Geometric Algebra for Physicists

Chris Doran • Anthony Lasenby

CAMBRIDGE

more information – www.cambridge.org/9780521715959

GEOMETRIC ALGEBRA FOR PHYSICISTS

Geometric algebra is a powerful mathematical language with applications across a range of subjects in physics and engineering. Written by two of the leading researchers in the field, this book is a complete guide to the current state of the subject.

Early chapters provide a self-contained development of geometric algebra and form the basis of an undergraduate lecture course. Topics covered include new techniques for handling rotations in arbitrary dimensions, and the links between rotations, bivectors and the structure of the Lie groups. Following chapters extend the concept of a complex analytic function theory to arbitrary dimensions. This has applications in quantum theory and electromagnetism. All four Maxwell equations are united into one single equation, and new techniques are discussed for its solution. Later chapters cover some advanced topics in physics, including non-Euclidean geometry, quantum entanglement and gauge theories. The final chapters describe the construction of a gauge theory of gravitation in Minkowski spacetime. Using the tools of geometric algebra, advanced applications such as black holes and cosmic strings are explored.

This book will be of interest to researchers working in the fields of geometry, relativity and quantum theory. It can also be used as a textbook for advanced undergraduate and graduate courses on the physical applications of geometric algebra.

CHRIS DORAN obtained his PhD from the University of Cambridge, having gained a distinction in Part III of his undergraduate degree. He was elected a Junior Research Fellow of Churchill College, Cambridge in 1993, was made a Lloyd's of London Fellow in 1996 and was the Schlumberger Interdisciplinary Research Fellow of Darwin College, Cambridge between 1997 and 2000. He is currently a Fellow of Sidney Sussex College, Cambridge and holds an EPSRC Advanced Fellowship. Dr Doran has published widely on aspects of mathematical physics and is currently researching applications of geometric algebra in engineering and computer science.

ANTHONY LASENBY is Professor of Astrophysics and Cosmology at the University of Cambridge, and is currently Head of the Astrophysics Group and the Mullard Radio Astronomy Observatory in the Cavendish Laboratory. He began his astronomical career with a PhD at Jodrell Bank, specialising in the Cosmic Microwave Background, which has been a major subject of his research ever since. After a brief period at the National Radio Astronomy Observatory in America, he moved from Manchester to Cambridge in 1984, and has been at the Cavendish since then. He is the author or coauthor of nearly 200 papers spanning a wide range of fields, from early universe cosmology to computer vision. His introduction to geometric algebra came in 1988, when he encountered the work of David Hestenes for the first time, and since then he has been developing geometric algebra techniques and employing them in his research in many areas.

GEOMETRIC ALGEBRA FOR PHYSICISTS

CHRIS DORAN
and
ANTHONY LASENBY
University of Cambridge

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo,
Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521715959

First published 2003
Reprinted 2003
First paperback edition published 2007
4th printing 2010

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-48022-1 hardback
ISBN 978-0-521-71595-9 paperback

Contents

Preface

The ideas and concepts of physics are best expressed in the language of mathematics. But this language is far from unique. Many different algebraic systems exist and are in use today, all with their own advantages and disadvantages. In this book we describe what we believe to be the most powerful available mathematical system developed to date. This is *geometric algebra*, which is presented as a new mathematical tool to add to your existing set as either a theoretician or experimentalist. Our aim is to introduce the new techniques via their applications, rather than as purely formal mathematics. These applications are diverse, and throughout we emphasise the unity of the mathematics underpinning each of these topics.

The history of geometric algebra is one of the more unusual tales in the development of mathematical physics. William Kingdon Clifford introduced his geometric algebra in the 1870s, building on the earlier work of Hamilton and Grassmann. It is clear from his writing that Clifford intended his algebra to describe the geometric properties of vectors, planes and higher-dimensional objects. But most physicists first encounter the algebra in the guise of the Pauli and Dirac matrix algebras of quantum theory. Few then contemplate using these unwieldy matrices for practical geometric computing. Indeed, some physicists come away from a study of Dirac theory with the view that Clifford's algebra is inherently quantum-mechanical. In this book we aim to dispel this belief by giving a straightforward introduction to this new and fundamentally different approach to vectors and vector multiplication. In this language much of the standard subject matter taught to physicists can be formulated in an elegant and highly condensed fashion. And the portability of the techniques we discuss enables us to reach a range of advanced topics with little extra work.

This book is intended to be of interest to both students and researchers in physics. The early chapters grew out of an undergraduate lecture course that we have run for a number of years in the Physics Department at Cambridge University. We are indebted to the students who attended the early versions of this course, and helped to shape the material into a form suitable for undergraduate tuition. These early chapters require little more than a basic knowledge of linear algebra and vector geometry, and some exposure to classical mechanics. More advanced physical concepts are introduced as the book progresses.

A number of themes run throughout this book. The first is that geometric algebra enables us to express fundamental physics in a language that is free from coordinates or indices. Coordinates are only introduced later, when the geometry of a given problem is clear. This approach gives many equations a degree of clarity which is lost in tensor algebra. A second theme is the way in which rotations are handled in geometric algebra through the use of *rotors*. This approach extends to arbitrary spaces the idea of using a complex phase to rotate in a plane. Rotor techniques can be applied in spaces of arbitrary signature and are particularly well suited to formulating Lorentz and conformal transformations. The latter are central to our treatment of non-Euclidean geometry. Rotors also provide a framework for studying Lie groups and Lie algebras, and are essential to our discussion of gauge theories.

The third theme is the invertibility of the geometric product of vectors, which makes it possible to divide by a vector. This idea extends to the vector derivative, which has an inverse in the form a first-order Green's function. The vector derivative and its inverse enable us to extend complex analytic function theory to arbitrary dimensions. This theory is perfectly suited to electromagnetism, as all four Maxwell equations can be combined into a single spacetime equation involving the invertible vector derivative. The same vector derivative appears in the Dirac theory, and is central to the gauge treatment of gravitation which dominates the final two chapters of this book.

This book would not have been possible without the help and encouragement of a large number of people. We thank Stephen Gull for helping initiate much of the research described here, for his constant advice and criticism, and for use of a number of his figures. We also thank David Hestenes for all his work in shaping the modern subject of geometric algebra and for his constant encouragement. Special mention must be made of our many collaborators, in particular Joan Lasenby, Anthony Challinor, Leo Dorst, Tim Havel, Antony Lewis, Mark Ashdown, Frank Sommen, Shyamal Somaroo, Jeff Tomasi, Bill Fitzgerald, Youri Dabrowski and Mike Hobson. Special thanks also goes to Mike for his help with Latex and explaining the intricacies of the CUP style files. We thank the Physics Department of Cambridge University for the use of their facilities, and for the range of technical advice and expertise we regularly called on. Finally we thank everyone at Cambridge University Press who helped in the production of this book.

CD would also like to thank the EPSRC and Sidney Sussex College for their support, his friends and colleagues, all at Nomads HC, and above all Helen for not complaining about the lost evenings as I worked on this book. I promise to finish the decorating now it is complete.

AL thanks Joan and his children Robert and Alison for their constant enthusiasm and support, and their patience in the face of many explanations of topics from this book.

Cambridge *C.J.L. Doran*
July 2002 *A.N. Lasenby*

Notation

The subject of vector geometry in general, and geometric algebra in particular, suffers from a profusion of notations and conventions. In short, there is no single convention that is perfectly suited to the entire range of applications of geometric algebra. For example, many of the formulae and results given in this book involve arbitrary numbers of vectors and are valid in vector spaces of arbitrary dimensions. These formulae invariably look neater if one does not embolden all of the vectors in the expression. For this reason we typically choose to write vectors in a lower case italic script, a, and more general multivectors in upper case italic script, M. But in some applications, particularly mechanics and dynamics, one often needs to reserve lower case italic symbols for coordinates and scalars, and in these situations writing vectors in bold face is helpful. This convention in adopted in chapter 3.

For many applications it is useful to have a notation which distinguishes frame vectors from general vectors. In these cases we write the former in an upright font as $\{\mathsf{e}_i\}$. But this notation looks clumsy in certain settings, and is not followed rigorously in some of the later chapters. In this book our policy is to ensure that we adopt a consistent notation within each chapter, and any new or distinct features are explained either at the start of the chapter or at their point of introduction.

Some conventions are universally adopted throughout this book, and for convenience we have gathered together a number of these here.

(i) The geometric (or Clifford) algebra generated by the vector space of signature (p, q) is denoted $\mathcal{G}(p, q)$. In the first three chapters we employ the abbreviations $\mathcal{G}_2$ and $\mathcal{G}_3$ for the Euclidean algebras $\mathcal{G}(2, 0)$ and $\mathcal{G}(3, 0)$. In chapter 4 we use $\mathcal{G}_n$ to denote all algebras $\mathcal{G}(p, q)$ of total dimension n.

(ii) The geometric product of A and B is denoted by juxtaposition, AB.

(iii) The inner product is written with a centred dot, $A \cdot B$. The inner product is only employed between homogeneous multivectors.

(iv) The outer (exterior) product is written with a wedge, $A \wedge B$. The outer product is also only employed between homogeneous multivectors.

(v) Inner and outer products are always performed before geometric products. This enables us to remove unnecessary brackets. For example, the expression $a \cdot b\, c$ is to be read as $(a \cdot b)c$.

(vi) Angled brackets $\langle M \rangle_p$ are used to denote the result of projecting onto the terms in M of grade p. The subscript zero is dropped for the projection onto the scalar part.

(vii) The reverse of the multivector M is denoted either with a dagger, $M^\dagger$, or with a tilde, $\tilde{M}$. The latter is employed for applications in spacetime.

(viii) Linear functions are written in an upright font as $\mathsf{F}(a)$ or $\mathsf{h}(a)$. This helps to distinguish linear functions from multivectors. Some exceptions are encountered in chapters 13 and 14, where caligraphic symbols are used for certain tensors in gravitation. The adjoint of a linear function is denoted with a bar, $\bar{\mathsf{h}}(a)$.

(ix) Lie groups are written in capital, Roman font as in $\mathrm{SU}(n)$. The corresponding Lie algebra is written in lower case, $\mathrm{su}(n)$.

Further details concerning the conventions adopted in this book can be found in sections 2.5 and 4.1.

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.

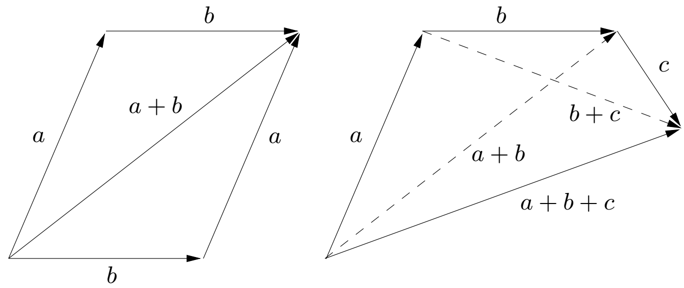

Figure 1.1 *A geometric picture of vector addition.* The result of $a + b$ is formed by adding the tail of b to the head of a. As is shown, the resultant vector $a + b$ is the same as $b + a$. This finds an algebraic expression in the statement that addition is commutative. In the right-hand diagram 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

$\lambda_1, \dots, \lambda_n$ (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 $\{a_1, \dots, a_n\}$ 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

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

The proof is straightforward:

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

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 $\{\mathsf{e}_1, \dots, \mathsf{e}_n\}$ 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

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

so that

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

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.

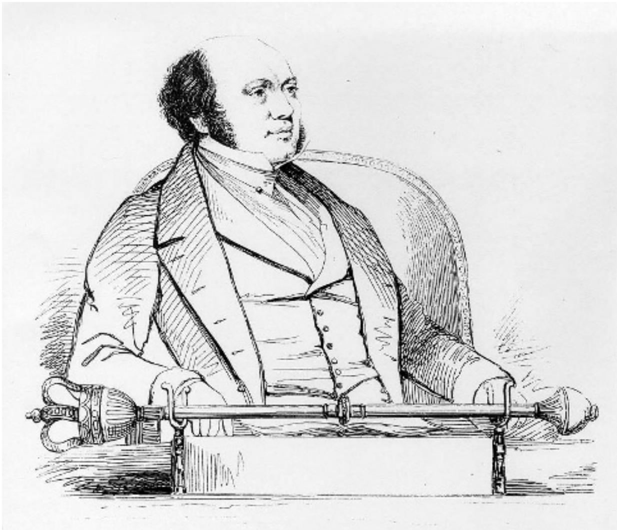

Figure 1.2 *William Rowan Hamilton 1805–1865.* Inventor of quaternions, and one of the key scientific figures of the nineteenth century. He spent many years frustrated at being unable to extend his theory of couples of numbers (complex numbers) to three dimensions. In the autumn of 1843 he returned to 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

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

where the alternating tensor ϵ_{ijk} is defined by

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is a cylic permutation of 123,} \\ -1 & \text{if } ijk \text{ is an anticylic permutation of 123,} \\ 0 & \text{otherwise.} \end{cases} \tag{1.33}$$

We recognise the preceding as defining the cross product of two vectors, $a\times b$. This has the following properties:

(i) $a\times b$ is perpendicular to the plane defined by a and b;
(ii) $a\times b$ has magnitude $|a||b|\sin(\theta)$;
(iii) the vectors a, b and $a\times b$ form a right-handed set.

These properties can alternatively be viewed as defining the cross product, and from them the algebraic definition can be recovered. This is achieved by starting with a right-handed orthonormal frame $\{\mathsf{e}_i\}$. For these we must have

$$\mathsf{e}_1\times\mathsf{e}_2 = \mathsf{e}_3 \qquad \text{etc.} \tag{1.34}$$

so that we can write

$$\mathsf{e}_i\times\mathsf{e}_j = \epsilon_{ijk}\mathsf{e}_k. \tag{1.35}$$

Expanding out a vector in terms of this basis recovers the formula

$$\begin{aligned} a\times b &= (a_i\mathsf{e}_i)\times(b_j\mathsf{e}_j) \\ &= a_ib_j(\mathsf{e}_i\times\mathsf{e}_j) \\ &= (\epsilon_{ijk}a_ib_j)\mathsf{e}_k. \end{aligned} \tag{1.36}$$

Hence the geometric definition recovers the algebraic one.

The cross product quickly proved itself to be invaluable to physicists, dramatically simplifying equations in dynamics and electromagnetism. In the latter part of the nineteenth century many physicists, most notably Gibbs, advocated abandoning quaternions altogether and just working with the individual scalar and cross products. We shall see in later chapters that Gibbs was misguided in some of his objections to the quaternion product, but his considerable reputation carried the day and by the 1900s quaternions had all but disappeared from mainstream physics.

1.6 The outer product

The cross product has one major failing — it only exists in three dimensions. In two dimensions there is nowhere else to go, whereas in four dimensions the concept of a vector orthogonal to a pair of vectors is not unique. To see this, consider four orthonormal vectors $\mathsf{e}_1,\ldots,\mathsf{e}_4$. If we take the pair e_1 and e_2 and attempt

Figure 1.3 *Hermann Gunther Grassmann (1809–1877)*, born in Stettin, Germany (now Szczecin, Poland). A German mathematician and schoolteacher, Grassmann was the third of his parents' twelve children and was born into a family of scholars. His father studied theology and became a minister, before switching to teaching mathematics and physics at the Stettin Gymnasium. Hermann followed in his father's footsteps, first studying theology, classical languages and literature at Berlin. After returning to Stettin in 1830 he turned his attention to mathematics and physics. Grassmann passed the qualifying examination to win a teaching certificate in 1839. This exam included a written assignment on the tides, for which he gave a simplified treatment of Laplace's work based upon a new geometric calculus that he had developed. By 1840 he had decided to concentrate on mathematics research. He published the first edition of his geometric calculus, the 300 page *Lineale Ausdehnungslehre* in 1844, the same year that Hamilton announced the discovery of the quaternions. His work did not achieve the same impact as the quaternions, however, and it was many years before his ideas were understood and appreciated by other mathematicians. Disappointed by this lack of interest, Grassmann turned his attention to linguistics and comparative philology, with greater immediate impact. He was an expert in Sanskrit and translated the *Rig-Veda* (1876–1877). He also formulated the linguistic law (named after him) stating that in Indo-European bases, successive syllables may not begin with aspirates. He died before he could see his ideas on geometry being adopted into mainstream mathematics.

to find a vector perpendicular to both of these, we see that any combination of e_3 and e_4 will do.

A suitable generalisation of the idea of the cross product was constructed by

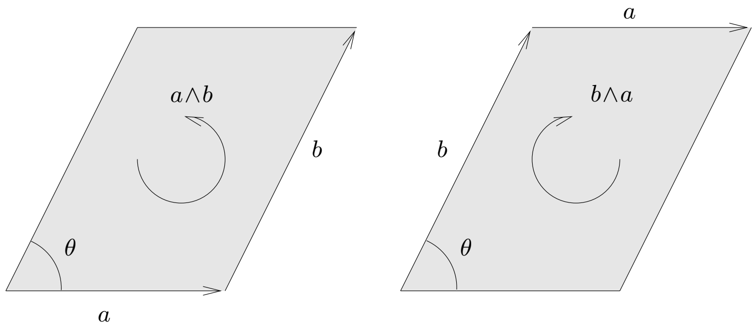

Figure 1.4 *The outer product.* The outer or wedge product of a and b returns a directed area element of area $|a||b|\sin(\theta)$. The orientation of the parallelogram is defined by whether the circuit a, b, $-a$, $-b$ is right-handed (anticlockwise) or left-handed (clockwise). Interchanging the order of the vectors reverses the orientation and introduces a minus sign in the product.

the remarkable German mathematician H.G. Grassmann (see figure 1.3). His work had its origin in the *Barycentrischer Calcul* of Möbius. There the author introduced expressions like AB for the line connecting the points A and B and ABC for the triangle defined by A, B and C. Möbius also introduced the crucial idea that the sign of the quantity should change if any two points are interchanged. (These *oriented* segments are now referred to as *simplices*.) It was Grassmann's leap of genius to realise that expressions like AB could actually be viewed as a product between vectors. He thus introduced the *outer* or *exterior product* which, in modern notation, we write as $a \wedge b$, or 'a wedge b'.

The outer product can be defined on any vector space and, geometrically, we are not forced to picture these vectors as displacements. Indeed, Grassmann was motivated by a *projective* viewpoint, where the elements of the vector space are interpreted as points, and the outer product of two points defines the line through the points. For our purposes, however, it is simplest to adopt a picture in which vectors represent directed line segments. The outer product then provides a means of encoding a plane, without relying on the notion of a vector perpendicular to it. The result of the outer product is therefore neither a scalar nor a vector. It is a new mathematical entity encoding an oriented plane and is called a *bivector*. It can be visualised as the parallelogram obtained by sweeping one vector along the other (figure 1.4). Changing the order of the vectors reverses the orientation of the plane. The magnitude of $a \wedge b$ is $|a||b|\sin(\theta)$, the same as the area of the plane segment swept out by the vectors.

The outer product of two vectors has the following algebraic properties:

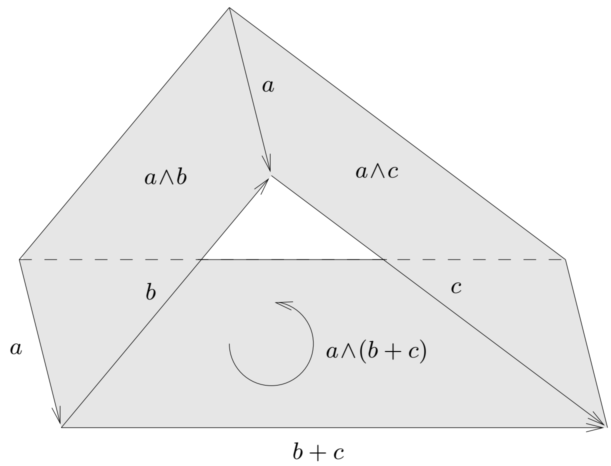

Figure 1.5 *A geometric picture of bivector addition.* In three dimensions any two non-parallel planes share a common line. If this line is denoted a, the two planes can be represented by $a \wedge b$ and $a \wedge c$. Bivector addition proceeds much like vector addition. The planes are combined at a common boundary and the resulting plane is defined by the initial and final edges, as opposed to the initial and final points for vector addition. The mathematical statement of this addition rule is the distributivity of the outer product over addition.

(i) The product is *antisymmetric*:

$$a\wedge b = -b\wedge a. \tag{1.37}$$

This has the geometric interpretation of reversing the orientation of the surface defined by a and b. It follows immediately that

$$a\wedge a = 0, \quad \text{for all vectors } a. \tag{1.38}$$

(ii) Bivectors form a linear space, the same way that vectors do. In two and three dimensions the addition of bivectors is easy to visualise. In higher dimensions this addition is not always so easy to visualise, because two planes need not share a common line.

(iii) The outer product is distributive over addition:

$$a\wedge(b + c) = a\wedge b + a\wedge c. \tag{1.39}$$

This helps to visualise the addition of bivectors which share a common line (see figure 1.5).

While it is convenient to visualise the outer product as a parallelogram, the actual shape of the object is not conveyed by the result of the product. This can be seen easily by defining $a' = a + \lambda b$ and forming

$$a' \wedge b = a \wedge b + \lambda b \wedge b = a \wedge b. \tag{1.40}$$

The same bivector can therefore be generated by many different pairs of vectors. In many ways it is better to replace the picture of a directed parallelogram with that of a directed circle. The circle defines both the plane and a handedness, and its area is equal to the magnitude of the bivector. This therefore conveys all of the information one has about the bivector, though it does make bivector addition harder to visualise.

1.6.1 Two dimensions

The outer product of any two vectors defines a plane, so one has to go to at least two dimensions to form an interesting product. Suppose then that $\{\mathsf{e}_1, \mathsf{e}_2\}$ are an orthonormal basis for the plane, and introduce the vectors

$$a = a_1 \mathsf{e}_1 + a_2 \mathsf{e}_2, \quad b = b_1 \mathsf{e}_1 + b_2 \mathsf{e}_2. \tag{1.41}$$

The outer product $a \wedge b$ contains

$$\begin{aligned} a \wedge b &= a_1 b_1 \mathsf{e}_1 \wedge \mathsf{e}_1 + a_1 b_2 \mathsf{e}_1 \wedge \mathsf{e}_2 + a_2 b_1 \mathsf{e}_2 \wedge \mathsf{e}_1 + a_2 b_2 \mathsf{e}_2 \wedge \mathsf{e}_2 \\ &= (a_1 b_2 - a_2 b_1) \mathsf{e}_1 \wedge \mathsf{e}_2, \end{aligned} \tag{1.42}$$

which recovers the imaginary part of the product of (1.18). The term therefore immediately has the expected magnitude $|a|\,|b| \sin(\theta)$. The coefficient of $\mathsf{e}_1 \wedge \mathsf{e}_2$ is positive if a and b have the same orientation as e_1 and e_2. The orientation is defined by traversing the boundary of the parallelogram defined by the vectors a, b, $-a$, $-b$ (see figure 1.4). By convention, we usually work with a right-handed set of reference axes (viewed from above). In this case the coefficient $a_1 b_2 - a_2 b_1$ will be positive if a and b also form a right-handed pair.

1.6.2 Three dimensions

In three dimensions the space of bivectors is also three-dimensional, because each bivector can be placed in a one-to-one correspondence with the vector perpendicular to it. Suppose that $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$ form a right-handed basis (see comments below), and the two vectors a and b are expanded in this basis as $a = a_i \mathsf{e}_i$ and $b = b_i \mathsf{e}_i$. The bivector $a \wedge b$ can then be decomposed in terms of an orthonormal frame of bivectors by

$$\begin{aligned} a \wedge b &= (a_i \mathsf{e}_i) \wedge (b_j \mathsf{e}_j) \\ &= (a_2 b_3 - b_3 a_2) \mathsf{e}_2 \wedge \mathsf{e}_3 + (a_3 b_1 - a_1 b_3) \mathsf{e}_3 \wedge \mathsf{e}_1 \\ &\quad + (a_1 b_2 - a_2 b_1) \mathsf{e}_1 \wedge \mathsf{e}_2. \end{aligned} \tag{1.43}$$

The components in this frame are therefore the same as those of the cross product. But instead of being the components of a vector perpendicular to a and b, they are the components of the bivector $a \wedge b$. It is this distinction which enables the outer product to be defined in any dimension.

1.6.3 Handedness

We have started to employ the idea of *handedness* without giving a satisfactory definition of it. The only space in which there is an unambiguous definition of handedness is three dimensions, as this is the space we inhabit and most of us can distinguish our left and right hands. This concept of 'left' and 'right' is a man-made convention adopted to make our life easier, and it extends to the concept of a frame in a straightforward way. Suppose that we are presented with three orthogonal vectors $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$. We align the 3 axis with the thumb of our right hand and then close our fist. If the direction in which our fist closes is the same as that formed by rotating from the 1 to the 2 axis, the frame is right-handed. If not, it is left-handed.

Swapping any pair of vectors swaps the handedness of a frame. Performing two such swaps returns us to the original handedness. In three dimensions this corresponds to a cyclic reordering, and ensures that the frames $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$, $\{\mathsf{e}_3, \mathsf{e}_1, \mathsf{e}_2\}$ and $\{\mathsf{e}_2, \mathsf{e}_3, \mathsf{e}_1\}$ all have the same orientation.

There is no agreed definition of a 'right-handed' orientation in spaces of dimensions other than three. All one can do is to make sure that any convention used is adopted consistently. In all dimensions the orientation of a set of vectors is changed if any two vectors are swapped. In two dimensions one does still tend to talk about right-handed axes, though the definition is dependent on the idea of looking down on the plane *from above*. The idea of above and below is not a feature of the plane itself, but depends on how we embed it in our three-dimensional world. There is no definition of left or right-handed which is intrinsic to the plane.

1.6.4 Extending the outer product

The preceding examples demonstrate that in arbitrary dimensions the components of $a \wedge b$ are given by

$$(a \wedge b)_{ij} = a_{[i} b_{j]} \tag{1.44}$$

where the [] denotes antisymmetrisation. Grassmann was able to take this idea further by defining an outer product for any number of vectors. The idea is a simple extension of the preceding formula. Expressed in an orthonormal frame, the components of the outer product on n vectors are the totally antisymmetrised products of the components of each vector. This definition has the useful property that the outer product is *associative*,

$$a \wedge (b \wedge c) = (a \wedge b) \wedge c. \tag{1.45}$$

For example, in three dimensions we have

$$a \wedge b \wedge c = (a_i \mathsf{e}_i) \wedge (b_j \mathsf{e}_j) \wedge (c_k \mathsf{e}_k) = \epsilon_{ijk} a_i b_j c_k \mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \mathsf{e}_3, \tag{1.46}$$

which represents a *directed volume* (see section 2.4).

A further feature of the antisymmetry of the product is that the outer product of any set of linearly dependent vectors vanishes. This means that statements like 'this vector lies on a given plane', or 'these two hypersurfaces share a common line' can be encoded algebraically in a simple manner. Equipped with these ideas, Grassmann was able to construct a system capable of handling geometric concepts in arbitrary dimensions.

Despite Grassmann's considerable achievement, the book describing his ideas, his *Lineale Ausdehnungslehre*, did not have any immediate impact. This was no doubt due largely to his relative lack of reputation (he was still a German schoolteacher when he wrote this work). It was over twenty years before anyone of note referred to Grassmann's work, and during this time Grassmann produced a second, extended version of the *Ausdehnungslehre*. In the latter part of the nineteenth century Grassmann's work started to influence leading figures like Gibbs and Clifford. Gibbs wrote a number of papers praising Grassmann's work and contrasting it favourably with the quaternion algebra. Clifford used Grassmann's work as the starting point for the development of his geometric algebra, the subject of this book.

Today, Grassmann's ideas are recognised as the first presentation of the abstract theory of vector spaces over the field of real numbers. Since his death, his work has given rise to the influential and fashionable areas of *differential forms* and *Grassmann variables*. The latter are anticommuting variables and are fundamental to the foundations of much of modern supersymmetry and superstring theory.

1.7 Notes

Descriptions of linear algebra and vector spaces can be found in most introductory textbooks of mathematics, as can discussions of the scalar and cross products and complex arithmetic. Quaternions, on the other hand, are much less likely to be mentioned. There is a large specialised literature on the quaternions, and a good starting point are the works of Altmann (1986, 1989). Altmann's paper on 'Hamilton, Rodriques and the quaternion scandal' (1989) is also a good introduction to the history of the subject.

The outer product is covered in most modern textbooks on geometry and physics, such as those by Nakahara (1990), Schutz (1980), and Gockeler & Schucker (1987). In most of these works, however, the exterior product is only treated in the context of differential forms. Applications to wider topics in geometry have been discussed by Hestenes (1991) and others. A useful summary in provided in the proceedings of the conference *Hermann Gunther Grassmann (1809–1877)*, edited by Schubring (1996). Grassmann's *Lineale Ausdehnungslehre* is also finally available in English translation due to Kannenberg (1995).

For those with a deeper interest in the history of mathematics and the development of vector algebra a good starting point is the set of books by Kline (1972). There are also biographies available of many of the key protagonists. Perhaps even more interesting is to return to their original papers and experience first hand the robust and often humorous language employed at the time. The collected works of J.W. Gibbs (1906) are particularly entertaining and enlightening, and contain a good deal of valuable historical information.

1.8 Exercises

1.1 Suppose that the two sets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_n\}$ form bases for the same vector space, and suppose initially that $m > n$. By establishing a contradiction, prove the *basis theorem* that all bases of a vector space have the same number of elements.

1.2 Demonstrate that the following define vector spaces:

(a) the set of all polynomials of degree less than or equal to n;
(b) all solutions of a given linear homogeneous ordinary differential equation;
(c) the set of all $n \times m$ matrices.

1.3 Prove that in Euclidean space $|a + b| \leq |a| + |b|$. When does equality hold?

1.4 Show that the unit quaternions $\{\pm 1, \pm i, \pm j \pm k\}$ form a discrete group.

1.5 The unit quaternions i, j, k are generators of rotations about their respective axes. Are rotations through either π or $\pi/2$ consistent with the equation $ijk = -1$?

1.6 Prove the following:

(a) $a\cdot(b\times c) = b\cdot(c\times a) = c\cdot(a\times b)$;
(b) $a\times(b\times c) = a\cdot c\, b - a\cdot b\, c$;
(c) $|a\times b| = |a|\,|b|\sin(\theta)$, where $a\cdot b = |a|\,|b|\cos(\theta)$.

1.7 Prove that the dimension of the space formed by the exterior product of m vectors drawn from a space of dimension n is

$$\frac{n(n-1)\cdots(n-m+1)}{1\cdot 2\cdots m} = \frac{n!}{(n-m)!m!}.$$

1.8 Prove that the n-fold exterior product of a set of n *dependent* vectors is zero.

1.9 A convex polygon in a plane is specified by the ordered set of points $\{x_0, x_1, \dots, x_n\}$. Prove that the directed area of the polygon is given by

$$A = \tfrac{1}{2}(x_0 \wedge x_1 + x_1 \wedge x_2 + \cdots + x_n \wedge x_0).$$

What is the significance of the sign? Can you extend the idea to a triangulated surface in three dimensions?

2

Geometric algebra in two and three dimensions

Geometric algebra was introduced in the nineteenth century by the English mathematician William Kingdon Clifford (figure 2.1). Clifford appears to have been one of the small number of mathematicians at the time to be significantly influenced by Grassmann's work. Clifford introduced his *geometric algebra* by uniting the inner and outer products into a single *geometric* product. This is associative, like Grassmann's product, but has the crucial extra feature of being *invertible*, like Hamilton's quaternion algebra. Indeed, Clifford's original motivation was to unite Grassmann's and Hamilton's work into a single structure. In the mathematical literature one often sees this subject referred to as *Clifford algebra*. We have chosen to follow the example of David Hestenes, and many other modern researchers, by returning to Clifford's original choice of name — *geometric algebra*. One reason for this is that the first published definition of the geometric product was due to Grassmann, who introduced it in the second *Ausdehnungslehre*. It was Clifford, however, who realised the great potential of this product and who was responsible for advancing the subject.

In this chapter we introduce the basics of geometric algebra in two and three dimensions in a way that is intended to appear natural and geometric, if somewhat informal. A more formal, axiomatic approach is delayed until chapter 4, where geometric algebra is defined in arbitrary dimensions. The meaning of the various terms in the algebra we define will be illustrated with familiar examples from geometry. In so doing we will also uncover how Hamilton's quaternions fit into geometric algebra, and understand where it was that Hamilton and his followers went wrong in their treatment of three-dimensional geometry. One of the most powerful applications of geometric algebra is to rotations, and these are considered in some detail in this chapter. It is well known that rotations in a plane can be efficiently handled with complex numbers. We will see how to extend this idea to rotations in three-dimensional space. This representation has many applications in classical and quantum physics.

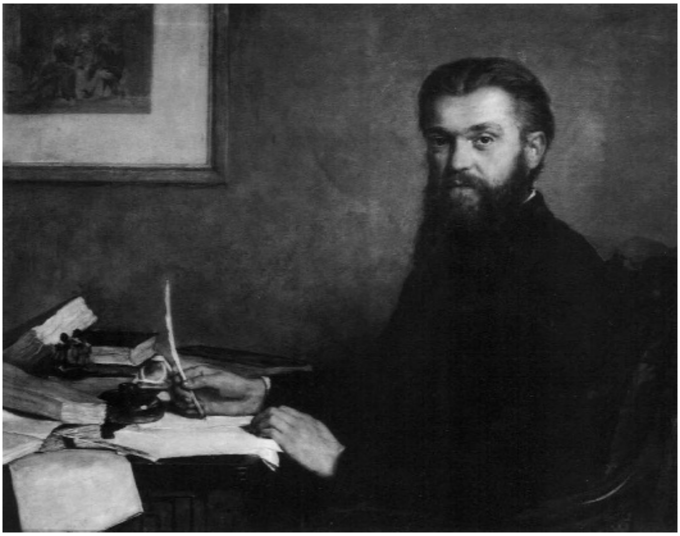

Figure 2.1 *William Kingdon Clifford 1845–1879.* Born in Exeter on 4 May 1845, his father was a justice of the peace and his mother died early in his life. After school he went to King's College, London and then obtained a scholarship to Trinity College, Cambridge, where he followed the likes of Thomson and Maxwell in becoming Second Wrangler. There he also achieved a reputation as a daring athlete, despite his slight frame. He was recommended for a fellowship at Trinity College by Maxwell, and in 1871 took the Professorship of Applied Mathematics at University College, London. He was made a Fellow of the Royal Society at the extremely young age of 29. He married Lucy in 1875, and their house became a fashionable meeting place for scientists and philosophers. As well as being one of the foremost mathematicians of his day, he was an accomplished linguist, philosopher and author of children's stories. Sadly, his insatiable appetite for physical and mental exercise was not matched by his physique, and in 1878 he was instructed to stop work and leave England for the Mediterranean. He returned briefly, only for his health to deteriorate further in the English climate. He left for Madeira, where he died on 3 March 1879 at the age of just 33. Further details of his life can be found in the book *Such Silver Currents* (Chisholm, 2002). Portrait by John Collier (©The Royal Society).

2.1 A new product for vectors

In chapter 1 we studied various products for vectors, including the symmetric scalar (or inner) product and the antisymmetric exterior (or outer) product. In two dimensions, we showed how to interpret the result of the complex product zw^* (section 1.3). The scalar term is the inner product of the two vectors representing the points in the complex plane, and the imaginary term records their directed area. Furthermore, the scalar term is symmetric, and the imaginary term is antisymmetric in the two arguments. Clifford's powerful idea was to generalise this product to arbitrary dimensions by replacing the imaginary term with the outer product. The result is the *geometric product* and is written simply as ab. The result is the sum of a scalar and a bivector, so

$$ab = a\cdot b + a\wedge b. \tag{2.1}$$

This sum of two distinct objects — a scalar and a bivector — looks strange at first and goes against the rule that one should only add like objects. This is the feature of geometric algebra that initially causes the greatest difficulty, in much the same way that $i^2 = -1$ initially unsettles most school children. So how is the sum on the right-hand side of equation (2.1) to be viewed? The answer is that it should be viewed in precisely the same way as the addition of a real and an imaginary number. The result is neither purely real nor purely imaginary — it is a mixture of two different objects which are combined to form a single complex number. Similarly, the addition of a scalar to a bivector enables us to keep track of the separate components of the product ab. The advantages of this are precisely the same as the advantages of complex arithmetic over working with the separate real and imaginary parts. This analogy between *multivectors* in geometric algebra and complex numbers is more than a mere pedagogical device. As we shall discover, geometric algebra encompasses both complex numbers and quaternions. Indeed, Clifford's achievement was to generalise complex arithmetic to spaces of arbitrary dimensions.

From the symmetry and antisymmetry of the terms on the right-hand side of equation (2.1) we see that

$$ba = b\cdot a + b\wedge a = a\cdot b - a\wedge b. \tag{2.2}$$

It follows that

$$a\cdot b = \tfrac{1}{2}(ab + ba) \tag{2.3}$$

and

$$a\wedge b = \tfrac{1}{2}(ab - ba). \tag{2.4}$$

We can thus define the inner and outer products in terms of the geometric product. This forms the starting point for an axiomatic development of geometric algebra, which is presented in chapter 4.

If we form the product of a and the parallel vector λa we obtain

$$a(\lambda a) = \lambda a\cdot a + \lambda a\wedge a = \lambda a\cdot a, \tag{2.5}$$

which is therefore a pure scalar. It follows similarly that a^2 is a scalar, so we can write $a^2 = |a|^2$ for the square of the length of a vector. If instead a and b

are perpendicular vectors, their product is

$$ab = a\cdot b + a\wedge b = a\wedge b \tag{2.6}$$

and so is a pure bivector. We also see that

$$ba = b\cdot a + b\wedge a = -a\wedge b = -ab, \tag{2.7}$$

which shows us that *orthogonal vectors anticommute*. The geometric product between general vectors encodes the relative contributions of both their parallel and perpendicular components, summarising these in the separate scalar and bivector terms.

2.2 An outline of geometric algebra

Clifford went further than just allowing scalars to be added to bivectors. He defined an algebra in which elements of any type could be added or multiplied together. This is what he called a *geometric algebra*. Elements of a geometric algebra are called *multivectors* and these form a linear space — scalars can be added to bivectors, and vectors, etc. Geometric algebra is a *graded* algebra, and elements of the algebra can be broken up into terms of different *grade*. The scalar objects are assigned grade-0, the vectors grade-1, the bivectors grade-2 and so on. Essentially, the grade of the object is the dimension of the hyperplane it specifies. The term 'grade' is preferred to 'dimension', however, as the latter is regularly employed for the size of a linear space. We denote the operation of projecting onto the terms of a chosen grade by $\langle\ \rangle_r$, so $\langle ab\rangle_2$ denotes the grade-2 (bivector) part of the geometric product ab. That is,

$$\langle ab\rangle_2 = a\wedge b. \tag{2.8}$$

The subscript 0 on the scalar term is usually suppressed, so we also have

$$\langle ab\rangle_0 = \langle ab\rangle = a\cdot b. \tag{2.9}$$

Arbitrary multivectors can also be multiplied together with the geometric product. To do this we first extend the geometric product of two vectors to an arbitrary number of vectors. This is achieved with the additional rule that the geometric product is *associative*:

$$a(bc) = (ab)c = abc. \tag{2.10}$$

The associativity property enables us to remove the brackets and write the product as abc. Arbitrary multivectors can now be written as sums of products of vectors. The geometric product of multivectors therefore inherits the two main properties of the product for vectors, which is to say it is associative:

$$A(BC) = (AB)C = ABC, \tag{2.11}$$

and distributive over addition:

$$A(B + C) = AB + AC. \tag{2.12}$$

Here $A, B, \ldots, C$ denote multivectors containing terms of arbitrary grade.

The associativity property ensures that it is now possible to divide by vectors, thus realising Hamilton's goal. Suppose that we know that $ab = C$, where C is some combination of a scalar and bivector. We find that

$$Cb = (ab)b = a(bb) = ab^2, \tag{2.13}$$

so we can define $b^{-1} = b/b^2$, and recover a from

$$a = Cb^{-1}. \tag{2.14}$$

This ability to divide by vectors gives the algebra considerable power.

As an example of these axioms in action, consider forming the square of the bivector $a \wedge b$. The properties of the geometric product allow us to write

$$\begin{aligned}(a\wedge b)(a\wedge b) &= (ab - a\cdot b)(a\cdot b - ba) \\ &= -ab^2a - (a\cdot b)^2 + a\cdot b(ab + ba) \\ &= (a\cdot b)^2 - a^2b^2 \\ &= -a^2b^2\sin^2(\theta),\end{aligned} \tag{2.15}$$

where we have assumed that $a\cdot b = |a|\,|b|\cos(\theta)$. The magnitude of the bivector $a\wedge b$ is therefore equal to the area of the parallelogram with sides defined by a and b. Manipulations such as these are commonplace in geometric algebra, and can provide simplified proofs of a number of useful results.

2.3 Geometric algebra of the plane

The easiest way to understand the geometric product is by example, so consider a two-dimensional space (a plane) spanned by two orthonormal vectors e_1 and e_2. These basis vectors satisfy

$$\mathsf{e}_1{}^2 = \mathsf{e}_2{}^2 = 1, \qquad \mathsf{e}_1\cdot\mathsf{e}_2 = 0. \tag{2.16}$$

The final entity present in the algebra is the bivector $\mathsf{e}_1 \wedge \mathsf{e}_2$. This is the highest grade element in the algebra, since the outer product of a set of dependent vectors is always zero. The highest grade element in a given algebra is usually called the *pseudoscalar*, and its grade coincides with the dimension of the underlying vector space.

The full algebra is spanned by the basis set

$$\begin{array}{ccc} 1 & \{\mathsf{e}_1, \mathsf{e}_2\} & \mathsf{e}_1 \wedge \mathsf{e}_2 \\ \text{1 scalar} & \text{2 vectors} & \text{1 bivector} \end{array}. \tag{2.17}$$

We denote this algebra $\mathcal{G}_2$. Any multivector can be decomposed in this basis, and sums and products can be calculated in terms of this basis. For example, suppose that the multivectors A and B are given by

$$A = \alpha_0 + \alpha_1 \mathsf{e}_1 + \alpha_2 \mathsf{e}_2 + \alpha_3 \mathsf{e}_1 \wedge \mathsf{e}_2,$$
$$B = \beta_0 + \beta_1 \mathsf{e}_1 + \beta_2 \mathsf{e}_2 + \beta_3 \mathsf{e}_1 \wedge \mathsf{e}_2,$$

then their sum $S = A + B$ is given by

$$S = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)\mathsf{e}_1 + (\alpha_2 + \beta_2)\mathsf{e}_2 + (\alpha_3 + \beta_3)\mathsf{e}_1 \wedge \mathsf{e}_2. \tag{2.18}$$

This result for the addition of multivectors is straightforward and unsurprising. Matters become more interesting, however, when we start forming products.

2.3.1 The bivector and its products

To study the properties of the bivector $\mathsf{e}_1 \wedge \mathsf{e}_2$ we first recall that for orthogonal vectors the geometric product is a pure bivector:

$$\mathsf{e}_1 \mathsf{e}_2 = \mathsf{e}_1 \cdot \mathsf{e}_2 + \mathsf{e}_1 \wedge \mathsf{e}_2 = \mathsf{e}_1 \wedge \mathsf{e}_2, \tag{2.19}$$

and that orthogonal vectors anticommute:

$$\mathsf{e}_2 \mathsf{e}_1 = \mathsf{e}_2 \wedge \mathsf{e}_1 = -\mathsf{e}_1 \wedge \mathsf{e}_2 = -\mathsf{e}_1 \mathsf{e}_2. \tag{2.20}$$

We can now form products in which $\mathsf{e}_1 \mathsf{e}_2$ multiplies vectors from the left and the right. First from the left we find that

$$(\mathsf{e}_1 \wedge \mathsf{e}_2)\mathsf{e}_1 = (-\mathsf{e}_2 \mathsf{e}_1)\mathsf{e}_1 = -\mathsf{e}_2 \mathsf{e}_1 \mathsf{e}_1 = -\mathsf{e}_2 \tag{2.21}$$

and

$$(\mathsf{e}_1 \wedge \mathsf{e}_2)\mathsf{e}_2 = (\mathsf{e}_1 \mathsf{e}_2)\mathsf{e}_2 = \mathsf{e}_1 \mathsf{e}_2 \mathsf{e}_2 = \mathsf{e}_1. \tag{2.22}$$

If we assume that e_1 and e_2 form a right-handed pair, we see that left-multiplication by the bivector rotates vectors 90° clockwise (i.e. in a negative sense). Similarly, acting from the right

$$\mathsf{e}_1(\mathsf{e}_1 \mathsf{e}_2) = \mathsf{e}_2, \qquad \mathsf{e}_2(\mathsf{e}_1 \mathsf{e}_2) = -\mathsf{e}_1. \tag{2.23}$$

So right multiplication rotates 90° anticlockwise — a positive sense.

The final product in the algebra to consider is the square of the bivector $\mathsf{e}_1 \wedge \mathsf{e}_2$:

$$(\mathsf{e}_1 \wedge \mathsf{e}_2)^2 = \mathsf{e}_1 \mathsf{e}_2 \mathsf{e}_1 \mathsf{e}_2 = -\mathsf{e}_1 \mathsf{e}_1 \mathsf{e}_2 \mathsf{e}_2 = -1. \tag{2.24}$$

Geometric considerations have led naturally to a quantity which squares to -1. This fits with the fact that two successive left (or right) multiplications of a vector by $\mathsf{e}_1 \mathsf{e}_2$ rotates the vector through 180°, which is equivalent to multiplying by -1. The fact that we now have a firm geometric picture for objects whose algebraic square is -1 opens up the possibility of providing a geometric interpretation for the unit imaginary employed throughout physics, a theme which will be explored further in this book.

2.3.2 Multiplying multivectors

Now that all of the individual products have been found, we can compute the product of the two general multivectors A and B of equation (2.18),

$$AB = M = \mu_0 + \mu_1 \mathsf{e}_1 + \mu_2 \mathsf{e}_2 + \mu_3 \mathsf{e}_1 \mathsf{e}_2, \tag{2.25}$$

where

$$\begin{aligned}
\mu_0 &= \alpha_0\beta_0 + \alpha_1\beta_1 + \alpha_2\beta_2 - \alpha_3\beta_3, \\
\mu_1 &= \alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_3\beta_2 - \alpha_2\beta_3, \\
\mu_2 &= \alpha_0\beta_2 + \alpha_2\beta_0 + \alpha_1\beta_3 - \alpha_3\beta_1, \\
\mu_3 &= \alpha_0\beta_3 + \alpha_3\beta_0 + \alpha_1\beta_2 - \alpha_2\beta_1.
\end{aligned} \tag{2.26}$$

The full product shown here is actually rarely used, but writing it out explicitly does emphasise some of its key features. The product is always well defined, and the algebra is closed under it. Indeed, the product could easily be made an intrinsic part of a computer language, in the same way that complex arithmetic is already intrinsic to some languages. The basis vectors can also be represented with matrices, for example

$$\boldsymbol{E}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \boldsymbol{E}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.27}$$

(Verifying that these satisfy the required algebraic relations is left as an exercise.) Geometric algebras in general are associative algebras, so it is always possible to construct a matrix representation for them. The problem with this is that the matrices hide the geometric content of the elements they represent. Much of the mathematical literature does focus on matrix representations, and for this work the term *Clifford algebra* is appropriate. For the applications in this book, however, the underlying geometry is the important feature of the algebra and matrix representations are usually redundant. *Geometric algebra* is a much more appropriate name for this subject.

2.3.3 Connection with complex numbers

It is clear that there is a close relationship between geometric algebra in two dimensions and the algebra of complex numbers. The unit bivector squares to -1 and generates rotations through $90°$. The combination of a scalar and a bivector, which is formed naturally via the geometric product, can therefore be viewed as a complex number. We write this as

$$Z = u + v\mathsf{e}_1\mathsf{e}_2 = u + Iv, \tag{2.28}$$

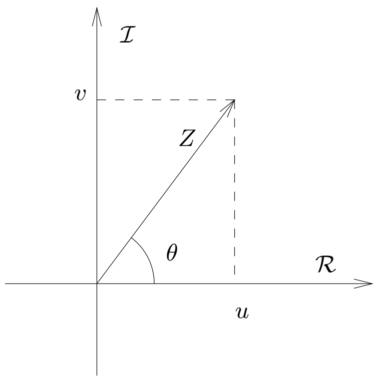

Figure 2.2 *The Argand diagram.* The complex number $Z = u + iv$ represents a vector in the complex plane, with Cartesian components u and v. The polar decomposition into $|Z|\exp(i\theta)$ can alternatively be viewed as an instruction to rotate 1 through θ and dilate by $|Z|$.

where

$$I = \mathsf{e}_1 \wedge \mathsf{e}_2, \qquad I^2 = -1. \tag{2.29}$$

Throughout we employ the symbol I for the pseudoscalar of the algebra of interest. That is why we have used it here, rather than the tempting alternative i. The latter is seen often in the literature, but the i symbol has the problem of suggesting an element which commutes with all others, which is not necessarily a property of the pseudoscalar.

Complex numbers serve a dual purpose in two dimensions. They generate rotations and dilations through their polar decomposition $|Z|\exp(i\theta)$, and they also represent vectors as points on the Argand diagram (see figure 2.2). But in the geometric algebra $\mathcal{G}_2$ complex numbers are replaced by scalar + bivector combinations, whereas vectors are grade-1 objects,

$$x = u\mathsf{e}_1 + v\mathsf{e}_2. \tag{2.30}$$

Is there a natural map between x and the multivector Z? The answer is simple — pre-multiply by e_1,

$$\mathsf{e}_1 x = u + v\mathsf{e}_1\mathsf{e}_2 = u + Iv = Z. \tag{2.31}$$

That is all there is to it! The role of the preferred vector e_1 is clear — it is the real axis. Using this product vectors in a plane can be interchanged with complex numbers in a natural manner.

If we now consider the complex conjugate of Z, $Z^\dagger = u - iv$, we see that

$$Z^\dagger = u + v\mathsf{e}_2\mathsf{e}_1 = x\mathsf{e}_1, \tag{2.32}$$

which has simply reversed the order of the geometric product of x and e_1. This operation of reversing the order of products is one of the fundamental operations performed in geometric algebra, and is called *reversion* (see section 2.5). Suppose now that we introduce a second complex number W, with vector equivalent y:

$$W = \mathsf{e}_1 y. \tag{2.33}$$

The complex product $ZW^\dagger = W^\dagger Z$ now becomes

$$W^\dagger Z = y\mathsf{e}_1\mathsf{e}_1 x = yx, \tag{2.34}$$

which returns the geometric product yx. This is as expected, as the complex product was used to suggest the form of the geometric product.

2.3.4 Rotations

Since we know how to rotate complex numbers, we can use this to find a formula for rotating vectors in a plane. We know that a positive rotation through an angle ϕ for a complex number Z is achieved by

$$Z \mapsto Z' = \mathrm{e}^{i\phi} Z, \tag{2.35}$$

where i is the standard unit imaginary (see figure 2.3). Again, we now view Z as a combination of a scalar and a pseudoscalar in $\mathcal{G}_2$ and so replace i with I. The exponential of $I\phi$ is defined by power series in the normal way, so we still have

$$\mathrm{e}^{I\phi} = \sum_{n=0}^{\infty} \frac{(I\phi)^n}{n!} = \cos\phi + I\,\sin\phi. \tag{2.36}$$

Suppose that Z' has the vector equivalent x',

$$x' = \mathsf{e}_1 Z'. \tag{2.37}$$

We now have a means of rotating the vector directly by writing

$$x' = \mathsf{e}_1 \mathrm{e}^{I\phi} Z = \mathsf{e}_1 \mathrm{e}^{I\phi} \mathsf{e}_1 x. \tag{2.38}$$

But

$$\begin{aligned} \mathsf{e}_1 \mathrm{e}^{I\phi} \mathsf{e}_1 &= \mathsf{e}_1(\cos\phi + I\,\sin\phi)\mathsf{e}_1 \\ &= \cos\phi - I\,\sin\phi = \mathrm{e}^{-I\phi}, \end{aligned} \tag{2.39}$$

where we have employed the result that I *anticommutes* with vectors. We therefore arrive at the formulae

$$x' = \mathrm{e}^{-I\phi} x = x\mathrm{e}^{I\phi}, \tag{2.40}$$

which achieve a rotation of the vector x in the I plane, through an angle ϕ. In section 2.7 we show how to extend this idea to arbitrary dimensions. The

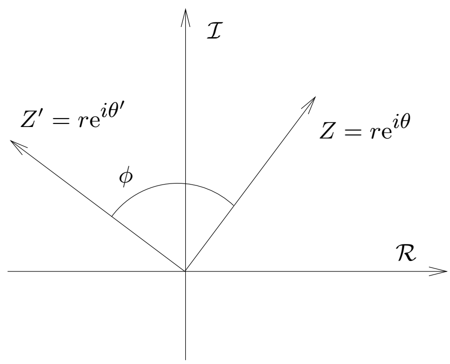

Figure 2.3 *A rotation in the complex plane.* The complex number Z is multiplied by the phase term $\exp(I\phi)$, the effect of which is to replace θ by $\theta' = \theta + \phi$.

change of sign in the exponential acting from the left and right of the vector x is to be expected. We saw earlier that left-multiplication by I generated left-handed rotations, and right-multiplication generated right-handed rotations. As the overall rotation is right-handed, the sign of I must be negative when acting from the left.

This should illustrate that geometric algebra fully encompasses complex arithmetic, and we will see later that complex analysis is fully incorporated as well. The beauty of the geometric algebra formulation is that it shows immediately how to extend the ideas of complex analysis to higher dimensions, a problem which had troubled mathematicians for many years. The key to this is the separation of the two roles of complex numbers by treating vectors as grade-1 objects, and the quantities acting on them (the complex numbers) as combinations of grade-0 and grade-2 objects. These two roles generalise differently in higher dimensions and, once one sees this, extending complex analysis becomes straightforward.

2.4 The geometric algebra of space

The geometric algebra of three-dimensional space is a remarkably powerful tool for solving problems in geometry and classical mechanics. It describes vectors, planes and volumes in a single algebra, which contains all of the familiar vector operations. These include the vector cross product, which is revealed as a disguised form of bivector. The algebra also provides a very clear and compact method for encoding rotations, which is considerably more powerful than working with matrices.

We have so far constructed the geometric algebra of a plane. We now add a third vector e_3 to our two-dimensional set $\{\mathsf{e}_1, \mathsf{e}_2\}$. All three vectors are assumed to be orthonormal, so they all *anticommute*. From these three basis vectors we generate the independent bivectors

$$\{\mathsf{e}_1\mathsf{e}_2, \mathsf{e}_2\mathsf{e}_3, \mathsf{e}_3\mathsf{e}_1\}.$$

This is the expected number of independent planes in space. There is one further term to consider, which is the product of all three vectors:

$$(\mathsf{e}_1\mathsf{e}_2)\mathsf{e}_3 = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_3. \tag{2.41}$$

This results in a grade-3 object, called a *trivector*. It corresponds to sweeping the bivector $\mathsf{e}_1 \wedge \mathsf{e}_2$ along the vector e_3, resulting in a three-dimensional volume element (see section 2.4.3). The trivector represents the unique volume element in three dimensions. It is the highest grade element and is unique up to scale (or volume) and handedness (sign). This is again called the *pseudoscalar* for the algebra.

In three dimensions there are no further directions to add, so the algebra is spanned by

$$\begin{array}{cccc} 1 & \{\mathsf{e}_i\} & \{\mathsf{e}_i \wedge \mathsf{e}_j\} & \mathsf{e}_1\mathsf{e}_2\mathsf{e}_3 \\ \text{1 scalar} & \text{3 vectors} & \text{3 bivectors} & \text{1 trivector} \end{array} \tag{2.42}$$

This basis defines a graded linear space of total dimension $8 = 2^3$. We call this algebra $\mathcal{G}_3$. Notice that the dimensions of each subspace are given by the binomial coefficients.

2.4.1 Products of vectors and bivectors

Our expanded algebra gives us a number of new products to consider. We start by considering the product of a vector and a bivector. We have already looked at this in two dimensions, and found that a normalised bivector rotates vectors in its plane by 90°. Each of the basis bivectors in equation (2.42) shares the properties of the single bivector studied previously for two dimensions. So

$$(\mathsf{e}_1\mathsf{e}_2)^2 = (\mathsf{e}_2\mathsf{e}_3)^2 = (\mathsf{e}_3\mathsf{e}_1)^2 = -1 \tag{2.43}$$

and each bivector generates 90° rotations in its own plane.

The geometric product for vectors extends to all objects in the algebra, so we can form expressions such as aB, where a is a vector and B is a bivector. Now that our algebra contains a trivector $\mathsf{e}_1(\mathsf{e}_2 \wedge \mathsf{e}_3)$, we see that the result of the product aB can contain both vector and trivector terms, the latter arising if a does not lie fully in the B plane. To understand the properties of the product aB we first decompose a into terms in and out of the plane,

$$a = a_{\parallel} + a_{\perp}, \tag{2.44}$$

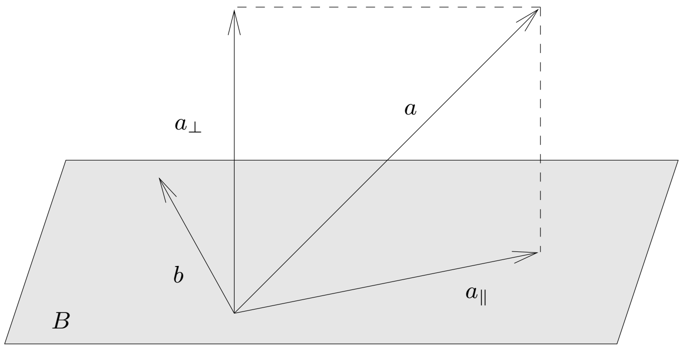

Figure 2.4 *A vector and a bivector.* The vector a can be written as the sum of a term in the plane B and a term perpendicular to the plane, so that $a = a_\parallel + a_\perp$. The bivector B can be written as $a_\parallel \wedge b$, where b is perpendicular to $a_\parallel$.

as shown in figure 2.4. We can now write $aB = (a_\parallel + a_\perp)B$. Suppose that we also write

$$B = a_\parallel \wedge b = a_\parallel b, \tag{2.45}$$

where b is orthogonal to $a_\parallel$ in the B plane. It is always possible to find such a vector b. We now see that

$$a_\parallel B = a_\parallel(a_\parallel b) = {a_\parallel}^2 b \tag{2.46}$$

and so is a vector. This is clear in that the product of a plane with a vector in the plane must remain in the plane. On the other hand

$$a_\perp B = a_\perp(a_\parallel \wedge b) = a_\perp a_\parallel b, \tag{2.47}$$

which is the product of three orthogonal (anticommuting) vectors and so is a trivector. As expected, the product of a vector and a bivector will in general contain vector and trivector terms.

To explore this further let us form the product of the vector a with the bivector $b \wedge c$. From the associative and distributive properties of the geometric product we have

$$a(b \wedge c) = a\tfrac{1}{2}(bc - cb) = \tfrac{1}{2}(abc - acb). \tag{2.48}$$

We now use the rearrangement

$$ab = 2a\cdot b - ba \tag{2.49}$$

to write

$$a(b\wedge c) = (a\cdot b)c - (a\cdot c)b - \tfrac{1}{2}(bac - cab)$$
$$= 2(a\cdot b)c - 2(a\cdot c)b + \tfrac{1}{2}(bc - cb)a, \tag{2.50}$$

so that

$$a(b\wedge c) - (b\wedge c)a = 2(a\cdot b)c - 2(a\cdot c)b. \tag{2.51}$$

The right-hand side of this equation is a vector, so the antisymmetrised product of a vector with a bivector is another vector. Since this operation is grade-lowering, we give it the dot symbol again and write

$$a\cdot B = \tfrac{1}{2}(aB - Ba), \tag{2.52}$$

where B is an arbitrary bivector. The preceding rearrangement means that we have proved one of the most useful results in geometric algebra,

$$a\cdot(b\wedge c) = a\cdot b\, c - a\cdot c\, b. \tag{2.53}$$

Returning to equation (2.46) we see that we must have

$$a\cdot B = a_{\|}B = a_{\|}\cdot B. \tag{2.54}$$

So the effect of taking the inner product of a vector with a bivector is to project onto the component of the vector in the plane, and then rotate this through 90° and dilate by the magnitude of B. We can also confirm that

$$a\cdot B = {a_{\|}}^2 b = -(a_{\|}b)a_{\|} = -B\cdot a, \tag{2.55}$$

as expected.

The remaining part of the product of a vector and a bivector returns a grade-3 trivector. This product is denoted with a wedge since it is grade-raising, so

$$a\wedge(b\wedge c) = \tfrac{1}{2}\big(a(b\wedge c) + (b\wedge c)a\big). \tag{2.56}$$

A few lines of algebra confirm that this outer product is associative,

$$\begin{aligned} a\wedge(b\wedge c) &= \tfrac{1}{2}\big(a(b\wedge c) + (b\wedge c)a\big) \\ &= \tfrac{1}{4}\big(abc - acb + bca - cba\big) \\ &= \tfrac{1}{4}\big(2(a\wedge b)c + bac + bca + 2c(a\wedge b) - cab - acb\big) \\ &= \tfrac{1}{2}\big((a\wedge b)c + c(a\wedge b) + b(c\cdot a) - (c\cdot a)b\big) \\ &= (a\wedge b)\wedge c, \end{aligned} \tag{2.57}$$

so we can unambiguously write the result as $a \wedge b \wedge c$. The product $a \wedge b \wedge c$ is therefore associative and antisymmetric on all pairs of vectors, and so is precisely Grassmann's exterior product (see section 1.6). This demonstrates that

Grassmann's exterior product sits naturally within geometric algebra. From equation (2.47) we have

$$a \wedge B = a_\perp B = a_\perp \wedge B, \tag{2.58}$$

so the effect of the exterior product with a bivector is to project onto the component of the vector perpendicular to the plane, and return a volume element (a trivector). We can confirm simply that this product is symmetric in its vector and bivector arguments:

$$a \wedge B = a_\perp \wedge a_\parallel \wedge b = -a_\parallel \wedge a_\perp \wedge b = a_\parallel \wedge b \wedge a_\perp = B \wedge a. \tag{2.59}$$

The full product of a vector and a bivector can now be written as

$$aB = a \cdot B + a \wedge B, \tag{2.60}$$

where the dot is generalised to mean the *lowest* grade part of the product, while the wedge means the *highest* grade part of the product. In a similar manner to the geometric product of vectors, the separate dot and wedge products can be written in terms of the geometric product as

$$\begin{aligned} a \cdot B &= \tfrac{1}{2}(aB - Ba), \\ a \wedge B &= \tfrac{1}{2}(aB + Ba). \end{aligned} \tag{2.61}$$

But pay close attention to the signs in these formulae, which are the opposite way round to the case of two vectors. The full product of a vector and a bivector wraps up the separate vector and trivector terms in the single product aB. The advantage of this is again that the full product is invertible.

2.4.2 The bivector algebra

Our three independent bivectors also give us another new product to consider. We already know that squaring a bivector results in a scalar. But if we multiply together two bivectors representing orthogonal planes we find that, for example,

$$(\mathsf{e}_1 \wedge \mathsf{e}_2)(\mathsf{e}_2 \wedge \mathsf{e}_3) = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_2\mathsf{e}_3 = \mathsf{e}_1\mathsf{e}_3, \tag{2.62}$$

resulting in a third bivector. We also find that

$$(\mathsf{e}_2 \wedge \mathsf{e}_3)(\mathsf{e}_1 \wedge \mathsf{e}_2) = \mathsf{e}_3\mathsf{e}_2\mathsf{e}_2\mathsf{e}_1 = \mathsf{e}_3\mathsf{e}_1 = -\mathsf{e}_1\mathsf{e}_3, \tag{2.63}$$

so the product of orthogonal bivectors is antisymmetric. The symmetric contribution vanishes because the two planes are perpendicular.

If we introduce the following labelling for the basis bivectors:

$$B_1 = \mathsf{e}_2\mathsf{e}_3, \quad B_2 = \mathsf{e}_3\mathsf{e}_1, \quad B_3 = \mathsf{e}_1\mathsf{e}_2, \tag{2.64}$$

we find that their product satisfies

$$B_i B_j = -\delta_{ij} - \epsilon_{ijk} B_k. \tag{2.65}$$

There is a clear analogy with the geometric product of vectors here, in that the symmetric part is a scalar, whereas the antisymmetric part is a bivector. In higher dimensions it turns out that the symmetrised product of two bivectors can have grade-0 and grade-4 terms (which we will ultimately denote with the dot and wedge symbols). The antisymmetrised product is always a bivector, and bivectors form a closed algebra under this product.

The basis bivectors satisfy

$$B_1{}^2 = B_2{}^2 = B_3{}^2 = -1 \tag{2.66}$$

and

$$B_1 B_2 = -B_2 B_1, \quad \text{etc.} \tag{2.67}$$

These are the properties of the generators of the quaternion algebra (see section 1.4). This observation helps to sort out some of the problems encountered with the quaternions. Hamilton attempted to identify pure quaternions (null scalar part) with vectors, but we now see that they are actually *bivectors*. This causes problems when looking at how objects transform under reflections. Hamilton also imposed the condition $ijk = -1$ on his unit quaternions, whereas we have

$$B_1 B_2 B_3 = \mathsf{e}_2\mathsf{e}_3\mathsf{e}_3\mathsf{e}_1\mathsf{e}_1\mathsf{e}_2 = +1. \tag{2.68}$$

To set up an isomorphism we must flip a sign somewhere, for example in the y component:

$$i \leftrightarrow B_1, \quad j \leftrightarrow -B_2, \quad k \leftrightarrow B_3. \tag{2.69}$$

This shows us that the quaternions are a *left-handed* set of bivectors, whereas Hamilton and others attempted to view the i, j, k as a right-handed set of vectors. Not surprisingly, this was a potential source of great confusion and meant one had to be extremely careful when applying quaternions in vector algebra.

2.4.3 The trivector

Given three vectors, a, b and c, the trivector $a\wedge b\wedge c$ is formed by sweeping $a\wedge b$ along the vector c (see figure 2.5). The result can be represented pictorially as an oriented parallelepiped. As with bivectors, however, the picture should not be interpreted too literally. The trivector $a \wedge b \wedge c$ does not contain any shape information. It just records a volume and an orientation.

The various algebraic properties of trivectors have straightforward geometric interpretations. The same oriented volume is obtained by sweeping $a\wedge b$ along c or $b \wedge c$ along a. The mathematical expression of this is that the outer product is associative, $a \wedge (b \wedge c) = (a \wedge b) \wedge c$. The trivector $a \wedge b \wedge c$ changes sign under interchange of any pair of vectors, which follows immediately from the

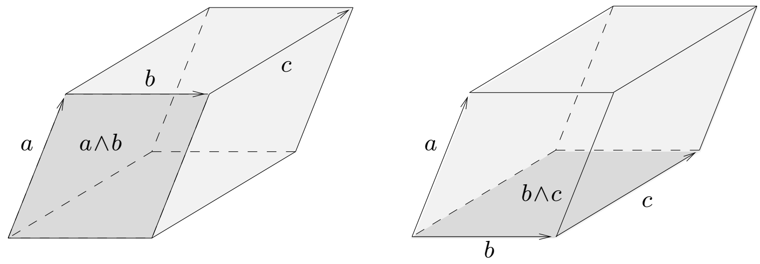

Figure 2.5 *The trivector.* The trivector $a \wedge b \wedge c$ can be viewed as the oriented parallelepiped obtained from sweeping the bivector $a \wedge b$ along the vector c. In the left-hand diagram the bivector $a \wedge b$ is swept along c. In the right-hand one $b \wedge c$ is swept along a. The result is the same in both cases, demonstrating the equality $a \wedge b \wedge c = b \wedge c \wedge a$. The associativity of the outer product is also clear from such diagrams.

antisymmetry of the exterior product. The geometric picture of this is that swapping any two vectors reverses the orientation by which the volume is swept out. Under two successive interchanges of pairs of vectors the trivector returns to itself, so

$$a \wedge b \wedge c = c \wedge a \wedge b = b \wedge c \wedge a. \tag{2.70}$$

This is also illustrated in figure 2.5.

The unit right-handed pseudoscalar for space is given the standard symbol I, so

$$I = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_3, \tag{2.71}$$

where the $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$ are any right-handed frame of orthonormal vectors. If a left-handed set of orthonormal vectors is multiplied together the result is $-I$. Given an arbitrary set of three vectors we must have

$$a \wedge b \wedge c = \alpha I, \tag{2.72}$$

where α is a scalar. It is not hard to show that $|\alpha|$ is the volume of the parallelepiped with sides defined by a, b and c. The sign of α encodes whether the set $\{a, b, c\}$ forms a right-handed or left-handed frame. In three dimensions this fully accounts for the information in the trivector.

Now consider the product of the vector e_1 and the pseudoscalar,

$$\mathsf{e}_1 I = \mathsf{e}_1(\mathsf{e}_1\mathsf{e}_2\mathsf{e}_3) = \mathsf{e}_2\mathsf{e}_3. \tag{2.73}$$

This returns a bivector — the plane perpendicular to the original vector (see figure 2.6). The product of a grade-1 vector with the grade-3 pseudoscalar is therefore a grade-2 bivector. Multiplying from the left we find that

$$I\mathsf{e}_1 = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_3\mathsf{e}_1 = -\mathsf{e}_1\mathsf{e}_2\mathsf{e}_1\mathsf{e}_3 = \mathsf{e}_2\mathsf{e}_3. \tag{2.74}$$

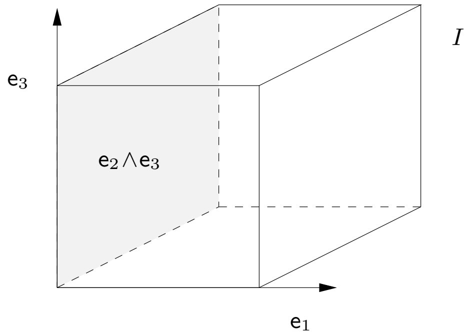

Figure 2.6 *A vector and a trivector.* The result of multiplying the vector e_1 by the trivector I is the plane $\mathsf{e}_1(\mathsf{e}_1\mathsf{e}_2\mathsf{e}_3) = \mathsf{e}_2\mathsf{e}_3$. This is the plane perpendicular to the e_1 vector.

The result is therefore independent of order, and this holds for any basis vector. It follows that the pseudoscalar commutes with all vectors in three dimensions:

$$Ia = aI. \tag{2.75}$$

This is always the case for the pseudoscalar in spaces of odd dimension. In even dimensions, the pseudoscalar anticommutes with all vectors, as we have already seen in two dimensions.

We can now express each of our basis bivectors as the product of the pseudoscalar and a *dual* vector:

$$\mathsf{e}_1\mathsf{e}_2 = I\mathsf{e}_3, \qquad \mathsf{e}_2\mathsf{e}_3 = I\mathsf{e}_1, \qquad \mathsf{e}_3\mathsf{e}_1 = I\mathsf{e}_2. \tag{2.76}$$

This operation of multiplying by the pseudoscalar is called a *duality* transformation and was originally introduced by Grassmann. Again, we can write

$$aI = a\cdot I \tag{2.77}$$

with the dot used to denote the lowest grade term in the product. The result of this can be understood as a projection — projecting onto the component of I perpendicular to a.

We next form the square of the pseudoscalar:

$$I^2 = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_3\mathsf{e}_1\mathsf{e}_2\mathsf{e}_3 = \mathsf{e}_1\mathsf{e}_2\mathsf{e}_1\mathsf{e}_2 = -1. \tag{2.78}$$

So the pseudoscalar commutes with all elements and squares to -1. It is therefore a further candidate for a unit imaginary. In some physical applications this is the correct one to use, whereas for others it is one of the bivectors. The properties of I in three dimensions make it particularly tempting to replace it with the symbol i, and this is common practice in much of the literature. This convention can still lead to confusion, however, and is not adopted in this book.

Finally, we consider the product of a bivector and the pseudoscalar:

$$I(\mathsf{e}_1 \wedge \mathsf{e}_2) = I\mathsf{e}_1\mathsf{e}_2\mathsf{e}_3\mathsf{e}_3 = II\mathsf{e}_3 = -\mathsf{e}_3. \tag{2.79}$$

So the result of the product of I with the bivector formed from e_1 and e_2 is $-\mathsf{e}_3$, that is, minus the vector perpendicular to the $\mathsf{e}_1 \wedge \mathsf{e}_2$ plane. This provides a definition of the vector cross product as

$$a \times b = -I(a \wedge b). \tag{2.80}$$

The vector cross product is largely redundant now that we have the exterior product and duality at our disposal. For example, consider the result for the double cross product. We form

$$\begin{aligned} a \times (b \times c) &= -Ia \wedge (-I(b \wedge c)) \\ &= \tfrac{1}{2} I\big(aI(b \wedge c) - (b \wedge c)Ia\big) \\ &= -a \cdot (b \wedge c). \end{aligned} \tag{2.81}$$

We have already calculated the expansion of the final line, which turns out to be the first example of a much more general, and very useful, formula.

Equation (2.80) shows how the cross product of two vectors is a disguised bivector, the bivector being mapped to a vector by a duality operation. It is now clear why the product only exists in three dimensions — this is the only space for which the dual of a bivector is a vector. We will have little further use for the cross product and will rarely employ it from now on. This means we can also do away with the awkward distinction between polar and axial vectors. Instead we just talk in terms of vectors and bivectors. Both may belong to three-dimensional linear spaces, but they are quite different objects with distinct algebraic properties.

2.4.4 The Pauli algebra

The full geometric product for vectors can be written

$$\mathsf{e}_i\mathsf{e}_j = \mathsf{e}_i \cdot \mathsf{e}_j + \mathsf{e}_i \wedge \mathsf{e}_j = \delta_{ij} + I\epsilon_{ijk}\mathsf{e}_k. \tag{2.82}$$

This may be familiar to many — it is the Pauli algebra of quantum mechanics! The Pauli matrices therefore form a matrix representation of the geometric algebra of space. The Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.83}$$

These matrices satisfy

$$\sigma_i\sigma_j = \delta_{ij}\mathsf{I} + i\epsilon_{ijk}\sigma_k, \tag{2.84}$$

where I is the 2×2 identity matrix. Historically, these matrices were discovered by Pauli in his investigations of the quantum theory of spin. The link with geometric algebra ('Clifford algebra' in the quantum theory textbooks) was only made later.

Surprisingly, though the link with the geometric algebra of space is now well established, one seldom sees the Pauli matrices referred to as a representation for the algebra of a set of vectors. Instead they are almost universally referred to as the components of a single vector in 'isospace'. A handful of authors (most notably David Hestenes) have pointed out the curious nature of this interpretation. Such discussion remains controversial, however, and will only be touched on in this book. As with all arguments over interpretations of quantum mechanics, how one views the Pauli matrices has little effect on the predictions of the theory.

The fact that the Pauli matrices form a matrix representation of $\mathcal{G}_3$ provides an alternative way of performing multivector manipulations. This method is usually slower, but can sometimes be used to advantage, particularly in programming languages where complex arithmetic is built in. Working directly with matrices does obscure geometric meaning, and is usually best avoided.

2.5 Conventions

A number of conventions help to simplify expressions in geometric algebra. For example, expressions such as $(a \cdot b)c$ and $I(a \wedge b)$ demonstrate that it would be useful to have a convention which allows us to remove the brackets. We thus introduce the operator ordering convention that in the absence of brackets, *inner and outer products are performed before geometric products.* This can remove significant numbers of unnecessary brackets. For example, we can safely write

$$I(a\wedge b) = I\, a\wedge b. \tag{2.85}$$

and

$$(a\cdot b)c = a\cdot b\, c. \tag{2.86}$$

In addition, unless brackets specify otherwise, inner products are performed before outer products,

$$a\cdot b\, c\wedge d = (a\cdot b)c\wedge d. \tag{2.87}$$

A simple notation for the result of projecting out the elements of a multivector that have a given grade is also invaluable. We denote this with angled brackets $\langle\,\rangle_r$, where r is the grade onto which we want to project. With this notation we can write, for example,

$$a\wedge b = \langle a\wedge b\rangle_2 = \langle ab\rangle_2. \tag{2.88}$$

The final expression holds because $a \wedge b$ is the sole grade-2 component of the geometric product ab. This notation can be extremely useful as it often enables inner and outer products to be replaced by geometric products, which are usually simpler to manipulate. The operation of taking the scalar part of a product is often needed, and it is conventional for this to drop the subscript zero and simply write

$$\langle M \rangle = \langle M \rangle_0. \tag{2.89}$$

The scalar part of any pair of multivectors is symmetric:

$$\langle AB \rangle = \langle BA \rangle. \tag{2.90}$$

It follows that the scalar part satisfies the cyclic reordering property

$$\langle AB \cdots C \rangle = \langle B \cdots CA \rangle, \tag{2.91}$$

which is frequently employed in manipulations.

An important operation in geometric algebra is that of *reversion*, which reverses the order of vectors in any product. There are two conventions for this in common usage. One is the dagger symbol, $A^\dagger$, used for Hermitian conjugation in matrix algebra. The other is to use a tilde, $\tilde{A}$. In three-dimensional applications the dagger symbol is often employed, as the reverse operation returns the same result as Hermitian conjugation of the Pauli matrix representation of the algebra. In spacetime physics, however, the tilde symbol is the better choice as the dagger is reserved for a different (frame-dependent) operation in relativistic quantum mechanics. For the remainder of this chapter we will use the dagger symbol, as we will concentrate on applications in three dimensions.

Scalars and vectors are invariant under reversion, but bivectors change sign:

$$(\mathsf{e}_1\mathsf{e}_2)^\dagger = \mathsf{e}_2\mathsf{e}_1 = -\mathsf{e}_1\mathsf{e}_2. \tag{2.92}$$

Similarly, we see that

$$I^\dagger = \mathsf{e}_3\mathsf{e}_2\mathsf{e}_1 = \mathsf{e}_1\mathsf{e}_3\mathsf{e}_2 = -\mathsf{e}_1\mathsf{e}_2\mathsf{e}_3 = -I. \tag{2.93}$$

A general multivector in $\mathcal{G}_3$ can be written

$$M = \alpha + a + B + \beta I, \tag{2.94}$$

where a is a vector, B is a bivector and α and β are scalars. From the above we see that the reverse of M, $M^\dagger$, is

$$M^\dagger = \alpha + a - B - \beta I. \tag{2.95}$$

As stated above, this operation has the same effect as Hermitian conjugation applied to the Pauli matrices.

We have now introduced a number of terms, some of which have overlapping meaning. It is useful at this point to refer to multivectors which only contain terms of a single grade as *homogeneous*. The term *inner* product is reserved for the lowest grade part of the geometric product of two homogeneous multivectors. For two homogeneous multivectors of the same grade the inner product and scalar product reduce to the same thing. The terms *exterior* and *outer* products are interchangeable, though we will tend to prefer the latter for its symmetry with the inner product. The inner and outer products are also referred to colloquially as the *dot* and *wedge* products. We have followed convention in referring to the highest grade element in a geometric algebra as the *pseudoscalar*. This is a convenient name, though one must be wary that in tensor analysis the term can mean something subtly different. Both *directed volume element* and *volume form* are good alternative names, but we will stick with *pseudoscalar* in this book.

2.6 Reflections

The full power of geometric algebra begins to emerge when we consider reflections and rotations. We start with an arbitrary vector a and a unit vector n ($n^2 = 1$), and resolve a into parts parallel and perpendicular to n. This is achieved simply by forming

$$\begin{aligned} a &= n^2 a \\ &= n(n\cdot a + n\wedge a) \\ &= a_{\parallel} + a_{\perp}, \end{aligned} \tag{2.96}$$

where

$$a_{\parallel} = a\cdot n\, n, \qquad a_{\perp} = n\, n\wedge a. \tag{2.97}$$

The formula for $a_{\parallel}$ is certainly the projection of a onto n, and the remaining term must be the perpendicular component (sometimes called the rejection). We can check that $a_{\perp}$ is perpendicular to n quite simply:

$$n\cdot a_{\perp} = \langle nn\, n\wedge a\rangle = \langle n\wedge a\rangle = 0. \tag{2.98}$$

This is a simple example of how using the projection onto grade operator to replace inner and outer products with geometric products can simplify derivations.

The result of reflecting a in the plane orthogonal to n is the vector $a' = a_{\perp} - a_{\parallel}$ (see figure 2.7). This can be written

$$\begin{aligned} a' = a_{\perp} - a_{\parallel} &= n\, n\wedge a - a\cdot n\, n \\ &= -n\cdot a\, n - n\wedge a\, n \\ &= -nan. \end{aligned} \tag{2.99}$$

This formula is already more compact than can be written down without the geometric product. The best one can do with just the inner product is the

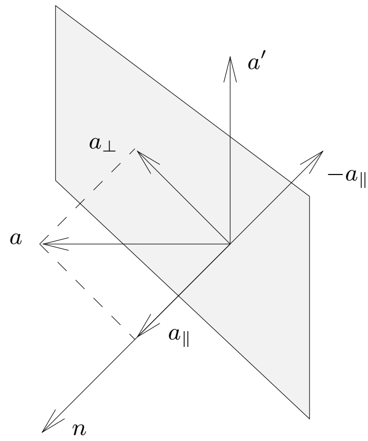

Figure 2.7 *A reflection.* The vector a is reflected in the (hyper)plane perpendicular to n. This is the way to describe reflections in arbitrary dimensions. The result a' is formed by reversing the sign of $a_\|$, the component of a in the n direction.

equivalent expression

$$a' = a - 2a\cdot n\, n. \tag{2.100}$$

The compression afforded by the geometric product becomes increasingly impressive as reflections are compounded together. The formula

$$a' = -nan \tag{2.101}$$

is valid is spaces of any dimension — it is a quite general formula for a reflection.

We should check that our formula for the reflection has the desired property of leaving lengths and angles unchanged. To do this we need only verify that the scalar product between vectors is unchanged if both are reflected, which is achieved with a simple rearrangement:

$$(-nan)\cdot(-nbn) = \langle(-nan)(-nbn)\rangle = \langle nabn\rangle = \langle abnn\rangle = a\cdot b. \tag{2.102}$$

In this manipulation we have made use of the cyclic reordering property of the scalar part of a geometric product, as defined in equation (2.91).

2.6.1 Complex conjugation

In two dimensions we saw that the vector x is mapped to a complex number Z by

$$Z = \mathsf{e}_1 x, \quad x = \mathsf{e}_1 Z. \tag{2.103}$$

The complex conjugate $Z^\dagger$ is the reverse of this, $Z^\dagger = x\mathsf{e}_1$, so maps to the vector

$$x' = \mathsf{e}_1 Z^\dagger = \mathsf{e}_1 x \mathsf{e}_1. \tag{2.104}$$

This can be converted into the formula for a reflection if we remember that the two-dimensional pseudoscalar $I = \mathsf{e}_1\mathsf{e}_2$ anticommutes with all vectors and squares to -1. We therefore have

$$x' = -\mathsf{e}_1 I I x \mathsf{e}_1 = -\mathsf{e}_1 I x \mathsf{e}_1 I = -\mathsf{e}_2 x \mathsf{e}_2. \tag{2.105}$$

This is precisely the expected relation for a reflection in the line perpendicular to e_2, which is to say a reflection in the real axis.

2.6.2 Reflecting bivectors

Now suppose that we form the bivector $B = a \wedge b$ and reflect both of these vectors in the plane perpendicular to n. The result is

$$B' = (-nan)\wedge(-nbn). \tag{2.106}$$

This simplifies as follows:

$$\begin{aligned}(-nan)\wedge(-nbn) &= \tfrac{1}{2}(nannbn - nbnnan)\\ &= \tfrac{1}{2}n(ab - ba)n\\ &= nBn.\end{aligned} \tag{2.107}$$

The effect of sandwiching a multivector between a vector, nMn, always preserves the grade of the multivector M. We will see how to prove this in general when we have derived a few more results for manipulating inner and outer products. The resulting formula nBn shows that bivectors are subject to the same transformation law as vectors, *except for a change in sign.* This is the origin of the conventional distinction between polar and axial vectors. Axial vectors are usually generated by the cross product, and we saw in section 2.4.3 that the cross product generates a bivector, and then dualises it back to a vector. But when the two vectors in the cross product are reflected, the bivector they form is reflected according to (2.107). The dual vector IB is subject to the same transformation law, since

$$I(nBn) = n(IB)n, \tag{2.108}$$

and so does not transform as a (polar) vector. In many texts this can be a source of much confusion. But now we have a much healthier alternative: banish all talk of axial vectors in favour of bivectors. We will see in later chapters that all of the main examples of 'axial' vectors in physics (angular velocity, angular momentum, the magnetic field etc.) are better viewed as bivectors.

2.6.3 Trivectors and handedness

The final object to try reflecting in three dimensions is the trivector $a \wedge b \wedge c$. We first write

$$(-nan)\wedge(-nbn)\wedge(-ncn) = \langle(-nan)(-nbn)(-ncn)\rangle_3 \\ = -\langle nabcn\rangle_3, \tag{2.109}$$

which follows because the only way to form a trivector from the geometric product of three vectors is through the exterior product of all three. Now the product abc can only contain a vector and trivector term. The former cannot give rise to an overall trivector, so we are left with

$$(-nan)\wedge(-nbn)\wedge(-ncn) = -\langle na\wedge b\wedge cn\rangle_3. \tag{2.110}$$

But any trivector in three dimensions is a multiple of the pseudoscalar I, which commutes with all vectors, so we are left with

$$(-nan)\wedge(-nbn)\wedge(-ncn) = -a\wedge b\wedge c. \tag{2.111}$$

The overall effect is simply to flip the sign of the trivector, which is a way of stating that reflections have determinant -1. This means that if all three vectors in a right-handed triplet are reflected in some plane, the resulting triplet is left handed (and vice versa).

2.7 Rotations

Our starting point for the treatment of rotations is the result that *a rotation in the plane generated by two unit vectors m and n is achieved by successive reflections in the (hyper)planes perpendicular to m and n.* This is illustrated in figure 2.8. Any component of a perpendicular to the $m\wedge n$ plane is unaffected, and simple trigonometry confirms that the angle between the initial vector a and the final vector c is twice the angle between m and n. (The proof of this is left as an exercise.) The result of the successive reflections is therefore to rotate through 2θ in the $m\wedge n$ plane, where $m\cdot n = \cos(\theta)$.

So how does this look using geometric algebra? We first form

$$b = -mam \tag{2.112}$$

and then perform a second reflection to obtain

$$c = -nbn = -n(-mam)n = nmamn. \tag{2.113}$$

This is starting to look extremely simple! We define

$$R = nm, \tag{2.114}$$

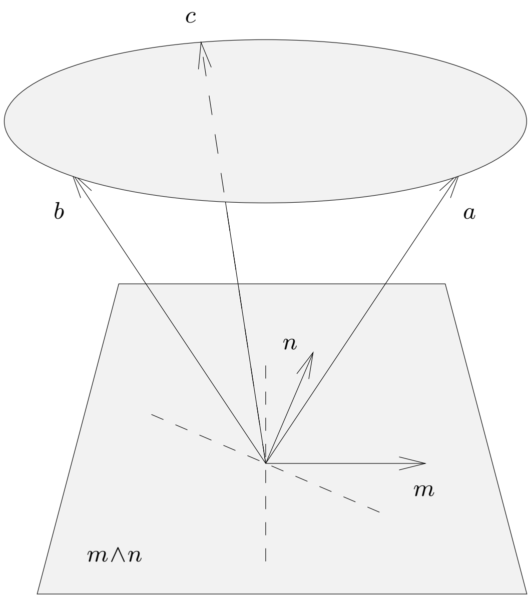

Figure 2.8 *A rotation from two reflections.* The vector b is the result of reflecting a in the plane perpendicular to m, and c is the result of reflecting b in the plane perpendicular to n.

so that we can now write the result of the rotation as

$$c = RaR^{\dagger}. \tag{2.115}$$

This transformation $a \mapsto RaR^{\dagger}$ is a totally general way of handling rotations. In deriving this transformation the dimensionality of the space of vectors was never specified, so the transformation law must work in all spaces, *whatever their dimension*. The rule also works for *any grade* of multivector!

2.7.1 Rotors

The quantity $R = nm$ is called a *rotor* and is one of the most important objects in applications of geometric algebra. Immediately, one can see the importance of the *geometric* product in both (2.114) and (2.115), which tells us that rotors provide a way of handling rotations that is unique to geometric algebra. To study the properties of the rotor R we first write

$$R = nm = n\cdot m + n\wedge m = \cos(\theta) + n\wedge m. \tag{2.116}$$

We already calculated the magnitude of the bivector $m \wedge n$ in equation (2.15), where we obtained

$$(n\wedge m)(n\wedge m) = -\sin^2(\theta). \tag{2.117}$$

We therefore define the *unit* bivector B in the $m\wedge n$ plane by

$$B = \frac{m\wedge n}{\sin(\theta)}, \qquad B^2 = -1. \tag{2.118}$$

The reason for this choice of orientation ($m \wedge n$ rather than $n \wedge m$) is to ensure that the rotation has the orientation specified by the generating bivector, as can be seen in figure 2.8. In terms of the bivector B we now have

$$R = \cos(\theta) - B\sin(\theta), \tag{2.119}$$

which is simply the polar decomposition of a complex number, with the unit imaginary replaced by the unit bivector B. We can therefore write

$$R = \exp(-B\theta), \tag{2.120}$$

with the exponential defined in terms of its power series in the normal way. (The power series for the exponential is absolutely convergent for any multivector argument.)

Now recall that our formula was for a rotation through 2θ. If we want to rotate through θ, the appropriate rotor is

$$R = \exp(-B\theta/2), \tag{2.121}$$

which gives the formula

$$a \mapsto a' = \mathrm{e}^{-B\theta/2} a \mathrm{e}^{B\theta/2} \tag{2.122}$$

for a rotation through θ in the B plane, with handedness determined by B (see figure 2.9). This description encourages us to think of rotations taking place *in a plane*, and as such gives equations which are valid in any dimension. The more traditional idea of rotations taking place around an axis is an entirely three-dimensional concept which does not generalise.

Since the rotor R is a geometric product of two unit vectors, we see immediately that

$$RR^\dagger = nm(nm)^\dagger = nmmn = 1 = R^\dagger R. \tag{2.123}$$

This provides a quick proof that our formula has the correct property of preserving lengths and angles. Suppose that $a' = RaR^\dagger$ and $b' = RbR^\dagger$, then

$$\begin{aligned} a'\cdot b' &= \tfrac{1}{2}(RaR^\dagger RbR^\dagger + RbR^\dagger RaR^\dagger) \\ &= \tfrac{1}{2}R(ab+ba)R^\dagger \\ &= a\cdot b\, RR^\dagger \\ &= a\cdot b. \end{aligned} \tag{2.124}$$

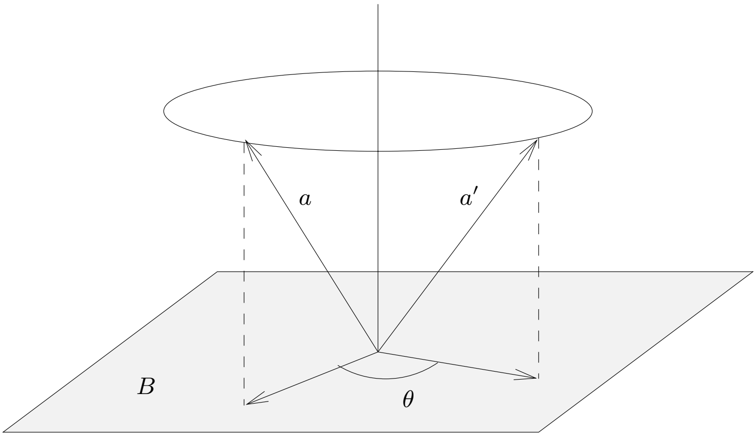

Figure 2.9 *A rotation in three dimensions.* The vector a is rotated to $a' = RaR^\dagger$. The rotor R is defined by $R = \exp(-B\theta/2)$, which describes the rotation directly in terms of the plane and angle. The rotation has the orientation specified by the bivector B.

We can also see that the inverse transformation is given by

$$a = R^\dagger a' R. \tag{2.125}$$

The proof is straightforward:

$$R^\dagger a' R = R^\dagger R a R^\dagger R = a. \tag{2.126}$$

The usefulness of rotors provides ample justification for adding up terms of different grades. The rotor R on its own has no geometric significance, which is to say that no meaning should be attached to the separate scalar and bivector terms. When R is written in the form $R = \exp(-B\theta/2)$, however, the bivector B has clear geometric significance, as does the vector formed from $RaR^\dagger$. This illustrates a central feature of geometric algebra, which is that both geometrically meaningful objects (vectors, planes etc.) and the elements that act on them (in this case rotors) are represented in the same algebra.

2.7.2 Constructing a rotor

Suppose that we wish to rotate the unit vector a into another unit vector b, leaving all vectors perpendicular to a and b unchanged. This is accomplished by a reflection perpendicular to the unit vector n half-way between a and b followed by a reflection in the plane perpendicular to b (see figure 2.10). The vector n is

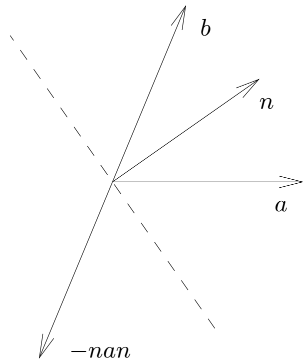

Figure 2.10 *A rotation from a to b.* The vector a is rotated onto b by first reflecting in the plane perpendicular to n, and then in the plane perpendicular to b. The vectors a, b and n all have unit length.

given by

$$n = \frac{(a+b)}{|a+b|}, \tag{2.127}$$

which reflects a into $-b$. Combining this with the reflection in the plane perpendicular to b we arrive at the rotor

$$R = bn = \frac{1+ba}{|a+b|} = \frac{1+ba}{\sqrt{2(1+b\cdot a)}}, \tag{2.128}$$

which represents a simple rotation in the $a\wedge b$ plane. This formula shows us that

$$Ra = \frac{a+b}{\sqrt{2(1+b\cdot a)}} = a\frac{1+ab}{\sqrt{2(1+b\cdot a)}} = aR^\dagger. \tag{2.129}$$

It follows that we can write

$$RaR^\dagger = R^2 a = aR^{\dagger\,2}. \tag{2.130}$$

This is always possible for vectors in the plane of rotation. Returning to the polar form $R = \exp(-B\theta/2)$, where B is the $a \wedge b$ plane, we see that

$$R^2 = \exp(-B\theta), \tag{2.131}$$

so we can rotate a onto b with the formula

$$b = \mathrm{e}^{-B\theta} a = a\mathrm{e}^{B\theta}. \tag{2.132}$$

This is precisely the form found in the plane using complex numbers, and was the source of much of the confusion over the use of quaternions for rotations. Hamilton thought that a single-sided transformation law of the form $a \mapsto Ra$ should be the correct way to encode a rotation, with the full angle appearing in the exponential. He thought that this was the natural generalisation of the complex number representation. But we can see now that this formula only works for vectors *in the plane of rotation.* The correct formula for all vectors is the double-sided, half-angle formula $a \mapsto RaR^\dagger$. This formula ensures that given a vector c perpendicular to the $a \wedge b$ plane we have

$$Rc = c\frac{1+ba}{\sqrt{2(1+b\cdot a)}} = \frac{1+ba}{\sqrt{2(1+b\cdot a)}}c = Rc, \tag{2.133}$$

so that

$$RcR^\dagger = cRR^\dagger = c, \tag{2.134}$$

and the vector is unrotated. The single-sided law does not have this property. Correctly identifying the double-sided transformation law means that unit bivectors such as

$$\mathsf{e}_1\mathsf{e}_2 = \mathrm{e}^{\mathsf{e}_1\mathsf{e}_2\pi/2} \tag{2.135}$$

are generators of rotations through π, and not $\pi/2$. The fact that unit bivectors square to -1 is consistent with this because, acting double sidedly, the rotor -1 is the identity operation. More generally, R and $-R$ generate the same rotation, so there is a two-to-one map between rotors and rotations. (Mathematicians talk of the rotors providing a *double-cover* representation of the rotation group.)

2.7.3 Rotating multivectors

Suppose that the two vectors forming the bivector $B = a\wedge b$ are both rotated. What is the expression for the resulting bivector? To find this we form

$$\begin{aligned} B' = a'\wedge b' &= \tfrac{1}{2}(RaR^\dagger RbR^\dagger - RbR^\dagger RaR^\dagger) \\ &= \tfrac{1}{2}R(ab-ba)R^\dagger \\ &= Ra\wedge bR^\dagger \\ &= RBR^\dagger, \end{aligned} \tag{2.136}$$

where we have used the rotor normalisation formula $R^\dagger R = 1$. Bivectors are rotated using precisely the same formula as vectors! The same turns out to be true for all geometric multivectors, and this is one of the most attractive features of geometric algebra. In section 4.2 we prove that the transformation $A \mapsto RAR^\dagger$ preserves the grade of the multivector on which the rotors act. For applications in three dimensions we only need check this result for the trivector case, as we have already demonstrated it for vectors and bivectors. The pseudoscalar in three dimensions, I, commutes with all other terms in the algebra, so we have

$$RIR^\dagger = IRR^\dagger = I, \tag{2.137}$$

which is certainly grade-preserving. This result is one way of saying that rotations have determinant +1. We now have a means of rotating all geometric objects in three dimensions. In chapter 3 we will take full advantage of this when studying rigid-body dynamics.

2.7.4 Rotor composition law

Having seen how individual rotors are used to represent rotations, we now look at their composition law. Let the rotor R_1 transform the vector a into a vector b:

$$b = R_1 a R_1^\dagger. \tag{2.138}$$

Now rotate b into another vector c, using a rotor R_2. This requires

$$c = R_2 b R_2^\dagger = R_2 R_1 a R_1^\dagger R_2^\dagger = R_2 R_1 a (R_2 R_1)^\dagger, \tag{2.139}$$

so that if we write

$$c = R a R^\dagger, \tag{2.140}$$

then the composite rotor is given by

$$R = R_2 R_1. \tag{2.141}$$

This is the *group combination rule* for rotors. Rotors form a group because the product of two rotors is a third rotor, as can be checked from

$$R_2 R_1 (R_2 R_1)^\dagger = R_2 R_1 R_1^\dagger R_2^\dagger = R_2 R_2^\dagger = 1. \tag{2.142}$$

In three dimensions the fact that the multivector R contains only even-grade elements and satisfies $RR^\dagger = 1$ is sufficient to ensure that R is a rotor. The fact that rotors form a continuous group (called a *Lie group*) is a subject we will return to later in this book.

Rotors are the exception to the rule that all multivectors are subject to a double-sided transformation law. Rotors are already mixed-grade objects, so multiplying on the left (or right) by another rotor does not take us out of the space of rotors. All geometric entities, such as lines and planes, are single-grade objects, and their grades cannot be changed by a rotation. They are therefore all subject to a double-sided transformation law. Again, this brings us back to the central theme that both geometric objects and the operators acting on them are contained in a single algebra.

The composition rule (2.141) has a surprising consequence. Suppose that the rotor R_1 is kept fixed, and we set $R_2 = \exp(-B\theta/2)$. We now take the vector c on a 2π excursion back to itself. The final rotor R is

$$R = \mathrm{e}^{-B\pi} R_1 = -R_1. \tag{2.143}$$

The rotor has changed sign under a 2π rotation! This is usually viewed as a quantum-mechanical phenomenon related to the existence of fermions. But we can now see that the result is classical and is simply a consequence of our rotor description of rotations. (The relationship between rotors and fermion wavefunctions is discussed in chapter 8.) A geometric interpretation of the distinction between R and $-R$ is provided by the *direction* in which a rotation is performed. Suppose we want to rotate e_1 onto e_2. The rotor to achieve this is

$$R(\theta) = \mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\theta/2}. \tag{2.144}$$

If we rotate in a positive sense through $\pi/2$ the final rotor is given by

$$R(\pi/2) = \frac{1}{\sqrt{2}}(1 - \mathsf{e}_1\mathsf{e}_2). \tag{2.145}$$

If we rotate in the negative (clockwise) sense, however, the final rotor is

$$R(-3\pi/2) = -\frac{1}{\sqrt{2}}(1 - \mathsf{e}_1\mathsf{e}_2) = -R(\pi/2). \tag{2.146}$$

So, while R and $-R$ define the same absolute rotation (and the same rotation matrix), their different signs can be employed to record information about the handedness of the rotation.

The rotor composition rule provides a simple formula for the compound effect of two rotations. Suppose that we have

$$R_1 = \mathrm{e}^{-B_1\theta_1/2}, \quad R_2 = \mathrm{e}^{-B_2\theta_2/2}, \tag{2.147}$$

where both B_1 and B_2 are unit bivectors. The product rotor is

$$\begin{aligned} R =& \big(\cos(\theta_2/2) - \sin(\theta_2/2)B_2\big)\big(\cos(\theta_1/2) - \sin(\theta_1/2)B_1\big) \\ =& \cos(\theta_2/2)\cos(\theta_1/2) - \big(\cos(\theta_2/2)\sin(\theta_1/2)B_1 + \cos(\theta_1/2)\sin(\theta_2/2)B_2\big) \\ &+ \sin(\theta_2/2)\sin(\theta_1/2)B_1B_2. \end{aligned} \tag{2.148}$$

So if we write $R = R_2R_1 = \exp(-B\theta/2)$, where B is a new unit bivector, we immediately see that

$$\cos(\theta/2) = \cos(\theta_2/2)\cos(\theta_1/2) + \sin(\theta_2/2)\sin(\theta_1/2)\langle B_1B_2\rangle \tag{2.149}$$

and

$$\begin{aligned} \sin(\theta/2)B = \cos(\theta_2/2)\sin(\theta_1/2)B_1 &+ \cos(\theta_1/2)\sin(\theta_2/2)B_2 \\ &- \sin(\theta_2/2)\sin(\theta_1/2)\langle B_1B_2\rangle_2. \end{aligned} \tag{2.150}$$

These half-angle relations for rotations were first discovered by the mathematician Rodriguez, three years before the invention of the quaternions! It is well known that these provide a simple means of calculating the compound effect of two rotations. Numerically, it is usually even simpler to just multiply the rotors directly and not worry about calculating any trigonometric functions.

2.7.5 Euler angles

A standard way to parameterise rotations is via the three Euler angles $\{\phi, \theta, \psi\}$. These are defined to rotate an initial set of axes, $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$, onto a new set $\{\mathsf{e}_1', \mathsf{e}_2', \mathsf{e}_3'\}$ (often denoted x, y, z and x', y', z' respectively). First we rotate about the e_3 axis — i.e. in the $\mathsf{e}_1\mathsf{e}_2$ plane — anticlockwise through an angle ϕ. The rotor for this is

$$R_\phi = \mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\phi/2}. \tag{2.151}$$

Next we rotate about the axis formed by the transformed e_1 axis through an amount θ. The plane for this is

$$IR_\phi \mathsf{e}_1 R_\phi^\dagger = R_\phi \mathsf{e}_2\mathsf{e}_3 R_\phi^\dagger. \tag{2.152}$$

The rotor is therefore

$$R_\theta = \exp\bigl(-R_\phi \mathsf{e}_2\mathsf{e}_3 R_\phi^\dagger \theta/2\bigr) = R_\phi \mathrm{e}^{-\mathsf{e}_2\mathsf{e}_3\theta/2} R_\phi^\dagger. \tag{2.153}$$

The intermediate rotor is now

$$R' = R_\theta R_\phi = \mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\phi/2}\mathrm{e}^{-\mathsf{e}_2\mathsf{e}_3\theta/2}. \tag{2.154}$$

Note the order! Finally, we rotate about the transformed e_3 axis through an angle ψ. The appropriate plane is now

$$IR'\mathsf{e}_3 R'^\dagger = R'\mathsf{e}_1\mathsf{e}_2 R'^\dagger \tag{2.155}$$

and the rotor is

$$R_\psi = \exp\bigl(-R'\mathsf{e}_1\mathsf{e}_2 R'^\dagger \psi/2\bigr) = R'\mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\psi/2} R'^\dagger. \tag{2.156}$$

The resultant rotor is therefore

$$R = R_\psi R' = \mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\phi/2}\mathrm{e}^{-\mathsf{e}_2\mathsf{e}_3\theta/2}\mathrm{e}^{-\mathsf{e}_1\mathsf{e}_2\psi/2}, \tag{2.157}$$

which has decoupled very nicely and is really quite simple — it is much easier to visualise and work with than the equivalent matrix formula! Now that we have geometric algebra at our disposal we will, in fact, have little cause to use the Euler angles in calculations.

2.8 Notes

In this chapter we have given a lengthy introduction to geometric algebra in two and three dimensions. The latter algebra is generated entirely by three basis vectors $\{\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3\}$ subject to the rule that $\mathsf{e}_i\mathsf{e}_j + \mathsf{e}_j\mathsf{e}_i = 2\delta_{ij}$. This simple rule generates an algebra of remarkable power and richness which we will explore in following chapters.

There is a large literature on the geometric algebra of three-dimensional space and its applications in physics. The most complete text is *New Foundations* *for Classical Mechanics* by David Hestenes (1999). Hestenes has also written many papers on the subject, most of which are listed in the bibliography at the end of this book. Other introductory papers have been written by Gull, Lasenby and Doran (1993a), Doran et al. (1996a) and Vold (1993a, 1993b). Clifford's *Mathematical Papers* (1882) are also of considerable interest. The use of geometric algebra for handling rotations is very common in the fields of engineering and computer science, though often purely in the guise of the quaternion algebra. Searching one of the standard scientific databases with the keyword 'quaternions' returns too many papers to begin to list here.

2.9 Exercises

2.1 From the properties of the geometric product, show that the symmetrised product of two vectors satisfies the properties of a scalar product, as listed in section 1.2.

2.2 By expanding the bivector $a \wedge b$ in terms of geometric products, prove that it anticommutes with both a and b, but commutes with any vector perpendicular to the $a \wedge b$ plane.

2.3 Verify that the $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ matrices of equation (2.27) satisfy the correct multiplication relations to form a representation of $\mathcal{G}_2$. Use these to verify equations (2.26).

2.4 Construct the multiplication table generated by the orthonormal vectors e_1, e_2 and e_3. Do these generate a (finite) group?

2.5 Prove that all of the following forms are equivalent expressions of the vector cross product:

$$a \times b = -I a \wedge b = b \cdot (Ia) = -a \cdot (Ib).$$

Interpret each form geometrically. Hence establish that

$$a \times (b \times c) = -a \cdot (b \wedge c) = -(a \cdot b\, c - a \cdot c\, b)$$

and

$$a \cdot (b \times c) = [a, b, c] = a \wedge b \wedge c\, I^{-1}.$$

2.6 Prove that the effect of successive reflections in the planes perpendicular to the vectors m and n results in a rotation through twice the angle between m and n.

2.7 What is the reverse of $RaR^{\dagger}$, where a is a vector? Which objects in three dimensions have this property, and why must the result be another vector?

2.8 Show that the rotor

$$R = \frac{1 + ba}{|a + b|}$$

can also be written as $\exp(-B\theta/2)$, where B is the unit bivector in the $a \wedge b$ plane and θ is the angle between a and b.

2.9 The Cayley–Klein parameters are a set of four real numbers α, β, γ and δ subject to the normalisation condition

$$\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1.$$

These can be used to paramaterise an arbitrary rotation matrix as follows:

$$\boldsymbol{U} = \begin{pmatrix} \alpha^2 + \beta^2 - \gamma^2 - \delta^2 & 2(\beta\gamma + \alpha\delta) & 2(\beta\delta - \alpha\gamma) \\ 2(\beta\gamma - \alpha\delta) & \alpha^2 - \beta^2 + \gamma^2 - \delta^2 & 2(\gamma\delta + \alpha\beta) \\ 2(\beta\delta + \alpha\gamma) & 2(\gamma\delta - \alpha\beta) & \alpha^2 - \beta^2 - \gamma^2 + \delta^2 \end{pmatrix}.$$

Can you relate the Cayley–Klein parameters to the rotor description?

2.10 Show that the set of all rotors forms a continuous group. Can you identify the group manifold?

3

Classical mechanics

In this chapter we study the use of geometric algebra in classical mechanics. We will assume that readers already have a basic understanding of the subject, as a complete presentation of classical mechanics with geometric algebra would require an entire book. Such a book has been written, *New Foundations for Classical Mechanics* by David Hestenes (1999), which looks in detail at many of the topics discussed here. Our main focus in this chapter is to areas where geometric algebra offers some immediate benefits over traditional methods. These include motion in a central force and rigid-body rotations, both of which are dealt with in some detail. More advanced topics in Lagrangian and Hamiltonian dynamics are covered in chapter 12, and relativistic dynamics is covered in chapter 5.

Classical mechanics was one of the areas of physics that prompted the development of many of the mathematical techniques routinely used today. This is particularly true of vector analysis, and it is now common to see classical mechanics described using an abstract vector notation. Many of the formulae in this chapter should be completely familiar from such treatments. A key difference comes in adopting the outer product of vectors in place of the cross product. This means, for example, that angular momentum and torque both become bivectors. The outer product is clearer conceptually, but on its own it does not bring any calculational advantages. The main new computational tool we have at our disposal is the geometric product, and here we highlight a number of examples of its use.

In this chapter we have chosen to write all vectors in a bold font. This is conventional for three-dimensional physics and many of the formulae presented below look unnatural if this notation is not followed. Bivectors and other general multivectors are left in regular font, which helps to distinguish them from vectors.

3.1 Elementary principles

We start by considering a point particle with a trajectory $\boldsymbol{x}(t)$ described as a function of time. Here $\boldsymbol{x}$ is the position vector relative to some origin and the time t is taken as some absolute 'Newtonian' standard on which all observers agree. The particle has velocity

$$\boldsymbol{v} = \dot{\boldsymbol{x}} = \frac{d\boldsymbol{x}}{dt}, \tag{3.1}$$

where the overdot denotes differentiation with respect to time t. If the particle has mass m, then the momentum $\boldsymbol{p}$ is defined by $\boldsymbol{p} = m\boldsymbol{v}$. Newton's second law of motion states that

$$\dot{\boldsymbol{p}} = \boldsymbol{f}, \tag{3.2}$$

where the vector $\boldsymbol{f}$ is the force acting on the particle. Usually the mass m is constant and we recover the familiar expression $\boldsymbol{f} = m\boldsymbol{a}$, where $\boldsymbol{a}$ is the acceleration

$$\boldsymbol{a} = \frac{d^2\boldsymbol{x}}{dt^2}. \tag{3.3}$$

The case of constant mass is assumed throughout this chapter. The path for a single particle is then determined by a second-order differential equation (assuming $\boldsymbol{f}$ does not depend on higher derivatives).

The work done by the force $\boldsymbol{f}$ on a particle is defined by the line integral

$$W_{12} = \int_{t_1}^{t_2} \boldsymbol{f}\cdot\boldsymbol{v}\,dt = \int_1^2 \boldsymbol{f}\cdot d\boldsymbol{s}. \tag{3.4}$$

The final form here illustrates that the integral is independent of how the path is parameterised. From Newton's second law we have

$$W_{12} = m\int_{t_1}^{t_2} \dot{\boldsymbol{v}}\cdot\boldsymbol{v}\,dt = \frac{m}{2}\int_{t_1}^{t_2} \frac{d}{dt}(v^2)\,dt, \tag{3.5}$$

where $v = |\boldsymbol{v}| = \sqrt{(\boldsymbol{v}^2)}$. It follows that the work done is equal to the change in kinetic energy T, where

$$T = \tfrac{1}{2}mv^2. \tag{3.6}$$

In the case where the work is independent of the path from point 1 to point 2 the force is said to be *conservative*, and can be written as the gradient of a potential:

$$\boldsymbol{f} = -\boldsymbol{\nabla} V. \tag{3.7}$$

For conservative forces the work also evaluates to

$$W_{12} = -\int_1^2 d\boldsymbol{s}\cdot\boldsymbol{\nabla} V = V_1 - V_2 \tag{3.8}$$

and the total energy $E = T + V$ is conserved.

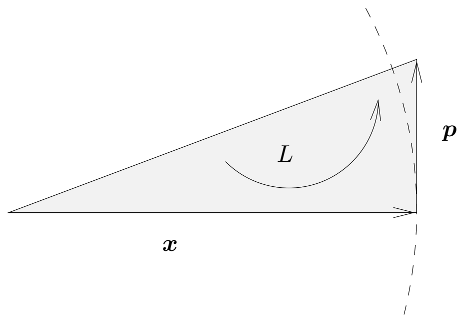

Figure 3.1 *Angular momentum.* The particle sweeps out the plane $L = \boldsymbol{x} \wedge \boldsymbol{p}$. The angular momentum should be directly related to the area swept out (cf. Kepler's second law), so is naturally encoded as a bivector. The position vector $\boldsymbol{x}$ depends on the choice of origin.

3.1.1 Angular momentum

Angular momentum is traditionally discussed in terms of the cross product, even though it is quite clear that what is required is a way of encoding the area swept out by a particle as it moves relative to some origin (see figure 3.1). We saw in chapter 2 that the exterior product provides this, and that the more traditional cross product is a derived concept based on the three-dimensional result that every directed plane has a unique normal. We therefore have no hesitation in dispensing with the traditional definition of angular momentum as an axial vector, and replace it with a bivector. So, if a particle has momentum $\boldsymbol{p}$ and position vector $\boldsymbol{x}$ from some origin, we define the angular momentum of the particle about the origin as the bivector

$$L = \boldsymbol{x} \wedge \boldsymbol{p}. \tag{3.9}$$

This definition does not alter the steps involved in computing L since the components are the same as those of the cross product. We will see, however, that the freedom we have to now use the geometric product can speed up derivations. The definition of angular momentum as a bivector maintains a clear distinction with vector quantities such as position and velocity, removing the need for the rather awkward definitions of polar and axial vectors. The definition of L as a bivector also fits neatly with the rotor description of rotations, as we shall see later in this chapter.

If we differentiate L we obtain

$$\frac{dL}{dt} = \boldsymbol{v} \wedge (m\boldsymbol{v}) + \boldsymbol{x} \wedge (m\boldsymbol{a}) = \boldsymbol{x} \wedge \boldsymbol{f}. \tag{3.10}$$

We define the torque N about the origin as the *bivector*

$$N = \boldsymbol{x}\wedge\boldsymbol{f}, \tag{3.11}$$

so that the torque and angular momentum are related by

$$\frac{dL}{dt} = N. \tag{3.12}$$

The idea of the torque being a bivector is also natural as torques act over a plane. The plane in question is defined by the vector $\boldsymbol{f}$ and the chosen origin, so both L and N depend on the origin. Recall also that bivectors are additive, much like vectors, so the result of applying two torques is found by adding the respective bivectors.

The angular momentum bivector can be written in an alternative way by first defining $r = |\boldsymbol{x}|$ and writing

$$\boldsymbol{x} = r\hat{\boldsymbol{x}}. \tag{3.13}$$

We therefore have

$$\dot{\boldsymbol{x}} = \frac{d}{dt}(r\hat{\boldsymbol{x}}) = \dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}, \tag{3.14}$$

so that

$$L = m\boldsymbol{x}\wedge(\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}) = mr\hat{\boldsymbol{x}}\wedge(\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}}) = mr^2\,\hat{\boldsymbol{x}}\wedge\dot{\hat{\boldsymbol{x}}}. \tag{3.15}$$

But since $\hat{\boldsymbol{x}}^2 = 1$ we must have

$$0 = \frac{d}{dt}\hat{\boldsymbol{x}}^2 = 2\hat{\boldsymbol{x}}\cdot\dot{\hat{\boldsymbol{x}}}. \tag{3.16}$$

We can therefore eliminate the outer product in equation (3.15) and write

$$L = mr^2\hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}} = -mr^2\dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}}, \tag{3.17}$$

which is useful in a number of problems.

3.1.2 Systems of particles

The preceding definitions generalise easily to systems of particles. For these it is convenient to distinguish between internal and external forces, so the force on the ith particles is

$$\sum_j \boldsymbol{f}_{ji} + \boldsymbol{f}_i^e = \dot{\boldsymbol{p}}_i. \tag{3.18}$$

Here $\boldsymbol{f}_i^e$ is the external force and $\boldsymbol{f}_{ij}$ is the force on the jth particle due to the ith particle. We assume that $\boldsymbol{f}_{ii} = 0$. Newton's third law (in its weak form) states that

$$\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ji}. \tag{3.19}$$

This is not obeyed by all forces, but is assumed to hold for the forces considered in this chapter. Summing the force equation over all particles we find that

$$\sum_i m_i \boldsymbol{a}_i = \sum_i \boldsymbol{f}_i^e + \sum_{i,j} \boldsymbol{f}_{ij} = \sum_i \boldsymbol{f}_i^e. \tag{3.20}$$

All of the internal forces cancel as a consequence of the third law. We define the centre of mass $\boldsymbol{X}$ by

$$\boldsymbol{X} = \frac{1}{M} \sum_i m_i \boldsymbol{x}_i, \tag{3.21}$$

where M is the total mass

$$M = \sum_i m_i. \tag{3.22}$$

The position of the centre of mass is governed by the force law

$$M \frac{d^2 \boldsymbol{X}}{dt^2} = \sum_i \boldsymbol{f}_i^e = \boldsymbol{f}^e \tag{3.23}$$

and so only responds to the total external force on the system. The total momentum of the system is defined by

$$\boldsymbol{P} = \sum_i \boldsymbol{p}_i = M \frac{d\boldsymbol{X}}{dt} \tag{3.24}$$

and is conserved if the total external force is zero.

The total angular momentum about the chosen origin is found by summing the individual bivector contributions,

$$L = \sum_i \boldsymbol{x}_i \wedge \boldsymbol{p}_i. \tag{3.25}$$

The rate of change of L is governed by

$$\dot{L} = \sum_i \boldsymbol{x}_i \wedge \dot{\boldsymbol{p}}_i = \sum_i \boldsymbol{x}_i \wedge \boldsymbol{f}_i^e + \sum_{i,j} \boldsymbol{x}_i \wedge \boldsymbol{f}_{ji}. \tag{3.26}$$

The final term is a double sum containing pairs of terms going as

$$\boldsymbol{x}_i \wedge \boldsymbol{f}_{ji} + \boldsymbol{x}_j \wedge \boldsymbol{f}_{ij} = (\boldsymbol{x}_i - \boldsymbol{x}_j) \wedge \boldsymbol{f}_{ji}. \tag{3.27}$$

The *strong* form of Newton's third law states that the interparticle force $\boldsymbol{f}_{ij}$ is directed along the vector $\boldsymbol{x}_i - \boldsymbol{x}_j$ between the two particles. This law is obeyed by a sufficiently wide range of forces to make it a useful restriction. (The most notable exception to this law is electromagnetism.) Under this restriction the total angular momentum satisfies

$$\frac{dL}{dt} = N^e, \tag{3.28}$$

where N^e is the total external torque. If the applied external torque is zero, and the strong law of action and reaction is obeyed, then the total angular momentum is conserved.

A useful expression for the angular momentum is obtained by introducing a set of position vectors relative to the centre of mass. We write

$$\boldsymbol{x}_i = \boldsymbol{x}_i' + \boldsymbol{X}, \tag{3.29}$$

so that

$$\sum_i m_i \boldsymbol{x}_i' = 0. \tag{3.30}$$

The velocity of the ith particle is now

$$\boldsymbol{v}_i = \boldsymbol{v}_i' + \boldsymbol{v}, \tag{3.31}$$

where $\boldsymbol{v} = \dot{\boldsymbol{X}}$ is the velocity of the centre of mass. The total angular momentum contains four terms:

$$L = \sum_i \big(\boldsymbol{X}\wedge m_i\boldsymbol{v} + \boldsymbol{x}_i'\wedge m_i\boldsymbol{v}_i' + m_i\boldsymbol{x}_i'\wedge\boldsymbol{v} + \boldsymbol{X}\wedge m_i\boldsymbol{v}_i'\big). \tag{3.32}$$

The final two terms both contain factors of $\sum m_i\boldsymbol{x}_i'$ and so vanish, leaving

$$L = \boldsymbol{X}\wedge\boldsymbol{P} + \sum_i \boldsymbol{x}_i'\wedge\boldsymbol{p}_i'. \tag{3.33}$$

The total angular momentum is therefore the sum of the angular momentum of the centre of mass about the origin, plus the angular momentum of the system about the centre of mass. In many cases it is possible to chose the origin so that the centre of mass is at rest, in which case L is simply the total angular momentum about the centre of mass. Similar considerations hold for the kinetic energy, and it is straightforward to show that

$$T = \sum_i \tfrac{1}{2}m_i\boldsymbol{v}_i^2 = \tfrac{1}{2}M\boldsymbol{v}^2 + \tfrac{1}{2}\sum_i m_i{\boldsymbol{v}_i'}^2. \tag{3.34}$$

3.2 Two-body central force interactions

One of the most significant applications of the preceding ideas is to a system of two point masses moving under the influence of each other. The force acting between the particles is directed along the vector between them, and all external forces are assumed to vanish. It follows that both the total momentum $\boldsymbol{P}$ and angular momentum L are conserved.

We suppose that the particles have positions $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and masses m_1 and m_2. Newton's second law for the central force problem takes the form

$$m_1\ddot{\boldsymbol{x}}_1 = \boldsymbol{f}, \tag{3.35}$$

$$m_2\ddot{\boldsymbol{x}}_2 = -\boldsymbol{f}, \tag{3.36}$$

where $\boldsymbol{f}$ is the interparticle force. We define the relative separation vector $\boldsymbol{x}$ by

$$\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2. \tag{3.37}$$

This vector satisfies

$$m_1 m_2 \ddot{\boldsymbol{x}} = (m_1 + m_2)\boldsymbol{f}. \tag{3.38}$$

We accordingly define the *reduced mass* μ by

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}, \tag{3.39}$$

so that the final force equation can be written as

$$\mu \ddot{\boldsymbol{x}} = \boldsymbol{f}. \tag{3.40}$$

The two-body problem has now been reduced to an equivalent single-body equation. The strong form of the third law assumed here means that the force $\boldsymbol{f}$ is directed along $\boldsymbol{x}$, so we can write $\boldsymbol{f}$ as $f\hat{\boldsymbol{x}}$.

We next re-express the total angular momentum in terms of the centre of mass $\boldsymbol{X}$ and the relative vector $\boldsymbol{x}$. We start by writing

$$m_1 \boldsymbol{x}_1 = m_1 \boldsymbol{X} + \mu \boldsymbol{x}, \quad m_2 \boldsymbol{x}_2 = m_2 \boldsymbol{X} - \mu \boldsymbol{x}. \tag{3.41}$$

It follows that the total angular momentum L_t is given by

$$\begin{aligned} L_t &= m_1 \boldsymbol{x}_1 \wedge \dot{\boldsymbol{x}}_1 + m_2 \boldsymbol{x}_2 \wedge \dot{\boldsymbol{x}}_2 \\ &= M \boldsymbol{X} \wedge \dot{\boldsymbol{X}} + \mu \boldsymbol{x} \wedge \dot{\boldsymbol{x}}. \end{aligned} \tag{3.42}$$

We have assumed that there are no external forces acting, so both L_t and $\boldsymbol{P}$ are conserved. It follows that the internal angular momentum is also conserved and we write this as

$$L = \mu \boldsymbol{x} \wedge \dot{\boldsymbol{x}}. \tag{3.43}$$

Since L is constant, the motion of the particles is confined to the L plane. The trajectory of $\boldsymbol{x}$ must also sweep out area at a constant rate, since this is how L is defined. For planetary motion this is Kepler's second law, though he did not state it in quite this form. Kepler treated the sun as the origin, whereas L should be defined relative to the centre of mass.

The internal kinetic energy is

$$T = \tfrac{1}{2}\mu \dot{\boldsymbol{x}}^2 = \tfrac{1}{2}\mu(\dot{r}\hat{\boldsymbol{x}} + r\dot{\hat{\boldsymbol{x}}})^2 = \tfrac{1}{2}\mu \dot{r}^2 + \tfrac{1}{2}\mu r^2 \dot{\hat{\boldsymbol{x}}}^2. \tag{3.44}$$

From equation (3.17) we see that

$$L^2 = -\mu^2 r^4 \hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}}\dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}} = -\mu^2 r^4 \dot{\hat{\boldsymbol{x}}}^2. \tag{3.45}$$

We therefore define the constant l as the magnitude of L, so

$$l = \mu r^2 |\dot{\hat{\boldsymbol{x}}}|. \tag{3.46}$$

The kinetic energy can now be written as a function of r and $\dot{r}$ only:

$$T = \frac{\mu \dot{r}^2}{2} + \frac{l^2}{2\mu r^2}. \tag{3.47}$$

The force $\boldsymbol{f}$ is conservative and can be written in terms of a potential $V(r)$ as

$$\boldsymbol{f} = f\hat{\boldsymbol{x}} = -\boldsymbol{\nabla} V(r), \tag{3.48}$$

where

$$f = -\frac{dV}{dr}. \tag{3.49}$$

Since the force is conservative the total energy is conserved, so

$$E = \frac{\mu \dot{r}^2}{2} + \frac{l^2}{2\mu r^2} + V(r) \tag{3.50}$$

is a constant. For a given potential $V(r)$ this equation can be integrated to find the evolution of r. The full motion can then be recovered from L.

3.2.1 Inverse-square forces

The most important example of a two-body central force interaction is that described by an inverse-square force law. This case is encountered in gravitation and electrostatics and has been analysed in considerable detail by many authors (see the end of this chapter for suggested additional reading). In this section we review some of the key features of this system, highlighting the places where geometric algebra offers something new. An alternative approach to this problem is discussed in section 3.3.

Writing $f = -k/r^2$ the basic equation to solve is

$$\mu \ddot{\boldsymbol{x}} = -\frac{k}{r^2}\hat{\boldsymbol{x}} = -\frac{k}{r^3}\boldsymbol{x}. \tag{3.51}$$

The sign of k determines whether the force is attractive or repulsive (positive for attractive). This is a second-order vector differential equation, so we expect there to be two constant vectors in the solution — one for the initial position and one for the velocity. We already know that the angular momentum L is a constant of motion, and we can write this as

$$L = \mu r^2 \hat{\boldsymbol{x}}\dot{\hat{\boldsymbol{x}}} = -\mu r^2 \dot{\hat{\boldsymbol{x}}}\hat{\boldsymbol{x}}. \tag{3.52}$$

It follows that

$$L\dot{\boldsymbol{v}} = -\frac{k}{\mu r^2} L\hat{\boldsymbol{x}} = k\dot{\hat{\boldsymbol{x}}}, \tag{3.53}$$

which we can write in the form

$$\frac{d}{dt}(L\boldsymbol{v} - k\hat{\boldsymbol{x}}) = 0. \tag{3.54}$$

Eccentricity	Energy	Orbit
$e > 1$	$E > 0$	Hyperbola
$e = 1$	$E = 0$	Parabola
$e < 1$	$E < 0$	Ellipse
$e = 0$	$E = -\mu k^2/(2l^2)$	Circle

Table 3.1 *Classification of orbits for an inverse-square force law.*

The motion is therefore described by the simple equation

$$L\boldsymbol{v} = k(\hat{\boldsymbol{x}} + \boldsymbol{e}), \tag{3.55}$$

where the *eccentricity vector* $\boldsymbol{e}$ is a second vector constant of motion. This vector is also known in various contexts as the Laplace vector and as the Runge–Lenz vector. From its definition we can see that $\boldsymbol{e}$ must lie in the L plane.

To find a direct equation for the trajectory we first write

$$L\boldsymbol{v}\boldsymbol{x} = L(\boldsymbol{v}\cdot\boldsymbol{x} + \boldsymbol{v}\wedge\boldsymbol{x}) = \frac{1}{\mu}L\tilde{L} + \boldsymbol{v}\cdot\boldsymbol{x}\,L = k(r + \boldsymbol{e}\boldsymbol{x}). \tag{3.56}$$

The scalar part of this equation gives

$$r = \frac{l^2}{k\mu(1 + \boldsymbol{e}\cdot\hat{\boldsymbol{x}})}. \tag{3.57}$$

This equation specifies a conic surface in three dimensions with symmetry axis $\boldsymbol{e}$. The surface is formed by rotating a two-dimensional conic about this axis. Since the motion takes place entirely within the L plane the motion is described by a conic. That is, the trajectory $\boldsymbol{x}(t)$ is one of a hyperbola, parabola, ellipse or circle. The generic cases are ellipses for bound orbits and hyperbolae for free states. The cases of parabolic and circular orbits are exceptional as they require precise values of $|\boldsymbol{e}|$ (table 3.1).

In L and $\boldsymbol{e}$ we have found five of the six constants of motion (we only have two arbitrary constants in $\boldsymbol{e}$ as it is constrained to lie in the L plane). The final constant specifies where on the conic we start at time $t = 0$. We know that the energy is also a constant of motion, so it should be possible to express the energy directly in terms of L and $\boldsymbol{e}$. From equation (3.51) we see that the potential energy must go as k/r, provided we set the arbitrary constant so that $V = 0$ at infinity. The full energy is therefore given by

$$E = \frac{\mu}{2}\boldsymbol{v}^2 - \frac{k}{r}. \tag{3.58}$$

To simplify this we first form

$$L\boldsymbol{v}\boldsymbol{v}\tilde{L} = l^2 v^2 = k^2(\hat{\boldsymbol{x}} + \boldsymbol{e})^2. \tag{3.59}$$

It follows that

$$E = \frac{\mu k^2}{2l^2}(e^2 + 1 + 2\hat{\boldsymbol{x}}\cdot\boldsymbol{e}) - \frac{k}{r} = \frac{\mu k^2}{2l^2}(e^2 - 1), \tag{3.60}$$

where $e = |\boldsymbol{e}|$ is the eccentricity. The sign of the energy is governed entirely by e. Since the potential is set to zero at infinity, all bound states must have negative energy and hence an eccentricity $e < 1$. The limiting case of $e = 1$ describes a parabola (table 3.1).

3.2.2 Motion in time for elliptic orbits

Many methods can be used to find the trajectory as a function of time and these are discussed widely in the literature. Here we describe one of the simplest, which serves to highlight the essential difficulty of this problem. An alternative solution, which more fully exploits the techniques of geometric algebra, is described in section 3.3. From the energy equation we see that

$$\mu^2\dot{r}^2 = 2\mu E - \frac{l^2}{r^2} + \frac{2\mu k}{r}, \tag{3.61}$$

so t is given by

$$t = \mu\int_{r_0}^{r_1} \frac{r\,dr}{(2\mu k r + 2\mu E r^2 - l^2)^{1/2}}. \tag{3.62}$$

Evaluating this integral results in a rather complicated function of r, the general form of which is hard to invert and not very helpful. More useful formulae are obtained by specialising to one form of orbit. For bound problems we are interested in elliptic orbits for which E is negative. For these orbits it is useful to introduce the *semi-major axis* a defined by

$$a = \tfrac{1}{2}(r_1 + r_2) = -\frac{k}{2E}, \tag{3.63}$$

where r_1 and r_2 are the maximum and minimum values of r respectively. In terms of this we can write

$$2\mu k r + 2\mu E r^2 - l^2 = -\frac{\mu k}{a}(r^2 - 2ar) - l^2 = \frac{\mu k}{a}\big(a^2e^2 - (r-a)^2\big). \tag{3.64}$$

We now introduce a new variable Ψ, the *eccentric anomaly*, defined by

$$r = a\big(1 - e\cos(\Psi)\big). \tag{3.65}$$

In terms of this we find

$$t = \left(\frac{\mu a^3}{k}\right)^{1/2}\int_{\Psi_0}^{\Psi_1}\big(1 - e\cos(\Psi)\big)\,d\Psi, \tag{3.66}$$

so if we choose $t = 0$ to correspond to closest approach we have

$$\omega t = \Psi - e\sin(\Psi), \tag{3.67}$$

where

$$\omega^2 = \frac{k}{\mu a^3}. \tag{3.68}$$

Equations (3.65) and (3.67) provide a parametric solution relating r and t. This solution highlights the fact that the equation relating t and r is transcendental and does not have a simple closed form. The time taken for one orbit is $2\pi/\omega$, so the orbital period τ is related to the major axis a by

$$\tau^2 = \frac{4\pi^2\mu}{k}a^3. \tag{3.69}$$

This gives us the third of Kepler's three laws of planetary motion, that the square of the period is proportional to the cube of the major axis.

3.3 Celestial mechanics and perturbations

By far the most important application of the Newtonian theory of gravitation is to the motion of the planets in the solar system. This is a complicated subject of considerable historical and current importance, and we will only touch on a few applications. Detailed calculation of the motions of all of the planets in the solar system still represents a major computational challenge. Aside from the obvious problem of having to calculate the gravitational effects of every planet on every other planet, further effects must also be incorporated. These can include deviations of the shapes of the planets from spherical, the effects of tidal forces and ultimately general relativistic corrections.

A significant number of problems in celestial mechanics are best treated using perturbation theory. In this technique orbits are calculated as a series of ever smaller deviations from Kepler orbits. Since the Kepler orbit is specified entirely by L and $\boldsymbol{e}$, we should first form equations for these in the presence of a perturbing force. We modify the force law to read

$$\mu\ddot{\boldsymbol{x}} = -\frac{k}{r^3}\boldsymbol{x} + \boldsymbol{f}, \tag{3.70}$$

and assume that $\boldsymbol{f}$ is always small compared with the inverse-square term. The angular momentum L now satisfies

$$\dot{L} = \boldsymbol{x}\wedge\boldsymbol{f}, \tag{3.71}$$

so L is now only conserved if $\boldsymbol{f}$ is also a central force. With the eccentricity vector still defined by equation (3.55), we find that

$$k\dot{\boldsymbol{e}} = \dot{L}\cdot\boldsymbol{v} + \frac{1}{\mu}L\cdot\boldsymbol{f}. \tag{3.72}$$

Only five of the six equations for L and $\boldsymbol{e}$ are independent, as we always have $L\wedge\boldsymbol{e} = 0$.

For many problems the variation in L and $\boldsymbol{e}$ is slow compared to the orbital period. For these a useful approximation is obtained by finding the orbital average of $\boldsymbol{f}$ over one cycle, with L and $\boldsymbol{e}$ held constant. The quantities L and $\boldsymbol{e}$ are then assumed to vary slowly under the influence of the time-averaged force. Results for the orbital averages of numerous quantities can be found tabulated in many textbooks and are discussed in the exercises at the end of this chapter.

3.3.1 Example — general relativistic perturbations

Later in this book we will study how general relativity modifies the Newtonian view of gravity. For particles moving in a central potential, the modification is quite simple and can be handled efficiently using perturbation theory. The modified force law is

$$\ddot{\boldsymbol{x}} = -\frac{GM}{r^2}\left(1 + \frac{3l^2}{\mu^2 c^2 r^2}\right)\hat{\boldsymbol{x}}, \tag{3.73}$$

where c is the speed of light and we have replaced k by the gravitational expression $GM\mu$. (A small subtlety is that the derivatives here are with respect to proper time, but this does not affect our reasoning.) The force is still central, so the angular momentum L is still conserved. The eccentricity vector satisfies the simple equation

$$\dot{\boldsymbol{e}} = \frac{3l^2}{\mu^3 c^2 r^4} L\cdot\hat{\boldsymbol{x}}. \tag{3.74}$$

For bound orbits this gives rise to a precession of the major axis (see figure 3.2). The quantity of most interest is the amount $\boldsymbol{e}$ changes in one orbit. To get an approximate result for this we use the time-averaging idea and assume that the orbit is precisely elliptical. We therefore have

$$\Delta\boldsymbol{e} = -\frac{3l^2}{\mu^3 c^2} L \int_0^T dt \frac{\hat{\boldsymbol{x}}}{r^4}, \tag{3.75}$$

where T is the orbital period. Evaluating this integral is left as an exercise, and the final result is

$$\Delta\boldsymbol{e} = \frac{6\pi GM}{a(1-e^2)c^2}\boldsymbol{e}\cdot\hat{L}, \tag{3.76}$$

where $\hat{L} = L/l$. This gives a precession of $\boldsymbol{e}$ with the orientation of L, which corresponds to an advance (figure 3.2). For Mercury this gives rise to the famous advance in the perihelion of 43 arcseconds per century, which was finally explained by general relativity.

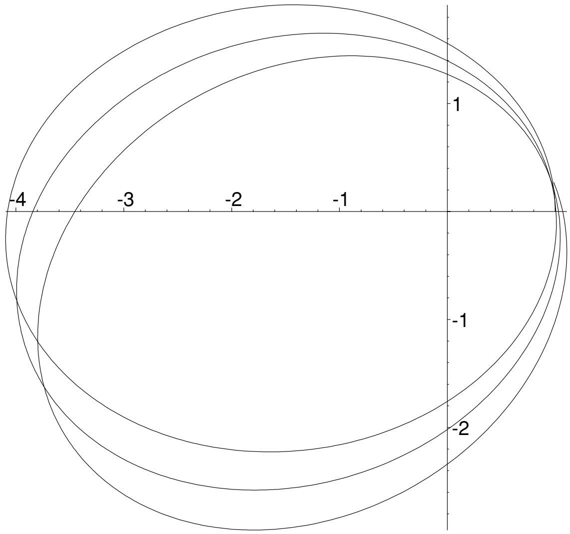

Figure 3.2 *Orbital precession.* The plot shows a modified orbit as predicted by general relativity. The ellipse precesses round in the same direction as the orbital motion. The parameters have been chosen to exaggerate the precession effect.

3.3.2 Spinor equations

An alternative method for analysing the Kepler problem is through the use of 'spinors'. These will be defined more carefully in later chapters, but in two and three dimensions they can be viewed as elements of the subalgebra of $\mathcal{G}_2$ and $\mathcal{G}_3$ consisting entirely of even elements. In two dimensions a spinor can therefore be identified with a complex number. The position vector $\boldsymbol{x}$ in two dimensions can be formed through a rotation and dilation via the polar decomposition

$$\boldsymbol{x} = \mathsf{e}_1 r \exp(\theta \mathsf{e}_1 \mathsf{e}_2) = r \exp(-\theta \mathsf{e}_1 \mathsf{e}_2) \mathsf{e}_1, \tag{3.77}$$

where $\{\mathsf{e}_1, \mathsf{e}_2\}$ denote a right-handed orthonormal frame and we assume that the vector lies in the $\mathsf{e}_1\mathsf{e}_2$ plane. We know from chapter 2 that the rotation formula only extends to higher dimensions if a double-sided prescription is adopted, so we write the vector $\boldsymbol{x}$ as

$$\boldsymbol{x} = U \mathsf{e}_1 U^\dagger = U^2 \mathsf{e}_1 = \mathsf{e}_1 {U^\dagger}^2. \tag{3.78}$$

In writing this we have placed all of the dynamics in the complex number U.

For the Kepler problem it turns out that the equation for U is considerably easier than that for $\boldsymbol{x}$. We assume that the plane of L is given by $\mathsf{e}_1\mathsf{e}_2$, and start by forming

$$r = |\boldsymbol{x}| = UU^\dagger. \tag{3.79}$$

(Recall that, for a scalar + bivector combination in two dimensions, the reverse operator is the same as complex conjugation.) On differentiating we find that

$$\dot{\boldsymbol{x}} = 2\dot{U}U\mathsf{e}_1, \tag{3.80}$$

hence

$$2r\dot{U} = \dot{\boldsymbol{x}}\mathsf{e}_1 U^\dagger = \dot{\boldsymbol{x}}U\mathsf{e}_1. \tag{3.81}$$

We now introduce the new variable s defined by

$$\frac{d}{ds} = r\frac{d}{dt}, \qquad \frac{dt}{ds} = r. \tag{3.82}$$

In terms of this

$$2\frac{dU}{ds} = \dot{\boldsymbol{x}}U\mathsf{e}_1 \tag{3.83}$$

and

$$2\frac{d^2U}{ds^2} = r\ddot{\boldsymbol{x}}U\mathsf{e}_1 + \dot{\boldsymbol{x}}\frac{dU}{ds}\mathsf{e}_1 = U\big(\ddot{\boldsymbol{x}}\boldsymbol{x} + \tfrac{1}{2}\dot{\boldsymbol{x}}^2\big). \tag{3.84}$$

Now suppose we have motion in a central inverse-square force:

$$\mu\ddot{\boldsymbol{x}} = -k\frac{\boldsymbol{x}}{r^3}. \tag{3.85}$$

The equation for U becomes

$$\frac{d^2U}{ds^2} = \frac{1}{2\mu}U\left(\tfrac{1}{2}\mu\dot{\boldsymbol{x}}^2 - \frac{k}{r}\right) = \frac{E}{2\mu}U, \tag{3.86}$$

which is simply the equation for harmonic motion! This has a number of advantages. First of all, the equation is easy to solve. If we set

$$\omega^2 = -\frac{E}{2\mu} \tag{3.87}$$

then the general solution is

$$U = A\exp(\hat{L}\omega s) + B\exp(-\hat{L}\omega s), \tag{3.88}$$

where A and B are constants and $\hat{L}$ is the unit bivector for the plane of motion. The motion is illustrated in figure 3.3. The particle trajectory maps out an ellipse with the origin at one focus, whereas U defines an ellipse with the origin at the centre. The particle completes two orbits for each full cycle of U.

Further advantages of formulating the dynamics in terms of U are that the equation for U is linear, so is better suited to perturbation theory, and that there is no singularity at $r = 0$, which provides better numerical stability. (Removing this singularity is called 'regularization'.) In addition, equation (3.86) is universal

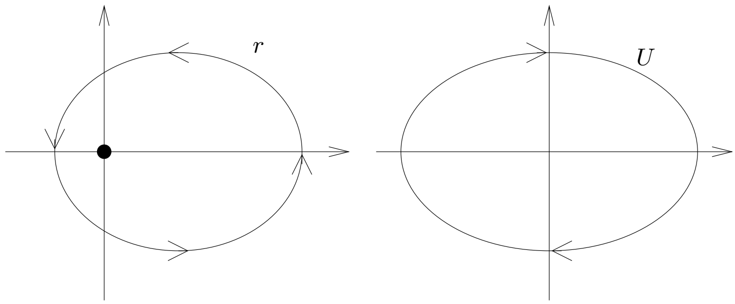

Figure 3.3 *Solution to the Kepler problem.* The particle orbit is shown on the left, and the corresponding spinor on the right. The particle completes two orbits every time U completes one cycle, since U and $-U$ describe the same position.

— it holds for $E > 0$ and $E < 0$. The solution when $E > 0$ simply has trigonometric functions replaced by exponentials. This universality is important, because perturbations can often send bound orbits into unbound ones.

For the method to be truly powerful, however, it must extend to three dimensions. The relevant formula in three dimensions is

$$\boldsymbol{x} = U \mathsf{e}_1 U^\dagger, \tag{3.89}$$

where U is a general even element. This means that U has four degrees of freedom now, whereas only three are required to specify $\boldsymbol{x}$. We are therefore free to impose a further additional constraint on U, which we will use to ensure the equations take on a convenient form. The quantity $UU^\dagger$ is still a scalar in three dimensions, so we have

$$r = UU^\dagger = U^\dagger U. \tag{3.90}$$

We next form $\dot{\boldsymbol{x}}$:

$$\dot{\boldsymbol{x}} = \dot{U} \mathsf{e}_1 U^\dagger + U \mathsf{e}_1 \dot{U}^\dagger. \tag{3.91}$$

We would like this to equal $2\dot{U}\mathsf{e}_1 U^\dagger$ for the preceding analysis to follow through. For this to hold we require

$$\dot{U}\mathsf{e}_1 U^\dagger - U \mathsf{e}_1 \dot{U}^\dagger = \dot{U}\mathsf{e}_1 U^\dagger - (\dot{U}\mathsf{e}_1 U^\dagger)^\dagger = 0. \tag{3.92}$$

The quantity $\dot{U}\mathsf{e}_1 U^\dagger$ only contains odd grade terms (grade-1 and grade-3). If we subtract its reverse, all that remains is the trivector (pseudoscalar) term. We therefore require that

$$\langle \dot{U}\mathsf{e}_1 U^\dagger \rangle_3 = 0, \tag{3.93}$$

which we adopt as our extra condition on U. With this condition satisfied we have

$$2\frac{dU}{ds} = \dot{\boldsymbol{x}}U\mathsf{e}_1 \tag{3.94}$$

and

$$2\frac{d^2U}{ds^2} = \big(\ddot{\boldsymbol{x}}\boldsymbol{x} + \tfrac{1}{2}\dot{\boldsymbol{x}}^2\big)U. \tag{3.95}$$

For an inverse-square force law we therefore recover the same harmonic oscillator equation. In the presence of a perturbing force we have

$$2\mu\frac{d^2U}{ds^2} - EU = \boldsymbol{f}\boldsymbol{x}U = r\boldsymbol{f}U\mathsf{e}_1. \tag{3.96}$$

This equation for U can be handled using standard techniques from perturbation theory. The equation was first found (in matrix form) by Kustaanheimo and Stiefel in 1964 . The analysis was refined and cast in its present form by Hestenes (1999).

3.4 Rotating systems and rigid-body motion

Rigid bodies can be viewed as another example of a system of particles, where now the effect of the internal forces is to keep all of the interparticle distances fixed. For such systems the internal forces can be ignored once one has found a set of dynamical variables that enforce the rigid-body constraint. The problem then reduces to solving for the motion of the centre of mass and for the angular momentum in the presence of any external forces or torques. Suitable variables are a vector $\boldsymbol{x}(t)$ for the centre of mass, and a set of variables to describe the attitude of the rigid body in space. Many forms exist for the latter variables, but here we will concentrate on parameterising the attitude of the rigid body with a *rotor*. Before applying this idea to rigid-body motion, we first look at the description of rotating frames with rotors.

3.4.1 Rotating frames

Suppose that the frame of vectors $\{\mathsf{f}_k\}$ is rotating in space. These can be related to a fixed orthonormal frame $\{\mathsf{e}_k\}$ by the time-dependent rotor $R(t)$:

$$\mathsf{f}_k(t) = R(t)\mathsf{e}_k R^\dagger(t). \tag{3.97}$$

The angular velocity vector $\boldsymbol{\omega}$ is traditionally defined by the formula

$$\dot{\mathsf{f}}_k = \boldsymbol{\omega}\times\mathsf{f}_k, \tag{3.98}$$

where the cross denotes the vector cross product. From section 2.4.3 we know that the cross product is related to the inner product with a bivector by

$$\boldsymbol{\omega}\times\mathsf{f}_k = (-I\boldsymbol{\omega})\cdot\mathsf{f}_k = \mathsf{f}_k\cdot(I\boldsymbol{\omega}). \tag{3.99}$$

We are now used to the idea that angular momentum is best viewed as a bivector, and we must expect the same to be true for angular velocity. We therefore define the angular velocity bivector Ω by

$$\Omega = I\boldsymbol{\omega}. \tag{3.100}$$

This choice ensures that the rotation has the orientation implied by Ω.

To see how Ω is related to the rotor R we start by differentiating equation (3.97):

$$\dot{\mathsf{f}}_k = \dot{R}\mathsf{e}_k R^\dagger + R\mathsf{e}_k \dot{R}^\dagger = \dot{R}R^\dagger \mathsf{f}_k + \mathsf{f}_k R\dot{R}^\dagger. \tag{3.101}$$

From the normalisation equation $RR^\dagger = 1$ we find that

$$0 = \frac{d}{dt}(RR^\dagger) = \dot{R}R^\dagger + R\dot{R}^\dagger. \tag{3.102}$$

Since differentiation and reversion are interchangeable operations we now have

$$\dot{R}R^\dagger = -R\dot{R}^\dagger = -(\dot{R}R^\dagger)^\dagger. \tag{3.103}$$

The quantity $\dot{R}R^\dagger$ is equal to minus its own reverse and has even grade, so must be a pure bivector. The equation for f_k now becomes

$$\dot{\mathsf{f}}_k = \dot{R}R^\dagger \mathsf{f}_k - \mathsf{f}_k \dot{R}R^\dagger = (2\dot{R}R^\dagger)\cdot \mathsf{f}_k. \tag{3.104}$$

Comparing this with equation (3.99) and equation (3.100) we see that $2\dot{R}R^\dagger$ must equal minus the angular velocity bivector Ω, so

$$2\dot{R}R^\dagger = -\Omega. \tag{3.105}$$

The dynamics is therefore contained in the single *rotor equation*

$$\dot{R} = -\tfrac{1}{2}\Omega R. \tag{3.106}$$

The reversed form of this is also useful:

$$\dot{R}^\dagger = \tfrac{1}{2}R^\dagger \Omega. \tag{3.107}$$

Equations of this type are surprisingly ubiquitous in physics. In the more general setting, rotors are viewed as elements of a *Lie group*, and the bivectors form their *Lie algebra*. We will have more to say about this in chapter 11.

3.4.2 Constant Ω

For the case of constant Ω equation (3.106) integrates immediately to give

$$R = \mathrm{e}^{-\Omega t/2} R_0, \tag{3.108}$$

which is the rotor for a constant frequency rotation in the positive sense in the Ω plane. The frame rotates according to

$$\mathsf{f}_k(t) = \mathrm{e}^{-\Omega t/2} R_0 \mathsf{e}_k R_0^\dagger \mathrm{e}^{\Omega t/2}. \tag{3.109}$$

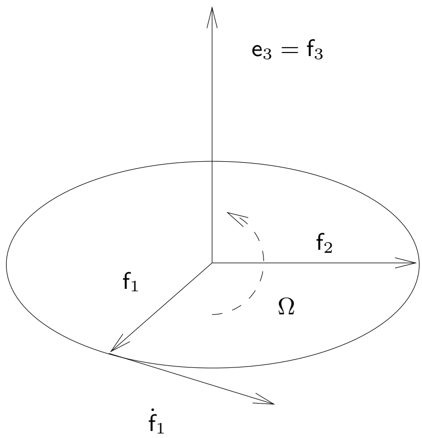

Figure 3.4 *Orientation of the angular velocity bivector.* Ω has the orientation of $f_1 \wedge \dot{f}_1$. It must therefore have orientation $+e_1 \wedge e_2$ when $\boldsymbol{\omega} = e_3$.

The constant term R_0 describes the orientation of the frame at $t = 0$, relative to the $\{e_k\}$ frame.

As an example, consider the case of motion about the e_3 axis (figure 3.4). We have

$$\Omega = \omega I e_3 = \omega e_1 e_2, \tag{3.110}$$

and for convenience we set $R_0 = 1$. The motion is described by

$$f_k(t) = \exp\Big(-\tfrac{1}{2} e_1 e_2 \omega t\Big) e_k \exp\Big(\tfrac{1}{2} e_1 e_2 \omega t\Big), \tag{3.111}$$

so that the f_1 axis rotates as

$$f_1 = e_1 \exp(e_1 e_2 \omega t) = \cos(\omega t) e_1 + \sin(\omega t) e_2. \tag{3.112}$$

This defines a *right-handed* (anticlockwise) rotation in the $e_1 e_2$ plane, as prescribed by the orientation of Ω.

3.4.3 Rigid-body motion

Suppose that a rigid body is moving through space. To describe the position in space of any part of the body, we need to specify the position of the centre of mass, and the vector to the point in the body from the centre of mass. The latter can be encoded in terms of a rotation from a fixed 'reference' body onto the body in space (figure 3.5). We let $\boldsymbol{x}_0$ denote the position of the centre of mass and $\boldsymbol{y}_i(t)$ denote the position (in space) of a point in the body. These are related by

$$\boldsymbol{y}_i(t) = R(t) \boldsymbol{x}_i R^\dagger(t) + \boldsymbol{x}_0(t), \tag{3.113}$$

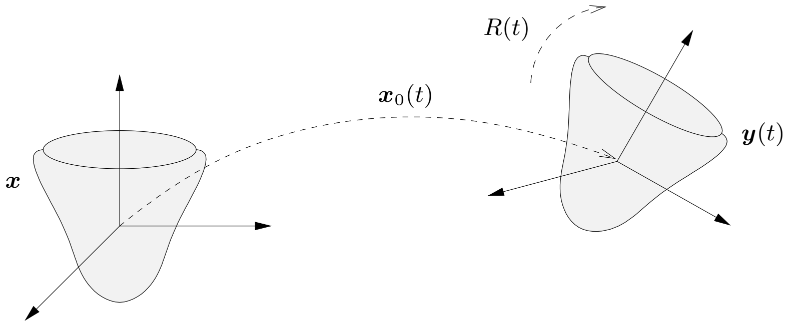

Figure 3.5 *Description of a rigid body.* The vector $\boldsymbol{x}_0(t)$ specifies the position of the centre of mass, relative to the origin. The rotor $R(t)$ defines the orientation of the body, relative to a fixed copy imagined to be placed at the origin. $\boldsymbol{x}$ is a vector in the reference body, and $\boldsymbol{y}$ is the vector in space of the equivalent point on the moving body.

where $\boldsymbol{x}_i$ is a fixed constant vector in the reference copy of the body. In this manner we have placed all of the rotational motion in the time-dependent rotor $R(t)$.

The velocity of the point $\boldsymbol{y} = R\boldsymbol{x}R^\dagger + \boldsymbol{x}_0$ is

$$\begin{aligned}\boldsymbol{v}(t) &= \dot{R}\boldsymbol{x}R^\dagger + R\boldsymbol{x}\dot{R}^\dagger + \dot{\boldsymbol{x}}_0 \\ &= -\tfrac{1}{2}\Omega R\boldsymbol{x}R^\dagger + \tfrac{1}{2}R\boldsymbol{x}R^\dagger\Omega + \boldsymbol{v}_0 \\ &= (R\boldsymbol{x}R^\dagger)\cdot\Omega + \boldsymbol{v}_0,\end{aligned} \tag{3.114}$$

where $\boldsymbol{v}_0$ is the velocity of the centre of mass. The bivector Ω defines the plane of rotation in space. This plane will lie at some orientation relative to the current position of the rigid body. For studying the motion it turns out to be extremely useful to transform the rotation plane back into the fixed, reference copy of the body. Since bivectors are subject to the same rotor transformation law as vectors we define the 'body' angular velocity Ω_B by

$$\Omega_B = R^\dagger \Omega R. \tag{3.115}$$

In terms of the body angular velocity the rotor equation becomes

$$\dot{R} = -\tfrac{1}{2}\Omega R = -\tfrac{1}{2}R\Omega_B, \qquad \dot{R}^\dagger = \tfrac{1}{2}\Omega_B R^\dagger. \tag{3.116}$$

The velocity of the body is now re-expressed as

$$\boldsymbol{v}(t) = R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger + \boldsymbol{v}_0, \tag{3.117}$$

which will turn out to be the more convenient form. (We have used the operator ordering conventions of section 2.5 to suppress unnecessary brackets in writing $R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger$ in place of $R(\boldsymbol{x}\cdot\Omega_B)R^\dagger$.)

To calculate the momentum of the rigid body we need the masses of each of the constituent particles. It is easier at this point to go to a continuum approximation and introduce a density $\rho = \rho(\boldsymbol{x})$. The position vector $\boldsymbol{x}$ is taken relative to the centre of mass, so we have

$$\int d^3x\, \rho = M \quad \text{and} \quad \int d^3x\, \rho \boldsymbol{x} = 0. \tag{3.118}$$

The momentum of the rigid body is simply

$$\int d^3x\, \rho \boldsymbol{v} = \int d^3x\, \rho (R\, \boldsymbol{x} \cdot \Omega_B\, R^\dagger + \boldsymbol{v}_0) = M \boldsymbol{v}_0, \tag{3.119}$$

so is specified entirely by the motion of the centre of mass. This is the continuum version of the result of section 3.1.2.

3.4.4 The inertia tensor

The next quantity we require is the angular momentum bivector L for the body about its centre of mass. We therefore form

$$\begin{aligned} L &= \int d^3x\, \rho (\boldsymbol{y} - \boldsymbol{x}_0) \wedge \boldsymbol{v} \\ &= \int d^3x\, \rho (R \boldsymbol{x} R^\dagger) \wedge (R\, \boldsymbol{x} \cdot \Omega_B\, R^\dagger + \boldsymbol{v}_0) \\ &= R \left(\int d^3x\, \rho \boldsymbol{x} \wedge (\boldsymbol{x} \cdot \Omega_B) \right) R^\dagger. \end{aligned} \tag{3.120}$$

The integral inside the brackets refers only to the fixed copy and so defines a time-independent function of Ω_B. This is the reason for working with Ω_B instead of the space angular velocity Ω. We define the *inertia tensor* $\mathcal{I}(B)$ by

$$\mathcal{I}(B) = \int d^3x\, \rho \boldsymbol{x} \wedge (\boldsymbol{x} \cdot B). \tag{3.121}$$

This is a linear function mapping bivectors to bivectors. This way of writing linear functions may be unfamiliar to those used to seeing tensors labelled with indices, but the notation is the natural extension to linear functions of the index-free approach advocated in this book. The linearity of the map is easy to check:

$$\begin{aligned} \mathcal{I}(\lambda A + \mu B) &= \int d^3x\, \rho \boldsymbol{x} \wedge \big(\boldsymbol{x} \cdot (\lambda A + \mu B) \big) \\ &= \int d^3x\, \rho \big(\lambda \boldsymbol{x} \wedge (\boldsymbol{x} \cdot A) + \mu \boldsymbol{x} \wedge (\boldsymbol{x} \cdot B) \big) \\ &= \lambda \mathcal{I}(A) + \mu \mathcal{I}(B). \end{aligned} \tag{3.122}$$

The fact that the inertia tensor maps bivectors to bivectors, rather than vectors to vectors, is also a break from tradition. This viewpoint is very natural given our earlier comments about the merits of bivectors over axial vectors, and provides a

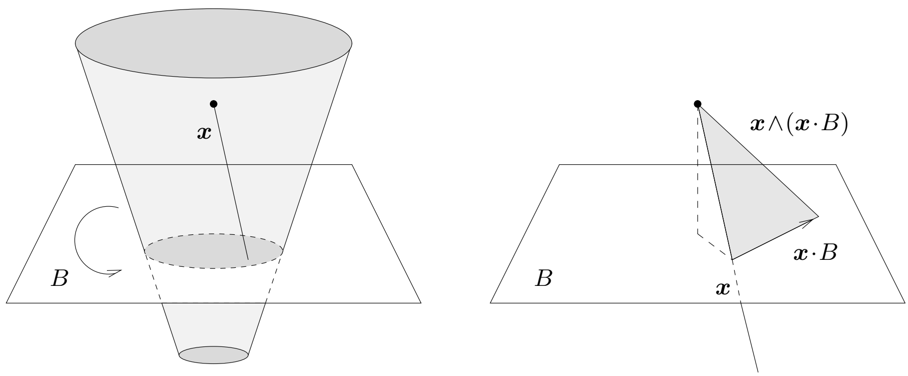

Figure 3.6 *The inertia tensor.* The inertia tensor $\mathcal{I}(B)$ is a linear function mapping its bivector argument B onto a bivector. It returns the total angular momentum about the centre of mass for rotation in the B plane.

clear geometric picture of the tensor (figure 3.6). Since both vectors and bivectors belong to a three-dimensional linear space, there is no additional complexity introduced in this new picture.

To understand the effect of the inertia tensor, suppose that the body rotates in the B plane at a fixed rate $|B|$, and we place the origin at the centre of mass (which is fixed). The velocity of the vector $\boldsymbol{x}$ is simply $\boldsymbol{x}\cdot B$, and the momentum density at this point is $\rho\boldsymbol{x}\cdot B$, as shown in figure 3.6. The angular momentum density bivector is therefore $\boldsymbol{x}\wedge(\rho\boldsymbol{x}\cdot B)$, and integrating this over the entire body returns the total angular momentum bivector for rotation in the B plane.

In general, the total angular momentum will not lie in the same plane as the angular velocity. This is one reason why rigid-body dynamics can often seem quite counterintuitive. When we see a body rotating, our eyes naturally pick out the angular *velocity* by focusing on the vector the body rotates around. Deciding the plane of the angular momentum is less easy, particularly if the internal mass distribution is hidden from us. But it is the angular momentum that responds directly to external torques, not the angular velocity, and this can have some unexpected consequences.

We have calculated the inertia tensor about the centre of mass, but bodies rotating around a fixed axis can be forced to rotate about any point. A useful theorem relates the inertia tensor about an arbitrary point to one about the centre of mass. Suppose that we want the inertia tensor relative to the point $\boldsymbol{a}$, where $\boldsymbol{a}$ is a vector taken from the centre of mass. Returning to the definition of equation (3.121) we see that we need to compute

$$\mathcal{I}_a(B) = \int d^3x\,\rho(\boldsymbol{x}-\boldsymbol{a})\wedge\big((\boldsymbol{x}-\boldsymbol{a})\cdot B\big). \qquad (3.123)$$

This integral evaluates to give

$$\begin{aligned}\mathcal{I}_a(B) &= \int d^3x\,\rho\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B) - \boldsymbol{x}\wedge(\boldsymbol{a}\cdot B) - \boldsymbol{a}\wedge(\boldsymbol{x}\cdot B) + \boldsymbol{a}\wedge(\boldsymbol{a}\cdot B)\big)\\ &= \mathcal{I}(B) + M\boldsymbol{a}\wedge(\boldsymbol{a}\cdot B).\end{aligned} \tag{3.124}$$

The inertia tensor relative to $\boldsymbol{a}$ is simply the inertia tensor about the centre of mass, plus the tensor for a point mass M at position $\boldsymbol{a}$.

3.4.5 Principal axes

So far we have only given an abstract specification of the inertia tensor. For most calculations it is necessary to introduce a set of basis vectors fixed in the body. As we are free to choose the directions of these vectors, we should ensure that this choice simplifies the equations of motion as much as possible. To see how to do this, consider the $\{\mathsf{e}_i\}$ frame and define the matrix $\mathcal{I}_{ij}$ by

$$\mathcal{I}_{ij} = -(I\mathsf{e}_i)\cdot\mathcal{I}(I\mathsf{e}_j). \tag{3.125}$$

This defines a *symmetric* matrix, as follows from the result

$$A\cdot(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot B)) = \langle A\boldsymbol{x}(\boldsymbol{x}\cdot B)\rangle = \langle (A\cdot\boldsymbol{x})\boldsymbol{x}B\rangle = B\cdot(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot A)). \tag{3.126}$$

(This sort of manipulation, where one uses the projection onto grade to replace inner and outer products by geometric products, is very common in geometric algebra.) This result ensures that

$$\begin{aligned}\mathcal{I}_{ij} &= -\int d^3x\,\rho(I\mathsf{e}_i)\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\mathsf{e}_j))\big)\\ &= -\int d^3x\,\rho(I\mathsf{e}_j)\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\mathsf{e}_i))\big) = \mathcal{I}_{ji}.\end{aligned} \tag{3.127}$$

It follows that the matrix $\mathcal{I}_{ij}$ will be diagonal if the $\{\mathsf{e}_i\}$ frame is chosen to coincide with the eigendirections of the inertia tensor. These directions are called the *principal axes*, and we always choose our frame along these directions.

The matrix $\mathcal{I}_{ij}$ is also positive-(semi)definite, as can be seen from

$$\begin{aligned}a_i a_j \mathcal{I}_{ij} &= -\int d^3x\,\rho(I\boldsymbol{a})\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot(I\boldsymbol{a}))\big)\\ &= \int d^3x\,\rho\big(\boldsymbol{x}\cdot(I\boldsymbol{a})\big)^2 \geq 0.\end{aligned} \tag{3.128}$$

It follows that all of the eigenvalues of $\mathcal{I}_{ij}$ must be positive (or possibly zero for the case of point or line masses). These eigenvalues are the principal moments of inertia and are crucial in specifying the properties of a rigid body. We denote these $\{i_1, i_2, i_3\}$, so that

$$\mathcal{I}_{jk} = \delta_{jk} i_k \quad \text{(no sum)}. \tag{3.129}$$

(It is more traditional to use a capital I for the moments of inertia, but this symbol is already employed for the pseudoscalar.) If two or three of the principal moments are the same the principal axes are not uniquely specified. In this case one simply chooses one orthonormal set of eigenvectors from the degenerate family of possibilities.

Returning to the index-free presentation, we see that the principal axes satisfy

$$\mathcal{I}(I\mathsf{e}_j) = I\mathsf{e}_k \mathcal{I}_{jk} = i_j I\mathsf{e}_j, \tag{3.130}$$

where again there is no sum implied between eigenvectors and their associated eigenvalue in the final expression. To calculate the effect of the inertia tensor on an arbitrary bivector B we decompose B in terms of the principal axes as

$$B = B_j I\mathsf{e}_j. \tag{3.131}$$

It follows that

$$\mathcal{I}(B) = \sum_{j=1}^{3} i_j B_j I\mathsf{e}_j = i_1 B_1 \mathsf{e}_2\mathsf{e}_3 + i_2 B_2 \mathsf{e}_3\mathsf{e}_1 + i_3 B_3 \mathsf{e}_1\mathsf{e}_2. \tag{3.132}$$

The fact that for most bodies the principal moments are not equal demonstrates that $\mathcal{I}(B)$ will not lie in the same plane as B, unless B is perpendicular to one of the principal axes.

A useful result for calculating the inertia tensor is that the principal axes of a body always coincide with symmetry axes, if any are present. This simplifies the calculation of the inertia tensor for a range of standard bodies, the results for which can be found in some of the books listed at the end of this chapter.

3.4.6 Kinetic energy and angular momentum

To calculate the kinetic energy of the body from the velocity of equation (3.114) we form the integral

$$\begin{aligned} T &= \tfrac{1}{2}\int d^3x\,\rho(R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger + \boldsymbol{v}_0)^2 \\ &= \tfrac{1}{2}\int d^3x\,\rho\big((\boldsymbol{x}\cdot\Omega_B)^2 + 2\boldsymbol{v}_0\cdot(R\,\boldsymbol{x}\cdot\Omega_B\,R^\dagger) + \boldsymbol{v}_0^2\big) \\ &= \tfrac{1}{2}\int d^3x\,\rho(\boldsymbol{x}\cdot\Omega_B)^2 + M\boldsymbol{v}_0^2. \end{aligned} \tag{3.133}$$

Again, there is a clean split into a rotational contribution and a term due to the motion of the centre of mass. Concentrating on the former, we use the manipulation

$$(\boldsymbol{x}\cdot\Omega_B)^2 = \langle \boldsymbol{x}\cdot\Omega_B \boldsymbol{x}\Omega_B\rangle = -\Omega_B\cdot\big(\boldsymbol{x}\wedge(\boldsymbol{x}\cdot\Omega_B)\big) \tag{3.134}$$

to write the rotational contribution as

$$-\tfrac{1}{2}\Omega_B\cdot\Big(\int d^3x\,\rho\boldsymbol{x}\wedge(\boldsymbol{x}\cdot\Omega_B)\Big) = -\tfrac{1}{2}\Omega_B\cdot\mathcal{I}(\Omega_B). \tag{3.135}$$

The minus sign is to be expected because bivectors all have negative squares. The sign can be removed by reversing one of the bivectors to construct a positive-definite product. The total kinetic energy is therefore

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \tfrac{1}{2}\Omega_B^\dagger\cdot\mathcal{I}(\Omega_B). \tag{3.136}$$

The inertia tensor is constructed from the point of view of the fixed body. From equation (3.120) we see that the angular momentum in space is obtained by rotating the body angular momentum $\mathcal{I}(\Omega_B)$ onto the space configuration, that is,

$$L = R\mathcal{I}(\Omega_B)\,R^\dagger. \tag{3.137}$$

We can understand this expression as follows. Suppose that a body rotates in space with angular velocity Ω. At a given instant we carry out a fixed rotation to align everything back with the fixed reference configuration. This reference copy then has angular velocity $\Omega_B = R^\dagger\Omega R$. The inertia tensor (fixed in the reference copy) returns the angular momentum, given an input angular velocity. The result of this is then rotated forwards onto the body in space, to return L.

The space and body angular velocities are related by $\Omega = R\Omega_B R^\dagger$, so the kinetic energy can be written in the form

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \tfrac{1}{2}\Omega^\dagger\cdot L. \tag{3.138}$$

We now introduce components $\{\omega_k\}$ for both Ω and Ω_B by writing

$$\Omega = \sum_{k=1}^3 \omega_k I\mathsf{f}_k, \quad \Omega_B = \sum_{k=1}^3 \omega_k I\mathsf{e}_k. \tag{3.139}$$

In terms of these we recover the standard expression

$$T = \tfrac{1}{2}M\boldsymbol{v}_0^2 + \sum_{k=1}^3 \tfrac{1}{2}i_k\omega_k^2. \tag{3.140}$$

3.4.7 Equations of motion

The equations of motion are $\dot{L} = N$, where N is the external torque. The inertia tensor is time-independent since it only refers to the static 'reference' copy of the rigid body, so we find that

$$\begin{aligned}\dot{L} &= \dot{R}\mathcal{I}(\Omega_B)R^\dagger + R\mathcal{I}(\Omega_B)\dot{R}^\dagger + R\mathcal{I}(\dot{\Omega}_B)R^\dagger \\ &= R\big(\mathcal{I}(\dot{\Omega}_B) - \tfrac{1}{2}\Omega_B\mathcal{I}(\Omega_B) + \tfrac{1}{2}\mathcal{I}(\Omega_B)\Omega_B\big)R^\dagger. \end{aligned}\tag{3.141}$$

At this point it is extremely useful to have a symbol to denote one-half of the commutator of two bivectors. The standard symbol for this is the cross, $\times$, so we define the *commutator product* by

$$A\times B = \tfrac{1}{2}(AB - BA). \tag{3.142}$$

This notation does raise the possibility of confusion with the vector cross product, but as the latter is not needed any more this should not pose a problem. The commutator product is so ubiquitous in applications that it needs its own symbol, and the cross is particularly convenient as it correctly conveys the antisymmetry of the product. In section 4.1.3 we prove that the commutator of any two bivectors results in a third bivector. This is easily confirmed in three dimensions by expressing both bivectors in terms of their dual vectors.

With the commutator product at our disposal the equations of motion are now written concisely as

$$\dot{L} = R\big(\mathcal{I}(\dot{\Omega}_B) - \Omega_B \times \mathcal{I}(\Omega_B)\big)R^\dagger. \tag{3.143}$$

The typical form of the rigid-body equations is recovered by expanding in terms of components. In terms of these we have

$$\begin{aligned}\dot{L} &= R\left(\sum_{k=1}^{3} i_k \dot{\omega}_k I\mathsf{e}_k - \sum_{j,k=1}^{3} i_k \omega_j \omega_k (I\mathsf{e}_j)\times(I\mathsf{e}_k)\right)R^\dagger \\ &= \sum_{k=1}^{3} \dot{\omega}_k I\mathsf{f}_k + \sum_{j,k,l=1}^{3} \epsilon_{jkl} i_k \omega_j \omega_k I\mathsf{f}_l.\end{aligned} \tag{3.144}$$

If we let N_k denote the components of the torque N in the rotating f_k frame,

$$N = \sum_{k=1}^{3} N_k I\mathsf{f}_k, \tag{3.145}$$

we recover the Euler equations of motion for a rigid body:

$$\begin{aligned} i_1\dot{\omega}_1 - \omega_2\omega_3(i_2 - i_3) &= N_1, \\ i_2\dot{\omega}_2 - \omega_3\omega_1(i_3 - i_1) &= N_2, \\ i_3\dot{\omega}_3 - \omega_1\omega_2(i_1 - i_2) &= N_3. \end{aligned} \tag{3.146}$$

Various methods can be used to solve these equations and are described in most mechanics textbooks. Here we will simply illustrate some features of the equations, and describe a solution method which does not resort to the explicit coordinate equations.

3.4.8 Torque-free motion

The torque-free equation $\dot{L} = 0$ reduces to

$$\mathcal{I}(\dot{\Omega}_B) - \Omega_B \times \mathcal{I}(\Omega_B) = 0. \tag{3.147}$$

This is a first-order constant coefficient differential equation for the bivector Ω_B. Closed form solutions exist, but before discussing some of these it is useful to consider the conserved quantities. Throughout this section we ignore any overall motion of the centre of mass of the rigid body. Since $\dot{L} = 0$ both the kinetic energy and the magnitude of L are constant. To exploit this we introduce the components

$$L_k = i_k \omega_k, \quad L = \sum_{k=1}^{3} L_k I \mathsf{f}_k. \tag{3.148}$$

These are the components of L in the rotating f_k frame. So, even though L is constant, the components L_k are time-dependent. In terms of these components the magnitude of L is

$$LL^\dagger = L_1^2 + L_2^2 + L_3^3 \tag{3.149}$$

and the kinetic energy is

$$T = \frac{L_1^2}{2i_1} + \frac{L_2^2}{2i_2} + \frac{L_3^2}{2i_3}. \tag{3.150}$$

Both $|L|$ and T are constants of motion, which imposes two constraints on the three components L_k. A useful way to visualise this is to think in terms of a vector $\boldsymbol{l}$ with components L_k:

$$\boldsymbol{l} = \sum_{k=1}^{3} L_k \mathsf{e}_k = -IR^\dagger L R. \tag{3.151}$$

This is the vector perpendicular to $R^\dagger L R$ — a rotating vector in the fixed reference body. Conservation of $|L|$ means that $\boldsymbol{l}$ is constrained to lie on a sphere, and conservation of T restricts $\boldsymbol{l}$ to the surface of an ellipsoid. Possible paths for $\boldsymbol{l}$ for a given rigid body are therefore defined by the intersections of a sphere with a family of ellipsoids (governed by T). For the case of unequal principal moments these orbits are non-degenerate. Examples of these orbits are shown in figure 3.7. This figure shows that orbits around the axes with the smallest and largest principal moments are stable, whereas around the middle axis the orbits are unstable. Any small change in the energy of the body will tend to throw it into a very different orbit if the orbit of $\boldsymbol{l}$ approaches close to e_2.

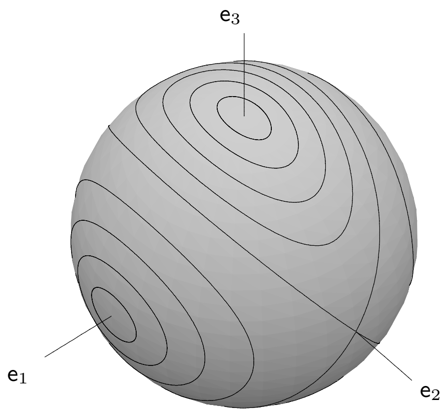

Figure 3.7 *Angular momentum orbits.* The point described by the vector $\boldsymbol{l}$ simultaneously lies on the surface of a sphere and an ellipse. The figure shows possible paths on the sphere for $\boldsymbol{l}$ in the case of $i_1 < i_2 < i_3$, with the 3 axis vertical.

3.4.9 The symmetric top

The full analytic solution for torque-free motion is complicated and requires elliptic functions. If the body has a single symmetry axis, however, the solution is quite straightforward. In this case the body has two equal moments of inertia, $i_1 = i_2$, and the third principal moment i_3 is assumed to be different. With this assignment e_3 is the symmetry axis of the body. The action of the inertia tensor on Ω_B is

$$\begin{aligned}\mathcal{I}(\Omega_B) &= i_1\omega_1\mathsf{e}_2\mathsf{e}_3 + i_1\omega_2\mathsf{e}_3\mathsf{e}_1 + i_3\omega_3\mathsf{e}_1\mathsf{e}_2 \\ &= i_1\Omega_B + (i_3 - i_1)\omega_3 I\mathsf{e}_3,\end{aligned} \tag{3.152}$$

so we can write $\mathcal{I}(\Omega_B)$ in the compact form

$$\mathcal{I}(\Omega_B) = i_1\Omega_B + (i_3 - i_1)(\Omega_B\wedge\mathsf{e}_3)\mathsf{e}_3. \tag{3.153}$$

(This type of expression offers many advantages over the alternative 'dyad' notation.) The torque-free equations of motion are now

$$\mathcal{I}(\dot{\Omega}_B) = \Omega_B\times\mathcal{I}(\Omega_B) = (i_3 - i_1)\Omega_B\times\big((\Omega_B\wedge\mathsf{e}_3)\mathsf{e}_3\big). \tag{3.154}$$

Since $\Omega_B\wedge\mathsf{e}_3$ is a trivector, we can dualise the final term and write

$$\mathcal{I}(\dot{\Omega}_B) = -(i_3 - i_1)\mathsf{e}_3\wedge\big((\Omega_B\wedge\mathsf{e}_3)\Omega_B\big). \tag{3.155}$$

It follows that

$$\mathsf{e}_3\wedge\mathcal{I}(\dot{\Omega}_B) = 0 = i_3\dot{\omega}_3 I, \tag{3.156}$$

which shows that ω_3 is a constant. This result can be read off directly from the Euler equations, but it is useful to see how it can be derived without dropping down to the individual component equations. The ability to do this becomes ever more valuable as the complexity of the equations increases.

Next we use the result that

$$\begin{aligned} i_1\Omega_B &= \mathcal{I}(\Omega_B) - (i_3 - i_1)(\Omega_B \wedge \mathsf{e}_3)\mathsf{e}_3 \\ &= \mathcal{I}(\Omega_B) + (i_1 - i_3)\omega_3 I\mathsf{e}_3 \end{aligned} \tag{3.157}$$

to write

$$\Omega = R\Omega_B R^\dagger = \frac{1}{i_1}L + \frac{i_1 - i_3}{i_1}\omega_3 R I\mathsf{e}_3 R^\dagger. \tag{3.158}$$

Our rotor equation now becomes

$$\dot{R} = -\tfrac{1}{2}\Omega R = -\frac{1}{2i_1}(LR + R(i_1 - i_3)\omega_3 I\mathsf{e}_3). \tag{3.159}$$

The right-hand side of this equation involves two constant bivectors, one multiplying R to the left and the other to the right. We therefore define the two bivectors

$$\Omega_l = \frac{1}{i_1}L, \qquad \Omega_r = \omega_3\frac{i_1 - i_3}{i_1}I\mathsf{e}_3, \tag{3.160}$$

so that the rotor equation becomes

$$\dot{R} = -\tfrac{1}{2}\Omega_l R - \tfrac{1}{2}R\Omega_r. \tag{3.161}$$

This equation integrates immediately to give

$$R(t) = \exp(-\tfrac{1}{2}\Omega_l t)R_0 \exp(-\tfrac{1}{2}\Omega_r t). \tag{3.162}$$

This fully describes the motion of a symmetric top. It shows that there is an 'internal' rotation in the $\mathsf{e}_1\mathsf{e}_2$ plane (the symmetry plane of the body). This is responsible for the precession of a symmetric top. The constant rotor R_0 defines the attitude of the rigid body at $t = 0$ and can be set to 1. The resultant body is then rotated in the plane of its angular momentum to obtain the final attitude in space.

3.5 Notes

Much of this chapter follows *New Foundations for Classical Mechanics* by David Hestenes (1999), which gives a comprehensive account of the applications to classical mechanics of geometric algebra in three dimensions. Readers are encouraged to compare the techniques used in this chapter with more traditional methods, a good description of which can be found in *Classical Mechanics* by Goldstein (1950), or *Analytical Mechanics* by Hand & Finch (1998). The standard reference for the Kustaanheimo–Stiefel equation is *Linear and Regular Celestial Mechanics* by Stiefel and Scheifele (1971). Many authors have explored this technique, particularly in the quaternionic framework. These include Hestenes' 'Celestial mechanics with geometric algebra' (1983) and the papers by Aramanovitch (1995) and Vrbik (1994, 1995).

3.6 Exercises

3.1 An elliptical orbit in an inverse-square force law is parameterised in terms of a scalar + pseudoscalar quantity U by $\boldsymbol{x} = U^2 \mathsf{e}_1$. Prove that U can be written

$$U = A_0 \mathrm{e}^{I\omega s} + B_0 \mathrm{e}^{-I\omega s},$$

where $dt/ds = r$, $r = |\boldsymbol{x}| = UU^\dagger$ and I is the unit bivector for the plane. What is the value of ω? Find the conditions on A_0 and B_0 such that at time $t = 0$, $s = 0$ and the particle lies on the positive e_1 axis with velocity in the positive e_2 direction. For which value of s does the velocity point in the $-\mathsf{e}_1$ direction? Find the values for the shortest and longest diameters of the ellipse, and verify that we can write

$$U = \sqrt{a(1+e)}\cos(\omega s) - \sqrt{a(1-e)} I \sin(\omega s),$$

where e is the eccentricity and a is the semi-major axis.

3.2 For elliptical orbits the semi-major axis a is defined by $a = \frac{1}{2}(r_1 + r_2)$, where r_1 and r_2 are the distances of closest and furthest approach. Prove that

$$\frac{l^2}{k\mu} = a(1 - e^2).$$

Hence show that we can write

$$r = \frac{a(1-e^2)}{1 + e\cos(\theta)},$$

where $e\cos(\theta) = \boldsymbol{e}\cdot\hat{\boldsymbol{x}}$. The eccentricity vector points to the point of closest approach. Why would we expect the orbital average of $\hat{\boldsymbol{x}}/r^4$ to also point in this direction? Prove that

$$\int_0^T dt \frac{\hat{\boldsymbol{x}}}{r^4} = \hat{\boldsymbol{e}} \frac{\mu}{la^2(1-e^2)^2} \int_0^{2\pi} \bigl(1 + e\cos(\theta)\bigr)^2 \cos(\theta)\, d\theta$$

and evaluate the integral.

3.3 A particle in three dimensions moves along a curve $\boldsymbol{x}(t)$ such that $|\boldsymbol{v}|$ is constant. Show that there exists a bivector Ω such that

$$\dot{\boldsymbol{v}} = \Omega\cdot\boldsymbol{v},$$

and give an explicit formula for Ω. Is this bivector unique?

3.4 Suppose that we measure components of the position vector $\boldsymbol{x}$ in a rotating frame $\{\mathsf{f}_i\}$. By referring this frame to a fixed frame, show that the components of $\boldsymbol{x}$ are given by

$$x_i = \mathsf{e}_i \cdot (R^\dagger \boldsymbol{x} R).$$

By differentiating this expression twice, prove that we can write

$$\mathsf{f}_i \ddot{x}_i = \ddot{\boldsymbol{x}} + \Omega \cdot (\Omega \cdot \boldsymbol{x}) + 2\Omega \cdot \dot{\boldsymbol{x}} + \dot{\Omega} \cdot \boldsymbol{x}.$$

Hence deduce expressions for the centrifugal, Coriolis and Euler forces in terms of the angular velocity bivector Ω.

3.5 Show that the inertia tensor satisfies the following properties:

$$\begin{aligned} \text{linearity:} \qquad & \mathcal{I}(\lambda A + \mu B) = \lambda \mathcal{I}(A) + \mu \mathcal{I}(B) \\ \text{symmetry:} \qquad & \langle A \mathcal{I}(B) \rangle = \langle \mathcal{I}(A) B \rangle. \end{aligned}$$

3.6 Prove that the inertia tensor $\mathcal{I}(B)$ for a solid cylinder of height h and radius a can be written

$$\mathcal{I}(B) = \frac{Mh^2}{12}(B - B \wedge \mathsf{e}_3\, \mathsf{e}_3) + \frac{Ma^2}{4}(B + B \wedge \mathsf{e}_3\, \mathsf{e}_3),$$

where e_3 is the symmetry axis.

3.7 For a torque-free symmetric top prove that the angular momentum, viewed back in the reference copy, rotates around the symmetry axis at an angular frequency ω, where

$$\omega = \omega_3 \frac{i_3 - i_1}{i_1}.$$

Show that the angle between the symmetry axis and the vector $\boldsymbol{l} = -IL$ is given by

$$\cos(\theta) = \frac{i_3 \omega}{l},$$

where $l^2 = \boldsymbol{l}^2 = LL^\dagger$. Hence show that the symmetry axis rotates in space in the L plane at an angular frequency ω', where

$$\omega' = \frac{i_3 \omega_3}{i_1 \cos(\theta)}.$$

4

Foundations of geometric algebra

In chapter 2 we introduced geometric algebra in two and three dimensions. We now turn to a discussion of the full, axiomatic framework for geometric algebra in arbitrary dimensions, with arbitrary signature. This will involve some duplication of material from chapter 2, but we hope that this will help reinforce some of the key concepts. Much of the material in this chapter is of primary relevance to those interested in the full range of applications of geometric algebra. Those interested solely in applications to space and spacetime may want to skip some of the material below, as both of these algebras are treated in a self-contained manner in chapters 2 and 5 respectively. The material on frames and linear algebra is important, however, and a knowledge of this is assumed for applications in gravitation.

The fact that geometric algebra can be applied in spaces of arbitrary dimensions is crucial to the claim that it is a mathematical tool of universal applicability. The framework developed here will enable us to extend geometric algebra to the study of relativistic dynamics, phase space, single and multiparticle quantum theory, Lie groups and manifolds. This chapter also highlights some of the new algebraic techniques we now have at our disposal. Many derivations can be simplified through judicious use of the geometric product at various intermediate steps. This is true even if the initial and final expressions contain only inner and outer products.

Many key relations in physics involve linear mappings between one space and another. In this chapter we also explore how geometric algebra simplifies the rich subject of linear transformations. We start with simple mappings between vectors in the same space and study their properties in a very general, basis-free framework. In later chapters this framework is extended to encompass functions between different spaces, and multilinear functions where the argument of the function can consist of one or more multivectors.