

Preface

It seems to have been decided that undergraduate mathematics today rests on two foundations: calculus and linear algebra. These may not be the best foundations for, say, number theory or combinatorics, but they serve quite well for undergraduate analysis and several varieties of undergraduate algebra and geometry. The really *perfect* sequel to calculus and linear algebra, however, would be a blend of the two—a subject in which calculus throws light on linear algebra and vice versa. Look no further! This perfect blend of calculus and linear algebra is Lie theory (named to honor the Norwegian mathematician Sophus Lie—pronounced “Lee”). So why is Lie theory not a standard undergraduate topic?

The problem is that, until recently, Lie theory was a subject for mature mathematicians or else a tool for chemists and physicists. There was no Lie theory for novice mathematicians. Only in the last few years have there been serious attempts to write Lie theory books for undergraduates. These books broke through to the undergraduate level by making some sensible compromises with generality; they stick to matrix groups and mainly to the classical ones, such as rotation groups of n -dimensional space.

In this book I stick to similar subject matter. The classical groups are introduced via a study of rotations in two, three, and four dimensions, which is also an appropriate place to bring in complex numbers and quaternions. From there it is only a short step to studying rotations in real, complex, and quaternion spaces of any dimension. In so doing, one has introduced the classical simple Lie groups, in their most geometric form, using only basic linear algebra. Then calculus intervenes to find the tangent spaces of the classical groups—their Lie algebras—and to move back and forth between the group and its algebra via the log and exponential functions. Again, the basics suffice: single-variable differentiation and the Taylor series for e^x and $\log(1+x)$.

Where my book diverges from the others is at the next level, the miraculous level where one discovers that the (curved) structure of a Lie group is almost completely captured by the structure of its (flat) Lie algebra. At this level, the other books retain many traces of the sophisticated approach to Lie theory. For example, they rely on deep ideas from outside Lie theory, such as the inverse function theorem, existence theorems for ODEs, and representation theory. Even inside Lie theory, they depend on the Killing form and the whole root system machine to prove simplicity of the classical Lie algebras, and they use everything under the sun to prove the Campbell–Baker–Hausdorff theorem that lifts structure from the Lie algebra to the Lie group. But actually, proving simplicity of the classical Lie algebras can be done by basic matrix arithmetic, and there is an amazing elementary proof of Campbell–Baker–Hausdorff due to Eichler [1968].

The existence of these little-known elementary proofs convinced me that a naive approach to Lie theory is possible and desirable. The aim of this book is to carry it out—developing the central concepts and results of Lie theory by the simplest possible methods, mainly from single-variable calculus and linear algebra. Familiarity with elementary group theory is also desirable, but I provide a crash course on the basics of group theory in Sections 2.1 and 2.2.

The naive approach to Lie theory is due to von Neumann [1929], and it is now possible to streamline it by using standard results of undergraduate mathematics, particularly the results of linear algebra. Of course, there is a downside to naiveté. It is probably not powerful enough to prove some of the results for which Lie theory is famous, such as the classification of the simple Lie algebras and the discovery of the five exceptional algebras.¹ To compensate for this lack of technical power, the end-of-chapter discussions introduce important results beyond those proved in the book, as part of an informal sketch of Lie theory and its history. It is also true that the naive methods do not afford the same insights as more sophisticated methods. But they offer another insight that is often undervalued—some important theorems are not as difficult as they look! I think that all mathematics students appreciate this kind of insight.

In any case, my approach is not entirely naive. A certain amount of topology is essential, even in basic Lie theory, and in Chapter 8 I take

¹I say so from painful experience, having entered Lie theory with the aim of understanding the exceptional groups. My opinion now is that the Lie theory that precedes the classification is a book in itself.

the opportunity to develop all the appropriate concepts from scratch. This includes everything from open and closed sets to simple connectedness, so the book contains in effect a minicourse on topology, with the rich class of multidimensional examples that Lie theory provides. Readers already familiar with topology can probably skip this chapter, or simply skim it to see how Lie theory influences the subject. (Also, if time does not permit covering the whole book, then the end of Chapter 7 is a good place to stop.)

I am indebted to Wendy Baratta, Simon Goberstein, Brian Hall, Rohan Hewson, Chris Hough, Nathan Jolly, David Kramer, Jonathan Lough, Michael Sun, Marc Ryser, Abe Shenitzer, Paul Stanford, Fan Wu and the anonymous referees for many corrections and comments. As usual, my wife, Elaine, served as first proofreader; my son Robert also served as the model for Figure 8.7. Thanks go to Monash University for the opportunity to teach courses from which this book has grown, and to the University of San Francisco for support while writing it.

Finally, a word about my title. Readers of a certain age will remember the book *Naive Set Theory* by Paul Halmos—a lean and lively volume covering the parts of set theory that all mathematicians ought to know. Paul Halmos (1916–2006) was my mentor in mathematical writing, and I dedicate this book to his memory. While not attempting to emulate his style (which is inimitable), I hope that *Naive Lie Theory* can serve as a similar introduction to Lie groups and Lie algebras. Lie theory today has become the subject that all mathematicians ought to know something about, so I believe the time has come for a naive, but mathematical, approach.

John Stillwell

University of San Francisco, December 2007

Monash University, February 2008

2

Groups

PREVIEW

This chapter begins by reviewing some basic group theory—subgroups, quotients, homomorphisms, and isomorphisms—in order to have a basis for discussing Lie groups in general and simple Lie groups in particular.

We revisit the group \mathbb{S}^3 of unit quaternions, this time viewing its relation to the group $\text{SO}(3)$ as a 2-to-1 homomorphism. It follows that \mathbb{S}^3 is *not a simple group*. On the other hand, $\text{SO}(3)$ is *simple*, as we show by a direct geometric proof.

This discovery motivates much of Lie theory. There are infinitely many simple Lie groups, and most of them are generalizations of rotation groups in some sense. However, deep ideas are involved in identifying the simple groups and in showing that we have enumerated them all.

To show why it is not easy to identify all the simple Lie groups we make a special study of $\text{SO}(4)$, the rotation group of \mathbb{R}^4 . Like $\text{SO}(3)$, $\text{SO}(4)$ can be described with the help of quaternions. But a rotation of \mathbb{R}^4 generally depends on *two* quaternions, and this gives $\text{SO}(4)$ a special structure, related to the *direct product* of \mathbb{S}^3 with itself. In particular, it follows that $\text{SO}(4)$ is *not simple*.

2.1 Crash course on groups

For readers who would like a reminder of the basic properties of groups, here is a crash course, oriented toward the kind of groups studied in this book. Even those who have not seen groups before will be familiar with the computational tricks—such as canceling by multiplying by the inverse—since they are the same as those used in matrix computations.

First, a group G is a set with “product” and “inverse” operations, and an identity element 1 , with the following three basic properties:

$$\begin{aligned} g_1(g_2g_3) &= (g_1g_2)g_3 && \text{for all } g_1, g_2, g_3 \in G, \\ g1 &= 1g = g && \text{for all } g \in G, \\ gg^{-1} &= g^{-1}g = 1 && \text{for all } g \in G. \end{aligned}$$

It should be mentioned that 1 is the *unique* element g' such that $gg' = g$ for all $g \in G$, because multiplying the equation $gg' = g$ on the left by g^{-1} gives $g' = 1$. Similarly, for each $g \in G$, g^{-1} is the unique element g'' such that $gg'' = 1$.

The above notation for “product,” “inverse,” and “identity” is called *multiplicative notation*. It is used (sometimes with I , e , or $\mathbf{1}$ in place of 1) for groups of numbers, quaternions, matrices, and all other groups whose operation is called “product.” There are a few groups whose operation is called “sum,” such as \mathbb{R}^n under vector addition. For these we use *additive notation*: $g_1 + g_2$ for the “sum” of $g_1, g_2 \in G$, $-g$ for the inverse of $g \in G$, and 0 (or $\mathbf{0}$) for the identity of G . Additive notation is used only when G is *abelian*, that is, when $g_1 + g_2 = g_2 + g_1$ for all $g_1, g_2 \in G$.

Since groups are generally not abelian, we have to speak of multiplying h by g “on the left” or “on the right,” because gh and hg are generally different. If we multiply *all* members g' of a group G on the left by a particular $g \in G$, we get back all the members of G , because for any $g'' \in G$ there is a $g' \in G$ such that $gg' = g''$ (namely $g' = g^{-1}g''$).

Subgroups and cosets

To study a group G we look at the groups H contained in it, the *subgroups* of G . For each subgroup H of G we have a decomposition of G into disjoint pieces called the (left or right) *cosets* of H in G . The left cosets (which we stick with, for the sake of consistency) are the sets of the form

$$gH = \{gh : h \in H\}.$$

Thus H itself is the coset for $g = 1$, and in general a coset gH is “ H translated by g ,” though one cannot usually take the word “translation” literally. One example for which this is literally true is G the plane \mathbb{R}^2 of points (x,y) under vector addition, and H the subgroup of points $(0,y)$. In this case we use additive notation and write the coset of (x,y) as

$$(x,y) + H = \{(x,y) : y \in \mathbb{R}\}, \quad \text{where } x \text{ is constant.}$$

Then H is the y -axis and the coset $(x,y) + H$ is H translated by the vector (x,y) (see Figure 2.1). This example also illustrates how a group G decom-

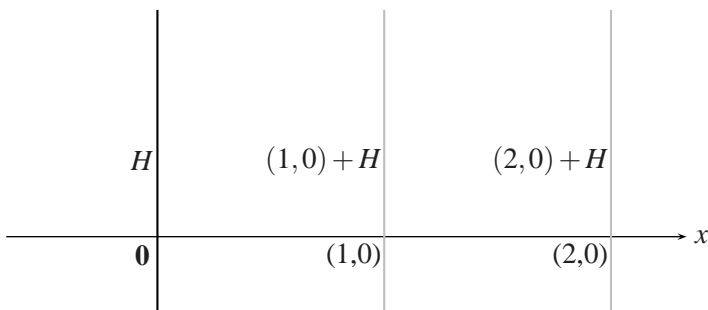


Figure 2.1: Subgroup H of \mathbb{R}^2 and cosets.

poses into *disjoint* cosets (decomposing the plane into parallel lines), and that different $g \in G$ can give the same coset gH . For example, $(1,0) + H$ and $(1,1) + H$ are both the vertical line $x = 1$.

Each coset gH is in 1-to-1 correspondence with H because we get back each $h \in H$ from $gh \in gH$ by multiplying on the left by g^{-1} . Different cosets are disjoint because if $g \in g_1H$ and $g \in g_2H$ then

$$g = g_1h_1 = g_2h_2 \quad \text{for some } h_1, h_2 \in H,$$

and therefore $g_1 = g_2h_2h_1^{-1}$. But then

$$g_1H = g_2h_2h_1^{-1}H = g_2(h_2h_1^{-1}H) = g_2H$$

because $h_2h_1^{-1} \in H$ and therefore $h_2h_1^{-1}H = H$ by the remark at the end of the last subsection (that multiplying a group by one of its members gives back the group). Thus if two cosets have an element in common, they are identical.

This algebraic argument has surprising geometric consequences; for example, a filling of \mathbb{S}^3 by disjoint circles known as the *Hopf fibration*. Figure 2.2 shows some of the circles, projected stereographically into \mathbb{R}^3 . The circles fill nested torus surfaces, one of which is shown in gray.

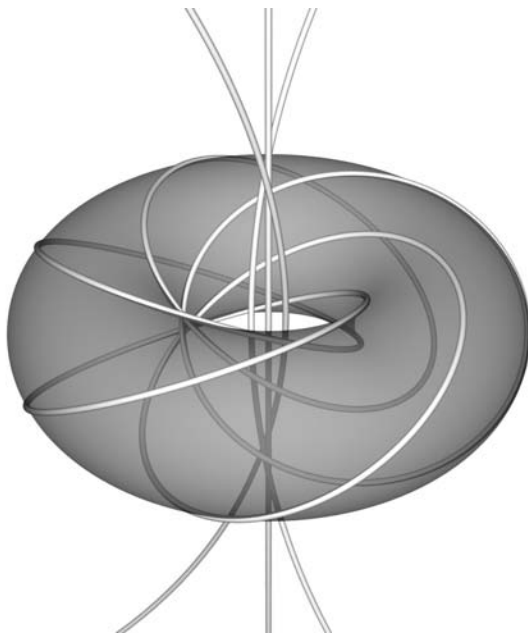


Figure 2.2: Some circles in the Hopf fibration.

Proposition: \mathbb{S}^3 can be decomposed into disjoint congruent circles.

Proof. As we saw in Section 1.3, the quaternions $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ of unit length satisfy

$$a^2 + b^2 + c^2 + d^2 = 1,$$

and hence they form a 3-sphere \mathbb{S}^3 . The unit quaternions also form a group G , because the product and inverse of unit quaternions are also unit quaternions, by the multiplicative property of absolute value.

One subgroup H of G consists of the unit quaternions of the form $\cos \theta + \mathbf{i} \sin \theta$, and these form a unit circle in the plane spanned by 1 and \mathbf{i} . It follows that any coset qH is also a unit circle, because multiplication by a quaternion q of unit length is an isometry, as we saw in Section 1.4. Since the cosets qH fill the whole group and are disjoint, we have a decomposition of the 3-sphere into unit circles. \square

Exercises

An important nonabelian group (in fact, it is the simplest example of a nonabelian Lie group) is the group of functions of the form

$$f_{a,b}(x) = ax + b, \quad \text{where } a, b \in \mathbb{R} \text{ and } a > 0.$$

The group operation is function composition.

2.1.1 If $f_{a,b}(x) = f_{a_2,b_2}(f_{a_1,b_1}(x))$, work out a, b in terms of a_1, b_1, a_2, b_2 , and check that they are the same as the a, b determined by

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_2 & b_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 0 & 1 \end{pmatrix}.$$

2.1.2 Also show that the inverse function $f_{a,b}^{-1}(x)$ exists, and that it corresponds to the inverse matrix

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}^{-1}.$$

This correspondence between functions and matrices is a *matrix representation* of the group of functions $f_{a,b}$. We have already seen examples of matrix representations of groups—such as the rotation groups in two and three dimensions—and, in fact, most of the important Lie groups can be represented by matrices.

The unit complex numbers, $\cos \theta + i \sin \theta$, form a group $\text{SO}(2)$ that we began to study in Section 1.1. We now investigate its subgroups.

2.1.3 Other than the trivial group $\{1\}$, what is the smallest subgroup of $\text{SO}(2)$?

2.1.4 Show that there is exactly one n -element subgroup of $\text{SO}(2)$, for each natural number n , and list its members.

2.1.5 Show that the union R of all the finite subgroups of $\text{SO}(2)$ is also a subgroup (the group of “rational rotations”).

2.1.6 If z is a complex number not in the group R described in Exercise 2.1.5, show that the numbers $\dots, z^{-2}, z^{-1}, 1, z, z^2, \dots$ are all distinct, and that they form a subgroup of $\text{SO}(2)$.

2.2 Crash course on homomorphisms

Normal subgroups

Since $hg \neq gh$ in general, it can also be that $gH \neq Hg$, where

$$Hg = \{hg : h \in H\}$$

is the *right* coset of H . If $gH = Hg$ for all $g \in G$, we say that H is a *normal* subgroup of G . An equivalent statement is that H equals

$$g^{-1}Hg = \{g^{-1}hg : h \in H\} \quad \text{for each } g \in G.$$

(Because of this, it would be more sensible to call H “self-conjugate,” but unfortunately the overused word “normal” has stuck.)

The good thing about a normal subgroup H is that its cosets themselves form a group when “multiplied” by the rule that “the coset of g_1 , times the coset of g_2 , equals the coset of g_1g_2 ”:

$$g_1H \cdot g_2H = g_1g_2H.$$

This rule makes sense for a normal subgroup H because if $g'_1H = g_1H$ and $g'_2H = g_2H$ then $g'_1g'_2H = g_1g_2H$ as follows:

$$\begin{aligned} g'_1g'_2H &= g'_1Hg'_2 && \text{since } g'_2H = Hg'_2 \text{ by normality,} \\ &= g_1Hg'_2 && \text{since } g'_1H = g_1H \text{ by assumption,} \\ &= g_1g'_2H && \text{since } g'_2H = Hg'_2 \text{ by normality,} \\ &= g_1g_2H && \text{since } g'_2H = g_2H \text{ by assumption.} \end{aligned}$$

The group of cosets is called the *quotient group of G by H* , and is written G/H . (When G and H are finite, the size of G/H is indeed the size of G divided by the size of H .) We reiterate that the quotient group G/H exists *only* when H is a normal subgroup. Another, more efficient, way to describe this situation is in terms of *homomorphisms*: structure-preserving maps from one group to another.

Homomorphisms and isomorphisms

When H is a normal subgroup of G , the map $\varphi : G \rightarrow G/H$ defined by

$$\varphi(g) = gH \quad \text{for all } g \in G$$

preserves products in the sense that

$$\varphi(g_1g_2) = \varphi(g_1) \cdot \varphi(g_2).$$

This follows immediately from the definition of product of cosets, because

$$\varphi(g_1g_2) = g_1g_2H = g_1H \cdot g_2H = \varphi(g_1) \cdot \varphi(g_2).$$

In general, a map $\varphi : G \rightarrow G'$ of one group into another is called a *homomorphism* (from the Greek for “similar form”) if it preserves products. A group homomorphism indeed preserves group structure, because it not only preserves products, but also the identity and inverses. Here is why:

- Since $g = 1g$ for any $g \in G$, we have

$$\varphi(g) = \varphi(1g) = \varphi(1)\varphi(g) \quad \text{because } \varphi \text{ preserves products.}$$

Multiplying both sides on the right by $\varphi(g)^{-1}$ then gives $1 = \varphi(1)$.

- Since $1 = gg^{-1}$ for any $g \in G$, we have

$$1 = \varphi(1) = \varphi(gg^{-1}) = \varphi(g)\varphi(g^{-1})$$

because φ preserves products.

This says that $\varphi(g^{-1}) = \varphi(g)^{-1}$, because the inverse of $\varphi(g)$ is unique.

Thus the image $\varphi(G)$ is of “similar” form to G , but we say that G' is *isomorphic* (of the “same form”) to G only when the map φ is 1-to-1 and onto (in which case we call φ an *isomorphism*). In general, $\varphi(G)$ is only a shadow of G , because many elements of G may map to the same element of G' . The case furthest from isomorphism is that in which φ sends *all* elements of G to 1.

Any homomorphism φ of G onto G' can be viewed as the special type $\varphi : G \rightarrow G/H$. The appropriate normal subgroup H of G is the so-called *kernel* of φ :

$$H = \ker \varphi = \{g \in G : \varphi(g) = 1\}.$$

Then G' is isomorphic to the group $G/\ker \varphi$ of cosets of $\ker \varphi$ because:

1. $\ker \varphi$ is a group, because

$$\begin{aligned} h_1, h_2 \in \ker \varphi &\Rightarrow \varphi(h_1) = \varphi(h_2) = 1 \\ &\Rightarrow \varphi(h_1)\varphi(h_2) = 1 \\ &\Rightarrow \varphi(h_1h_2) = 1 \\ &\Rightarrow h_1h_2 \in \ker \varphi \end{aligned}$$

and

$$\begin{aligned}
 h \in \ker \varphi &\Rightarrow \varphi(h) = 1 \\
 &\Rightarrow \varphi(h)^{-1} = 1 \\
 &\Rightarrow \varphi(h^{-1}) = 1 \\
 &\Rightarrow h^{-1} \in \ker \varphi.
 \end{aligned}$$

2. $\ker \varphi$ is a normal subgroup of G , because, for any $g \in G$,

$$\begin{aligned}
 h \in \ker \varphi &\Rightarrow \varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g^{-1}) = \varphi(g)1\varphi(g)^{-1} = 1 \\
 &\Rightarrow ghg^{-1} \in \ker \varphi.
 \end{aligned}$$

Hence $g(\ker \varphi)g^{-1} = \ker \varphi$, that is, $\ker \varphi$ is normal.

3. Each $g' = \varphi(g) \in G'$ corresponds to the coset $g(\ker \varphi)$.

In fact, $g(\ker \varphi) = \varphi^{-1}(g')$, because

$$\begin{aligned}
 k \in \varphi^{-1}(g') &\Leftrightarrow \varphi(k) = g' \quad (\text{definition of } \varphi^{-1}) \\
 &\Leftrightarrow \varphi(k) = \varphi(g) \\
 &\Leftrightarrow \varphi(g)^{-1}\varphi(k) = 1 \\
 &\Leftrightarrow \varphi(g^{-1}k) = 1 \\
 &\Leftrightarrow g^{-1}k \in \ker \varphi \\
 &\Leftrightarrow k \in g(\ker \varphi).
 \end{aligned}$$

4. Products of elements of $g'_1, g'_2 \in G'$ correspond to products of the corresponding cosets:

$$g'_1 = \varphi(g_1), g'_2 = \varphi(g_2) \Rightarrow \varphi^{-1}(g'_1) = g_1(\ker \varphi), \varphi^{-1}(g'_2) = g_2(\ker \varphi)$$

by step 3. But also

$$\begin{aligned}
 g'_1 = \varphi(g_1), g'_2 = \varphi(g_2) &\Rightarrow g'_1 g'_2 = \varphi(g_1)\varphi(g_2) = \varphi(g_1 g_2) \\
 &\Rightarrow \varphi^{-1}(g'_1 g'_2) = g_1 g_2(\ker \varphi),
 \end{aligned}$$

also by step 3. Thus the product $g'_1 g'_2$ corresponds to $g_1 g_2(\ker \varphi)$, which is the product of the cosets corresponding to g'_1 and g'_2 respectively.

To sum up: a group homomorphism φ of G onto G' gives a 1-to-1 correspondence between G' and $G/(\ker \varphi)$ that preserves products, that is, G' is isomorphic to $G/(\ker \varphi)$.

This result is called the *fundamental homomorphism theorem for groups*.

The det homomorphism

An important homomorphism for real and complex matrix groups G is the *determinant* map

$$\det : G \rightarrow \mathbb{C}^\times,$$

where \mathbb{C}^\times denotes the multiplicative group of nonzero complex numbers. The determinant map is a homomorphism because \det is multiplicative— $\det(AB) = \det(A)\det(B)$ —a fact well known from linear algebra.

The kernel of \det , consisting of the matrices with determinant 1, is therefore a normal subgroup of G . Many important Lie groups arise in precisely this way, as we will see in Chapter 3.

Simple groups

A many-to-1 homomorphism of a group G maps it onto a group G' that is “simpler” than G (or, at any rate, not more complicated than G). For this reason, groups that admit no such homomorphism, other than the homomorphism sending all elements to 1, are called *simple*. Equivalently, *a nontrivial group is simple if it contains no normal subgroups other than itself and the trivial group*.

One of the main goals of group theory in general, and Lie group theory in particular, is to find all the simple groups. We find the first interesting example in the next section.

Exercises

2.2.1 Check that $z \mapsto z^2$ is a homomorphism of \mathbb{S}^1 . What is its kernel? What are the cosets of the kernel?

2.2.2 Show directly (that is, without appealing to Exercise 2.2.1) that pairs $\{\pm z_\alpha\}$, where $z_\alpha = \cos \alpha + i \sin \alpha$, form a group G when pairs are multiplied by the rule

$$\{\pm z_\alpha\} \cdot \{\pm z_\beta\} = \{\pm(z_\alpha z_\beta)\}.$$

Show also that the function $\varphi : \mathbb{S}^1 \rightarrow G$ that sends both $z_\alpha, -z_\alpha \in \mathbb{S}^1$ to the pair $\{\pm z_\alpha\}$ is a 2-to-1 homomorphism.

2.2.3 Show that $z \mapsto z^2$ is a well-defined map from G onto \mathbb{S}^1 , where G is the group described in Exercise 2.2.2, and that this map is an isomorphism.

The space that consists of the pairs $\{\pm z_\alpha\}$ of opposite (or “antipodal”) points on the circle is called the *real projective line* \mathbb{RP}^1 . Thus the above exercises

show that the real projective line has a natural group structure, under which it is isomorphic to the circle group \mathbb{S}^1 .

In the next section we will consider the *real projective space* \mathbb{RP}^3 , consisting of the antipodal point pairs $\{\pm q\}$ on the 3-sphere \mathbb{S}^3 . These pairs likewise have a natural product operation, which makes \mathbb{RP}^3 a group—in fact, it is the group $\text{SO}(3)$ of rotations of \mathbb{R}^3 . We will show that \mathbb{RP}^3 is *not* the same group as \mathbb{S}^3 , because $\text{SO}(3)$ is simple and \mathbb{S}^3 is not.

We can see right now that \mathbb{S}^3 is not simple, by finding a nontrivial normal subgroup.

2.2.4 Show that $\{\pm 1\}$ is a normal subgroup of \mathbb{S}^3 .

However, it turns out that $\{\pm 1\}$ is the *only* nontrivial normal subgroup of \mathbb{S}^3 . In particular, the subgroup \mathbb{S}^1 that we found in Section 2.1 is not normal.

2.2.5 Show that \mathbb{S}^1 is not a normal subgroup of \mathbb{S}^3 .

2.3 The groups $\text{SU}(2)$ and $\text{SO}(3)$

The group $\text{SO}(2)$ of rotations of \mathbb{R}^2 about O can be viewed as a geometric object, namely the *unit circle* in the plane, as we observed in Section 1.1.

The unit circle, \mathbb{S}^1 , is the first in the series of *unit n -spheres* \mathbb{S}^n , the n th of which consists of the points at distance 1 from the origin in \mathbb{R}^{n+1} . Thus \mathbb{S}^2 is the ordinary sphere, consisting of the points at distance 1 from the origin in \mathbb{R}^3 . Unfortunately (for those who would like an example of an easily visualized but nontrivial Lie group) there is no rule for multiplying points that makes \mathbb{S}^2 a Lie group. In fact, the only other Lie group among the n -spheres is \mathbb{S}^3 . As we saw in Section 1.3, it becomes a group when its points are viewed as unit quaternions, under the operation of quaternion multiplication.

The group \mathbb{S}^3 of unit quaternions can also be viewed as the group of 2×2 complex matrices of the form

$$Q = \begin{pmatrix} a + di & -b - ci \\ b - ci & a - di \end{pmatrix}, \quad \text{where} \quad \det(Q) = 1,$$

because these are precisely the quaternions of absolute value 1. Such matrices are called *unitary*, and the group \mathbb{S}^3 is also known as the *special unitary group* $\text{SU}(2)$. Unitary matrices are the complex counterpart of orthogonal matrices, and we study the analogy between the two in Chapters 3 and 4.

The group $\text{SU}(2)$ is closely related to the group $\text{SO}(3)$ of rotations of \mathbb{R}^3 . As we saw in Section 1.5, rotations of \mathbb{R}^3 correspond 1-to-1 to

the pairs $\pm t$ of antipodal unit quaternions, the rotation being induced on $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ by the conjugation map $q \mapsto t^{-1}qt$. Also, the group operation of $SO(3)$ corresponds to quaternion multiplication, because if one rotation is induced by conjugation by t_1 , and another by conjugation by t_2 , then conjugation by t_1t_2 induces the product rotation (first rotation followed by the second). Of course, we multiply *pairs* $\pm t$ of quaternions by the rule

$$(\pm t_1)(\pm t_2) = \pm t_1t_2.$$

We therefore identify $SO(3)$ with the group \mathbb{RP}^3 of unit quaternion pairs $\pm t$ under this product operation. The map $\varphi: SU(2) \rightarrow SO(3)$ defined by $\varphi(t) = \{\pm t\}$ is a 2-to-1 homomorphism, because the two elements t and $-t$ of $SU(2)$ go to the single pair $\pm t$ in $SO(3)$. Thus $SO(3)$ looks “simpler” than $SU(2)$ because $SO(3)$ has only one element where $SU(2)$ has two. Indeed, $SO(3)$ is “simpler” because $SU(2)$ is not simple—it has the normal subgroup $\{\pm 1\}$ —and $SO(3)$ is. We now prove this famous property of $SO(3)$ by showing that $SO(3)$ has no nontrivial normal subgroup.

Simplicity of $SO(3)$. *The only nontrivial subgroup of $SO(3)$ closed under conjugation is $SO(3)$ itself.*

Proof. Suppose that H is a nontrivial subgroup of $SO(3)$, so H includes a nontrivial rotation, say the rotation h about axis l through angle α .

Now suppose that H is normal, so H also includes all elements $g^{-1}hg$ for $g \in SO(3)$. If g moves axis l to axis m , then $g^{-1}hg$ is the rotation about axis m through angle α . (In detail, g^{-1} moves m to l , h rotates through angle α about l , then g moves l back to m .) Thus the normal subgroup H includes the rotations through angle α about *all possible axes*.

Now a rotation through α about P , followed by rotation through α about Q , equals rotation through angle θ about R , where R and θ are as shown in Figure 2.3. As in Exercise 1.5.6, we obtain the rotation about P by successive reflections in the great circles PR and PQ , and then the rotation about Q by successive reflections in the great circles PQ and QR . In this sequence of four reflections, the reflections in PQ cancel out, leaving the reflections in PR and QR that define the rotation about R .

As P varies continuously over some interval of the great circle through P and Q , θ varies continuously over some interval. (R may also vary, but this does not matter.) It follows that θ takes some value of the form

$$\frac{m\pi}{n}, \quad \text{where } m \text{ is odd,}$$

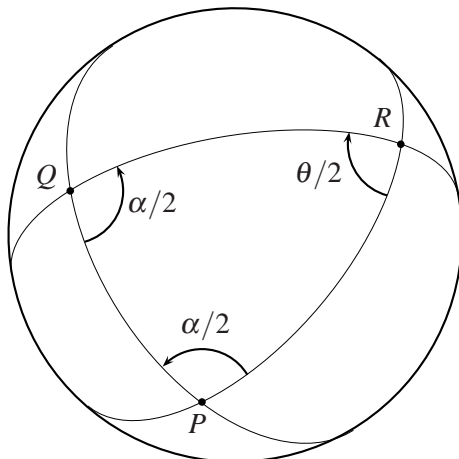


Figure 2.3: Angle of the product rotation.

because such numbers are dense in \mathbb{R} . The n -fold product of this rotation also belongs to H , and it is a rotation about R through $m\pi$, where m is odd. The latter rotation is simply rotation through π , so H includes rotations through π about any point on the sphere (by conjugation with a suitable g again).

Finally, taking the product of rotations with $\alpha/2 = \pi/2$ in Figure 2.3, it is clear that we can get a rotation about R through any angle θ between 0 and 2π . Hence H includes all the rotations in $SO(3)$. \square

Exercises

Like $SO(2)$, $SO(3)$ contains some finite subgroups. It contains all the finite subgroups of $SO(2)$ in an obvious way (as rotations of \mathbb{R}^3 about a fixed axis), but also three more interesting subgroups called the *polyhedral groups*. Each polyhedral group is so called because it consists of the rotations that map a regular polyhedron into itself.

Here we consider the group of 12 rotations that map a *regular tetrahedron* into itself. We consider the tetrahedron whose vertices are alternate vertices of the unit cube in $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$, where the cube has center at O and edges parallel to the \mathbf{i} , \mathbf{j} , and \mathbf{k} axes (Figure 2.4).

First, let us see why there are indeed 12 rotations that map the tetrahedron into itself. To do this, observe that the position of the tetrahedron is completely determined when we know

- Which of the four faces is in the position of the front face in Figure 2.4.

- Which of the three edges of that face is at the bottom of the front face in Figure 2.4.

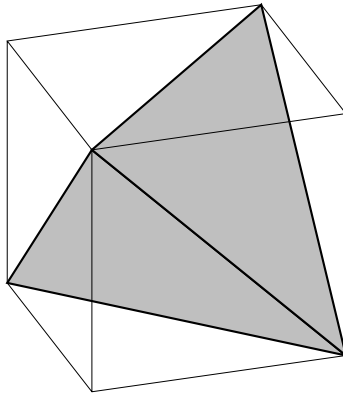


Figure 2.4: The tetrahedron and the cube.

- 2.3.1** Explain why this observation implies 12 possible positions of the tetrahedron, and also explain why all these positions can be obtained by rotations.
- 2.3.2** Similarly, explain why there are 24 rotations that map the cube into itself (so the rotation group of the tetrahedron is different from the rotation group of the cube).

The 12 rotations of the tetrahedron are in fact easy to enumerate with the help of Figure 2.5. As is clear from the figure, the tetrahedron is mapped into itself by two types of rotation:

- A $1/2$ turn about each line through the centers of opposite edges.
- A $1/3$ turn about each line through a vertex and the opposite face center.

- 2.3.3** Show that there are 11 distinct rotations among these two types. What rotation accounts for the 12th position of the tetrahedron?

Now we make use of the quaternion representation of rotations from Section 1.5. Remember that a rotation about axis u through angle θ corresponds to the quaternion pair $\pm q$, where

$$q = \cos \frac{\theta}{2} + u \sin \frac{\theta}{2}.$$

- 2.3.4** Show that the identity, and the three $1/2$ turns, correspond to the four quaternion pairs $\pm 1, \pm \mathbf{i}, \pm \mathbf{j}, \pm \mathbf{k}$.

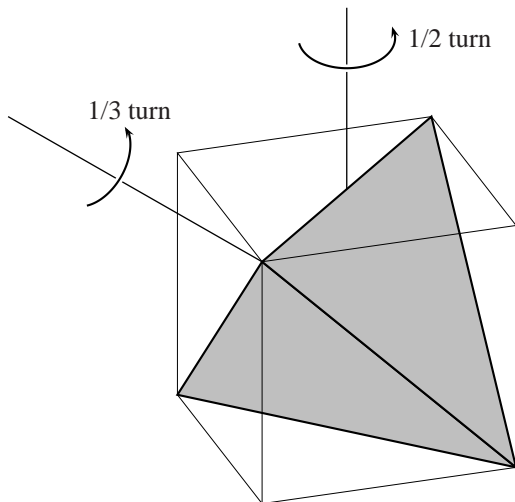


Figure 2.5: The tetrahedron and axes of rotation.

2.3.5 Show that the $1/3$ turns correspond to the eight antipodal pairs among the 16 quaternions

$$\pm \frac{1}{2} \pm \frac{\mathbf{i}}{2} \pm \frac{\mathbf{j}}{2} \pm \frac{\mathbf{k}}{2}.$$

The 24 quaternions obtained in Exercises 2.3.4 and 2.3.5 form an exceptionally symmetric configuration in \mathbb{R}^4 . They are the vertices of a regular figure called the 24-cell, copies of which form a “tiling” of \mathbb{R}^4 .

2.4 Isometries of \mathbb{R}^n and reflections

In this section we take up an idea that appeared briefly in the exercises for Section 1.5: the representation of isometries as products of reflections. There we showed that certain isometries of \mathbb{R}^2 and \mathbb{R}^3 are products of reflections. Here we represent isometries of \mathbb{R}^n as products of reflections, and in the next section we use this result to describe the rotations of \mathbb{R}^4 .

We actually prove that *any isometry of \mathbb{R}^n that fixes O is the product of reflections in hyperplanes through O* , and then specialize to *orientation-preserving isometries*. A *hyperplane H through O* is an $(n-1)$ -dimensional subspace of \mathbb{R}^n , and *reflection in H* is the linear map of \mathbb{R}^n that fixes the elements in H and reverses the vectors orthogonal to H .

Reflection representation of isometries. Any isometry of \mathbb{R}^n that fixes O is the product of at most n reflections in hyperplanes through O .

Proof. We argue by induction on n . For $n = 1$ the result is the obvious one that the only isometries of \mathbb{R} fixing O are the identity and the map $x \mapsto -x$, which is reflection in O .

Now suppose that the result is true for $n = k - 1$ and that f is an isometry of \mathbb{R}^k fixing O . If f is not the identity, suppose that $v \in \mathbb{R}^k$ is such that $f(v) = w \neq v$. Then the reflection r_u in the hyperplane orthogonal to $u = v - w$ maps the subspace $\mathbb{R}u$ of real multiples of u onto itself and the map $r_u f$ (“ f followed by r_u ”) is the identity on the subspace $\mathbb{R}u$ of \mathbb{R}^k .

The restriction of $r_u f$ to the \mathbb{R}^{k-1} orthogonal to $\mathbb{R}u$ is, by induction, the product of $\leq k - 1$ reflections. It follows that $f = r_u g$, where g is the product of $\leq k - 1$ reflections.

Therefore, f is the product of $\leq k$ reflections, and the result is true for all n by induction. \square

It follows in particular that any orientation-preserving isometry of \mathbb{R}^3 is the product of 0 or 2 reflections (because the product of an odd number of reflections reverses orientation). Thus any such isometry is a rotation about an axis passing through O .

This theorem is sometimes known as the *Cartan–Dieudonné theorem*, after a more general theorem proved by Cartan [1938], and generalized further by Dieudonné. Cartan’s theorem concerns “reflections” in spaces with real or complex coordinates, and Dieudonné’s extends it to spaces with coordinates from finite fields.

Exercises

Assuming that reflections are linear, the representation of isometries as products of reflections shows that all isometries fixing the origin are linear maps. In fact, there is nice direct proof that all such isometries (including reflections) are linear, pointed out to me by Marc Ryser. We suppose that f is an isometry that fixes O , and that u and v are any points in \mathbb{R}^n .

2.4.1 Prove that f preserves straight lines and midpoints of line segments.

2.4.2 Using the fact that $u + v$ is the midpoint of the line joining $2u$ and $2v$, and Exercise 2.4.1, show that $f(u + v) = f(u) + f(v)$.

2.4.3 Also prove that $f(ru) = rf(u)$ for any real number r .

It is also true that reflections have determinant -1 , hence the determinant detects the “reversal of orientation” effected by a reflection.

2.4.4 Show that reflection in the hyperplane orthogonal to a coordinate axis has determinant -1 , and generalize this result to any reflection.

2.5 Rotations of \mathbb{R}^4 and pairs of quaternions

A linear map is called *orientation-preserving* if its determinant is positive, and *orientation-reversing* otherwise. Reflections are linear and orientation-reversing, so a product of reflections is orientation-preserving if and only if it contains an even number of terms. We define a *rotation of \mathbb{R}^n about O* to be an orientation-preserving isometry that fixes O .

Thus it follows from the Cartan–Dieudonné theorem that any rotation of \mathbb{R}^4 is the product of 0, 2, or 4 reflections. The exact number is not important here—what we really want is a way to represent reflections by quaternions, as a stepping-stone to the representation of rotations by quaternions. Not surprisingly, each reflection is specified by the quaternion orthogonal to the hyperplane of reflection. More surprisingly, a rotation is specified by just *two* quaternions, regardless of the number of reflections needed to compose it. Our proof follows Conway and Smith [2003], p. 41.

Quaternion representation of reflections. *Reflection of $\mathbb{H} = \mathbb{R}^4$ in the hyperplane through O orthogonal to the unit quaternion u is the map that sends each $q \in \mathbb{H}$ to $-u\bar{q}u$.*

Proof. First observe that the map $q \mapsto -u\bar{q}u$ is an isometry. This is because

- $q \mapsto -\bar{q}$ reverses the real part of q and keeps the imaginary part fixed, hence it is reflection in the hyperplane spanned by \mathbf{i} , \mathbf{j} , and \mathbf{k} .
- Multiplication on the left by the unit quaternion u is an isometry by the argument in Section 1.4, and there is a similar argument for multiplication on the right.

Next notice that the map $q \mapsto -u\bar{q}u$ sends

$$\begin{aligned}vu \text{ to } -u\overline{(vu)}u &= -u\bar{u}\bar{v}u \quad \text{because } \overline{(vu)} = \bar{u}\bar{v}, \\ &= -\bar{v}u \quad \text{because } u\bar{u} = |u|^2 = 1.\end{aligned}$$

In particular, the map sends u to $-u$, so vectors parallel to u are reversed. And it sends $\mathbf{i}u$ to $\mathbf{i}u$, because $\bar{\mathbf{i}} = -\mathbf{i}$, and similarly $\mathbf{j}u$ to $\mathbf{j}u$ and $\mathbf{k}u$ to $\mathbf{k}u$. Thus the vectors $\mathbf{i}u$, $\mathbf{j}u$, and $\mathbf{k}u$, which span the hyperplane orthogonal to u , are fixed. Hence the map $q \mapsto -u\bar{q}u$ is reflection in this hyperplane. \square

Quaternion representation of rotations. Any rotation of $\mathbb{H} = \mathbb{R}^4$ about O is a map of the form $q \mapsto vqw$, where v and w are unit quaternions.

Proof. It follows from the quaternion representation of reflections that the result of successive reflections in the hyperplanes orthogonal to the unit quaternions u_1, u_2, \dots, u_{2n} is the map

$$q \mapsto u_{2n} \cdots \overline{u_3} u_2 \overline{u_1} q \overline{u_1} u_2 \overline{u_3} \cdots u_{2n},$$

because an even number of sign changes and conjugations makes no change. The pre- and postmultipliers are in general two different unit quaternions, $u_{2n} \cdots \overline{u_3} u_2 \overline{u_1} = v$ and $\overline{u_1} u_2 \overline{u_3} \cdots u_{2n} = w$, say, so the general rotation of \mathbb{R}^4 is a map of the form

$$q \mapsto vqw, \quad \text{where } v \text{ and } w \text{ are unit quaternions.}$$

Conversely, any map of this form is a rotation, because *multiplication of $\mathbb{H} = \mathbb{R}^4$ on either side by a unit quaternion is an orientation-preserving isometry*. We already know that multiplication by a unit quaternion is an isometry, by Section 1.4. And it preserves orientation by the following argument.

Multiplication of $\mathbb{H} = \mathbb{R}^4$ by a unit quaternion

$$v = \begin{pmatrix} a + id & -b - ic \\ b - ic & a - id \end{pmatrix}, \quad \text{where } a^2 + b^2 + c^2 + d^2 = 1,$$

is a linear transformation of \mathbb{R}^4 with matrix

$$R_v = \left(\begin{array}{cc|cc} a & -d & -b & c \\ d & a & -c & -b \\ \hline b & c & a & d \\ -c & b & -d & a \end{array} \right),$$

where the 2×2 submatrices represent the complex-number entries in v . It can be checked that $\det(R_v) = 1$. So multiplication by v , on either side, preserves orientation. \square

Exercises

The following exercises study the rotation $q \mapsto \mathbf{i}q$ of $\mathbb{H} = \mathbb{R}^4$, first expressing it as a product of “plane rotations”—of the planes spanned by $1, \mathbf{i}$ and \mathbf{j}, \mathbf{k} respectively—then breaking it down to a product of four reflections.

- 2.5.1** Check that $q \mapsto \mathbf{i}q$ sends 1 to \mathbf{i} , \mathbf{i} to -1 and \mathbf{j} to \mathbf{k} , \mathbf{k} to $-\mathbf{j}$. How many points of \mathbb{R}^4 are fixed by this map?
- 2.5.2** Show that the rotation that sends 1 to \mathbf{i} , \mathbf{i} to -1 and leaves \mathbf{j} , \mathbf{k} fixed is the product of reflections in the hyperplanes orthogonal to $u_1 = \mathbf{i}$ and $u_2 = (\mathbf{i} - 1)/\sqrt{2}$.
- 2.5.3** Show that the rotation that sends \mathbf{j} to \mathbf{k} , \mathbf{k} to $-\mathbf{j}$ and leaves 1 , \mathbf{i} fixed is the product of reflections in the hyperplanes orthogonal to $u_3 = \mathbf{k}$ and $u_4 = (\mathbf{k} - \mathbf{j})/\sqrt{2}$.

It follows, by the formula $q \mapsto -u\bar{q}u$ for reflection, that the product of rotations in Exercises 2.5.2 and 2.5.3 is the product

$$q \mapsto u_4\bar{u}_3u_2\bar{u}_1 q \bar{u}_1u_2\bar{u}_3u_4$$

of reflections in the hyperplanes orthogonal to u_1, u_2, u_3, u_4 respectively.

- 2.5.4** Check that $u_4\bar{u}_3u_2\bar{u}_1 = \mathbf{i}$ and $\bar{u}_1u_2\bar{u}_3u_4 = 1$, so the product of the four reflections is indeed $q \mapsto \mathbf{i}q$.

2.6 Direct products of groups

Before we analyze rotations of \mathbb{R}^4 from the viewpoint of group theory, it is desirable to review the concept of *direct product* or *Cartesian product* of groups.

Definition. If A and B are groups then their *direct product* $A \times B$ is the set of ordered pairs (a, b) , where $a \in A$ and $b \in B$, under the “product of pairs” operation defined by

$$(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2).$$

It is easy to check that this product operation is associative, that the identity element of $A \times B$ is the pair $(1_A, 1_B)$, where 1_A is the identity of A and 1_B is the identity of B , and that (a, b) has inverse (a^{-1}, b^{-1}) . Thus $A \times B$ is indeed a group.

Many important groups are nontrivial direct products; that is, they have the form $A \times B$ where neither A nor B is the trivial group $\{1\}$. For example:

- The group \mathbb{R}^2 , under vector addition, is the direct product $\mathbb{R} \times \mathbb{R}$. More generally, \mathbb{R}^n is the n -fold direct product $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$.

- If A and B are groups of $n \times n$ matrices, then the matrices of the form

$$\begin{pmatrix} a & \mathbf{0} \\ \mathbf{0} & b \end{pmatrix}, \quad \text{where } a \in A \quad \text{and} \quad b \in B,$$

make up a group isomorphic to $A \times B$ under matrix multiplication, where $\mathbf{0}$ is the $n \times n$ zero matrix. This is because of the so-called *block multiplication* of matrices, according to which

$$\begin{pmatrix} a_1 & \mathbf{0} \\ \mathbf{0} & b_1 \end{pmatrix} \begin{pmatrix} a_2 & \mathbf{0} \\ \mathbf{0} & b_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 & \mathbf{0} \\ \mathbf{0} & b_1 b_2 \end{pmatrix}.$$

- It follows, from the previous item, that \mathbb{R}^n is isomorphic to a $2n \times 2n$ matrix group, because \mathbb{R} is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \quad \text{where } x \in \mathbb{R}.$$

- The group $\mathbb{S}^1 \times \mathbb{S}^1$ is a group called the (two-dimensional) *torus* \mathbb{T}^2 . More generally, the n -fold direct product of \mathbb{S}^1 factors is called the *n -dimensional torus* \mathbb{T}^n .

We call $\mathbb{S}^1 \times \mathbb{S}^1$ a torus because its elements (θ, ϕ) , where $\theta, \phi \in \mathbb{S}^1$, can be viewed as the points on the torus surface (Figure 2.6).

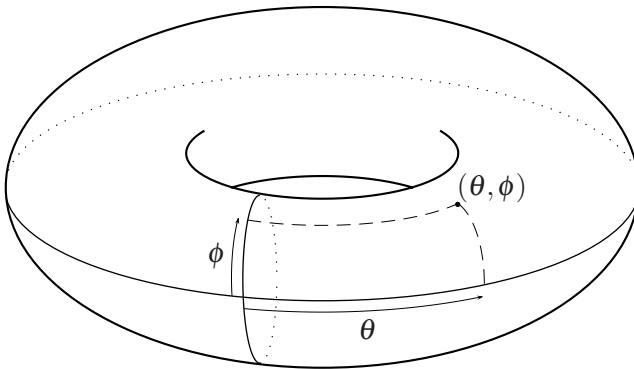


Figure 2.6: The torus $\mathbb{S}^1 \times \mathbb{S}^1$.

Since the groups \mathbb{R} and \mathbb{S}^1 are abelian, the same is true of all their direct products $\mathbb{R}^m \times \mathbb{T}^n$. It can be shown that the latter groups include *all* the connected abelian matrix Lie groups.

Exercises

If we let x_1, x_2, x_3, x_4 be the coordinates along mutually orthogonal axes in \mathbb{R}^4 , then it is possible to “rotate” the x_1 and x_2 axes while keeping the x_3 and x_4 axes fixed.

2.6.1 Write a 4×4 matrix for the transformation that rotates the (x_1, x_2) -plane through angle θ while keeping the x_3 - and x_4 -axes fixed.

2.6.2 Write a 4×4 matrix for the transformation that rotates the (x_3, x_4) -plane through angle ϕ while keeping the x_1 - and x_2 -axes fixed.

2.6.3 Observe that the rotations in Exercise 2.6.1 form an \mathbb{S}^1 , as do the rotations in Exercise 2.6.2, and deduce that $\text{SO}(4)$ contains a subgroup isomorphic to \mathbb{T}^2 .

The groups of the form $\mathbb{R}^m \times \mathbb{T}^n$ may be called “generalized cylinders,” based on the simplest example $\mathbb{R} \times \mathbb{S}^1$.

2.6.4 Why is it appropriate to call the group $\mathbb{R} \times \mathbb{S}^1$ a cylinder?

The notation \mathbb{S}^n is unfortunately *not* compatible with the direct product notation (at least not the way the notation \mathbb{R}^n is).

2.6.5 Explain why $\mathbb{S}^3 = \text{SU}(2)$ is not the same group as $\mathbb{S}^1 \times \mathbb{S}^1 \times \mathbb{S}^1$.

2.7 The map from $\text{SU}(2) \times \text{SU}(2)$ to $\text{SO}(4)$

In Section 2.5 we showed that the rotations of \mathbb{R}^4 are precisely the maps $q \mapsto vqw$, where v and w run through all the unit quaternions. Since v^{-1} is a unit quaternion if and only if v is, it is equally valid to represent each rotation of \mathbb{R}^4 by a map of the form $q \mapsto v^{-1}qw$, where v and w are unit quaternions. The latter representation is more convenient for what comes next.

The pairs of unit quaternions (v, w) form a group under the operation defined by

$$(v_1, w_1) \cdot (v_2, w_2) = (v_1 v_2, w_1 w_2),$$

where the products $v_1 v_2$ and $w_1 w_2$ on the right side are ordinary quaternion products. Since the v come from the group $\text{SU}(2)$ of unit quaternions, and the w likewise, the group of pairs (v, w) is the direct product $\text{SU}(2) \times \text{SU}(2)$ of $\text{SU}(2)$ with itself.

The map that sends each pair $(v, w) \in \text{SU}(2) \times \text{SU}(2)$ to the rotation $q \mapsto v^{-1}qw$ in $\text{SO}(4)$ is a *homomorphism* $\varphi : \text{SU}(2) \times \text{SU}(2) \rightarrow \text{SO}(4)$. This is because

- the product of the map $q \mapsto v_1^{-1}qw_1$ corresponding to (v_1, w_1)
- with the map $q \mapsto v_2^{-1}qw_2$ corresponding to (v_2, w_2)
- is the map $q \mapsto v_2^{-1}v_1^{-1}qw_1w_2$,
- which is the map $q \mapsto (v_1v_2)^{-1}q(w_1w_2)$ corresponding to the product (v_1v_2, w_1w_2) of (v_1, w_1) and (v_2, w_2) .

This homomorphism is onto $SO(4)$, because each rotation of \mathbb{R}^4 can be expressed in the form $q \mapsto v^{-1}qw$, but one might expect it to be very many-to-one, since many pairs (v, w) of unit quaternions conceivably give the same rotation. Surprisingly, this is not so. The representation of rotations by pairs is “unique up to sign” in the following sense: *if (v, w) gives a certain rotation, the only other pair that gives the same rotation is $(-v, -w)$.*

To prove this, it suffices to prove that the kernel of the homomorphism $\varphi : SU(2) \times SU(2) \rightarrow SO(4)$ has two elements.

Size of the kernel. *The homomorphism $\varphi : SU(2) \times SU(2) \rightarrow SO(4)$ is 2-to-1, because its kernel has two elements.*

Proof. Suppose that (v, w) is in the kernel, so $q \mapsto v^{-1}qw$ is the identity rotation. In particular, this rotation fixes 1, so

$$v^{-1}1w = 1; \quad \text{hence} \quad v = w.$$

Thus the map is in fact $q \mapsto v^{-1}qv$, which we know (from Section 1.5) fixes the real axis and rotates the space of pure imaginary quaternions. Only if $v = 1$ or $v = -1$ does the map fix everything; hence the kernel of φ has only two elements, $(1, 1)$ and $(-1, -1)$.

The left cosets of the kernel are therefore the 2-element sets

$$(v, w)(\pm 1, \pm 1) = (\pm v, \pm w),$$

and each coset corresponds to a distinct rotation of \mathbb{R}^4 , by the fundamental homomorphism theorem of Section 2.2. \square

This theorem shows that $SO(4)$ is “almost” the same as $SU(2) \times SU(2)$, and the latter is far from being a simple group. For example, the subgroup of pairs $(v, 1)$ is a nontrivial normal subgroup, but clearly not the whole of $SU(2) \times SU(2)$. This gives us a way to show that $SO(4)$ is not simple.

SO(4) is not simple. *There is a nontrivial normal subgroup of SO(4), not equal to SO(4).*

Proof. The subgroup of pairs $(v, 1) \in \text{SU}(2) \times \text{SU}(2)$ is normal; in fact, it is the kernel of the map $(v, w) \mapsto (1, w)$, which is clearly a homomorphism.

The corresponding subgroup of SO(4) consists of maps of the form $q \mapsto v^{-1}q1$, which likewise form a normal subgroup of SO(4). But this subgroup is not the whole of SO(4). For example, it does not include the map $q \mapsto qw$ for any $w \neq \pm 1$, by the “unique up to sign” representation of rotations by pairs (v, w) . \square

Exercises

An interesting subgroup $\text{Aut}(\mathbb{H})$ of SO(4) consists of the continuous *automorphisms* of $\mathbb{H} = \mathbb{R}^4$. These are the continuous bijections $\rho : \mathbb{H} \rightarrow \mathbb{H}$ that preserve the quaternion sum and product, that is,

$$\rho(p+q) = \rho(p) + \rho(q), \quad \rho(pq) = \rho(p)\rho(q) \quad \text{for any } p, q \in \mathbb{H}.$$

It is easy to check that, for each unit quaternion u , the ρ that sends $q \mapsto u^{-1}qu$ is an automorphism (first exercise), so it follows from Section 1.5 that $\text{Aut}(\mathbb{H})$ includes the SO(3) of rotations of the 3-dimensional subspace $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ of pure imaginary quaternions. The purpose of this set of exercises is to show that *all* continuous automorphisms of \mathbb{H} are of this form, so $\text{Aut}(\mathbb{H}) = \text{SO}(3)$.

2.7.1 Check that $q \mapsto u^{-1}qu$ is an automorphism of \mathbb{H} for any unit quaternion u .

Now suppose that ρ is *any* automorphism of \mathbb{H} .

2.7.2 Use the preservation of sums by an automorphism ρ to deduce in turn that

- ρ preserves 0, that is, $\rho(0) = 0$,
- ρ preserves differences, that is, $\rho(p - q) = \rho(p) - \rho(q)$.

2.7.3 Use preservation of products to deduce that

- ρ preserves 1, that is, $\rho(1) = 1$,
- ρ preserves quotients, that is, $\rho(p/q) = \rho(p)/\rho(q)$ for $q \neq 0$.

2.7.4 Deduce from Exercises 2.7.2 and 2.7.3 that $\rho(m/n) = m/n$ for any integers m and $n \neq 0$. This implies $\rho(r) = r$ for any real r , and hence that ρ is a *linear* map of \mathbb{R}^4 . Why?

Thus we now know that a continuous automorphism ρ is a linear bijection of \mathbb{R}^4 that preserves the real axis, and hence ρ maps $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ onto itself. It remains to show that the restriction of ρ to $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ is a rotation, that is, an orientation-preserving isometry, because we know from Section 1.5 that rotations of $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$ are of the form $q \mapsto u^{-1}qu$.

2.7.5 Prove in turn that

- ρ preserves conjugates, that is, $\rho(\bar{q}) = \overline{\rho(q)}$,
- ρ preserves distance,
- ρ preserves inner product in $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$,
- $\rho(p \times q) = \rho(p) \times \rho(q)$ in $\mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$, and hence ρ preserves orientation.

The appearance of $\text{SO}(3)$ as the automorphism group of the quaternion algebra \mathbb{H} suggests that the automorphism group of the octonion algebra \mathbb{O} might also be of interest. It turns out to be a 14-dimensional group called G_2 —the first of the exceptional Lie groups mentioned (along with \mathbb{O}) in Section 1.6. This link between \mathbb{O} and the exceptional groups was pointed out by Cartan [1908].

2.8 Discussion

The concept of simple group emerged around 1830 from Galois’s theory of equations. Galois showed that each polynomial equation has a finite group of “symmetries” (permutations of its roots that leave its coefficients invariant), and that the equation is solvable only if its group decomposes in a certain way. In particular, the general quintic equation is *not* solvable because its group contains the nonabelian simple group A_5 —the group of even permutations of five objects. The same applies to the general equation of any degree greater than 5, because A_n , the group of even permutations of n objects, is simple for any $n \geq 5$.

With this discovery, Galois effectively closed the classical theory of equations, but he opened the (much larger) theory of groups. Specifically, by exhibiting the nontrivial infinite family A_n for $n \geq 5$, he raised the problem of finding and classifying all finite simple groups. This problem is much deeper than anyone could have imagined in the time of Galois, because it depends on solving the corresponding problem for *continuous* groups, or Lie groups as we now call them.

Around 1870, Sophus Lie was inspired by Galois theory to develop an analogous theory of differential equations and their “symmetries,” which generally form continuous groups. As with polynomial equations, simple groups raise an obstacle to solvability. However, at that time it was not clear what the generalization of the group concept from finite to continuous should be. Lie understood continuous groups to be groups generated by “infinitesimal” elements, so he thought that the rotation group of \mathbb{R}^3 should

include “infinitesimal rotations.” Today, we separate out the “infinitesimal rotations” of \mathbb{R}^3 in a structure called $\mathfrak{so}(3)$, the *Lie algebra of $\text{SO}(3)$* . The concept of simplicity also makes sense for $\mathfrak{so}(3)$, and is somewhat easier to establish. Indeed, the infinitesimal elements of any continuous group G form a structure \mathfrak{g} now called the *Lie algebra of G* , which captures most of the structure of G but is easier to handle. We discuss “infinitesimal elements,” and their modern counterparts, further in Section 4.3.

It was a stroke of luck (or genius) that Lie decided to look at infinitesimal elements, because it enabled him to prove simplicity for whole infinite families of Lie algebras in one fell swoop. (As we will see later, most of the corresponding continuous groups are not *quite* simple, and one has to tease out certain small subgroups and quotient by them.) Around 1885 Lie proved results so general that they cover all but a finite number of simple Lie algebras—namely, those of the exceptional groups mentioned at the end of Chapter 1 (see Hawkins [2000], pp. 92–98).

In the avalanche of Lie’s results, the special case of $\mathfrak{so}(3)$ and $\text{SO}(3)$ seems to have gone unnoticed. It gradually came to light as twentieth-century books on Lie theory started to work out special cases of geometric interest by way of illustration. In the 1920s, quantum physics also directed attention to $\text{SO}(3)$, since rotations in three dimensions are physically significant. Still, it is remarkable that a purely geometric argument for the simplicity of $\text{SO}(3)$ took so long to emerge. Perhaps its belated appearance is due to its *topological* content, namely, the step that depends purely on continuity. The argument hinges on the fact that θ is a continuous function of distance along the great circle PQ , and that such a function takes every value between its extreme values: the so-called *intermediate value theorem*.

The theory of continuity (topology) came after the theory of continuous groups—not surprisingly, since one does not bother to develop a theory of continuity before seeing that it has some content—and applications of topology to group theory were rare before the 1920s. In this book we will present further isolated examples of continuity arguments in Sections 3.2, 3.8, and 7.5 before taking up topology systematically in Chapter 8.

Another book with a strongly geometric treatment of $\text{SO}(3)$ is Berger [1987]. Volume I of Berger, p. 169, has a simplicity proof for $\text{SO}(3)$ similar to the one given here, and it is extended to a simplicity result about $\text{SO}(n)$, for $n \geq 5$, on p. 170: $\text{SO}(2m+1)$ is simple and the only nontrivial normal subgroup of $\text{SO}(2m)$ is $\{\pm 1\}$. We arrive at the same result by a different

route in Section 7.5. (Our route is longer, but it also takes in the complex and quaternion analogues of $SO(n)$.) Berger treats $SO(4)$ with the help of quaternions on p. 190 of his Volume II, much as we have done here. The quaternion representation of rotations of \mathbb{R}^4 was another of Cayley's discoveries, made in 1855.

Lie observed the anomalous structure of $SO(4)$ at the infinitesimal level. He mentions it, in scarcely recognizable form, on p. 683 of Volume III of his 1893 book *Theorie der Transformationsgruppen*. The anomaly of $SO(4)$ is hidden in some modern treatments of Lie theory, where the concept of simplicity is superseded by the more general concept of *semisimplicity*. All simple groups are semisimple, and $SO(4)$ is semisimple, so an anomaly is removed by relaxing the concept of "simple" to "semisimple." However, the concept of semisimplicity makes little sense before one has absorbed the concept of simplicity, and our goal in this book is to understand the simple groups, notwithstanding the anomaly of $SO(4)$.