphic** at α if the function may be represented by a power series:

$$\sum_{n=0}^{\infty} a_n(x-\alpha)^n$$

which converges on an interval around α.

The function is said to be real analytic on a set V if it is real analytic at every point in V.

Note that functions which are defined by a power series are real analytic on their open interval of convergence.

2.6 Properties of Complex Numbers

The following are some miscellaneous properties about complex numbers.

Proposition 2.6.1. $e^n = 1$ *if and only if* n *is an integer multiple of* $2i\pi$*.*

2.7 The Complex Logarithm

Recall how to express complex numbers in exponential form:

$$a + bi = \sqrt{a^2 + b^2}e^{i\arctan(\frac{b}{a})} = re^{i\theta}$$

Any complex number can be expressed in the form e^z for some $z \in \mathbb{C}$. However, e^z is $2\pi i$-periodic, meaning that $e^z = e^{z+2\pi i}$. Thus, since this function is not injective, its inverse is not a well defined function. However, we can say that z is a **logarithm** of w if $w = e^z$. This logarithm function doesn't produce a unique output, so there can be multiple logarithms of the same number. Since e^z is always nonzero, only nonzero numbers can have a logarithm.

The complex logarithm can be defined as a power series.

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

But where did this come from? Recall the following from calculus:

$$\frac{d}{dx}log(1-x) = \text{-}\frac{1}{1-x} = \text{-}\sum_{n=0}^{\infty} x^n$$

If you integrate it term-wise, you get

$$\log(1-x) = \int \text{-}\sum_{n=0}^{\infty} x^n dx = \text{-}\sum_{n=1}^{\infty}\int x^n dx = \text{-}\sum_{n=0}^{\infty}\frac{x^{n+1}}{n+1} = \text{-}\sum_{n=1}^{\infty}\frac{x^n}{n}$$

Now, if you substitute $z = 1 - x$. you get

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

Lemma 2.7.1. *The function* $\log z$ *is defined and analytic and is the inverse of* e^z *in a circle of radius 1 about* $z = 1$*.*

Proof. Recall proposition 2.5.1. When $z = 1$, this function clearly converges because all the terms are 0. In this case, $\alpha = 1$ and thus the radius of convergence is 1. Therefore $\log z$ converges uniformly and absolutely in a circle of radius 1 about $z = 1$. Since it converges, and clearly can be represented by a power series, $\log z$ is defined and analytic on its interval of convergence.

Since this log coincides with the logarithm on the natural numbers, it is the inverse of e^x. ∎

3 Matrix Lie Groups

Definition. The **general linear group** denoted $GL(n : \mathbb{F})$ is the set of all invertable $n \times n$ matrices with entries from a scalar field $\mathbb{F}$.

$$GL(n, \mathbb{F}) = \{m \in M_n(\mathbb{F}) : det(m) \neq 0\}$$

Definition. Let A_m be a sequence of complex matrices in $M_n(\mathbb{C})$. We say that A_m **converges** to a matrix A if each of its entries converge to the corresponding entry in A.

Definition. A subgroup G of $GL(n; \mathbb{C})$ is a **matrix Lie group** if it has the following property: If A_m is a sequence of matrices in G which converge to A then either $A \in G$ or $A \notin GL(n; \mathbb{C})$.

This is equivalent to saying that G is a closed subgroup of $GL(n : \mathbb{C})$.

3.1 Examples of Matrix Lie Groups

The special and general linear groups, along with the orthogonal, unitary, and symplectic groups make up the classical groups, which are each defined as follows.

Orthogonal Group $O(n)$

Definition. The set of all orthogonal real matrices forms a closed subgroup of $GL(n; \mathbb{C})$ and is called the **orthogonal group** denoted $O(n)$.

Definition. The set of all orthogonal real matrices with determinant equal to 1 also forms a closed subgroup of $GL(n; \mathbb{C})$ and is called the **special orthogonal group** denoted $SO(n)$.

The elements of $SO(n)$ act as rotations in real space.

$O(n)$ and $SO(n)$ are both matrix Lie groups.

Generalized Orthogonal Group

Let n and k be positive integers. Define a symmetric bilinear form $[\cdot, \cdot]_{n,k}$ on $\mathbb{R}^{n+k}$ as follows:

$$[x, y]_{n,k} = x_1y_1 + ... + x_ny_n - x_{n+1}y_{n+1} - ... - x_{n+k}y_{n+k}$$

Definition. The set of all real $(n + k) \times (n + k)$ matrices which preserve the symmetric bilinear form defined above is called the **generalized orthogonal group** $O(n; k)$.

The generalized orthogonal group $O(n; k)$ is a subgroup of $GL(n + k, \mathbb{R})$ and a matrix Lie group.

Unitary Groups $U(n)$

Definition. The set of unitary matrices forms a subgroup of $GL(n;\mathbb{C})$ and is called the **unitary group** denoted $U(n)$.

The unitary group is closed and is thus a matrix Lie Group.

Definition. The set of unitary matrices with determinant equal to 1 also forms a subgroup of $GL(n;\mathbb{C})$ and is called the **special unitary group** denoted $SU(n)$.

The special unitary group is also a matrix Lie group. As you will soon notice, most groups have a "special" version which is just the subgroup of elements with determinant equal to 1. This will be denoted by adding an "S" in front of whatever the usual notation is for that group.

The Symplectic Group $Sp(n;\mathbb{C})$

Consider the skew-symmetric bilinear form B on C^{2n} defined as follows:

$$B[x,y] = \sum_{k=1}^{n} x_k y_{n+k} - x_{n+k} y_k$$

This is equivalent to $\langle x, Jy\rangle$ for $J = \begin{pmatrix} 0 & I \\ \text{-}I & 0 \end{pmatrix}$.

Definition. The set of all $2n \times 2n$ matrices which preserve B is called the **symplectic group**.

The symplectic group is a subgroup of $GL(2n;\mathbb{C})$ and is a matrix Lie group.

Definition. There is also the **compact symplectic group** defined as

$$Sp(n) = Sp(n,\mathbb{C}) \cap U(2n)$$

This is the group of all unitary matrices which preserve the bilinear form B.

The Heisenberg Group H

Definition. Consider the set of all matrices of the form

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{for } a,b,c \in \mathbb{R}$$

The set of all such matrices is called the **Heisenberg Group** which is denoted H.

H is a subgroup of $GL(3;\mathbb{R})$.

3.2 Topological Properties of Matrix Lie Groups

Compactness

A matrix Lie group $G \subseteq GL(n;\mathbb{C})$ is **compact** if it is compact in the usual topological sense as a subset of $M_n(\mathbb{C})$.

Because of the Heine-Borel theorem above, a matrix Lie group G is compact if and only if it is closed (as a subset of $M_n(\mathbb{C})$) and bounded. Recall that G is closed (as a subset of $M_n(\mathbb{C})$) if every convergent sequence of matrices converges to an element of G. G is bounded if there exists a constant C which is greater than the absolute value of every entry of every matrix in G.

Connectedness

Definition. A matrix Lie group G is **connected** if for every two matrices $A, B \in G$ there exists a continuous path between them. More precisely, there exists a continuous function $P(t)$ for $a \leq t \leq b$ such that $P(a) = A$ and $P(b) = B$.

Definition. For any matrix Lie group G, its **identity component**, denoted G_0, is the set of $A \in G$ such that there is a continuous path from A to the indentity.

The notion of *connected* for matrix Lie groups is what is called *path connected* in topology, which in topology is different from *connected*.

Theorem 3.2.1. *If G is a matrix Lie group, the identity component G_0 is a normal subgroup of G.*

Normal Subgroups Interlude

Definition. Let G be a group and $N \subseteq G$. The element $g \in G$ is said to **normalize** N if:

$$gNg^{-1} = N$$

.

Definition. Let G be a group and let N be a subgroup of G. N is a **normal subgroup** of G if any element of G normalizes N. That is

$$N \trianglelefteq G \Leftrightarrow \ (\forall g \in G)\ gNg^{-1} = N$$

Back to Connectedness

Proposition 3.2.1. *The groups $GL(n:\mathbb{C})$ and $SL(n:\mathbb{C})$ are connected for all $n \geq 1$.*

Proof. The proof of this makes use of the following fact: Every matrix is similar to an upper triangular matrix. Recall that matrices A and B are similar if there exists a matrix C such that $A = CBC^{\text{-}1}$.

Let $A = CBC^{\text{-}1} \in GL(n;\mathbb{C})$ where B is an upper triangular matrix. Reduce the upper part to 0 by multiplying it by $(1-t)$ as t goes to 1. Thus we have a path from A to $CDC^{\text{-}1}$ where D is a diagonal matrix with nonzero entries. Thus we can continuously make each of them closer to 1 until we have $CIC^{\text{-}1} = I$. Thus we have defined a continuous path from A to the identity. Since this can be done for any matrix A, $GL(n:\mathbb{C})$ is connected. ■

Simple Connectedness

Definition. A matrix Lie group G is **simply connected** if it is connected and every loop can be shrunk to a point in G.

Basically there are no holes. More precisely, let G be a connected matrix Lie group. G is simply connected if for every path (loop) A(t) such that $0 \leq t \leq 1$, $A(t) \in G$ for all t, and $A(0) = A(1)$, there exists a continuous function $A(s,t)$ with the following properties:

- $0 \leq s \leq 1$, $0 \leq t \leq 1$
- For all s, $A(s,0) = A(s,1)$.
- $A(0,t) = A(t)$
- For all t, $A(1,t) = A(1,0)$

$A(t)$ is a loop. $A(s,t)$ is a family of loops which shrink to a single point as s reaches 0.

3.3 Real Projective Spaces

Definition. The **real projective space** of dimension n denoted $\mathbb{R}P^n$ is the set of lines through the origin in $\mathbb{R}^{n+1}$.

Since each line may be identified by the two by the two antipodal points where it intersects the unit sphere S^n, we can consider $\mathbb{R}P^n$ to just be this collection of points.

There is a natural map $\pi : S^n \to \mathbb{R}P^n$ sending $u \in S^n$ to $\{u, \text{-}u\}$.

Define a distance function on $\mathbb{R}P^n$ as follows:

$$d(\{u, \text{-}u\}, \{v, \text{-}v\}) = \min\{d(u, v), d(u, \text{-}v)\}$$

This distance is the distance between the closest places that the two lines intersect the circle.

Proposition 3.3.1. *$\mathbb{R}P^n$ is not simply connected.*

Proof. Let u be any unit vector in $\mathbb{R}^{n+1}$, and $B(t)$ be any path in S^n connecting u to $\text{-}u$ Consider $A(t) := \pi(B(t))$, where π is the map defined above. Note that the following holds:

$$A(0) = \pi(B(0)) = \pi(u) = \pi(\text{-}u) = \pi(B(1)) = A(1)$$

Thus $A(t)$ is a loop. We want to show that $A(t)$ cannot be shrunk continuously to a point.
Suppose for contradiction that $A(t)$ could be shrunk continuously to a point. Then, by the definition, there would exist a continuous map $A(s, t)$ which fits the requirements in the definition of simple connectedness. In particular, $A(s, 0) = A(s, 1)$ for all s. Consider $B(s, t)$ such that $A(s, t) = \pi(B(s, t))$. Since $A(s, 0) = A(s, 1)$ for all s, $B(s, 0) = \pm B(s, 1)$ for all s. However, since $B(t)$ is a path from u to $\text{-}u$, we know that $B(0, 0) = \text{-}B(0, 1)$. Then in order for $B(s, t)$ to be continuous, we must have $B(s, 0) = \text{-}B(s, 1)$ for all s. It follows that $B(1, t)$ is not a single point, but instead a non constant path. Thus $A(1, t)$ must also be not a single point. This contradicts our assumption that $A(s, t)$ shrinks to a point when s approaches 1. Thus $\mathbb{R}P^n$ is not simply connected. ■

3.4 Topology of $SO(3)$

Proposition 3.4.1. *There is a continuous bijection between $SO(3)$ and $\mathbb{R}P^3$.*

Proof. Let B^3 be the closed ball of radius π in 3-dimensional space. We will first construct an map Φ from B^3 to $SO(3)$.
Note that the elements of $SO(3)$ can each be represented as follows: For any unit vector v and angle θ, let $R_{v,\theta}$ denote the rotation about the plane orthogonal to v by angle θ. Each element of $SO(3)$, other than the identity and 90 degree rotations, can be represented uniquely as $R_{v,\theta}$ for some $0 \leq \theta \leq \pi$. The identity would be $R_{v,0}$ for any vector v. For rotations about an angle of π, the following is true: $R_{v,\pi} = R_{\text{-}v,\pi}$ and thus those angles can be represented in two ways.
Define $\Phi : B^3 \to SO(3)$ as follows:

$$\Phi(u) = \begin{cases} R_{\hat{u},||u||} & \text{if } u \neq 0 \\ I & \text{if } u = 0 \end{cases}$$

This map can be thought of as mapping any vector in B^3 to the rotation orthogonal to the vector, by an angle equal to the length of the vector. Note that this map is continuous since as $||u||$ reaches 0, the transformation approaches the identity.

This map is injective except for at the boundary, since at the boundary, $||u|| = \pi$ and a rotation by π clockwise is the same as a rotation by π counterclockwise. Thus, antipodal points map to the same transformation.

This map is, however, surjective. Any element of $SO(3)$ can be represented by a rotation $R_{u,\theta}$, and every $R_{u,\theta}$ where $\theta \neq 0$ is mapped to by Φ. Since $R_{u,0} = I$ for all u, and I is also mapped to, every element of $SO(3)$ gets mapped to. ∎

3.5 Homomorphisms

Recall that in group theory a homomorphism is just a map between groups that preserves the operation. This is an $\mathcal{L}$-embedding, but specific to the language of groups.

Definition. Let G and H be matrix Lie groups. A map $\Phi : G \to H$ is a **Lie group homomorphism** if it is a continuous group homomorphism.

Definition. If Φ is a bijective Lie group homomorphism whose inverse is continuous, then Φ is a **Lie group isomorphism**.

Note: The condition that Φ be continuous is a technicality, since homomorphisms between Lie groups tend to be continuous anyway.

4 The Matrix Exponential

4.1 The Exponential of a Matrix

Recall that for a natural number $x \in \mathbb{R}$, the following identity holds:

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

We use this identity to define exponentiation on matrices.

Definition. If X is an $n \times n$ matrix, we define the **exponential** of X, denoted e^x or $\exp X$ by:

$$\boxed{e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}}$$

Proposition 4.1.1. *The series above converges for all $X \in M_n(\mathbb{C})$.*

For this proof we will first need to define the norm of a matrix.
Recall that for a complex number $z = a+bi$, the **norm** of z, denoted $|z|$ is equal to $\sqrt{a^2+b^2}$. For elements x of $\mathbb{C}^n$, the norm is defined as

$$|x| = \sqrt{|x_1|^2 + |x_2|^2 + ... + |x_n|^2} = \sqrt{\sum_{i \leq n} |x_i|^2}$$

The norm of an element of $M_n(\mathbb{C})$ is defined similarly, where an element of $M_n(\mathbb{C})$ is treated as if it were an element of $\mathbb{C}^{n^2}$

Definition. F For $X \in M_n(\mathbb{C})$, define the **Hilbert-Schmidt norm** as follows:

$$||X|| = \sqrt{\sum_{j,k \leq n} |X_{j,k}|^2}$$

This is the square root of the sum of squares of the norms of each entry of the matrix.
The Hilbert-Schmidt norm can also be calculated as follows:

$$||X|| = \sqrt{\text{trace}(X * X)}$$

This norm satisfies the following inequalities:

$$||X+Y|| \leq ||X|| + ||Y|| \qquad \textit{Triangle Inequality}$$
$$||XY|| \leq ||X||\ ||Y|| \qquad \textit{Cauchy-Schwarz Inequality}$$

An important point to note is that if X_m is a sequence of matrices, then X_m converges to X if and only if $|X_m - X|$ approaches 0 as $m \to \infty$.
And now we can do the proof of the proposition.

Proof. By Cauchy-Schwarz:

$$||X^m|| \leq ||X||^m$$

for $m \geq 1$, and thus it follows that

$$\sum_{m=0}^{\infty} ||\frac{X^m}{m!}|| = ||I|| + \sum_{m=1}^{\infty} \frac{||X^m||}{m!} < \infty$$

Since the norm of this sequence converges, the sequence must converge absolutely. ■

Proposition 4.1.2. *e^X is a continuous function of X.*

Proof. Each of the partial sums of $e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}$ are continuous, since X^n is a continuous function of X. Therefore, e^X is continuous.
Next we want to show that the series converges. Let $S_R = \{X \in M_n(\mathbb{C}) : ||X|| < R\}$ for each $R \in \mathbb{R}$. By the Weierstrass M-test, e^X converges on the set S_R for each $R \in \mathbb{R}$. Since $\bigcup_{R\in\mathbb{R}} S_R = M_n(\mathbb{C})$, we have that e^X converges on the set of all matrices.
Since e^X converges on the set of all matrices and it's partial sums are continuous, it is continuous on $M_n(\mathbb{C})$.

■

The following are some **properties** about the matrix exponential:

1. $e^0 = I$
2. $(e^X)^* = e^{X^*}$
3. $(e^X)^{-1} = e^{-x}$
4. $e^{(\alpha+\beta)X} = e^{\alpha X}e^{\beta X}$
5. $XY = YX \Rightarrow e^{X+Y} = e^X e^Y = e^y e^X$
6. $e^{CXC^{-1}} = Ce^X C^{-1}$

Proposition 4.1.3. *Let X be an $n \times n$ complex matrix. Then e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

$$\frac{d}{dt}e^{tX} = Xe^{tX} = e^{tX}X$$

In particular,

$$\frac{d}{dt}e^{tX}\bigg|_{t=0} = X$$

4.2 Computing the Exponential

There are three categories a matrix $X \in M_n(\mathbb{C})$ could fall into. The matrix X may be diagonizable, nilpotent, or neither. Thus we will case it as such.

Case 1: X is diagonizable

Recall that if X is diagonizable, it can be put into the form:

$$X = C \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} C^{-1}$$

Where $\lambda_1, ..., \lambda_n$ are the eigenvalues of X. Thus, by property 6 in the previous section,

$$e^X = e^{CDC^{-1}} = Ce^D C^{-1} = C \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix} C^{-1}$$

Thus, computing the matrix exponential of a diagonizable matrix is as simple as diagonalizing it, then exponentiating the eigenvalues.

Case 2: X is nilpotent

Definition. An $n \times n$ matrix X is **nilpotent** if there exists a positive integer n such that $X^n = 0$.

Note that if $X^n = 0$, then for all $m \geq n$, we have $X^m = 0$ as well. Thus, in this case, e^X is a finite series and can be computed explicitly.

Case 3: X is neither diagonizable nor nilpotent

Let X be an arbitrary matrix. By the SN Decomposition Theorem, there exist $S, N \in M_n(\mathbb{C})$ such that S is diagonalizable, N is nilpotent, $SN = NS$, and $X = S + N$. Thus since S and N commute, the following holds

$$e^X = e^{S+N} = e^S e^N$$

The exponentials e^S and e^N are both computable using the teqniques in cases 1 and 2. Thus, we have covered all types of matrices.

4.3 The Matrix Logarithm

Definition. For an $n \times n$ matrix A define the **matrix logarithm** $\log A$ as follows:

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

whenever the series converges.

Lemma 4.3.1. *The series* $\log A$ *converges when* $||A - I|| < 1$.

Proof. Since the complex-valued function has radius of convergence 1, and $||(A - I^n|| \leq ||A - I||^n$, the function converges when $||A - I|| < 1$ ■

Theorem 4.3.1. *The function*

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

is defined and continuous on the set of all $n \times n$ *matrices* A *such that* $||A - I|| < 1$.

Additionally, for all such A*, the following holds:*

$$e^{\log A} = A$$

And for X *such that* $||X|| < \log 2$*:*

$$||e^X - I|| < 1 \text{ and } \log e^X = X$$

Proof. The first claim is true by the lemma and the proposition 2.5.1.
The proof of the second two claims are too complicated so I won't go into those. ■

Proposition 4.3.1. *There exists a constant* c *such that for all* $n \times n$ *matrices* B *such that* $||B|| < \frac{1}{2}$*, we have*

$$||\log(B + I) - B|| \leq c||B||^2$$

Proof. Note that

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

$$= \sum_{n=1}^{\infty} (\text{-}1)^{n+1} \frac{B^n}{n} - (\text{-}1)^{1+1} \frac{B^1}{1} = \sum_{n=2}^{\infty} (\text{-}1)^{n+1} \frac{(B)^n}{n} = B^2 \sum_{n=1}^{\infty} (\text{-}1)^{n+1} \frac{(B)^{n-2}}{n}$$

Thus we have

$$||\log(I - B) - B|| = ||B||^2 \sum_{n=1}^{\infty} (\text{-}1)^{n+1} \frac{||B||^{n-2}}{n} \leq ||B||^2 \sum_{n=1}^{\infty} (\text{-}1)^{n+1} \frac{(\frac{1}{2})^{n-2}}{n}$$

It can be easily checked that the last sum converges. Thus the statement holds. ■

The reason this proposition is important is because of the following more concise version:

$$\log(I + B) = B + O(||B||^2)$$

where $O(||B||^2)$ is a term that grows by a constant times $||B||^2$ for all sufficiently small values of B. This will be useful in the next section when we discuss further properties of the exponential.

4.4 Further Properties of the Exponential

In this section we will give some miscellaneous important facts that will hopefully make it more clear why we care about the matrix exponential.

Theorem 4.4.1 (Lie Product Formula)**.** *Let X and Y be $n \times n$ complex matrices.*

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

Proof. The proof is a whole bunch of gross arithmetic and relies on proposition 4.3.1. It is left as an exercise to the reader to read page 36 of the Lie Groups textbook. ■

Theorem 4.4.2. *For any $X \in M_n(\mathbb{C})$, we have*

$$\det(e^X) = e^{traceX}$$

Definition. A function $A : \mathbb{R} \to GL(n;\mathbb{C})$ is called a **one-parameter subgroup** of $GL(n, \mathbb{C})$ if

1. A is continuous
2. A(0) = I
3. A(t+s) = A(t)A(s)

Theorem 4.4.3. *If A is a one-parameter subgroup of $GL(n;\mathbb{C})$, then there exists a unique $n \times n$ complex matrix X such that*

$$A(t) = e^{tX}$$

5 Lie Algebras

5.1 Lie Algebras of Matrix Groups

Lie algebras are super important for studying Lie groups. They're another type of simpler structure and there is one associated with each Lie group. Lie algebras are defined as follows:

Definition. Let G be a matrix Lie group. The **Lie algebra** of G denoted $\mathfrak{g}$ is the set of all matrices X such that $e^{tX} \in G$ for all $t \in \mathbb{R}$.

Equivalently, $X \in \mathfrak{g}$ if the one parameter subgroup generated by X lies in G.

General Linear Groups

Proposition 5.1.1. *The Lie algebra of $GL(n;\mathbb{C})$ denoted $\text{gl}(n;\mathbb{C})$ is the set of all $n \times n$ complex matrices. Additionally, the Lie algebra of $GL(n;\mathbb{R})$ denoted $\text{gl}(n;\mathbb{R})$ is the set of all $n \times n$ real matrices.*

Proof. If X is any $n \times n$ complex matrix, then by property 3, e^X is invertable. Any $n \times n$ complex matrix is in $\text{gl}(n;\mathbb{C})$.

If X is any $n \times n$ real matrix, then e^{tX} is invertable and real. Also, if e^{tX} is real for all $t \in \mathbb{R}$ then X is real, since $X = \frac{d}{dt} e^{tX} \mid_{t=0}$. Thus the Lie algebra of $GL(n;\mathbb{R})$ is $M_n(\mathbb{R})$ ∎

Proposition 5.1.2. *IF G is a subgroup of $GL(n,\mathbb{R})$, then the Lie algebra of G must consist of only real numbers.*

This proposition is proven in the previous discussion.

The Special Linear Groups

Proposition 5.1.3. *The Lie algebra of $SL(n;\mathbb{C})$, denoted $\text{sl}(n,\mathbb{C})$ is the set of $n \times n$ complex matrices with determinant 0. Similarly, the lie algebra of $SL(n,\mathbb{R})$, denoted $\text{sl}(n,\mathbb{R})$ is the set of real matrices with determinant 0.*

Proof. Recall theorem 4.4.2: $\det(e^X) = e^{\text{trace}(X)}$. Let $X \in M_n(\mathbb{C})$. If we know that e^{tX} is in $SL(n;\mathbb{C})$ for all $t \in \mathbb{R}$, this implies $\det(e^{tX}) = 1$. Thus $e^{t\text{trace}(X)} = 1$. Thus, by proposition 2.6.1, $t\text{trace}(X)$ must be an integer multiple of $2i\pi$ for all t. This is only possible if $\text{trace}(X) = 0$. Thus any element of $\text{sl}(n,\mathbb{C})$ must have trace equal to 0.

In the other direction, let X be an $n \times n$ complex matrix with trace 0. Then $\det(e^{tX}) = e^{t\text{trace}(X)} = e^0 = 1$ for all $t \in \mathbb{R}$. Thus for any X with trace 0, e^{tX} is in $SL(n,\mathbb{C})$. Thus $X \in \text{sl}(n,\mathbb{C})$. Thus $\text{sl}(n,\mathbb{C})$ is equal to the set of $n \times n$ matrices with trace 0. ∎

The Unitary Groups

Proposition 5.1.4. *The Lie algebra of $U(n)$, denoted* $\mathrm{u}(n)$*, is the space of all $n \times n$ complex matrices X such that $X^* = -X$.*

Proof. Recall proposition 1.1.1: U is unitary if and only if $U^* = U^{\text{-}1}$. Thus, $e^{tX} \in U(n)$, if and only if $(e^{tX})^* = (e^{tX})^{\text{-}1}$ which is equivalent to the condition $e^{tX^*} = e^{t(\text{-}X)}$. If $X^* = \text{-}X$, this is clearly true and thus, all such X are in $\mathrm{u}(n)$.

In the other direction, if $e^{tX^*} = e^{\text{-}tX}$, then $\frac{d}{dt}|_{t=0} e^{tX^*} = \frac{d}{dt}|_{t=0} e^{\text{-}tX}$, and thus $tX^* = t(\text{-}X)$ for all t, which is only true when $X^* = \text{-}X$. ■

The Orthogonal Group

Proposition 5.1.5. *The Lie algebra of $O(n)$, denoted* $\mathrm{so}(n)$*, is the set of all $n \times n$ real matrices X such that $X^T = \text{-}X$.*

Proof. Recall that by proposition 1.1.2, an $n \times n$ real matrix R is orthogonal if and only if $R^T = R^{\text{-}1}$. Thus, a matric X is in $\mathrm{so}(n)$ if and only if $(e^{tX})^T = (e^{tC})^{\text{-}1}$ for all t. This is equivalent to the condition $e^{tX^T} = e^{\text{-}tX}$. As shown in the proof of proposition 5.1.4, this is the case if and only if $X^T = \text{-}X$. ■

Also note that the trace of any element of $\mathrm{so}(n)$ must be 0, because since $X^T = \text{-}X$, the entries along the diagonal must all be 0.

Proposition 5.1.6. *The Lie algebras of $O(n)$ and $SO(n)$ are the same.*

Proof. By proposition 5.2.1, the exponential any matrix which is in the Lie Algebra is automatically in the identity component of the Lie group. Since $SO(n)$ is the identity component of $O(n)$, the Lie algebras must be the same.

Alternatively, since the trace of any matrix in $\mathrm{so}(n)$ is 0, the determinant of e^{tX} for any $X \in \mathrm{so}(n)$ must be 1. Thus, for any element of the lie Algebra, its corresponding Lie group element must have determinant 1. ■

Everything just stated also holds for $O(n, \mathbb{C})$ and $SO(n, C)$. Note that $\mathrm{so}(n, C) \neq \mathrm{su}(n)$.

The Generalized Orthogonal Group

First note that proposition 5.1.6 applies to the generalised orthogonal group as well, and thus the Lie algebras of $O(n; k)$ and $SO(n; k)$ are the same.

Recall from exercise 1, that a matrix A is in $O(n; k)$ if and only if

$$gA^T a = A^{\text{-}1} \quad \text{for} \quad g = \begin{pmatrix} I_n & 0 \\ 0 & \text{-}I_k \end{pmatrix}$$

.

Proposition 5.1.7. *The Lie algebra of $O(n;k)$ (and $SO(n;k)$), denoted* $\mathrm{so}(n;k)$ *is the set of all $n \times n$ real matrices X such that $gX^Tg = \text{-}X$.*

Proof. First note that the condition $g(e^{tX})^Tg = (e^{tX})^{\text{-}1}$ is equivalent to $e^{tgX^Tg} = e^{t(\text{-}X)}$. If this condition holds for X, then differentiating both sides gives is $gX^Tg = \text{-}X$. If $gX^Tg = \text{-}X$ then clearly the condition holds. ■

The Heisenberg Group

Proposition 5.1.8. *The Lie Algebra of the Heisenberg group is the set of all $n \times n$ strictly upper triangular matrices.*

Proof. First note that the following holds. If

$$X = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \quad \text{then} \quad e^X = \begin{pmatrix} 1 & a & b + \frac{ac}{2} \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

Thus, if X is strictly upper triangular, then X is in the Lie Algebra of H.

In the other direction, let X be an arbitrary matrix in the Lie algebra of H. Then, for all t, the entries on or below the diagonal must be 0 or 1. Thus, by taking the derivative, $\frac{d}{dt}|_{t=0}e^{tX} = X$ must be strictly upper triangular. ■

5.2 Properties of the Lie Algebra

Proposition 5.2.1. *Let G be a matrix Lie group and let X be an element if its Lie algebra. Then* e^X *is in the identity component of G.*

Proof. Note that when $t = 0$, then $e^{tX} = I$, and when $t = 1$ then $e^{tX} = e^X$. Since for all t, e^{tx} is continuous and in G, this defines a path from e^X to the identity component. ■

Proposition 5.2.2. *Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. Let $X \in \mathfrak{g}$ anf $A \in G$. Then $AXA^{\text{-}1} \in \mathfrak{g}$.*

Proof. Since $X \in \mathfrak{g}$, $e^{tX} \in G$ for all t. Since G is closed, $Ae^{tX}A^{\text{-}1} \in G$ for all t. Thus $e^{tAXA^{\text{-}1}} \in G$ for all t. Thus $AXA^{\text{-}1} \in \mathfrak{g}$. ■

Theorem 5.2.1. *Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra. Let $X, Y \in \mathfrak{g}$. Then:*

1. *$sX \in \mathfrak{g}$ for any $s \in \mathbb{R}$*
2. *$X + Y \in \mathfrak{g}$*

3. $XY - YX \in \mathfrak{g}$

Properties 1 and 2 show that $\mathfrak{g}$ is a real vector space, by showing that it is closed under vector addition and scalar multiplication. Property 3 shows that a Lie algebra defined in relation to a Lie group is in fact a Lie algebra in the abstract sense described in section 5.4.

Proof. Let G be a matrix Lie group and $\mathfrak{g}$ its Lie algebra.

Proof of 1. Let $s \in \mathbb{R}$ and $X \in \mathfrak{g}$. Then $e^{tX} \in G$ for all $t \in \mathbb{R}$. Thus $e^{s(\frac{t}{s})X} \in G$ for all $\frac{t}{s} \in \mathbb{R}$. Thus $tX \in \mathfrak{g}$.

proof of 2. Let $X, Y \in \mathfrak{g}$. Then $e^{t(X+Y)} \in G$. Recall Theorem 4.4.1:

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

Thus we simply need to show that $\lim \left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m$ is in G. Since $\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m$ is in G for all m, this defines a convergent sequence in G which converges to e^{X+Y}, which is invertable.

proof of 3.
Recall that $\frac{d}{dt} e^{tX}|_{t=0} = X$. Thus, by the product rule,

$$\frac{d}{dt}(e^{tX} Y e^{-tX})\bigg|_{t=0} = (XY)e^0 + (e^0 Y)(-X) = XY - YX$$

By proposition 5.2.2, $e^{tX} Y e^{-tX}$ is in $\mathfrak{g}$ for all $t \in \mathbb{R}$. Since $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$, it is topologically closed, meaning convergent sequences in $\mathfrak{g}$ must converge to something in $\mathfrak{g}$ Thus

$$\lim_{h\to 0} \frac{e^{hX} Y e^{-hX} - Y}{h} = \frac{d}{dt}(e^{tX} Y e^{-tX})\bigg|_{t=0} = XY - YX$$

must be in $\mathfrak{g}$. ∎

Definition. A subspace of $M_n(\mathbb{C})$ is called a **complex subspace** if its scalar field is the complex numbers.

Note that having complex entries is not sufficient.

Definition. A matrix Lie group is said to be **complex** if its Lie algebra is a complex subspace of $M_n(\mathbb{C})$.

5.3 Lie Algebra Homomorphisms

Definition. Given two matrices A and B, define the **bracket** (or commutator) as follows:

$$[A, B] = AB - BA$$

By theorem 5.2.1 point 3, the Lie algebra of a matrix Lie group will be closed under brackets.

Proposition 5.3.1. *For all matrices X, Y,*

$$[X,Y] = \frac{d}{dt}e^{tX}Ye^{-tX}|_{t=0}$$

Theorem 5.3.1. *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then there exists a unique linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\Phi(e^X) = e^{\phi(X)}$$

for all $X \in \mathfrak{g}$. This map has the following properties:

1. $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A^{-1})$
2. $\phi([X,Y]) = [\phi(X), \phi(Y)]$
3. $\phi(X) = \frac{d}{dt}\Phi(e^{tX})|_{t=0}$

Proof. Since Φ is a continuous group homomorphism, for a fixed X, $\Phi(e^{tX})$ will be continuous and will map 0 to the identity. Thus $\Phi(e^{tX})$ is a one parameter subgroup of H for any fixed X. Theorem 4.4.3 states that for any one parameter subgroup, there is a matrix Z such that e^{tZ} is the same map. Thus, for any $X \in \mathfrak{g}$, there exists a matrix Z such that $\Phi(e^{tX}) = e^{tZ}$. Since e^{tZ} must be in H for all t, thus $Z \in \mathfrak{h}$.
We define $\phi : \mathfrak{g} \to \mathfrak{h} : X \mapsto Z$. We will now show that this map has the desired properties.

property 0: $\Phi(e^X) = e^{\phi(X)}$.
Note that this is true by how ϕ is defined.

property 0.5: $\phi(sX) = s\phi(X)$ for all $s \in \mathbb{R}$.
This is immediate since $\phi(e^{tX}) = e^{tZ}$ thus $\phi(e^{tsX}) = e^{tsZ}$.

property 1: $\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A^{-1})$
Note that by properties 1 and 0,

$$e^{t\phi(AXA^{-1})} = e^{\phi(tAXA^{-1})} = \Phi(e^{tAXA^{-1}})$$

And then by some arithmetic the following is true:

$$\Phi(e^{tAXA^{-1}}) = \Phi(Ae^{tX}A^{-1}) = \Phi(A)\Phi(e^{tX})\Phi(A^{-1}) = \Phi(Ae^{tX}A^{-1}) = \Phi(A)e^{t\phi(X)}\Phi(A^{-1})$$

thus

$$e^{t\phi(AXA^{-1})} = \Phi(A)e^{t\phi(X)}\Phi(A^{-1})$$

Differentiating both sides and evaluating at $t = 0$ gives us

$$\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A^{-1})$$

property 2: $\phi([X,Y]) = [\phi(X),\phi(Y)]$
Recall that

$$[X,Y] = \frac{d}{dt}e^{tX}Ye^{-tX}\mid_{t=0}$$

Thus by some more arithmetic,

$$\phi([X,Y]) = \phi\left(\frac{d}{dt}e^{tX}Ye^{-tX}\mid_{t=0}\right) = \frac{d}{dt}\phi\left(e^{tX}Ye^{-tX}\right)\mid_{t=0}$$

$$= \frac{d}{dt}\Phi(e^{tX})\phi(Y)\Phi(e^{-tX})\mid_{t=0} = \frac{d}{dt}e^{t\phi(X)}\phi(Y)e^{-t\phi(X)})\mid_{t=0} = [\phi(X),\phi(Y)]$$

In this argument we used the fact that the derivative commutes with a linear transformation (prop 1.4.1).

property 3: $\phi(X) = \frac{d}{dt}(\Phi(e^{tx}))|_{t=0}$
This follows from our definition of ϕ. ∎

Given Φ, a common way to compute ϕ is by using property 3 of theorem 5.2.3.

Definition. A Lie algebra map $\phi : \mathfrak{g} \to \mathfrak{h}$ with the following property

$$\phi([X,Y]) = [\phi(X),\phi(Y)]$$

for all $X, Y \in g$, is called a **Lie algebra homomorphism**.

Theorem 5.3.1 states that every Lie group homomorphism gives rise to an associated Lie algebra homomorphism.

Theorem 5.3.2. *Suppose that G, H and K are matrix Lie groups and that $\Phi : H \to K$ and $\Psi : G \to H$ are matrix Lie group homomorphisms. Let $\Lambda : G \to K$ be the composition of Φ and Ψ. Let ϕ, ψ and λ be the corresponding Lie algebra maps. Then*

$$\lambda(X) = \phi(\psi(X))$$

This states that the Lie algebra homomorphism associated to the composition of two Lie group homomorphisms, is the composition of the two associated Lie algebra homomorphisms.

Proposition 5.3.2. *If $\Phi : G \to H$ is a Lie group homomorphism and $\phi : \mathfrak{g} \to \mathfrak{h}$ is the associated Lie algebra homomorphism, then the kernel of Φ is a closed normal subgroup of G and its associated Lie algebra is the kernel of ϕ.*

This is the Lie group extension of a standard algebraic fact.

Definition. Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then for each $A \in G$, define a linear map called the **adjoint mapping** as follows:

$$\mathrm{Ad}_A = AXA^{-1}$$

Definition. Let $\mathfrak{g}$ be a matrix Lie algebra. We denote the group of all invertable linear transformations of $\mathfrak{g}$ by $GL(\mathfrak{g})$.

Proposition 5.3.3. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. For each $A \in G$, Ad_A is an invertable linear transformation of $\mathfrak{g}$ with inverse $\mathrm{Ad}_{A^{\text{-}1}}$.*

Proposition 5.3.4. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. The map $Ad : G \to GL(\mathfrak{g}) : A \mapsto \mathrm{Ad}_A$ is a group homomorphism.*

Proposition 5.3.5. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. For each $A \in G$, Ad_A satisfies*

$$\mathrm{Ad}_A([X, Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$$

for all $X, Y \in G$.

Let $\mathfrak{g}$ be a Lie algebra with dimension k. Since $\mathfrak{g}$ and $\mathbb{R}^k$ are both vector spaces, $GL(\mathfrak{g})$ and $GL(k, \mathbb{R})$ are essentially the same.

Note that the map $Ad : G \to GL(\mathfrak{g})$ is continuous and is thus a Lie group homomorphism. Thus $GL(\mathfrak{g})$ is a matrix Lie group with an associated Lie algebra, $\mathrm{gl}(\mathfrak{g})$. Since $Ad : G \to GL(\mathfrak{g})$ is a Lie group homomorphism, by theorem 5.2.2, there is an associated Lie algebra homomorphism $ad : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$ with the property that

$$e^{\mathrm{ad}_X} = Ad(e^X)$$

Proposition 5.3.6. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Let $Ad : G \to Gl(\mathfrak{g})$ be the Lie group homomorphism defined above and let $ad : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$ be the associated Lie algebra homomorphism. Then for all $X, Y \in \mathfrak{g}$,*

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

Proof. Recall that by point 3 of Theorem 5.3.1,

$$\mathrm{ad}_X = \frac{d}{dt} Ad(e^{tX})|_{t=0}$$

Thus, by some arithmetic,

$$\mathrm{ad}_X(Y) = \frac{d}{dt} Ad(e^{tX})(Y)|_{t=0} = \frac{d}{dt} e^{tX} Y e^{\text{-}tX}|_{t=0} = [X, Y]$$

∎

Proposition 5.3.7. *For any X in $M_n(\mathbb{C})$, let $\mathrm{ad}_X : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be given by $\mathrm{ad}_X(Y) = [X, Y]$. Then for any $Y \in M_n(\mathbb{C})$ we have*

$$e^{\mathrm{ad}_X} Y = \mathrm{Ad}_{e^X} Y = e^X Y e^{\text{-}X}$$

5.4 The Exponential Mapping

Definition. If G is a matrix Lie group with lie algebra $\mathfrak{g}$, then the **exponential mapping** for G is the map

$$\exp : \mathfrak{g} \to G$$

This is just the regular exponential with domain restricted to $\mathfrak{g}$.

Though every matrix in $GL(n;\mathbb{C})$ can be expressed as the exponential of some matrix, this may not be true about this restriction because said matrix may not be in $\mathfrak{g}$. Thus this map may not be surjective.It also may not be injective. However, as we will see, it is locally bijective.

Theorem 5.4.1. *Let $U_\varepsilon := \{X \in M_n(\mathbb{C}) : ||X|| < \varepsilon\}$. This is like an open ball of radius ε around 0. Let $V_\varepsilon := \exp[U_\varepsilon]$ be its image. Suppose $G \subset GL(n;\mathbb{C})$ is a matrix Lie group with Lie algebra $\mathfrak{g}$. Then there exists an $0 < \varepsilon < \ln 2$ such that for all $A \in V_\varepsilon$, we have $A \in G$ if and only if $\log A \in \mathfrak{g}$.*

This theorem states that for any matrix Lie group G, there is a small enough radius around the origin such that the exponential mapping for G is bijective within that radius.
In order to prove this theorem, we will first need to prove a lemma.

Lemma 5.4.1. *Suppose that B_m is a sequence of elements that converge to I. Let $Y_m = \log(B_m)$. This is defined as B_m gets sufficiently close to I. Note that Y_m must converge to 0, but suppose that Y_m is nonzero for each m and that the sequence $\frac{Y_m}{||Y_m||}$ converges to $A \in M_n(\mathbb{C})$. Then $A \in \mathfrak{g}$.*

Proof. To prove this we simply need to show that $e^{tA} \in G$ for all $t \in \mathbb{R}$. As $m \to \infty$, we have $\left(\frac{t}{||Y_m||}\right) Y_m \to tA$ Since $||Y_m||$ converges to 0, we have find integers k_m such that $k_m||Y_m|| \to t$. Then

$$e^{k_m Y_m} = e^{k_m||Y_m||(Y_m/||Y_m||)} \to e^{tA}$$

However,

$$e^{k_m Y_m} = (e^{Y_m} e^{k_m}) = (B_m)^{k_m} \in G$$

Since t was arbitrary, we can find this for any t, and thus $A \in \mathfrak{g}$. ■

And now onto the proof of the theorem:

Proof. We may think of $M_n(\mathbb{C})$ as $C^{n^2} \cong \mathbb{R}^{2n^2}$, and then consider the usual bilinear form on $\mathbb{R}^{2n^2}$. Let D denote the orthogonal compliment of $\mathfrak{g}$. (See section 1.5)
Consider the map $\Phi : M_n(\mathbb{C}) \to M_n(/\mathbb{C})$ given by

$$\Phi(X + Y) = e^X e^Y$$

Where $X \in \mathfrak{g}$ and $Y \in D$.

We may consider Φ as a map $\mathbb{R}^{2n^2} \to \mathbb{R}^{2n^2}$ and thus we can compute its derivative:

$$\frac{d}{dt}\Phi(tX, 0)|_{t=0} = X \qquad \frac{d}{dt}\Phi(0, tY)|_{t=0} = Y$$

Thus the derivative of Φ at 0 is the identity, which is invertable. Recall the inverse function theorem. Therefore Φ has a local inverse in a neighborhood around I.

Ok, now for the actual proof. We are trying to prove that for some ε, if $A \in V_\varepsilon \cap G$, then $\log A \in \mathfrak{g}$. Assume for contradiction, this is not the case. Then there exists a sequence $A_m \in G$ such that $A_m \to I$, but for all m, we have $\log A_m \notin \mathfrak{g}$.

Because we know $\Phi^{\text{-}1}$ is defined around I, we have

$$A_m = e^{X_m} e^{Y_m}$$

for some sequences Xm and Y_m in G, which both go to 0. However, we must have $Y_m \neq 0$ for all m, otherwise $\log A_m = X_m \in G$. We may choose a subsequence of the Y_m such that $\frac{Y_m}{||Y_m||} \to Y$ where $||Y|| = 1$. Thus, by the lemma $Y \in \mathfrak{g}$. However, since the unit sphere in D is compact, the sequence Y_m must converge to an element of D. This is a contradiction.

Thus, the exponential mapping is bijective locally around 0. ■

Consequences of Theorem 5.4.1

Corollary 5.4.1. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then there exists a neighborhood around 0 called U which the exponential mapping maps homeomorphically to a neighborhood V around I.*

Proof. We already know that exp as well as log are both continuous around 0. This is all we need for exp to be a homeomorphism. ■

Corollary 5.4.2. *If G is a connected matrix Lie group then every element A of G can be written as*

$$A = e^{X_1} e^{X_2} \cdots e^{X_m}$$

for some $X_1, \ldots X_m \in \mathfrak{g}$.

Proof. Since G is connected, let $A(t)$ be a continuous path from I to A. Let V be the neighborhood around I in G such that the logarithm is defined and is in $\mathfrak{g}$.

We can choose a sequence of numbers ranging from 0 to 1: $0 = t_0 < t_1 < \cdots < t_m = 1$ close enough together such that

$$A(t_k)^{\text{-}1} A(t_{k-1}) \in V$$

For all $k \leq m$ Then

$$\left(A(t_1)^{\text{-}0} A(t_1)\right) \left(A(t_1)^{\text{-}1} A(t_2)\right) \ldots \left(A(t_{m-1})^{\text{-}1} A(t_m)\right) = A(t_0)^{\text{-}1} A(t_m) = I^{\text{-}1} A = A$$

Thus let X_k for $k \leq m$ be as follows

$$X_k = \log\left(A(t_k)^{-1}A(t_{k-1})\right)$$

Note that $X_k \in \mathfrak{g}$ since $A(t_k)^{-1}A(t_{k-1} \in V$. Then

$$A = e^{X_1}e^{X_2}\dots e^{X_m}$$

■

Corollary 5.4.3. *Suppose G, H are a matrix Lie groups, G is connected, and Φ_1, Φ_2 are Lie group homomorphisms of G to H. Let ϕ_1, ϕ_2 be the corresponding Lie algebra homomorphisms. If $\phi_1 = \phi_2$ then $\Phi_1 = \Phi_2$.*

Proof. Let g be any element of G. Since G is connected, Corollary 5.4.2 holds and thus g can be expressed as $g = e^{X_1}e^{X_2}\dots e^{X_m}$ with $X_i \in \mathfrak{g}$. Then

$$\Phi_1(g) = \Phi_1(e^{X_1})\Phi_1(e^{X_2})\dots\Phi_1(e^{X_m}) = e^{\phi_1(X_1)}e^{\phi_1(X_2)}\dots e^{\phi_1(X_m)}$$
$$= e^{\phi_2(X_1)}e^{\phi_2(X_2)}\dots e^{\phi_2(X_m)} = \Phi_2(e^{X_1})\Phi_2(e^{X_2})\dots\Phi_2(e^{X_m}) = \Phi_2(g)$$

■

Corollary 5.4.4. *Every continuous homomorphism between two matrix Lie groups is smooth.*

5.5 The General Lie Algebra

This section is about Lie algebras not necessarily associated to a matrix Lie group.

Note: We will only be dealing with finite-dimensional Lie algebras so every time "Lie algebra" is mentioned there is an implied "finite dimensional" before it.

Definition. A bilinear form $[\cdot,\cdot]$ on a vector space V has the **Jacobi identity** if

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

for all $X, Y, Z \in V$.

Definition. A **Lie algebra** is a finite-dimensional real or complex vector space $\mathfrak{g}$ along with a map $[\cdot,\cdot]$ from $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ with the following properties:

1. $[\cdot,\cdot]$ is bilinear
2. $[\cdot,\cdot]$ is skew-symmetric
3. The Jacobi identity holds for $[\cdot,\cdot]$

Proposition 5.5.1. *Let V be a finite-dimensional real or complex vector space and let* $\mathrm{gl}(V)$ *denote the space of all linear maps from $V \to V$. Then* $\mathrm{gl}(V)$ *along with the bracket operation is a Lie algebra.*

Proof. We will prove this by verifying each of the three conditions above holds.

1. $[\cdot, \cdot]$ is bilinear.

   We will show that it is bilinear by holding each side constant and verifying that the resulting function is in fact linear. Let $X \in V$. We will show that $[X, \cdot]$ is linear by verifying that it satisfies additivity and homogeneity.

   Additivity: Let $A, B \in V$.

   $$[X, A+B] = X(A+B) - (A+B)X = XA + XB - (AX + BX) = [X, A] + [X, B]$$

   Homogeneity: Let $A \in V$ and $c \in \mathbb{F}$.

   $$[X, cA] = X(cA) - (cA)X = c(XA - AX) = c[X, A]$$

   Thus $[X, \cdot]$ is a linear function. A symmetric argument would show the same for $[\cdot, Y]$.

2. $[\cdot, \cdot]$ is skew-symmetric.

   Let $X, Y \in V$.

   $$[X, Y] = XY - YX = \text{-}(YX - XY) = \text{-}[Y, X]$$

3. The Jacobi identity holds for $[\cdot, \cdot]$.

   $$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] =$$
   $$XYZ - XZY - YZX + ZYX + YZX - YXZ - ZXY + XZY + ZXY - ZYX - XYZ + YXZ$$
   $$= 0$$

   It all cancels out.

∎

Corollary 5.5.1. *The spaces $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ are each Lie algebras with respect to the bracket operation.*

Definition. Let $\mathfrak{g}$ be a Lie algebra. A set $\mathfrak{h} \subset \mathfrak{g}$ is a **subalgebra** of $\mathfrak{g}$ if it is a vector space and is closed under the corresponding map $[\cdot, \cdot]$.

A subalgebra is a Lie algebra.

Definition. If $\mathfrak{g}$ and $\mathfrak{g}$ are Lie algebra, then a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is called a **Lie algebra homomorphism** if

$$\phi([X, Y]) = [\phi(X), \phi(Y)]$$

.

If this map is bijective, it is called a **Lie algebra isomorphism**.

If a Lie algebra isomorphism is from a Lie algebra to itself it is called a **Lie algebra automorphism**

The inverse of a Lie algebra isomorphism is also a Lie algebra isomorphism.

Theorem 5.5.1. *Every finite-dimensional real Lie algebra is isomorphic to a subalgebra of* $gl(n;\mathbb{R})$.
Similarly, every finite-dimensional complex Lie algebra is isomorphic to a subalgebra of $gl(n;\mathbb{C})$.

We will not prove this.
We will now define the more general form of the map ad which was defined earlier for the Lie algebra of a matrix Lie group.

Definition. Let $\mathfrak{g}$ be a Lie algebra. For $X \in \mathfrak{g}$ define a linear map $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ by

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

Thus the map $X \to \mathrm{ad}_X$ is a linear map from $\mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$. This is linear as proven in the poof of proposition 5.5.1.

Proposition 5.5.2. *If* $\mathfrak{g}$ *is a Lie algebra then*

$$\mathrm{ad}_{[X,Y]} = \mathrm{ad}_X \mathrm{ad}_Y - \mathrm{ad}_Y \mathrm{ad}_X = [\mathrm{ad}_X, \mathrm{ad}_Y]$$

and thus $\mathrm{ad} : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g})$ *is a Lie algebra homomorphism.*

Proof. observe that the following holds by the Jacobi identity

$$[[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]]$$

thus let $Z \in \mathfrak{g}$.

$$\mathrm{ad}_{[X,Y]}(Z) = [[X, Y], Z] = [X, [Y, Z]] - [Y, [X, Z]] = [\mathrm{ad}_X, \mathrm{ad}_Y](Z)$$

■

Definition. If $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are Lie algebras, we can define the **direct sum** of $\mathfrak{g}_1$ and $\mathfrak{g}_2$, denoted $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ by taking their direct sum in the vector space sense and defining the bracket operation entrywise as follows

$$[(X_1, X_2), (Y_1, Y_2)] = ([X_1, Y_1], [X_2, Y_2])$$

Thus, if $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are Lie algebras then so is their direct product $\mathfrak{g}_1 \oplus \mathfrak{g}_2$.

5.6 Complixification of Real Lie Algebras

Definition. If V is a finite-dimensional real vector space, then the **complexification** of V denoted $V_{\mathbb{C}}$ is the space of

$$v_1 + v_2 i$$

for all $v_1, v_2 \in V$. This becomes a real vector space if we view it as V^2 or a complex vector space if we define

$$i(v_1 + v_2 i) = v_1 i - v_2$$

Note that V is a real subspace of $V_{\mathbb{C}}$.

Proposition 5.6.1. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra and let $\mathfrak{g}_C$ be its complexification as a vector space. Then the bracket operation has a unique extension to $\mathfrak{g}_{\mathbb{C}}$ making it a complex Lie algebra. The bracket operation on $\mathfrak{g}_{\mathbb{C}}$ is defined as follows:*

$$[X_1 + X_2 i, Y_1 + Y_2 i] = ([X_1, Y_1] - [X_2, Y_2]) + ([X_1, Y_2] - [Y_2, X_1])i$$

Proposition 5.6.2. *The following isomorphisms hold:*

$$\begin{aligned} \mathrm{gl}(n;\mathbb{R})_{\mathbb{C}} &\cong \mathrm{gl}(n;\mathbb{C}) \\ \mathrm{u}(n)_{\mathbb{C}} &\cong gl(n,\mathbb{C}) \\ \mathrm{su}(n)_{\mathbb{C}} &\cong \mathrm{sl}(n;\mathbb{C}) \\ \mathrm{sl}(n,\mathbb{R})_{\mathbb{C}} &\cong \mathrm{sl}(n;\mathbb{C}) \end{aligned}$$

6 Properties of Lie Groups and their Lie Algebras

6.1 The Baker-Campbell-Hausdorff Formula

The Baker-Campbell-Hausdorff Formula, or BCH,is a formula for expressing $\log(e^X e^Y)$ in Lie algebraic terms, that is in terms of iterated commutators. This will be useful for proving results in later sectons.

First we will define a function g.

$$g(z) = \frac{\log z}{1 - \frac{1}{z}}$$

This function is defined and holomorphic (analytic) in the disk $\{|z-1| < 1\}$, which is the disk of radius 1 around 1. Since g is holomorphic, it can be expressed as a power series of the following form:

$$g(z) = \sum_{m=0}^{\infty} a_m (z-1)^m$$

This series has radius of convergence 1.

Let V be a vector space. We will extend g to be defined on the set of linear transformations of V as follows. First, we define the Hilbert-Schmidt Norm by considering $V \cong C^n$. Thus for all linear operators on V such that $||A - I|| < 1$, we define

$$g(A) = \sum_{m=0}^{\infty} a_m (A-I)^m$$

Now that we have defined g, we can state BCH.

Theorem 6.1.1 (Baker-Campbell-Hausdorff Formula)**.** *For all $n \times n$ complex matrices X and Y such that $||X||$ and $||Y||$ are small, we have*

$$\boxed{\log(e^X e^Y) = X + \int_0^1 g(e^{\mathrm{ad}_X} e^{t\mathrm{ad}_Y})(Y)}$$

6.2 Group Versus Lie Algebra Homomorphisms

Recall that given matrix Lie groups G and H with lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, and a Lie group homomorphism $\Phi : G \to H$, there is a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ such that

$$\Phi(e^X) = e^{\phi(X)}$$

This was theorem 5.3.1. In this section we will prove the following: Given that G is simply connected, the converse is true.

Theorem 6.2.1. *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. If G is simply connected, then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that*

$$\Phi(e^X) = e^{\phi(X)}$$

.

In order to prove this theorem, we will need to construct a local homomorphism from ϕ.

Definition. If G and H are matrix Lie group, a **local homomorphism** of G to H is a pair (U, f) where U is a path connected neighborhood around the idenity of G and $f : U \to H$ is a continuous map such that $f(AB) = f(A)f(B)$ for all A, B such that $A, B, AB \in U$.

Note that U does not need to be a subgroup. A local homomorphism is as much of a homomorphism as a function could be on a domain which isn't closed under its operation.

Proposition 6.2.1. *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ and let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie group homomorphism. Define $U_\varepsilon \subset G$ by*

$$U_\varepsilon = \{A \in G : ||A - I|| < I \text{ and } ||\log A|| < \varepsilon\}$$

then there exists an $\varepsilon > 0$ such that $f : U_\varepsilon \to H$ given by

$$f(A) = e^{\phi(\log A)}$$

is a homomorphism.

Proof. Note that by theorem 5.4.2, there exists an ε small enough such taht if $||\log A|| < \varepsilon$ then $\log A \in \mathfrak{g}$. Chose such an ε small enough such that both theorem 5.4.1 applies and that for all $A, B \in U_\varepsilon$, BCH applies to $X := \log A \in \mathfrak{g}$ and $Y := \log B \in \mathfrak{g}$ and to $\phi(X) \in \mathfrak{h}$ and $\phi(Y) \in \mathfrak{h}$. Then if $AB \in U_\varepsilon$ then we have

$$f(AB) = f(e^X e^Y) = e^{\phi(\log(e^X e^Y))}$$

We can compute $\log(e^X e^Y)$ using BCH and then apply ϕ. Since ϕ is a Lie algebra homomorphism, after ϕ is applied, all quantities involving X and Y will simply change to be $\phi(X)$ and $\phi(Y)$. Thus

$$\phi\left(\log(e^X e^Y)\right) = \phi(X) + \int_0^1 \sum_{m=0}^{\infty} a_m (e^{\mathrm{ad}_{\phi(X)}} e^{\mathrm{ad}_{t\phi(Y)}} - I)^m \phi(Y) dt = \log\left(e^{\phi(X)} e^{\phi(Y)}\right)$$

Then we have

$$f(AB) = e^{\log(e^{\phi(X)} e^{\phi(Y)})} = e^{\phi(X)} e^{\phi(Y)} = f(A)f(B)$$

as desired. ■

Theorem 6.2.2. *Let G and H be matrix Lie groups and let G be simply connected. If (U, f) is a local homomorphism of G onto H, then there exists a unique global Lie group homomorphism $\Phi : G \to H$ such that Φ agrees with f on U.*

Proof. This proof will come in 4 steps:

Step 1: Define Φ along a path.

Let $A \in G$. Since G is simply connected and thus connected, there exists a path from A to I.

Definition. We call a partition $0 = t_0 < t_1 < \cdots < t_m = 1$ of $[0, 1]$ a **good partition** if the intervals are small enough such that for all s and t in the same subinterval, we have

$$A(s)A(t)^{-1} \in U$$

Corollary 5.4.2 showed that such partitions always exist.

Let $t_0, \ldots, t_m$ be a good partition of the path from A to I. Since $A(t_0) = I = A(t_0)^{-1}$ and $A(t_1)A(t_0)^{-1} \in U$, we have that $A(t_1) \in U$. We can write A as follows:

$$A = (A(t_m)A(t_{m-1})) \ldots (A(t_1)A(t_0))$$

Since Φ is supposed to be a homomorphism which agrees with f near the identity, it is reasonable to define

$$\Phi(A) = f\left(A(t_m)A(t_{m-1})\right) \ldots f\left(A(t_1)A(t_0)\right)$$

Step 2: Independence of partition

For any good partition, if we insert another point s between t_j and t_{j+1}, the partition points only get closer together so we clearly still have a good partition. In our definition of $\Phi(A)$, this would replace the factor $f(A(t_j)A(t_{j+1})^{-1})$ with $f(A(t_j)A(s)^{-1})f(A(s)A(t_{j+1})^{-1})$. This is okay though because since f is a local homomorphism and since $A(t_j)A(s)^{-1}, A(s)A(t_{j+1})^{-1}$ and $A(t_j)A(s)^{-1}A(s)A(t_{j+1})^{-1} = A(t_j)A(t_{j+1})^{-1}$ are all in U, then the following holds:

$$f(A(t_j)A(s)^{-1})f(A(s)A(t_{j+1})^{-1}) = f(A(t_j)A(s)^{-1}A(s)A(t_{j+1})^{-1}) = f(A(t_j)A(t_{j+1})^{-1})$$

And thus adding another point to the partition doesn't change the value of $\Phi(A)$.

One can extend this argument to see that adding any finite number of points to a partition will not change the value of $\Phi(A)$.

Let p and q be good partitions and let u be their union. Since adding points to p to get u won't change the value of $\Phi(A)$, you'd get the same value with p or u. The same argument holds for q and u. Thus p and q give the same value for $\Phi(A)$.

Step 3: Independence of path

This will be the step where we use simple connectedness. Let $A_0(t)$ and $A_1(t)$ be two paths joining I to A. Since G is simply connected, we know that $A_0(t)$ and $A_1(t)$ are homotopic. (see section 2.1)

In this context, $A_0 : [0,1] \to G$ and $A_1 : [0,1] \to G$ being homotopic means that there exists a continuous map (homotopy) $H : [0,1] \times [0,1] \to G$ such that

$$H(0,t) = A_0(t) \quad \text{and} \quad H(1,t) = A_1(t)$$

for all $t \in [0,1]$, and also

$$H(s,0) = 1 \quad \text{and} \quad H(s,1) = A$$

for all $s \in [0,1]$

Note that the interval $[0,1] \times [0,1]$ is compact since it is closed and bounded. Thus there exists an integer N (which will represent the number of partition pieces of H) sufficiently large such that for all (s,t) and (s',t') if $|s-s'| < \frac{2}{N}$ and $|t-t'| < \frac{2}{N}$ then we have $H(s,t)H(s',t')^{-1} \in U$. This says that as long as we partition the $[0,1]$ intervals into at least N even segments, then we are twice as close as we need to be to be a good partition. This means that even if we took out one point, the partition would still be good. This will be necessary later in this step.

We will now proceed to deform A_0 a little bit at a time to get A. We will do this by defining a sequence $B_{k,l}$ of paths such that $k \leq N-1$ and $B \leq N$. Define the paths as follows:

$$B_{k,l}(t) = \begin{cases} H(\frac{k+1}{N}, t) & \text{if } t \leq \frac{l-1}{N} \\ H(\frac{k+1}{N} - (t - \frac{l-1}{N}), t) & \text{if } \frac{l-1}{N} < t < \frac{l}{N} \\ H(\frac{k}{N}, t) & \text{if } t \geq \frac{l}{N} \end{cases}$$

In words, what this means is that when t is less that $\frac{l-1}{N}$, then $B_{k,l}$ coincides with $H(\frac{k+1}{N}, t)$. When t is between $\frac{l-1}{N}$ and $\frac{l}{N}$, then t coincides with H along a diagonal bath in the s,t plane. Then, once $B_{k,l}$ reaches $t = \frac{l}{N}$, then $B_{k,l}$ coincides with $H(\frac{k}{N}, t)$. Visually this looks like the diagram below:

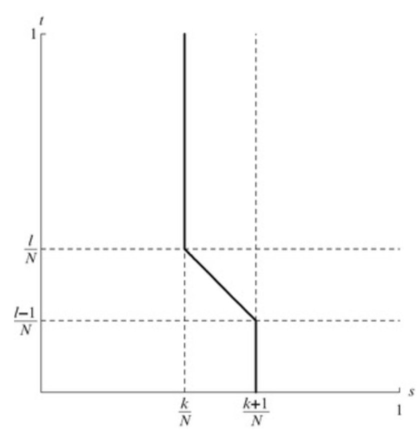

Figure 1: The path $B_{k,l}$

We may think of this sequence of paths as deforming A_0 to A_1 in steps. First we have $A_0 = B_{0,0}$, then $B_{0,1}, \dots B_{0,N}, B_{1,0}, \dots B_{N,N} = A_1$. We want to show that our N is large enough that these steps are close enough that the value of $\Phi(A)$ doesn't change as we move along each step.

Note that for a fixed $k < l$, $B_{k,l}(t) = B_{k,l+1}(t)$ for all t except for when $\frac{l-1}{N} < t < \frac{l+1}{N}$. Recall that the value of $\Phi(A)$ is only dependent on the values of $A(t)$ at the partition points of a good partition. By construction, the following partition is good:

$$0, \frac{1}{N}, \dots, \frac{l-1}{N}, \frac{l}{N}, \frac{l+1}{N}, \dots, 1$$

The values of $B_{k,l}(t)$ and $B_{k,l+1}(t)$ are the same at each of these partition points except for $\frac{l}{N}$. However, we have chosen N such that we may remove a partition point and still have a good partition. Thus remove $\frac{l}{N}$, and we have a good partition where $B_{k,l}(t)$ agrees with $B_{k,l+1}(t)$ on the value of $\Phi(A)$.

The same argument works for a fixed l. Thus the paths $B_{k,l}$ all have the same value of $\Phi(A)$. Namely $A_0(t)$ and $A_1(t)$ give the same value for $\Phi(A)$. Thus, we know that $\Phi(A)$ is independent of path.

It is obvious that Φ is a homomorphism and agrees with ϕ on U. Thus, Φ as we defined, is our desired result. ■

And now, onto the proof of theorem 6.2.1.

Proof. Lets remind ourselves of the theorem:

Theorem 6.2.1 *Let G and H be matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. If G is simply connected, then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that*

$$\Phi(e^X) = e^{\phi(X)}$$

.

Let G and H be matrix lie groups as stated in the theorem with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ and a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$. There are two things we need to prove: existence and uniqueness.

Existence: By proposition 6.2.1, there exists a local homomorphism f from a neighborhood U of G to H, defined by $f(A) = e^{\phi(\log A)}$. Thus, by theorem 6.2.2, there exists a global Lie group homomorphism $\Phi : G \to H$ such that Φ agrees with f in the neighborhood U.

Let $X \in G$. There exists an m large enough that $\frac{X}{m} \in U$. Thus

$$\Phi(e^{\frac{X}{m}}) = f(e^{\frac{X}{m}}) = e^{\phi(\frac{X}{m})}$$

Since Φ is a homomorphism, we have :

$$\Phi(e^X) = \Phi(e^{\frac{X}{m}})^m = e^{\phi(X)}$$

as desired.

Uniqueness: Let Φ_1 and Φ_2 be two such Lie group homomorphisms. Let $A \in G$. We can express A as in corollary 5.4.2: $A = e^{X_1} \cdots e^{X_m}$ and then we have

$$\Phi_1(A) = \Phi_2(A) = e^{\phi(X_1)} \cdots e^{\phi(X_m)}$$

Thus such a Φ must be unique. ■

We will conclude this section with a few results of this theorem

Corollary 6.2.1. *Suppose G and H are simply-connected matrix Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. If $\mathfrak{g}$ is isomorphic to $\mathfrak{h}$ then G is isomorphic to H.*

Proof. Let $\phi : \mathfrak{g} \to \mathfrak{h}$ be the Lie algebra homomorphism. By theorem 5.6, there is an associated Lie group homomorphism Φ. Let $\psi = \phi^{-1}$ and let Ψ be the Lie group homomorphism associated to ψ. Since $\phi \circ \psi = \mathrm{id}$, $\Phi \circ \Psi$ must also be the identity by Theorem 5.3.2. Thus $\Psi = \Phi^{-1}$ and Φ is a Lie group isomorphism. ■

Theorem 6.2.3. *Let G be a simply connected Lie group with Lie algebra $\mathfrak{g}$. Suppose that $\mathfrak{g}$ decomposes as the direct sum of two of its subalgebras. In other words, suppose that there exists Lie algebras $\mathfrak{h}_1$ and $\mathfrak{h}_2$ such that $\mathfrak{g} \cong \mathfrak{h}_1 \oplus \mathfrak{h}_2$. Then G has two simply connected closed subgroups, H_1 and H_2, whose Lie algebras are $\mathfrak{h}_1$ and $\mathfrak{h}_2$ and $G \cong H_1 \times H_2$.*

Proof. There are three things we need to prove. (1) The associated Lie groups to $\mathfrak{h}_1$ and $\mathfrak{h}_2$, H_1 and H_2 are in fact closed, connected subgroups, (2) H_1 and H_2 are simply connected, (3) G is isomorphic to the product of H_1 and H_2.

(1) Consider the Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{g}$ that sends elements of $\mathfrak{g}$ to their component in $\mathfrak{h}_1$. Since G is simply connected, theorem 6.2.1 applies and applies and thus there is a an associated lie group homomorphism $\Phi : G \to G$. By proposition 5.3.2, the Lie algebra of $\ker(\Phi)$ is $\ker(\phi)$. Note that $\ker(\phi)$ is $\mathfrak{h}_2$ since everything in $\mathfrak{h}_2$ has no $\mathfrak{h}_1$ component, and thus maps to the identity. Let H_2 be the identity component of $\ker(\Phi)$. Since Φ is continuous, $\ker(\Phi)$ is closed and so is it's identity component H_2. Thus H_2 is a closed, connected subgroup of G. The same argument can be done for H_1.

(2) Suppose that $A(t)$ is a loop in H_1. Since G is simply connected, there is a homotopy $A(s,t)$ shrinking $A(t)$ to as point in G. Note that ϕ acts as the identity on $\mathfrak{h}_1$. Thus Φ is the identity on H_1. Thus, Φ maps $A(s,t)$ to a homotopy taking $A(t)$ to a point in H_1. Thus H_1 is simply connected. A similar argument shows the same for H_2.

(3) Since $\mathfrak{g} \cong \mathfrak{h}_1 \oplus \mathfrak{h}_2$, elements of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ commute. It follows that elements of H_1 commute with elements of H_2. Thus we have a Lie group homomorphism $\Psi : H_1 \times H_2 \to G$ given by $\Psi(A,B) = AB$. Let $X \in \mathfrak{h}_1$ and $Y \in \mathfrak{h}_2$. Since elements of $\mathfrak{h}_1$ and $\mathfrak{h}_2$ commute, $e^X e^Y = e^{X+Y} \in G$, and this ϕ is just the original homomorphism from $\mathfrak{h}_1 \oplus \mathfrak{h}_2 \to G$. Since G is simply connected, by corollary 6.2.1, $G \cong H_1 \times H_2$. ■

6.3 Universal Covers

Theorem 6.2.1 states that if G is simply connected, then exponentiating the Lie algebra homomorphism gives rise to a Lie group homomorphism. However, if G is not simply connected, we may look for another group $\tilde{G}$ which is simply connected and shares a Lie algebra with G.

Definition. Let G be a connecred matrix Lie group. Then a **universal cover** of G is a simply connected matrix Lie group H together with a Lie group homomorphism $\Phi : G \to H$ such that the associated Lie algebra map is an isomorphism. The homomorphism Φ is called a **covering map**.

As we will see with the following proposition, universal covers are unique up to isomorphism.

Proposition 6.3.1. *If G is a connected matrix Lie group and (H_1, Φ_1) and (H_2, Φ_2) are universal covers of G, then there exists a Lie group homomorphism $\Psi : H_1 \to H_2$ such that $\Phi_1 \circ \Psi = \Phi_1$.*

Thus, it makes sense to speak of *the* universal cover of G.

Also, if H is a simply connected Lie group and $\phi : \mathfrak{h} \to \mathfrak{g}$ is a Lie algebra isomorphism, then by theorem 6.2.1, we can construct an associated Lie algebra homomorphism $\Phi : H \to G$ such that (H, Φ) is a universal cover of G.

We may define a universal cover more informally as follows: The universal cover of a matrix Lie group G is a simply connected matrix Lie group $\tilde{G}$ that shares a Lie algebra with G.

6.4 Subgroups and Subalgebras

This section is to provide an answer to the following question: If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, and $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, dies there exist a matrix Lie group $H \subset G$ such that the Lie algebra of H is $\mathfrak{h}$?

If exp were a homeomorphism between $\mathfrak{g}$ and G and if BCH worked globally, then the answer would be yes.