1

Introduction

The goal of expressing geometrical relationships through algebraic equations has dominated much of the development of mathematics. This line of thinking goes back to the ancient Greeks, who constructed a set of geometric laws to describe the world as they saw it. Their view of geometry was largely unchallenged until the eighteenth century, when mathematicians discovered new geometries with different properties from the Greeks' *Euclidean* geometry. Each of these new geometries had distinct algebraic properties, and a major preoccupation of nineteenth century mathematicians was to place these geometries within a unified algebraic framework. One of the key insights in this process was made by W.K. Clifford, and this book is concerned with the implications of his discovery.

Before we describe Clifford's discovery (in chapter 2) we have gathered together some introductory material of use throughout this book. This chapter revises basic notions of vector spaces, emphasising pictorial representations of the underlying algebraic rules — a theme which dominates this book. The material is presented in a way which sets the scene for the introduction of Clifford's product, in part by reflecting the state of play when Clifford conducted his research. To this end, much of this chapter is devoted to studying the various products that can be defined between vectors. These include the scalar and vector products familiar from three-dimensional geometry, and the complex and quaternion products. We also introduce the *outer* or *exterior* product, though this is covered in greater depth in later chapters. The material in this chapter is intended to be fairly basic, and those impatient to uncover Clifford's insight may want to jump straight to chapter 2. Readers unfamiliar with the outer product are encouraged to read this chapter, however, as it is crucial to understanding Clifford's discovery.

1.1 Vector (linear) spaces

At the heart of much of geometric algebra lies the idea of vector, or linear spaces. Some properties of these are summarised here and assumed throughout this book. In this section we talk in terms of *vector* spaces, as this is the more common term. For all other occurrences, however, we prefer to use the term *linear* space. This is because the term '*vector*' has a very specific meaning within geometric algebra (as the grade-1 elements of the algebra).

1.1.1 Properties

Vector spaces are defined in terms of two objects. These are the vectors, which can often be visualised as directions in space, and the scalars, which are usually taken to be the real numbers. The vectors have a simple addition operation rule with the following obvious properties:

(i) Addition is *commutative*:

$$a + b = b + a. \tag{1.1}$$

(ii) Addition is *associative*:

$$a + (b + c) = (a + b) + c. \tag{1.2}$$

This property enables us to write expressions such as $a + b + c$ without ambiguity.

(iii) There is an identity element, denoted 0:

$$a + 0 = a. \tag{1.3}$$

(iv) Every element a has an inverse $-a$:

$$a + (-a) = 0. \tag{1.4}$$

For the case of directed line segments each of these properties has a clear geometric equivalent. These are illustrated in figure 1.1.

Vector spaces also contain a multiplication operation between the scalars and the vectors. This has the property that for any scalar λ and vector a, the product λa is also a member of the vector space. Geometrically, this corresponds to the dilation operation. The following further properties also hold for any scalars λ, μ and vectors a and b:

(i) $\lambda(a + b) = \lambda a + \lambda b$;
(ii) $(\lambda + \mu)a = \lambda a + \mu a$;
(iii) $(\lambda\mu)a = \lambda(\mu a)$;
(iv) if $1\lambda = \lambda$ for all scalars λ then $1a = a$ for all vectors a.

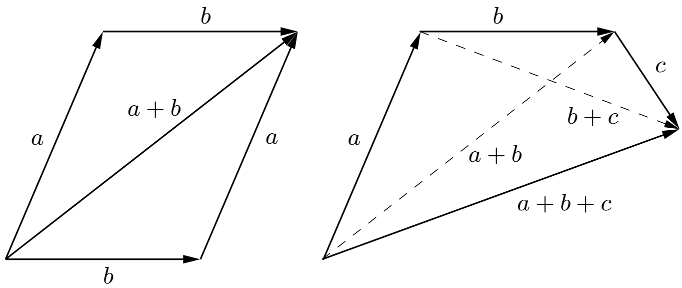

Figure 1.1 *A geometric picture of vector addition.* The result of $a+b$ is formed by adding the tail of b to the head of a. As is shown, the resultant vector $a+b$ is the same as $b+a$. This finds an algebraic expression in the statement that addition is commutative. In the right-hand diagram the vector $a+b+c$ is constructed two different ways, as $a+(b+c)$ and as $(a+b)+c$. The fact that the results are the same is a geometric expression of the associativity of vector addition.

The preceding set of rules serves to define a vector space completely. Note that the + operation connecting scalars is different from the + operation connecting the vectors. There is no ambiguity, however, in using the same symbol for both.

The following two definitions will be useful later in this book:

(i) Two vector spaces are said to be *isomorphic* if their elements can be placed in a one-to-one correspondence which preserves sums, and there is a one-to-one correspondence between the scalars which preserves sums and products.

(ii) If $\mathcal{U}$ and $\mathcal{V}$ are two vector spaces (sharing the same scalars) and all the elements of $\mathcal{U}$ are contained in $\mathcal{V}$, then $\mathcal{U}$ is said to form a *subspace* of $\mathcal{V}$.

1.1.2 Bases and dimension

The concept of dimension is intuitive for simple vector spaces — lines are one-dimensional, planes are two-dimensional, and so on. Equipped with the axioms of a vector space we can proceed to a formal definition of the dimension of a vector space. First we need to define some terms.

(i) A vector b is said to be a *linear combination* of the vectors $a_1, \dots, a_n$ if scalars $\lambda_1, \dots, \lambda_n$ can be found such that

$$b = \lambda_1 a_1 + \cdots + \lambda_n a_n = \sum_{i=1}^{n} \lambda_i a_i. \tag{1.5}$$

(ii) A set of vectors $\{a_1, \dots, a_n\}$ is said to be *linearly dependent* if scalars

$\lambda_1, \ldots, \lambda_n$ (not all zero) can be found such that

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0. \tag{1.6}$$

If such a set of scalars cannot be found, the vectors are said to be *linearly independent*.

(iii) A set of vectors $\{a_1, \ldots, a_n\}$ is said to *span* a vector space $\mathcal{V}$ if every element of $\mathcal{V}$ can be expressed as a linear combination of the set.

(iv) A set of vectors which are both linearly independent and span the space $\mathcal{V}$ are said to form a *basis* for $\mathcal{V}$.

These definitions all carry an obvious, intuitive picture if one thinks of vectors in a plane or in three-dimensional space. For example, it is clear that two independent vectors in a plane provide a basis for all vectors in that plane, whereas any three vectors in the plane are linearly dependent. These axioms and definitions are sufficient to prove the *basis theorem*, which states that *all bases of a vector space have the same number of elements.* This number is called the *dimension* of the space. Proofs of this statement can be found in any textbook on linear algebra, and a sample proof is left to work through as an exercise. Note that any two vector spaces of the same dimension and over the same field are isomorphic.

The axioms for a vector space define an abstract mathematical entity which is already well equipped for studying problems in geometry. In so doing we are not compelled to interpret the elements of the vector space as displacements. Often different interpretations can be attached to isomorphic spaces, leading to different types of geometry (affine, projective, finite, *etc.*). For most problems in physics, however, we need to be able to do more than just add the elements of a vector space; we need to multiply them in various ways as well. This is necessary to formalise concepts such as angles and lengths and to construct higher-dimensional surfaces from simple vectors.

Constructing suitable products was a major concern of nineteenth century mathematicians, and the concepts they introduced are integral to modern mathematical physics. In the following sections we study some of the basic concepts that were successfully formulated in this period. The culmination of this work, Clifford's *geometric product*, is introduced separately in chapter 2. At various points in this book we will see how the products defined in this section can all be viewed as special cases of Clifford's geometric product.

1.2 The scalar product

Euclidean geometry deals with concepts such as lines, circles and perpendicularity. In order to arrive at Euclidean geometry we need to add two new concepts to our vector space. These are distances between points, which allow us to define a circle, and angles between vectors so that we can say that two lines are perpendicular. The introduction of a scalar product achieves both of these goals.

Given any two vectors a, b, the scalar product $a \cdot b$ is a rule for obtaining a number with the following properties:

(i) $a \cdot b = b \cdot a$;
(ii) $a \cdot (\lambda b) = \lambda(a \cdot b)$;
(iii) $a \cdot (b + c) = a \cdot b + a \cdot c$;
(iv) $a \cdot a > 0$, unless $a = 0$.

(When we study relativity, this final property will be relaxed.) The introduction of a scalar product allows us to define the length of a vector, $|a|$, by

$$|a| = \sqrt{}(a \cdot a). \tag{1.7}$$

Here, and throughout this book, the positive square root is always implied by the $\sqrt{}$ symbol. The fact that we now have a definition of lengths and distances means that we have specified a *metric space*. Many different types of metric space can be constructed, of which the simplest are the *Euclidean* spaces we have just defined.

The fact that for Euclidean space the inner product is positive-definite means that we have a Schwarz inequality of the form

$$|a \cdot b| \leq |a|\,|b|. \tag{1.8}$$

The proof is straightforward:

$$\begin{aligned} (a + \lambda b) \cdot (a + \lambda b) \geq 0 \qquad \forall \lambda \\ \Rightarrow\ a \cdot a + 2\lambda a \cdot b + \lambda^2 b \cdot b \geq 0 \qquad \forall \lambda \\ \Rightarrow (a \cdot b)^2 \leq a \cdot a\, b \cdot b, \end{aligned} \tag{1.9}$$

where the last step follows by taking the discriminant of the quadratic in λ. Since all of the numbers in this inequality are positive we recover (1.8). We can now define the *angle* θ between a and b by

$$a \cdot b = |a||b| \cos(\theta). \tag{1.10}$$

Two vectors whose scalar product is zero are said to be *orthogonal*. It is usually convenient to work with bases in which all of the vectors are mutually orthogonal. If all of the basis vectors are further normalised to have unit length, they are said to form an *orthonormal* basis. If the set of vectors $\{\mathsf{e}_1, \ldots, \mathsf{e}_n\}$ denote such a basis, the statement that the basis is orthonormal can be summarised as

$$\mathsf{e}_i \cdot \mathsf{e}_j = \delta_{ij}. \tag{1.11}$$

Here the δ_{ij} is the Kronecker delta function, defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{1.12}$$

We can expand any vector a in this basis as

$$a = \sum_{i=1}^{n} a_i \mathsf{e}_i = a_i \mathsf{e}_i, \tag{1.13}$$

where we have started to employ the *Einstein summation convention* that pairs of indices in any expression are summed over. This convention will be assumed throughout this book. The $\{a_i\}$ are the *components* of the vector a in the $\{\mathsf{e}_i\}$ basis. These are found simply by

$$a_i = \mathsf{e}_i \cdot a. \tag{1.14}$$

The scalar product of two vectors $a = a_i \mathsf{e}_i$ and $b = b_i \mathsf{e}_i$ can now written simply as

$$a \cdot b = (a_i \mathsf{e}_i) \cdot (b_j \mathsf{e}_j) = a_i b_j \, \mathsf{e}_i \cdot \mathsf{e}_j = a_i b_j \delta_{ij} = a_i b_i. \tag{1.15}$$

In spaces where the inner product is not positive-definite, such as Minkowski spacetime, there is no equivalent version of the Schwarz inequality. In such cases it is often only possible to define an 'angle' between vectors by replacing the cosine function with a cosh function. In these cases we can still introduce orthonormal frames and use these to compute scalar products. The main modification is that the Kronecker delta is replaced by η_{ij} which again is zero if $i \neq j$, but can take values ± 1 if $i = j$.

1.3 Complex numbers

The scalar product is the simplest product one can define between vectors, and once such a product is defined one can formulate many of the key concepts of Euclidean geometry. But this is by no means the only product that can be defined between vectors. In two dimensions a new product can be defined via complex arithmetic. A complex number can be viewed as an ordered pair of real numbers which represents a direction in the complex plane, as was realised by Wessel in 1797. Their product enables complex numbers to perform geometric operations, such as rotations and dilations. But suppose that we take the complex number $z = x + iy$ and square it, forming

$$z^2 = (x + iy)^2 = x^2 - y^2 + 2xyi. \tag{1.16}$$

In terms of vector arithmetic, neither the real nor imaginary parts of this expression have any geometric significance. A more geometrically useful product is defined instead by

$$zz^* = (x + iy)(x - iy) = x^2 + y^2, \tag{1.17}$$

which returns the square of the length of the vector. A product of two vectors in a plane, z and $w = u + vi$, can therefore be constructed as

$$zw^* = (x + iy)(u - iv) = xu + vy + i(uy - vx). \tag{1.18}$$

The real part of the right-hand side recovers the scalar product. To understand the imaginary term consider the polar representation

$$z = |z|e^{i\theta}, \qquad w = |w|e^{i\phi} \tag{1.19}$$

so that

$$zw^* = |z||w|e^{i(\theta - \phi)}. \tag{1.20}$$

The imaginary term has magnitude $|z||w|\sin(\theta - \phi)$, where $\theta - \phi$ is the angle between the two vectors. The magnitude of this term is therefore the area of the parallelogram defined by z and w. The sign of the term conveys information about the *handedness* of the area element swept out by the two vectors. This will be defined more carefully in section 1.6.

We thus have a satisfactory interpretation for both the real and imaginary parts of the product zw^*. The surprising feature is that these are still both parts of a complex number. We thus have a second interpretation for complex addition, as a sum between scalar objects and objects representing plane segments. The advantages of adding these together are precisely the advantages of working with complex numbers as opposed to pairs of real numbers. This is a theme to which we shall return regularly in following chapters.

1.4 Quaternions

The fact that complex arithmetic can be viewed as representing a product for vectors in a plane carries with it a further advantage — it allows us to divide by a vector. Generalising this to three dimensions was a major preoccupation of the physicist W.R. Hamilton (see figure 1.2). Since a complex number $x + iy$ can be represented by two rectangular axes on a plane it seemed reasonable to represent directions in space by a triplet consisting of one real and two complex numbers. These can be written as $x + iy + jz$, where the third term jz represents a third axis perpendicular to the other two. The complex numbers i and j have the properties that $i^2 = j^2 = -1$. The norm for such a triplet would then be

$$(x + iy + jz)(x - iy - jz) = (x^2 + y^2 + z^2) - yz(ij + ji). \tag{1.21}$$

The final term is problematic, as one would like to recover the scalar product here. The obvious solution to this problem is to set $ij = -ji$ so that the last term vanishes.

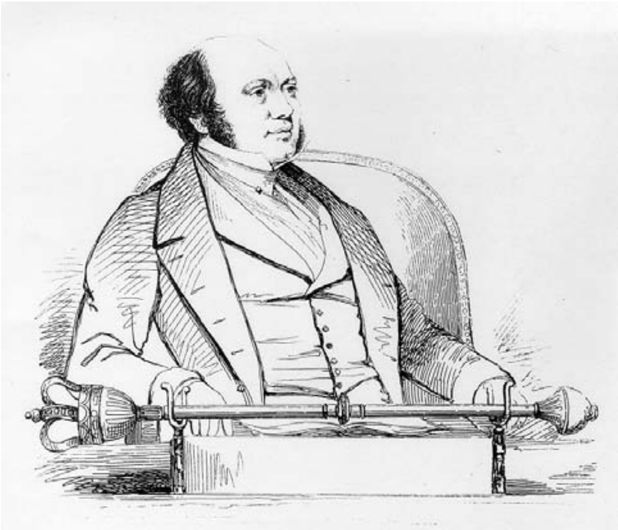

Figure 1.2 *William Rowan Hamilton 1805–1865.* Inventor of quaternions, and one of the key scientific figures of the nineteenth century. He spent many years frustrated at being unable to extend his theory of couples of numbers (complex numbers) to three dimensions. In the autumn of 1843 he returned to this problem, quite possibly prompted by a visit he received from the young German mathematician Eisenberg. Among Eisenberg's papers was the observation that matrices form the elements of an algebra that was much like ordinary arithmetic except that multiplication was non-commutative. This was the vital step required to find the quaternion algebra. Hamilton arrived at this algebra on 16 October 1843 while out walking with his wife, and carved the equations in stone on Brougham Bridge. His discovery of quaternions is perhaps the best-documented mathematical discovery ever.

The anticommutative law $ij = -ji$ ensures that the norm of a triplet behaves sensibly, and also that multiplication of triplets in a plane behaves in a reasonable manner. The same is not true for the general product of triplets, however. Consider

$$\begin{aligned}(a + ib + jc)(x + iy + jz) = (ax - by - cz) + i(ay + bx)\\ + j(az + cx) + ij(bz - cy).\end{aligned} \tag{1.22}$$

Setting $ij = -ji$ is no longer sufficient to remove the ij term, so the algebra does not close. The only thing for Hamilton to do was to set $ij = k$, where k is some unknown, and see if it could be removed somehow. While walking along the Royal Canal he suddenly realised that if his triplets were instead made up of four terms he would be able to close the algebra in a simple, symmetric way.

To understand his discovery, consider

$$(a+ib+jc+kd)(a-ib-jc-kd) \\ = a^2+b^2+c^2+d^2(-k^2)-bd(ik+ki)-cd(jk+kj), \quad (1.23)$$

where we have assumed that $i^2=j^2=-1$ and $ij=-ji$. The expected norm of the above product is $a^2+b^2+c^2+d^2$, which is obtained by setting $k^2=-1$ and $ik=-ki$ and $jk=-kj$. So what values do we use for jk and ik? These follow from the fact that $ij=k$, which gives

$$ik = i(ij) = (ii)j = -j \quad (1.24)$$

and

$$kj = (ij)j = -i. \quad (1.25)$$

Thus the multiplication rules for quaternions are

$$i^2 = j^2 = k^2 = -1 \quad (1.26)$$

and

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \quad (1.27)$$

These can be summarised neatly as $i^2=j^2=k^2=ijk=-1$. It is a simple matter to check that these multiplication laws define a closed algebra.

Hamilton was so excited by his discovery that the very same day he obtained leave to present a paper on the quaternions to the Royal Irish Academy. The subsequent history of the quaternions is a fascinating story which has been described by many authors. Some suggested material for further reading is given at the end of this chapter. In brief, despite the many advantages of working with quaternions, their development was blighted by two major problems.

The first problem was the status of vectors in the algebra. Hamilton identified vectors with *pure quaternions*, which had a null scalar part. On the surface this seems fine — pure quaternions define a three-dimensional vector space. Indeed, Hamilton invented the word '*vector*' precisely for these objects and this is the origin of the now traditional use of i, j and k for a set of orthonormal basis vectors. Furthermore, the full product of two pure quaternions led to the definition of the extremely useful cross product (see section 1.5). The problem is that the product of two pure vectors does not return a new pure vector, so the vector part of the algebra does not close. This means that a number of ideas in complex analysis do not extend easily to three dimensions. Some people felt that this meant that the full quaternion product was of little use, and that the scalar and vector parts of the product should be kept separate. This criticism misses the point that the quaternion product is *invertible*, which does bring many advantages.

The second major difficulty encountered with quaternions was their use in describing rotations. The irony here is that quaternions offer the clearest way of handling rotations in three dimensions, once one realises that they provide a 'spin-1/2' representation of the rotation group. That is, if a is a vector (a pure quaternion) and R is a unit quaternion, a new vector is obtained by the *double-sided* transformation law

$$a' = RaR^*, \tag{1.28}$$

where the $*$ operation reverses the sign of all three 'imaginary' components. A consequence of this is that each of the basis quaternions i, j and k generates rotations through π. Hamilton, however, was led astray by the analogy with complex numbers and tried to impose a single-sided transformation of the form $a' = Ra$. This works if the axis of rotation is perpendicular to a, but otherwise does not return a pure quaternion. More damagingly, it forces one to interpret the basis quaternions as generators of rotations through $\pi/2$, which is simply wrong!

Despite the problems with quaternions, it was clear to many that they were a useful mathematical system worthy of study. Tait claimed that quaternions 'freed the physicist from the constraints of coordinates and allowed thoughts to run in their most natural channels' — a theme we shall frequently meet in this book. Quaternions also found favour with the physicist James Clerk Maxwell, who employed them in his development of the theory of electromagnetism. Despite these successes, however, quaternions were weighed down by the increasingly dogmatic arguments over their interpretation and were eventually displaced by the hybrid system of vector algebra promoted by Gibbs.

1.5 The cross product

Two of the lasting legacies of the quaternion story are the introduction of the idea of a vector, and the cross product between two vectors. Suppose we form the product of two pure quaternions a and b, where

$$a = a_1 i + a_2 j + a_3 k, \qquad b = b_1 i + b_2 j + b_3 k. \tag{1.29}$$

Their product can be written

$$ab = -a_i b_i + c, \tag{1.30}$$

where c is the pure quaternion

$$c = (a_2 b_3 - a_3 b_2) i + (a_3 b_1 - a_1 b_3) j + (a_1 b_2 - a_2 b_1) k. \tag{1.31}$$

Writing $c = c_1 i + c_2 j + c_3 k$ the component relation can be written as

$$c_i = \epsilon_{ijk} a_j b_k, \tag{1.32}$$