Chapter 2

Basics of Classical Lie Groups: The Exponential Map, Lie Groups, and Lie Algebras

Le rôle prépondérant de la théorie des groupes en mathématiques a été longtemps insoupçonné; il y a quatre-vingts ans, le nom même de groupe était ignoré. C'est Galois qui, le premier, en a eu une notion claire, mais c'est seulement depuis les travaux de Klein et surtout de Lie que l'on a commencé à voir qu'il n'y a presque aucune théorie mathématique où cette notion ne tienne une place importante.
—Henri Poincaré

2.1 The Exponential Map

The inventors of Lie groups and Lie algebras (starting with Lie!) regarded Lie groups as groups of symmetries of various topological or geometric objects. Lie algebras were viewed as the "infinitesimal transformations" associated with the symmetries in the Lie group.

For example, the group $\mathbf{SO}(n)$ of rotations is the group of orientation-preserving isometries of the Euclidean space $\mathbb{E}^n$.

The Lie algebra $\mathfrak{so}(n, \mathbb{R})$ consisting of real skew symmetric $n \times n$ matrices is the corresponding set of infinitesimal rotations.

The geometric link between a Lie group and its Lie algebra is the fact that the Lie algebra can be viewed as the tangent space to the Lie group at the identity.

There is a map from the tangent space to the Lie group, called the *exponential map*.

The Lie algebra can be considered as a linearization of the Lie group (near the identity element), and the exponential map provides the "delinearization," i.e., it takes us back to the Lie group.

These concepts have a concrete realization in the case of groups of matrices, and for this reason we begin by studying the behavior of the exponential map on matrices.

We begin by defining the exponential map on matrices and proving some of its properties.

The exponential map allows us to "linearize" certain algebraic properties of matrices.

It also plays a crucial role in the theory of linear differential equations with constant coefficients.

But most of all, as we mentioned earlier, it is a stepping stone to Lie groups and Lie algebras.

At first, we define *manifolds* as embedded submanifolds of $\mathbb{R}^N$, and we define linear Lie groups, using the famous result of Cartan (apparently actually due to Von Neumann) that a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$ is a manifold, and thus, a Lie group.

This way, *Lie algebras* can be "computed" using tangent vectors to curves of the form $t \mapsto A(t)$, where $A(t)$ is a matrix.

Given an $n \times n$ (real or complex) matrix $A = (a_{i,\,j})$, we would like to define the exponential e^A of A as the sum of the series

$$e^A = I_n + \sum_{p \geq 1} \frac{A^p}{p!} = \sum_{p \geq 0} \frac{A^p}{p!},$$

letting $A^0 = I_n$.

The following lemma shows that the above series is indeed absolutely convergent.

Lemma 2.1.1 *Let* $A = (a_{i\,j})$ *be a (real or complex)* $n \times n$ *matrix, and let*

$$\mu = \max\{|a_{i\,j}| \mid 1 \leq i, j \leq n\}.$$

If $A^p = (a^p_{i\,j})$*, then*

$$|a^p_{i\,j}| \leq (n\mu)^p$$

for all i, j*,* $1 \leq i, j \leq n$*. As a consequence, the* n^2 *series*

$$\sum_{p \geq 0} \frac{a^p_{i\,j}}{p!}$$

converge absolutely, and the matrix

$$e^A = \sum_{p \geq 0} \frac{A^p}{p!}$$

is a well-defined matrix.

It is instructive to compute explicitly the exponential of some simple matrices. As an example, let us compute the exponential of the real skew symmetric matrix

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

We find that

$$e^A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Thus, e^A is a rotation matrix!

This is a general fact. If A is a skew symmetric matrix, then e^A is an orthogonal matrix of determinant $+1$, i.e., a rotation matrix.

Furthermore, every rotation matrix is of this form; i.e., the exponential map from the set of skew symmetric matrices to the set of rotation matrices is surjective.

In order to prove these facts, we need to establish some properties of the exponential map.

But before that, let us work out another example showing that the exponential map is not always surjective.

Let us compute the exponential of a real 2×2 matrix with null trace of the form

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

We need to find an inductive formula expressing the powers A^n. Observe that

$$A^2 = (a^2 + bc)I_2 = -\det(A)I_2.$$

If $a^2 + bc = 0$, we have

$$e^A = I_2 + A.$$

If $a^2 + bc < 0$, let $\omega > 0$ be such that $\omega^2 = -(a^2 + bc)$.

Then, $A^2 = -\omega^2 I_2$, and we get

$$e^A = \cos\omega\, I_2 + \frac{\sin\omega}{\omega} A.$$

If $a^2 + bc > 0$, let $\omega > 0$ be such that $\omega^2 = (a^2 + bc)$. Then $A^2 = \omega^2 I_2$, and we get

$$e^A = \cosh\omega\, I_2 + \frac{\sinh\omega}{\omega} A,$$

where $\cosh\omega = (e^\omega + e^{-\omega})/2$ and $\sinh\omega = (e^\omega - e^{-\omega})/2$.

It immediately verified that in all cases,

$$\det\left(e^A\right) = 1.$$

This shows that the exponential map is a function from the set of 2×2 matrices with null trace to the set of 2×2 matrices with determinant 1.

This function is not surjective. Indeed, $tr(e^A) = 2\cos\omega$ when $a^2 + bc < 0$, $tr(e^A) = 2\cosh\omega$ when $a^2 + bc > 0$, and $tr(e^A) = 2$ when $a^2 + bc = 0$.

As a consequence, for any matrix A with null trace,

$$\mathrm{tr}\left(e^A\right) \geq -2,$$

and any matrix B with determinant 1 and whose trace is less than -2 is not the exponential e^A of any matrix A with null trace. For example,

$$B = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix},$$

where $a < 0$ and $a \neq -1$, is not the exponential of any matrix A with null trace.

A fundamental property of the exponential map is that if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then the eigenvalues of e^A are $e^{\lambda_1}, \ldots, e^{\lambda_n}$. For this we need two lemmas.

Lemma 2.1.2 *Let A and U be (real or complex) matrices, and assume that U is invertible. Then*

$$e^{UAU^{-1}} = Ue^A U^{-1}.$$

Say that a square matrix A is an *upper triangular matrix* if it has the following shape,

$$\begin{pmatrix} a_{1\,1} & a_{1\,2} & a_{1\,3} & \ldots & a_{1\,n-1} & a_{1\,n} \\ 0 & a_{2\,2} & a_{2\,3} & \ldots & a_{2\,n-1} & a_{2\,n} \\ 0 & 0 & a_{3\,3} & \ldots & a_{3\,n-1} & a_{3\,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1\,n-1} & a_{n-1\,n} \\ 0 & 0 & 0 & \ldots & 0 & a_{n\,n} \end{pmatrix},$$

i.e., $a_{i\,j} = 0$ whenever $j < i$, $1 \leq i, j \leq n$.

Lemma 2.1.3 *Given any complex $n \times n$ matrix A, there is an invertible matrix P and an upper triangular matrix T such that*

$$A = PTP^{-1}.$$

Remark: If E is a Hermitian space, the proof of Lemma 2.1.3 can be easily adapted to prove that there is an *orthonormal* basis $(u_1, \ldots, u_n)$ with respect to which the matrix of f is upper triangular.

In terms of matrices, this means that there is a unitary matrix U and an upper triangular matrix T such that $A = UTU^*$. This is usually known as *Schur's Lemma*.

Using this result, we can immediately rederive the fact that if A is a Hermitian matrix, then there is a unitary matrix U and a real diagonal matrix D such that $A = UDU^*$.

If $A = PTP^{-1}$ where T is upper triangular, note that the diagonal entries on T are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A.

Indeed, A and T have the same characteristic polynomial. This is because if A and B are any two matrices such that $A = PBP^{-1}$, then

$$\begin{aligned}
\det(A - \lambda\, I) &= \det(PBP^{-1} - \lambda\, P\, IP^{-1}),\\
&= \det(P(B - \lambda\, I)P^{-1}),\\
&= \det(P)\det(B - \lambda\, I)\det(P^{-1}),\\
&= \det(P)\det(B - \lambda\, I)\det(P)^{-1},\\
&= \det(B - \lambda\, I).
\end{aligned}$$

Furthermore, it is well known that the determinant of a matrix of the form

$$\begin{pmatrix}
\lambda_1 - \lambda & a_{1\,2} & a_{1\,3} & \dots & a_{1\,n-1} & a_{1\,n}\\
0 & \lambda_2 - \lambda & a_{2\,3} & \dots & a_{2\,n-1} & a_{2\,n}\\
0 & 0 & \lambda_3 - \lambda & \dots & a_{3\,n-1} & a_{3\,n}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
0 & 0 & 0 & \dots & \lambda_{n-1} - \lambda & a_{n-1\,n}\\
0 & 0 & 0 & \dots & 0 & \lambda_n - \lambda
\end{pmatrix}$$

is $(\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$.

Thus the eigenvalues of $A = PTP^{-1}$ are the diagonal entries of T. We use this property to prove the following lemma:

Lemma 2.1.4 *Given any complex $n \times n$ matrix A, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then $e^{\lambda_1}, \ldots, e^{\lambda_n}$ are the eigenvalues of e^A. Furthermore, if u is an eigenvector of A for λ_i, then u is an eigenvector of e^A for e^{λ_i}.*

As a consequence, we can show that

$$\det(e^A) = e^{\mathrm{tr}(A)},$$

where $\mathrm{tr}(A)$ is the *trace of A*, i.e., the sum $a_{1\,1} + \cdots + a_{n\,n}$ of its diagonal entries, which is also equal to the sum of the eigenvalues of A.

This is because the determinant of a matrix is equal to the product of its eigenvalues, and if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then by Lemma 2.1.4, $e^{\lambda_1}, \ldots, e^{\lambda_n}$ are the eigenvalues of e^A, and thus

$$\det\left(e^A\right) = e^{\lambda_1} \cdots e^{\lambda_n} = e^{\lambda_1 + \cdots + \lambda_n} = e^{\mathrm{tr}(A)}.$$

This shows that e^A is always an invertible matrix, since e^z is never zero for every $z \in \mathbb{C}$.

In fact, the inverse of e^A is e^{-A}, but we need to prove another lemma.

This is because it is generally not true that

$$e^{A+B} = e^A e^B,$$

unless A and B commute, i.e., $AB = BA$.

Lemma 2.1.5 *Given any two complex $n \times n$ matrices A, B, if $AB = BA$, then*

$$e^{A+B} = e^A e^B.$$

Now, using Lemma 2.1.5, since A and $-A$ commute, we have

$$e^A e^{-A} = e^{A+-A} = e^{0_n} = I_n,$$

which shows that the inverse of e^A is e^{-A}.

We will now use the properties of the exponential that we have just established to show how various matrices can be represented as exponentials of other matrices.

First, we review some more or less standard results about matrices.

2.2 Normal, Symmetric, Skew Symmetric, Orthogonal, Hermitian, Skew Hermitian, and Unitary Matrices

First, we consider real matrices.

Definition 2.2.1 Given a real $m \times n$ matrix A, the *transpose* $A^\top$of A is the $n \times m$ matrix $A^\top = (a^\top_{i,\,j})$ defined such that

$$a^\top_{i,\,j} = a_{j,\,i}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. A real $n \times n$ matrix A is

1. *normal* iff
$$A\,A^\top = A^\top A,$$
2. *symmetric* iff
$$A^\top = A,$$
3. *skew symmetric* iff
$$A^\top = -A,$$
4. *orthogonal* iff
$$A\,A^\top = A^\top A = I_n.$$

Theorem 2.2.2 *For every normal matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PD\,P^\top$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix}$$

such that each block D_i is either a one-dimensional matrix (i.e., a real scalar) or a two-dimensional matrix of the form

$$D_i = \begin{pmatrix} \lambda_i & -\mu_i \\ \mu_i & \lambda_i \end{pmatrix}$$

where $\lambda_i, \mu_i \in \mathbb{R}$, with $\mu_i > 0$.

Theorem 2.2.3 *For every symmetric matrix A, there is an orthogonal matrix P and a diagonal matrix D such that $A = PD\,P^{\top}$, where D is of the form*

$$D = \begin{pmatrix} \lambda_1 & & \ldots & \\ & \lambda_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & \lambda_n \end{pmatrix}$$

where $\lambda_i \in \mathbb{R}$.

Theorem 2.2.4 *For every skew symmetric matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = P D P^\top$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \dots & \\ & D_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & D_p \end{pmatrix}$$

such that each block D_i is either 0 or a two-dimensional matrix of the form

$$D_i = \begin{pmatrix} 0 & -\mu_i \\ \mu_i & 0 \end{pmatrix}$$

where $\mu_i \in \mathbb{R}$, with $\mu_i > 0$. In particular, the eigenvalues of A are pure imaginary of the form $i\mu_i$, or 0.

Theorem 2.2.5 *For every orthogonal matrix A, there is an orthogonal matrix P and a block diagonal matrix D such that $A = PD\,P^{\top}$, where D is of the form*

$$D = \begin{pmatrix} D_1 & & \ldots & \\ & D_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & D_p \end{pmatrix}$$

such that each block D_i is either 1, -1, or a two-dimensional matrix of the form

$$D_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

where $0 < \theta_i < \pi$.

In particular, the eigenvalues of A are of the form $\cos\theta_i \pm i\sin\theta_i$, or 1, or -1.

If $\det(A) = +1$ (A is a *rotation matrix*), then the number of -1's must be even.

In this case, we can pair the entries -1 as matrices

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

which amounts to allowing $\theta_i = \pi$. So, a rotation matrix can be written as $A = PDP^\top$ for some orthogonal matrix, P, and some block diagonal matrix

$$D = \begin{pmatrix} D_1 & & \ldots & \\ & D_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & D_p \end{pmatrix},$$

where each block D_i is either 1 or a two-dimensional matrix of the form

$$D_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

where $0 < \theta_i \leq \pi$. Furthermore, we can always regroup the 1's together.

We now consider complex matrices.

Definition 2.2.6 Given a complex $m \times n$ matrix A, the *transpose* $A^\top$ of A is the $n \times m$ matrix $A^\top = (a^\top_{i,j})$ defined such that

$$a^\top_{i,j} = a_{j,i}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. The *conjugate* $\overline{A}$ of A is the $m \times n$ matrix $\overline{A} = (b_{i,j})$ defined such that

$$b_{i,j} = \overline{a}_{i,j}$$

for all i, j, $1 \leq i \leq m$, $1 \leq j \leq n$. Given an $n \times n$ complex matrix A, the *adjoint* A^* of A is the matrix defined such that

$$A^* = \overline{(A^\top)} = (\overline{A})^\top.$$

A complex $n \times n$ matrix A is

1. *normal* iff
$$AA^* = A^*A,$$
2. *Hermitian* iff
$$A^* = A,$$
3. *skew Hermitian* iff
$$A^* = -A,$$
4. *unitary* iff
$$AA^* = A^*A = I_n.$$

Theorem 2.2.7 *For every complex normal matrix A, there is a unitary matrix U and a diagonal matrix D such that $A = UDU^*$. Furthermore, if A is Hermitian, D is a real matrix, if A is skew Hermitian, then the entries in D are pure imaginary or zero, and if A is unitary, then the entries in D have absolute value 1.*

2.3 The Lie Groups $\mathbf{GL}(n,\mathbb{R})$, $\mathbf{SL}(n,\mathbb{R})$, $\mathbf{O}(n)$, $\mathbf{SO}(n)$, the Lie Algebras $\mathfrak{gl}(n,\mathbb{R})$, $\mathfrak{sl}(n,\mathbb{R})$, $\mathfrak{o}(n)$, $\mathfrak{so}(n)$, and the Exponential Map

The set of real invertible $n \times n$ matrices forms a group under multiplication, denoted by $\mathbf{GL}(n,\mathbb{R})$.

The subset of $\mathbf{GL}(n,\mathbb{R})$ consisting of those matrices having determinant $+1$ is a subgroup of $\mathbf{GL}(n,\mathbb{R})$, denoted by $\mathbf{SL}(n,\mathbb{R})$.

It is also easy to check that the set of real $n \times n$ *orthogonal matrices* forms a group under multiplication, denoted by $\mathbf{O}(n)$.

The subset of $\mathbf{O}(n)$ consisting of those matrices having determinant $+1$ is a subgroup of $\mathbf{O}(n)$, denoted by $\mathbf{SO}(n)$. We will also call matrices in $\mathbf{SO}(n)$ *rotation matrices*.

Staying with easy things, we can check that the set of real $n \times n$ matrices with null trace forms a vector space under addition, and similarly for the set of skew symmetric matrices.

Definition 2.3.1 The group $\mathbf{GL}(n, \mathbb{R})$ is called the *general linear group*, and its subgroup $\mathbf{SL}(n, \mathbb{R})$ is called the *special linear group*. The group $\mathbf{O}(n)$ of orthogonal matrices is called the *orthogonal group*, and its subgroup $\mathbf{SO}(n)$ is called the *special orthogonal group* (or *group of rotations*). The vector space of real $n \times n$ matrices with null trace is denoted by $\mathfrak{sl}(n, \mathbb{R})$, and the vector space of real $n \times n$ skew symmetric matrices is denoted by $\mathfrak{so}(n)$.

Remark: The notation $\mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{so}(n)$ is rather strange and deserves some explanation. The groups $\mathbf{GL}(n, \mathbb{R})$, $\mathbf{SL}(n, \mathbb{R})$, $\mathbf{O}(n)$, and $\mathbf{SO}(n)$ are more than just groups.

They are also topological groups, which means that they are topological spaces (viewed as subspaces of $\mathbb{R}^{n^2}$) and that the multiplication and the inverse operations are continuous (in fact, smooth).

Furthermore, they are smooth real manifolds.

The real vector spaces $\mathfrak{sl}(n)$ and $\mathfrak{so}(n)$ are what is called *Lie algebras*.

However, we have not defined the algebra structure on $\mathfrak{sl}(n, \mathbb{R})$ and $\mathfrak{so}(n)$ yet.

The algebra structure is given by what is called the *Lie bracket*, which is defined as

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

Lie algebras are associated with Lie groups.

What is going on is that the Lie algebra of a Lie group is its tangent space at the identity, i.e., the space of all tangent vectors at the identity (in this case, I_n).

In some sense, the Lie algebra achieves a "linearization" of the Lie group.

The exponential map is a map from the Lie algebra to the Lie group, for example,

$$\exp\colon \mathfrak{so}(n) \to \mathbf{SO}(n)$$

and

$$\exp\colon \mathfrak{sl}(n, \mathbb{R}) \to \mathbf{SL}(n, \mathbb{R}).$$

The exponential map often allows a parametrization of the Lie group elements by simpler objects, the Lie algebra elements.

One might ask, what happened to the Lie algebras $\mathfrak{gl}(n, \mathbb{R})$ and $\mathfrak{o}(n)$ associated with the Lie groups $\mathbf{GL}(n, \mathbb{R})$ and $\mathbf{O}(n)$?

We will see later that $\mathfrak{gl}(n, \mathbb{R})$ is the set of *all* real $n \times n$ matrices, and that $\mathfrak{o}(n) = \mathfrak{so}(n)$.

The properties of the exponential map play an important role in studying a Lie group.

For example, it is clear that the map

$$\exp \colon \mathfrak{gl}(n, \mathbb{R}) \to \mathbf{GL}(n, \mathbb{R})$$

is well-defined, but since every matrix of the form e^A has a positive determinant, exp is not surjective.

Similarly, since
$$\det(e^A) = e^{\mathrm{tr}(A)},$$
the map
$$\exp\colon \mathfrak{sl}(n, \mathbb{R}) \to \mathbf{SL}(n, \mathbb{R})$$
is well-defined. However, we showed in Section 2.1 that it is not surjective either.

As we will see in the next theorem, the map
$$\exp\colon \mathfrak{so}(n) \to \mathbf{SO}(n)$$
is well-defined and surjective.

The map
$$\exp\colon \mathfrak{o}(n) \to \mathbf{O}(n)$$
is well-defined, but it is not surjective, since there are matrices in $\mathbf{O}(n)$ with determinant -1.

Remark: The situation for matrices over the field $\mathbb{C}$ of complex numbers is quite different, as we will see later.

We now show the fundamental relationship between $\mathbf{SO}(n)$ and $\mathfrak{so}(n)$.

Theorem 2.3.2 *The exponential map*

$$\exp\colon \mathfrak{so}(n) \to \mathbf{SO}(n)$$

is well-defined and surjective.

When $n = 3$ (and A is skew symmetric), it is possible to work out an explicit formula for e^A.

For any 3×3 real skew symmetric matrix

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

letting $\theta = \sqrt{a^2 + b^2 + c^2}$ and

$$B = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix},$$

we have the following result known as *Rodrigues's formula* (1840):

Lemma 2.3.3 *The exponential map* exp: $\mathfrak{so}(3) \to \mathbf{SO}(3)$ *is given by*

$$e^A = \cos\theta\, I_3 + \frac{\sin\theta}{\theta}A + \frac{(1-\cos\theta)}{\theta^2}B,$$

or, equivalently, by

$$e^A = I_3 + \frac{\sin\theta}{\theta}A + \frac{(1-\cos\theta)}{\theta^2}A^2$$

if $\theta \neq 0$, *with* $e^{0_3} = I_3$.

The above formulae are the well-known formulae expressing a rotation of axis specified by the vector (a, b, c) and angle θ.

Since the exponential is surjective, it is possible to write down an explicit formula for its inverse (but it is a multivalued function!).

This has applications in kinematics, robotics, and motion interpolation.

2.4 Symmetric Matrices, Symmetric Positive Definite Matrices, and the Exponential Map

Recall that a real symmetric matrix is called *positive* (or *positive semidefinite*) if its eigenvalues are all positive or zero, and *positive definite* if its eigenvalues are all strictly positive.

We denote the vector space of real symmetric $n \times n$ matrices by $\mathbf{S}(n)$, the set of symmetric positive matrices by $\mathbf{SP}(n)$, and the set of symmetric positive definite matrices by $\mathbf{SPD}(n)$.

The next lemma shows that every symmetric positive definite matrix A is of the form e^B for some unique symmetric matrix B.

The set of symmetric matrices is a vector space, but it is not a Lie algebra because the Lie bracket $[A, B]$ is not symmetric unless A and B commute, and the set of symmetric (positive) definite matrices is not a multiplicative group, so this result is of a different flavor as Theorem 2.3.2.

Lemma 2.4.1 *For every symmetric matrix B, the matrix e^B is symmetric positive definite. For every symmetric positive definite matrix A, there is a unique symmetric matrix B such that $A = e^B$.*

Lemma 2.4.1 can be reformulated as stating that the map $\exp\colon \mathbf{S}(n) \to \mathbf{SPD}(n)$ is a bijection.

It can be shown that it is a homeomorphism.

In the case of invertible matrices, the polar form theorem can be reformulated as stating that there is a bijection between the topological space $\mathbf{GL}(n, \mathbb{R})$ of real $n \times n$ invertible matrices (also a group) and $\mathbf{O}(n) \times \mathbf{SPD}(n)$.

As a corollary of the polar form theorem and Lemma 2.4.1, we have the following result:

For every invertible matrix A there is a unique orthogonal matrix R and a unique symmetric matrix S such that

$$A = R\, e^S.$$

Thus, we have a bijection between $\mathbf{GL}(n, \mathbb{R})$ and $\mathbf{O}(n) \times \mathbf{S}(n)$.

But $\mathbf{S}(n)$ itself is isomorphic to $\mathbb{R}^{n(n+1)/2}$. Thus, there is a bijection between $\mathbf{GL}(n, \mathbb{R})$ and $\mathbf{O}(n) \times \mathbb{R}^{n(n+1)/2}$.

It can also be shown that this bijection is a homeomorphism.

This is an interesting fact. Indeed, this homeomorphism essentially reduces the study of the topology of $\mathbf{GL}(n, \mathbb{R})$ to the study of the topology of $\mathbf{O}(n)$.

This is nice, since it can be shown that $\mathbf{O}(n)$ is compact.

In $A = R\, e^S$, if $\det(A) > 0$, then R must be a rotation matrix (i.e., $\det(R) = +1$), since $\det\left(e^S\right) > 0$.

In particular, if $A \in \mathbf{SL}(n, \mathbb{R})$, since $\det(A) = \det(R) = +1$, the symmetric matrix S must have a null trace, i.e., $S \in \mathbf{S}(n) \cap \mathfrak{sl}(n, \mathbb{R})$.

Thus, we have a bijection between $\mathbf{SL}(n, \mathbb{R})$ and $\mathbf{SO}(n) \times (\mathbf{S}(n) \cap \mathfrak{sl}(n, \mathbb{R}))$.

2.5 The Lie Groups $\mathbf{GL}(n, \mathbb{C})$, $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, $\mathbf{SU}(n)$, the Lie Algebras $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{u}(n)$, $\mathfrak{su}(n)$, and the Exponential Map

The set of complex invertible $n \times n$ matrices forms a group under multiplication, denoted by $\mathbf{GL}(n, \mathbb{C})$.

The subset of $\mathbf{GL}(n, \mathbb{C})$ consisting of those matrices having determinant $+1$ is a subgroup of $\mathbf{GL}(n, \mathbb{C})$, denoted by $\mathbf{SL}(n, \mathbb{C})$.

It is also easy to check that the set of complex $n \times n$ unitary matrices forms a group under multiplication, denoted by $\mathbf{U}(n)$.

The subset of $\mathbf{U}(n)$ consisting of those matrices having determinant $+1$ is a subgroup of $\mathbf{U}(n)$, denoted by $\mathbf{SU}(n)$.

We can also check that the set of complex $n \times n$ matrices with null trace forms a real vector space under addition, and similarly for the set of skew Hermitian matrices and the set of skew Hermitian matrices with null trace.

Definition 2.5.1 The group $\mathbf{GL}(n, \mathbb{C})$ is called the *general linear group*, and its subgroup $\mathbf{SL}(n, \mathbb{C})$ is called the *special linear group*. The group $\mathbf{U}(n)$ of unitary matrices is called the *unitary group*, and its subgroup $\mathbf{SU}(n)$ is called the *special unitary group*. The real vector space of complex $n \times n$ matrices with null trace is denoted by $\mathfrak{sl}(n, \mathbb{C})$, the real vector space of skew Hermitian matrices is denoted by $\mathfrak{u}(n)$, and the real vector space $\mathfrak{u}(n) \cap \mathfrak{sl}(n, \mathbb{C})$ is denoted by $\mathfrak{su}(n)$.

Remarks:

(1) As in the real case, the groups $\mathbf{GL}(n, \mathbb{C})$, $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, and $\mathbf{SU}(n)$ are also topological groups (viewed as subspaces of $\mathbb{R}^{2n^2}$), and in fact, smooth real manifolds. Such objects are called *(real) Lie groups*.

The real vector spaces $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{u}(n)$, and $\mathfrak{su}(n)$ are *Lie algebras* associated with $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, and $\mathbf{SU}(n)$.

The algebra structure is given by the *Lie bracket*, which is defined as

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

(2) It is also possible to define complex Lie groups, which means that they are topological groups and smooth *complex* manifolds. It turns out that $\mathbf{GL}(n, \mathbb{C})$ and $\mathbf{SL}(n, \mathbb{C})$ are complex manifolds, but not $\mathbf{U}(n)$ and $\mathbf{SU}(n)$.

One should be very careful to observe that even though the Lie algebras $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{u}(n)$, and $\mathfrak{su}(n)$ consist of matrices with complex coefficients, we view them as *real* vector spaces. The Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ is also a complex vector space, but $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$ are not! Indeed, if A is a skew Hermitian matrix, iA is *not* skew Hermitian, but Hermitian!

Again the Lie algebra achieves a "linearization" of the Lie group. In the complex case, the Lie algebras $\mathfrak{gl}(n, \mathbb{C})$ is the set of *all* complex $n \times n$ matrices, but $\mathfrak{u}(n) \neq \mathfrak{su}(n)$, because a skew Hermitian matrix does not necessarily have a null trace.

The properties of the exponential map also play an important role in studying complex Lie groups.

For example, it is clear that the map

$$\exp\colon \mathfrak{gl}(n, \mathbb{C}) \to \mathbf{GL}(n, \mathbb{C})$$

is well-defined, but this time, it is surjective! One way to prove this is to use the Jordan normal form. Similarly, since

$$\det\left(e^{A}\right) = e^{\mathrm{tr}(A)},$$

the map

$$\exp\colon \mathfrak{sl}(n, \mathbb{C}) \to \mathbf{SL}(n, \mathbb{C})$$

is well-defined, but it is not surjective! As we will see in the next theorem, the maps

$$\exp\colon \mathfrak{u}(n) \to \mathbf{U}(n)$$

and

$$\exp\colon \mathfrak{su}(n) \to \mathbf{SU}(n)$$

are well-defined and surjective.

Theorem 2.5.2 *The exponential maps*

$$\exp\colon \mathfrak{u}(n) \to \mathbf{U}(n) \quad \textit{and} \quad \exp\colon \mathfrak{su}(n) \to \mathbf{SU}(n)$$

are well-defined and surjective.

We now extend the result of Section 2.4 to Hermitian matrices.

2.6 Hermitian Matrices, Hermitian Positive Definite Matrices, and the Exponential Map

Recall that a Hermitian matrix is called *positive* (or *positive semidefinite*) if its eigenvalues are all positive or zero, and *positive definite* if its eigenvalues are all strictly positive.

We denote the real vector space of Hermitian $n \times n$ matrices by $\mathbf{H}(n)$, the set of Hermitian positive matrices by $\mathbf{HP}(n)$, and the set of Hermitian positive definite matrices by $\mathbf{HPD}(n)$.

The next lemma shows that every Hermitian positive definite matrix A is of the form e^B for some unique Hermitian matrix B.

As in the real case, the set of Hermitian matrices is a real vector space, but it is not a Lie algebra because the Lie bracket $[A, B]$ is not Hermitian unless A and B commute, and the set of Hermitian (positive) definite matrices is not a multiplicative group.

Lemma 2.6.1 *For every Hermitian matrix B, the matrix e^B is Hermitian positive definite. For every Hermitian positive definite matrix A, there is a unique Hermitian matrix B such that $A = e^B$.*

Lemma 2.6.1 can be reformulated as stating that the map $\exp\colon \mathbf{H}(n) \to \mathbf{HPD}(n)$ is a bijection. In fact, it can be shown that it is a homeomorphism.

In the case of complex invertible matrices, the polar form theorem can be reformulated as stating that there is a bijection between the topological space $\mathbf{GL}(n, \mathbb{C})$ of complex $n \times n$ invertible matrices (also a group) and $\mathbf{U}(n) \times \mathbf{HPD}(n)$.

As a corollary of the polar form theorem and Lemma 2.6.1, we have the following result: For every complex invertible matrix A, there is a unique unitary matrix U and a unique Hermitian matrix S such that

$$A = U\, e^S.$$

Thus, we have a bijection between $\mathbf{GL}(n, \mathbb{C})$ and $\mathbf{U}(n) \times \mathbf{H}(n)$.

But $\mathbf{H}(n)$ itself is isomorphic to $\mathbb{R}^{n^2}$, and so there is a bijection between $\mathbf{GL}(n, \mathbb{C})$ and $\mathbf{U}(n) \times \mathbb{R}^{n^2}$.

It can also be shown that this bijection is a homeomorphism. This is an interesting fact.

Indeed, this homeomorphism essentially reduces the study of the topology of $\mathbf{GL}(n, \mathbb{C})$ to the study of the topology of $\mathbf{U}(n)$.

This is nice, since it can be shown that $\mathbf{U}(n)$ is compact (as a real manifold).

In the polar decomposition $A = Ue^S$, we have $|\det(U)| = 1$, since U is unitary, and $\mathrm{tr}(S)$ is real, since S is Hermitian (since it is the sum of the eigenvalues of S, which are real), so that $\det\left(e^S\right) > 0$.

Thus, if $\det(A) = 1$, we must have $\det\left(e^S\right) = 1$, which implies that $S \in \mathbf{H}(n) \cap \mathfrak{sl}(n, \mathbb{C})$.

Thus, we have a bijection between $\mathbf{SL}(n, \mathbb{C})$ and $\mathbf{SU}(n) \times (\mathbf{H}(n) \cap \mathfrak{sl}(n, \mathbb{C}))$.

In the next section we study the group $\mathbf{SE}(n)$ of affine maps induced by orthogonal transformations, also called rigid motions, and its Lie algebra.

We will show that the exponential map is surjective. The groups $\mathbf{SE}(2)$ and $\mathbf{SE}(3)$ play play a fundamental role in robotics, dynamics, and motion planning.

2.7 The Lie Group $\mathbf{SE}(n)$ and the Lie Algebra $\mathfrak{se}(n)$

First, we review the usual way of representing affine maps of $\mathbb{R}^n$ in terms of $(n+1) \times (n+1)$ matrices.

Definition 2.7.1 The set of affine maps ρ of $\mathbb{R}^n$, defined such that

$$\rho(X) = RX + U,$$

where R is a rotation matrix ($R \in \mathbf{SO}(n)$) and U is some vector in $\mathbb{R}^n$, is a group under composition called the group of *direct affine isometries, or rigid motions*, denoted by $\mathbf{SE}(n)$.

Every rigid motion can be represented by the $(n+1) \times (n+1)$ matrix

$$\begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix}$$

in the sense that

$$\begin{pmatrix} \rho(X) \\ 1 \end{pmatrix} = \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix}$$

iff

$$\rho(X) = RX + U.$$

Definition 2.7.2 The vector space of real $(n+1)\times(n+1)$ matrices of the form

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix},$$

where Ω is a skew symmetric matrix and U is a vector in $\mathbb{R}^n$, is denoted by $\mathfrak{se}(n)$.

Remark: The group $\mathbf{SE}(n)$ is a Lie group, and its Lie algebra turns out to be $\mathfrak{se}(n)$.

We will show that the exponential map $\exp\colon \mathfrak{se}(n) \to \mathbf{SE}(n)$ is surjective. First, we prove the following key lemma.

Lemma 2.7.3 *Given any $(n+1) \times (n+1)$ matrix of the form*

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix}$$

where Ω is any matrix and $U \in \mathbb{R}^n$,

$$A^k = \begin{pmatrix} \Omega^k & \Omega^{k-1}U \\ 0 & 0 \end{pmatrix},$$

where $\Omega^0 = I_n$. As a consequence,

$$e^A = \begin{pmatrix} e^\Omega & VU \\ 0 & 1 \end{pmatrix},$$

where

$$V = I_n + \sum_{k \geq 1} \frac{\Omega^k}{(k+1)!}.$$

We can now prove our main theorem.

We will need to prove that V is invertible when Ω is a skew symmetric matrix.

It would be tempting to write V as

$$V = \Omega^{-1}(e^{\Omega} - I).$$

Unfortunately, for odd n, a skew symmetric matrix of order n is not invertible! Thus, we have to find another way of proving that V is invertible.

Remark: We have

$$V = I_n + \sum_{k\geq 1} \frac{\Omega^k}{(k+1)!} = \int_0^1 e^{\Omega t} dt.$$

This can be used to give a more explicit formula for V if we have an explicit formula for $e^{\Omega t}$ (see below for $n = 3$).

Theorem 2.7.4 *The exponential map*

$$\exp\colon \mathfrak{se}(n) \to \mathbf{SE}(n)$$

is well-defined and surjective.

In the case $n = 3$, given a skew symmetric matrix

$$\Omega = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

letting $\theta = \sqrt{a^2 + b^2 + c^2}$, it it easy to prove that if $\theta = 0$, then

$$e^A = \begin{pmatrix} I_3 & U \\ 0 & 1 \end{pmatrix},$$

and that if $\theta \neq 0$ (using the fact that $\Omega^3 = -\theta^2 \Omega$), then

$$e^\Omega = I_3 + \frac{\sin\theta}{\theta}\Omega + \frac{(1-\cos\theta)}{\theta^2}\Omega^2$$

and

$$V = I_3 + \frac{(1-\cos\theta)}{\theta^2}\Omega + \frac{(\theta-\sin\theta)}{\theta^3}\Omega^2.$$

We finally reach the best vista point of our hike, the formal definition of (linear) Lie groups and Lie algebras.