14

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**

14.1 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 maps 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. On the way to Lie algebras, we derive the classical "Rodrigues-like" formulae for rotations and for rigid motions in $\mathbb{R}^2$ and $\mathbb{R}^3$. We give an elementary proof that the exponential map is surjective for both $\mathbf{SO}(n)$ and $\mathbf{SE}(n)$, not using any topology, just our normal forms for matrices.

The last section gives a quick introduction to Lie groups and Lie algebras. 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. This section is inspired from Artin [5], Chevalley [31], Marsden and Ratiu [120], Curtis [38], Howe [91], and Sattinger and Weaver [147].

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 problem is, Why is it well-defined? The following lemma shows that the above series is indeed absolutely convergent.

Lemma 14.1.1 *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

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

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.

Proof. The proof is by induction on p. For $p = 0$, we have $A^0 = I_n$, $(n\mu)^0 = 1$, and the lemma is obvious. Assume that

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

for all i, j, $1 \leq i, j \leq n$. Then we have

$$\left|a^{p+1}_{i\,j}\right| = \left|\sum_{k=1}^{n} a^p_{i\,k} a_{k\,j}\right| \leq \sum_{k=1}^{n} \left|a^p_{i\,k}\right| \left|a_{k\,j}\right| \leq \mu \sum_{k=1}^{n} \left|a^p_{i\,k}\right| \leq n\mu(n\mu)^p = (n\mu)^{p+1},$$

for all i, j, $1 \leq i, j \leq n$. For every pair (i, j) such that $1 \leq i, j \leq n$, since

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

the series

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

is bounded by the convergent series

$$e^{n\mu} = \sum_{p\geq 0} \frac{(n\mu)^p}{p!},$$

and thus it is absolutely convergent. This shows that

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

is well defined. $\square$

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 need to find an inductive formula expressing the powers A^n. Let us observe that

$$\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}^2 = -\theta^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then, letting

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

we have

$$\begin{aligned} A^{4n} &= \theta^{4n} I_2, \\ A^{4n+1} &= \theta^{4n+1} J, \end{aligned}$$

$$\begin{aligned} A^{4n+2} &= -\theta^{4n+2} I_2, \\ A^{4n+3} &= -\theta^{4n+3} J, \end{aligned}$$

and so

$$e^A = I_2 + \frac{\theta}{1!}J - \frac{\theta^2}{2!}I_2 - \frac{\theta^3}{3!}J + \frac{\theta^4}{4!}I_2 + \frac{\theta^5}{5!}J - \frac{\theta^6}{6!}I_2 - \frac{\theta^7}{7!}J + \cdots.$$

Rearranging the order of the terms, we have

$$e^A = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right) I_2 + \left(\frac{\theta}{1!} - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right) J.$$

We recognize the power series for $\cos\theta$ and $\sin\theta$, and thus

$$e^A = \cos\theta I_2 + \sin\theta J,$$

that is

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

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, $A^2 = -\omega^2 I_2$. We get

$$e^A = I_2 + \frac{A}{1!} - \frac{\omega^2}{2!}I_2 - \frac{\omega^2}{3!}A + \frac{\omega^4}{4!}I_2 + \frac{\omega^4}{5!}A - \frac{\omega^6}{6!}I_2 - \frac{\omega^6}{7!}A + \cdots.$$

Rearranging the order of the terms, we have

$$e^A = \left(1 - \frac{\omega^2}{2!} + \frac{\omega^4}{4!} - \frac{\omega^6}{6!} + \cdots\right) I_2 + \frac{1}{\omega}\left(\omega - \frac{\omega^3}{3!} + \frac{\omega^5}{5!} - \frac{\omega^7}{7!} + \cdots\right) A.$$

We recognize the power series for $\cos\omega$ and $\sin\omega$, and thus

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

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

$$e^A = I_2 + \frac{A}{1!} + \frac{\omega^2}{2!} I_2 + \frac{\omega^2}{3!} A + \frac{\omega^4}{4!} I_2 + \frac{\omega^4}{5!} A + \frac{\omega^6}{6!} I_2 + \frac{\omega^6}{7!} A + \cdots.$$

Rearranging the order of the terms, we have

$$e^A = \left(1 + \frac{\omega^2}{2!} + \frac{\omega^4}{4!} + \frac{\omega^6}{6!} + \cdots\right) I_2 + \frac{1}{\omega}\left(\omega + \frac{\omega^3}{3!} + \frac{\omega^5}{5!} + \frac{\omega^7}{7!} + \cdots\right) A.$$

If we recall that $\cosh\omega = \left(e^\omega + e^{-\omega}\right)/2$ and $\sinh\omega = \left(e^\omega - e^{-\omega}\right)/2$, we recognize the power series for $\cosh\omega$ and $\sinh\omega$, and thus

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

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,

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

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

$$e^{UAU^{-1}} = Ue^AU^{-1}.$$

Proof. A trivial induction shows that

$$UA^pU^{-1} = (UAU^{-1})^p,$$

and thus

$$\begin{aligned} e^{UAU^{-1}} &= \sum_{p\geq 0} \frac{(UAU^{-1})^p}{p!} = \sum_{p\geq 0} \frac{UA^pU^{-1}}{p!} \\ &= U\left(\sum_{p\geq 0} \frac{A^p}{p!}\right) U^{-1} = Ue^AU^{-1}. \end{aligned}$$

□

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

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

Proof. We prove by induction on n that if $f\colon \mathbb{C}^n \to \mathbb{C}^n$ is a linear map, then there is a basis $(u_1, \ldots, u_n)$ with respect to which f is represented by an upper triangular matrix. For $n = 1$ the result is obvious. If $n > 1$, since $\mathbb{C}$ is algebraically closed, f has some eigenvalue $\lambda_1 \in \mathbb{C}$, and let u_1 be an eigenvector for λ_1. We can find $n-1$ vectors $(v_2, \ldots, v_n)$ such that $(u_1, v_2, \ldots, v_n)$ is a basis of $\mathbb{C}^n$, and let W be the subspace of dimension $n-1$ spanned by $(v_2, \ldots, v_n)$. In the basis $(u_1, v_2 \ldots, v_n)$, the matrix of f is of the form

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

since its first column contains the coordinates of $\lambda_1 u_1$ over the basis $(u_1, v_2, \ldots, v_n)$. Letting $p\colon \mathbb{C}^n \to W$ be the projection defined such that $p(u_1) = 0$ and $p(v_i) = v_i$ when $2 \leq i \leq n$, the linear map $g\colon W \to W$ defined as the restriction of $p \circ f$ to W is represented by the $(n-1) \times (n-1)$ matrix $(a_{i\,j})_{2\leq i,j\leq n}$ over the basis $(v_2, \ldots, v_n)$. By the induction hypothesis, there is a basis $(u_2, \ldots, u_n)$ of W such that g is represented by an upper triangular matrix $(b_{i\,j})_{1\leq i,j\leq n-1}$.

However,

$$\mathbb{C}^n = \mathbb{C}u_1 \oplus W,$$

and thus $(u_1, \ldots, u_n)$ is a basis for $\mathbb{C}^n$. Since p is the projection from $\mathbb{C}^n = \mathbb{C}u_1 \oplus W$ onto W and $g\colon W \to W$ is the restriction of $p \circ f$ to W, we have

$$f(u_1) = \lambda_1 u_1$$

and

$$f(u_{i+1}) = a_{1\,i} u_1 + \sum_{j=1}^{n-1} b_{i\,j} u_{j+1}$$

for some $a_{1\,i} \in \mathbb{C}$, when $1 \leq i \leq n-1$. But then the matrix of f with respect to $(u_1, \ldots, u_n)$ is upper triangular. Thus, there is a change of basis matrix P such that $A = PTP^{-1}$ where T is upper triangular. $\square$

Remark: If E is a Hermitian space, the proof of Lemma 14.1.3 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

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

is $(\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$, and thus the eigenvalues of $A = PTP^{-1}$ are the diagonal entries of T. We use this property to prove the following lemma.

Lemma 14.1.4 *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}.*

Proof. By Lemma 14.1.3 there is an invertible matrix P and an upper triangular matrix T such that

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

By Lemma 14.1.2,

$$e^{PTP^{-1}} = Pe^T P^{-1}.$$

However, we showed that A and T have the same eigenvalues, which are the diagonal entries $\lambda_1, \ldots, \lambda_n$ of T, and $e^A = e^{PTP^{-1}} = Pe^T P^{-1}$ and e^T have the same eigenvalues, which are the diagonal entries of e^T. Clearly, the diagonal entries of e^T are $e^{\lambda_1}, \ldots, e^{\lambda_n}$. Now, if u is an eigenvector of A for the eigenvalue λ, a simple induction shows that u is an eigenvector of A^n for the eigenvalue λ^n, from which is follows that u is an eigenvector of e^A for e^λ. □

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 14.1.4, $e^{\lambda_1}, \ldots, e^{\lambda_n}$ 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 null 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., $AB = BA$. We need to prove this last fact.

Lemma 14.1.5 *Given any two complex $n \times n$ matrices A, B, if $AB = BA$, then*

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

Proof. Since $AB = BA$, we can expand $(A + B)^p$ using the binomial formula:

$$(A + B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{p-k},$$

and thus

$$\frac{1}{p!}(A+B)^p = \sum_{k=0}^{p} \frac{A^k B^{p-k}}{k!(p-k)!}.$$

Note that for any integer $N \geq 0$, we can write

$$\begin{aligned}\sum_{p=0}^{2N} \frac{1}{p!}(A+B)^p &= \sum_{p=0}^{2N}\sum_{k=0}^{p} \frac{A^k B^{p-k}}{k!(p-k)!}\\ &= \left(\sum_{p=0}^{N} \frac{A^p}{p!}\right)\left(\sum_{p=0}^{N} \frac{B^p}{p!}\right) + \sum_{\substack{\max(k,l) > N\\ k+l \leq 2N}} \frac{A^k}{k!}\frac{B^l}{l!},\end{aligned}$$

where there are $N(N+1)$ pairs (k,l) in the second term. Letting

$$\|A\| = \max\{|a_{i\,j}| \mid 1 \leq i,j \leq n\}, \quad \|B\| = \max\{|b_{i\,j}| \mid 1 \leq i,j \leq n\},$$

and $\mu = \max(\|A\|, \|B\|)$, note that for every entry $c_{i\,j}$ in $\left(A^k/k!\right)\left(B^l/l!\right)$ we have

$$|c_{i\,j}| \leq n\frac{(n\mu)^k}{k!}\frac{(n\mu)^l}{l!} \leq \frac{(n^2\mu)^{2N}}{N!}.$$

As a consequence, the absolute value of every entry in

$$\sum_{\substack{\max(k,l) > N\\ k+l \leq 2N}} \frac{A^k}{k!}\frac{B^l}{l!}$$

is bounded by

$$N(N+1)\frac{(n^2\mu)^{2N}}{N!},$$

which goes to 0 as $N \mapsto \infty$. From this, it immediately follows that

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

□

Now, using Lemma 14.1.5, 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.

14.2 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

First, we recall some basic facts and definitions. 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 14.2.1 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.[1] Such objects are called *Lie groups*. 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

[1] We refrain from defining manifolds right now, not to interupt the flow of intuitive ideas.

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 14.1 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 14.2.2 *The exponential map*

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

is well-defined and surjective.

Proof. First, we need to prove that if A is a skew symmetric matrix, then e^A is a rotation matrix. For this, first check that

$$\left(e^A\right)^\top = e^{A^\top}.$$

Then, since $A^\top = -A$, we get

$$\left(e^A\right)^\top = e^{A^\top} = e^{-A},$$

and so

$$\left(e^A\right)^\top e^A = e^{-A}e^A = e^{-A+A} = e^{0_n} = I_n,$$

and similarly,

$$e^A \left(e^A\right)^\top = I_n,$$

showing that e^A is orthogonal. Also,

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

and since A is real skew symmetric, its diagonal entries are 0, i.e., $\operatorname{tr}(A) = 0$, and so $\det(e^A) = +1$.

For the surjectivity, we will use Theorem 11.4.4 and Theorem 11.4.5. Theorem 11.4.4 says that for every skew symmetric matrix A there is an orthogonal matrix P such that $A = PD\,P^\top$, where D is a block diagonal matrix 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

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

where $\theta_i \in \mathbb{R}$, with $\theta_i > 0$. Theorem 11.4.5 says that for every orthogonal matrix R there is an orthogonal matrix P such that $R = PE\,P^\top$, where E is a block diagonal matrix of the form

$$E = \begin{pmatrix} E_1 & & \dots & \\ & E_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & E_p \end{pmatrix}$$

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

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

If R is a rotation matrix, there is an even number of -1's and they can be grouped into blocks of size 2 associated with $\theta = \pi$. Let D be the block matrix associated with E in the obvious way (where an entry 1 in E is associated with a 0 in D). Since by Lemma 14.1.2

$$e^A = e^{PDP^{-1}} = Pe^DP^{-1},$$

and since D is a block diagonal matrix, we can compute e^D by computing the exponentials of its blocks. If $D_i = 0$, we get $E_i = e^0 = +1$, and if

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

we showed earlier that

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

exactly the block E_i. Thus, $E = e^D$, and as a consequence,

$$e^A = e^{PDP^{-1}} = Pe^D P^{-1} = PEP^{-1} = PE\,P^\top = R.$$

This shows the surjectivity of the exponential. $\square$

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 14.2.3 *The exponential map* $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ *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* $e^{0_3} = I_3$.

Proof sketch. First, prove that

$$A^2 = -\theta^2 I + B,$$
$$AB = BA = 0.$$

From the above, deduce that

$$A^3 = -\theta^2 A,$$

and for any $k \geq 0$,

$$\begin{aligned} A^{4k+1} &= \theta^{4k} A, \\ A^{4k+2} &= \theta^{4k} A^2, \\ A^{4k+3} &= -\theta^{4k+2} A, \\ A^{4k+4} &= -\theta^{4k+2} A^2. \end{aligned}$$

Then prove the desired result by writing the power series for e^A and regrouping terms so that the power series for cos and sin show up. $\square$

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.

14.3 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 null, 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 14.2.2.

Lemma 14.3.1 *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 $A = e^B$.*

Proof. We showed earlier that

$$\left(e^B\right)^\top = e^{B^\top}.$$

If B is a symmetric matrix, then since $B^\top = B$, we get

$$\left(e^B\right)^\top = e^{B^\top} = e^B,$$

and e^B is also symmetric. Since the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the symmetric matrix B are real and the eigenvalues of e^B are $e^{\lambda_1}, \ldots, e^{\lambda_n}$, and since $e^\lambda > 0$ if $\lambda \in \mathbb{R}$, e^B is positive definite.

If A is symmetric positive definite, by Theorem 11.4.3 there is an orthogonal matrix P such that $A = PD\,P^\top$, where D is a diagonal matrix

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

where $\lambda_i > 0$, since A is positive definite. Letting

$$L = \begin{pmatrix} \log\lambda_1 & & \dots & \\ & \log\lambda_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & \log\lambda_n \end{pmatrix},$$

it is obvious that $e^L = D$, with $\log\lambda_i \in \mathbb{R}$, since $\lambda_i > 0$.

Let

$$B = PL\,P^\top.$$

By Lemma 14.1.2, we have

$$e^B = e^{PL\,P^\top} = e^{PLP^{-1}} = Pe^L P^{-1} = Pe^L\,P^\top = PD\,P^\top = A.$$

Finally, we prove that if B_1 and B_2 are symmetric and $A = e^{B_1} = e^{B_2}$, then $B_1 = B_2$. Since B_1 is symmetric, there is an orthonormal basis $(u_1, \dots, u_n)$ of eigenvectors of B_1. Let $\mu_1, \dots, \mu_n$ be the corresponding eigenvalues. Similarly, there is an orthonormal basis $(v_1, \dots, v_n)$ of eigenvectors of B_2. We are going to prove that B_1 and B_2 agree on the basis $(v_1, \dots, v_n)$, thus proving that $B_1 = B_2$.

Let μ be some eigenvalue of B_2, and let $v = v_i$ be some eigenvector of B_2 associated with μ. We can write

$$v = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Since v is an eigenvector of B_2 for μ and $A = e^{B_2}$, by Lemma 14.1.4

$$A(v) = e^\mu v = e^\mu \alpha_1 u_1 + \cdots + e^\mu \alpha_n u_n.$$

On the other hand,

$$A(v) = A(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_1 A(u_1) + \cdots + \alpha_n A(u_n),$$

and since $A = e^{B_1}$ and $B_1(u_i) = \mu_i u_i$, by Lemma 14.1.4 we get

$$A(v) = e^{\mu_1}\alpha_1 u_1 + \cdots + e^{\mu_n}\alpha_n u_n.$$

Therefore, $\alpha_i = 0$ if $\mu_i \neq \mu$. Letting

$$I = \{i \mid \mu_i = \mu,\ i \in \{1, \dots, n\}\},$$

we have

$$v = \sum_{i \in I} \alpha_i u_i.$$

Now,

$$\begin{aligned} B_1(v) = B_1\left(\sum_{i\in I}\alpha_i u_i\right) &= \sum_{i\in I}\alpha_i B_1(u_i) = \sum_{i\in I}\alpha_i\mu_i u_i \\ &= \sum_{i\in I}\alpha_i\mu u_i = \mu\left(\sum_{i\in I}\alpha_i u_i\right) = \mu v, \end{aligned}$$

since $\mu_i = \mu$ when $i \in I$. Since v is an eigenvector of B_2 for μ,

$$B_2(v) = \mu v,$$

which shows that

$$B_1(v) = B_2(v).$$

Since the above holds for every eigenvector v_i, we have $B_1 = B_2$. $\square$

Lemma 14.3.1 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 (Theorem 12.1.3) and Lemma 14.3.1, 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 $\mathbf{SO}(n) \times (\mathbf{S}(n) \cap \mathfrak{sl}(n, \mathbb{R}))$.

We can also use the results of Section 11.4 to show that the exponential map is a surjective map from the skew Hermitian matrices to the unitary matrices.

14.4 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 14.4.1 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 14.4.2 *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.

Proof. First, we need to prove that if A is a skew Hermitian matrix, then e^A is a unitary matrix. For this, first check that

$$\left(e^A\right)^* = e^{A^*}.$$

Then, since $A^* = -A$, we get

$$\left(e^A\right)^* = e^{A^*} = e^{-A},$$

and so

$$\left(e^A\right)^* e^A = e^{-A}e^A = e^{-A+A} = e^{0_n} = I_n,$$

and similarly, $e^A\left(e^A\right)^* = I_n$, showing that e^A is unitary. Since

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

if A is skew Hermitian and has null trace, then $\det(e^A) = +1$.

For the surjectivity we will use Theorem 11.4.7. First, assume that A is a unitary matrix. By Theorem 11.4.7, there is a unitary matrix U and a diagonal matrix D such that $A = UDU^*$. Furthermore, since A is unitary, the entries $\lambda_1, \ldots, \lambda_n$ in D (the eigenvalues of A) have absolute value $+1$. Thus, the entries in D are of the form $\cos\theta + i\sin\theta = e^{i\theta}$. Thus, we can assume that D is a diagonal matrix of the form

$$D = \begin{pmatrix} e^{i\theta_1} & & \ldots & \\ & e^{i\theta_2} & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & e^{i\theta_p} \end{pmatrix}.$$

If we let E be the diagonal matrix

$$E = \begin{pmatrix} i\theta_1 & & \ldots & \\ & i\theta_2 & \ldots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \ldots & i\theta_p \end{pmatrix}$$

it is obvious that E is skew Hermitian and that

$$e^E = D.$$

Then, letting $B = UEU^*$, we have

$$e^B = A,$$

and it is immediately verified that B is skew Hermitian, since E is.

If A is a unitary matrix with determinant $+1$, since the eigenvalues of A are $e^{i\theta_1}, \ldots, e^{i\theta_p}$ and the determinant of A is the product

$$e^{i\theta_1} \cdots e^{i\theta_p} = e^{i(\theta_1 + \cdots + \theta_p)}$$

of these eigenvalues, we must have

$$\theta_1 + \cdots + \theta_p = 0,$$

and so, E is skew Hermitian and has zero trace. As above, letting

$$B = UEU^*,$$

we have

$$e^B = A,$$

where B is skew Hermitian and has null trace. $\square$

We now extend the result of Section 14.3 to Hermitian matrices.

14.5 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 null, 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 14.5.1 *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 $A = e^B$.*

Proof. It is basically the same as the proof of Theorem 14.5.1, except that a Hermitian matrix can be written as $A = UDU^*$, where D is a real diagonal matrix and U is unitary instead of orthogonal. $\square$

Lemma 14.5.1 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 14.5.1, 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.

14.6 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 14.6.1 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 14.6.2 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 14.6.3 *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

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

Proof. A trivial induction on k shows that

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

Then we have

$$\begin{aligned} e^A &= \sum_{k \geq 0} \frac{A^k}{k!}, \\ &= I_{n+1} + \sum_{k \geq 1} \frac{1}{k!} \begin{pmatrix} \Omega^k & \Omega^{k-1}U \\ 0 & 0 \end{pmatrix}, \\ &= \begin{pmatrix} I_n + \sum_{k \geq 0} \frac{\Omega^k}{k!} & \sum_{k \geq 1} \frac{\Omega^{k-1}}{k!} U \\ 0 & 1 \end{pmatrix}, \\ &= \begin{pmatrix} e^{\Omega} & VU \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

□

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. However, observe that we have the following useful fact:

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

This is what we will use in Theorem 14.6.4 to prove surjectivity.

Theorem 14.6.4 *The exponential map*

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

is well-defined and surjective.

Proof. Since Ω is skew symmetric, e^{Ω} is a rotation matrix, and by Theorem 14.2.2, the exponential map

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

is surjective. Thus, it remains to prove that for every rotation matrix R, there is some skew symmetric matrix Ω such that $R = e^{\Omega}$ and

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

is invertible. By Theorem 11.4.4, for every skew symmetric matrix Ω there is an orthogonal matrix P such that $\Omega = PD\,P^{\top}$, where D is a block diagonal matrix 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

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

where $\theta_i \in \mathbb{R}$, with $\theta_i > 0$. Actually, we can assume that $\theta_i \neq k2\pi$ for all $k \in \mathbb{Z}$, since when $\theta_i = k2\pi$ we have $e^{D_i} = I_2$, and D_i can be replaced by two one-dimensional blocks each consisting of a single zero. To compute V, since $\Omega = PD\,P^{\top} = PDP^{-1}$, observe that

$$\begin{aligned} V &= I_n + \sum_{k\geq 1} \frac{\Omega^k}{(k+1)!} \\ &= I_n + \sum_{k\geq 1} \frac{PD^kP^{-1}}{(k+1)!} \\ &= P\left(I_n + \sum_{k\geq 1} \frac{D^k}{(k+1)!}\right)P^{-1} \\ &= PWP^{-1}, \end{aligned}$$

where

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

We can compute

$$W = I_n + \sum_{k\geq 1} \frac{D^k}{(k+1)!} = \int_0^1 e^{Dt}dt,$$

by computing

$$W = \begin{pmatrix} W_1 & & \dots & \\ & W_2 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \dots & W_p \end{pmatrix}$$

by blocks. Since

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

when D_i is a 2×2 skew symmetric matrix and $W_i = \int_0^1 e^{D_i t} dt$, we get

$$W_i = \begin{pmatrix} \int_0^1 \cos(\theta_i t) dt & \int_0^1 - \sin(\theta_i t) dt \\ \int_0^1 \sin(\theta_i t) dt & \int_0^1 \cos(\theta_i t) dt \end{pmatrix} = \frac{1}{\theta_i} \begin{pmatrix} \sin(\theta_i t) \mid_0^1 & \cos(\theta_i t) \mid_0^1 \\ -\cos(\theta_i t) \mid_0^1 & \sin(\theta_i t) \mid_0^1 \end{pmatrix},$$

that is,

$$W_i = \frac{1}{\theta_i} \begin{pmatrix} \sin\theta_i & -(1 - \cos\theta_i) \\ 1 - \cos\theta_i & \sin\theta_i \end{pmatrix},$$

and $W_i = 1$ when $D_i = 0$. Now, in the first case, the determinant is

$$\frac{1}{\theta_i^2} \left((\sin\theta_i)^2 + (1 - \cos\theta_i)^2 \right) = \frac{2}{\theta_i^2} (1 - \cos\theta_i),$$

which is nonzero, since $\theta_i \neq k2\pi$ for all $k \in \mathbb{Z}$. Thus, each W_i is invertible, and so is W, and thus, $V = PWP^{-1}$ is invertible. $\square$

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.

14.7 Finale: Lie Groups and Lie Algebras

In this section we attempt to define precisely Lie groups and Lie algebras. One of the reasons that Lie groups are nice is that they have a differential structure, which means that the notion of tangent space makes sense at any point of the group. Furthermore, the tangent space at the identity happens to have some algebraic structure, that of a Lie algebra. Roughly, the tangent space at the identity provides a "linearization" of the Lie group, and it turns out that many properties of a Lie group are reflected in its Lie algebra, and that the loss of information is not too severe. The challenge that we are facing is that unless our readers are already familiar with manifolds, the amount of basic differential geometry required to define Lie groups and Lie algebras in full generality is overwhelming.

Fortunately, all the Lie groups that we need to consider are subspaces of $\mathbb{R}^N$ for some sufficiently large N. In fact, they are all isomorphic to subgroups of $\mathbf{GL}(N, \mathbb{R})$ for some suitable N, even $\mathbf{SE}(n)$, which is isomorphic to a subgroup of $\mathbf{SL}(n+1)$. Such groups are called *linear Lie groups* (or *matrix groups*). Since the groups under consideration are subspaces of $\mathbb{R}^N$, we do not need the definition of an abstract manifold. We just have to define embedded submanifolds (also called submanifolds) of $\mathbb{R}^N$ (in the case of $\mathbf{GL}(n, \mathbb{R})$, $N = n^2$). This is the path that we will follow.

In general, the difficult part in proving that a subgroup of $\mathbf{GL}(n, \mathbb{R})$ is a Lie group is to prove that it is a manifold. Fortunately, there is a characterization of the linear groups that obviates much of the work. This characterization rests on two theorems. First, a Lie subgroup H of a Lie group G (where H is an embedded submanifold of G) is closed in G (see Warner [176], Chapter 3, Theorem 3.21, page 97). Second, a theorem of Von Neumann and Cartan asserts that a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$ is an embedded submanifold, and thus, a Lie group (see Warner [176], Chapter 3, Theorem 3.42, page 110). Thus, a linear Lie group is a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$.

Since our Lie groups are subgroups (or isomorphic to subgroups) of $\mathbf{GL}(n, \mathbb{R})$ for some suitable n, it is easy to define the Lie algebra of a Lie group using curves. This approach to define the Lie algebra of a matrix group is followed by a number of authors, such as Curtis [38]. However, Curtis is rather cavalier, since he does not explain why the required curves actually exist, and thus, according to his definition, Lie algebras could be the trivial vector space! Although we will not prove the theorem of Von Neumann and Cartan, we feel that it is important to make clear why the definitions make sense, i.e., why we are not dealing with trivial objects.

A small annoying technical problem will arise in our approach, the problem with discrete subgroups. If A is a subset of $\mathbb{R}^N$, recall that A inherits a topology from $\mathbb{R}^N$ called the *subspace topology*, and defined such that a subset V of A is open if

$$V = A \cap U$$

for some open subset U of $\mathbb{R}^N$. A point $a \in A$ is said to be *isolated* if there is there is some open subset U of $\mathbb{R}^N$ such that

$$\{a\} = A \cap U,$$

in other words, if $\{a\}$ is an open set in A.

The group $\mathbf{GL}(n, \mathbb{R})$ of real invertible $n \times n$ matrices can be viewed as a subset of $\mathbb{R}^{n^2}$, and as such, it is a topological space under the subspace topology (in fact, a dense open subset of $\mathbb{R}^{n^2}$). One can easily check that multiplication and the inverse operation are continuous, and in fact smooth (i.e., C^∞-continuously differentiable). This makes $\mathbf{GL}(n, \mathbb{R})$ a *topological group*. Any subgroup G of $\mathbf{GL}(n, \mathbb{R})$ is also a topological space under the subspace topology. A subgroup G is called a *discrete subgroup* if it has some isolated point. This turns out to be equivalent to the fact that every point of G is isolated, and thus, G has the discrete topology (every subset of G is open). Now, because $\mathbf{GL}(n, \mathbb{R})$ is Hausdorff, it can be shown that every discrete subgroup of $\mathbf{GL}(n, \mathbb{R})$ is closed (which means that its complement is open). Thus, discrete subgroups of $\mathbf{GL}(n, \mathbb{R})$ are Lie groups! But these are not very interesting Lie groups, and so we will consider only closed subgroups of $\mathbf{GL}(n, \mathbb{R})$ that are not discrete.

Let us now review the definition of an embedded submanifold. For simplicity, we restrict our attention to smooth manifolds. For detailed presentations, see DoCarmo [51, 52], Milnor [127], Marsden and Ratiu [120], Berger and Gostiaux [14], or Warner [176]. For the sake of brevity, we use the terminology *manifold* (but other authors would say *embedded submanifolds*, or something like that).

The intuition behind the notion of a smooth manifold in $\mathbb{R}^N$ is that a subspace M is a manifold of dimension m if every point $p \in M$ is contained in some open subset set U of M (in the subspace topology) that can be parametrized by some function $\varphi\colon \Omega \to U$ from some open subset Ω of the origin in $\mathbb{R}^m$, and that φ has some nice properties that allow the definition of smooth functions on M and of the tangent space at p. For this, φ has to be at least a homeomorphism, but more is needed: φ must be smooth, and the derivative $\varphi'(0_m)$ at the origin must be injective (letting $0_m = \underbrace{(0, \ldots, 0)}_{m}$).

Definition 14.7.1 Given any integers N, m, with $N \geq m \geq 1$, an *m-dimensional smooth manifold in $\mathbb{R}^N$, for short a manifold*, is a nonempty subset M of $\mathbb{R}^N$ such that for every point $p \in M$ there are two open subsets $\Omega \subseteq \mathbb{R}^m$ and $U \subseteq M$, with $p \in U$, and a smooth function $\varphi\colon \Omega \to \mathbb{R}^N$ such that φ is a homeomorphism between Ω and $U = \varphi(\Omega)$, and $\varphi'(t_0)$ is injective, where $t_0 = \varphi^{-1}(p)$. The function $\varphi\colon \Omega \to U$ is called a *(local)* *parametrization of* M *at* p. If $0_m \in \Omega$ and $\varphi(0_m) = p$, we say that $\varphi\colon \Omega \to U$ is *centered at* p.

Recall that $M \subseteq \mathbb{R}^N$ is a topological space under the subspace topology, and U is some open subset of M in the subspace topology, which means that $U = M \cap W$ for some open subset W of $\mathbb{R}^N$. Since $\varphi\colon \Omega \to U$ is a homeomorphism, it has an inverse $\varphi^{-1}\colon U \to \Omega$ that is also a homeomorphism, called a *(local) chart*. Since $\Omega \subseteq \mathbb{R}^m$, for every point $p \in M$ and every parametrization $\varphi\colon \Omega \to U$ of M at p, we have $\varphi^{-1}(p) = (z_1, \ldots, z_m)$ for some $z_i \in \mathbb{R}$, and we call $z_1, \ldots, z_m$ the *local coordinates of* p *(w.r.t.* φ^{-1}*)*. We often refer to a manifold M without explicitly specifying its dimension (the integer m).

Intuitively, a chart provides a "flattened" local map of a region on a manifold. For instance, in the case of surfaces (2-dimensional manifolds), a chart is analogous to a planar map of a region on the surface. For a concrete example, consider a map giving a planar representation of a country, a region on the earth, a curved surface.

Remark: We could allow $m = 0$ in definition 14.7.1. If so, a manifold of dimension 0 is just a set of isolated points, and thus it has the discrete topology. In fact, it can be shown that a discrete subset of $\mathbb{R}^N$ is countable. Such manifolds are not very exciting, but they do correspond to discrete subgroups.

Example 14.1 The unit sphere S^2 in $\mathbb{R}^3$ defined such that

$$S^2 = \left\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\right\}$$

is a smooth 2-manifold, because it can be parametrized using the following two maps φ_1 and φ_2:

$$\varphi_1\colon (u, v) \mapsto \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2+1}\right)$$

and

$$\varphi_2\colon (u, v) \mapsto \left(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{1-u^2-v^2}{u^2+v^2+1}\right).$$

The map φ_1 corresponds to the inverse of the stereographic projection from the north pole $N = (0, 0, 1)$ onto the plane $z = 0$, and the map φ_2 corresponds to the inverse of the stereographic projection from the south pole $S = (0, 0, -1)$ onto the plane $z = 0$, as illustrated in Figure 14.1. We leave as an exercise to check that the map φ_1 parametrizes $S^2 - \{N\}$ and that the map φ_2 parametrizes $S^2 - \{S\}$ (and that they are smooth, homeomorphisms, etc.). Using φ_1, the open lower hemisphere is parametrized by the open disk of center O and radius 1 contained in the plane $z = 0$.

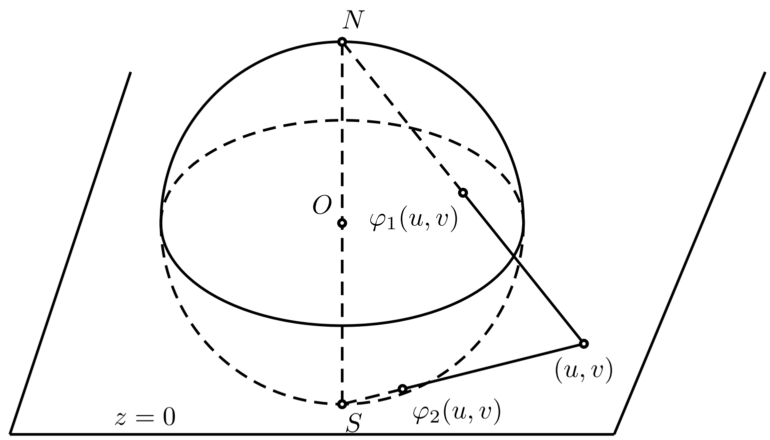

Figure 14.1. Inverse stereographic projections

The chart φ_1^{-1} assigns local coordinates to the points in the open lower hemisphere. If we draw a grid of coordinate lines parallel to the x and y axes inside the open unit disk and map these lines onto the lower hemisphere using φ_1, we get curved lines on the lower hemisphere. These "coordinate lines" on the lower hemisphere provide local coordinates for every point on the lower hemisphere. For this reason, older books often talk about *curvilinear coordinate systems* to mean the coordinate lines on a surface induced by a chart. We urge our readers to define a manifold structure on a torus. This can be done using four charts.

Every open subset of $\mathbb{R}^N$ is a manifold in a trivial way. Indeed, we can use the inclusion map as a parametrization. In particular, $\mathbf{GL}(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, since its complement is closed (the set of invertible matrices is the inverse image of the determinant function, which is continuous). Thus, $\mathbf{GL}(n, \mathbb{R})$ is a manifold. We can view $\mathbf{GL}(n, \mathbb{C})$ as a subset of $\mathbb{R}^{(2n)^2}$ using the embedding defined as follows: For every complex $n \times n$ matrix A, construct the real $2n \times 2n$ matrix such that every entry $a + ib$ in A is replaced by the 2×2 block

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

where $a, b \in \mathbb{R}$. It is immediately verified that this map is in fact a group isomorphism. Thus, we can view $\mathbf{GL}(n, \mathbb{C})$ as a subgroup of $\mathbf{GL}(2n, \mathbb{R})$, and as a manifold in $\mathbb{R}^{(2n)^2}$.

A 1-manifold is called a *(smooth) curve*, and a 2-manifold is called a *(smooth) surface* (although some authors require that they also be connected).

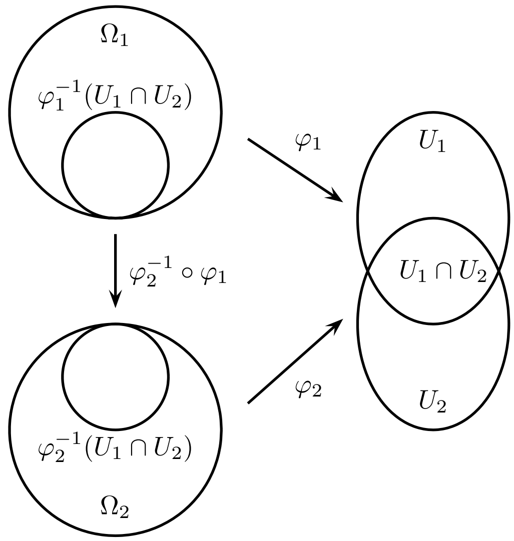

Figure 14.2. Parametrizations and transition functions

The following two lemmas provide the link with the definition of an abstract manifold. The first lemma is easily shown using the inverse function theorem.

Lemma 14.7.2 *Given an m-dimensional manifold M in $\mathbb{R}^N$, for every $p \in M$ there are two open sets $\Omega, W \subseteq \mathbb{R}^N$ with $0_N \in \Omega$ and $p \in M \cap W$, and a smooth diffeomorphism $\varphi\colon \Omega \to W$, such that $\varphi(0_N) = p$ and*

$$\varphi(\Omega \cap (\mathbb{R}^m \times \{0_{N-m}\})) = M \cap W.$$

The next lemma is easily shown from Lemma 14.7.2. It is a key technical result used to show that interesting properties of maps between manifolds do not depend on parametrizations.

Lemma 14.7.3 *Given an m-dimensional manifold M in $\mathbb{R}^N$, for every $p \in M$ and any two parametrizations $\varphi_1\colon \Omega_1 \to U_1$ and $\varphi_2\colon \Omega_2 \to U_2$ of M at p, if $U_1 \cap U_2 \neq \emptyset$, the map $\varphi_2^{-1} \circ \varphi_1\colon \varphi_1^{-1}(U_1 \cap U_2) \to \varphi_2^{-1}(U_1 \cap U_2)$ is a smooth diffeomorphism.*

The maps $\varphi_2^{-1} \circ \varphi_1\colon \varphi_1^{-1}(U_1 \cap U_2) \to \varphi_2^{-1}(U_1 \cap U_2)$ are called *transition maps*. Lemma 14.7.3 is illustrated in Figure 14.2.

Let us review the definitions of a smooth curve in a manifold and the tangent vector at a point of a curve.

Definition 14.7.4 Let M be an m-dimensional manifold in $\mathbb{R}^N$. A *smooth curve γ in M* is any function $\gamma\colon I \to M$ where I is an open interval in $\mathbb{R}$ and such that for every $t \in I$, letting $p = \gamma(t)$, there is some parametrization $\varphi\colon \Omega \to U$ of M at p and some open interval $]t - \epsilon,\, t + \epsilon[\subseteq I$ such that the curve $\varphi^{-1} \circ \gamma\colon]t - \epsilon,\, t + \epsilon[\to \mathbb{R}^m$ is smooth.

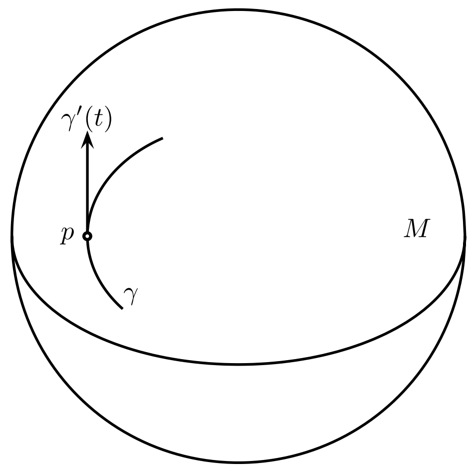

Figure 14.3. Tangent vector to a curve on a manifold

Using Lemma 14.7.3, it is easily shown that Definition 14.7.4 does not depend on the choice of the parametrization $\varphi\colon \Omega \to U$ at p.

Lemma 14.7.3 also implies that γ viewed as a curve $\gamma\colon I \to \mathbb{R}^N$ is smooth. Then the *tangent vector to the curve* $\gamma\colon I \to \mathbb{R}^N$ *at* t, denoted by $\gamma'(t)$, is the value of the derivative of γ at t (a vector in $\mathbb{R}^N$) computed as usual:

$$\gamma'(t) = \lim_{h \mapsto 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Given any point $p \in M$, we will show that the set of tangent vectors to all smooth curves in M through p is a vector space isomorphic to the vector space $\mathbb{R}^m$. The tangent vector at p to a curve γ on a manifold M is illustrated in Figure 14.3.

Given a smooth curve $\gamma\colon I \to M$, for any $t \in I$, letting $p = \gamma(t)$, since M is a manifold, there is a parametrization $\varphi\colon \Omega \to U$ such that $\varphi(0_m) = p \in U$ and some open interval $J \subseteq I$ with $t \in J$ and such that the function

$$\varphi^{-1} \circ \gamma\colon J \to \mathbb{R}^m$$

is a smooth curve, since γ is a smooth curve. Letting $\alpha = \varphi^{-1} \circ \gamma$, the derivative $\alpha'(t)$ is well-defined, and it is a vector in $\mathbb{R}^m$. But $\varphi \circ \alpha\colon J \to M$ is also a smooth curve, which agrees with γ on J, and by the chain rule,

$$\gamma'(t) = \varphi'(0_m)(\alpha'(t)),$$

since $\alpha(t) = 0_m$ (because $\varphi(0_m) = p$ and $\gamma(t) = p$). Observe that $\gamma'(t)$ is a vector in $\mathbb{R}^N$. Now, for every vector $v \in \mathbb{R}^m$, the curve $\alpha\colon J \to \mathbb{R}^m$ defined such that

$$\alpha(u) = (u - t)v$$

for all $u \in J$ is clearly smooth, and $\alpha'(t) = v$. This shows that the set of tangent vectors at t to all smooth curves (in $\mathbb{R}^m$) passing through 0_m is the entire vector space $\mathbb{R}^m$. Since every smooth curve $\gamma: I \to M$ agrees with a curve of the form $\varphi \circ \alpha: J \to M$ for some smooth curve $\alpha: J \to \mathbb{R}^m$ (with $J \subseteq I$) as explained above, and since it is assumed that $\varphi'(0_m)$ is injective, $\varphi'(0_m)$ maps the vector space $\mathbb{R}^m$ injectively to the set of tangent vectors to γ at p, as claimed. All this is summarized in the following definition.

Definition 14.7.5 Let M be an m-dimensional manifold in $\mathbb{R}^N$. For every point $p \in M$, the *tangent space T_pM at p* is the set of all vectors in $\mathbb{R}^N$ of the form $\gamma'(0)$, where $\gamma: I \to M$ is any smooth curve in M such that $p = \gamma(0)$. The set T_pM is a vector space isomorphic to $\mathbb{R}^m$. Every vector $v \in T_pM$ is called a *tangent vector to M at p*.

We can now define Lie groups (postponing defining smooth maps).

Definition 14.7.6 A *Lie group* is a nonempty subset G of $\mathbb{R}^N$ ($N \geq 1$) satisfying the following conditions:

(a) G is a group.

(b) G is a manifold in $\mathbb{R}^N$.

(c) The group operation $\cdot : G \times G \to G$ and the inverse map ${}^{-1}: G \to G$ are smooth.

(Smooth maps are defined in Definition 14.7.10). It is immediately verified that $\mathbf{GL}(n, \mathbb{R})$ is a Lie group. Since all the Lie groups that we are considering are subgroups of $\mathbf{GL}(n, \mathbb{R})$, the following definition is in order.

Definition 14.7.7 A *linear Lie group* is a subgroup G of $\mathbf{GL}(n, \mathbb{R})$ (for some $n \geq 1$) which is a smooth manifold in $\mathbb{R}^{n^2}$.

Let $\mathbf{M}(n, \mathbb{R})$ denote the set of all real $n \times n$ matrices (invertible or not). If we recall that the exponential map

$$\exp: A \mapsto e^A$$

is well defined on $\mathbf{M}(n, \mathbb{R})$, we have the following crucial theorem due to Von Neumann and Cartan.

Theorem 14.7.8 *A closed subgroup G of $\mathbf{GL}(n, \mathbb{R})$ is a linear Lie group. Furthermore, the set $\mathfrak{g}$ defined such that*

$$\mathfrak{g} = \{X \in \mathbf{M}(n, \mathbb{R}) \mid e^{tX} \in G \text{ for all } t \in \mathbb{R}\}$$

is a vector space equal to the tangent space T_IG at the identity I, and $\mathfrak{g}$ is closed under the Lie bracket $[-, -]$ defined such that $[A, B] = AB - BA$ for all $A, B \in \mathbf{M}(n, \mathbb{R})$.

Theorem 14.7.8 applies even when G is a discrete subgroup, but in this case, $\mathfrak{g}$ is trivial (i.e., $\mathfrak{g} = \{0\}$). For example, the set of nonnull reals $\mathbb{R}^* =$

$\mathbb{R} - \{0\} = \mathbf{GL}(1, \mathbb{R})$ is a Lie group under multiplication, and the subgroup

$$H = \{2^n \mid n \in \mathbb{Z}\}$$

is a discrete subgroup of $\mathbb{R}^*$. Thus, H is a Lie group. On the other hand, the set $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ of nonnull rational numbers is a multiplicative subgroup of $\mathbb{R}^*$, but it is not closed, since $\mathbb{Q}$ is dense in $\mathbb{R}$.

The proof of Theorem 14.7.8 involves proving that when G is not a discrete subgroup, there is an open subset $\Omega \subseteq \mathbf{M}(n, \mathbb{R})$ such that $0_{n,n} \in \Omega$, an open subset $W \subseteq \mathbf{M}(n, \mathbb{R})$ such that $I \in W$, and that $\exp\colon \Omega \to W$ is a diffeomorphism such that

$$\exp(\Omega \cap \mathfrak{g}) = W \cap G.$$

If G is closed and not discrete, we must have $m \geq 1$, and $\mathfrak{g}$ has dimension m.

With the help of Theorem 14.7.8 it is now very easy to prove that $\mathbf{SL}(n)$, $\mathbf{O}(n)$, $\mathbf{SO}(n)$, $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, and $\mathbf{SU}(n)$ are Lie groups. We can also prove that $\mathbf{SE}(n)$ is a Lie group as follows. Recall that we can view every element of $\mathbf{SE}(n)$ as a real $(n+1) \times (n+1)$ matrix

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

where $R \in \mathbf{SO}(n)$ and $U \in \mathbb{R}^n$. In fact, such matrices belong to $\mathbf{SL}(n+1)$. This embedding of $\mathbf{SE}(n)$ into $\mathbf{SL}(n+1)$ is a group homomorphism, since the group operation on $\mathbf{SE}(n)$ corresponds to multiplication in $\mathbf{SL}(n+1)$:

$$\begin{pmatrix} RS & RV + U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S & V \\ 0 & 1 \end{pmatrix}.$$

Note that the inverse is given by

$$\begin{pmatrix} R^{-1} & -R^{-1}U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R^{\top} & -R^{\top}U \\ 0 & 1 \end{pmatrix}.$$

Also note that the embedding shows that as a manifold, $\mathbf{SE}(n)$ is diffeomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ (given a manifold M_1 of dimension m_1 and a manifold M_2 of dimension m_2, the product $M_1 \times M_2$ can be given the structure of a manifold of dimension $m_1 + m_2$ in a natural way). Thus, $\mathbf{SE}(n)$ is a Lie group with underlying manifold $\mathbf{SO}(n) \times \mathbb{R}^n$, and in fact, a subgroup of $\mathbf{SL}(n+1)$.

Even though $\mathbf{SE}(n)$ is diffeomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ as a manifold, it is *not* isomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ as a group, because the group multiplication on $\mathbf{SE}(n)$ is not the multiplication on $\mathbf{SO}(n) \times \mathbb{R}^n$. Instead, $\mathbf{SE}(n)$ is a *semidirect product* of $\mathbf{SO}(n)$ and $\mathbb{R}^n$; see Chapter 2, Problem 2.19.

Returning to Theorem 14.7.8, the vector space $\mathfrak{g}$ is called the *Lie algebra* of the Lie group G. Lie algebras are defined as follows.

Definition 14.7.9 A *(real) Lie algebra* $\mathcal{A}$ is a real vector space together with a bilinear map $[\cdot, \cdot]: \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ called the *Lie bracket* on $\mathcal{A}$ such that the following two identities hold for all $a, b, c \in \mathcal{A}$:

$$[a,\, a] = 0,$$

and the so-called *Jacobi identity*

$$[a,\, [b,\, c]] + [c,\, [a,\, b]] + [b,\, [c,\, a]] = 0.$$

It is immediately verified that $[b, a] = -[a, b]$.

In view of Theorem 14.7.8, the vector space $\mathfrak{g} = T_I G$ associated with a Lie group G is indeed a Lie algebra. Furthermore, the exponential map $\exp: \mathfrak{g} \to G$ is well-defined. In general, exp is neither injective nor surjective, as we observed earlier. Theorem 14.7.8 also provides a kind of recipe for "computing" the Lie algebra $\mathfrak{g} = T_I G$ of a Lie group G. Indeed, $\mathfrak{g}$ is the tangent space to G at I, and thus we can use curves to compute tangent vectors. Actually, for every $X \in T_I G$, the map

$$\gamma_X: t \mapsto e^{tX}$$

is a smooth curve in G, and it is easily shown that $\gamma_X'(0) = X$. Thus, we can use these curves. As an illustration, we show that the Lie algebras of $\mathbf{SL}(n)$ and $\mathbf{SO}(n)$ are the matrices with null trace and the skew symmetric matrices.

Let $t \mapsto R(t)$ be a smooth curve in $\mathbf{SL}(n)$ such that $R(0) = I$. We have $\det(R(t)) = 1$ for all $t \in]-\epsilon,\, \epsilon\,[$. Using the chain rule, we can compute the derivative of the function

$$t \mapsto \det(R(t))$$

at $t = 0$, and we get

$$\det{}_I'(R'(0)) = 0.$$

It is an easy exercise to prove that

$$\det{}_I'(X) = \mathrm{tr}(X),$$

and thus $\mathrm{tr}(R'(0)) = 0$, which says that the tangent vector $X = R'(0)$ has null trace. Another proof consists in observing that $X \in \mathfrak{sl}(n, \mathbb{R})$ iff

$$\det(e^{tX}) = 1$$

for all $t \in \mathbb{R}$. Since $\det(e^{tX}) = e^{\mathrm{tr}(tX)}$, for $t = 1$, we get $\mathrm{tr}(X) = 0$, as claimed. Clearly, $\mathfrak{sl}(n, \mathbb{R})$ has dimension $n^2 - 1$.

Let $t \mapsto R(t)$ be a smooth curve in $\mathbf{SO}(n)$ such that $R(0) = I$. Since each $R(t)$ is orthogonal, we have

$$R(t)\, R(t)^\top = I$$

for all $t \in]-\epsilon,\, \epsilon\,[$. Taking the derivative at $t = 0$, we get

$$R'(0)\, R(0)^\top + R(0)\, R'(0)^\top = 0,$$

but since $R(0) = I = R(0)^\top$, we get

$$R'(0) + R'(0)^\top = 0,$$

which says that the tangent vector $X = R'(0)$ is skew symmetric. Since the diagonal elements of a skew symmetric matrix are null, the trace is automatically null, and the condition $\det(R) = 1$ yields nothing new. This shows that $\mathfrak{o}(n) = \mathfrak{so}(n)$. It is easily shown that $\mathfrak{so}(n)$ has dimension $n(n-1)/2$.

As a concrete example, the Lie algebra $\mathfrak{so}(3)$ of $\mathbf{SO}(3)$ is the real vector space consisting of all 3×3 real skew symmetric matrices. Every such matrix is of the form

$$\begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}$$

where $b, c, d \in \mathbb{R}$. The Lie bracket $[A, B]$ in $\mathfrak{so}(3)$ is also given by the usual commutator, $[A, B] = AB - BA$.

We can define an isomorphism of Lie algebras $\psi\colon (\mathbb{R}^3, \times) \to \mathfrak{so}(3)$ by the formula

$$\psi(b, c, d) = \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}.$$

It is indeed easy to verify that

$$\psi(u \times v) = [\psi(u),\ \psi(v)].$$

It is also easily verified that for any two vectors $u = (b, c, d)$ and $v = (b', c', d')$ in $\mathbb{R}^3$

$$\psi(u)(v) = u \times v.$$

The exponential map $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is given by Rodrigues's formula (see Lemma 14.2.3):

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

or equivalently by

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

if $\theta \neq 0$, where

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

$\theta = \sqrt{b^2 + c^2 + d^2}$, $B = A^2 + \theta^2 I_3$, and with $e^{0_3} = I_3$.

Using the above methods, it is easy to verify that the Lie algebras $\mathfrak{gl}(n,\mathbb{R})$, $\mathfrak{sl}(n,\mathbb{R})$, $\mathfrak{o}(n)$, and $\mathfrak{so}(n)$, are respectively $\mathbf{M}(n,\mathbb{R})$, the set of matrices with null trace, and the set of skew symmetric matrices (in the last two cases). A similar computation can be done for $\mathfrak{gl}(n,\mathbb{C})$, $\mathfrak{sl}(n,\mathbb{C})$, $\mathfrak{u}(n)$, and $\mathfrak{su}(n)$, confirming the claims of Section 14.4. It is easy to show that $\mathfrak{gl}(n,\mathbb{C})$ has dimension $2n^2$, $\mathfrak{sl}(n,\mathbb{C})$ has dimension $2(n^2-1)$, $\mathfrak{u}(n)$ has dimension n^2, and $\mathfrak{su}(n)$ has dimension n^2-1.

For example, the Lie algebra $\mathfrak{su}(2)$ of $\mathbf{SU}(2)$ (or S^3) is the real vector space consisting of all 2×2 (complex) skew Hermitian matrices of null trace. Every such matrix is of the form

$$i(d\sigma_1 + c\sigma_2 + b\sigma_3) = \begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix},$$

where $b,c,d \in \mathbb{R}$, and $\sigma_1,\sigma_2,\sigma_3$ are the Pauli spin matrices (see Section 8.1), and thus the matrices $i\sigma_1, i\sigma_2, i\sigma_3$ form a basis of the Lie algebra $\mathfrak{su}(2)$. The Lie bracket $[A,B]$ in $\mathfrak{su}(2)$ is given by the usual commutator, $[A,B]=AB-BA$.

It is easily checked that the vector space $\mathbb{R}^3$ is a Lie algebra if we define the Lie bracket on $\mathbb{R}^3$ as the usual cross product $u\times v$ of vectors. Then we can define an isomorphism of Lie algebras $\varphi\colon(\mathbb{R}^3,\times)\to\mathfrak{su}(2)$ by the formula

$$\varphi(b,c,d) = \frac{i}{2}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{2}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix}.$$

It is indeed easy to verify that

$$\varphi(u\times v) = [\varphi(u),\, \varphi(v)].$$

Returning to $\mathfrak{su}(2)$, letting $\theta=\sqrt{b^2+c^2+d^2}$, we can write

$$d\sigma_1 + c\sigma_2 + b\sigma_3 = \begin{pmatrix} b & -ic+d \\ ic+d & -b \end{pmatrix} = \theta A,$$

where

$$A = \frac{1}{\theta}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{\theta}\begin{pmatrix} b & -ic+d \\ ic+d & -b \end{pmatrix},$$

so that $A^2=I$, and it can be shown that the exponential map $\exp\colon\mathfrak{su}(2)\to\mathbf{SU}(2)$ is given by

$$\exp(i\theta A) = \cos\theta\,\mathbf{1} + i\sin\theta\, A.$$

In view of the isomorphism $\varphi\colon(\mathbb{R}^3,\times)\to\mathfrak{su}(2)$, where

$$\varphi(b,c,d) = \frac{1}{2}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} = i\frac{\theta}{2}A,$$

the exponential map can be viewed as a map $\exp\colon (\mathbb{R}^3, \times) \to \mathbf{SU}(2)$ given by the formula

$$\exp(\theta v) = \left[\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\, v\right],$$

for every vector θv, where v is a unit vector in $\mathbb{R}^3$ and $\theta \in \mathbb{R}$. In this form, $\exp(\theta v)$ is a quaternion corresponding to a rotation of axis v and angle θ.

As we showed, $\mathbf{SE}(n)$ is a Lie group, and its lie algebra $\mathfrak{se}(n)$ described in Section 14.6 is easily determined as the subalgebra of $\mathfrak{sl}(n+1)$ consisting of all matrices of the form

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}$$

where $B \in \mathfrak{so}(n)$ and $U \in \mathbb{R}^n$. Thus, $\mathfrak{se}(n)$ has dimension $n(n+1)/2$. The Lie bracket is given by

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix}\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} BC - CB & BV - CU \\ 0 & 0 \end{pmatrix}.$$

We conclude by indicating the relationship between homomorphisms of Lie groups and homomorphisms of Lie algebras. First, we need to explain what is meant by a smooth map between manifolds.

Definition 14.7.10 Let M_1 (m_1-dimensional) and M_2 (m_2-dimensional) be manifolds in $\mathbb{R}^N$. A function $f\colon M_1 \to M_2$ is *smooth* if for every $p \in M_1$ there are parametrizations $\varphi\colon \Omega_1 \to U_1$ of M_1 at p and $\psi\colon \Omega_2 \to U_2$ of M_2 at $f(p)$ such that $f(U_1) \subseteq U_2$ and

$$\psi^{-1} \circ f \circ \varphi\colon \Omega_1 \to \mathbb{R}^{m_2}$$

is smooth.

Using Lemma 14.7.3, it is easily shown that Definition 14.7.10 does not depend on the choice of the parametrizations $\varphi\colon \Omega_1 \to U_1$ and $\psi\colon \Omega_2 \to U_2$. A smooth map f between manifolds is a *smooth diffeomorphism* if f is bijective and both f and f^{-1} are smooth maps.

We now define the derivative of a smooth map between manifolds.

Definition 14.7.11 Let M_1 (m_1-dimensional) and M_2 (m_2-dimensional) be manifolds in $\mathbb{R}^N$. For any smooth function $f\colon M_1 \to M_2$ and any $p \in M_1$, the function $f'_p\colon T_pM_1 \to T_{f(p)}M_2$, called the *tangent map of f at p, or derivative of f at p, or differential of f at p*, is defined as follows: For every $v \in T_pM_1$ and every smooth curve $\gamma\colon I \to M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = v$,

$$f'_p(v) = (f \circ \gamma)'(0).$$

The map f'_p is also denoted by df_p or T_pf. Doing a few calculations involving the facts that

$$f \circ \gamma = (f \circ \varphi) \circ (\varphi^{-1} \circ \gamma) \quad \text{and} \quad \gamma = \varphi \circ (\varphi^{-1} \circ \gamma)$$

and using Lemma 14.7.3, it is not hard to show that $f'_p(v)$ does not depend on the choice of the curve γ. It is easily shown that f'_p is a linear map.

Finally, we define homomorphisms of Lie groups and Lie algebras and see how they are related.

Definition 14.7.12 Given two Lie groups G_1 and G_2, a *homomorphism (or map) of Lie groups* is a function $f\colon G_1 \to G_2$ that is a homomorphism of groups and a smooth map (between the manifolds G_1 and G_2). Given two Lie algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, a *homomorphism (or map) of Lie algebras* is a function $f\colon \mathcal{A}_1 \to \mathcal{A}_2$ that is a linear map between the vector spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ and that preserves Lie brackets, i.e.,

$$f([A, B]) = [f(A), f(B)]$$

for all $A, B \in \mathcal{A}_1$.

An *isomorphism of Lie groups* is a bijective function f such that both f and f^{-1} are maps of Lie groups, and an *isomorphism of Lie algebras* is a bijective function f such that both f and f^{-1} are maps of Lie algebras. It is immediately verified that if $f\colon G_1 \to G_2$ is a homomorphism of Lie groups, then $f'_I\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism of Lie algebras. If some additional assumptions are made about G_1 and G_2 (for example, connected, simply connected), it can be shown that f is pretty much determined by f'_I.

Alert readers must have noticed that we only defined the Lie algebra of a linear group. In the more general case, we can still define the Lie algebra $\mathfrak{g}$ of a Lie group G as the tangent space T_IG at the identity I. The tangent space $\mathfrak{g} = T_IG$ is a vector space, but we need to define the Lie bracket. This can be done in several ways. We explain briefly how this can be done in terms of so-called adjoint representations. This has the advantage of not requiring the definition of left-invariant vector fields, but it is still a little bizarre!

Given a Lie group G, for every $a \in G$ we define *left translation* as the map $L_a\colon G \to G$ such that $L_a(b) = ab$ for all $b \in G$, and *right translation* as the map $R_a\colon G \to G$ such that $R_a(b) = ba$ for all $b \in G$. The maps L_a and R_a are diffeomorphisms, and their derivatives play an important role. The inner automorphisms $R_{a^{-1}} \circ L_a$ (also written as $R_{a^{-1}}L_a$) also play an important role. Note that

$$R_{a^{-1}}L_a(b) = aba^{-1}.$$

The derivative

$$(R_{a^{-1}}L_a)'_I\colon \mathfrak{g} \to \mathfrak{g}$$

of $R_{a^{-1}}L_a$ at I is an isomorphism of Lie algebras, denoted by $\mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$. The map $a \mapsto \mathrm{Ad}_a$ is a map of Lie groups

$$\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g}),$$

called the *adjoint representation of* G (where $\mathbf{GL}(\mathfrak{g})$ denotes the Lie group of all bijective linear maps on $\mathfrak{g}$).

In the case of a linear group, one can verify that

$$\mathrm{Ad}(a)(X) = \mathrm{Ad}_a(X) = aXa^{-1}$$

for all $a \in G$ and all $X \in \mathfrak{g}$. The derivative

$$\mathrm{Ad}'_I \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

of Ad at I is map of Lie algebras, denoted by $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, called the *adjoint representation of* $\mathfrak{g}$ (where $\mathfrak{gl}(\mathfrak{g})$ denotes the Lie algebra of all linear maps on $\mathfrak{g}$).

In the case of a linear group, it can be verified that

$$\mathrm{ad}(A)(B) = [A,\, B]$$

for all $A, B \in \mathfrak{g}$. One can also check that the Jacobi identity on $\mathfrak{g}$ is equivalent to the fact that ad preserves Lie brackets, i.e., ad is a map of Lie algebras:

$$\mathrm{ad}([A,\, B]) = [\mathrm{ad}(A),\, \mathrm{ad}(B)]$$

for all $A, B \in \mathfrak{g}$ (where on the right, the Lie bracket is the commutator of linear maps on $\mathfrak{g}$). Thus, we recover the Lie bracket from ad.

This is the key to the definition of the Lie bracket in the case of a general Lie group (not just a linear Lie group). We define the Lie bracket on $\mathfrak{g}$ as

$$[A,\, B] = \mathrm{ad}(A)(B).$$

To be complete, we would have to define the exponential map $\exp\colon \mathfrak{g} \to G$ for a general Lie group. For this we would need to introduce some left-invariant vector fields induced by the derivatives of the left translations, and integral curves associated with such vector fields.

This is not hard, but we feel that it is now time to stop our introduction to Lie groups and Lie algebras, even though we have not even touched many important topics, for instance vector fields and differential foms. Readers who wish to learn more about Lie groups and Lie algebras should consult (more or less listed in order of difficulty) Curtis [38], Sattinger and Weaver [147], and Marsden and Ratiu [120]. The excellent lecture notes by Carter, Segal, and Macdonald [30] constitute a very efficient (although somewhat terse) introduction to Lie algebras and Lie groups. Classics such as Weyl [180] and Chevalley [31] are definitely worth consulting, although the presentation and the terminology may seem a bit old fashioned. For more advanced texts, one may consult Abraham and Marsden [1], Warner [176], Sternberg [161], Bröcker and tom Dieck [22], and Knapp [102]. For those who read French, Mneimné and Testard [128] is very clear and quite thorough, and uses very little differential geometry, although it is more advanced than Curtis. Chapter 1, by Bryant, in Freed and Uhlenbeck [24] is also worth reading, but the pace is fast, and Chapters 7 and 8 of Fulton and Harris [69] are very good, but familiarity with manifolds is assumed.

14.8 Applications of Lie Groups and Lie Algebras

Some applications of Lie groups and Lie algebras to robotics and motion planning are discussed in Selig [155] and Murray, Li, and Sastry [131]. Applications to physics are discussed in Sattinger and Weaver [147] and Marsden and Ratiu [120].

The fact that the exponential maps $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ and $\exp\colon \mathfrak{se}(3) \to \mathbf{SE}(3)$ are surjective is important in robotics applications. Indeed, some matrices associated with joints arising in robot kinematics can be written as exponentials $e^{\theta \mathbf{s}}$, where θ is a joint angle and $\mathbf{s} \in \mathfrak{se}(3)$ is the so-called *joint screw* (see Selig [155], Chapter 4). One should also observe that if a rigid motion (R, b) is used to define the position of a rigid body, then the velocity of a point p is given by $(R'p + b')$. In other words, the element (R', b') of the Lie algebra $\mathfrak{se}(3)$ is a sort of velocity vector.

The surjectivity of the exponential map $\exp\colon \mathfrak{se}(3) \to \mathbf{SE}(3)$ implies that there is a map $\log\colon \mathbf{SE}(3) \to \mathfrak{se}(3)$, although it is multivalued. Still, this log "function" can be used to perform motion interpolation. For instance, given two rigid motions $B_1, B_2 \in \mathbf{SE}(3)$ specifying the position of a rigid body B, we can compute $\log(B_1)$ and $\log(B_2)$, which are just elements of the Euclidean space $\mathfrak{se}(3)$, form the linear interpolant $(1-t)\log(B_1)+t\log(B_2)$, and then apply the exponential map to get an interpolating rigid motion

$$e^{(1-t)\log(B_1)+t\log(B_2)}.$$

Of course, this can also be done for a sequence of rigid motions $B_1, \ldots, B_n$, where $n > 2$, and instead of using affine interpolation between two consecutive positions, a polynomial spline can be used to interpolate between the $\log(B_i)$'s in $\mathfrak{se}(3)$. This approach has been investigated by Kim, M.-J., Kim, M.-S. and Shin [98, 99], and Park and Ravani [133, 134].

R.S. Ball published a treatise on the theory of screws in 1900 [8]. Basically, Ball's screws are rigid motions, and his instantaneous screws correspond to elements of the Lie algebra $\mathfrak{se}(3)$ (they are rays in $\mathfrak{se}(3)$). A *screw system* is simply a subspace of $\mathfrak{se}(3)$. Such systems were first investigated by Ball [8]. The first heuristic classification of screw systems was given by Hunt [92]. Screw systems play an important role in kinematics, see McCarthy [124] and Selig [155], Chapter 8.

Lie groups and Lie algebras are also a key ingredient in the use of symmetries in motion, to reduce the number of parameters in the equations of motion, and in optimal control. Such applications are described in a very exciting paper by Marsden and Ostrowski [119] (see also the references in this paper).

14.9 Problems

Problem 14.1 Given a Hermitian space E, for every linear map $f: E \to E$, 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 $A = UTU^*$.

Remark: This extension of Lemma 14.1.3 is usually known as *Schur's lemma*.

Problem 14.2 Prove that the torus obtained by rotating a circle of radius b contained in a plane containing the z-axis and whose center is on a circle of center O and radius b in the xy-plane is a manifold by giving four parametrizations. What are the conditions required on a, b?
Hint. What about

$$\begin{aligned} x &= a\cos\theta + b\cos\theta\cos\varphi, \\ y &= a\sin\theta + b\sin\theta\cos\varphi, \\ z &= b\sin\varphi? \end{aligned}$$

Problem 14.3 (a) Prove that the maps φ_1 and φ_2 parametrizing the sphere are indeed smooth and injective, that $\varphi_1'(u, v)$ and $\varphi_2'(u, v)$ are injective, and that φ_1 and φ_2 give the sphere the structure of a manifold.

(b) Prove that the map $\psi_1: \Delta(1) \to S^2$ defined such that

$$\psi_1(x, y) = \left(x,\, y,\, \sqrt{1 - x^2 - y^2}\right),$$

where $\Delta(1)$ is the unit open disk, is a parametrization of the open upper hemisphere. Show that there are five other similar parametrizations, which, together with ψ_1, make S^2 into a manifold.

Problem 14.4 Use Lemma 14.7.3 to prove that Definition 14.7.4 does not depend on the choice of the parametrization $\varphi: \Omega \to U$ at p.

Problem 14.5 Given a linear Lie group G, for every $X \in T_I G$, letting γ be the smooth curve in G

$$\gamma_X : t \mapsto e^{tX},$$

prove that $\gamma_X'(0) = X$.

Problem 14.6 Prove that

$$\det{}_I'(X) = \mathrm{tr}(X).$$

Hint. Find the directional derivative

$$\lim_{t \to 0} \frac{\det(I + tX) - \det(I)}{t}.$$

Problem 14.7 Confirm that $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{u}(n)$, and $\mathfrak{su}(n)$, are the vector spaces of matrices described in Section 14.4. Prove that $\mathfrak{gl}(n, \mathbb{C})$ has dimension $2n^2$, $\mathfrak{sl}(n, \mathbb{C})$ has dimension $2(n^2 - 1)$, $\mathfrak{u}(n)$ has dimension n^2, and $\mathfrak{su}(n)$ has dimension $n^2 - 1$.

Problem 14.8 Prove that the map $\varphi\colon (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$ defined by the formula

$$\varphi(b, c, d) = \frac{i}{2}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{2}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix}$$

is an isomorphism of Lie algebras. If

$$A = \frac{1}{\theta}\begin{pmatrix} b & -ic + d \\ ic + d & -b \end{pmatrix},$$

where $\theta = \sqrt{b^2 + c^2 + d^2}$, prove that the exponential map $\exp\colon \mathfrak{su}(2) \to \mathbf{SU}(2)$ is given by

$$\exp(i\theta A) = \cos\theta\, \mathbf{1} + i \sin\theta\, A.$$

Problem 14.9 Prove that Definition 14.7.10 does not depend on the parametrizations $\varphi\colon \Omega_1 \to U_1$ and $\psi\colon \Omega_2 \to U_2$.

Problem 14.10 In Definition 14.7.11, prove that $f'_p(v)$ does not depend on the choice of the curve γ, and that f'_p is a linear map.

Problem 14.11 In the case of a linear group, prove that

$$\mathrm{Ad}(a)(X) = \mathrm{Ad}_a(X) = aXa^{-1}$$

for all $a \in G$ and all $X \in \mathfrak{g}$.

Problem 14.12 In the case of a linear group, prove that

$$\mathrm{ad}(A)(B) = [A,\, B]$$

for all $A, B \in \mathfrak{g}$.

Check that the Jacobi identity on $\mathfrak{g}$ is equivalent to the fact that ad preserves Lie brackets, i.e., ad is a map of Lie algebras:

$$\mathrm{ad}([A,\, B]) = [\mathrm{ad}(A),\, \mathrm{ad}(B)]$$

for all $A, B \in \mathfrak{g}$ (where on the right, the Lie bracket is the commutator of linear maps on $\mathfrak{g}$).

Problem 14.13 Consider the Lie algebra $\mathfrak{su}(2)$, whose basis is the Pauli spin matrices $\sigma_1, \sigma_2, \sigma_3$ (see Chapter 6, Section 8.1). The map $\mathrm{ad}(X)$ is a linear map for every $X \in \mathfrak{g}$, since $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. Compute the matrices representing $\mathrm{ad}(\sigma_1)$, $\mathrm{ad}(\sigma_2)$, $\mathrm{ad}(\sigma_3)$.

Problem 14.14 (a) Consider the affine maps ρ of $\mathbb{A}^2$ defined such that

$$\rho\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix},$$

where $\theta, u, v \in \mathbb{R}$.

Given any map ρ as above, letting

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{and} \quad U = \begin{pmatrix} u \\ v \end{pmatrix},$$

ρ can be represented by the 3×3 matrix

$$A = \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & u \\ \sin\theta & \cos\theta & v \\ 0 & 0 & 1 \end{pmatrix}$$

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

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

Prove that these maps are affine bijections and that they form a group, denoted by $\mathbf{SE}(2)$ (the *direct affine isometries, or rigid motions*, of $\mathbb{A}^2$). Prove that such maps preserve the inner product of $\mathbb{R}^2$, i.e., that for any four points $a, b, c, d \in \mathbb{A}^2$,

$$\rho(\mathbf{ac}) \cdot \rho(\mathbf{bd}) = \mathbf{ac} \cdot \mathbf{bd}.$$

If $\theta \neq k2\pi$ ($k \in \mathbb{Z}$), prove that ρ has a unique fixed point c_ρ, and that w.r.t. any frame with origin c_ρ, ρ is a rotation of angle θ and of center c_ρ.

(b) Let us now consider the set of matrices of the form

$$\begin{pmatrix} 0 & -\theta & u \\ \theta & 0 & v \\ 0 & 0 & 0 \end{pmatrix}$$

where $\theta, u, v \in \mathbb{R}$. Verify that this set of matrices is a vector space isomorphic to $(\mathbb{R}^3, +)$. This vector space is denoted by $\mathfrak{se}(2)$. Show that in general, $AB \neq BA$.

(c) Given a matrix

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

letting

$$\Omega = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} u \\ v \end{pmatrix}$$

we can write

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

Prove that

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

where $\Omega^0 = I_2$. Prove that if $\theta = k2\pi$ $(k \in \mathbb{Z})$, then

$$e^A = \begin{pmatrix} I_2 & U \\ 0 & 1 \end{pmatrix},$$

and that if $\theta \neq k2\pi$ $(k \in \mathbb{Z})$, then

$$e^A = \begin{pmatrix} \cos\theta & -\sin\theta & \frac{u}{\theta}\sin\theta + \frac{v}{\theta}(\cos\theta - 1) \\ \sin\theta & \cos\theta & \frac{u}{\theta}(-\cos\theta + 1) + \frac{v}{\theta}\sin\theta \\ 0 & 0 & 1 \end{pmatrix}.$$

Hint. Letting $V = \Omega^{-1}(e^\Omega - I_2)$, prove that

$$V = I_2 + \sum_{k \geq 1} \frac{\Omega^k}{(k+1)!}$$

and that

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

Another proof consists in showing that

$$A^3 = -\theta^2 A,$$

and that

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

(d) Prove that e^A is a direct affine isometry in $\mathbf{SE}(2)$. If $\theta \neq k2\pi$ $(k \in \mathbb{Z})$, prove that V is invertible, and thus prove that the exponential map $\exp\colon \mathfrak{se}(2) \to \mathbf{SE}(2)$ is surjective. How do you need to restrict θ to get an injective map?

Remark: Rigid motions can be used to describe the motion of rigid bodies in the plane. Given a fixed Euclidean frame $(O, (e_1, e_2))$, we can assume that some moving frame $(C, (u_1, u_2))$ is attached (say glued) to a rigid body B (for example, at the center of gravity of B) so that the position and orientation of B in the plane are completely (and uniquely) determined by some rigid motion

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

where U specifies the position of C w.r.t. O, and R specifies the orientation (i.e., angle) of B w.r.t. the fixed frame $(O, (e_1, e_2))$. Then, a motion of B in the plane corresponds to a curve in the space $\mathbf{SE}(2)$. The space $\mathbf{SE}(2)$ is topologically quite complex (in particular, it is "curved"). The exponential map allows us to work in the simpler (noncurved) Euclidean space $\mathfrak{se}(2)$. Thus, given a sequence of "snapshots" of B, say $B_0, B_1, \ldots, B_m$, we can try to find an interpolating motion (a curve in $\mathbf{SE}(2)$) by finding a simpler curve in $\mathfrak{se}(2)$ (say, a B-spline) using the inverse of the exponential map. Of course, it is desirable that the interpolating motion be reasonably smooth and "natural." Computer animations of such motions can be easily implemented.

Problem 14.15 (a) Consider the set of affine maps ρ of $\mathbb{A}^3$ defined such that

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

where R is a rotation matrix (an orthogonal matrix of determinant $+1$) and U is some vector in $\mathbb{R}^3$. Every such a map can be represented by the 4×4 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

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

Prove that these maps are affine bijections and that they form a group, denoted by $\mathbf{SE}(3)$ (the *direct affine isometries, or rigid motions*, of $\mathbb{A}^3$). Prove that such maps preserve the inner product of $\mathbb{R}^3$, i.e., that for any four points $a, b, c, d \in \mathbb{A}^3$,

$$\rho(\mathbf{ac}) \cdot \rho(\mathbf{bd}) = \mathbf{ac} \cdot \mathbf{bd}.$$

Prove that these maps do not always have a fixed point.

(b) Let us now consider the set of 4×4 matrices of the form

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

where Ω is a skew symmetric matrix

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

and U is a vector in $\mathbb{R}^3$.

Verify that this set of matrices is a vector space isomorphic to $(\mathbb{R}^6, +)$. This vector space is denoted by $\mathfrak{se}(3)$. Show that in general, $AB \neq BA$.

(c) Given a matrix

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

as in (b), prove that

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

where $\Omega^0 = I_3$. Given

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

let $\theta = \sqrt{a^2 + b^2 + c^2}$. Prove that if $\theta = k2\pi$ $(k \in \mathbb{Z})$, then

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

and that if $\theta \neq k2\pi$ $(k \in \mathbb{Z})$, then

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

where

$$V = I_3 + \sum_{k \geq 1} \frac{\Omega^k}{(k+1)!}.$$

(d) Prove that

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

and

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

Hint. Use the fact that

$$\Omega^3 = -\theta^2\Omega.$$

(e) Prove that e^A is a direct affine isometry in $\mathbf{SE}(3)$. Prove that V is invertible.
Hint. Assume that the inverse of V is of the form

$$W = I_3 + a\Omega + b\Omega^2,$$

and show that $a, b,$ are given by a system of linear equations that always has a unique solution.

Prove that the exponential map $\exp\colon \mathfrak{se}(3) \to \mathbf{SE}(3)$ is surjective. You may use the fact that $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is surjective, where

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

Remark: Rigid motions can be used to describe the motion of rigid bodies in space. Given a fixed Euclidean frame $(O, (e_1, e_2, e_3))$, we can assume that some moving frame $(C, (u_1, u_2, u_3))$ is attached (say glued) to a rigid body B (for example, at the center of gravity of B) so that the position and orientation of B in space are completely (and uniquely) determined by some rigid motion

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

where U specifies the position of C w.r.t. O, and R specifies the orientation of B w.r.t. the fixed frame $(O, (e_1, e_2, e_3))$. Then a motion of B in space corresponds to a curve in the space $\mathbf{SE}(3)$. The space $\mathbf{SE}(3)$ is topologically quite complex (in particular, it is "curved"). The exponential map allows us to work in the simpler (noncurved) Euclidean space $\mathfrak{se}(3)$. Thus, given a sequence of "snapshots" of B, say $B_0, B_1, \ldots, B_m$, we can try to find an interpolating motion (a curve in $\mathbf{SE}(3)$) by finding a simpler curve in $\mathfrak{se}(3)$ (say, a B-spline) using the inverse of the exponential map. Of course, it is desirable that the interpolating motion be reasonably smooth and "natural." Computer animations of such motions can be easily implemented.

Problem 14.16 Let A and B be the 4×4 matrices

$$A = \begin{pmatrix} 0 & -\theta_1 & 0 & 0 \\ \theta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\theta_2 \\ 0 & 0 & \theta_2 & 0 \end{pmatrix}$$

and

$$B = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 \end{pmatrix}$$

where $\theta_1, \theta_2 \geq 0$. (i) Compute A^2, and prove that

$$B = e^A,$$

where

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

letting $A^0 = I_n$. Use this to prove that for every orthogonal 4×4 matrix B there is a skew symmetric matrix A such that

$$B = e^A.$$

(ii) Given a skew symmetric 4×4 matrix A, prove that there are two skew symmetric matrices A_1 and A_2 and some $\theta_1, \theta_2 \geq 0$ such that

$$\begin{aligned}
A &= A_1 + A_2, \\
A_1^3 &= -\theta_1^2 A_1, \\
A_2^3 &= -\theta_2^2 A_2, \\
A_1 A_2 &= A_2 A_1 = 0, \\
\operatorname{tr}(A_1^2) &= -2\theta_1^2, \\
\operatorname{tr}(A_2^2) &= -2\theta_2^2,
\end{aligned}$$

and where $A_i = 0$ if $\theta_i = 0$ and $A_1^2 + A_2^2 = -\theta_1^2 I_4$ if $\theta_2 = \theta_1$.

Using the above, prove that

$$e^A = I_4 + \frac{\sin\theta_1}{\theta_1} A_1 + \frac{\sin\theta_2}{\theta_2} A_2 + \frac{(1-\cos\theta_1)}{\theta_1^2} A_1^2 + \frac{(1-\cos\theta_2)}{\theta_2^2} A_2^2.$$

(iii) Given an orthogonal 4×4 matrix B, prove that there are two skew symmetric matrices A_1 and A_2 and some $\theta_1, \theta_2 \geq 0$ such that

$$B = I_4 + \frac{\sin\theta_1}{\theta_1} A_1 + \frac{\sin\theta_2}{\theta_2} A_2 + \frac{(1-\cos\theta_1)}{\theta_1^2} A_1^2 + \frac{(1-\cos\theta_2)}{\theta_2^2} A_2^2,$$

where

$$\begin{aligned}
A_1^3 &= -\theta_1^2 A_1, \\
A_2^3 &= -\theta_2^2 A_2, \\
A_1 A_2 &= A_2 A_1 = 0, \\
\operatorname{tr}(A_1^2) &= -2\theta_1^2, \\
\operatorname{tr}(A_2^2) &= -2\theta_2^2,
\end{aligned}$$

and where $A_i = 0$ if $\theta_i = 0$ and $A_1^2 + A_2^2 = -\theta_1^2 I_4$ if $\theta_2 = \theta_1$. Prove that

$$\begin{aligned}
\frac{1}{2}(B - B^\top) &= \frac{\sin\theta_1}{\theta_1} A_1 + \frac{\sin\theta_2}{\theta_2} A_2, \\
\frac{1}{2}(B + B^\top) &= I_4 + \frac{(1-\cos\theta_1)}{\theta_1^2} A_1^2 + \frac{(1-\cos\theta_2)}{\theta_2^2} A_2^2, \\
\operatorname{tr}(B) &= 2\cos\theta_1 + 2\cos\theta_2.
\end{aligned}$$

(iv) Prove that if $\sin\theta_1 = 0$ or $\sin\theta_2 = 0$, then A_1, A_2, and the $\cos\theta_i$ can be computed from B. Prove that if $\theta_2 = \theta_1$, then

$$B = \cos\theta_1 I_4 + \frac{\sin\theta_1}{\theta_1}(A_1 + A_2),$$

and $\cos\theta_1$ and $A_1 + A_2$ can be computed from B.

(v) Prove that

$$\frac{1}{4}\mathrm{tr}\left(\left(B - B^\top\right)^2\right) = 2\cos^2\theta_1 + 2\cos^2\theta_2 - 4.$$

Prove that $\cos\theta_1$ and $\cos\theta_2$ are solutions of the equation

$$x^2 - sx + p = 0,$$

where

$$s = \frac{1}{2}\mathrm{tr}(B), \quad p = \frac{1}{8}\left(\mathrm{tr}(B)\right)^2 - \frac{1}{16}\mathrm{tr}\left(\left(B - B^\top\right)^2\right) - 1.$$

Prove that we also have

$$\cos^2\theta_1\cos^2\theta_2 = \det\left(\frac{1}{2}\left(B + B^\top\right)\right).$$

If $\sin\theta_i \neq 0$ for $i = 1, 2$ and $\cos\theta_2 \neq \cos\theta_1$, prove that the system

$$\begin{aligned}\frac{1}{2}\left(B - B^\top\right) &= \frac{\sin\theta_1}{\theta_1}A_1 + \frac{\sin\theta_2}{\theta_2}A_2,\\ \frac{1}{4}\left(B + B^\top\right)\left(B - B^\top\right) &= \frac{\sin\theta_1\cos\theta_1}{\theta_1}A_1 + \frac{\sin\theta_2\cos\theta_2}{\theta_2}A_2\end{aligned}$$

has a unique solution for A_1 and A_2.

(vi) Prove that $A = A_1 + A_2$ has an orthonormal basis of eigenvectors such that the first two are a basis of the plane w.r.t. which B is a rotation of angle θ_1, and the last two are a basis of the plane w.r.t. which B is a rotation of angle θ_2.

Remark: I do not know a simple way to compute such an orthonormal basis of eigenvectors of $A = A_1 + A_2$, but it should be possible!