2

Lie Algebras and the Exponential Mapping

2.1 The Matrix Exponential

The exponential of a matrix plays a crucial role in the theory of Lie groups. The exponential enters into the definition of the Lie algebra of a matrix Lie group (Section 2.5) and is the mechanism for passing information from the Lie algebra to the Lie group. Since many computations are done much more easily at the level of the Lie algebra, the exponential is indispensable in studying (matrix) Lie groups.

Let X be an $n \times n$ real or complex matrix. We wish to define the exponential of X, denoted e^X or $\exp X$, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}. \tag{2.1}$$

We will follow the convention of using letters such as X and Y for the variable in the matrix exponential.

Proposition 2.1. *For any $n \times n$ real or complex matrix X, the series (2.1) converges. The matrix exponential e^X is a continuous function of X.*

Before proving this, let us review some elementary analysis. Recall that the norm of a vector $x = (x_1, \ldots, x_n)$ in $\mathbb{C}^n$ is defined to be

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2}.$$

We now define the norm of a matrix by thinking of the space $M_n(\mathbb{C})$ of all $n \times n$ matrices as $\mathbb{C}^{n^2}$. This means that we define

$$\|X\| = \left(\sum_{k,l=1}^{n} |X_{kl}|^2 \right)^{1/2}. \tag{2.2}$$

This norm satisfies the inequalities

$$\|X+Y\| \leq \|X\| + \|Y\|, \tag{2.3}$$
$$\|XY\| \leq \|X\| \, \|Y\| \tag{2.4}$$

for all $X, Y \in M_n(\mathbb{C})$. The first of these inequalities is the triangle inequality and is a standard result from elementary analysis. The second of these inequalities follows from the Schwarz inequality (Exercise 1). If X_m is a sequence of matrices, then it is easy to see that X_m converges to a matrix X in the sense of Definition 1.3 if and only if $\|X_m - X\| \to 0$ as $m \to \infty$.

The norm (2.2) is called the **Hilbert–Schmidt** norm. There is another commonly used norm on the space of matrices, called the **operator norm**, whose definition is not relevant to us. It is easily shown that convergence in the Hilbert–Schmidt norm is equivalent to convergence in the operator norm. (This is true because we work with linear operators on the *finite-dimensional* space $\mathbb{C}^n$.) Furthermore, the operator norm also satisfies (2.3) and (2.4). Thus, it matters little whether we use the operator norm or the Hilbert–Schmidt norm.

A sequence X_m of matrices is said to be a **Cauchy sequence** if

$$\|X_m - X_l\| \to 0$$

as $m, l \to \infty$. Thinking of the space $M_n(\mathbb{C})$ of matrices as $\mathbb{C}^{n^2}$ and using a standard result from analysis, we have the following.

Proposition 2.2. *If X_m is a Cauchy sequence in $M_n(\mathbb{C})$, then there exists a unique matrix X such that X_m converges to X.*

That is, every Cauchy sequence in $M_n(\mathbb{C})$ converges.

Now, consider an infinite series whose terms are matrices:

$$X_0 + X_1 + X_2 + \cdots. \tag{2.5}$$

If

$$\sum_{m=0}^{\infty} \|X_m\| < \infty,$$

then the series (2.5) is said to **converge absolutely**. If a series converges absolutely, then it is not hard to show that the partial sums of the series form a Cauchy sequence, and, hence, by Proposition 2.2, the series converges. That is, any series which converges absolutely also converges. (The converse is not true; a series of matrices can converge without converging absolutely.)

We now turn to the proof of Proposition 2.1.

Proof. In light of (2.4), we see that

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

and, hence,

$$\sum_{m=0}^{\infty}\left\|\frac{X^m}{m!}\right\| \leq \sum_{m=0}^{\infty}\frac{\|X\|^m}{m!} = e^{\|X\|} < \infty.$$

Thus, the series (2.1) converges absolutely, and so it converges.

To show continuity, note that since X^m is a continuous function of X, the partial sums of (2.1) are continuous. However, it is easy to see that (2.1) converges uniformly on each set of the form $\{\|X\| \leq R\}$, and so the sum is, again, continuous. □

We now list some elementary properties of the matrix exponential.

Proposition 2.3. *Let X and Y be arbitrary $n \times n$ matrices. Then, we have the following:*

1. $e^0 = I$.
2. $\left(e^X\right)^* = e^{X^*}$.
3. e^X *is invertible and* $\left(e^X\right)^{-1} = e^{-X}$.
4. $e^{(\alpha+\beta)X} = e^{\alpha X}e^{\beta X}$ *for all* α *and* β *in* $\mathbb{C}$.
5. *If* $XY = YX$, *then* $e^{X+Y} = e^X e^Y = e^Y e^X$.
6. *If* C *is invertible, then* $e^{CXC^{-1}} = Ce^XC^{-1}$.
7. $\left\|e^X\right\| \leq e^{\|X\|}$.

It is *not* true in general that $e^{X+Y} = e^X e^Y$, although, by Point 4, it is true if X and Y commute. This is a crucial point, which we will consider in detail later. (See the Lie product formula in Section 2.4 and the Baker–Campbell–Hausdorff formula in Chapter 3.)

Proof. Point 1 is obvious and Point 2 follows from taking term-by-term adjoints of the series for e^X. Points 3 and 4 are special cases of Point 5. To verify Point 5, we simply multiply the power series term by term. (It is left to the reader to verify that this is legal.) Thus,

$$e^X e^Y = \left(I + X + \frac{X^2}{2!} + \cdots\right)\left(I + Y + \frac{Y^2}{2!} + \cdots\right).$$

Multiplying this out and collecting terms where the power of X plus the power of Y equals m, we get

$$e^X e^Y = \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{X^k}{k!}\frac{Y^{m-k}}{(m-k)!} = \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}X^kY^{m-k}. \qquad (2.6)$$

Now, because (and *only* because) X and Y commute,

$$(X+Y)^m = \sum_{k=0}^{m}\frac{m!}{k!(m-k)!}X^kY^{m-k},$$

and, thus, (2.6) becomes

$$e^X e^Y = \sum_{m=0}^{\infty} \frac{1}{m!}(X+Y)^m = e^{X+Y}.$$

To prove Point 6, simply note that

$$\left(CXC^{-1}\right)^m = CX^mC^{-1}$$

and, thus, the two sides of Point 6 are equal term by term.

Point 7 is evident from the proof of Proposition 2.1. □

Proposition 2.4. *Let X be a $n \times n$ complex matrix. Then, e^{tX} is a smooth curve in $M_n(\mathbb{C})$ and*

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

In particular,

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

Proof. Differentiate the power series for e^{tX} term by term. (This is permitted because, for each i and j, $\left(e^{tX}\right)_{ij}$ is given by a convergent power series in t, and it is a standard theorem that one can differentiate power series term by term.) □

2.2 Computing the Exponential of a Matrix

We consider here methods for exponentiating general matrices. A special method for exponentiating 2×2 matrices is described in Exercises 6 and 7.

2.2.1 Case 1: X is diagonalizable

Suppose that X is an $n \times n$ real or complex matrix and that X is diagonalizable over $\mathbb{C}$; that is, there exists an invertible complex matrix C such that $X = CDC^{-1}$, with

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

It is easily verified that e^D is the diagonal matrix with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_n}$, and so in light of Proposition 2.3, we have

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

Thus, if we can explicitly diagonalize X, we can explicitly compute e^X. Note that if X is real, then although C may be complex and the λ_k's may be complex, e^X must come out to be real, since each term in the series (2.1) is real.

For example, take

$$X = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}.$$

Then, the eigenvectors of X are $\begin{pmatrix} 1 \\ i \end{pmatrix}$ and $\begin{pmatrix} i \\ 1 \end{pmatrix}$, with eigenvalues $-ia$ and ia, respectively. Thus, the invertible matrix

$$C = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

maps the basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to the eigenvectors of X, and so (check) $C^{-1}XC$ is a diagonal matrix D. Thus, $X = CDC^{-1}$ and

$$\begin{aligned} e^X &= \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{pmatrix} \begin{pmatrix} 1/2 & -i/2 \\ -i/2 & 1/2 \end{pmatrix} \\ &= \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}. \end{aligned} \tag{2.7}$$

Note that explicitly if X (and hence a) is real, then e^X is real. See Exercise 6 for an alternative method of calculation.

2.2.2 Case 2: X is nilpotent

An $n \times n$ matrix X is said to be **nilpotent** if $X^m = 0$ for some positive integer m. Of course, if $X^m = 0$, then $X^l = 0$ for all $l > m$. In this case, the series (2.1), which defines e^X, terminates after the first m terms, and so can be computed explicitly.

For example, let us compute e^X, where

$$X = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that

$$X^2 = \begin{pmatrix} 0 & 0 & ac \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and that $X^3 = 0$. Thus,

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

2.2.3 Case 3: X arbitrary

A general matrix X may be neither nilpotent nor diagonalizable. However, by Theorem B.6, every matrix X can be written (uniquely) in the form $X = S + N$, with S diagonalizable, N nilpotent, and $SN = NS$. Then, since N and S commute,

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

and e^S and e^N can be computed as in the two previous subsections.

For example, take

$$X = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

Then,

$$X = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}.$$

The two terms clearly commute (since the first one is a multiple of the identity), and, so,

$$e^X = \begin{pmatrix} e^a & 0 \\ 0 & e^a \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^a & e^a b \\ 0 & e^a \end{pmatrix}.$$

2.3 The Matrix Logarithm

We wish to define a matrix logarithm, which should be an inverse function (to the extent possible) to the matrix exponential. Let us recall the situation for the logarithm of complex numbers, in order to see what is reasonable to expect in the matrix case. Since e^z is never zero, only nonzero numbers can have a logarithm. Every nonzero complex number can be written as e^z for some z, but the z is not unique. There is no continuous way to define the logarithm on the set of all nonzero complex numbers. The situation for matrices is similar. For any $X \in M_n(\mathbb{C})$, e^X is invertible; therefore, only invertible matrices can possibly have a logarithm. We will see (Theorem 2.9) that every invertible matrix can be written as e^X, for some $X \in M_n(\mathbb{C})$. However, the X is not unique and there is no continuous way to define the matrix logarithm on the set of all invertible matrices.

The simplest way to define the matrix logarithm is by a power series. We recall how this works in the complex case.

Lemma 2.5. *The function*

$$\log z = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m} \tag{2.8}$$

is defined and analytic in a circle of radius 1 about $z = 1$.

For all z with $|z - 1| < 1$,

$$e^{\log z} = z.$$

For all u with $|u| < \log 2$, $|e^u - 1| < 1$ and

$$\log e^u = u.$$

Proof. The usual logarithm for real, positive numbers satisfies

$$\frac{d}{dx}\log(1-x) = \frac{-1}{1-x} = -\left(1 + x + x^2 + \cdots\right)$$

for $|x| < 1$. Integrating term by term and noting that $\log 1 = 0$ gives

$$\log(1-x) = -\left(x + \tfrac{x^2}{2} + \tfrac{x^3}{3} + \cdots\right).$$

Taking $z = 1 - x$ (so that $x = 1 - z$), we have

$$\begin{aligned}\log z &= -\left((1-z) + \frac{(1-z)^2}{2} + \frac{(1-z)^3}{3} + \cdots\right)\\ &= \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(z-1)^m}{m}.\end{aligned}$$

This series has radius of convergence 1 and defines a complex analytic function on the set $\{|z-1| < 1\}$, which coincides with the usual logarithm for real z in the interval $(0, 2)$. Now, $\exp(\log z) = z$ for $z \in (0, 2)$, and by analyticity, this identity continues to hold on the whole set $\{|z-1| < 1\}$. (That is to say, the functions $z \to \exp(\log z)$ and $z \to z$ are both complex analytic functions and they agree on the interval $(0, 2)$; therefore they must agree on the whole disk $\{|z-1| < 1\}$.)

On the other hand, if $|u| < \log 2$, then

$$|e^u - 1| = \left|u + \tfrac{u^2}{2!} + \cdots\right| \leq |u| + \frac{|u|^2}{2!} + \cdots = e^{|u|} - 1 < 1.$$

Thus, $\log(\exp u)$ makes sense for all such u. Since $\log(\exp u) = u$ for real u with $|u| < \log 2$, it follows by analyticity that $\log(\exp u) = u$ for all complex numbers with $|u| < \log 2$. ☐

Definition 2.6. *For any $n \times n$ matrix A, define $\log A$ by*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \tag{2.9}$$

whenever the series converges.

Since the complex-valued series (2.8) has radius of convergence 1 and since $\|(A-I)^m\| \leq \|A-I\|^m$, the matrix-valued series (2.9) will converge if $\|A - I\| < 1$. However, in contrast to the complex-valued case, the series (2.9) may converge even if $\|A - I\| > 1$, since $\|(A - I)^m\|$ may be strictly smaller than $\|A - I\|^m$. For example, if $A - I$ is nilpotent, then (2.9) terminates and, thus, converges. (See Exercise 8.) Nevertheless, we will mostly content ourselves with considering the case $\|A - I\| < 1$.

Theorem 2.7. *The function*

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

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

For all A with $\|A - I\| < 1$,

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

For all X with $\|X\| < \log 2$, $\left\|e^X - I\right\| < 1$ and

$$\log e^X = X.$$

Proof. Since $\|(A - I)^m\| \leq \|(A - I)\|^m$ and since the series (2.8) has radius of convergence 1, the series (2.9) converges absolutely for all A with $\|A - I\| < 1$. The proof of continuity is essentially the same as for the exponential.

We will now show that $\exp(\log A) = A$ for all A with $\|A - I\| < 1$. We do this by considering two cases.

Case 1: A is diagonalizable.

Suppose that $A = CDC^{-1}$, with D diagonal. Then, $A - I = CDC^{-1} - I = C(D - I)C^{-1}$. It follows that $(A - I)^m$ is of the form

$$(A - I)^m = C \begin{pmatrix} (z_1 - 1)^m & & 0 \\ & \ddots & \\ 0 & & (z_n - 1)^m \end{pmatrix} C^{-1},$$

where $z_1, \ldots, z_n$ are the eigenvalues of A.

Now, if $\|A - I\| < 1$, then it is not hard to show (Exercise 2) that each eigenvalue z_k of A must satisfy $|z_k - 1| < 1$. Thus,

$$\sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A - I)^m}{m} = C \begin{pmatrix} \log z_1 & & 0 \\ & \ddots & \\ 0 & & \log z_n \end{pmatrix} C^{-1},$$

and by Lemma 2.5,

$$e^{\log A} = C \begin{pmatrix} e^{\log z_1} & & 0 \\ & \ddots & \\ 0 & & e^{\log z_n} \end{pmatrix} C^{-1} = A.$$

Case 2: A is not diagonalizable.

If A is not diagonalizable, then, using Theorem B.7, it is not difficult to construct a sequence A_m of diagonalizable matrices with $A_m \to A$. (See Exercise 5.) If $\|A - I\| < 1$, then $\|A_m - I\| < 1$ for all sufficiently large m. By Case 1, $\exp(\log A_m) = A_m$, and, so, by the continuity of exp and log, $\exp(\log A) = A$.

Thus, we have shown that $\exp(\log A) = A$ for all A with $\|A - I\| < 1$. Now, the same argument as in the complex case shows that if $\|X\| < \log 2$, then $\left\|e^X - I\right\| < 1$. The same two-case argument shows that $\log(\exp X) = X$ for all such X. □

Proposition 2.8. *There exists a constant c such that for all $n \times n$ matrices B with $\|B\| < \frac{1}{2}$,*

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

Proof. Note that

$$\log(I + B) - B = \sum_{m=2}^{\infty} (-1)^{m+1} \frac{B^m}{m} = B^2 \sum_{m=2}^{\infty} (-1)^{m+1} \frac{B^{m-2}}{m}$$

so that

$$\|\log(I + B) - B\| \leq \|B\|^2 \sum_{m=2}^{\infty} \frac{\left(\frac{1}{2}\right)^{m-2}}{m}.$$

This is what we want. (It is easily verified that the sum in the last expression is convergent.) □

We may restate the proposition in a more concise way by saying that

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

where $O(\|B\|^2)$ denotes a quantity of order $\|B\|^2$ (i.e., a quantity that is bounded by a constant times $\|B\|^2$ for all sufficiently small values of $\|B\|$).

We conclude this section with a result that, although we will not use it elsewhere, is worth recording. The proof is sketched in Exercises 8 and 9.

Theorem 2.9. *Every invertible $n \times n$ matrix can be expressed as e^X for some $X \in M_n(\mathbb{C})$.*

2.4 Further Properties of the Matrix Exponential

In this section, we give several additional results involving the exponential of a matrix that will be important in our study of Lie algebras.

Theorem 2.10 (Lie Product Formula). *Let X and Y be $n \times n$ complex matrices. Then,*

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

This theorem has a big brother, called the Trotter product formula, which gives the same result in the case where X and Y are suitable unbounded operators on an infinite-dimensional Hilbert space. The Trotter product formula is described, for example, in Reed and Simon (1980), Section VIII.8.

Proof. If we multiply the power series for $e^{\frac{X}{m}}$ and $e^{\frac{Y}{m}}$, all but three of the terms will involve $1/m^2$ or higher powers of $1/m$. Thus,

$$e^{\frac{X}{m}}e^{\frac{Y}{m}} = I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right).$$

Now, since $e^{\frac{X}{m}}e^{\frac{Y}{m}} \to I$ as $m \to \infty$, $e^{\frac{X}{m}}e^{\frac{Y}{m}}$ is in the domain of the logarithm for all sufficiently large m. By Proposition 2.8,

$$\begin{aligned}
\log\left(e^{\frac{X}{m}}e^{\frac{Y}{m}}\right) &= \log\left(I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right) \\
&= \frac{X}{m} + \frac{Y}{m} + O\left(\left\|\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right\|^2\right) \\
&= \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right).
\end{aligned}$$

Exponentiating the logarithm then gives

$$e^{\frac{X}{m}}e^{\frac{Y}{m}} = \exp\left(\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right)$$

and, therefore ,

$$\left(e^{\frac{X}{m}}e^{\frac{Y}{m}}\right)^m = \exp\left(X + Y + O\left(\frac{1}{m}\right)\right).$$

Thus, by the continuity of the exponential, we conclude that

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

which is the Lie product formula. □

Recall (Section B.5) that the trace of a matrix is defined as the sum of its diagonal entries and that similar matrices have the same trace.

Theorem 2.11. *For any $X \in M_n(\mathbb{C})$, we have*

$$\det\left(e^X\right) = e^{\operatorname{trace}(X)}.$$

Proof. There are three cases, as in Section 2.2.

Case 1: X is diagonalizable. Suppose there is a complex invertible matrix C such that

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

Then,

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

Thus, $\operatorname{trace}(X) = \sum \lambda_i$ and $\det(e^X) = \prod e^{\lambda_i} = e^{\sum \lambda_i}$.

Case 2: X is nilpotent. If X is nilpotent, then by Theorem B.7, there is an invertible matrix C such that

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

In that case (it is easy to see), e^X will be upper triangular, with ones on the diagonal:

$$e^X = C \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} C^{-1}.$$

Thus, if X is nilpotent, $\operatorname{trace}(X) = 0$ and $\det(e^X) = 1$.

Case 3: X is arbitrary. As pointed out in Section 2.2, every matrix X can be written as the sum of two commuting matrices S and N, with S diagonalizable (over $\mathbb{C}$) and N nilpotent. Since S and N commute, $e^X = e^S e^N$. So, by the two previous cases,

$$\det\left(e^X\right) = \det\left(e^S\right)\det\left(e^N\right) = e^{\operatorname{trace}(S)} e^{\operatorname{trace}(N)} = e^{\operatorname{trace}(X)},$$

which is what we want. (Note that $\operatorname{trace}(N) = 0$ and $\operatorname{trace}(S) = \operatorname{trace}(X)$.) □

Definition 2.12. *A function $A : \mathbb{R} \to \mathsf{GL}(n;\mathbb{C})$ is called a* ***one-parameter subgroup*** *of $\mathsf{GL}(n;\mathbb{C})$ if*

1. *A is continuous,*
2. *$A(0) = I$,*
3. *$A(t+s) = A(t)A(s)$ for all $t, s \in \mathbb{R}$.*

Theorem 2.13 (One-Parameter Subgroups). *If A is a one-parameter subgroup of $\mathsf{GL}(n;\mathbb{C})$, then there exists a unique $n \times n$ complex matrix X such that*

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

By taking $n=1$, and noting that $\mathsf{GL}(1;\mathbb{C}) \cong \mathbb{C}^*$, this theorem provides a method of solving Exercise 19 in Chapter 1.

Proof. The uniqueness is immediate, since if there is such an X, then $X = \frac{d}{dt}A(t)\big|_{t=0}$. So, we need only worry about existence.

Let B_ε be the open ball of radius ε about zero in $M_n(\mathbb{C})$; that is, $B_\varepsilon = \{X \in M_n(\mathbb{C}) |\, \|X\| < \varepsilon\}$. Assume that $\varepsilon < \log 2$. Then, we have shown that "exp" takes B_ε injectively into $M_n(\mathbb{C})$, with continuous inverse that we denote "log." Now, let $U = \exp(B_{\varepsilon/2})$, which is an open set in $\mathsf{GL}(n;\mathbb{C})$.

Lemma 2.14. *Every $g \in U$ has a unique square root h in U, given by $h = \exp(\frac{1}{2}\log g)$.*

Proof. Let $X = \log g$. Then, $h = \exp(X/2)$ is a square root of g, since $h^2 = \exp(X) = g$. Suppose $h' \in U$ satisfies $(h')^2 = g$. Let $Y = \log h'$; then, $\exp(Y) = h'$ and $\exp(2Y) = (h')^2 = g = \exp(X)$. We have that $Y \in B_{\varepsilon/2}$ and, thus, $2Y \in B_\varepsilon$, and also that $X \in B_{\varepsilon/2} \subset B_\varepsilon$. Since exp is injective on B_ε and $\exp(2Y) = \exp(X)$, we must have $2Y = X$. Thus, $h' = \exp(Y) = \exp(X/2) = h$. This shows the uniqueness of the square root in U. ∎

Returning to the proof of Theorem 2.13, the continuity of A guarantees that there exists $t_0 > 0$ such that $A(t) \in U$ for all t with $|t| \leq t_0$. Then, let $X = \frac{1}{t_0}\log(A(t_0))$, so that $t_0 X = \log(A(t_0))$. Then, $t_0 X \in B_{\varepsilon/2}$ and $A(t_0) = \exp(t_0 X)$. Then, $A(t_0/2)$ is in U and $A(t_0/2)^2 = A(t_0)$. By the lemma, $A(t_0)$ has a *unique* square root in U, and that unique square root is $\exp(t_0 X/2)$. So, we must have $A(t_0/2) = \exp(t_0 X/2)$.

Applying this argument repeatedly, we conclude that

$$A(t_0/2^k) = \exp(t_0 X/2^k)$$

for all positive integers k. Then, for any integer m, we have $A(mt_0/2^k) = A(t_0/2^k)^m = \exp(mt_0 X/2^k)$. This means that $A(t) = \exp(tX)$ for all real numbers t of the form $t = mt_0/2^k$, and the set of such t's is dense in $\mathbb{R}$. Since both $\exp(tX)$ and $A(t)$ are continuous, it follows that $A(t) = \exp(tX)$ for all real numbers t. ∎

2.5 The Lie Algebra of a Matrix Lie Group

The Lie algebra is an indispensable tool in studying matrix Lie groups. On the one hand, Lie algebras are simpler than matrix Lie groups, because (as we will see) the Lie algebra is a linear space. Thus, we can understand much about Lie algebras just by doing linear algebra. On the other hand, the Lie algebra of a matrix Lie group contains much information about that group. (See, for example, Theorem 2.27 in Section 2.7, and the Baker–Campbell–Hausdorff Formula (Chapter 3).) Thus, many questions about matrix Lie groups can be answered by considering a similar but easier problem for the Lie algebra.

Definition 2.15. *Let G be a matrix Lie group. The* ***Lie algebra of*** *G, denoted $\mathfrak{g}$, is the set of all matrices X such that e^{tX} is in G for all real numbers t.*

This means that X is in $\mathfrak{g}$ if and only if the one-parameter subgroup generated by X lies in G. Note that even though G is a subgroup of $\mathsf{GL}(n;\mathbb{C})$ (and not necessarily of $\mathsf{GL}(n;\mathbb{R})$), we do *not* require that e^{tX} be in G for all complex numbers t, but only for all *real* numbers t. Also, it is definitely not enough to have just e^X in G. That is, it is easy to give an example of an X and a G such that $e^X \in G$ but such that $e^{tX} \notin G$ for some real values of t (Exercise 10). Such an X is not in the Lie algebra of G.

There is an abstract notion of a Lie algebra (not necessarily associated to any group), which is described in Section 2.8. The results of Section 2.6 will show that $\mathfrak{g}$ is, indeed, a Lie algebra in that sense.

It is customary to use lowercase Gothic (Fraktur) characters such as $\mathfrak{g}$ and $\mathfrak{h}$ to refer to Lie algebras.

We will show in Section 2.7 that every matrix Lie group is an embedded submanifold of $\mathsf{GL}(n;\mathbb{C})$. We will then show that $\mathfrak{g}$ is the tangent space to G at the identity. See Corollary 2.35. This means that $\mathfrak{g}$ can alternatively be defined as the set of all derivatives of smooth curves through the identity in G.

2.5.1 Physicists' Convention

Physicists are accustomed to considering the map $X \to e^{iX}$ instead of $X \to e^X$. Thus, a physicist would think of the Lie algebra of G as the set of all matrices X such that $e^{itX} \in G$ for all real numbers t. In the physics literature, the Lie algebra is frequently referred to as the space of "infinitesimal group elements." The physics literature does not always distinguish clearly between a matrix Lie group and its Lie algebra.

Before examining general properties of the Lie algebra, let us compute the Lie algebras of the matrix Lie groups introduced in the previous chapter.

2.5.2 The general linear groups

If X is any $n \times n$ complex matrix, then by Proposition 2.3, e^{tX} is invertible. Thus, the Lie algebra of $\mathsf{GL}(n;\mathbb{C})$ is the space of all $n \times n$ complex matrices. This Lie algebra is denoted $\mathsf{gl}(n;\mathbb{C})$.

If X is any $n \times n$ real matrix, then e^{tX} will be invertible and real. On the other hand, if e^{tX} is real for all real numbers t, then $X = \frac{d}{dt}e^{tX}\big|_{t=0}$ will also be real. Thus, the Lie algebra of $\mathsf{GL}(n;\mathbb{R})$ is the space of all $n \times n$ real matrices, denoted $\mathsf{gl}(n;\mathbb{R})$.

Note that the preceding argument shows that if G is a subgroup of $\mathsf{GL}(n;\mathbb{R})$, then the Lie algebra of G must consist entirely of real matrices. We will use this fact when appropriate in what follows.

2.5.3 The special linear groups

Recall Theorem 2.11: $\det(e^X) = e^{\operatorname{trace}(X)}$. Thus, if $\operatorname{trace}(X) = 0$, then $\det(e^{tX}) = 1$ for all real numbers t. On the other hand, if X is any $n \times n$ matrix such that $\det(e^{tX}) = 1$ for all t, then $e^{t\operatorname{trace}(X)} = 1$ for all t. This means that $t\operatorname{trace}(X)$ is an integer multiple of $2\pi i$ for all t, which is only possible if $\operatorname{trace}(X) = 0$. Thus, the Lie algebra of $\mathsf{SL}(n;\mathbb{C})$ is the space of all $n \times n$ complex matrices with trace zero, denoted $\mathsf{sl}(n;\mathbb{C})$.

Similarly, the Lie algebra of $\mathsf{SL}(n;\mathbb{R})$ is the space of all $n \times n$ real matrices with trace zero, denoted $\mathsf{sl}(n;\mathbb{R})$.

2.5.4 The unitary groups

Recall that a matrix U is unitary if and only if $U^* = U^{-1}$. Thus, e^{tX} is unitary if and only if

$$\left(e^{tX}\right)^* = \left(e^{tX}\right)^{-1} = e^{-tX}. \tag{2.10}$$

By Point 2 of Proposition 2.3, $\left(e^{tX}\right)^* = e^{tX^*}$, and so (2.10) becomes

$$e^{tX^*} = e^{-tX}. \tag{2.11}$$

Clearly, a sufficient condition for (2.11) to hold is that $X^* = -X$. On the other hand, if (2.11) holds for all t, then by differentiating at $t = 0$, we see that $X^* = -X$ is necessary.

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

By combining the two previous computations, we see that the Lie algebra of $\mathsf{SU}(n)$ is the space of all $n \times n$ complex matrices X such that $X^* = -X$ and $\operatorname{trace}(X) = 0$, denoted $\mathsf{su}(n)$.

2.5.5 The orthogonal groups

The identity component of $\mathsf{O}(n)$ is just $\mathsf{SO}(n)$. Since (Proposition 2.16) the exponential of a matrix in the Lie algebra is automatically in the identity component, the Lie algebra of $\mathsf{O}(n)$ is the same as the Lie algebra of $\mathsf{SO}(n)$.

Now, an $n \times n$ real matrix R is orthogonal if and only if $R^{tr} = R^{-1}$. So, given an $n \times n$ real matrix X, e^{tX} is orthogonal if and only if $(e^{tX})^{tr} = (e^{tX})^{-1}$, or

$$e^{tX^{tr}} = e^{-tX}. \tag{2.12}$$

Clearly, a sufficient condition for this to hold is that $X^{tr} = -X$. If (2.12) holds for all t, then by differentiating at $t = 0$, we must have $X^{tr} = -X$.

Thus, the Lie algebra of $\mathsf{O}(n)$, as well as the Lie algebra of $\mathsf{SO}(n)$, is the space of all $n \times n$ real matrices X with $X^{tr} = -X$, denoted $\mathsf{so}(n)$. Note that the condition $X^{tr} = -X$ forces the diagonal entries of X to be zero, and, so, necessarily the trace of X is zero.

The same argument shows that the Lie algebra of $\mathsf{SO}(n;\mathbb{C})$ is the space of $n \times n$ complex matrices satisfying $X^{tr} = -X$, denoted $\mathsf{so}(n;\mathbb{C})$. This is not the same as $\mathsf{su}(n)$.

2.5.6 The generalized orthogonal groups

A matrix A is in $\mathsf{O}(n;k)$ if and only if $A^{tr}gA = g$, where g is the $(n+k)\times(n+k)$ diagonal matrix with the first n diagonal entries equal to one and the last k diagonal entries equal to minus one. This condition is equivalent to the condition $g^{-1}A^{tr}g = A^{-1}$, or, since explicitly $g^{-1} = g$, $gA^{tr}g = A^{-1}$. Now, if X is an $(n+k) \times (n+k)$ real matrix, then e^{tX} is in $\mathsf{O}(n;k)$ if and only if

$$ge^{tX^{tr}}g = e^{tgX^{tr}g} = e^{-tX}.$$

This condition holds for all real t if and only if $gX^{tr}g = -X$. Thus, the Lie algebra of $\mathsf{O}(n;k)$, which is the same as the Lie algebra of $\mathsf{SO}(n;k)$, consists of all $(n+k) \times (n+k)$ real matrices X with $gX^{tr}g = -X$. This Lie algebra is denoted $\mathsf{so}(n;k)$.

(In general, the group $\mathsf{SO}(n;k)$ will not be connected, in contrast to the group $\mathsf{SO}(n)$. The identity component of $\mathsf{SO}(n;k)$, which is also the identity component of $\mathsf{O}(n;k)$, is denoted $\mathsf{SO}(n;k)_e$. The Lie algebra of $\mathsf{SO}(n;k)_e$ is the same as the Lie algebra of $\mathsf{SO}(n;k)$.)

2.5.7 The symplectic groups

These are denoted $\mathsf{sp}(n;\mathbb{R})$, $\mathsf{sp}(n;\mathbb{C})$, and $\mathsf{sp}(n)$. The calculation of these Lie algebras is similar to that of the generalized orthogonal groups, and I will just record the result here. Let J be the matrix in the definition of the symplectic groups. Then, $\mathsf{sp}(n;\mathbb{R})$ is the space of $2n \times 2n$ real matrices X such that $JX^{tr}J = X$, $\mathsf{sp}(n;\mathbb{C})$ is the space of $2n \times 2n$ complex matrices satisfying the same condition, and $\mathsf{sp}(n) = \mathsf{sp}(n;\mathbb{C}) \cap \mathsf{u}(2n)$. A simple calculation shows that the elements of $\mathsf{sp}(n;\mathbb{C})$ are precisely the $2n \times 2n$ matrices of the form

$$\begin{pmatrix} A & B \\ C & -A^{tr} \end{pmatrix},$$

where A is an arbitrary $n \times n$ matrix and B and C are arbitrary *symmetric* matrices.

2.5.8 The Heisenberg group

Recall that the Heisenberg group H is the group of all 3×3 real matrices A of the form

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.13}$$

with $a, b, c \in \mathbb{R}$. Recall also that in Section 2.2, Case 2, we computed the exponential of a matrix of the form

$$X = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix} \tag{2.14}$$

and saw that e^X was in H. On the other hand, if X is any matrix such that e^{tX} is of the form (2.13), then all of the entries of $X = \frac{d}{dt} e^{tX}\big|_{t=0}$ which are on or below the diagonal must be zero, so that X is of form (2.14).

Thus, the Lie algebra of the Heisenberg group is the space of all 3×3 real matrices that are strictly upper triangular.

2.5.9 The Euclidean and Poincaré groups

Recall that the Euclidean group $\mathsf{E}(n)$ is (or can be thought of as) the group of $(n+1) \times (n+1)$ real matrices of the form

$$\begin{pmatrix} & & & x_1 \\ & R & & \vdots \\ & & & x_n \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

with $R \in \mathsf{O}(n)$. Now, if X is an $(n+1) \times (n+1)$ real matrix such that e^{tX} is in $\mathsf{E}(n)$ for all t, then $X = \frac{d}{dt} e^{tX}\big|_{t=0}$ must be zero along the bottom row:

$$X = \begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_n \\ 0 & \cdots & 0 \end{pmatrix} \tag{2.15}$$

Our goal, then, is to determine which matrices of the form (2.15) are actually in the Lie algebra of the Euclidean group. A simple computation shows that for $n \geq 1$,

$$\begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_n \\ 0 & \cdots & 0 \end{pmatrix}^n = \begin{pmatrix} & Y^n & Y^{n-1}y \\ & & \\ 0 & \cdots & 0 \end{pmatrix},$$

where y is the column vector with entries $y_1, \ldots, y_n$. It follows that if X is as in (2.15), then e^{tX} is of the form

$$e^{tX} = \begin{pmatrix} & & & * \\ & e^{tY} & & \vdots \\ & & & * \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Now, we have already established that e^{tY} is in $\mathsf{O}(n)$ for all t if and only if $Y^{tr} = -Y$. Thus, we see that the Lie algebra of $\mathsf{E}(n)$ is the space of all $(n+1) \times (n+1)$ real matrices of the form (2.15) with Y satisfying $Y^{tr} = -Y$.

A similar argument shows that the Lie algebra of $\mathsf{P}(n;1)$ is the space of all $(n+2) \times (n+2)$ real matrices of the form

$$\begin{pmatrix} & & y_1 \\ & Y & \vdots \\ & & y_{n+1} \\ 0 & \cdots & 0 \end{pmatrix},$$

with $Y \in \mathsf{so}(n;1)$.

2.6 Properties of the Lie Algebra

We will now establish various basic properties of the Lie algebra of a matrix Lie group. The reader is invited to verify by direct calculation that these general properties hold for the examples computed in the previous section.

Proposition 2.16. *Let G be a matrix Lie group, and X an element of its Lie algebra. Then, e^X is an element of the identity component of G.*

Proof. By definition of the Lie algebra, e^{tX} lies in G for all real t. However, as t varies from 0 to 1, e^{tX} is a continuous path connecting the identity to e^X. □

Proposition 2.17. *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Let X be an element of $\mathfrak{g}$, and A an element of G. Then, AXA^{-1} is in $\mathfrak{g}$.*

Proof. This is immediate, since, by Proposition 2.3,

$$e^{t(AXA^{-1})} = Ae^{tX}A^{-1},$$

and, thus, $Ae^{tX}A^{-1} \in G$ for all t. □

Theorem 2.18. *Let G be a matrix Lie group, $\mathfrak{g}$ its Lie algebra, and X and Y elements of $\mathfrak{g}$. Then*

1. *$sX \in \mathfrak{g}$ for all real numbers s,*
2. *$X + Y \in \mathfrak{g}$,*
3. *$XY - YX \in \mathfrak{g}$.*

If one follows the physics convention for the definition of the Lie algebra, then condition 3 should be replaced with the condition $-i\,(XY - YX) \in \mathfrak{g}$. Properties 1 and 2 show that $\mathfrak{g}$ is a real vector space, (i.e., a real subspace of the space of $M_n(\mathbb{C})$). Property 3 shows that $\mathfrak{g}$ is, in fact, a Lie algebra in the abstract sense described in Section 2.8. Note that Property 1 applies only to *real* numbers s (compare Definition 2.20).

Proof. Point 1 is immediate, since $e^{t(sX)} = e^{(ts)X}$, which must be in G if X is in $\mathfrak{g}$. Point 2 is easy to verify *if* X and Y commute, since, in that case, $e^{t(X+Y)} = e^{tX}e^{tY}$. If X and Y do not commute, this argument does not work. However, the Lie product formula states that

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

Because X and Y are in the Lie algebra, $e^{tX/m}$ and $e^{tY/m}$ are in G, as is $\left(e^{tX/m}e^{tY/m}\right)^m$, since G is a group. However, because G is a matrix Lie group, the limit of things in G must be again in G, provided that the limit is invertible. Since $e^{t(X+Y)}$ is automatically invertible, we conclude that it must be in G. This shows that $X+Y$ is in $\mathfrak{g}$.

Now for Point 3. Recall (Proposition 2.4) that $\frac{d}{dt}e^{tX}\big|_{t=0} = X$. It follows that $\frac{d}{dt}e^{tX}Y\big|_{t=0} = XY$, and, hence, by the product rule (Exercise 3),

$$\begin{aligned}\frac{d}{dt}\left(e^{tX}Ye^{-tX}\right)\bigg|_{t=0} &= (XY)e^0 + (e^0Y)(-X)\\ &= XY - YX.\end{aligned}$$

Now, by Proposition 2.17, $e^{tX}Ye^{-tX}$ is in $\mathfrak{g}$ for all t. Furthermore, we have (by Points 1 and 2) established that $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$. This means, in particular, that $\mathfrak{g}$ is a topologically closed subset of $M_n(\mathbb{C})$. It follows that

$$XY - YX = \lim_{h\to 0}\frac{e^{hX}Ye^{-hX} - Y}{h}$$

belongs to $\mathfrak{g}$. □

Definition 2.19. *Given two $n \times n$ matrices A and B, the* ***bracket*** *(or* ***commutator****) of A and B, denoted $[A, B]$, is defined to be*

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

According to Theorem 2.18, the Lie algebra of any matrix Lie group is closed under brackets.

It is important to note that even if the elements of G have complex entries, the Lie algebra $\mathfrak{g}$ of G is not necessarily a complex vector space. That is, for X in $\mathfrak{g}$, iX may not be in $\mathfrak{g}$. For example, elements of $\mathsf{SU}(n)$ will, in general, have complex entries (i.e., $\mathsf{SU}(n)$ is not contained in $\mathsf{GL}(n;\mathbb{R})$). Nevertheless, if X is in the Lie algebra $\mathsf{su}(n)$, then $X^* = -X$ and, so, $(iX)^* = iX$. This means that iX is not in $\mathsf{su}(n)$ unless X is zero.

Definition 2.20. *A matrix Lie group G is said to be* ***complex*** *if its Lie algebra $\mathfrak{g}$ is a complex subspace of $M_n(\mathbb{C})$ (i.e., if $iX \in \mathfrak{g}$ for all $X \in \mathfrak{g}$).*

Examples of complex groups are $\mathsf{GL}(n;\mathbb{C})$, $\mathsf{SL}(n;\mathbb{C})$, $\mathsf{SO}(n;\mathbb{C})$, and $\mathsf{Sp}(n;\mathbb{C})$. The condition in Definition 2.20 is equivalent to the condition that G be a complex submanifold of $\mathsf{GL}(n;\mathbb{C})$. (See Appendix C.)

We return now to the setting of general, not necessarily complex, matrix Lie groups. The following very important theorem tells us that a Lie group homomorphism between two Lie groups gives rise in a natural way to a map between the corresponding Lie algebras. In particular, this will tell us that two isomorphic Lie groups have "the same" Lie algebras (i.e., the Lie algebras are isomorphic in the sense of Section 2.8). See Exercise 12.

Theorem 2.21. *Let G and H be matrix Lie groups, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Suppose that $\Phi : G \to H$ is a Lie group homomorphism. Then, there exists a unique real linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\Phi(e^X) = e^{\phi(X)} \tag{2.16}$$

for all $X \in \mathfrak{g}$. The map ϕ has following additional properties:

1. *$\phi\left(AXA^{-1}\right) = \Phi(A)\phi(X)\Phi(A)^{-1}$, for all $X \in \mathfrak{g}$, $A \in G$*
2. *$\phi([X,Y]) = [\phi(X), \phi(Y)]$, for all $X, Y \in \mathfrak{g}$*
3. *$\phi(X) = \frac{d}{dt}\Phi(e^{tX})\big|_{t=0}$, for all $X \in \mathfrak{g}$*

Suppose that G, H, and K are matrix Lie groups and $\Phi : H \to K$ and $\Psi : G \to H$ are Lie group homomorphisms. Let $\Lambda : G \to K$ be the composition of Φ and Ψ, $\Lambda(A) = \Phi(\Psi(A))$. Let ϕ, ψ, and λ be the associated Lie algebra maps. Then,

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

In practice, given a Lie group homomorphism Φ, the way one goes about computing ϕ is by using Property 3. Of course, since ϕ is (real) linear, it suffices to compute ϕ on a basis for $\mathfrak{g}$. In the language of differentiable manifolds, Property 3 says that ϕ is the derivative (or differential) of Φ at the identity, which is the standard definition of ϕ. (See also Corollary 2.35 in Section 2.7.)

A linear map with Property 2 is called a **Lie algebra homomorphism**. (See Section 2.8.) This theorem says that every Lie group homomorphism gives rise to a Lie algebra homomorphism. We will see eventually that the converse is true *under certain circumstances*. Specifically, suppose that G and H are Lie groups and that $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism. If G is *simply connected*, then there exists a unique Lie group homomorphism $\Phi : G \to H$ such that Φ and ϕ are related as in Theorem 2.21. (The proof of this deep result is in Chapter 3.) We now proceed with the proof of Theorem 2.21.

Proof. The proof is similar to the proof of Theorem 2.18. Since Φ is a continuous group homomorphism, $\Phi(e^{tX})$ will be a one-parameter subgroup of H, for each $X \in \mathfrak{g}$. Thus, by Theorem 2.13, there is a unique matrix Z such that

$$\Phi\left(e^{tX}\right) = e^{tZ} \tag{2.17}$$

for all $t \in \mathbb{R}$. This Z must lie in $\mathfrak{h}$ since $e^{tZ} = \Phi\left(e^{tX}\right) \in H$.

We now define $\phi(X) = Z$ and check in several steps that ϕ has the required properties.

Step 1: $\Phi(e^X) = e^{\phi(X)}$.
This follows from (2.17) and our definition of ϕ, by putting $t = 1$.

Step 2: $\phi(sX) = s\phi(X)$ *for all* $s \in \mathbb{R}$.
This is immediate, since if $\Phi(e^{tX}) = e^{tZ}$, then $\Phi(e^{tsX}) = e^{tsZ}$.

Step 3: $\phi(X+Y) = \phi(X) + \phi(Y)$.
By Steps 1 and 2,

$$e^{t\phi(X+Y)} = e^{\phi[t(X+Y)]} = \Phi\left(e^{t(X+Y)}\right).$$

By the Lie product formula and the fact that Φ is a continuous homomorphism, we have

$$\begin{aligned} e^{t\phi(X+Y)} &= \Phi\Big(\lim_{m\to\infty}\left(e^{tX/m}e^{tY/m}\right)^m\Big) \\ &= \lim_{m\to\infty}\left(\Phi\left(e^{tX/m}\right)\Phi(e^{tY/m})\right)^m. \end{aligned}$$

However, we then have

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

Differentiating this result at $t = 0$ gives the desired result.

Step 4: $\phi\left(AXA^{-1}\right) = \Phi(A)\phi(X)\Phi(A)^{-1}$.
By Steps 1 and 2,

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

Using a property of the exponential and Step 1, this becomes

$$\begin{aligned} \exp t\phi(AXA^{-1}) &= \Phi\big(Ae^{tX}A^{-1}\big) = \Phi(A)\Phi(e^{tX})\Phi(A)^{-1} \\ &= \Phi(A)e^{t\phi(X)}\Phi(A)^{-1}. \end{aligned}$$

Differentiating this at $t = 0$ gives the desired result.

Step 5: $\phi([X,Y]) = [\phi(X), \phi(Y)]$.
Recall from the proof of Theorem 2.18 that

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

Hence,

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

where we have used the fact that a derivative commutes with a linear transformation.

Now, by Step 4,

$$\begin{aligned} \phi([X,Y]) &= \frac{d}{dt} \Phi(e^{tX}) \phi(Y) \Phi(e^{-tX}) \bigg|_{t=0} \\ &= \frac{d}{dt} e^{t\phi(X)} \phi(Y) e^{-t\phi(X)} \bigg|_{t=0} \\ &= [\phi(X), \phi(Y)]. \end{aligned}$$

Step 6: $\phi(X) = \frac{d}{dt} \Phi(e^{tX})\big|_{t=0}$.

This follows from (2.17) and our definition of ϕ.

Step 7: *ϕ is the unique real linear map such that* $\Phi(e^X) = e^{\phi(X)}$.

Suppose that ψ is another such map. Then,

$$e^{t\psi(X)} = e^{\psi(tX)} = \Phi(e^{tX})$$

so that

$$\psi(X) = \frac{d}{dt} \Phi(e^{tX}) \bigg|_{t=0}.$$

Thus, by Step 6, ψ coincides with ϕ.

Step 8: $\lambda = \phi \circ \psi$.

For any $X \in \mathfrak{g}$,

$$\Lambda\left(e^{tX}\right) = \Phi\big(\Psi\big(e^{tX}\big)\big) = \Phi\left(e^{t\psi(X)}\right) = e^{t\phi(\psi(X))}.$$

Thus, $\lambda(X) = \phi(\psi(X))$. □

Definition 2.22 (The Adjoint Mapping). *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Then, for each $A \in G$, define a linear map $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ by the formula*

$$\mathrm{Ad}_A(X) = AXA^{-1}.$$

Proposition 2.23. *Let G be a matrix Lie group, with Lie algebra $\mathfrak{g}$. Let $\mathsf{GL}(\mathfrak{g})$ denote the group of all invertible linear transformations of $\mathfrak{g}$. Then, for each $A \in G$, Ad_A is an invertible linear transformation of $\mathfrak{g}$ with inverse $\mathrm{Ad}_{A^{-1}}$, and the map $A \to \mathrm{Ad}_A$ is a group homomorphism of G into $\mathsf{GL}(\mathfrak{g})$. Furthermore, for each $A \in G$, Ad_A satisfies $\mathrm{Ad}_A([X,Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$ for all $X, Y \in \mathfrak{g}$.*

Proof. Easy. Note that Proposition 2.17 guarantees that $\mathrm{Ad}_A(X)$ is actually in $\mathfrak{g}$ for all $X \in \mathfrak{g}$. $\square$

Since $\mathfrak{g}$ is a real vector space with some dimension k, $\mathsf{GL}(\mathfrak{g})$ is essentially the same as $\mathsf{GL}(k;\mathbb{R})$. Thus, we will regard $\mathsf{GL}(\mathfrak{g})$ as a matrix Lie group. It is easy to show that $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ is continuous, and so is a Lie group homomorphism. By Theorem 2.21, there is an associated real linear map $X \to \mathrm{ad}_X$ from the Lie algebra of G to the Lie algebra of $\mathsf{GL}(\mathfrak{g})$ (i.e., from $\mathfrak{g}$ to $\mathsf{gl}(\mathfrak{g})$), with the property that

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

Here, $\mathsf{gl}(\mathfrak{g})$ is the Lie algebra of $\mathsf{GL}(\mathfrak{g})$, namely the space of all linear maps of $\mathfrak{g}$ to itself.

Proposition 2.24. *Let G be a matrix Lie group, let $\mathfrak{g}$ be its Lie algebra, and let $\mathrm{Ad} : G \to \mathsf{GL}(\mathfrak{g})$ be the Lie group homomorphism defined above. Let $\mathrm{ad} : \mathfrak{g} \to \mathsf{gl}(\mathfrak{g})$ be the associated Lie algebra map. Then, for all $X, Y \in \mathfrak{g}$*

$$\mathrm{ad}_X(Y) = [X, Y]. \tag{2.18}$$

Proof. Recall that by Point 3 of Theorem 2.21, ad can be computed as follows:

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

Thus,

$$\begin{aligned}\mathrm{ad}_X(Y) &= \left.\frac{d}{dt}\mathrm{Ad}(e^{tX})(Y)\right|_{t=0} = \left.\frac{d}{dt}e^{tX}Ye^{-tX}\right|_{t=0} \\ &= [X, Y],\end{aligned}$$

which is what we wanted to prove. $\square$

We have proved, as a consequence of Theorem 2.21 and Proposition 2.24, the following result, which we will make use of later.

Proposition 2.25. *For any X in $M_n(\mathbb{C})$, let $\mathrm{ad}_X : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be given by $\mathrm{ad}_X Y = [X, Y]$. Then, for any Y in $M_n(\mathbb{C})$, we have*

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

This result can also be proved by direct calculation—see Exercise 19.

2.7 The Exponential Mapping

Definition 2.26. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, then the **exponential mapping for** G is the map*

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

That is, the exponential mapping for G is the matrix exponential restricted to the Lie algebra $\mathfrak{g}$ of G. We have shown (Theorem 2.9) that every matrix in $\mathsf{GL}(n;\mathbb{C})$ is the exponential of some $n\times n$ matrix. Nevertheless, if $G \subset \mathsf{GL}(n;\mathbb{C})$ is a closed subgroup, there may exist A in G such that there is no X *in the Lie algebra* $\mathfrak{g}$ *of* G with $\exp X = A$. Consider, for example, the matrix

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

in $\mathsf{SL}(2;\mathbb{C})$. I claim that there exists no $X \in \mathsf{sl}(2;\mathbb{C})$ with $\exp X = A$. To see this, consider an arbitrary matrix X in $\mathsf{sl}(2;\mathbb{C})$. Since $\operatorname{trace}(X) = 0$, the eigenvalues of X are negatives of each other. There are then two possibilities. First, the eigenvalues of X could both be zero. In that case, $\exp X$ will have 1 as an eigenvalue and, so, $\exp X \neq A$. Second, the eigenvalues of X could be of the form $(\lambda, -\lambda)$, with λ being a nonzero complex number. In that case, X has *distinct* eigenvalues and is, therefore, diagonalizable. It follows that $\exp X$ is also diagonalizable. However, A is not diagonalizable. (The eigenvalues of A are -1 and -1; if it were diagonalizable it would have to be $-I$.) This shows that $\exp X \neq A$. (See also Exercises 26, 27, 29, 30, and 31.)

We see, then, that the exponential mapping for a matrix Lie group G does not necessarily map $\mathfrak{g}$ onto G. Furthermore, the exponential mapping may not be one-to-one on $\mathfrak{g}$. Nevertheless, it provides a crucial mechanism for passing information between the group and the Lie algebra. Indeed, we will see (Corollary 2.29) that the exponential mapping is *locally* one-to-one and onto, a result that will be essential, for example, in Chapter 3.

Theorem 2.27. *For* $0 < \varepsilon < \ln 2$, *let* $U_\varepsilon = \{X \in M_n(\mathbb{C}) \mid \|X\| < \varepsilon\}$ *and let* $V_\varepsilon = \exp(U_\varepsilon)$. *Suppose* $G \subset \mathsf{GL}(n;\mathbb{C})$ *is a matrix Lie group with Lie algebra* $\mathfrak{g}$. *Then there exists* $\varepsilon \in (0, \ln 2)$ *such that for all* $A \in V_\varepsilon$, A *is in* G *if and only if* $\log A$ *is in* $\mathfrak{g}$.

The condition $\varepsilon < \ln 2$ guarantees (Theorem 2.7) that for all $X \in V_\varepsilon$, $\log(\exp X)$ is defined and equal to X.

Note that if $X = \log A$ is in $\mathfrak{g}$, then $A = \exp X$ is in G. Thus, the content of the theorem is that for some ε, having A in $V_\varepsilon \cap G$ implies that $\log A$ must be in $\mathfrak{g}$. There are several important consequences of this theorem, described after the proof.

Proof. We begin with a lemma.

Lemma 2.28. *Suppose* B_m *are elements of* G *and that* $B_m \to I$. *Let* $Y_m = \log B_m$, *which is defined for all sufficiently large* m. *Suppose that* Y_m *is nonzero for all* m *and that* $Y_m / \|Y_m\| \to Y \in M_n(\mathbb{C})$. *Then,* $Y \in \mathfrak{g}$.

Proof. To show that $Y \in \mathfrak{g}$, we must show that $\exp tY \in G$ for all $t \in \mathbb{R}$. As $m \to \infty$, $(t/\|Y_m\|)\, Y_m \to tY$. Note that since $B_m \to I$, $Y_m \to 0$, and, so, $\|Y_m\| \to 0$. Thus, we can find integers k_m such that $(k_m \|Y_m\|) \to t$. Then,

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

However, $\exp(k_m Y_m) = \exp(Y_m)^{k_m} = (B_m)^{k_m} \in G$ and G is closed, and we conclude that $\exp(tY) \in G$. □

Let us think of $M_n(\mathbb{C})$ as $\mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$. Then, $\mathfrak{g}$ is a subspace of $\mathbb{R}^{2n^2}$. Let D denote the orthogonal complement of $\mathfrak{g}$ with respect to the usual inner product on $\mathbb{R}^{2n^2}$. Consider the map $\Phi : \mathfrak{g} \oplus D \to \mathsf{GL}(n;\mathbb{C})$ given by

$$\Phi(X, Y) = e^X e^Y.$$

Of course, we can identify $\mathfrak{g} \oplus D$ with $\mathbb{R}^{2n^2}$. Moreover, $\mathsf{GL}(n;\mathbb{C})$ is an open subset of $M_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$. Thus, we can regard Φ as a map from $\mathbb{R}^{2n^2}$ to itself.

Now, using the properties of the matrix exponential, we see that

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

This shows that the derivative of Φ at the point $0 \in \mathbb{R}^{2n^2}$ is the identity. (Recall that the derivative at a point of a function from $\mathbb{R}^{2n^2}$ to itself is a linear map of $\mathbb{R}^{2n^2}$ to itself, in this case the identity map.) In particular, the derivative of Φ at 0 is invertible. Thus, the inverse function theorem says that Φ has a continuous local inverse, defined in a neighborhood of I.

Now, as we have remarked, what we need to prove is that for some ε, $A \in V_\varepsilon \cap G$ implies $\log A \in \mathfrak{g}$. Suppose this is not the case. Then we can find a sequence A_m in G such that $A_m \to I$ as $m \to \infty$ and such that for all m, $\log A_m \notin \mathfrak{g}$. Using the local inverse of the map Φ, we can write A_m (for all sufficiently large m) as

$$A_m = e^{X_m} e^{Y_m}, \quad X_m \in \mathfrak{g},\, Y_m \in D,$$

in such a way that X_m and Y_m tend to zero as m tends to infinity. We must have $Y_m \neq 0$, since otherwise we would have $\log A_m = X_m \in \mathfrak{g}$.

Now, let $B_m = \exp(-X_m)A_m = \exp(Y_m)$. Then, B_m is in G and $B_m \to I$ as $m \to \infty$. Since the unit sphere in D is compact, we can choose a subsequence of the Y_m's (still called Y_m) so that $Y_m / \|Y_m\|$ converges to some $Y \in D$, with $\|Y\| = 1$. Then, by the lemma, $Y \in \mathfrak{g}$. This is a contradiction, because D is the orthogonal complement of $\mathfrak{g}$. Thus, there must be some ε such that $\log A \in \mathfrak{g}$ for all A in $V_\varepsilon \cap G$. □

Corollary 2.29. *If G is a matrix Lie group with Lie algebra $\mathfrak{g}$, there exists a neighborhood U of 0 in $\mathfrak{g}$ and a neighborhood V of I in G such that the exponential mapping takes U homeomorphically onto V.*

Proof. Let ε be such that Theorem 2.27 holds and set $U = U_\varepsilon \cap \mathfrak{g}$ and $V = V_\varepsilon \cap G$. The theorem implies that exp takes U onto V. Furthermore, exp is a homeomorphism of U onto V, since there is a continuous inverse map, namely, the restriction of the matrix logarithm to V. □

Definition 2.30. *If U and V are as in Corollary 2.29, then the inverse map $\exp^{-1} : V \to \mathfrak{g}$ is called the **logarithm** for G.*

Corollary 2.31. *If G is a connected matrix Lie group, then every element A of G can be written in the form*

$$A = e^{X_1} e^{X_2} \cdots e^{X_m} \tag{2.19}$$

for some $X_1, X_2, \ldots, X_m$ in $\mathfrak{g}$.

Even if G is connected, it is definitely *not* the case in general that every element of G can be written as single exponential, $A = \exp X$ (with $X \in \mathfrak{g}$), as the example given earlier in this section shows.

Proof. Since G is connected, we can find a continuous path $A(t)$ in G with $A(0) = I$ and $A(1) = A$. Let V be a neighborhood of I in G as in Corollary 2.29, so that every element of V is the exponential of an element of $\mathfrak{g}$. A standard argument using the compactness of the interval $[0, 1]$ shows that we can pick a sequence of numbers $t_0, \ldots, t_m$ with $0 = t_0 < t_1 < \cdots < t_m = 1$ such that

$$A_{t_{k-1}}^{-1} A_{t_k} \in V$$

for all $k = 1, \ldots, m$. Then,

$$A = (A_{t_0}^{-1} A_{t_1})(A_{t_1}^{-1} A_{t_2}) \cdots (A_{t_{m-1}}^{-1} A_{t_m}).$$

If we choose $X_k \in \mathfrak{g}$ with $\exp X_k = A_{t_{k-1}}^{-1} A_{t_k}$ $(k = 1, \ldots, m)$, we have

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

□

Corollary 2.32. *Suppose G is a connected matrix Lie group, H is a matrix Lie group, and Φ_1 and Φ_2 are Lie group homomorphisms of G into H. Let ϕ_1 and ϕ_2 be the associated Lie algebra homomorphisms. If $\phi_1 = \phi_2$, then $\Phi_1 = \Phi_2$.*

Proof. Let g be any element of G. Since G is connected, Corollary 2.31 tells us that g can be written as $g = e^{X_1} e^{X_2} \cdots e^{X_n}$, with $X_i \in \mathfrak{g}$. Then,

$$\begin{aligned}
\Phi_1(g) &= \Phi_1(e^{X_1}) \cdots \Phi_1(e^{X_n}) \\
&= e^{\phi_1(X_1)} \cdots e^{\phi_1(X_n)} \\
&= e^{\phi_2(X_1)} \cdots e^{\phi_2(X_n)} \\
&= \Phi_2(e^{X_1}) \cdots \Phi_2(e^{X_n}) \\
&= \Phi_2(g).
\end{aligned}$$

□

We are now in a position to obtain Theorem 1.19 of Chapter 1 as a consequence of Theorem 2.27.

Corollary 2.33. *Every matrix Lie group G is a smooth embedded submanifold of $M_n(\mathbb{C})$ and, hence, a Lie group.*

Proof. Let $\varepsilon \in (0, \ln 2)$ be such that Theorem 2.27 holds. Then for any $A_0 \in G$, consider the neighborhood $A_0 V_\varepsilon$ of A_0 in $M_n(\mathbb{C})$. Note that $A \in A_0 V_\varepsilon$ if and only if $A_0^{-1} A \in V_\varepsilon$. Define a local coordinate system on $A_0 V_\varepsilon$ by writing each $A \in A_0 V_\varepsilon$ as $A = A_0 \exp X$, for $X \in U_\varepsilon \subset M_n(\mathbb{C})$. It follows from Theorem 2.27 that (for $A \in A_0 V_\varepsilon$) $A \in G$ if and only if $X \in \mathfrak{g}$. This means that in this local coordinate system defined near A_0, G looks like the subspace $\mathfrak{g}$ of $M_n(\mathbb{C})$. Since we can find such local coordinates near any point A_0 in G, G is an embedded submanifold of $M_n(\mathbb{C})$. This shows, as discussed in Section C.2.6, that G is a Lie group. $\square$

Corollary 2.33 implies that a matrix Lie group G is necessarily *locally* path-connected. It follows that G is connected (in the usual topological sense) if and only if it is path-connected. Thus our definition of connectedness in Section 1.7 (which was actually *path*-connectedness) is equivalent to the usual topological definition.

Corollary 2.34. *Every continuous homomorphism between two matrix Lie groups is smooth.*

Proof. Given $A \in G$, we write nearby elements $B \in G$ (as in the proof of Corollary 2.33) as $B = A \exp X$, $X \in \mathfrak{g}$. Then,

$$\Phi(B) = \Phi(A)\Phi(\exp X) = \Phi(A)\exp(\phi(X)).$$

This says that in exponential coordinates near A, Φ is a composition of the linear map ϕ, the exponential mapping, and multiplication on the left by $\Phi(A)$, all of which are smooth. This shows that Φ is smooth near any point $A \in G$. $\square$

Corollary 2.35. *Suppose $G \subset \mathsf{GL}(n;\mathbb{C})$ is a matrix Lie group with Lie algebra $\mathfrak{g}$. Then, a matrix X is in $\mathfrak{g}$ if and only if there exists a smooth curve γ in $M_n(\mathbb{C})$ such that 1) $\gamma(t)$ lies in G for all t; 2) $\gamma(0) = I$; 3) $d\gamma/dt|_{t=0} = X$. Thus, $\mathfrak{g}$ is the tangent space at the identity to G.*

See Proposition C.3 for a description of the tangent space of an embedded submanifold in terms of derivatives of smooth curves.

Proof. If X is in $\mathfrak{g}$, then we may take $\gamma(t) = \exp(tX)$ and then $\gamma(0) = I$ and $d\gamma/dt|_{t=0} = X$. In the other direction, suppose that $\gamma(t)$ is a smooth curve in G with $\gamma(0) = I$. Then, by Theorem 2.27, $\log(\gamma(t))$ is in $\mathfrak{g}$ for all sufficiently small t. Now, $\mathfrak{g}$ is a real subspace of $M_n(\mathbb{C})$ and, therefore, also a topologically closed subset of $M_n(\mathbb{C})$. Thus, the quantity

$$\left.\frac{d\log(\gamma(t))}{dt}\right|_{t=0} = \lim_{h\to 0}\frac{\log(\gamma(h)) - 0}{h}$$

is again in $\mathfrak{g}$. However,

$$\log(\gamma(t)) = (\gamma(t) - I) - \frac{(\gamma(t) - I)^2}{2} + \frac{(\gamma(t) - I)^3}{3} + \cdots .$$

If we differentiate this term by term (it is not hard to see that this is permitted) and apply the product rule, all terms but the first will give zero. (For example, the derivative of the second term is $-\frac{1}{2}[(d\gamma/dt)(\gamma(t) - I) + (\gamma(t) - I)(d\gamma/dt)]$, which is equal to zero at $t = 0$.) Thus, we obtain that

$$\left.\frac{d\log(\gamma(t))}{dt}\right|_{t=0} = \left.\frac{d\gamma}{dt}\right|_{t=0} \in \mathfrak{g}.$$

□

2.8 Lie Algebras

We now consider the abstract notion of a Lie algebra, not necessarily given to us as the Lie algebra of a matrix Lie group. Proposition 2.37 shows that the Lie algebra of a matrix Lie group is indeed a Lie algebra in the abstract sense.

Definition 2.36. *A* ***finite-dimensional real or complex Lie algebra*** *is a finite-dimensional real or complex vector space* $\mathfrak{g}$*, together with a map* $[\cdot,\cdot]$ *from* $\mathfrak{g} \times \mathfrak{g}$ *into* $\mathfrak{g}$*, with the following properties:*

1. $[\cdot,\cdot]$ *is bilinear.*
2. $[X, Y] = -[Y, X]$ *for all* $X, Y \in \mathfrak{g}$.
3. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ *for all* $X, Y, Z \in \mathfrak{g}$.

Condition 2 is called "skew symmetry." Condition 3 is called the **Jacobi identity**. Note also that Condition 2 implies that $[X, X] = 0$ for all $X \in \mathfrak{g}$. We will deal only with finite-dimensional Lie algebras and will from now on interpret "Lie algebra" as "finite-dimensional Lie algebra."

It should be emphasized here that $\mathfrak{g}$ can be *any* vector space (not necessarily a space of matrices) and that the "bracket" operation $[\cdot,\cdot]$ can be *any* bilinear, skew-symmetric map that satisfies the Jacobi identity. In particular, $[X, Y]$ is not necessarily equal to $XY - YX$; indeed, the expression $XY - YX$ does not even make sense in general, since $\mathfrak{g}$ does not necessarily have a product operation defined on it. For example, let $\mathfrak{g} = \mathbb{R}^3$ and define $[x, y]$ to be $x \times y$, where $\times$ is the cross product (vector product). This operation is, clearly, bilinear and skew-symmetric, and it can be checked that is satisfies the Jacobi identity. There is, so far as I can see, no product operation "xy" on $\mathbb{R}^3$ such that $x \times y = xy - yx$.

Although the bracket operation in a Lie algebra does not have to be given to us as $[X, Y] = XY - YX$, it is possible to construct Lie algebras in this way. That is to say, if $\mathcal{A}$ is an associative algebra and we define $[\cdot,\cdot] : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ by $[X, Y] = XY - YX$, then this operation does, indeed, make $\mathcal{A}$ into a Lie algebra. This operation is clearly bilinear and skew-symmetric, and it is a simple computation to check, using the associativity of $\mathcal{A}$, the Jacobi identity. For any Lie algebra, the Jacobi identity means that the bracket operation *behaves as if* it were $XY - YX$, even if it is not actually defined this way. Indeed, it can be shown that every Lie algebra $\mathfrak{g}$ can be embedded into some associative algebra $\mathcal{A}$ in such a way that the bracket on $\mathfrak{g}$ corresponds to the operation $XY - YX$ in $\mathcal{A}$.

If $\mathfrak{g}$ is a Lie algebra, we can think of the bracket operation as making $\mathfrak{g}$ into an algebra in the general sense. This algebra, however, is not associative. The Jacobi identity is to be thought of as a substitute for associativity.

Proposition 2.37. *The space $M_n(\mathbb{R})$ of all $n \times n$ real matrices is a real Lie algebra with respect to the bracket operation $[A, B] = AB - BA$. The space $M_n(\mathbb{C})$ of all $n \times n$ complex matrices is a complex Lie algebra with respect to the same bracket operation.*

Let V be a finite-dimensional real or complex vector space, and let $\mathsf{gl}(V)$ denote the space of linear maps of V into itself. Then, $\mathsf{gl}(V)$ becomes a real or complex Lie algebra with the bracket operation $[A, B] = AB - BA$.

Proof. The only nontrivial point is the Jacobi identity. The only way to prove this is to write everything out and see, and this is best left to the reader. Note that each double bracket generates 4 terms, for a total of 12. Each of the six orderings of $\{X, Y, Z\}$ occurs twice, once with a plus sign and once with a minus sign. Note that the associativity of the matrix product is essential to the proof. $\square$

Definition 2.38. *A **subalgebra** of a real or complex Lie algebra $\mathfrak{g}$ is a subspace $\mathfrak{h}$ of $\mathfrak{g}$ such that $[H_1, H_2] \in \mathfrak{h}$ for all H_1 and $H_2 \in \mathfrak{h}$. If $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{h}$ is a real subspace of $\mathfrak{g}$ which is closed under brackets, then $\mathfrak{h}$ is said to be a **real subalgebra** of $\mathfrak{g}$.*

*If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, then a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is called a **Lie algebra homomorphism** if $\phi([X, Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$. If, in addition, ϕ is one-to-one and onto, then ϕ is called a **Lie algebra isomorphism**. A Lie algebra isomorphism of a Lie algebra with itself is called a **Lie algebra automorphism**.*

A subalgebra of a Lie algebra is, again, a Lie algebra. A real subalgebra of a complex Lie algebra is a real Lie algebra. The inverse of a Lie algebra isomorphism is, again, a Lie algebra isomorphism.

Proposition 2.39. *The Lie algebra $\mathfrak{g}$ of a matrix Lie group G is a real Lie algebra.*

Proof. By Theorem 2.18, $\mathfrak{g}$ is a real subalgebra of the space $M_n(\mathbb{C})$ of all complex matrices and is, thus, a real Lie algebra. $\square$

Theorem 2.40 (Ado). *Every finite-dimensional real Lie algebra is isomorphic to a subalgebra of* $\mathsf{gl}(n;\mathbb{R})$. *Every finite-dimensional complex Lie algebra is isomorphic to a complex subalgebra of* $\mathsf{gl}(n;\mathbb{C})$.

This deep theorem is proved, for example, in Varadarajan (1974). The proof is beyond the scope of this book and requires a careful examination of the structure of complex Lie algebras. The theorem tells us that every Lie algebra is (isomorphic to) a Lie algebra of matrices. This is in contrast to the situation for Lie groups, where most, but not all, Lie groups are matrix Lie groups—see Section C.3.

We now introduce the abstract Lie algebra version of the map "ad," which we introduced earlier for the Lie algebra of a matrix Lie group.

Definition 2.41. *Let* $\mathfrak{g}$ *be a Lie algebra. For* $X \in \mathfrak{g}$, *define a linear map* $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ *by*

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

*Thus, "*ad*" (i.e., the map* $X \to \mathrm{ad}_X$*) can be viewed as a linear map from* $\mathfrak{g}$ *into* $\mathsf{gl}(\mathfrak{g})$, *where* $\mathsf{gl}(\mathfrak{g})$ *denotes the space of linear operators from* $\mathfrak{g}$ *to* $\mathfrak{g}$.

Since $\mathrm{ad}_X(Y)$ is just $[X, Y]$, it might seem foolish to introduce the additional "ad" notation. However, thinking of $[X, Y]$ as a linear map in Y for each fixed X gives a somewhat different perspective. In any case, the "ad" notation is extremely useful in some situations. For example, instead of writing

$$[X, [X, [X, [X, Y]]]],$$

we can now write

$$(\mathrm{ad}_X)^4\,(Y).$$

This sort of notation will be essential in Chapter 3.

Proposition 2.42. *If* $\mathfrak{g}$ *is a Lie algebra, then*

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

that is, ad: $\mathfrak{g} \to \mathsf{gl}(\mathfrak{g})$ *is a Lie algebra homomorphism.*

Proof. Observe that

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

whereas

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

So, we want to show that

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

or, equivalently,

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

which is exactly the Jacobi identity. $\square$

2.8.1 Structure constants

Let $\mathfrak{g}$ be a finite-dimensional real or complex Lie algebra, and let $X_1, \ldots, X_n$ be a basis for $\mathfrak{g}$ (as a vector space). Then, for each i and j, $[X_i, X_j]$ can be written uniquely in the form

$$[X_i, X_j] = \sum_{k=1}^{n} c_{ijk} X_k.$$

The constants c_{ijk} are called the **structure constants** of $\mathfrak{g}$ (with respect to the chosen basis). Clearly, the structure constants determine the bracket operation on $\mathfrak{g}$. In some of the literature, the structure constants play an important role, although we will not have much necessity to use them in this book. (They appear mainly in Appendix D, where the quantities ε_{ijk} are the structure constants for the Lie algebra $\mathsf{so}(3)$.) In the physics literature, the structure constants are defined as $[X_i, X_j] = \sqrt{-1} \sum_k c_{ijk} X_k$, reflecting the factor of $\sqrt{-1}$ difference between the physics definition of the Lie algebra and our own.

The structure constants satisfy the following two conditions:

$$\begin{aligned} c_{ijk} + c_{jik} &= 0, \\ \sum_m (c_{ijm} c_{mkl} + c_{jkm} c_{mil} + c_{kim} c_{mjl}) &= 0 \end{aligned}$$

for all i, j, k, l. The first of these conditions comes from the skew symmetry of the bracket, and the second comes from the Jacobi identity. (The reader is invited to verify these conditions for himself.)

2.8.2 Direct sums

If $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are Lie algebras, we can define the **direct sum** of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ as follows. We consider the direct sum of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ in the vector space sense, and we define a bracket operation on $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ by

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

It is straightforward to verify that this operation satisfies the Jacobi identity and makes $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ into a Lie algebra. If $G_1 \subset \mathsf{GL}(n_1; \mathbb{C})$ and $G_2 \subset \mathsf{GL}(n_2; \mathbb{C})$ are matrix Lie groups and $G_1 \times G_2$ is their direct product (regarded as a subgroup of $\mathsf{GL}(n_1 + n_2; \mathbb{C})$ in the obvious way), then it is easily verified that the Lie algebra of $G_1 \times G_2$ is isomorphic to $\mathfrak{g}_1 \oplus \mathfrak{g}_2$.

2.9 The Complexification of a Real Lie Algebra

Definition 2.43. *If V is a finite-dimensional real vector space, then the* ***complexification*** *of V, denoted $V_{\mathbb{C}}$, is the space of formal linear combinations*

$$v_1 + iv_2,$$

with $v_1, v_2 \in V$*. This becomes a real vector space in the obvious way and becomes a complex vector space if we define*

$$i(v_1 + iv_2) = -v_2 + iv_1.$$

We could more pedantically define $V_{\mathbb{C}}$ to be the space of ordered pairs (v_1, v_2) with $v_1, v_2 \in V$, but this is notationally cumbersome. It is straightforward to verify that the above definition really makes $V_{\mathbb{C}}$ into a complex vector space. We will regard V as a real subspace of $V_{\mathbb{C}}$ in the obvious way.

Proposition 2.44. *Let* $\mathfrak{g}$ *be a finite-dimensional real Lie algebra and* $\mathfrak{g}_{\mathbb{C}}$ *its complexification (as a real vector space). Then, the bracket operation on* $\mathfrak{g}$ *has a unique extension to* $\mathfrak{g}_{\mathbb{C}}$ *which makes* $\mathfrak{g}_{\mathbb{C}}$ *into a complex Lie algebra. The complex Lie algebra* $\mathfrak{g}_{\mathbb{C}}$ *is called the* ***complexification*** *of the real Lie algebra* $\mathfrak{g}$.

Proof. The uniqueness of the extension is obvious, since if the bracket operation on $\mathfrak{g}_{\mathbb{C}}$ is to be bilinear, then it must be given by

$$[X_1 + iX_2, Y_1 + iY_2] = ([X_1, Y_1] - [X_2, Y_2]) + i\,([X_1, Y_2] + [X_2, Y_1])\,. \tag{2.20}$$

To show existence, we must now check that (2.20) is really bilinear and skew symmetric and that it satisfies the Jacobi identity. It is clear that (2.20) is *real* bilinear, and skew-symmetric. The skew symmetry means that if (2.20) is complex linear in the first factor, it is also complex linear in the second factor. Thus, we need only show that

$$[i(X_1 + iX_2), Y_1 + iY_2] = i\,[X_1 + iX_2, Y_1 + iY_2]\,. \tag{2.21}$$

The left-hand side of (2.21) is

$$[-X_2 + iX_1, Y_1 + iY_2] = (-\,[X_2, Y_1] - [X_1, Y_2]) + i\,([X_1, Y_1] - [X_2, Y_2])\,,$$

whereas the right-hand side of (2.21) is

$$\begin{aligned} &i\,\{([X_1, Y_1] - [X_2, Y_2]) + i\,([X_2, Y_1] + [X_1, Y_2])\} \\ &= (-\,[X_2, Y_1] - [X_1, Y_2]) + i\,([X_1, Y_1] - [X_2, Y_2])\,, \end{aligned}$$

and, indeed, these are equal.

It remains to check the Jacobi identity. Of course, the Jacobi identity holds if X, Y, and Z are in $\mathfrak{g}$. However, observe that the expression on the left-hand side of the Jacobi identity is (complex!) linear in X for fixed Y and Z. It follows that the Jacobi identity holds if X is in $\mathfrak{g}_{\mathbb{C}}$, and Y and Z are in $\mathfrak{g}$. The same argument then shows that we can extend to Y in $\mathfrak{g}_{\mathbb{C}}$, and then to Z in $\mathfrak{g}_{\mathbb{C}}$. Thus, the Jacobi identity holds in $\mathfrak{g}_{\mathbb{C}}$. ∎

Proposition 2.45. *The Lie algebras* $\mathsf{gl}(n;\mathbb{C})$*,* $\mathsf{sl}(n;\mathbb{C})$*,* $\mathsf{so}(n;\mathbb{C})$*, and* $\mathsf{sp}(n;\mathbb{C})$ *are complex Lie algebras. In addition, we have the following isomorphisms of complex Lie algebras:*

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

Proof. From the computations in the previous section, we see easily that the specified Lie algebras are, in fact, complex subalgebras of $\mathsf{gl}(n;\mathbb{C})$ and hence are complex Lie algebras.

Now, $\mathsf{gl}(n;\mathbb{C})$ is the space of all $n\times n$ complex matrices, whereas $\mathsf{gl}(n;\mathbb{R})$ is the space of all $n\times n$ real matrices. Clearly, then, every $X\in\mathsf{gl}(n;\mathbb{C})$ can be written uniquely in the form X_1+iX_2, with $X_1,X_2\in\mathsf{gl}(n;\mathbb{R})$. This gives us a complex vector space isomorphism of $\mathsf{gl}(n;\mathbb{R})_{\mathbb{C}}$ with $\mathsf{gl}(n;\mathbb{C})$, and it is a triviality to check that this is a Lie algebra isomorphism.

On the other hand, $\mathsf{u}(n)$ is the space of all $n\times n$ complex skew-self-adjoint matrices. However, if X is any $n\times n$ complex matrix, then

$$X=\frac{X-X^*}{2}+i\frac{X+X^*}{2i},$$

where $(X-X^*)/2$ and $(X+X^*)/2i$ are both skew. Thus, X can be written as a skew matrix plus i times a skew matrix, and it is easy to see that this decomposition is unique. Thus, every X in $\mathsf{gl}(n;\mathbb{C})$ can be written uniquely as X_1+iX_2, with X_1 and X_2 in $\mathsf{u}(n)$. It follows that $\mathsf{u}(n)_{\mathbb{C}}\cong\mathsf{gl}(n;\mathbb{C})$. If X has trace zero, then so do X_1 and X_2, which shows that $\mathsf{su}(n)_{\mathbb{C}}\cong\mathsf{sl}(n;\mathbb{C})$.

The verification of the remaining isomorphisms is similar and is left as an exercise to the reader. □

Note that $\mathsf{u}(n)_{\mathbb{C}}\cong\mathsf{gl}(n;\mathbb{R})_{\mathbb{C}}\cong\mathsf{gl}(n;\mathbb{C})$. However, $\mathsf{u}(n)$ is *not* isomorphic to $\mathsf{gl}(n;\mathbb{R})$, except when $n=1$. The real Lie algebras $\mathsf{u}(n)$ and $\mathsf{gl}(n;\mathbb{R})$ are called **real forms** of the complex Lie algebra $\mathsf{gl}(n;\mathbb{C})$. A given complex Lie algebra may have several nonisomorphic real forms. See Exercise 17.

Physicists do not always clearly distinguish between a matrix Lie group and its (real) Lie algebra, or between a real Lie algebra and its complexification. Thus, for example, some references in the physics literature to $\mathsf{SU}(2)$ actually refer to the complexified Lie algebra, $\mathsf{sl}(2;\mathbb{C})$.

2.10 Exercises

1. The Schwarz inequality from elementary analysis tells us that for all $u=(u_1,\ldots,u_n)$ and $v=(v_1,\ldots,v_n)$ in $\mathbb{C}^n$, we have

$$|u_1v_1 + \cdots + u_nv_n|^2 \leq \left(\sum_{k=1}^{n} |u_k|^2\right)\left(\sum_{k=1}^{n} |v_k|^2\right).$$

Use this to verify that $\|XY\| \leq \|X\| \, \|Y\|$ for all $X, Y \in M_n(\mathbb{C})$, where the norm $\|X\|$ of a matrix X is defined by (2.2).

2. Show that for $X \in M_n(\mathbb{C})$ and any orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathbb{C}^n$, $\|X\|^2 = \sum_{j,k=1}^{n} |\langle u_j, Xu_k\rangle|^2$, where $\|X\|$ is defined by (2.2). Now show that if v is an eigenvector for X with eigenvalue λ, then $|\lambda| \leq \|X\|$.
3. *The product rule.* Recall that a matrix-valued function $A(t)$ is said to be smooth if each $A_{ij}(t)$ is smooth. The derivative of such a function is defined as
$$\left(\frac{dA}{dt}\right)_{ij} = \frac{dA_{ij}}{dt}$$
or, equivalently,
$$\frac{d}{dt}A(t) = \lim_{h \to 0} \frac{A(t+h) - A(t)}{h}.$$
Let $A(t)$ and $B(t)$ be two such functions. Prove that $A(t)B(t)$ is again smooth and that
$$\frac{d}{dt}\left[A(t)B(t)\right] = \frac{dA}{dt}B(t) + A(t)\frac{dB}{dt}.$$
4. Show that for all $X \in M_n(\mathbb{C})$,
$$\lim_{m \to \infty} \left[I + \frac{X}{m}\right]^m = e^X.$$
5. Using Theorem B.7, show that every $n \times n$ complex matrix A is the limit of a sequence of diagonalizable matrices.
Hint: If the characteristic polynomial of A has n distinct roots, then A is diagonalizable.
6. Show that every 2×2 matrix X with trace zero satisfies
$$X^2 = -\det(X)I.$$
If X is 2×2 with trace zero, show by direct calculation using the power series for the exponential that
$$e^X = \cos\left(\sqrt{\det X}\right)I + \frac{\sin\sqrt{\det X}}{\sqrt{\det X}}X. \tag{2.22}$$
Use this to give an alternative derivation of the result in (2.7).
Notes: Since the functions $\cos\theta$ and $\sin\theta/\theta$ are even functions of θ, the value of (2.22) is independent of the choice of the square root of $\det X$. The value of the coefficient of X in (2.22) is to be interpreted as 1 when $\det X = 0$, in accordance with the limit $\lim_{\theta \to 0} \sin\theta/\theta = 1$.

7. Use the result of Exercise 6 to compute the exponential of the matrix
$$X = \begin{pmatrix} 4 & 3 \\ -1 & 2 \end{pmatrix}.$$
Hint: Write X as the sum of a multiple of the identity and a matrix with trace zero.
8. A matrix A is said to be **unipotent** if $A - I$ is nilpotent (i.e., if A is of the form $A = I + N$, with N nilpotent). Note that $\log A$ is defined whenever A is unipotent, because the series in Definition 2.6 terminates.
(a) Show that if A is unipotent, then $\log A$ is nilpotent.
(b) Show that if X is nilpotent, then e^X is unipotent.
(c) Show that if A is unipotent, then $\exp(\log A) = A$ and that if X is nilpotent, then $\log(\exp X) = X$.
Hint: Let $A(t) = I + t(A - I)$. Show that $\exp(\log(A(t)))$ depends polynomially on t and that $\exp(\log(A(t))) = A(t)$ for all sufficiently small t.
9. Show that every invertible $n \times n$ matrix A can be written as $A = e^X$ for some $X \in M_n(\mathbb{C})$.
Hint: Theorem B.5 implies that A is similar to a block-diagonal matrix in which each block is of the form $\lambda I + N_\lambda$, with N_λ being nilpotent. Use this result and Exercise 8.
10. Give an example of a matrix Lie group G and a matrix X such that $e^X \in G$, but $X \notin \mathfrak{g}$.
11. Suppose G is a matrix Lie group in $\mathsf{GL}(n; \mathbb{C})$ and let $\mathfrak{g}$ be its Lie algebra. Suppose that A is in G and that $\|A - I\| < 1$, so that the power series for $\log A$ is convergent. Is it necessarily the case that $\log A$ is in $\mathfrak{g}$? Prove or give a counterexample.
12. Show that two isomorphic matrix Lie groups have isomorphic Lie algebras.
13. *The Lie algebra* $\mathsf{so}(3;1)$. Write out explicitly the general form of a 4×4 real matrix in $\mathsf{so}(3;1)$.
14. Verify directly that Proposition 2.17 and Theorem 2.18 hold for the Lie algebra of $\mathsf{SU}(n)$.
15. *The Lie algebra* $\mathsf{su}(2)$. Show that the following matrices form a basis for the real Lie algebra $\mathsf{su}(2)$:
$$E_1 = \tfrac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix};\ E_2 = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix};\ E_3 = \tfrac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$
Compute $[E_1, E_2]$, $[E_2, E_3]$, and $[E_3, E_1]$. Show that there is an invertible linear map $\phi : \mathsf{su}(2) \to \mathbb{R}^3$ such that $\phi([X, Y]) = \phi(X) \times \phi(Y)$ for all $X, Y \in \mathsf{su}(2)$, where $\times$ denotes the cross product (vector product) on $\mathbb{R}^3$.
16. *The Lie algebras* $\mathsf{su}(2)$ *and* $\mathsf{so}(3)$. Show that the real Lie algebras $\mathsf{su}(2)$ and $\mathsf{so}(3)$ are isomorphic.
Note: Nevertheless, the corresponding *groups* $\mathsf{SU}(2)$ and $\mathsf{SO}(3)$ are not isomorphic. (Rather, $\mathsf{SO}(3)$ is isomorphic to $\mathsf{SU}(2)/\{I, -I\}$.)

17. *The Lie algebras* $\mathsf{su}(2)$ *and* $\mathsf{sl}(2;\mathbb{R})$. Show that $\mathsf{su}(2)$ and $\mathsf{sl}(2;\mathbb{R})$ are not isomorphic Lie algebras, even though $\mathsf{su}(2)_{\mathbb{C}} \cong \mathsf{sl}(2;\mathbb{R})_{\mathbb{C}} \cong \mathsf{sl}(2;\mathbb{C})$.
    Hint: Using Exercise 15, show that $\mathsf{su}(2)$ has no two-dimensional subalgebras.
18. Let G be a matrix Lie group and let $\mathfrak{g}$ be its Lie algebra. For each $A \in G$, show that Ad_A is a Lie algebra automorphism of $\mathfrak{g}$.
19. ("Ad" and "ad") Let X and Y be $n \times n$ matrices. Show by induction that

$$(\mathrm{ad}_X)^m (Y) = \sum_{k=0}^{m} \binom{m}{k} X^k Y (-X)^{m-k},$$

where

$$(\mathrm{ad}_X)^m (Y) = \underbrace{[X, \cdots [X, [X, Y]] \cdots]}_{m}.$$

Now, show by direct computation that

$$e^{\mathrm{ad}_X}(Y) = \mathrm{Ad}_{e^X}(Y) = e^X Y e^{-X}.$$

Assume that it is legal to multiply power series term by term. (This result was obtained indirectly in Proposition 2.25.)
    Hint: Recall that Pascal's Triangle gives a relationship between numbers of the form $\binom{m+1}{k}$ and numbers of the form $\binom{m}{k}$.
20. If $\mathfrak{g}$ is a Lie algebra, then a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is called an **ideal** if $[X, H] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and $H \in \mathfrak{h}$. If $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism, show that $\ker \phi$ is an ideal in $\mathfrak{g}_1$.
21. Classify up to isomorphism all one-dimensional and two-dimensional real Lie algebras. (There is one isomorphism class of one-dimensional algebras and two isomorphism classes of two-dimensional algebras.)
22. Show that for any Lie algebra $\mathfrak{g}$ and any X in $\mathfrak{g}$, ad_X is a derivation of $\mathfrak{g}$; that is,

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

for all Y and Z in $\mathfrak{g}$.
23. *The complexification of a real Lie algebra*. Let $\mathfrak{g}$ be a real Lie algebra, $\mathfrak{g}_{\mathbb{C}}$ its complexification, and $\mathfrak{h}$ an arbitrary complex Lie algebra. Show that every real Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$ extends uniquely to a complex Lie algebra homomorphism of $\mathfrak{g}_{\mathbb{C}}$ into $\mathfrak{h}$. (This is the **universal property** of the complexification of a real Lie algebra. This property can be used as an alternative definition of the complexification.)
24. If $\mathfrak{g}$ is a Lie algebra, the **center** of $\mathfrak{g}$ is the set of all $Z \in \mathfrak{g}$ such that $[X, Z] = 0$ for all $X \in \mathfrak{g}$. Show that the center of $\mathfrak{g}$ is an ideal (as defined in Exercise 20).
25. Suppose that G is a connected, commutative matrix Lie group with Lie algebra $\mathfrak{g}$. Show that the exponential mapping for G maps $\mathfrak{g}$ onto G.

26. *The exponential mapping for the Heisenberg group*. Show that the exponential mapping from the Lie algebra of the Heisenberg group to the Heisenberg group is one-to-one and onto.
27. *The exponential mapping for* $\mathsf{U}(n)$. Show that the exponential mapping from $\mathsf{u}(n)$ to $\mathsf{U}(n)$ is onto, but not one-to-one. (Note that this shows that $\mathsf{U}(n)$ is connected.)
    Hint: Every unitary matrix has an orthonormal basis of eigenvectors.
28. Consider the space $\mathsf{gl}(n;\mathbb{C})$ of all $n \times n$ complex matrices. As usual, for $X \in \mathsf{gl}(n;\mathbb{C})$, define $\mathrm{ad}_X : \mathsf{gl}(n;\mathbb{C}) \to \mathsf{gl}(n;\mathbb{C})$ by $\mathrm{ad}_X(Y) = [X, Y]$. Suppose that X is a diagonalizable matrix. Show, then, that ad_X is diagonalizable as an operator on $\mathsf{gl}(n;\mathbb{C})$.
    Hint: Consider first the case where X is actually diagonal.
    Note: The problem of diagonalizing ad_X is an important one that we will encounter again in Chapter 6, when we consider semisimple Lie algebras.
29. Show explicitly that $\exp : \mathsf{so}(3) \to \mathsf{SO}(3)$ is onto.
    Hint: Using Exercise 16 from Chapter 1, show that in a suitable orthonormal basis, R is of the form

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}.$$

30. *The exponential mapping for* $\mathsf{SL}(2;\mathbb{R})$. Show that the image of the exponential mapping for $\mathsf{SL}(2;\mathbb{R})$ consists of precisely those matrices $A \in \mathsf{SL}(2;\mathbb{R})$ such that $\operatorname{trace}(A) > -2$, together with the matrix $-I$ (which has trace -2). To do this, consider the possibilities for the eigenvalues of a matrix in the Lie algebra $\mathsf{sl}(2;\mathbb{R})$ and in the group $\mathsf{SL}(2;\mathbb{R})$. In the Lie algebra, show that the eigenvalues are of the form $(\lambda, -\lambda)$ or $(i\lambda, -i\lambda)$, with λ real. In the group, show that the eigenvalues are of the form $(a, 1/a)$ or $(-a, -1/a)$, with a real and positive, or of the form $(e^{i\theta}, e^{-i\theta})$, with θ real. The case of a repeated eigenvalue ($(0,0)$ in the Lie algebra and $(1,1)$ or $(-1,-1)$ in the group) will have to be treated separately using the Jordan canonical form (Section B.4).
    Show that the image of the exponential mapping is not dense in $\mathsf{SL}(2;\mathbb{R})$.
31. Determine the image of the exponential mapping for $\mathsf{SL}(2;\mathbb{C})$. Is the image of the exponential mapping dense in $\mathsf{SL}(2;\mathbb{C})$?

http://www.springer.com/978-0-387-40122-5

Lie Groups, Lie Algebras, and Representations
An Elementary Introduction
Hall, B.
2003, XIV, 354 p., Hardcover
ISBN: 978-0-387-40122-5