Chapter 1

Matrices

In this chapter, we define quaternionic numbers and discuss basic algebraic properties of matrices, including the correspondence between matrices and linear transformations. We begin with a visual example that motivates the topic of matrix groups.

1. Rigid motions of the sphere: a motivating example

The simplest interesting matrix group, called $SO(3)$, can be described in the following (admittedly imprecise) way:

$$SO(3) = \text{all positions of a globe on a fixed stand.}$$

Three elements of $SO(3)$ are pictured in Figure 1. Though the globe always occupies the same place in space, the three elements differ in the directions where various countries face.

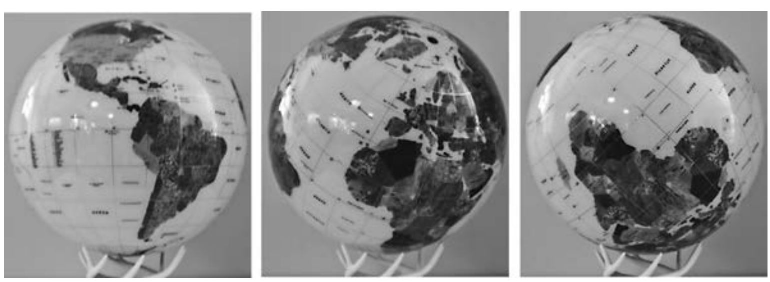


Figure 1. Three elements of $SO(3)$.

Let's (arbitrarily) call the first picture "the identity". Every other element of $SO(3)$ is achieved, starting with the identity, by physically moving the globe in some way. $SO(3)$ becomes a group under composition of motions (since different motions might place the globe in the same position, think about why this group operation is well-defined). Several questions come to mind.

Question 1.1. *Is $SO(3)$ an abelian group?*

The North Pole of the globe faces up in the identity position. Rotating the globe around the axis through the North Pole provides a "circle's worth" of elements of $SO(3)$ for which the North Pole faces up. Similarly, there is a circle's worth of elements of $SO(3)$ for which the North Pole is located as in picture 2, or at any other point of the globe. Any element of $SO(3)$ is achieved, starting with the identity, by first moving the North Pole to the correct position and then rotating about the axis through its new position. It is therefore natural to ask:

Question 1.2. *Is there a natural bijection between $SO(3)$ and the product $S^2 \times S^1 = \{(p, \theta) \mid p \in S^2, \theta \in S^1\}$?*

Here S^2 denotes the sphere (the surface of the globe) and S^1 denotes the circle, both special cases of the general definition of an **n-dimensional sphere**:

$$S^n = \{(x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots + x_{n+1}^2 = 1\}.$$

Graphics programmers, who model objects moving and spinning in space, need an efficient way to represent the rotation of such objects. A bijection between $SO(3)$ and $S^2 \times S^1$ would help, allowing any rotation to be coded using only three real numbers – two that locate a point of S^2 (latitude and longitude) and one angle that locates a point of S^1. If no such bijection exists, can we nevertheless understand the shape of $SO(3)$ sufficiently well to somehow parameterize its elements via three real numbers?

One is tempted to refer to elements of $SO(3)$ as "rotations" of the sphere, but perhaps there are motions more complicated than rotations.

Question 1.3. *Can every element of $SO(3)$ be achieved, starting with the identity, by rotating through some angle about some single axis?*

If so, then for any element of $SO(3)$, there must be a pair of antipodal points of the globe in their identity position.

You might borrow your roommate's basketball and use visual intuition to guess the correct answers to Questions 1.1, 1.2 and 1.3. But our definition of $SO(3)$ is probably too imprecise to lead to rigorous proofs of your answers. We will return to these questions after developing the algebraic background needed to define $SO(3)$ in a more precise way, as a group of matrices.

2. Fields and skew-fields

A matrix is an array of numbers, but what type of numbers? Matrices of real numbers and matrices of complex numbers are familiar. Are there other good choices? We need to add, multiply and invert matrices, so we must choose a number system with a notion of addition, multiplication, and division; in other words, we must choose a field or a skew-field.

Definition 1.4. *A **skew-field** is a set, $\mathbb{K}$, together with operations called addition (denoted "$+$") and multiplication (denoted "$\cdot$") satisfying:*

(1) $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(b + c) \cdot a = b \cdot a + c \cdot a$.

(2) $\mathbb{K}$ *is an abelian group under addition, with identity denoted as "0".*

(3) $\mathbb{K} - \{0\}$ *is a group under multiplication, with identity denoted as "1".*

*A skew-field in which multiplication is commutative ($a \cdot b = b \cdot a$ for all $a, b \in \mathbb{K}$) is called a **field**.*

The real numbers, $\mathbb{R}$, and the rational numbers, $\mathbb{Q}$, are fields. The plane $\mathbb{R}^2$ is NOT a field under the operations of component-wise addition and multiplication:

$$(a,b) + (c,d) = (a+c, b+d)$$
$$(a,b) \cdot (c,d) = (ac, bd),$$

because, for example, the element $(5,0)$ does not have a multiplicative inverse; that is, no element times $(5,0)$ equals $(1,1)$, which is the only possible multiplicative identity element. A similar argument shows that for $n > 1$, $\mathbb{R}^n$ is not a field under component-wise addition and multiplication.

In order to make $\mathbb{R}^2$ into a field, we use component-wise addition, but we define a more clever multiplication operation:

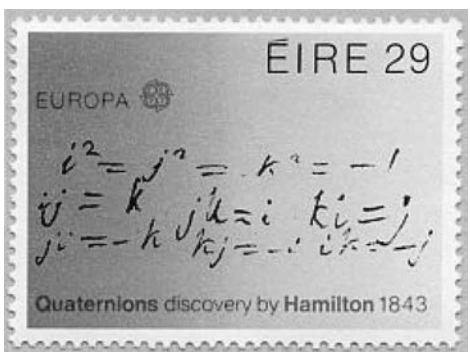
$$(a,b) \cdot (c,d) = (ac - bd, ad + bc).$$

If we denote $(a,b) \in \mathbb{R}^2$ symbolically as $a+b\mathbf{i}$, then this multiplication operation becomes familiar complex multiplication:

$$(a + b\mathbf{i}) \cdot (c + d\mathbf{i}) = (ac - bd) + (ad + bc)\mathbf{i}.$$

It is straightforward to check that $\mathbb{R}^2$ is a field under these operations; it is usually denoted $\mathbb{C}$ and called **the complex numbers**.

3. The quaternions



Is it possible to contrive a multiplication operation which, together with component-wise addition, makes $\mathbb{R}^n$ into a skew-field for $n > 2$? This is an important and difficult question. In 1843 Hamilton discovered that the answer is yes for $n = 4$.

To describe this multiplication rule, we will denote an element $(a, b, c, d) \in \mathbb{R}^4$ symbolically as $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. We then define a multiplication rule for the symbols $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. The symbol "1" acts as expected:

$$\mathbf{i} \cdot 1 = 1 \cdot \mathbf{i} = \mathbf{i}, \qquad \mathbf{j} \cdot 1 = 1 \cdot \mathbf{j} = \mathbf{j} \qquad \mathbf{k} \cdot 1 = 1 \cdot \mathbf{k} = \mathbf{k}.$$

The other three symbols square to -1:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1.$$

Finally, the product of two of $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ equals plus or minus the third:

$$\begin{array}{ccc} \mathbf{i} \cdot \mathbf{j} = \mathbf{k}, & \mathbf{j} \cdot \mathbf{k} = \mathbf{i}, & \mathbf{k} \cdot \mathbf{i} = \mathbf{j}, \\ \mathbf{j} \cdot \mathbf{i} = -\mathbf{k}, & \mathbf{k} \cdot \mathbf{j} = -\mathbf{i}, & \mathbf{i} \cdot \mathbf{k} = -\mathbf{j}. \end{array}$$

This sign convention can be remembered using Figure 2.

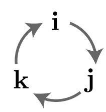


Figure 2. The quaternionic multiplication rule.

This multiplication rule for $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ extends linearly to a multiplication on all of $\mathbb{R}^4$. For example,

$$\begin{aligned} (2 + 3\mathbf{k}) \cdot (\mathbf{i} + 7\mathbf{j}) &= 2\mathbf{i} + 14\mathbf{j} + 3\mathbf{k}\mathbf{i} + 21\mathbf{k}\mathbf{j} \\ &= 2\mathbf{i} + 14\mathbf{j} + 3\mathbf{j} - 21\mathbf{i} \\ &= -19\mathbf{i} + 17\mathbf{j}. \end{aligned}$$

The product of two arbitrary elements has the following formula:

$$\begin{aligned} (1.1) \quad & (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \cdot (x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k}) \\ &= (ax - by - cz - dw) + (ay + bx + cw - dz)\mathbf{i} \\ &\quad + (az + cx + dy - bw)\mathbf{j} + (aw + dx + bz - cy)\mathbf{k}. \end{aligned}$$

The set $\mathbb{R}^4$, together with component-wise addition and the above-described multiplication operation, is denoted as $\mathbb{H}$ and called the **quaternions**. The quaternions have proven to be fundamental in several areas of math and physics. They are almost as important and as natural as the real and complex numbers.

To prove that $\mathbb{H}$ is a skew-field, the only difficult step is verifying that every non-zero element has a multiplicative inverse. For this, it is useful to define the **conjugate** and the **norm** of an arbitrary element $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$ as follows:

$$\bar{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$$
$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

It is straightforward to check that $q \cdot \bar{q} = \bar{q} \cdot q = |q|^2$ and therefore that $\frac{\bar{q}}{|q|^2}$ is a multiplicative inverse of q.

The rule for multiplying two quaternions with no $\mathbf{j}$ or $\mathbf{k}$ components agrees with our multiplication rule in $\mathbb{C}$. We therefore have skew-field inclusions:

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}.$$

Any real number commutes with every element of $\mathbb{H}$. In Exercise 1.18, you will show that only real numbers have this property. In particular, every non-real complex number fails to commute with some elements of $\mathbb{H}$.

Any complex number can be expressed as $z = a + b\mathbf{i}$ for some $a, b \in \mathbb{R}$. Similarly, any quaternion can be expressed as $q = z + w\mathbf{j}$ for some $z, w \in \mathbb{C}$, since:

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} = (a + b\mathbf{i}) + (c + d\mathbf{i})\mathbf{j}.$$

This analogy between $\mathbb{R} \subset \mathbb{C}$ and $\mathbb{C} \subset \mathbb{H}$ is often useful.

In this book, the elements of matrices are always either real, complex, or quaternionic numbers. Other fields, like $\mathbb{Q}$ or the finite fields, are used in other branches of mathematics but for our purposes would lead to a theory of matrices with insufficient geometric structure. We want groups of matrices to have algebraic and geometric properties, so we restrict to skew-fields that look like $\mathbb{R}^n$ for some n. This way, groups of matrices are subsets of Euclidean spaces and therefore inherit geometric notions like distances and tangent vectors.

But is there a multiplication rule that makes $\mathbb{R}^n$ into a skew-field for values of n other than $1, 2$ and 4? Do other (substantially different) multiplication rules for $\mathbb{R}^1, \mathbb{R}^2$ and $\mathbb{R}^4$ exist? Can $\mathbb{R}^4$ be made into a field rather than just a skew-field? The answer to all of these questions is NO. More precisely, Frobenius proved in 1877 that

$\mathbb{R}, \mathbb{C}$ and $\mathbb{H}$ are the only associative real division algebras, up to the natural notion of equivalence [**4**].

Definition 1.5. *An associative* ***real division algebra*** *is a real vector space, $\mathbb{K}$, with a multiplication rule, that is a skew-field under vector-addition and multiplication, such that for all $a \in \mathbb{R}$ and all $q_1, q_2 \in \mathbb{K}$:*

$$a(q_1 \cdot q_2) = (aq_1) \cdot q_2 = q_1 \cdot (aq_2).$$

The final hypothesis relates multiplication and scalar multiplication. It insures that $\mathbb{K}$ has a subfield isomorphic to $\mathbb{R}$, namely, all scalar multiples of the multiplicative identity 1.

We will not prove Frobenius' theorem; we require it only for reassurance that we are not omitting any important number systems from our discussion. There is an important multiplication rule for $\mathbb{R}^8$, called **octonian** multiplication, but it is not associative, so it makes $\mathbb{R}^8$ into something weaker than a skew-field. We will not consider the octonians.

In this book, $\mathbb{K}$ always denotes one of $\{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$, except where stated otherwise.

4. Matrix operations

In this section, we briefly review basic notation and properties of matrices. Let $M_{m,n}(\mathbb{K})$ denote the set of all m by n matrices with entries in $\mathbb{K}$. For example,

$$M_{2,3}(\mathbb{C}) = \left\{ \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \end{pmatrix} \middle| z_{ij} \in \mathbb{C} \right\}.$$

Denote the space $M_{n,n}(\mathbb{K})$ of **square matrices** as simply $M_n(\mathbb{K})$. If $A \in M_{m,n}(\mathbb{K})$, then A_{ij} denotes the element in row i and column j of A.

Addition of same-dimension matrices is defined component-wise, so that

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

The product of $A \in M_{m,n}(\mathbb{K})$ and $B \in M_{n,l}(\mathbb{K})$ is the element $AB \in M_{m,l}(\mathbb{K})$ defined by the familiar formula:

$$(AB)_{ij} = (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } B) = \sum_{s=1}^{n} A_{is} B_{sj}. \tag{1.2}$$

Matrix multiplication is not generally commutative.

Denote a **diagonal matrix** as in this example:

$$\text{diag}(1,2,3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

The **identity matrix** is:

$$I = \text{diag}(1, ..., 1).$$

The **transpose** of $A \in M_{m,n}(\mathbb{K})$ is the matrix $A^T \in M_{n,m}$ obtained by interchanging the rows and columns of A, so that:

$$(A^T)_{ij} = A_{ji}.$$

For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

It is straightforward to check that if $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ then

$$(A \cdot B)^T = B^T \cdot A^T \tag{1.3}$$

for any matrices A and B of compatible dimensions to be multiplied.

Matrix multiplication and addition interact as follows:

Proposition 1.6. *For all $A, B, C \in M_n(\mathbb{K})$,*

(1) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.
(2) $(A + B) \cdot C = A \cdot C + B \cdot C$ *and* $C \cdot (A + B) = C \cdot A + C \cdot B$.
(3) $A \cdot I = I \cdot A = A$.

The **trace** of a square matrix $A \in M_n(\mathbb{K})$ is defined as the sum of its diagonal entries:

$$\text{trace}(A) = A_{11} + \cdots + A_{nn}.$$

When $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, we have the familiar property for $A, B \in M_n(\mathbb{K})$:

$$\text{trace}(AB) = \text{trace}(BA). \tag{1.4}$$

Since multiplication in $\mathbb{H}$ is not commutative, this property is false even in $M_1(\mathbb{H})$.

When $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, the **determinant** function,

$$\det : M_n(\mathbb{K}) \to \mathbb{K},$$

is familiar. It can be defined recursively by declaring that the determinant of $A \in M_1(\mathbb{K})$ equals its single element, and the determinant of $A \in M_{n+1}(\mathbb{K})$ is defined in terms of determinants of elements of $M_n(\mathbb{K})$ by the expansion of minors formula:

$$\det(A) = \sum_{j=1}^{n+1} (-1)^{j+1} \cdot A_{1j} \cdot \det(A[1, j]), \tag{1.5}$$

where $A[i, j] \in M_n(\mathbb{K})$ is the matrix obtained by crossing out row i and column j from A. For example,

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} [2, 1] = \begin{pmatrix} b & c \\ h & i \end{pmatrix}.$$

Thus, the determinant of a 3×3 matrix is:

$$\begin{aligned} \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= a \cdot \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \cdot \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} \\ &\quad + c \cdot \det \begin{pmatrix} d & e \\ g & h \end{pmatrix} \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \\ &= aei + bfg + cdh - (afh + bdi + ceg). \end{aligned}$$

It is clear that $\det(I) = 1$. In a linear algebra course, one proves that for all $A, B \in M_n(\mathbb{K})$,

$$\det(A \cdot B) = \det(A) \cdot \det(B). \tag{1.6}$$

We postpone defining the determinant of a quaternionic matrix until the next chapter. Exercise 1.5 at the end of this chapter demonstrates why Equation 1.5 is insufficient when $\mathbb{K} = \mathbb{H}$.

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. When $a \in \mathbb{K}$ and $A \in M_{n,m}(\mathbb{K})$, we define $a \cdot A \in M_{n,m}(\mathbb{K})$ to be the result of left-multiplying the elements of A by a:

$$(a \cdot A)_{ij} = a \cdot A_{ij}.$$

This operation is called **left scalar multiplication**. The operations of matrix addition and left scalar multiplication make $M_{n,m}(\mathbb{K})$ into a *left vector space* over $\mathbb{K}$.

Definition 1.7. *A* ***left vector space*** *over a skew-field $\mathbb{K}$ is a set M with an addition operation from $M \times M$ to M (denoted $A, B \mapsto A+B$) and scalar multiplication operation from $\mathbb{K} \times M$ to M (denoted $a, A \mapsto a \cdot A$) such that M is an abelian group under addition, and for all $a, b \in \mathbb{K}$ and all $A, B \in M$,*

(1) $a \cdot (b \cdot A) = (a \cdot b) \cdot A.$
(2) $1 \cdot A = A.$
(3) $(a + b) \cdot A = a \cdot A + b \cdot A.$
(4) $a \cdot (A + B) = a \cdot A + a \cdot B.$

This exactly matches the familiar definition of a vector space. Familiar terminology for vector spaces over fields, like **subspaces**, **bases**, **linear independence**, and **dimension**, make sense for left vector spaces over skew-fields. For example:

Definition 1.8. *A subset W of a left vector space V over a skew-field $\mathbb{K}$ is called a* ***$\mathbb{K}$-subspace*** *(or just a* ***subspace****) if for all $a, b \in \mathbb{K}$ and all $A, B \in W$, $a \cdot A + b \cdot B \in W$.*

If we had instead chosen *right* scalar multiplication in $M_{n,m}(\mathbb{K})$, defined as $(A \cdot a)_{ij} = A_{ij} \cdot a$, then $M_{n,m}(\mathbb{K})$ would have become a *right vector space* over $\mathbb{K}$. In a right vector space, scalar multiplication is denoted $a, A \mapsto A \cdot a$. Properties (2) through (4) of Definition 1.7 must be re-written to reflect this notational change. Property (1) is special because the change is more than just notational:

(1') $(A \cdot b) \cdot a = A \cdot (b \cdot a).$

Do you see the difference? The net effect of multiplying A by b and then by a is to multiply A by ab in a left vector space, or by ba in a right vector space.

When $\mathbb{K}$ is a field, the difference between a left and a right vector space over $\mathbb{K}$ is an irrelevant notational distinction, so one speaks simply of "vector spaces". But when $\mathbb{K} = \mathbb{H}$, it makes an essential difference that we are henceforth adopting the convention of left scalar multiplication, and thereby choosing to regard $M_{n,m}(\mathbb{H})$ as a left vector space over $\mathbb{H}$.

5. Matrices as linear transformations

One cornerstone of a linear algebra course is the discovery that matrices correspond to linear transformations, and vice versa. We now review that discovery. Extra care is needed when $\mathbb{K} = \mathbb{H}$.

Definition 1.9. *Suppose that V_1 and V_2 are left vector spaces over $\mathbb{K}$. A function $f : V_1 \to V_2$ is called $\mathbb{K}$-**linear** (or simply **linear**) if for all $a, b \in \mathbb{K}$ and all $X, Y \in V_1$,*

$$f(a \cdot X + b \cdot Y) = a \cdot f(X) + b \cdot f(Y).$$

It is natural to identify $\mathbb{K}^n = \{(q_1, ..., q_n) \mid q_i \in \mathbb{K}\}$ with $M_{1,n}(\mathbb{K})$ (horizontal single-row matrices) and thereby regard $\mathbb{K}^n$ as a left vector space over $\mathbb{K}$. Using this identification, there are two potential ways in which matrices might correspond to linear transformations from $\mathbb{K}^n$ to $\mathbb{K}^n$:

Definition 1.10. *If $A \in M_n(\mathbb{K})$, define $R_A : \mathbb{K}^n \to \mathbb{K}^n$ and define $L_A : \mathbb{K}^n \to \mathbb{K}^n$ such that for $X \in \mathbb{K}^n$,*

$$R_A(X) = X \cdot A \quad \textit{and} \quad L_A(X) = (A \cdot X^T)^T.$$

For example, if $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in M_2(\mathbb{R})$, then for $(x, y) \in \mathbb{R}^2$,

$$R_A(x, y) = \begin{pmatrix} x & y \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = (x + 3y, 2x + 4y), \text{ and}$$

$$L_A(x, y) = \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right)^T = \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix}^T = (x + 2y, 3x + 4y).$$

We first prove that *right* multiplication determines a one-to-one correspondence between linear functions from $\mathbb{K}^n$ to $\mathbb{K}^n$ and matrices.

Proposition 1.11.

(1) *For any* $A \in M_n(\mathbb{K})$, $R_A : \mathbb{K}^n \to \mathbb{K}^n$ *is* $\mathbb{K}$*-linear.*

(2) *Each* $\mathbb{K}$*-linear function from* $\mathbb{K}^n$ *to* $\mathbb{K}^n$ *equals* R_A *for some* $A \in M_n(\mathbb{K})$.

Proof. To prove (1), notice that for all $a, b \in \mathbb{K}$ and $X, Y \in \mathbb{K}^n$,

$$\begin{aligned} R_A(aX + bY) = (aX + bY) \cdot A &= a(X \cdot A) + b(Y \cdot A) \\ &= a \cdot R_A(X) + b \cdot R_A(Y). \end{aligned}$$

To prove (2), assume that $f : \mathbb{K}^n \to \mathbb{K}^n$ is $\mathbb{K}$-linear. Let $A \in M_n(\mathbb{K})$ denote the matrix whose i^{th} row is $f(e_i)$, where

$$e_1 = (1, 0, ..., 0), e_2 = (0, 1, 0, ..., 0), ..., e_n = (0, ..., 0, 1)$$

denotes the standard basis for $\mathbb{K}^n$. It's easy to see that $f(e_i) = R_A(e_i)$ for all $i = 1, ..., n$. Since f and R_A are both linear maps and they agree on a basis, we conclude that $f = R_A$. □

We see from the proof that the rows of $A \in M_n(\mathbb{K})$ are the images under R_A of $\{e_1, ..., e_n\}$. Similarly, the columns are the images under L_A.

Most linear algebra textbooks use the convention of identifying a matrix $A \in M_n(\mathbb{K})$ with the function $L_A : \mathbb{K}^n \to \mathbb{K}^n$. Unfortunately, this function is necessarily $\mathbb{K}$-linear only when $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Proposition 1.12. *Let* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

(1) *For any* $A \in M_n(\mathbb{K})$, $L_A : \mathbb{K}^n \to \mathbb{K}^n$ *is* $\mathbb{K}$*-linear.*

(2) *Each* $\mathbb{K}$*-linear function from* $\mathbb{K}^n$ *to* $\mathbb{K}^n$ *equals* L_A *for some* $A \in M_n(\mathbb{K})$.

Proposition 1.12 is an immediate corollary of Proposition 1.11 plus the following easily verified fact:

$$L_A = R_{A^T} \text{ for all } A \in M_n(\mathbb{R}) \text{ or } A \in M_n(\mathbb{C}).$$

Our previous decision to consider $\mathbb{H}^n$ as a *left* vector space over $\mathbb{H}$ forces us now to use the correspondence $A \leftrightarrow R_A$ between matrices and linear transformations (rather than $A \leftrightarrow L_A$), at least when we wish to include $\mathbb{K} = \mathbb{H}$ in our discussion.

Under either correspondence between matrices and transformations, matrix multiplication corresponds to composition of transformations, since:

$$L_A(L_B(X)) = L_{A\cdot B}(X) \text{ and } R_A(R_B(X)) = R_{B\cdot A}(X).$$

In a linear algebra course, this is one's first indication that the initially unmotivated definition of matrix multiplication is in fact quite natural.

6. The general linear groups

The set $M_n(\mathbb{K})$ is not a group under matrix multiplication because some matrices do not have multiplicative inverses. For example, if $A \in M_n(\mathbb{K})$ has all entries zero, then A has no multiplicative inverse; that is, there is no matrix B for which $AB = BA = I$. However, the elements of $M_n(\mathbb{K})$ which do have inverses form a very important group whose subgroups are the main topic of this text.

Definition 1.13. *The **general linear group** over* $\mathbb{K}$ *is:*

$$GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \exists B \in M_n(\mathbb{K}) \text{ such that } AB = BA = I\}.$$

Such a matrix B is the multiplicative inverse of A and is therefore denoted A^{-1}. As its name suggests, $GL_n(\mathbb{K})$ is a group under the operation of matrix multiplication (why?). The following more visual characterization of the general linear group is often useful:

Proposition 1.14.

$$GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid R_A : \mathbb{K}^n \to \mathbb{K}^n \text{ is a linear isomorphism}\}.$$

For $A \in M_n(\mathbb{K})$, R_A is always linear; it is called an isomorphism if it is invertible (or equivalently, surjective, or equivalently, injective). Thus, general linear matrices correspond to motions of $\mathbb{K}^n$ with no collapsing.

Proof. If $A \in GL_n(\mathbb{K})$ and B is such that $BA = I$, then

$$R_A \circ R_B = R_{BA} = R_I = \text{id (the identity)},$$

so R_A has inverse R_B.

Conversely, let $A \in M_n(\mathbb{K})$ be such that R_A is invertible. The map $(R_A)^{-1}$ is linear, which can be seen by applying R_A to both sides of the following equation:

$$(R_A)^{-1}(aX + bY) \stackrel{?}{=} a(R_A)^{-1}(X) + b(R_A)^{-1}(Y).$$

Since every linear map is represented by a matrix, $(R_A)^{-1} = R_B$ for some $B \in M_n(\mathbb{K})$. Therefore, $R_{BA} = R_A \circ R_B = \text{id}$, which implies $BA = I$. Similarly, $R_{AB} = R_B \circ R_A = \text{id}$, which implies $AB = I$. □

The following well-known fact from linear algebra provides yet another useful description of the general linear group, at least when $\mathbb{K} \neq \mathbb{H}$:

Proposition 1.15. *If* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, *then*

$$GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid \det(A) \neq 0\}.$$

In fact, the elements of the inverse of a matrix can be described explicitly in terms of the determinant of the matrix and its minors:

Proposition 1.16 (Cramer's rule)**.** *Let* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. *Using the notation of Equation* 1.5*,*

$$(A^{-1})_{ij} = (-1)^{i+j}\frac{\det(A[j,i])}{\det(A)}.$$

7. Change of basis via conjugation

In this section, we review a basic fact from linear algebra: a conjugate of a matrix represents the same linear transformation as the matrix, but in a different basis.

Let $\mathfrak{g}$ denote an n-dimensional (left) vector space over $\mathbb{K}$. Then $\mathfrak{g}$ is isomorphic to $\mathbb{K}^n$. In fact, there are many isomorphisms from $\mathfrak{g}$ to $\mathbb{K}^n$. For any ordered basis $V = \{v_1, ..., v_n\}$ of $\mathfrak{g}$, the following is an isomorphism:

$$(c_1v_1 + \cdots + c_nv_n) \mapsto (c_1, ..., c_n). \tag{1.7}$$

Every isomorphism from $\mathfrak{g}$ to $\mathbb{K}^n$ has this form for some ordered basis of $\mathfrak{g}$, so choosing an isomorphism amounts to choosing an ordered basis. In practice, there is typically no choice of basis that is more natural than the other choices. To convince yourself of this, consider the case where $\mathfrak{g}$ is an arbitrary subspace of $\mathbb{K}^m$ for some $m > n$.

Now suppose that $f : \mathfrak{g} \to \mathfrak{g}$ is a linear transformation. In order to identify f with a matrix, we must first choose an ordered basis V of $\mathfrak{g}$. We use this basis to identify $\mathfrak{g} \cong \mathbb{K}^n$ and thereby to regard f as a linear transformation from $\mathbb{K}^n$ to $\mathbb{K}^n$, which can be represented as R_A for some $A \in M_n(\mathbb{K})$. A crucial point is that A depends on the choice of ordered basis. To emphasize this dependence, we say that "A represents f in the basis V (via right-multiplication)." We would like to determine which matrix represents f in a different basis.

To avoid cumbersome notation, we will simplify this problem without really losing generality. Suppose that $f : \mathbb{K}^n \to \mathbb{K}^n$ is a linear transformation. We know that $f = R_A$ for some $A \in M_n(\mathbb{K})$. Translating this sentence into our new terminology, we say that "A represents f in the standard basis of $\mathbb{K}^n$," which is:

$$\{e_1 = (1, 0, ..., 0), e_2 = (0, 1, 0, ..., 0), ..., e_n = (0, ..., 0, 1)\}.$$

Now let $V = \{v_1, ..., v_n\}$ denote an arbitrary basis of $\mathbb{K}^n$. We seek the matrix that represents f in the basis V. First, we let $g \in GL_n(\mathbb{K})$ denote the matrix whose rows are $v_1, v_2, ..., v_n$. We call g the **change of basis matrix**. To understand why, notice that $e_i g = v_i$ for each $i = 1, ..., n$. So,

$$(c_1, ..., c_n) \cdot g = (c_1 e_1 + \cdots + c_n e_n) \cdot g = c_1 v_1 + \cdots + c_n v_n.$$

By Equation 1.7, the vector $c_1 v_1 + \cdots + c_n v_n \in \mathbb{K}^n$ is represented in the basis V as $(c_1, ..., c_n)$. Thus, $R_g : \mathbb{K}^n \to \mathbb{K}^n$ translates between V and the standard basis. For $X \in \mathbb{K}^n$, $R_g(X)$ represents in the standard basis the same vector that X represents in V. Further, $R_{g^{-1}}(X)$ represents in V the same vector that X represents in the standard basis.

Proposition 1.17. *gAg^{-1} represents f in the basis V.*

Proof. Let $X = (c_1, ..., c_n)$, which represents $c_1 v_1 + \cdots + c_n v_n$ in V. We must show that $R_{gAg^{-1}}(X)$ represents $(c_1 v_1 + \cdots + c_n v_n) \cdot A$ in

V. This follows from:

$$R_{gAg^{-1}}(X) = (c_1, ..., c_n)gAg^{-1} = (c_1v_1 + \cdots + c_nv_n)Ag^{-1}$$
$$= R_{g^{-1}}((c_1v_1 + \cdots + c_nv_n) \cdot A).$$

□

Proposition 1.17 can be summarized in the following way: for any $A \in M_n(\mathbb{K})$ and any $g \in GL_n(\mathbb{K})$, the matrix gAg^{-1} represents R_A in the basis $\{e_1g, ..., e_ng\}$.

The basic idea of the proof was simple enough: the transformation $R_{gAg^{-1}} = R_{g^{-1}} \circ R_A \circ R_g$ first translates into the standard basis, then performs the transformation associated to A, then translates back.

This key result requires only slight modification when representing linear transformations using *left* matrix multiplication when $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$: for any $A \in M_n(\mathbb{K})$ and any $g \in GL_n(\mathbb{K})$, the matrix $g^{-1}Ag$ represents L_A in the basis $\{ge_1, ..., ge_n\}$ (via left multiplication). The proof idea is the same: $L_{g^{-1}Ag} = L_{g^{-1}} \circ L_A \circ L_g$ first translates into the standard basis, then performs the transformation associated to A, then translates back.

8. Exercises

Ex. 1.1. Describe a natural 1-to-1 correspondence between elements of $SO(3)$ and elements of

$$T^1S^2 = \{(p, v) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid |p| = |v| = 1 \text{ and } p \perp v\},$$

which can be thought of as the collection of all unit-length vectors v tangent to all points p of S^2. Compare to Question 1.2.

Ex. 1.2. Prove Equation 1.3.

Ex. 1.3. Prove Equation 1.4.

Ex. 1.4. Let $A, B \in M_n(\mathbb{K})$. Prove that if $AB = I$, then $BA = I$.

Ex. 1.5. Suppose that the determinant of $A \in M_n(\mathbb{H})$ were defined as in Equation 1.5. Show for $A = \begin{pmatrix} \mathbf{i} & \mathbf{j} \\ \mathbf{i} & \mathbf{j} \end{pmatrix} \in M_2(\mathbb{H})$ that $\det(A) \neq 0$ but $R_A : \mathbb{H}^2 \to \mathbb{H}^2$ is not invertible.

Ex. 1.6. Find $B \in M_2(\mathbb{R})$ such that $R_B : \mathbb{R}^2 \to \mathbb{R}^2$ is a counter-clockwise rotation through an angle θ.

Ex. 1.7. Describe all elements $A \in GL_n(\mathbb{R})$ with the property that $AB = BA$ for all $B \in GL_n(\mathbb{R})$.

Ex. 1.8. Let $SL_2(\mathbb{Z})$ denote the set of all 2 by 2 matrices with integer entries and with determinant 1. Prove that $SL_2(\mathbb{Z})$ is a subgroup of $GL_2(\mathbb{R})$. Is $SL_n(\mathbb{Z})$ (defined analogously) a subgroup of $GL_n(\mathbb{R})$?

Ex. 1.9. Describe the product of two matrices in $M_6(\mathbb{K})$ which both have the form:

$$\begin{pmatrix} a & b & 0 & 0 & 0 & 0 \\ c & d & 0 & 0 & 0 & 0 \\ 0 & 0 & e & f & g & 0 \\ 0 & 0 & h & i & j & 0 \\ 0 & 0 & k & l & m & 0 \\ 0 & 0 & 0 & 0 & 0 & n \end{pmatrix}$$

Describe a general rule for the product of two matrices with the same **block form**.

Ex. 1.10. If $G_1 \subset GL_{n_1}(\mathbb{K})$ and $G_2 \subset GL_{n_2}(\mathbb{K})$ are subgroups, describe a subgroup of $GL_{n_1+n_2}(\mathbb{K})$ that is isomorphic to $G_1 \times G_2$.

Ex. 1.11. Show by example that for $A \in M_n(\mathbb{H})$, $L_A : \mathbb{H}^n \to \mathbb{H}^n$ is not necessarily $\mathbb{H}$-linear.

Ex. 1.12. Define the *real* and *imaginary* parts of a quaternion as follows:

$$\mathrm{Re}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = a$$
$$\mathrm{Im}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}.$$

Let $q_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$ and $q_2 = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$ be *purely imaginary* quaternions in $\mathbb{H}$. Prove that $-\mathrm{Re}(q_1 \cdot q_2)$ is their vector dot product in $\mathbb{R}^3 = \mathrm{span}\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ and $\mathrm{Im}(q_1 \cdot q_2)$ is their vector cross product.

Ex. 1.13. Prove that non-real elements $q_1, q_2 \in \mathbb{H}$ commute if and only if their imaginary parts are parallel; that is, $\mathrm{Im}(q_1) = \lambda \cdot \mathrm{Im}(q_2)$ for some $\lambda \in \mathbb{R}$.

Ex. 1.14. Characterize the pairs $q_1, q_2 \in \mathbb{H}$ which anti-commute, meaning that $q_1 q_2 = -q_2 q_1$.

Ex. 1.15. If $q \in \mathbb{H}$ satisfies $q\mathbf{i} = \mathbf{i}q$, prove that $q \in \mathbb{C}$.

Ex. 1.16. Prove that complex multiplication in $\mathbb{C} \cong \mathbb{R}^2$ does not extend to a multiplication operation on $\mathbb{R}^3$ that makes $\mathbb{R}^3$ into a real division algebra.

Ex. 1.17. Describe a subgroup of $GL_{n+1}(\mathbb{R})$ that is isomorphic to the group $\mathbb{R}^n$ under the operation of vector-addition.

Ex. 1.18. If $\lambda \in \mathbb{H}$ commutes with every element of $\mathbb{H}$, prove that $\lambda \in \mathbb{R}$.

Chapter 10

Homogeneous manifolds

Other than spheres and matrix groups, we haven't yet encountered many examples of manifolds. In this chapter, we will construct many more, including the projective spaces, which are of fundamental importance in many areas of physics and mathematics. Our goal is to describe new manifolds in simple and elegant ways rather than with messy equations. The only way to achieve this goal is to first generalize our definition of "manifold."

1. Generalized manifolds

In Chapter 7, we defined an *embedded manifold* to essentially mean a subset of some ambient Euclidean space that is locally identified with $\mathbb{R}^d$. The name indicates that it is "embedded" in the ambient Euclidean space. We will henceforth use the word "manifold" to mean a slightly more general type of object. We will define a (generalized) manifold to essentially mean a *set* that is locally identified with $\mathbb{R}^d$. There will no longer be an ambient Euclidean space. This generalization is motivated by several considerations:

(1) Cosmologists model the universe as a 3-dimensional manifold, but it would be unnatural for the model to be a subset of $\mathbb{R}^4$ or $\mathbb{R}^5$, since the universe is the whole universe.

(2) In Section 6 of Chapter 8, we defined $\mathbb{RP}^n$ as the set of all lines through the origin in $\mathbb{R}^{n+1}$, and we described a sense in which $\mathbb{RP}^n$ locally looks like $\mathbb{R}^n$, and therefore deserves to be called a manifold, even though it is a set of lines (rather than a set of points in some Euclidean space).

(3) In Section 1 of Chapter 9, we proved that

$$(\mathbb{R}^n, +)/\langle\{e_1, ..., e_n\}\rangle$$

is isomorphic to the torus, T^n. In fact, it's *smoothly* isomorphic, but for this assertion to make sense, we must somehow regard this coset space as a manifold. More generally, many important manifolds are best described as coset spaces of the form G/H, where G is a matrix group and $H \subset G$ is a closed subgroup. We must learn to regard a coset space as a manifold, even though it is a set of cosets (rather than a set of points in Euclidean space).

Definition 10.1. *A **manifold** of dimension d is set, M, together with a family of injective functions $\varphi_i : U_i \to M$, called **parametrizations** (where each U_i is an open set in $\mathbb{R}^d$) such that:*

(1) $\bigcup_i \varphi_i(U_i) = M$ *(the parametrizations cover all of M).*

(2) (Compatibility condition) *For any pair i, j of indices with $W = \varphi_i(U_i) \cap \varphi_j(U_j) \neq \emptyset$, the sets $\varphi_i^{-1}(W)$ and $\varphi_j^{-1}(W)$ are open in $\mathbb{R}^d$, and*

$$\varphi_j^{-1} \circ \varphi_i : \varphi_i^{-1}(W) \to \varphi_j^{-1}(W)$$

is a smooth function between these open sets.

(3) *The family of parametrizations is maximal with respect to conditions (1) and (2).*

*The family of parametrizations is called an **atlas** for M.*

The compatibility condition is illustrated in Figure 1. This condition is satisfied for the embedded manifolds defined in Chapter 7 (basically because $\varphi_j^{-1} \circ \varphi_i$ is a composition of two smooth functions, but the reader will check the details in Exercise 10.1).

Condition (3) insures, for example, that the restriction of any parametrization to any open subset of its domain is also a parametri-

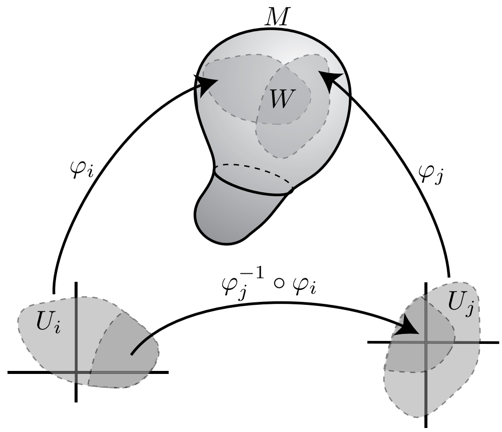


Figure 1. The change of parametrization, $\varphi_j^{-1} \circ \varphi_i$, must be smooth.

zation of the atlas. Our focus in this chapter is the construction of new examples of manifolds. For this activity, we can safely ignore condition (3); it will be enough to construct an atlas satisfying (1) and (2). This is because such an atlas can be uniquely completed to an atlas also satisfying (3). The completion is achieved by adding to it all parametrizations compatible with the original ones in the sense of (2). These newly added parametrizations will automatically also be compatible with each other (Exercise 10.2).

Since a manifold is locally identified (via parametrizations) with open sets in $\mathbb{R}^d$, one can do to a manifold most things that can be done to an open subset of Euclidean space. We'll begin by discussing how to do topological things:

Definition 10.2. *Let M be a manifold. A subset of M is called* ***open*** *(in M) if it equals a union of sets of the form $\varphi_i(A)$, where $A \subset U_i$ is an open subset of the domain, U_i, of a parametrization $\varphi_i : U_i \to M$. The collection of all open sets in M is called the* ***topology*** *of M.*

In particular, the image of any parametrization covering a point of M is a **neighborhood** of that point (which means an open set in M containing that point).

In Chapter 3, we observed that the following concepts could be defined purely in terms of topology: convergence, limit points, closed, compact, path-connected, and continuous. Therefore all of these definitions (previously stated for subsets of Euclidean space) make perfect sense for manifolds. For example, a function $f : M_1 \to M_2$ between manifolds is called *continuous* if for every set A that's open in M_2, $f^{-1}(A)$ is open in M_1.

We next use local parameterizations to define, among other things, the *derivative* of a function between manifolds:

Definition 10.3. *Let M_1 and M_2 be manifolds with dimensions d_1 and d_2 respectively. Let $f : M_1 \to M_2$ be a continuous function, and let $p \in M_1$. Choose parametrizations $\varphi_1 : U_1 \to V_1 \subset M_1$ and $\varphi_2 : U_2 \to V_2 \subset M_2$ covering p and $f(p)$ respectively. Assume that $f(V_1) \subset V_2$; if this is not already the case, it can be achieved by replacing U_1 with the smaller open set $\varphi_1^{-1}(V_1 \cap f^{-1}(V_2))$. Define $\phi = \varphi_2^{-1} \circ f \circ \varphi_1 : U_1 \to U_2$. Define $x = \varphi_1^{-1}(p) \in U_1$.*

(1) *f is called **differentiable** at p if ϕ is differentiable at x.*

(2) *A "**curve** in M_1 through p" is a function $\gamma : (-\epsilon, \epsilon) \to M_1$ (for some $\epsilon > 0$) such that $\gamma(0) = p$ and γ is differentiable at 0 (in the sense of part (2)).*

(3) *A "**tangent vector** to M_1 at p" is an equivalence class of curves in M_1 through p, with two curves considered equivalent if their compositions with φ_1^{-1} (which are curves in U_1 through x) have the same initial derivative.*

(4) *The "**tangent space** to M_1 at p," denoted T_pM_1, is the set of all tangent vectors to M_1 at p, endowed with the structure of a vector space (addition and scalar multiplication) via the identification:*

$$T_pM_1 \cong T_xU_1 \cong \mathbb{R}^{d_1} \qquad \text{described as:}$$
$$[\gamma] \leftrightarrow (\varphi_1^{-1} \circ \gamma)'(0),$$

where $[\gamma] \in T_pM_1$ denotes the equivalence class of the curve γ in M_1 through p.

(5) *(Assuming f is differentiable at p) the **derivative** of f at p is the linear function*

$$df_p : T_pM_1 \to T_{f(p)}M_2$$

defined as

$$df_p([\gamma]) = [f \circ \gamma].$$

Identifying $T_pM_1 \cong \mathbb{R}^{d_1}$ and $T_{f(p)}M_2 \cong \mathbb{R}^{d_2}$ as in part (5), df_p can equivalently be defined as the linear function identified with $d\phi_x : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$.

In Figure 2, an element $[\gamma] \in T_pM_1$ is illustrated as a vector in space, which is literally correct only for an embedded manifold in $\mathbb{R}^3$, but is nevertheless a useful way to picture tangent vectors generally.

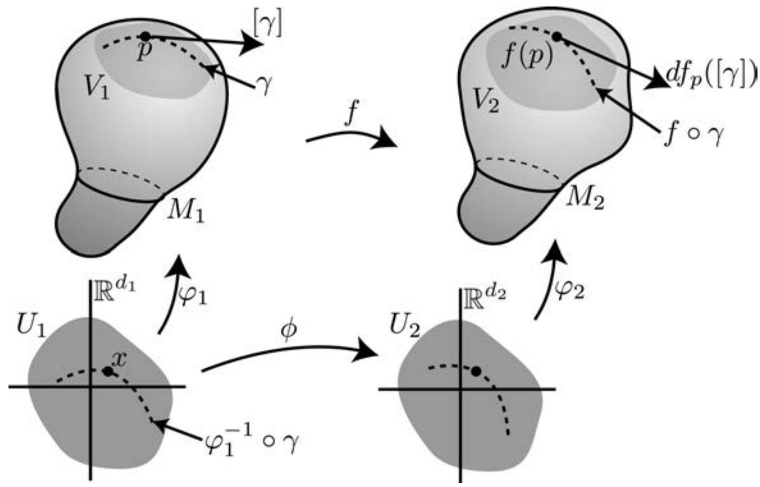


Figure 2. $df_p([\gamma]) = [f \circ \gamma]$.

We leave it to the reader to confirm that Definition 10.3 is well-defined; that is, replacing φ_1 and φ_2 with different parametrizations covering p and $f(p)$ respectively would NOT change:

- whether f is continuous (or differentiable) at p.
- which pairs of curves in M_1 through p are equivalent.

- the vector space operations (addition and scalar multiplication) on T_pM_1.

The compatibility condition from Definition 10.1 is contrived to insure all of these things.

The reader should also check that part (5) is well defined in the sense that the element $[f \circ \gamma] \in T_{f(p)}M_2$ does not depend on the representative, γ, of the equivalence class $[\gamma] \in T_pM_1$.

Definition 10.3 reduces to the previous definitions of these terms from Chapter 7 in the special case of embedded manifolds. For example, when $M_1 \subset \mathbb{R}^m$ is an embedded manifold, two curves in M_1 through p are equivalent if and only if their initial velocities equal the same element of $\mathbb{R}^m$. Therefore, elements of the newly defined T_pM_1 naturally correspond one-to-one with elements of the previously defined T_pM_1 via the correspondence $[\gamma] \leftrightarrow \gamma'(0)$.

A function $f : M_1 \to M_2$ between two manifolds is called **smooth** if for every $p \in M_1$, there exist parametrizations covering p and $f(p)$ such that ϕ (defined as in Definition 10.3) is smooth. For embedded manifolds, it requires some work to prove that this is equivalent to our previous definition of smoothness. Exactly as before, two manifolds are called **diffeomorphic** if there exists a smooth bijection between them whose inverse is also smooth.

The same strategy can be used to decide whether a vector field is smooth. A **vector field**, X, on a manifold, M, means a choice for each $p \in M$ of a vector $X(p) \in T_pM$. Given a parametrization, $\varphi : U \subset \mathbb{R}^d \to V \subset M$, the restriction of X to V is naturally associated with a function from U to $\mathbb{R}^d$. The vector field X is called **smooth** if each point of M is covered by a parametrization with respect to which this associated function is smooth.

The **Whitney Embedding Theorem** says that any manifold is diffeomorphic to an embedded manifold. This means that the class of generalized manifolds is really no more general than the class of embedded manifolds that we previously studied in Chapter 7. However, the specific manifolds we will study throughout this chapter would not look elegant or simple if one tried to describe them as embedded manifolds. Describing them as generalized manifolds is the way to go.

2. The projective spaces

Other than spheres, the most important manifolds in geometry are the projective spaces, which we will define and study in this section.

Definition 10.4. *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and $n \geq 1$. A* ***line*** *in $\mathbb{K}^{n+1}$ means a 1-dimensional $\mathbb{K}$-subspace of $\mathbb{K}^{n+1}$, that is, a set that has the form $\{\lambda \cdot v \mid \lambda \in \mathbb{K}\}$ for some non-zero vector $v \in \mathbb{K}^{n+1}$. The set*

$$\mathbb{KP}^n = \{\textit{all lines in } \mathbb{K}^{n+1}\}$$

is called ***real projective space*** *($\mathbb{RP}^n$),* ***complex projective space*** *($\mathbb{CP}^n$), or* ***quaternionic projective space*** *($\mathbb{HP}^n$).*

Under the identification $\mathbb{C}^{n+1} \cong \mathbb{R}^{2(n+1)}$, notice that any line in $\mathbb{C}^{n+1}$ gets identified with a 2-dimensional $\mathbb{R}$-subspace of $\mathbb{R}^{2(n+1)}$ and is therefore best visualized as a plane rather than a line. Similarly, a line in $\mathbb{H}^{n+1}$ gets identified with a 4-dimensional $\mathbb{R}$-subspace of $\mathbb{R}^{4(n+1)}$.

Example 10.5 (A manifold structure for $\mathbb{KP}^n$)**.** *For any given point $(z_0, z_1, ..., z_n) \in \mathbb{K}^{n+1}$ (other than the origin), the unique line containing this point will be denoted as $[z_0, z_1, ..., z_n] \in \mathbb{KP}^n$.*

For each $i = 0, 1, ..., n$, define $V_i = \{[z_0, z_1, ..., z_n] \in \mathbb{KP}^n \mid z_i \neq 0\}$ and define $\varphi_i : \mathbb{K}^n \to V_i$ to be the map that assigns the i^{th} coordinate to equal 1; that is:

$$\begin{aligned}
\varphi_0(z_1, z_2, ..., z_n) &= [1, z_1, z_2, ..., z_n],\\
\varphi_1(z_1, z_2, ..., z_n) &= [z_1, 1, z_2, ..., z_n],\\
&\vdots\\
\varphi_n(z_1, z_2, ..., z_n) &= [z_1, z_2, ..., z_n, 1].
\end{aligned}$$

These $n+1$ parametrizations are injective and together they cover all of $\mathbb{KP}^n$.

It remains to verify the compatibility condition of Definition 10.1. *For example, when* $i = 0$ *and* $j = 1$*,*

$$W = \varphi_0(\mathbb{K}^n) \cap \varphi_1(\mathbb{K}^n) = \{[z_0, z_1, ..., z_n] \in \mathbb{KP}^n \mid z_0, z_1 \neq 0\},$$

and therefore

$$\varphi_0^{-1}(W) = \varphi_1^{-1}(W) = \{(z_1, ..., z_n) \in \mathbb{K}^n \mid z_1 \neq 0\},$$

which is open. Furthermore,

$$\left(\varphi_1^{-1} \circ \varphi_0\right)(z_1, z_2, ..., z_n) = (z_1^{-1}, z_2, ..., z_n),$$

which is smooth. There is nothing special here about $i = 0, j = 1$*; condition (2) can be similarly verified for all other pairs of indices.*

Since these parametrizations locally identify $\mathbb{KP}^n$ *with* $\mathbb{K}^n$*, we see that dim*$(\mathbb{KP}^n) \in \{n, 2n, 4n\}$*, depending on* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$.

There is a noteworthy linguistic difference between an embedded manifold (Definition 7.14) and a general manifold (Definition 10.1). One may ask whether a subset of Euclidean space *is* an embedded manifold, whereas one may only ask whether an arbitrary set "has a natural manifold structure," meaning an atlas as in the definition. In the previous example, we didn't prove that $\mathbb{KP}^n$ *is* a manifold, but rather we made it into a manifold in a natural way. The symbol "$\mathbb{KP}^n$" will henceforth denote the set of lines in $\mathbb{K}^{n+1}$ together with the above-described manifold structure (atlas).

For $n \geq 2$, the spaces $\mathbb{CP}^n$ and $\mathbb{HP}^n$ are not diffeomorphic to any previously familiar manifolds, but $\mathbb{CP}^1$ and $\mathbb{HP}^1$ are very familiar:

Proposition 10.6. $\mathbb{CP}^1$ *is diffeomorphic to* S^2*, and* $\mathbb{HP}^1$ *is diffeomorphic to* S^4.

Proof. We will prove the first claim; the second is similar and is left to the reader in Exercise 10.9. We will show that the bijection $F : S^2 \to \mathbb{CP}^1$ defined so that for all unit-length vectors (x, y, z),

$$F(x, y, z) = \begin{cases} [1, 0] & \text{if } z = 1 \\ \left[\frac{2}{1-z}(x + \mathbf{i}y), 1\right] & \text{if } z \neq 1 \end{cases}$$

is a diffeomorphism. For inputs other than $(0,0,1)$, F is the composition of the following three diffeomorphisms:

$$\begin{array}{ccc} S^2 - \{(0,0,1)\} & & \mathbb{CP}^1 - \{[1,0]\} \\ {\scriptstyle (x,y,z)\mapsto\frac{2}{1-z}(x,y)}\big\downarrow & & \big\uparrow{\scriptstyle (a+b\mathbf{i})\mapsto[a+b\mathbf{i},1]} \\ \mathbb{R}^2 & \xrightarrow{(a,b)\mapsto(a+\mathbf{i}b)} & \mathbb{C} \end{array}$$

The left function is stereographic projection, defined in Exercise 7.1. For these inputs, the definition of smoothness is obviously satisfied with respect to the parametrizations determined by the vertical arrows in the above diagram.

It remains to choose parametrizations covering the exceptional point $p = (0,0,1)$ and its image $f(p) = [1,0]$, and to verify with respect to these parametrizations that F and its inverse satisfy the definition of smooth at p. For this, we choose the parametrizations determined by the vertical arrows in this diagram:

$$\begin{array}{ccc} S^2 - \{(0,0,-1)\} & \xrightarrow{F} & \mathbb{CP}^1 - \{[0,1]\} \\ {\scriptstyle (x,y,z)\mapsto\frac{2}{1+z}(x,y)}\big\downarrow & & \big\uparrow{\scriptstyle (a+b\mathbf{i})\mapsto[1,a+b\mathbf{i}]} \\ \mathbb{R}^2 & \xrightarrow{\phi} & \mathbb{C} \end{array}$$

The left function is the projection onto the plane $z = 1$ discussed in Exercise 7.1(3). The reader should check that $\phi : \mathbb{R}^2 \to \mathbb{C}$ (defined as the composition: up, then right, then down) is described by the formula

$$\phi(x,y) = \frac{1}{4}(x - \mathbf{i}y),$$

which is clearly smooth with smooth inverse. □

The next section is devoted to proving the following: if G is a matrix group and $H \subset G$ is a closed subgroup, then the coset space G/H has a natural manifold structure. As a clarifying illustration, we end the current section with a description of the natural manifold structure on one particular coset space:

Example 10.7 ($Sp(1)/T \cong S^2$). *Regard $Sp(1) \cong S^3 \subset \mathbb{H}$ as the group of unit-length quaternions. Its standard maximal torus is the circle:*

$$T = \{e^{\mathbf{i}\theta} \mid \theta \in [0, 2\pi)\} = S^3 \cap \mathbb{C},$$

where $\mathbb{C} \subset \mathbb{H}$. For any $q \in S^3$, the coset $q \cdot T \in S^3/T$ is:

$$q \cdot T = \{q \cdot e^{\mathbf{i}\theta} \mid \theta \in [0, 2\pi)\} = S^3 \cap \{q \cdot \lambda \mid \lambda \in \mathbb{C}\}, \tag{10.1}$$

which is also a circle; in fact, it is diffeomorphic to T via the diffeomorphism that left-multiplies by q. For example, $\mathbf{j} \cdot T$ is the circle in the $\{\mathbf{j}, \mathbf{k}\}$-plane of $\mathbb{H}$. In general, the cosets of any subgroup partition the group; in this case, it is interesting that the 3-dimensional sphere is naturally partitioned into a collection of disjoint circles.

These are not just any circles. It is apparent from Equation 10.1 *that each coset is a "complex great circle," which means the intersection of S^3 with a 1-dimensional $\mathbb{C}$-subspace of $\mathbb{H} \cong \mathbb{C}^2$ (see Exercise* 2.12 (2) *for a description of the identification $\mathbb{H} \cong \mathbb{C}^2$ that is appropriate here). The coset space S^3/T is thereby identified with the set of lines in $\mathbb{C}^2$, which is called $\mathbb{CP}^1$, and which by Proposition* 10.6 *is diffeomorphic to S^2.*

3. Coset spaces are manifolds

When G is a group and $H \subset G$ is a subgroup, recall that "G/H" and "$H\backslash G$" respectively denote the left and right *coset space*, i.e., the collection of (left/right) cosets. The left coset containing g will be denoted either as $g \cdot H$ or as $[g]$. The right coset containing g will be denoted either as $H \cdot g$ or as $[g]$. Algebra textbooks focus on the case where H is normal in G, so that the coset space is itself a group. We will not care whether H is normal in G, but we will instead assume that G and H are matrix groups; in this case, the coset space has a natural manifold structure:

Theorem 10.8. *Let G be a matrix group and $H \subset G$ a closed subgroup. Let $\mathfrak{h} \subset \mathfrak{g}$ denote their Lie algebras. Define*

$$\mathfrak{p} = \mathfrak{h}^{\perp} = \{V \in \mathfrak{g} \mid \langle V, X \rangle_{\mathbb{R}} = 0 \ \forall X \in \mathfrak{h}\} \ \text{and} \ \mathfrak{p}_{\epsilon} = \{V \in \mathfrak{p} \mid |V| < \epsilon\}.$$

For every $g \in G$, define the parametrization $\varphi_g : \mathfrak{p}_\epsilon \to G/H$ as:

$$\varphi_g(X) = [g \cdot \exp(X)].$$

If ϵ is sufficiently small, the family of parametrizations $\{\varphi_g \mid g \in G\}$ determines a manifold structure on the coset space G/H; that is, they are injective and they satisfy the compatibility condition of Definition 10.1*.*

The notation "G/H" will henceforth denote the left coset space together with the manifold structure described in the theorem. Notice that the dimension of the G/H equals

$$\dim(G/H) = \dim(\mathfrak{p}) = \dim(G) - \dim(H).$$

The theorem and and its proof can be modified in the obvious way to provide a manifold structure on the *right* coset space, with parametrizations defined as $\varphi_g(X) = [\exp(X) \cdot g]$. The notation "$H \backslash G$" will henceforth denote the right coset space together with this manifold structure.

Most of the work of the proof goes into the following lemma, which is of independent interest:

Lemma 10.9. *With the assumptions and notation of Theorem* 10.8*, there exists* $\epsilon > 0$ *such that the function* $\Psi : \mathfrak{p}_\epsilon \times H \to G$ *defined as*

$$\Psi(X, h) = \exp(X) \cdot h$$

is a diffeomorphism onto its image, which is an open set in G *containing* H*.*

Proof of Lemma 10.9. The derivative of Ψ at $(0, I)$ is the identity, so the Inverse Function Theorem (Theorem 7.22) implies that the restriction of Ψ to a sufficiently small neighborhood of $(0, I)$ is a diffeomorphism onto its image. There exists $\epsilon > 0$ and a neighborhood, U, of I in H such that the restriction of Ψ to $\mathfrak{p}_\epsilon \times U$ is a diffeomorphism onto its image (because ϵ and U can be chosen such that this domain lies inside the previously mentioned neighborhood). For any $h \in H$, the right-multiplication map $\mathcal{R}_h : G \to G$ is a diffeomorphism, so the restriction of Ψ to $\mathfrak{p}_\epsilon \times (U \cdot h)$ is also a diffeomorphism onto its image.

In summary, we have chosen a value $\epsilon > 0$ such that $\Psi : \mathfrak{p}_\epsilon \times H \to G$ is a *local* diffeomorphism at each point of its domain (which means its restriction to a neighborhood of that point is a diffeomorphism onto its image).

To prove the theorem, it remains to show that Ψ is injective for some possibly smaller choice of ϵ. Suppose to the contrary that no such ϵ exists; that is, for every $\epsilon > 0$, there exists a pair of *distinct* points $(X_1, h_1), (X_2, h_2) \in \mathfrak{p}_\epsilon \times H$ such that

$$\begin{aligned}\Psi(X_1, h_1) = \Psi(X_2, h_2) &\iff \exp(X_1) \cdot h_1 = \exp(X_2) \cdot h_2 \\ &\iff \exp(X_2)^{-1} \cdot \exp(X_1) = h_2 h_1^{-1} \in H.\end{aligned}$$

Theorem 7.1 applied to H (which is itself a matrix group because it is closed) implies there exists $A \in \mathfrak{h}$ with $\exp(A) = h_2 h_1^{-1}$, provided ϵ is sufficiently small. The norm of A can be made arbitrarily small by choosing ϵ arbitrarily small. In summary:

$$\exp(X_2)^{-1} \cdot \exp(X_1) = \exp(A). \tag{10.2}$$

Now consider the smooth function $F : \mathfrak{g} \to G$ defined as

$$F(Y) = \exp(Y^{\mathfrak{p}}) \cdot \exp(Y^{\mathfrak{h}}),$$

where the superscripts denote orthogonal projections onto those subspaces. The derivative of F at 0 is the identity map, so according to the Inverse Function Theorem (Theorem 7.22), F is a diffeomorphism when restricted to a sufficiently small neighborhood of the origin in $\mathfrak{g}$. Now re-write Equation 10.2 as:

$$\underbrace{\exp(X_1)}_{F(X_1+0)} = \underbrace{\exp(X_2) \cdot \exp(A)}_{F(X_2+A)}.$$

When ϵ is sufficiently small, this contradicts the injectivity of F. □

Proof of Theorem 10.8. Choose ϵ as in Lemma 10.9. For any $g \in G$, φ_g is injective because for any $X_1, X_2 \in \mathfrak{p}_\epsilon$:

$$\begin{aligned}\varphi_g(X_1) = \varphi_g(X_2) &\iff [g \cdot \exp(X_1)] = [g \cdot \exp(X_2)] \\ &\iff \exp(X_1)^{-1} \cdot \exp(X_2) = h \in H \\ &\iff \exp(X_2) = \exp(X_1) \cdot h \\ &\iff \Psi(X_2, I) = \Psi(X_1, h) \\ &\iff X_1 = X_2 \text{ and } h = I.\end{aligned}$$

It remains to verify the compatibility condition of Definition 10.1. Let $a, b \in G$ such that $W = \varphi_a(\mathfrak{p}_\epsilon) \cap \varphi_b(\mathfrak{p}_\epsilon) \neq \emptyset$. For $X \in \mathfrak{p}_\epsilon$, notice that $X \in \varphi_a^{-1}(W)$ if and only if there exists $(Y, h) \in \mathfrak{p}_\epsilon \times H$ such that:

$$a \cdot \exp(X) = b \cdot \exp(Y) \cdot h \iff \exp(X) = (a^{-1}b) \cdot \Psi(Y, h). \tag{10.3}$$

So $\varphi_a^{-1}(W)$ is open because it equals the pre-image under the continuous function $\exp : \mathfrak{p}_\epsilon \to G$ of the open set $(a^{-1}b) \cdot \text{Image}(\Psi)$.

Lemma 10.9 implies that the choice $(Y, h) \in \mathfrak{p}_\epsilon \times H$ in Equation 10.3 is uniquely and smoothly determined by X. The transition function $\varphi_b^{-1} \circ \varphi_a : \varphi_a^{-1}(W) \to \varphi_b^{-1}(W)$, which sends $X \mapsto Y$, is therefore smooth. □

Proposition 10.10. *With the assumptions and notation of Theorem 10.8, define $\pi : G \to G/H$ such that $\pi(g) = [g]$ for all $g \in G$.*

(1) *π is smooth.*
(2) *When M is a manifold, a function $f : G/H \to M$ is smooth if and only if $f \circ \pi$ is smooth.*

Proof. Exercise 10.10. □

4. Group actions

To understand the symmetries of a manifold, one must understand how groups act on it.

Definition 10.11. *A (left or right)* ***action*** *of a group G on a set M is a function*

$$\varphi : G \to \{\textit{the set of bijections from } M \textit{ to } M\}$$

such that for all $g_1, g_2 \in G$,

$$\begin{aligned} &\textit{(left action)} \quad \varphi(g_1 \cdot g_2) = \varphi(g_1) \circ \varphi(g_2), \\ &\textit{(right action)} \quad \varphi(g_1 \cdot g_2) = \varphi(g_2) \circ \varphi(g_1). \end{aligned}$$

An action of a group G on a vector space over $\mathbb{K}$ is called a ***$\mathbb{K}$-linear*** *action (or a* ***representation****) if each bijection in its image is $\mathbb{K}$-linear.*

An action of a matrix group G on a manifold M is called ***smooth*** *if the function $G \times M \to M$ that sends $(g, p) \mapsto \varphi(g)(p)$ is smooth (with respect to the natural manifold structure on $G \times M$ discussed in Exercise 10.5). This implies that each bijection in its image is a diffeomorphism.*

The set of bijections from M to M form a group under composition. A *left* action of G on M could equivalently be defined as a homomorphism from G to this group.

We have already encountered and studied the following important group actions:

Example 10.12. *Let $G \subset GL_n(\mathbb{K})$ be a matrix group with Lie algebra denoted $\mathfrak{g}$.*

(1) *The action of G on $\mathbb{K}^n$ by left multiplication ($\varphi(g) = L_g$) is a left action. It is $\mathbb{R}$-linear but not necessarily $\mathbb{K}$-linear.*
(2) *The action of G on $\mathbb{K}^n$ by right multiplication ($\varphi(g) = R_g$) is a $\mathbb{K}$-linear right action.*
(3) *If $G \subset \mathcal{O}_n(\mathbb{K})$, then the previous two actions of G on $\mathbb{K}^n$ restrict to smooth actions on the sphere of unit-length vectors in $\mathbb{K}^n$.*
(4) *The adjoint action of G on $\mathfrak{g}$ ($\varphi(g) = Ad_g$) is a left $\mathbb{R}$-linear action. If $G \subset \mathcal{O}_n(\mathbb{K})$, then it restricts to a smooth left action on the sphere of unit-length vectors in $\mathfrak{g}$.*
(5) *There are three natural smooth actions of G on G:*
    (a) *Left-multiplication ($\varphi(g) = \mathcal{L}_g$) is a smooth left action.*
    (b) *Right-multiplication ($\varphi(g) = \mathcal{R}_g$)is a smooth right action.*
    (c) *Conjugation ($\varphi(g) = C_g$) is a smooth left action.*

Group actions are fundamentally important to many fields of mathematics. The following general vocabulary will be useful as we explore their relevance to matrix groups:

Definition 10.13. *Let φ be a (left or right) action of a group G on a set M.*

(1) *For any $g \in G$ and $p \in M$, the element $\varphi(g)(p) \in M$ will also be denoted as "$g \star p$".*
(2) *For any $p \in M$, the set $G \star p = \{g \star p \mid g \in G\}$ is called the **orbit** containing p.*
(3) *The set of orbits is called the **orbit space** and is denoted as $G\backslash M$ (for a left action) or M/G (for a right action).*

(4) *φ is called **transitive** if there is only one orbit; that is, for any pair $p, q \in M$ there exists $g \in G$ such that $g \star p = q$.*

(5) *For any $p \in G$, the set $G_p = \{g \in G \mid g \star p = p\}$ is called the **stabilizer** of p.*

(6) *φ is called **free** if all stabilizers are trivial; in other words, $(g \star p = p) \Rightarrow (g = I)$.*

For a smooth action of a matrix group G on a manifold M, all stabilizers are *closed* subgroups of G (Exercise 10.14).

Example 10.14. *Let G be a matrix group and $H \subset G$ a closed subgroup. The action of H on G by right multiplication is a smooth free right action whose orbits are the left cosets, so the orbit space equals the coset space. The familiar notation "G/H" for this coset space is consistent with our new notation for a general orbit space. Furthermore, the action of G on the orbit space G/H defined as*

$$g_1 \star [g_2] = [g_1 \cdot g_2]$$

is a well-defined transitive left action. Proposition 10.10 *can be used to prove that this action is smooth (Exercise* 10.16*).*

Similarly, the action of H on G by left multiplication is a smooth free left action whose orbit space, $H \backslash G$, is the space of right cosets. The action of G on this coset space defined as $g_1 \star [g_2] = [g_2 \cdot g_1]$ is a well-defined transitive smooth right action.

In summary: If G is a matrix group and $H \subset G$ a closed subgroup, then either coset space (G/H or $H \backslash G$) is a special type of manifold – one on which a matrix group acts transitively.

Definition 10.15. *A manifold, M, is called **homogeneous** if there exists a transitive smooth action of a matrix group on M.*

Conversely, it turns out essentially that every homogeneous manifold is diffeomorphic to a coset space. We will sketch the proof of this assertion in the next section.

5. Homogeneous manifolds

In this section, as promised, we will sketch a proof that every homogeneous manifold is diffeomorphic to a coset space. To illustrate the key idea, it is useful to first pursue a different goal: to describe an unexpected way in which Theorem 10.8 can provide a homogeneous manifold structure on certain sets. The first set that we'd like to turn into a homogeneous manifold is:

Definition 10.16. *Let $m < n$ be positive integers. The set of m-dimensional $\mathbb{K}$-subspaces of $\mathbb{K}^n$ is denoted $G_m(\mathbb{K}^n)$ and is called the* ***Grassmann manifold*** *of m-planes in $\mathbb{K}^n$.*

The name suggests that it's a manifold, but it is not obvious how to construct an atlas. Luckily we won't have to. Instead, consider the single element $V_0 = \text{span}\{e_1, e_2, ..., e_m\} \in G_m(\mathbb{K}^n)$ (the span of the first m members of the standard orthonormal basis of $\mathbb{K}^n$). For any $g \in GL_n(\mathbb{K})$, notice that $R_g(V_0) = \{v \cdot g \mid v \in V_0\}$ is another element of $G_m(\mathbb{K}^n)$. But this observation doesn't quite allow us to identify $GL_n(\mathbb{K})$ with $G_m(\mathbb{K}^n)$ because the identification is not one-to-one. For example, $R_g(V_0) = V_0$ if and only if g lies in the following subgroup:

$$H_0 = \{g \in GL_n(\mathbb{K}) \mid \text{each of the first } m \text{ rows of } g \text{ is in } V_0\}. \tag{10.4}$$

That is, H_0 contains all the general linear matrices whose top-right block equals the m-by-$(n-m)$ zero matrix.

In fact, $R_{g_1}(V_0) = R_{g_2}(V_0)$ if and only if g_1 and g_2 lie in the same coset of $H_0\backslash GL_n(\mathbb{K})$. We therefore have a natural identification

$$G_m(\mathbb{K}^n) \cong H_0\backslash GL_n(\mathbb{K}).$$

Theorem 10.8 provides a homogeneous manifold structure on this coset space and (via this identification) also on $G_m(\mathbb{K}^n)$.

The key idea here is the following simple and general fact:

Lemma 10.17. *Let φ be a transitive left (respectively right) action of a group G on a set M, let $p_0 \in M$, and let $H = G_{p_0}$ be the stabilizer of p_0. Then the function $F : G/H \to M$ (respectively $F : H\backslash G \to M$) defined as*

$$F([g]) = g \star p_0$$

is a well-defined bijection.

Proof. Exercise 10.15. □

In either case, F is in fact an **equivariant** bijection, which means that it identifies the transitive actions of G on these two sets. This means, for example in the left version, that for each $g \in G$, the following diagram commutes:

$$\begin{array}{ccc} G/H & \xrightarrow{\tilde{\varphi}(g)} & G/H \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle F} \\ M & \xrightarrow{\varphi(g)} & M \end{array} \tag{10.5}$$

where $\tilde{\varphi}$ is the action of G on G/H defined as $\tilde{\varphi}(g)([x]) = [g \cdot x]$, while φ is the given action of G on M, denoted as $\varphi(g)(p) = g \star p$.

According to Exercise 10.12, no generality is lost in developing the general theory only for *left* actions. Nevertheless, certain examples are most naturally described in terms of right actions, including the next example.

Example 10.18 (Steifel manifolds)**.** *Let $m < n$ be positive integers. The **Steifel manifold** of m-frames in $\mathbb{K}^n$, denoted $S_m(\mathbb{K}^n)$, is defined as the set of all ordered lists of m linearly independent elements of $\mathbb{K}^n$.*

There is a natural transitive right action of $GL_n(\mathbb{K})$ on $S_m(\mathbb{K}^n)$ described as follows: if $g \in GL_n(\mathbb{K})$ and $F = (v_1, ..., v_m) \in S_m(\mathbb{K}^n)$, then $g \star F = (v_1 \cdot g, ..., v_m \cdot g) \in S_m(\mathbb{K}^n)$. The simplest element of $S_m(\mathbb{K}^n)$ is $F_0 = (e_1, ..., e_m)$ (the first m members of the standard orthonormal basis of $\mathbb{K}^n$ listed in natural order). Its stabilizer, H, equals the subgroup of all matrices in $GL_n(\mathbb{K})$ whose first m rows equal $e_1, e_2, ..., e_m$ (in that order). Lemma 10.17 *provides a natural identification $S_m(\mathbb{K}^n) \cong H \backslash GL_n(\mathbb{K})$. Since this coset space has the structure of a homogeneous manifold, the Steifel manifold inherits the same via this identification.*

The Grassmann and Steifel manifolds are previously unfamiliar sets that can be given the structure of homogeneous manifolds by identifying transitive actions on them. We now turn our attention to a smooth transitive action on a set M that already has a manifold structure. In this case, we wish to verify that the manifold structure that M inherits via the identification with a coset space is the same as its given manifold structure. More significantly, we wish to prove that every homogeneous manifold is a coset space. These wishes are granted by the following smooth version of Lemma 10.17:

Theorem 10.19. *In Lemma* 10.17*, if M is a manifold and G is a path connected matrix group and φ is smooth, then F is a diffeomorphism.*

We'll see in the proof that the "path-connected" hypothesis can be replaced with the weaker hypothesis that G has only countably many connected components.

Sketch of proof. We know from Lemma 10.17 that F is a well-defined bijection, while Proposition 10.10 implies that F is smooth. According to the Inverse Function Theorem for generalized manifolds (Exercise 10.4), it will suffice to prove that for any $g \in G$, the linear transformation $dF_{[g]} : T_{[g]}(G/H) \to T_{g\star p_0}M$ is an isomorphism.

In fact, it is enough to verify the special case when $g = I$. Why? Because the diagram in Equation 10.5 gives that for any $g \in G$,

$$F = \underbrace{\varphi(g)}_{\text{diffeo}} \circ F \circ \underbrace{\tilde{\varphi}(g)^{-1}}_{\text{diffeo}}.$$

So the chain rule for generalized manifolds (Exercise 10.3) gives:

$$dF_{[g]} = \underbrace{d\left(\varphi(g)\right)_{p_0}}_{\text{isomorphism}} \circ dF_{[I]} \circ \underbrace{d\left(\tilde{\varphi}(g)^{-1}\right)_{[g]}}_{\text{isomorphism}}.$$

Thus, $dF_{[g]}$ is an isomorphism if and only if $dF_{[I]}$ is an isomorphism.

We will next prove that $dF_{[I]} : T_{[I]}(G/H) \to T_{p_0}M$ is injective (has trivial kernel). Suppose to the contrary that $dF_{[I]}(X) = 0$ for some non-zero $X \in T_{[I]}(G/H)$. Let $\mathfrak{h}$ denote the Lie algebra of H and define $\mathfrak{p} = \mathfrak{h}^{\perp}$. Notice that

$$0 = dF_{[I]}(X) = \frac{d}{dt}\bigg|_{t=0} \exp(tX) \star p_0,$$

where on the right, we are regarding X as an element of $\mathfrak{p}$ via the natural identification $T_{[I]}(G/H) \cong \mathfrak{p}$ provided by Theorem 10.8. For all $t_0 \in \mathbb{R}$, we have:

$$\begin{aligned}
\frac{d}{dt}\Big|_{t=t_0} \exp(tX) \star p_0 &= \frac{d}{dt}\Big|_{t=0} \exp((t_0+t)X) \star p_0 \\
&= \frac{d}{dt}\Big|_{t=0} (\exp(t_0X)\cdot \exp(tX)) \star p_0 \\
&= \frac{d}{dt}\Big|_{t=0} \varphi\left(\exp(t_0X)\right)\left(\exp(tX)\star p_0\right) \\
&= d\left(\varphi(\exp(t_0X))\right)_{p_0}\left(\frac{d}{dt}\Big|_{t=0}\exp(tX)\star p_0\right) = 0.
\end{aligned}$$

Since the path $t \mapsto \exp(tX) \star p_0$ has zero derivative for all time, it is constant: $\exp(tX) \star p_0 = p_0$ for all $t \in \mathbb{R}$. Therefore $\exp(tX) \in H$ for all $t \in \mathbb{R}$. This implies that $X \in \mathfrak{h}$, which contradicts the hypothesis that X is a non-zero vector in $\mathfrak{p}$.

In summary, $dF_{[I]} : T_{[I]}(G/H) \to T_{p_0}M$ is injective, which would force it to also be surjective if we knew that $\dim(G/H) = \dim(M)$. It therefore remains to rule out the possibility that M has a larger dimension that G/H. This requires some topology, and we refer the reader to [**11**] or [**13**] for a proof. The path-connected hypothesis is used here, but can be replaced by the weaker hypothesis that G has countably many connected components. In fact, this hypothesis can be completely removed, provided one follows the common convention of adding two topological conditions to the definition of manifold (namely, "Hausdorff" and "second countable"). □

The previous theorem allows us to better understand many familiar manifolds by identifying them with coset spaces. To give an example, we must first adopt some notation conventions. First, we will always think of products of matrix groups as having the most obvious block-diagonal form, so if $H_1 \subset GL_{n_1}(\mathbb{K})$ and $H_2 \subset GL_{n_2}(\mathbb{K})$, then "$H_1 \times H_2$" will denote the following subgroup of $GL_{n_1+n_2}(\mathbb{K})$:

$$H_1 \times H_2 = \{\text{diag}(A, B) \mid A \in H_1, B \in H_2\}.$$

Further, when $H \subset GL_n(\mathbb{K})$ is a subgroup, we will denote:

$$S(H) = \{M \in H \mid \det(M) = 1\}.$$

For example, $S(O(n)) = SO(n)$, $S(U(n)) = SU(n)$ and $S(Sp(n)) = Sp(n)$. With this notation, we have:

Corollary 10.20. *In each of the following, $n \geq 1$ and "=" means the manifold is diffeomorphic to the coset space:*

(1) $\mathbb{KP}^n = \mathcal{O}_{n+1}(\mathbb{K})/\left(\mathcal{O}_1(\mathbb{K}) \times \mathcal{O}_n(\mathbb{K})\right)$
$= S\left(\mathcal{O}_{n+1}(\mathbb{K})\right)/S\left(\mathcal{O}_1(\mathbb{K}) \times \mathcal{O}_n(\mathbb{K})\right)$
In particular:

$$\begin{aligned}
\mathbb{RP}^n &= O(n+1)/\left(\{\pm 1\} \times O(n)\right) \\
&= SO(n+1)/S\left(\{\pm 1\} \times O(n)\right) \\
\mathbb{CP}^n &= U(n+1)/\left(U(1) \times U(n)\right) \\
&= SU(n+1)/S\left(U(1) \times U(n)\right) \\
\mathbb{HP}^n &= Sp(n+1)/\left(Sp(1) \times Sp(n)\right)
\end{aligned}$$

(2) $S^n = O(n+1)/(\{1\} \times O(n))$
$= SO(n+1)/\left(\{1\} \times SO(n)\right)$

(3) $S^{2n+1} = U(n+1)/\left(\{1\} \times U(n)\right)$
$= SU(n+1)/\left(\{1\} \times SU(n)\right)$

(4) $S^{4n+3} = Sp(n+1)/\left(\{1\} \times Sp(n)\right)$

Proof. In each example, written as "$M = G/H$," we must identify a smooth transitive action of G on M and a point of M whose stabilizer equals H. Since G/H is diffeomorphic to $H\backslash G$, it doesn't matter whether we identify a left or right action. In all cases, the action is the standard action of $\mathcal{O}_{n+1}(\mathbb{K})$ (or a subgroup thereof) on $\mathbb{K}^{n+1}$ (or the induced action on the lines in $\mathbb{K}^{n+1}$ or the induced action on the sphere of unit-length vectors in $\mathbb{K}^{n+1}$) for the appropriate choice of $\mathbb{K}$. Furthermore, in all cases, p_0 is either e_1 or $[e_1]$, where $e_1 = (1, 0, ..., 0)$ denotes the first member of the standard orthonormal basis of the $\mathbb{K}^{n+1}$. The transitivity of these actions follows from Exercise 3.18. □

6. Riemannian manifolds

Since a manifold is locally identified with Euclidean space, one can do to a manifold *almost* everything that one can do to Euclidean space. But there is a crucial exception: one cannot measure lengths of curves or distances between pairs of points. Attempting to do so using parametrizations would lead to answers that would depend on the choice of parametrization and therefore would not be well-defined.

These “metric” measurements require an additional structure that is specified in the following two definitions.

Definition 10.21. *Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{R}$. An **inner product** on $\mathcal{V}$ is a bilinear function $\mathcal{V} \times \mathcal{V} \to \mathbb{R}$, denoted as $v, w \mapsto \langle v, w\rangle$, satisfying the following properties:*

(1) (symmetric) $\langle v, w\rangle = \langle w, v\rangle$ *for all* $v, w \in \mathcal{V}$.

(2) (positive definite) $\langle v, v\rangle \geq 0$ *for all* $v \in \mathcal{V}$, *with equality if and only if* $v = 0$.

In other words, an inner product is a function that has the same algebraic properties of the standard inner product on $\mathbb{R}^n$, as enumerated in Section 1 of Chapter 3. Just as with the standard inner product, an arbitrary inner product empowers one to compute norms and angles as follows:

$$|v| = \sqrt{\langle v, v\rangle} \quad \text{and} \quad \angle(v, w) = \arccos\left(\frac{\langle v, w\rangle}{|v||w|}\right).$$

These definitions generalize Definition 3.2 and Equation 3.4 respectively.

If $\mathcal{V}$ is a subspace of $\mathbb{R}^n$ for some n, then the standard inner product on $\mathbb{R}^n$ restricts to an inner product on $\mathcal{V}$. But this is not the only inner product on $\mathcal{V}$. In fact, given any basis of $\mathcal{V}$, there exists a unique inner product on $\mathcal{V}$ with respect to which this basis is orthonormal (Exercise 10.26). Different inner products yield different calculations of norms and angles.

Definition 10.22. *Let M be a manifold. A **Riemannian metric** on M is a choice for each $p \in M$ of an inner product, $\langle\cdot,\cdot\rangle_p$, on T_pM that “varies smoothly with p” in the following sense: for any pair, X and Y, of smooth vector fields on M, the map $p \mapsto \langle X(p), Y(p)\rangle_p$ is a smooth function from M to $\mathbb{R}$. A manifold together with a Riemannian metric is called a **Riemannian manifold**.*

A Riemannian metric on M empowers one to compute norms of (and angles between) its tangent vectors at any point. For example, the norm of $v \in T_pM$ is defined as $|v|_p = \sqrt{\langle v, v\rangle_p}$. Building on this, one can define the **length** of a smooth path $\gamma : [a, b] \to M$ to equal

$\int_a^b |\gamma'(t)|_{\gamma(t)}\,\mathrm{dt}$. The **distance** between a pair of points $p, q \in M$ can then be defined as the infimum length of all smooth paths $\gamma : [a, b] \to M$ with $\gamma(a) = p$ and $\gamma(b) = q$. When M is path-connected, this infimum is always finite, and this distance function makes M into a metric space. The topology induced by this distance function is the same as the topology described in Definition 10.2. The reader will prove these assertions in Exercise 10.27.

The natural notion of equivalence for Riemannian manifolds is:

Definition 10.23. *A diffeomorphism $f : M_1 \to M_2$ between Riemannian manifolds is called an* ***isometry*** *if its derivative preserves inner products; that is, for all $p \in M_1$ and all $v, w \in T_pM_1$, we have*

$$\langle df_p(v), df_p(w)\rangle_{f(p)} = \langle v, w\rangle_p.$$

Two Riemannian manifolds are called ***isometric*** *if there exists an isometry between them.*

A diffeomorphism between Riemannian manifolds is an isometry if and only if it preserves distances (Exercise 10.28), so this definition is consistent with the common use of the word "isometry" within the study of metric spaces.

If $M \subset \mathbb{R}^m$ is an embedded manifold, then each of its tangent spaces is a subspace of $\mathbb{R}^m$. The natural **embedded Riemannian metric** on M just means the restriction to each tangent space of the standard inner product on $\mathbb{R}^m$. As previously mentioned, every matrix group $G \subset GL_n(\mathbb{K}) \subset M_n(\mathbb{K})$ is an embedded manifold and therefore inherits a natural embedded Riemannian metric. When $G \subset \mathcal{O}_n(\mathbb{K})$, this embedded Riemannian metric has special symmetry properties that were studied in Section 4 of Chapter 8.

The **Nash Embedding Theorem** states that every Riemannian manifold is isometric to an embedded manifold with the embedded Riemannian metric. This means that the class of generalized manifolds with arbitrary Riemannian metrics is really no more general than the class of embedded manifolds with embedded metrics. However, some of the Riemannian manifolds we will describe in this section would not look elegant or simple if one attempted to describe them as embedded.

Our next goal is to describe a natural Riemannian metric on certain coset spaces of the form G/H. The metric on G/H will be contrived such that the projection map $\pi : G \to G/H$ preserves the Riemannian metrics. What should this mean? A smooth function from a higher-dimensional Riemannian manifold to a lower-dimensional Riemannian manifold could never be an isometry, yet there is a meaningful sense in which it might preserve the Riemannian metrics:

Definition 10.24. *Let $f : M \to B$ be a smooth function between Riemannian manifolds with $k = dim(M) - dim(B) > 0$. Then f is called a **Riemannian submersion** if for all $p \in M$:*

(1) *$df_p : T_pM \to T_{f(p)}B$ is surjective (which implies that the kernel of df_p has dimension k).*
(2) *For all $v, w \in T_pM$ orthogonal to the kernel of df_p,*

$$\langle df_p(v), df_p(w)\rangle_{f(p)} = \langle v, w\rangle_p.$$

Theorem 10.25. *If $G \subset \mathcal{O}_n(\mathbb{K})$ is a matrix group and $H \subset G$ is a closed subgroup, there exists a unique Riemannian metric on G/H (called the **submersion metric**) such that the projection map $\pi : G \to G/H$ is a Riemannian submersion.*

The distance function induced by this Riemannian metric on G/H is quite natural. It can be shown that the distance between a pair of cosets in G/H can be computed by regarding the cosets as subsets of G and calculating the infimum distance within G between a point of the first coset and a point of the second.

Proof. Let $g \in G$ be an arbitrary element, so $[g] \in G/H$ is an arbitrary coset. The parametrization φ_g defined in Theorem 10.8 provides a natural identification $F_g : \mathfrak{p} \to T_{[g]}(G/H)$, namely:

$$F_g(X) = \frac{d}{dt}\bigg|_{t=0} [g \cdot \exp(tX)] \tag{10.6}$$

for all $X \in \mathfrak{p}$.

As in Section 4 of Chapter 8, we will denote by $\langle\cdot,\cdot\rangle_{\mathbb{R}}$ the inner product on $\mathfrak{g}$ coming from the embedded Riemannian metric on G. We will use the same notation for the restriction of this inner product to the subspace $\mathfrak{p} \subset \mathfrak{g}$. Via the identification F_g, this inner product on

$\mathfrak{p}$ naturally induces an inner product $T_{[g]}(G/H)$. In other words, an arbitrary pair of vectors in $T_{[g]}(G/H)$ can be written as $F_g(X), F_g(Y)$ for some $X, Y \in \mathfrak{p}$, and we define:

$$\langle F_g(X), F_g(Y)\rangle_{[g]} = \langle X, Y\rangle_{\mathbb{R}}.$$

We must prove that the inner product on $T_{[g]}(G/H)$ constructed in this manner does not depend on the choice of representative, g, of the coset $[g]$. An alternative representative would have the form $g \cdot h$ for some $h \in H$. It is straightforward to verify that for all $X \in \mathfrak{p}$:

$$F_{gh}(X) = F_g(\mathrm{Ad}_h X).$$

So the inner product is well-defined because for all $X, Y \in \mathfrak{p}$,

$$\begin{aligned}\langle F_{gh}(X), F_{gh}(Y)\rangle_{[gh]} &= \langle X, Y\rangle_{\mathbb{R}} \\ &= \langle \mathrm{Ad}_h X, \mathrm{Ad}_h Y\rangle_{\mathbb{R}} = \langle F_g(\mathrm{Ad}_h X), F_g(\mathrm{Ad}_h Y)\rangle_{[g]}.\end{aligned}$$

The second equality above reflects the Ad-invariant property of $\langle\cdot,\cdot\rangle_{\mathbb{R}}$ that was described in Section 4 of Chapter 8.

We leave to the reader in Exercise 10.29 the verification that this Riemannian metric varies smoothly with the point and that $\pi : G \to G/H$ is a Riemannian submersion. □

We end this section by describing the natural symmetry property of the submersion metrics constructed in the previous proof. For this, we must first specify what it means for an action to respect a Riemannian metric:

Definition 10.26. *A smooth action, φ, of a matrix group G on a Riemannian manifold M is called an* ***isometric action*** *if for every $g \in G$, the diffeomorphism $\varphi(g) : M \to M$ is an isometry. A Riemannian manifold, M, is called a* ***Riemannian homogeneous manifold*** *if there exists a transitive smooth isometric action of a matrix group on M.*

The Riemannian version of Example 10.14 is:

Proposition 10.27. *Under the assumptions of Theorem 10.25, the left action of G on G/H is an isometric action, so G/H is a Riemannian homogeneous manifold (with respect to the submersion metric).*

Proof. Let $g \in G$ and denote by $\mathcal{L}_g : (G/H) \to (G/H)$ the left-multiplication map, defined as $\mathcal{L}_g([x]) = [g \cdot x]$. We must prove that for all $x \in G$, the linear transformation

$$d(\mathcal{L}_g)_{[x]} : T_{[x]}(G/H) \to T_{[gx]}(G/H)$$

preserves inner products. As in the previous proof, $\mathfrak{p}$ is identified with both of these tangent spaces such that the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{p} & \xrightarrow{\text{identity}} & \mathfrak{p} \\ {\scriptstyle F_x}\downarrow & & \downarrow{\scriptstyle F_{gx}} \\ T_{[x]}(G/H) & \xrightarrow{d(\mathcal{L}_g)_{[x]}} & T_{[gx]}(G/H) \end{array} \tag{10.7}$$

Since the inner products on these tangent spaces are defined via these identifications, it follows that $d(\mathcal{L}_g)_{[x]}$ preserves inner products. □

Combining these results with Corollary 10.20 provides the structure of a Riemannian homogeneous manifold on all of the projective spaces: $\mathbb{RP}^n$, $\mathbb{CP}^n$, and $\mathbb{HP}^n$. Spheres and projective spaces are among the most natural Riemannian manifolds.

7. Lie groups

Essentially all of the theory of matrix groups in this book is also true for a generalization of matrix groups called *Lie groups*. In this section, we briefly overview (without proofs) the structures and theorems for matrix groups that carry over to Lie groups. We begin with the definition:

Definition 10.28. *A **Lie group** is a manifold, G, with a smooth group operation $G \times G \to G$.*

In other words, a Lie group is a manifold that is also a group. Many authors add to this definition the requirement that the "inverse map" $G \to G$, sending $g \mapsto g^{-1}$, is smooth; however, this turns out to be a consequence of the smoothness of the group operation $(g_1, g_2) \mapsto g_1 \cdot g_2$.

In Chapter 7, we proved that matrix groups are (embedded) manifolds. The group operation is smooth, so *matrix groups are Lie groups.*

All important structures of matrix groups carry over to Lie groups. For example, the Lie algebra, $\mathfrak{g}$, of a Lie group G is defined as you would expect:

$$\mathfrak{g} = T_I G.$$

For every $g \in G$, the conjugation map $C_g : G \to G$ sending $x \mapsto gxg^{-1}$ is smooth, so one can define:

$$Ad_g = d(C_g)_I : \mathfrak{g} \to \mathfrak{g}.$$

Next, the Lie bracket operation in $\mathfrak{g}$ is defined as you would expect: for $A, B \in \mathfrak{g}$,

$$[A, B] = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{a(t)} B,$$

where $a(t)$ is any differentiable path in G with $a(0) = I$ and with $a'(0) = A$. It turns out that this operation satisfies the familiar Lie bracket properties of Proposition 8.4. Next, the exponential map

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

is defined with inspiration from Proposition 6.10: For $A \in \mathfrak{g}$, the path $t \mapsto e^{tA}$ (at least for sufficiently small t) means the integral curve of the vector field on G whose value at $g \in G$ is $d(\mathcal{L}_g)_I(A) \in T_g G$, where $\mathcal{L}_g : G \to G$ denotes the map $x \mapsto g \cdot x$.

Moreover, every *compact* Lie group has a maximal torus, and the key facts from Chapter 9 generalize to this setting. In particular:

Proposition 10.29. *Let G be a path-connected compact Lie group, let T be a maximal torus of G, and let $\tau \subset \mathfrak{g}$ denote the Lie algebras of $T \subset G$.*

1. *For every $x \in G$ there exists $g \in G$ such that $x \in g \cdot T \cdot g^{-1}$.*
2. *Every maximal torus of G equals $g \cdot T \cdot g^{-1}$ for some $g \in G$.*
3. *If $x \in G$ commutes with every element of T, then $x \in T$.*
4. *If $X \in \tau$ commutes with every element of τ, then $X \in \tau$.*

In Chapter 9, we proved these claims for $SO(n)$, $U(n)$, $SU(n)$ and $Sp(n)$, with different proofs for each case. The proof techniques required for general compact Lie groups (or even general compact matrix groups) are a bit beyond the scope of this book.

Our previous results about coset spaces generalize to Lie groups. In particular, if G is a Lie group and $H \subset G$ is a closed subgroup, then the coset space G/H has a natural manifold structure with respect to which the left action of G on G/H, defined as $g\star[x] = [g\cdot x]$, is smooth and transitive. Conversely, if a path-connected Lie group G acts smoothly and transitively on a manifold M, then M is diffeomorphic to G/G_{p_0} for any $p_0 \in M$.

Lie groups are among the most fundamental objects in mathematics and physics, and the most significant Lie groups are matrix groups. In fact, Lie groups turn out to be only *slightly* more general than matrix groups. One indication of this is Ado's Theorem: the Lie algebra of any Lie group is isomorphic to the Lie algebra of a matrix group. Another indication is that every *compact* Lie group turns out to be smoothly isomorphic to a matrix group.

In fact, a major achievement of modern mathematics is the classification of compact Lie groups. The only such groups we have encountered so far are $SO(n)$, $O(n)$, $U(n)$, $SU(n)$, $Sp(n)$, and products of these, such as, for example, $SO(3)\times SO(5)\times SU(2)$. It turns out that there are not many more than these.

Theorem 10.30. *The Lie algebra of any compact Lie group is isomorphic to the Lie algebra of a product $G_1 \times G_2 \times \cdots \times G_k$, where each G_i is one of $\{SO(n), SU(n), Sp(n)\}$ for some n or is one of a list of five possible exceptions.*

The five "exceptional Lie groups" mentioned in the theorem are named:

G_2, which has dimension 14
F_4, which has dimension 54
E_6, which has dimension 78
E_7, which has dimension 133
E_8, which has dimension 248

We have seen that non-isomorphic Lie groups sometimes have isomorphic Lie algebras. For example, $U(n)$ is not on the list in Theorem 10.30 because it has the same Lie algebra as $SU(n)\times SO(2)$, by Exercise 4.21. The problem of determining all Lie groups with the same Lie algebra as $G_1\times G_2\times\cdots\times G_k$ is well-understood, but is also beyond the scope of this text. Aside from this detail, the theorem gives a complete classification of compact Lie groups!

Although we will not construct the exceptional Lie groups (which are all matrix groups) or fully prove Theorem 10.30, the next chapter will overview the key ideas of the proof. To pave the way, we will end this chapter with a key result: every compact Lie group G has a **bi-invariant metric**, which means a Riemannian metric for which the left and right actions of G on G are both isometric actions.

Although we will not fully prove this fact, we will describe how to reduce it to a more tractable assertion. Describing a Riemannian metric on a general manifold requires a lot of information: a different inner product must be chosen for each tangent space. Directly managing this much information is generally intractable. But on a Lie group, one can describe a metric simply by choosing an inner product on a *single* vector space, namely its Lie algebra:

Definition 10.31. *Let G be a Lie group with Lie algebra denoted $\mathfrak{g}$. Any inner product, $\langle\cdot,\cdot\rangle_I$, on $\mathfrak{g}$ determines a Riemannian metric on G in either of the following ways:*

(1) ***(Left-invariant metric)*** *For all $g\in G$ and $v,w\in T_gG$, define:*

$$\langle v,w\rangle_g=\langle d(\mathcal{L}_{g^{-1}})_g(v),d(\mathcal{L}_{g^{-1}})_g(v)\rangle_I,$$

where $\mathcal{L}_g:G\to G$ is defined as $\mathcal{L}_g(x)=g\cdot x$.

(2) ***(Right-invariant metric)*** *For all $g\in G$ and $v,w\in T_gG$, define:*

$$\langle v,w\rangle_g=\langle d(\mathcal{R}_{g^{-1}})_g(v),d(\mathcal{R}_{g^{-1}})_g(v)\rangle_I,$$

where $\mathcal{R}_g:G\to G$ is defined as $\mathcal{R}_g(x)=x\cdot g$.

The idea is that for each $g\in G$, $\mathfrak{g}$ is naturally identified with T_gG via either $d(\mathcal{L}_g)_I$ or $d(\mathcal{R}_g)_I$. Via either of these identifications, an inner product on $\mathfrak{g}$ induces an inner product on all other tangent spaces. These two choices are named after the symmetries they possess. With respect to a left-invariant metric on a Lie group G, the left action of G on G is isometric. Similarly, with respect to a right-invariant metric, the right action of G on G is isometric.

For a *compact* Lie group, we don't have to choose:

Theorem 10.32. *If G is a compact Lie group with Lie algebra denoted $\mathfrak{g}$, then there exists an inner product, $\langle\cdot,\cdot\rangle$, on $\mathfrak{g}$ satisfying the following equivalent conditions:*

(1) *(**Ad-invariance**) For all $g \in G$ and all $v, w \in \mathfrak{g}$,*

$$\langle Ad_g(v), Ad_g(w)\rangle = \langle v, w\rangle.$$

(2) *(**Bi-invariance**) The left-invariant Riemannian metric on G determined by $\langle\cdot,\cdot\rangle$ is the same as the right-invariant Riemannian metric (and is therefore called a **bi-invariant metric**).*

The equivalence of (1) and (2) is straightforward (Exercise 10.31), but it requires some new ideas to prove that an Ad-invariant inner product exists. We will not address the proof here.

For the compact matrix group $\mathcal{O}_n(\mathbb{K})$, the inner product inherited from the ambient Euclidean space, previously denoted $\langle\cdot,\cdot\rangle_{\mathbb{R}}$, was proven in Section 4 of Chapter 8 to be Ad-invariant by first re-describing it as: $\langle X, Y\rangle_{\mathbb{R}} = \text{Real}(X \cdot Y^*)$. Thus, Theorem 10.32 generalizes a familiar structure on $\mathcal{O}_n(\mathbb{K})$ to arbitrary compact Lie groups.

We end this section by stating without proof a powerful generalization of Theorem 10.8 and Theorem 10.25:

Theorem 10.33. *Given a smooth free isometric left action of a compact Lie group G on a Riemannian manifold M, the orbit space, $G\backslash M$, has a unique manifold structure and Riemannian metric such that the projection $\pi : M \to G\backslash M$ is a smooth Riemannian submersion.*

8. Exercises

Ex. 10.1. Prove that an embedded manifold (Definition 7.14) is a manifold (Definition 10.1).

Ex. 10.2. Prove the assertion in Section 1 that an atlas satisfying conditions (1) and (2) of Definition 10.1 can be uniquely completed to an atlas that also satisfies condition (3).

Ex. 10.3. Prove the chain rule (Proposition 7.21) for general (not necessarily embedded) manifolds.

Ex. 10.4. Prove the Inverse Function Theorem (Theorem 7.22) for general (not necessarily embedded) manifolds.

Ex. 10.5. If M_1 and M_2 are (generalized) manifolds, describe a natural manifold structure on their product, $M_1 \times M_2$, such that

$$\dim(M_1 \times M_2) = \dim(M_1) + \dim(M_2).$$

Prove that the projection $f_1 : M_1 \times M_2 \to M_1$, defined as $f(p_1, p_2) = p_1$, is smooth, as is the analogously defined projection f_2.

Ex. 10.6. State and prove generalizations of Exercises 7.11, 7.12 and 7.13 for general (not necessarily embedded) manifolds.

Ex. 10.7. Let X be a smooth vector field on a manifold, M. Prove that for *any* parametrization, $\varphi : U \subset \mathbb{R}^d \to V \subset M$, the naturally associated function from U to $\mathbb{R}^d$ is smooth.

Ex. 10.8. Prove that $\mathbb{RP}^1$ is diffeomorphic to $SO(2)$.

Ex. 10.9. Prove the assertion in Proposition 10.6 that $\mathbb{HP}^1$ is diffeomorphic to S^4.

Ex. 10.10. Prove Proposition 10.10.

Ex. 10.11. Let $H \subset K \subset G$ be nested closed matrix groups. Prove that the natural map $f : G/H \to G/K$, defined as $f(g \cdot H) = g \cdot K$, is smooth and that the preimage of any element of G/K is diffeomorphic to K/H.

Ex. 10.12. Let φ be a right action of a group G on a set M. Define $\overline{\varphi}(g) = \varphi(g^{-1})$ for all $g \in G$. Prove that $\overline{\varphi}$ is a left action of G on M.

Ex. 10.13. For a transitive action of a group G on a set M, prove that any two stabilizers are conjugate.

Ex. 10.14. For a smooth action of a matrix group G on a manifold M, prove that all stabilizers are *closed* subgroups of G.

Ex. 10.15. Prove Lemma 10.17.

Ex. 10.16. In Example 10.14, prove the assertion that the natural left action of G on G/H is smooth.

Ex. 10.17. For positive integers $m < n$, prove that the Grassmann manifolds $G_m(\mathbb{K}^n)$ and $G_{n-m}(\mathbb{K}^n)$ are diffeomorphic.

Ex. 10.18. With notation as in Corollary 10.20, find H such that

$$\mathbb{KP}^n = GL_n(\mathbb{K})/H = SL_n(\mathbb{K})/S(H).$$

Ex. 10.19. With notation as in Corollary 10.20, find H such that

$$T^1S^n = O(n+1)/H$$

(see Exercise 7.13 for the definition of the unit tangent bundle).

Ex. 10.20. With notation as in Corollary 10.20, show that:

$$G_m(\mathbb{R}^n) = O(n)/\left(O(m) \times O(n-m)\right).$$

Ex. 10.21 (Flag manifolds)**.** A **flag** in $\mathbb{C}^n$ means a collection of nested $\mathbb{C}$-subspaces:

$$V_1 \subset V_2 \subset \cdots \subset V_{n-1}$$

such that $\dim(V_i) = i$ for each i. Let F denote the set of all flags in $\mathbb{C}^n$. Verify that the left action of $SU(n)$ on F defined as

$$g \star (V_1 \subset V_2 \subset \cdots \subset V_{n-1}) = (g \cdot V_1 \subset g \cdot V_2 \subset \cdots \subset g \cdot V_{n-1})$$

is transitive and that every stabilizer is a maximal torus of $SU(n)$.

Ex. 10.22. One of the $n = 1$ cases of Corollary 10.20 says that $\mathbb{CP}^1 = SU(2)/S(U(1)\times U(1))$. How is this transitive action of $SU(2)$ on $\mathbb{CP}^1$ related to the adjoint action of $Sp(1)$ on S^2 whereby the double cover $Sp(1) \to SO(3)$ was constructed in Section 6 of Chapter 8?

Ex. 10.23. Prove that $S(U(1) \times U(n))$ is smoothly isomorphic to $U(n)$ and that $S(\{\pm 1\} \times O(n))$ is smoothly isomorphic to $O(n)$. *Comment: These groups are mentioned in Corollary* 10.20.

Ex. 10.24. For a homogeneous manifold $M = G/H$, the parametrization φ_I defined in Theorem 10.8 provides a natural identification of $\mathfrak{p}$ with $T_{[I]}(G/H)$ (this identification was named F_I in Equation 10.6). Explicitly describe $\mathfrak{p}$ for the following examples:

(1) $S^4 = SO(5)/(\{1\} \times SO(4))$.

(2) $\mathbb{CP}^2 = U(3)/(U(1) \times U(2))$.

(3) $\mathbb{CP}^2 = SU(3)/S(U(1) \times U(2))$.

(4) $G_2(\mathbb{R}^5) = GL_5(\mathbb{R})/H_0$ as in Equation 10.4.

(5) $G_2(\mathbb{R}^5) = O(5)/\left(O(2) \times O(3)\right)$ as in Exercise 10.20.

(6) $S_2(\mathbb{C}^5) = GL_5(\mathbb{C})/H$ as in Example 10.18.

Ex. 10.25. Let $\{e_1, e_2, e_3, e_4, e_5\}$ denote the standard orthonormal basis of $\mathbb{R}^5$. Let $V_0 = \text{span}\{e_1, e_2\} \subset \mathbb{R}^5$, so $V_0^\perp = \text{span}\{e_3, e_4, e_5\}$. Let $t \mapsto f(t)$ be a differentiable one-parameter family of linear transformations from V_0 to $V_0^\perp$, which can be expressed with respect to the above bases as a one-parameter family of matrices:

$$f(t) = \begin{pmatrix} f_{11}(t) & f_{12}(t) & f_{13}(t) \\ f_{21}(t) & f_{22}(t) & f_{23}(t) \end{pmatrix},$$

where each f_{ij} is a differentiable function with $f_{ij}(0) = 0$. Define:

$$V(t) = \text{the graph of } f(t) = \{v + f(t)(v) \mid v \in V_0\},$$

so that $t \mapsto V(t)$ is a path in $G_2(\mathbb{R}^5)$ whose initial derivative, $V'(0) \in T_{V_0}(G_2(\mathbb{R}^5))$, is determined by the six elements of $f'(0)$. How is this related to the six real numbers used to describe $\mathfrak{p}$ in parts (4) or (5) of Exercise 10.24?

Ex. 10.26. Let $\mathcal{V}$ be a finite-dimensional vector space over $\mathbb{R}$. Given any basis of $\mathcal{V}$, prove there exists a unique inner product on $\mathcal{V}$ with respect to which the basis is orthonormal.

Ex. 10.27. If M is a path-connected Riemannian manifold, prove that the distance between any pair of points is finite and that the distance function makes M into a metric space. Prove that the topology induced by this distance function is the same as the topology described in Definition 10.2.

Ex. 10.28. Prove that a diffeomorphism between Riemannian manifolds is an isometry if and only if it preserves distances.

Ex. 10.29. Complete the proof of Theorem 10.25 by proving that the Riemannian metric varies smoothly with the point and that $\pi : G \to G/H$ is a Riemannian submersion.

Ex. 10.30. If G is a matrix group and $H \subset G$ is a closed *normal* subgroup, prove that G/H is a Lie group.

Ex. 10.31. In Theorem 10.32, prove that (1) and (2) are equivalent.

Ex. 10.32. Prove that an Ad-invariant inner product, $\langle \cdot, \cdot \rangle$, on the Lie algebra of a Lie group G must satisfy the following **infinitesimal Ad-invariance** property: for all $A, B, C \in \mathfrak{g}$,

$$\langle [A, B], C \rangle = -\langle [A, C], B \rangle.$$

Hint: Copy the proof of Proposition 8.14*.*

Ex. 10.33 (An exotic sphere)**.** Consider the smooth left action of the matrix group $H = Sp(1) \times Sp(1)$ on the manifold $G = Sp(2)$ defined so that for all $q_1, q_2 \in Sp(1)$ and all $A \in Sp(2)$:

$$(q_1, q_2) \star A = \begin{pmatrix} q_1 & 0 \\ 0 & q_1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} \overline{q_2} & 0 \\ 0 & 1 \end{pmatrix}.$$

Prove that this action is free.

Comment: Theorem 10.33 *provides the orbit space with the structure of a* 7*-dimensional Riemannian manifold. This orbit space has been proven to be an* ***exotic sphere****: it is homeomorphic but not diffeomorphic to* S^7*.*