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.

 https://doi.org/10.1017/CBO9780511807497.006

4.1 Axiomatic development

We should now have an intuitive feel for the elements of a geometric algebra — the multivectors — and some of their multiplicative properties. The next step is to define a set of axioms and conventions which enable us to efficiently manipulate them. Geometric algebra can be defined using a number of axiomatic frameworks, all of which give rise to the same final algebra. In the main we will follow the approach first developed by Hestenes and Sobczyk and raise the geometric product to primary status in the algebra. The properties of the inner and outer products are then inherited from the full geometric product, and this simplifies proofs of a number of important results.

Our starting point is the vector space from which the entire algebra will be generated. Vectors (i.e. grade-1 multivectors) have a special status in the algebra, as the grading of the algebra is determined by them. Three main axioms govern the properties of the geometric product for vectors.

(i) The geometric product is associative:

$$a(bc) = (ab)c = abc. \tag{4.1}$$

(ii) The geometric product is distributive over addition:

$$a(b + c) = ab + ac. \tag{4.2}$$

(iii) The square of any vector is a real scalar: $a^2 \in \Re$.

The final axiom is the key one which distinguishes a geometric algebra from a general associative algebra. We do not force the scalar to be positive, so we can incorporate Minkowski spacetime without modification of our axioms. Nothing is assumed about the commutation properties of the geometric product — matrix multiplication is one picture to keep in mind. Indeed, one can always represent the geometric product in terms of products of suitably chosen matrices, but this does not bring any new insights into the properties of the geometric product.

By successively multiplying together vectors we generate the complete algebra. Elements of this algebra are called multivectors and are usually written in upper-case italic font. The space of multivectors is *linear over the real numbers*, so if λ and μ are scalars and A and B are multivectors $\lambda A + \mu B$ is also a multivector. We only consider the algebra over the reals as most occurrences of complex numbers in physics turn out to have a geometric origin. This geometric meaning can be lost if we admit a scalar unit imaginary. Any multivector can be written as a sum of geometric products of vectors. They too can be multiplied using the geometric product and this product inherits properties (i) and (ii) above. So, for multivectors A, B and C, we have

$$(AB)C = A(BC) = ABC \tag{4.3}$$

and

$$A(B+C) = AB + AC. \tag{4.4}$$

If we now form the square of the vector $a+b$ we find that

$$(a+b)^2 = (a+b)(a+b) = a^2 + ab + ba + b^2. \tag{4.5}$$

It follows that the symmetrised product of two vectors can be written

$$ab + ba = (a+b)^2 - a^2 - b^2, \tag{4.6}$$

and so must also be a scalar, by axiom (iii). We therefore *define* the inner product for vectors by

$$a \cdot b = \frac{1}{2}(ab + ba). \tag{4.7}$$

The remaining, antisymmetric part of the geometric product is defined as the exterior product and returns a bivector,

$$a \wedge b = \frac{1}{2}(ab - ba). \tag{4.8}$$

These definitions combine to give the familiar result

$$ab = a \cdot b + a \wedge b. \tag{4.9}$$

In forming this decomposition we have defined both the inner and outer products of vectors in terms of the geometric product. This contrasts with the common alternative of defining the geometric product in terms of separate inner and outer products. Some authors prefer this alternative because the (less familiar) geometric product is defined in terms of more familiar objects. The main drawback, however, is that work still remains to establish the main properties of the geometric product. In particular, it is far from obvious that the product is associative, which is invaluable for its use.

4.1.1 The outer product, grading and bases

In the preceding we defined the outer product of two vectors and asserted that this returns a bivector (a grade-2 multivector). This is the key to defining the grade operation for the entire algebra. To do this we first extend the definition of the outer product to arbitrary numbers of vectors. The outer (exterior) product of the vectors $a_1, \ldots, a_r$ is denoted by $a_1 \wedge a_2 \wedge \cdots \wedge a_r$ and is defined as the totally antisymmetrised sum of all geometric products:

$$a_1 \wedge a_2 \wedge \cdots \wedge a_r = \frac{1}{r!} \sum (-1)^{\epsilon} a_{k_1} a_{k_2} \cdots a_{k_r}. \tag{4.10}$$

The sum runs over every permutation $k_1, \dots, k_r$ of $1, \dots, r$, and $(-1)^\epsilon$ is $+1$ or -1 as the permutation $k_1, \dots, k_r$ is even or odd respectively. So, for example,

$$a_1 \wedge a_2 = \frac{1}{2!}(a_1 a_2 - a_2 a_1) \tag{4.11}$$

as required.

The antisymmetry of the outer product ensures that it vanishes if any two vectors are the same. It follows that the outer product vanishes if the vectors are linearly dependent, since in this case one vector can be written as a linear combination of the remaining vectors. The outer product therefore records the dimensionality of the object formed from a set of vectors. This is precisely what we mean by *grade*, so we define the outer product of r vectors as having grade r. Any multivector which can be written purely as the outer product of a set of vectors is called a *blade*. Any multivector can be expressed as a sum of blades, as can be verified by introducing an explicit basis. These blades all have definite grade and in turn define the grade or grades of the multivector.

We rarely need the full antisymmetrised expression when studying blades. Instead we can employ the result that *every blade can be written as a geometric product of orthogonal, anticommuting vectors*. The anticommutation of orthogonal vectors then takes care of the antisymmetry of the product. In Euclidean space this result is simple to prove using a form of Gram–Schmidt orthogonalisation. Given two vectors a and b we form

$$b' = b - \lambda a. \tag{4.12}$$

We then see that

$$a \wedge (b - \lambda a) = a \wedge b - \lambda a \wedge a = a \wedge b. \tag{4.13}$$

So the same bivector is obtained, whatever the value of λ (figure 4.1). The bivector encodes an oriented plane with magnitude determined by the area. Interchanging b and b' changes neither the orientation nor the magnitude, so returns the same bivector. We now form

$$a \cdot b' = a \cdot (b - \lambda a) = a \cdot b - \lambda a^2. \tag{4.14}$$

So if we set $\lambda = a \cdot b / a^2$ we have $a \cdot b' = 0$ and can write

$$a \wedge b = a \wedge b' = a b'. \tag{4.15}$$

One can continue in this manner and construct a complete set of orthogonal vectors generating the same outer product.

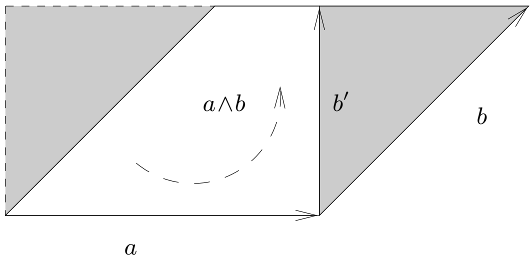

Figure 4.1 *The Gram–Schmidt process.* The outer product $a \wedge b$ is independent of shape of the parallelogram formed by a and b. The only information contained in $a \wedge b$ is the oriented plane and a magnitude. The vectors b and b' generate the same bivector, so we can choose b' orthogonal to a and write $a \wedge b = ab'$.

An alternative form for b' is quite revealing. We write

$$\begin{aligned} b' &= b - a^{-1} a \cdot b \\ &= a^{-1}(ab - a \cdot b) \\ &= a^{-1}(a \wedge b). \end{aligned} \tag{4.16}$$

This shows that b' is formed by rotating a through $90°$ in the $a \wedge b$ plane, and dilating by the appropriate amount. The algebraic form also makes it clear why $ab' = a \wedge b$, and gives a formula that extends simply to higher grades.

The above argument is fine for Euclidean space, but breaks down for spaces of mixed signature. The inverse $a^{-1} = a/a^2$ is not defined when a is null ($a^2 = 0$), so an alternative procedure is required. Fortunately this is a relatively straightforward exercise. We start with the set of r independent vectors $a_1, \ldots, a_r$ and form the $r \times r$ symmetric matrix

$$\mathsf{M}_{ij} = a_i \cdot a_j. \tag{4.17}$$

The symmetry of this matrix ensures that it can always be diagonalised with an orthogonal matrix R_{ij},

$$\mathsf{R}_{ik} \mathsf{M}_{kl} \mathsf{R}^{\mathrm{t}}_{lj} = \mathsf{R}_{ik} \mathsf{R}_{jl} \mathsf{M}_{kl} = \Lambda_{ij}. \tag{4.18}$$

Here Λ_{ij} is diagonal and, unless stated otherwise, the summation convention is employed. The matrix R_{ij} defines a new set of vectors via

$$e_i = \mathsf{R}_{ij} a_j. \tag{4.19}$$

These satisfy

$$\begin{aligned} e_i \cdot e_j &= (\mathsf{R}_{ik} a_k) \cdot (\mathsf{R}_{jl} a_l) \\ &= \mathsf{R}_{ik} \mathsf{R}_{jl} \mathsf{M}_{kl} \\ &= \Lambda_{ij}. \end{aligned} \tag{4.20}$$

The vectors $e_1, \ldots, e_r$ are therefore orthogonal and hence all anticommute. Their geometric product is therefore totally antisymmetric, and we have

$$\begin{aligned} e_1 e_2 \cdots e_r &= e_1 \wedge \cdots \wedge e_r \\ &= (\mathsf{R}_{1i} a_i) \wedge \cdots (\mathsf{R}_{rk} a_k) \\ &= \det\,(\mathsf{R}_{ij})\, a_1 \wedge a_2 \wedge \cdots \wedge a_r. \end{aligned} \tag{4.21}$$

The determinant appears here because of the total antisymmetry of the expression (see section 4.5.2). But since R_{ij} is an orthogonal matrix it has determinant ± 1, and by choosing the order of the $\{e_i\}$ vectors appropriately we can set the determinant of R_{ij} to 1. This ensures that we can always find a set of vectors such that

$$a_1 \wedge a_2 \wedge \cdots \wedge a_r = e_1 e_2 \cdots e_r. \tag{4.22}$$

This result will simplify the proofs of a number of results in this chapter.

For a given vector space, an orthonormal frame $\{\mathsf{e}_i\}, i = 1, \ldots, n$ provides a natural way to view the entire geometric algebra. We denote this algebra $\mathcal{G}_n$. Most of the results derived in this chapter are independent of signature, so in the following we let $\mathcal{G}_n$ denote the geometric algebra of a space of dimension n with arbitrary (non-degenerate) signature. One can also consider the degenerate case where some of the basis vectors are null, though we will not need such algebras in this book. The basis vectors build up to form a basis for the entire algebra as

$$1, \quad \mathsf{e}_i, \quad \mathsf{e}_i \mathsf{e}_j \ (i < j), \quad \mathsf{e}_i \mathsf{e}_j \mathsf{e}_k \ (i < j < k), \qquad \ldots. \tag{4.23}$$

The fact that the basis vectors anticommute ensures that each product in the basis set is totally antisymmetric. The product of r distinct basis vectors is then, by definition, a grade-r multivector. The basis (4.23) therefore naturally defines a basis for each of the grade-r subspaces of $\mathcal{G}_n$. We denote each of these subspaces by $\mathcal{G}_n^r$. The size of each subspace is given by the number of distinct combinations of r objects from a set of n. (The order is irrelevant, because of the total antisymmetry.) These are given by the binomial coefficients, so

$$\dim\big(\mathcal{G}_n^r\big) = \binom{n}{r}. \tag{4.24}$$

For example, we have already seen that in two dimensions the algebra contains terms of grade $0, 1, 2$ with each space having dimension $1, 2, 1$ respectively. Similarly in three dimensions the separate graded subspaces have dimension $1, 3, 3, 1$. The binomial coefficients always exhibit a mirror symmetry between the r and $n - r$ terms. This gives rise to the notion of duality, which is explained in section 4.1.4 where we explore the properties of the highest grade element of the algebra — the pseudoscalar.

The total dimension of the algebra is

$$\dim(\mathcal{G}_n) = \sum_{r=0}^{n} \dim(\mathcal{G}_n^r) = \sum_{r=0}^{n} \binom{n}{r} = (1+1)^n = 2^n. \tag{4.25}$$

One can see that the total size of the algebra quickly becomes very large. If one wanted to find a matrix representation of the algebra, the matrices would have to be of the order of $2^{n/2} \times 2^{n/2}$. For all but the lowest values of n these matrices become totally impractical for computations. This is one reason why matrix representations do not help much with understanding and using geometric algebra.

We have now defined the grade operation for our linear space $\mathcal{G}_n$. An arbitrary multivector A can be decomposed into a sum of pure grade terms

$$A = \langle A\rangle_0 + \langle A\rangle_1 + \cdots = \sum_r \langle A\rangle_r. \tag{4.26}$$

The operator $\langle\ \rangle_r$ projects onto the grade-r terms in the argument, so $\langle A\rangle_r$ returns the grade-r components in A. Multivectors containing terms of only one grade are called *homogeneous*. They are often written as A_r, so

$$\langle A_r\rangle_r = A_r. \tag{4.27}$$

Take care not to confuse the grading subscript in A_r with frame indices in expressions like $\{\mathsf{e}_k\}$. The context should always make clear which is intended. The grade-0 terms in $\mathcal{G}_n$ are the real scalars and commute with all other elements. We continue to employ the useful abbreviation

$$\langle A\rangle = \langle A\rangle_0 \tag{4.28}$$

for the operation of taking the scalar part.

An important feature of a geometric algebra is that *not all homogeneous multivectors are pure blades.* This is confusing at first, because we have to go to four dimensions before we reach our first counterexample. Suppose that $\{\mathsf{e}_1, \ldots, \mathsf{e}_4\}$ form an orthonormal basis for the Euclidean algebra $\mathcal{G}_4$. There are six independent basis bivectors in this algebra, and from these we can construct terms like

$$B = \alpha \mathsf{e}_1 \wedge \mathsf{e}_2 + \beta \mathsf{e}_3 \wedge \mathsf{e}_4, \tag{4.29}$$

where α and β are scalars. B is a pure bivector, so is homogeneous, but it cannot be reduced to a blade. That is, we cannot find two vectors a and b such that $B = a\wedge b$. The reason is that $\mathsf{e}_1\wedge\mathsf{e}_2$ and $\mathsf{e}_3\wedge\mathsf{e}_4$ do not share a common vector. This is not possible in three dimensions, because any two planes with a common origin share a common line. A four-dimensional bivector like B is therefore hard for us to visualise. There is a way to visualise B in three dimensions, however, and it is provided by *projective geometry.* This is described in chapter 10.

4.1.2 Further properties of the geometric product

The decomposition of the geometric product of two vectors into a scalar term and a bivector term has a natural extension to general multivectors. To establish the results of this section we make repeated use of the formula

$$ab = 2a \cdot b - ba \tag{4.30}$$

which we use to reorder expressions. As a first example, consider the case of a geometric product of vectors. We find that

$$\begin{aligned} aa_1a_2 \cdots a_r &= 2a \cdot a_1 \, a_2 \cdots a_r - a_1 a a_2 \cdots a_r \\ &= 2a \cdot a_1 \, a_2 \cdots a_r - 2a \cdot a_2 \, a_1 a_3 \cdots a_r + a_1 a_2 a a_3 \cdots a_r \\ &= 2 \sum_{k=1}^{r} (-1)^{k+1} a \cdot a_k \, a_1 a_2 \cdots \check{a}_k \cdots a_r + (-1)^r a_1 a_2 \cdots a_r a, \end{aligned} \tag{4.31}$$

where the check on $\check{a}_k$ denotes that this term is missing from the series. We continue to follow the conventions introduced in chapter 2 so, in the absence of brackets, inner products are performed before outer products, and both are performed before geometric products.

Suppose now that the vectors $a_1, \ldots, a_r$ are replaced by a set of anticommuting vectors $e_1, \ldots, e_r$. We find that

$$\frac{1}{2}\Big(ae_1e_2 \cdots e_r - (-1)^r e_1 e_2 \cdots e_r a\Big) = \sum_{k=1}^{r} (-1)^{k+1} a \cdot e_k \, e_1 e_2 \cdots \check{e}_k \cdots e_r. \tag{4.32}$$

The right-hand side contains a sum of terms formed from the product of $r-1$ anticommuting vectors, so has grade $r-1$. Since any grade-r multivector can be written as a sum of terms formed from anticommuting vectors, the combination on the left-hand side will always return a multivector of grade $r-1$. We therefore define the inner product between a vector a and a grade-r multivector A_r by

$$a \cdot A_r = \frac{1}{2}\Big(aA_r - (-1)^r A_r a\Big). \tag{4.33}$$

The inner product of a vector and a grade-r multivector results in a multivector with grade reduced by one.

The main work of this section is in establishing the properties of the remaining part of the product aA_r. For the case where A_r is a vector, the remaining term is the antisymmetric product, and so is a bivector. This turns out to be true in general — the remaining part of the geometric product returns the exterior product,

$$\frac{1}{2}\Big(a(a_1 \wedge a_2 \wedge \cdots \wedge a_r) + (-1)^r (a_1 \wedge a_2 \wedge \cdots \wedge a_r) a\Big) = a \wedge a_1 \wedge a_2 \wedge \cdots \wedge a_r. \tag{4.34}$$

We will prove this important result by induction. First, we write the blade as a geometric product of anticommuting vectors, so that the result we will establish becomes

$$\frac{1}{2}\Big(ae_1e_2\cdots e_r + (-1)^r e_1e_2\cdots e_r a\Big) = a\wedge e_1\wedge e_2\wedge\cdots\wedge e_r. \tag{4.35}$$

For $r = 1$ the result is true as the right-hand side defines the bivector $a \wedge e_1$. For $r > 1$ we proceed by writing

$$\begin{aligned} a\wedge e_1\wedge e_2\wedge\cdots\wedge e_r = {} & \frac{1}{r+1} ae_1e_2\cdots e_r \\ & + \frac{1}{r+1}\sum_{k=1}^{r}(-1)^k e_k(a\wedge e_1\wedge\cdots\wedge\check{e}_k\wedge\cdots\wedge e_r). \end{aligned} \tag{4.36}$$

This result is easily established by writing out all terms in the full antisymmetric product and gathering together the terms which start with the same vector. Next we assume that equation (4.35) holds for the case of an $r-1$ blade, and expand the term inside the sum as follows:

$$\begin{aligned} \sum_{k=1}^{r} & (-1)^k e_k(a\wedge e_1\wedge\cdots\wedge\check{e}_k\wedge\cdots\wedge e_r) \\ &= \frac{1}{2}\sum_{k=1}^{r}(-1)^k e_k\Big(ae_1\cdots\check{e}_k\cdots e_r + (-1)^{r-1}e_1\cdots\check{e}_k\cdots e_r a\Big) \\ &= \frac{1}{2}\sum_{k=1}^{r}(-1)^k e_k a e_1\cdots\check{e}_k\cdots e_r + \frac{r}{2}(-1)^r e_1\cdots e_r a \\ &= \sum_{k=1}^{r}(-1)^k (e_k\cdot a) e_1\cdots\check{e}_k\cdots e_r + \frac{r}{2}\Big(ae_1\cdots e_r + (-1)^r e_1\cdots e_r a\Big) \\ &= \frac{r-1}{2} ae_1\cdots e_r + \frac{r+1}{2}(-1)^r e_1\cdots e_r a, \end{aligned} \tag{4.37}$$

where we have used equation (4.32). Substituting this result into equation (4.36) then proves equation (4.35) for a grade-r blade, assuming it is true for a blade of grade $r-1$. Since the result is already established for $r = 1$, equation (4.34) holds for all blades and hence all multivectors.

We extend the definition of the wedge symbol by writing

$$a\wedge A_r = \frac{1}{2}\Big(aA_r + (-1)^r A_r a\Big). \tag{4.38}$$

With this definition we now have

$$aA_r = a\cdot A_r + a\wedge A_r, \tag{4.39}$$

which extends the decomposition of the geometric product in precisely the desired way. In equation (4.38) one can see how the geometric product can simplify many calculations. The left-hand side would, in general, require totally antisymmetrising all possible products. But the right-hand side only requires evaluating two products — an enormous saving! As we have established the grades of the separate inner and outer products, we also have

$$aA_r = \langle aA_r\rangle_{r-1} + \langle aA_r\rangle_{r+1}, \tag{4.40}$$

where

$$a\cdot A_r = \langle aA_r\rangle_{r-1}, \quad a\wedge A_r = \langle aA_r\rangle_{r+1}. \tag{4.41}$$

So, as expected, multiplication by a vector raises and lowers the grade of a multivector by 1.

A homogeneous multivector can be written as a sum of blades, and each blade can be written as a geometric product of anticommuting vectors. Applying the preceding decomposition, we establish that the product of two homogeneous multivectors decomposes as

$$A_rB_s = \langle A_rB_s\rangle_{|r-s|} + \langle A_rB_s\rangle_{|r-s|+2} + \cdots + \langle A_rB_s\rangle_{r+s}. \tag{4.42}$$

We retain the $\cdot$ and $\wedge$ symbols for the lowest and highest grade terms in this series:

$$\begin{aligned} A_r\cdot B_s &= \langle A_rB_s\rangle_{|r-s|}, \\ A_r\wedge B_s &= \langle A_rB_s\rangle_{r+s}. \end{aligned} \tag{4.43}$$

This is the most general use of the wedge symbol, and is consistent with the earlier definition as the antisymmetrised product of a set of vectors. We can check that the outer product is associative by forming

$$(A_r\wedge B_s)\wedge C_t = \langle A_rB_s\rangle_{r+s}\wedge C_t = \langle (A_rB_s)C_t\rangle_{r+s+t}. \tag{4.44}$$

Associativity of the outer product then follows from the fact that the geometric product is associative:

$$\langle (A_rB_s)C_t\rangle_{r+s+t} = \langle A_rB_sC_t\rangle_{r+s+t} = A_r\wedge B_s\wedge C_t. \tag{4.45}$$

In equation (4.32) we established a formula for the result for the inner product of a vector and a blade formed from orthogonal vectors. We now extend this to a more general result that is extremely useful in practice. We start by writing

$$a\cdot(a_1\wedge a_2\wedge\cdots\wedge a_r) = a\cdot\langle a_1a_2\cdots a_r\rangle_r, \tag{4.46}$$

where $a_1,\ldots,a_r$ are a general set of vectors. The geometric product $a_1a_2\cdots a_r$ can only contain terms of grade r, $r-2$, $\ldots$, so

$$\begin{aligned} &\frac{1}{2}\Big(aa_1a_2\cdots a_r - (-1)^r a_1a_2\cdots a_ra\Big) \\ &\qquad\qquad = a\cdot\langle a_1a_2\cdots a_r\rangle_r + a\cdot\langle a_1a_2\cdots a_r\rangle_{r-2} + \cdots . \end{aligned} \tag{4.47}$$

The term we are after is the $r-1$ grade part, so we have

$$a\cdot(a_1\wedge a_2\wedge\cdots\wedge a_r) = \frac{1}{2}\langle aa_1a_2\cdots a_r - (-1)^r a_1a_2\cdots a_ra\rangle_{r-1}. \tag{4.48}$$

We can now apply equation (4.31) inside the grade projection operator to form

$$
\begin{aligned}
a\cdot(a_1\wedge a_2\wedge\cdots\wedge a_r) &= \sum_{k=1}^{r}(-1)^{k+1} a\cdot a_k \langle a_1\cdots\check{a}_k\cdots a_r\rangle_{r-1} \\
&= \sum_{k=1}^{r}(-1)^{k+1} a\cdot a_k\, a_1\wedge\cdots\wedge\check{a}_k\wedge\cdots\wedge a_r.
\end{aligned}
\tag{4.49}
$$

The first two cases illustrate how the general formula behaves:

$$
\begin{aligned}
a\cdot(a_1\wedge a_2) &= a\cdot a_1\, a_2 - a\cdot a_2\, a_1, \\
a\cdot(a_1\wedge a_2\wedge a_3) &= a\cdot a_1\, a_2\wedge a_3 - a\cdot a_2\, a_1\wedge a_3 + a\cdot a_3\, a_1\wedge a_2.
\end{aligned}
\tag{4.50}
$$

The first case was established in chapter 2, where it was used to replace the formula for the double cross product of vectors in three dimensions.

4.1.3 The reverse, the scalar and the commutator product

Now that the grading is established, we can establish some general properties of the reversion operator, which was first introduced in chapter 2. The reverse of a product of vectors is defined by

$$
(ab\cdots c)^\dagger = c\cdots ba. \tag{4.51}
$$

For a blade the reverse can be formed by a series of swaps of anticommuting vectors, each resulting in a minus sign. The first vector has to swap past $r-1$ vectors, the second past $r-2$, and so on. This demonstrates that

$$
A_r^\dagger = (-1)^{r(r-1)/2} A_r. \tag{4.52}
$$

If we now consider the scalar part of a geometric product of two grade-r multivectors we find that

$$
\langle A_r B_r\rangle = \langle A_r B_r\rangle^\dagger = \langle B_r^\dagger A_r^\dagger\rangle = (-1)^{r(r-1)}\langle B_r A_r\rangle = \langle B_r A_r\rangle, \tag{4.53}
$$

so, for general A and B,

$$
\langle AB\rangle = \langle BA\rangle. \tag{4.54}
$$

It follows that

$$
\langle A\cdots BC\rangle = \langle CA\cdots B\rangle. \tag{4.55}
$$

This cyclic reordering property is frequently useful for manipulating expressions. The product in equation (4.54) is sometimes given the symbol $*$, so we write

$$
A*B = \langle AB\rangle. \tag{4.56}
$$

A further product of considerable importance in geometric algebra is the commutator product of two multivectors. This is denoted with a cross, $\times$, and is defined by

$$A\times B = \frac{1}{2}(AB - BA). \tag{4.57}$$

Care must be taken to include the factor of one-half, which is different to the standard commutator of two operators in quantum mechanics. The commutator product satisfies the *Jacobi identity*

$$A\times(B\times C) + B\times(C\times A) + C\times(A\times B) = 0, \tag{4.58}$$

which is easily seen by expanding out the products.

The commutator arises most frequently in equations involving bivectors. Given a bivector B and a vector a we have

$$B\times a = \frac{1}{2}(Ba - aB) = B\cdot a, \tag{4.59}$$

which therefore results in a second vector. Now consider the product of a bivector and a blade formed from anticommuting vectors. We have

$$\begin{aligned} B(e_1e_2\cdots e_r) &= 2(B\times e_1)e_2\cdots e_r + e_1Be_2\cdots e_r \\ &= 2(B\times e_1)e_2\cdots e_r + \cdots + 2e_1\cdots(B\times e_r) + e_1e_2\cdots e_rB. \end{aligned} \tag{4.60}$$

It follows that

$$B\times(e_1e_2\cdots e_r) = \sum_{i=1}^{r} e_1\cdots(B\cdot e_i)\cdots e_r. \tag{4.61}$$

The sum involves a series of terms which can only contain grades r and $r-2$. But if we form the reverse of the commutator product between a bivector and a homogeneous multivector, we find that

$$\begin{aligned} (B\times A_r)^\dagger &= \frac{1}{2}(BA_r - A_rB)^\dagger \\ &= \frac{1}{2}(-A_r^\dagger B + BA_r^\dagger) \\ &= (-1)^{r(r-1)/2}B\times A_r. \end{aligned} \tag{4.62}$$

It follows that $B\times A_r$ has the same properties under reversion as A_r. But multivectors of grade r and $r-2$ always behave differently under reversion. The commutator product in equation (4.61) must therefore result in a grade-r multivector. Since this is true of any grade-r basis element, it must be true of any homogeneous multivector. That is,

$$B\times A_r = \langle B\times A_r\rangle_r. \tag{4.63}$$

The commutator of a multivector with a bivector therefore preserves the grade of the multivector. Furthermore, the commutator of two bivectors must result in a third bivector. This is the basis for incorporating the theory of Lie groups into geometric algebra.

A similar argument to the preceding one shows that the symmetric product with a bivector must raise or lower the grade by 2. We can summarise this by writing

$$\begin{aligned} BA_r &= \langle BA_r \rangle_{r-2} + \langle BA_r \rangle_r + \langle BA_r \rangle_{r+2} \\ &= B \cdot A_r + B \times A_r + B \wedge A_r, \end{aligned} \tag{4.64}$$

where

$$\frac{1}{2}(BA_r - A_r B) = B \times A_r \tag{4.65}$$

and

$$\frac{1}{2}(BA_r + A_r B) = B \cdot A_r + B \wedge A_r. \tag{4.66}$$

It is assumed in these formulae that A_r has grade $r > 1$.

4.1.4 Pseudoscalars and duality

The exterior product of n vectors defines a grade-n blade. For a given vector space the highest grade element is unique, up to a magnitude. The outer product of n vectors is therefore a multiple of the unique *pseudoscalar* for $\mathcal{G}_n$. This is denoted I, and has two important properties. The first is that I is normalised to

$$|I^2| = 1. \tag{4.67}$$

The sign of I^2 depends on the size of space and the signature. It turns out that the pseudoscalar squares to -1 for the three algebras of most use in this book — those of the Euclidean plane and space, and of spacetime. But this is in no way a general property.

The second property of the pseudoscalar I is that it defines an *orientation*. For any ordered set of n vectors, their outer product will either have the same sign as I, or the opposite sign. Those with the same sign are assigned a positive orientation, and those with opposite sign have a negative orientation. The orientation is swapped by interchanging any pair of vectors. In three dimensions we always choose the pseudoscalar I such that it has the orientation specified by a right-handed set of vectors. In other spaces one just asserts a choice of I and then sticks to that choice consistently.

The product of the grade-n pseudoscalar I with a grade-r multivector A_r is a grade $n - r$ multivector. This operation is called a *duality* transformation. If A_r is a blade, IA_r returns the *orthogonal complement* of A_r. That is, the blade formed from the space of vectors not contained in A_r. It is clear why this has grade $n - r$. Every blade acts as a pseudoscalar for the space spanned by its generating vectors. So, even if we are working in three dimensions, we can treat the bivector e_1e_2 as a pseudoscalar for any manipulation taking place entirely in the e_1e_2 plane. This is often a very helpful idea.

In spaces of odd dimension, I commutes with all vectors and so commutes with all multivectors. In spaces of even dimension, I anticommutes with vectors and so anticommutes with all odd-grade multivectors. In all cases the pseudoscalar commutes with all even-grade multivectors in its algebra. We summarise this by

$$IA_r = (-1)^{r(n-1)} A_r I. \tag{4.68}$$

An important use of the pseudoscalar is for interchanging inner and outer products. For example, we have

$$\begin{aligned} a\cdot(A_r I) &= \frac{1}{2}\Big(aA_r I - (-1)^{n-r} A_r I a\Big) \\ &= \frac{1}{2}\Big(aA_r I - (-1)^{n-r}(-1)^{n-1} A_r a I\Big) \\ &= \frac{1}{2}\Big(aA_r + (-1)^{r} A_r a\Big) I \\ &= a\wedge A_r\, I. \end{aligned} \tag{4.69}$$

More generally, we can take two multivectors A_r and B_s, with $r+s \leq n$, and form

$$\begin{aligned} A_r\cdot(B_s I) &= \langle A_r B_s I\rangle_{|r-(n-s)|} \\ &= \langle A_r B_s I\rangle_{n-(r+s)} \\ &= \langle A_r B_s\rangle_{r+s} I \\ &= A_r\wedge B_s\, I. \end{aligned} \tag{4.70}$$

This type of interchange is very common in applications. Note how simple this proof is made by the application of the geometric product in the intermediate steps.

4.2 Rotations and reflections

In chapter 2 we showed that in three dimensions a reflection in the plane perpendicular to the unit vector n is performed by

$$a \mapsto a' = -nan. \tag{4.71}$$

This formula holds in arbitrary numbers of dimensions. Provided $n^2 = 1$, we see that n is transformed to

$$n \mapsto -nnn = -n, \tag{4.72}$$

whereas any vector $a_\perp$ perpendicular to n is mapped to

$$a_\perp \mapsto -na_\perp n = a_\perp nn = a_\perp. \tag{4.73}$$

So, for a vector a, the component parallel to n has its sign reversed, whereas the component perpendicular to n is unchanged. This is what we mean by a reflection in the hyperplane perpendicular to n.

Two successive reflections in the hyperplanes perpendicular to m and n result in a rotation in the $m \wedge n$ plane. This is encoded in the rotor

$$R = nm = \exp(-\hat{B}\theta/2) \tag{4.74}$$

where

$$\cos(\theta/2) = n \cdot m, \quad \hat{B} = \frac{m \wedge n}{\sin(\theta/2)}. \tag{4.75}$$

The rotor R generates a rotation through the by now familiar formula

$$a \mapsto a' = RaR^\dagger. \tag{4.76}$$

Rotations form a group, as the result of combining two rotations is a third rotation. The same must therefore be true of rotors. Suppose that R_1 and R_2 generate two distinct rotations. The combined rotations take a to

$$a \mapsto R_2(R_1 a R_1^\dagger)R_2^\dagger = R_2 R_1 a R_1^\dagger R_2^\dagger. \tag{4.77}$$

We therefore define the product rotor

$$R = R_2 R_1, \tag{4.78}$$

so that the result of the composite rotation is described by $RaR^\dagger$, as usual. The product R is a new rotor, and in general it will consist of geometric products of an even number of unit vectors,

$$R = lk \cdots nm. \tag{4.79}$$

We will adopt this as our definition of a rotor. The reversed rotor is

$$R^\dagger = mn \cdots kl. \tag{4.80}$$

The result of the map $a \mapsto RaR^\dagger$ returns a vector for any vector a, since

$$RaR^\dagger = lk \cdots \big(n(mam)n\big) \cdots kl \tag{4.81}$$

and each successive sandwich between a vector returns a new vector.

We can immediately establish the normalisation condition

$$RR^\dagger = lk \cdots nmmn \cdots kl = 1 = R^\dagger R. \tag{4.82}$$

In Euclidean spaces, where every vector has a positive square, this normalisation is automatic. In mixed signature spaces, like Minkowski spacetime, unit vectors can have $n^2 = \pm 1$. In this case the condition $RR^\dagger = 1$ is taken as a further condition satisfied by a rotor. In the case where R is the product of two rotors we can easily confirm that

$$RR^\dagger = R_2 R_1 (R_2 R_1)^\dagger = R_2 R_1 R_1^\dagger R_2^\dagger = 1. \tag{4.83}$$

The set of rotors therefore forms a *group*, called a rotor group. This is similar to the group of rotation matrices, though not identical due to the two-to-one map between rotors and rotation matrices. We will have more to say about the group properties of rotors in chapter 11.

In Euclidean spaces every rotor can be written as the exponential of a bivector,

$$R = \exp(-B/2). \tag{4.84}$$

The bivector B defines the plane or planes in which the rotation takes place. The sign ensures that the rotation has the orientation defined by B. In mixed signature spaces one can always write a rotor as $\pm\exp(B)$. In either case the effect of the rotor R on the vector a is

$$a \mapsto \exp(-B/2)a\exp(B/2). \tag{4.85}$$

We can prove that the right-hand side always returns a vector by considering a Taylor expansion of

$$a(\lambda) = \exp(-\lambda B/2)a\exp(\lambda B/2). \tag{4.86}$$

Differentiating the expression on the right produces the power series expansion

$$a(\lambda) = a + \lambda a\cdot B + \frac{\lambda^2}{2!}(a\cdot B)\cdot B + \cdots . \tag{4.87}$$

Since the inner product of a vector and a bivector always results in a new vector, each term in this expansion is a vector. Setting $\lambda = 1$ then demonstrates that equation (4.85) results in a new vector, defined by

$$\exp(-B/2)a\exp(B/2) = a + a\cdot B + \frac{1}{2!}(a\cdot B)\cdot B + \cdots . \tag{4.88}$$

4.2.1 Multivector transformations

Suppose now that every vector in a blade undergoes the same rotation. This is the sort of transformation implied if a plane or volume element is to be rotated. The r-blade A_r can be written

$$A_r = a_1 \wedge \cdots \wedge a_r = \frac{1}{r!}\sum(-1)^\epsilon a_{k_1}a_{k_2}\cdots a_{k_r}, \tag{4.89}$$

with the sum running over all permutations. If each vector in a geometric product is rotated, the result is the multivector

$$\begin{aligned}(Ra_1R^\dagger)(Ra_2R^\dagger)\cdots(Ra_rR^\dagger) &= Ra_1R^\dagger Ra_2R^\dagger\cdots Ra_rR^\dagger \\ &= Ra_1a_2\cdots a_rR^\dagger.\end{aligned} \tag{4.90}$$

This holds for each term in the antisymmetrised sum, so the transformation law for the blade A_r is simply

$$A_r \mapsto A_r' = RA_rR^\dagger. \tag{4.91}$$

Blades transform with the same simple law as vectors! All multivectors share the same transformation law regardless of grade when each component vector is rotated. This is one reason why the rotor formulation is so powerful. The alternative, tensor form would require an extra matrix for each additional vector.

4.3 Bases, frames and components

Any set of linearly independent vectors form a basis for the vectors in a geometric algebra. Such a set is often referred to as a *frame*. Repeated use of the outer product then builds up a basis for the entire algebra. In this section we use the symbols $\mathsf{e}_1, \ldots, \mathsf{e}_n$ or $\{\mathsf{e}_k\}$ to denote a frame for n-dimensional space. We do not restrict the frame to be orthonormal, so the $\{\mathsf{e}_k\}$ do not necessarily anticommute. The reason for the change of font for frame vectors, as opposed to general sets of vectors, is that use of frames nearly always implies reference to coordinates. It is natural write the coordinates of the vector a as a_i or a^i so, to avoid confusion with a set of vectors, we write the frame vectors in a different font.

The volume element for the $\{\mathsf{e}_k\}$ frame is defined by

$$E_n \equiv \mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \cdots \wedge \mathsf{e}_n. \tag{4.92}$$

The grade-n multivector E_n is a multiple of the pseudoscalar for the space spanned by the $\{\mathsf{e}_k\}$. The fact that the vectors are independent guarantees that $E_n \neq 0$. Associated with any arbitrary frame is a reciprocal frame $\{\mathsf{e}^k\}$ defined by the property

$$\mathsf{e}^i \cdot \mathsf{e}_j = \delta^i_j, \quad \forall i, j = 1 \ldots n. \tag{4.93}$$

The 'Kronecker δ', δ^i_j, has value $+1$ if $i = j$ and is zero otherwise. The reciprocal frame is constructed as follows:

$$\mathsf{e}^j = (-1)^{j-1} \mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \cdots \wedge \check{\mathsf{e}}_j \wedge \cdots \wedge \mathsf{e}_n \, E_n^{-1}, \tag{4.94}$$

where as usual the check on $\check{\mathsf{e}}_j$ denotes that this term is missing from the expression. The formula for e^j has a simple interpretation. The vector e^j must be perpendicular to all the vectors $\{\mathsf{e}_i, i \neq j\}$. To find this we form the exterior product of the $n-1$ vectors $\{\mathsf{e}_i, i \neq j\}$. The dual of this returns a vector perpendicular to all vectors in the subspace, and this duality is achieved by the factor of E_n. All that remains is to fix up the normalisation. For this we recall the duality results of section 4.1.4 and form

$$\mathsf{e}_1 \cdot \mathsf{e}^1 = \mathsf{e}_1 \cdot (\mathsf{e}_2 \wedge \cdots \wedge \mathsf{e}_n \, E_n^{-1}) = (\mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \cdots \wedge \mathsf{e}_n) E_n^{-1} = 1. \tag{4.95}$$

This confirms that the formula for the reciprocal frame is correct.

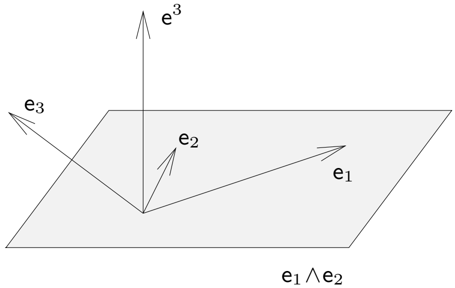

Figure 4.2 *The reciprocal frame.* The vectors e_1, e_2 and e_3 form a non-orthonormal frame for three-dimensional space. The vector e^3 is formed by constructing the $\mathsf{e}_1 \wedge \mathsf{e}_2$ plane, and forming the vector perpendicular to this plane. The length is fixed by demanding $\mathsf{e}^3 \cdot \mathsf{e}_3 = 1$.

4.3.1 Application — crystallography

An important application of the formula for a reciprocal frame is in crystallography. If a crystal contains some repeated structure defined by the vectors $\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3$, then constructive interference occurs for wavevectors whose difference satisfies

$$\Delta k = 2\pi(n_1\mathsf{e}^1 + n_2\mathsf{e}^2 + n_3\mathsf{e}^3), \tag{4.96}$$

where n_1, n_2, n_3 are integers. The reciprocal frame is defined by

$$\mathsf{e}^1 = \frac{\mathsf{e}_2 \wedge \mathsf{e}_3}{\mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \mathsf{e}_3}, \quad \mathsf{e}^2 = \frac{\mathsf{e}_3 \wedge \mathsf{e}_1}{\mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \mathsf{e}_3}, \quad \mathsf{e}^3 = \frac{\mathsf{e}_1 \wedge \mathsf{e}_2}{\mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \mathsf{e}_3}. \tag{4.97}$$

If we write

$$\mathsf{e}_1 \wedge \mathsf{e}_2 \wedge \mathsf{e}_3 = [\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3] I, \tag{4.98}$$

where I is the three-dimensional pseudoscalar and $[\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3]$ denotes the scalar triple product, we arrive at the standard formula

$$\mathsf{e}^1 = \frac{(\mathsf{e}_2 \wedge \mathsf{e}_3) I^{-1}}{[\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3]} = \frac{\mathsf{e}_2 \boldsymbol{\times} \mathsf{e}_3}{[\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3]}, \tag{4.99}$$

with similar results holding for e^2 and e^3. Here the bold cross $\boldsymbol{\times}$ denotes the vector cross product, not to be confused with the commutator product. Figure 4.2 illustrates the geometry involved in defining the reciprocal frame.

4.3.2 Components

The basis vectors $\{\mathsf{e}_k\}$ are linearly independent, so any vector a can be written uniquely in terms of this set as

$$a = a^i \mathsf{e}_i = a_i \mathsf{e}^i. \tag{4.100}$$

We continue to employ the summation convention and summed indices appear once as a superscript and once as a subscript. The set of scalars $(a^1, \ldots, a^n)$ are the *components* of the vector a in the $\{\mathsf{e}_k\}$ frame. To find the components we form

$$a\cdot\mathsf{e}^i = a^j\mathsf{e}_j\cdot\mathsf{e}^i = a^j\delta^i_j = a^i \tag{4.101}$$

and

$$a\cdot\mathsf{e}_i = a_j\mathsf{e}^j\cdot\mathsf{e}_i = a_j\delta^j_i = a_i. \tag{4.102}$$

These formulae explain the labelling scheme for the components. In many applications we are only interested in orthonormal frames in Euclidean space. In this case the frame and its reciprocal are equivalent, and there is no need for the distinct subscript and superscript indices. The notation is unavoidable in mixed signature spaces, however, and is very useful in differential geometry, so it is best to adopt it at the outset.

Combining the equations (4.100), (4.101) and (4.102) we see that

$$a\cdot\mathsf{e}_i\,\mathsf{e}^i = a\cdot\mathsf{e}^i\,\mathsf{e}_i = a. \tag{4.103}$$

This holds for any vector a in the space spanned by the $\{\mathsf{e}_k\}$. This result generalises simply to arbitrary multivectors. First, for the bivector $a\wedge b$ we have

$$\mathsf{e}_i\,\mathsf{e}^i\cdot(a\wedge b) = \mathsf{e}_i\,\mathsf{e}^i\cdot a\,b - \mathsf{e}_i\,\mathsf{e}^i\cdot b\,a = ab - ba = 2a\wedge b. \tag{4.104}$$

This extends for an arbitrary grade-r multivector A_r to give

$$\mathsf{e}_i\,\mathsf{e}^i\cdot A_r = rA_r. \tag{4.105}$$

Since $\mathsf{e}_i\mathsf{e}^i = n$, we also see that

$$\mathsf{e}_i\,\mathsf{e}^i\wedge A_r = \mathsf{e}_i(\mathsf{e}^i A_r - \mathsf{e}^i\cdot A_r) = (n-r)A_r. \tag{4.106}$$

Subtracting the two preceding results we obtain,

$$\mathsf{e}_i A_r \mathsf{e}^i = (-1)^r(n-2r)A_r. \tag{4.107}$$

The $\{\mathsf{e}_k\}$ basis extends easily to provide a basis for the entire algebra generated by the basis vectors. We can then decompose any multivector A into a set of components through

$$A_{i\cdots jk} = \langle(\mathsf{e}_k\wedge\mathsf{e}_j\cdots\wedge\mathsf{e}_i)A\rangle. \tag{4.108}$$

and

$$A = \sum_{i<j\cdots<k} A_{ij\cdots k}\mathsf{e}^i\wedge\cdots\wedge\mathsf{e}^j\wedge\mathsf{e}^k. \tag{4.109}$$

The components $A_{ij\cdots k}$ are totally antisymmetric on all indices and are usually referred to as the components of an *antisymmetric tensor*. We shall have more to say about tensors in following sections.

4.3.3 Application — recovering a rotor

As an application of the preceding results, suppose that we have two sets of vectors in three dimensions $\{\mathsf{e}_k\}$ and $\{\mathsf{f}_k\}$, $k = 1, 2, 3$. The vectors need not be orthonormal, but we know that the two sets are related by a rotation. The rotation is governed by the formula

$$\mathsf{f}_k = R\mathsf{e}_k R^\dagger \tag{4.110}$$

and we seek a simple expression for the rotor R. In three dimensions the rotor R can be written as

$$R = \exp(-B/2) = \alpha - \beta B, \tag{4.111}$$

where

$$\alpha = \cos(|B|/2), \quad \beta = \frac{\sin(|B|/2)}{|B|}. \tag{4.112}$$

The reverse is

$$R^\dagger = \exp(B/2) = \alpha + \beta B. \tag{4.113}$$

We therefore find that

$$\begin{aligned} \mathsf{e}_k R^\dagger \mathsf{e}^k &= \mathsf{e}_k(\alpha + \beta B)\mathsf{e}^k \\ &= 3\alpha - \beta B \\ &= 4\alpha - R^\dagger. \end{aligned} \tag{4.114}$$

We now form

$$\mathsf{f}_k \mathsf{e}^k = R\mathsf{e}_k R^\dagger \mathsf{e}^k = 4\alpha R - 1. \tag{4.115}$$

It follows that R is a scalar multiple of $1 + \mathsf{f}_k\mathsf{e}^k$. We therefore establish the simple formula

$$R = \frac{1 + \mathsf{f}_k\mathsf{e}^k}{|1 + \mathsf{f}_k\mathsf{e}^k|} = \frac{\psi}{\surd(\psi\tilde{\psi})}, \tag{4.116}$$

where $\psi = 1 + \mathsf{f}_k\mathsf{e}^k$. This compact formula recovers the rotor directly from the frame vectors. A problem arises if the rotation is through precisely 180°, in which case ψ vanishes. This case can be dealt with simply enough by considering the image of two of the three vectors.

4.4 Linear algebra

Many key relations in physics involve linear mappings between two, sometimes different, spaces. These are the subject of tensor analysis in the standard literature. Examples include the stress and strain tensors of elasticity, the conductivity tensor of electromagnetism and the inertia tensor of dynamics. If one has only met the study of linear transformations through tensor analysis, one could be forgiven for thinking that the subject cannot be discussed without a large dose of index notation. The indices refer to components of tensors in some frame, though the essence of tensor analysis is to establish a set of results which are independent of the choice of frame. In our opinion, this subject is much more simply dealt with if one can avoid specifying a frame until it is absolutely necessary. Perhaps unsurprisingly, it is geometric algebra that provides precisely the tools necessary to achieve such a development.

In this section we use capital, sans-serif symbols for linear functions. This helps to distinguish functions from their multivector argument. The dimension and signature of the vector space is arbitrary unless otherwise specified. We assume that readers are familiar with the basic properties of linear transformations in the guise of matrices. Suppose, then, that we are interested in a quantity F which maps vectors to vectors linearly in the same space. That is, if a is a vector in the space acted on by F, then $\mathsf{F}(a)$ lies in the same space. The linearity of F is expressed by

$$\mathsf{F}(\lambda a + \mu b) = \lambda \mathsf{F}(a) + \mu \mathsf{F}(b), \tag{4.117}$$

for scalars λ and μ and vectors a and b. Geometrically, we can think of F as an instruction to take a vector and rotate/dilate it to a new vector. No frame or components are required for such a picture. A simple example is provided by a rotation, which can be written as

$$\mathsf{R}(a) = RaR^{\dagger}, \tag{4.118}$$

where R is a rotor. It is a simple matter to confirm that this map is linear.

4.4.1 Extension to multivectors

Once one has formulated the action of a linear function on a vector, the obvious next step is to let the function act on a multivector. In this way we extend the action of a linear function to the full geometric algebra defined by the underlying vector space. Suppose that two vectors a and b are acted on by the linear function F. The bivector $a \wedge b$ then transforms to $\mathsf{F}(a) \wedge \mathsf{F}(b)$. We take this as the definition for the action of F on a bivector blade:

$$\mathsf{F}(a \wedge b) = \mathsf{F}(a) \wedge \mathsf{F}(b). \tag{4.119}$$

Since the right-hand side is the outer product of two vectors, it is also a bivector blade (see figure 4.3). The action on sums of blades is defined by the linearity of F:

$$\mathsf{F}(a \wedge b + c \wedge d) = \mathsf{F}(a \wedge b) + \mathsf{F}(c \wedge d). \tag{4.120}$$

Continuing in this manner, we define the action of F on an arbitrary blade by

$$\mathsf{F}(a \wedge b \wedge \cdots \wedge c) = \mathsf{F}(a) \wedge \mathsf{F}(b) \wedge \cdots \wedge \mathsf{F}(c). \tag{4.121}$$

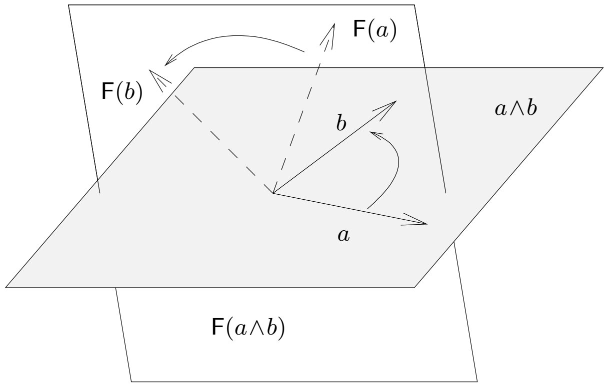

Figure 4.3 *The extended linear function.* The action of F on the bivector $a\wedge b$ results in the new plane $\mathsf{F}(a)\wedge\mathsf{F}(b)$. This is the definition of $\mathsf{F}(a\wedge b)$.

Extension by linearity then defines the action of F on arbitrary multivectors. By construction, F is both linear over multivectors,

$$\mathsf{F}(\lambda A+\mu B)=\lambda\mathsf{F}(A)+\mu\mathsf{F}(B), \tag{4.122}$$

and grade-preserving,

$$\mathsf{F}(A_r)=\langle\mathsf{F}(A_r)\rangle_r, \tag{4.123}$$

where A_r is a grade-r multivector. A simple example is provided by rotations. We have already established a formula for the result of rotating all of the vectors in a blade. For the extension of a rotation we therefore have

$$\begin{aligned}\mathsf{R}(a\wedge b\wedge\cdots\wedge c)&=(RaR^\dagger)\wedge(RbR^\dagger)\wedge\cdots\wedge(RcR^\dagger)\\&=R\,a\wedge b\wedge\cdots\wedge c\,R^\dagger.\end{aligned} \tag{4.124}$$

It follows that acting on an arbitrary multivector A we have

$$\mathsf{R}(A)=RAR^\dagger. \tag{4.125}$$

Again, it is simple to confirm that this has the expected properties.

4.4.2 The product

The product of two linear functions is formed by letting a second function act on the result of the first function. Thus the action of the product of F and G is defined by

$$(\mathsf{FG})(a)=\mathsf{F}\big(\mathsf{G}(a)\big)=\mathsf{FG}(a). \tag{4.126}$$

The final expression enables us to remove some brackets without any ambiguity. A price to pay for removing indices is that brackets are often required to show

how calculations are ordered. Any convention that enables brackets to be systematically dropped is then well worth adopting. It is straightforward to show that FG is a linear function if F and G are both linear:

$$\mathsf{FG}(\lambda a + \mu b) = \mathsf{F}(\lambda \mathsf{G}(a) + \mu \mathsf{G}(b)) = \lambda \mathsf{FG}(a) + \mu \mathsf{FG}(b). \tag{4.127}$$

Next we form the extension of a product function. Suppose that H is given by the product of F and G:

$$\mathsf{H}(a) = \mathsf{F}\big(\mathsf{G}(a)\big) = \mathsf{FG}(a). \tag{4.128}$$

It follows that

$$\begin{aligned} \mathsf{H}(a\wedge b\wedge\cdots\wedge c) &= \mathsf{F}\big(\mathsf{G}(a)\big)\wedge\mathsf{F}\big(\mathsf{G}(b)\big)\wedge\cdots\wedge\mathsf{F}\big(\mathsf{G}(c)\big) \\ &= \mathsf{F}\big(\mathsf{G}(a)\wedge\mathsf{G}(b)\wedge\cdots\wedge\mathsf{G}(c)\big) \\ &= \mathsf{F}\big(\mathsf{G}(a\wedge b\wedge\cdots\wedge c)\big), \end{aligned} \tag{4.129}$$

so the multilinear action of the product of two linear functions is the product of their exterior actions. In dealing with combinations of linear functions we can therefore write

$$\mathsf{H}(A) = \mathsf{FG}(A), \tag{4.130}$$

since the meaning of the right-hand side is unambiguous.

4.4.3 The adjoint

Given a linear function F, the adjoint, or transpose, $\bar{\mathsf{F}}$ is defined so that

$$a\cdot\bar{\mathsf{F}}(b) = \mathsf{F}(a)\cdot b, \tag{4.131}$$

for all vectors a and b. If F is a mapping from one vector space to another, then the adjoint function maps from the second space back to the first. In terms of an arbitrary frame $\{\mathsf{e}_k\}$ we have

$$\mathsf{e}_i\cdot\bar{\mathsf{F}}(a) = a\cdot\mathsf{F}(\mathsf{e}_i), \tag{4.132}$$

so we can construct the adjoint using

$$\mathrm{ad}(\mathsf{F})(a) = \bar{\mathsf{F}}(a) = \mathsf{e}^i\, a\cdot\mathsf{F}(\mathsf{e}_i). \tag{4.133}$$

The notation of a bar for the adjoint, rather than a superscript T or †, is slightly unconventional, though it does agree with that of Hestenes & Sobczyk (1984). The notation is very useful in handwritten work, where it is also convenient to denote the linear function with an underline. Some formulae relating functions and their adjoints have a neat symmetry when this overbar/underbar convention is followed.

The operation of taking the adjoint of the adjoint of a function returns the original function. This is verified by forming

$$\mathrm{ad}(\bar{\mathsf{F}})(a) = \mathsf{e}^i a\cdot\bar{\mathsf{F}}(\mathsf{e}_i) = \mathsf{e}^i\, \mathsf{e}_i\cdot\mathsf{F}(a) = \mathsf{F}(a). \tag{4.134}$$

The adjoint of a product of two functions is found as follows:

$$\begin{aligned}\mathrm{ad}(\mathsf{FG})(a) &= \mathsf{e}^i\, a\cdot\mathsf{FG}(\mathsf{e}_i) = \bar{\mathsf{F}}(a)\cdot\mathsf{G}(\mathsf{e}_i)\,\mathsf{e}^i \\ &= \bar{\mathsf{G}}\bar{\mathsf{F}}(a)\cdot\mathsf{e}_i\,\mathsf{e}^i = \bar{\mathsf{G}}\bar{\mathsf{F}}(a).\end{aligned} \tag{4.135}$$

The operation of taking the adjoint of a product therefore reverses the order in which the linear functions act. A *symmetric* function is one which is equal to its own adjoint, $\bar{\mathsf{F}} = \mathsf{F}$. Two particularly significant examples of symmetric functions are the functions $\mathsf{F}\bar{\mathsf{F}}$ and $\bar{\mathsf{F}}\mathsf{F}$. To verify that these are symmetric we form

$$\mathrm{ad}(\mathsf{F}\bar{\mathsf{F}}) = \mathrm{ad}(\bar{\mathsf{F}})\mathrm{ad}(\mathsf{F}) = \mathsf{F}\bar{\mathsf{F}}, \tag{4.136}$$

with a similar derivation holding for $\bar{\mathsf{F}}\mathsf{F}$. These functions will be met again later in this chapter.

The adjoint is still a linear function, so its extension to arbitrary multivectors is precisely as expected:

$$\bar{\mathsf{F}}(a\wedge b\wedge\cdots\wedge c) = \bar{\mathsf{F}}(a)\wedge\bar{\mathsf{F}}(b)\wedge\cdots\wedge\bar{\mathsf{F}}(c). \tag{4.137}$$

If we now consider two bivectors $a_1 \wedge a_2$ and $b_1 \wedge b_2$, we find that

$$\begin{aligned}(a_1\wedge a_2)\cdot\mathsf{F}(b_1\wedge b_2) &= a_1\cdot\mathsf{F}(b_2)\, a_2\cdot\mathsf{F}(b_1) - a_1\cdot\mathsf{F}(b_1)\, a_2\cdot\mathsf{F}(b_2) \\ &= \bar{\mathsf{F}}(a_1)\cdot b_2\, \bar{\mathsf{F}}(a_2)\cdot b_1 - \bar{\mathsf{F}}(a_1)\cdot b_1\, \bar{\mathsf{F}}(a_2)\cdot b_2 \\ &= \bar{\mathsf{F}}(a_1\wedge a_2)\cdot(b_1\wedge b_2).\end{aligned} \tag{4.138}$$

It follows that for two bivectors B_1 and B_2

$$B_1\cdot\bar{\mathsf{F}}(B_2) = \mathsf{F}(B_1)\cdot B_2. \tag{4.139}$$

This result extends for arbitrary multivectors to give

$$\langle A\bar{\mathsf{F}}(B)\rangle = \langle\mathsf{F}(A)B\rangle. \tag{4.140}$$

This is a special case of an even more general and powerful result. Consider the expression

$$\begin{aligned}\mathsf{F}(a\wedge b)\cdot c &= \mathsf{F}(a)\,\mathsf{F}(b)\cdot c - \mathsf{F}(b)\,\mathsf{F}(a)\cdot c \\ &= \mathsf{F}\big(a\, b\cdot\bar{\mathsf{F}}(c) - b\, a\cdot\bar{\mathsf{F}}(c)\big) \\ &= \mathsf{F}\big((a\wedge b)\cdot\bar{\mathsf{F}}(c)\big).\end{aligned} \tag{4.141}$$

Building up in this way we establish the useful results:

$$\begin{aligned}A_r\cdot\bar{\mathsf{F}}(B_s) &= \bar{\mathsf{F}}\big(\mathsf{F}(A_r)\cdot B_s\big) \qquad r\le s, \\ \mathsf{F}(A_r)\cdot B_s &= \mathsf{F}\big(A_r\cdot\bar{\mathsf{F}}(B_s)\big) \qquad r\ge s.\end{aligned} \tag{4.142}$$

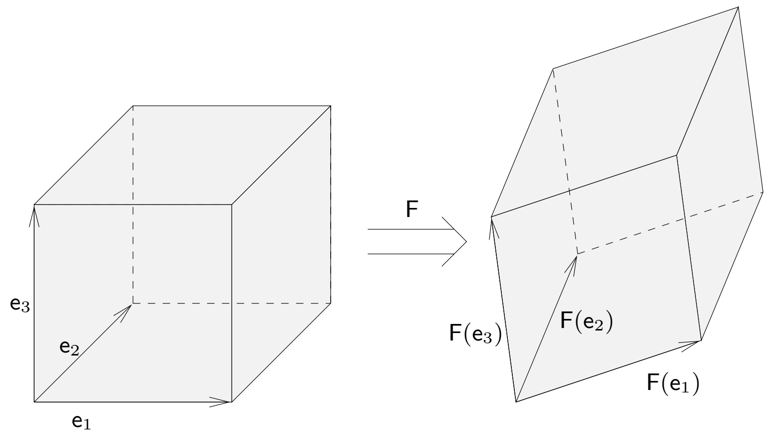

Figure 4.4 *The determinant.* The unit cube is transformed to a parallelepiped with sides $\mathsf{F}(\mathsf{e}_1)$, $\mathsf{F}(\mathsf{e}_2)$ and $\mathsf{F}(\mathsf{e}_3)$. The determinant is the volume scale factor, so is given by the volume of the parallelepiped, $\mathsf{F}(\mathsf{e}_1)\wedge\mathsf{F}(\mathsf{e}_2)\wedge\mathsf{F}(\mathsf{e}_3) = \mathsf{F}(I)$.

These reduce to equation (4.140) in the case when $r = s$. One way to think of these formulae is as follows. In the expression $\mathsf{F}(A_r)\cdot B_s$, with $r \geq s$, there are r separate applications of the function F on vectors. When the result is contracted with B_s, s of these applications are converted to adjoint functions $\bar{\mathsf{F}}$. The remaining $r - s$ applications act on the multivector $A_r \cdot \bar{\mathsf{F}}(B_s)$, which has grade $r - s$.

4.4.4 The determinant

Now that we have seen how a linear function defines an action on the entire geometric algebra, we can give a very compact definition of the determinant. The pseudoscalar for any space is unique up to scaling, and linear functions are grade-preserving, so we define

$$\mathsf{F}(I) = \det\,(\mathsf{F})\, I. \tag{4.143}$$

It should be immediately apparent that this definition of the determinant is much more compact and intuitive than the matrix definition (discussed later). The definition (4.143) shows clearly that the determinant is the volume scale factor for the operation F. In particular, acting on the unit hypercube, the result $\mathsf{F}(I)$ returns the directed volume of the resultant object (see figure 4.4).

As an example of the power of the geometric algebra definition, consider the product of two functions, F and G. From equation (4.130) it follows that

$$\det\,(\mathsf{FG})I = \mathsf{FG}(I) = \det\,(\mathsf{G})\,\mathsf{F}(I) = \det\,(\mathsf{F})\det\,(\mathsf{G})\,I, \tag{4.144}$$

which establishes that the determinant of the product of two functions is the product of their determinants. This is one of the key properties of the determinant, yet in conventional developments it is hard to prove. By contrast, the geometric algebra approach establishes the result in a few lines. Similarly, one can easily establish that the determinant of the adjoint is the same as that of the original function,

$$\det(\mathsf{F}) = \langle \mathsf{F}(I)I^{-1}\rangle = \langle I\bar{\mathsf{F}}(I^{-1})\rangle = \det(\bar{\mathsf{F}}). \tag{4.145}$$

Example 4.1

Consider the linear function

$$\mathsf{F}(a) = a + \alpha a\cdot f_1\, f_2, \tag{4.146}$$

where α is a scalar and f_1 and f_2 are a pair of arbitrary vectors. Construct the action of F on a general multivector and find its determinant.

We start by forming

$$\begin{aligned}\mathsf{F}(a\wedge b) &= (a + \alpha a\cdot f_1 f_2)\wedge(b + \alpha b\cdot f_1 f_2)\\ &= a\wedge b + \alpha(b\cdot f_1 a - a\cdot f_1 b)\wedge f_2\\ &= a\wedge b + \alpha\big((a\wedge b)\cdot f_1\big)\wedge f_2.\end{aligned} \tag{4.147}$$

It follows that

$$\mathsf{F}(A) = A + \alpha(A\cdot f_1)\wedge f_2. \tag{4.148}$$

The determinant is now calculated as follows:

$$\begin{aligned}\mathsf{F}(I) &= I + \alpha(I\cdot f_1)\wedge f_2\\ &= I + \alpha f_1\cdot f_2\, I,\end{aligned} \tag{4.149}$$

hence $\det(\mathsf{F}) = 1 + \alpha f_1\cdot f_2$.

4.4.5 The inverse

We now construct a simple, explicit formula for the inverse of a linear function. We start by considering a multivector B, lying entirely in the algebra defined by the pseudoscalar I. For these we have

$$\det(\mathsf{F})IB = \mathsf{F}(I)B = \mathsf{F}\big(I\bar{\mathsf{F}}(B)\big), \tag{4.150}$$

where we have used the adjoint formulae of equation (4.142). The inner product with a pseudoscalar is replaced with a geometric product, since no other grades are present in the full product. Replacing IB by A we find that

$$\det(\mathsf{F})A = \mathsf{F}\big(I\bar{\mathsf{F}}(I^{-1}A)\big) \tag{4.151}$$

with a similar result holding for the adjoint. It follows that

$$\begin{aligned} \mathsf{F}^{-1}(A) &= I\bar{\mathsf{F}}(I^{-1}A)\det(\mathsf{F})^{-1}, \\ \bar{\mathsf{F}}^{-1}(A) &= I\mathsf{F}(I^{-1}A)\det(\mathsf{F})^{-1}. \end{aligned} \tag{4.152}$$

These relations provide simple, explicit formulae for the inverse of a function. The derivation of these formulae is considerably quicker than anything available in traditional matrix/tensor analysis.

Example 4.2

Find the inverse of the function defined in equation (4.146).

With

$$\mathsf{F}(A) = A + \alpha(A\cdot f_1)\wedge f_2 \tag{4.153}$$

we have

$$\begin{aligned} \langle A_r \mathsf{F}(B_r)\rangle &= \langle A_r B_r\rangle + \alpha\langle A_r (B_r\cdot f_1)\wedge f_2\rangle \\ &= \langle A_r B_r\rangle + \alpha\langle f_2\cdot A_r B_r f_1\rangle, \end{aligned} \tag{4.154}$$

hence

$$\bar{\mathsf{F}}(A) = A + \alpha f_1\wedge(f_2\cdot A). \tag{4.155}$$

It follows that

$$\begin{aligned} \mathsf{F}^{-1}(A) &= \big(I^{-1}A + \alpha f_1\wedge(f_2\cdot(I^{-1}A))\big)(1+\alpha f_1\cdot f_2)^{-1} \\ &= (A + \alpha f_1\cdot(f_2\wedge A))(1+\alpha f_1\cdot f_2)^{-1} \\ &= A - \frac{\alpha}{1+\alpha f_1\cdot f_2} f_2\wedge(f_1\cdot A). \end{aligned} \tag{4.156}$$

Example 4.3

Find the inverse of the rotation

$$\mathsf{R}(a) = RaR^\dagger, \tag{4.157}$$

where R is a rotor.

We have already seen that the action of R on a general multivector is

$$\mathsf{R}(A) = RAR^\dagger \quad \text{and} \quad \bar{\mathsf{R}}(A) = R^\dagger AR \tag{4.158}$$

Hence

$$\det(\mathsf{R})I = RIR^\dagger = IRR^\dagger = I, \tag{4.159}$$

so $\det(\mathsf{R}) = 1$. It follows that

$$\mathsf{R}^{-1}(A) = IR^\dagger I^{-1}AR = R^\dagger AR = \bar{\mathsf{R}}(a), \tag{4.160}$$

so, as expected, the inverse of a rotation is the same as the adjoint. This is the definition of an orthogonal transformation.

4.4.6 Eigenvectors and eigenblades

We assume that readers are familiar with the concept of an eigenvalue and eigenvector of a matrix. All of the standard results for these have obvious counterparts in the geometric algebra framework. This subject will be explored more thoroughly in chapter 11. Here we give a simple outline, concentrating on the new concepts that geometric algebra offers. A linear function F has an eigenvector e if

$$\mathsf{F}(e) = \lambda e. \tag{4.161}$$

The scalar λ is the associated eigenvalue. It follows that

$$\det\,(\mathsf{F} - \lambda \mathsf{I}) = 0, \tag{4.162}$$

which defines a polynomial equation for λ. Techniques for finding eigenvalues and eigenvectors are discussed widely in the literature.

In general, the polynomial equation for λ will have complex roots. Traditional developments of the subject usually allow these and consider linear superpositions over the complex field. But if one starts with a real mapping between real vectors it is not clear that this formal complexification is useful. What one would like would be a more geometric classification of a general linear transformation. This is provided by the notion of an *eigenblade*. We extend the notion of an eigenvector to that of an eigenblade A_r satisfying

$$\mathsf{F}(A_r) = \lambda A_r, \tag{4.163}$$

where A_r is a grade-r blade and λ is real. One immediate example is the pseudoscalar, for which $\lambda = \det\,(\mathsf{F})$. More generally, each eigenblade determines an invariant subspace of the transformation.

As an example of the geometric clarity of the eigenblade concept, consider a function satisfying

$$\mathsf{F}(e_1) = \lambda e_2, \quad \mathsf{F}(e_2) = -\lambda e_1. \tag{4.164}$$

Traditionally, one might write that $e_1 \pm ie_2$ are eigenvectors with eigenvalues $\mp i\lambda$, where i is the unit imaginary. But the identity

$$\mathsf{F}(e_1 \wedge e_2) = \lambda^2 e_1 \wedge e_2 \tag{4.165}$$

identifies the plane $e_1 \wedge e_2$ as an eigenbivector of F. The role of the complex structure inherent in F is played by the unit bivector $e_1 \wedge e_2$. A linear function can have many distinct eigenbivectors, each acting as a distinct imaginary for its own plane. Replacing all of these by a single scalar imaginary throws away a considerable amount of useful information.

4.4.7 Symmetric and antisymmetric functions

An important aspect of the theory of linear functions is finding natural, *canonical*† expressions for a function. For symmetric functions in Euclidean space this form is via its spectral decomposition. If e_i and e_j are eigenvectors of a function, with eigenvalues λ_i and λ_j, we have (no sums implied)

$$e_i \cdot \mathsf{F}(e_j) = e_i \cdot (\lambda_j e_j) = \lambda_j e_i \cdot e_j. \tag{4.166}$$

But if F is symmetric, this also equals

$$\bar{\mathsf{F}}(e_i) \cdot e_j = \mathsf{F}(e_i) \cdot e_j = (\lambda_i e_i) \cdot e_j = \lambda_i e_i \cdot e_j. \tag{4.167}$$

It follows that

$$(\lambda_i - \lambda_j) e_i \cdot e_j = 0, \tag{4.168}$$

so eigenvectors of a symmetric function with distinct eigenvalues must be orthogonal.

If we admit the existence of complex eigenvectors and eigenvalues we also find that (no sums)

$$e^* \cdot \mathsf{F}(e) = \lambda e^* \cdot e = \mathsf{F}(e^*) \cdot e = \lambda^* e^* \cdot e. \tag{4.169}$$

So for any symmetric function we also have

$$(\lambda - \lambda^*) e^* \cdot e = 0. \tag{4.170}$$

Provided $e^* \cdot e \neq 0$ we can conclude that the eigenvalue, and hence the eigenvector, is real. In Euclidean space this inequality is always satisfied, and every symmetric function on an n-dimensional space has a spectral decomposition of the form

$$\mathsf{F}(a) = \lambda_1 \mathsf{P}_1(a) + \lambda_2 \mathsf{P}_2(a) + \cdots + \lambda_m \mathsf{P}_m(a). \tag{4.171}$$

Here $\lambda_1 < \lambda_2 < \cdots < \lambda_m$ are the m distinct eigenvalues ($m \leq n$) and the P_i are projections onto each of the invariant subspaces defined by the eigenvectors. For the case of a projection onto a one-dimensional space we have simply

$$\mathsf{P}_i(a) = a \cdot e_i \, e_i. \tag{4.172}$$

The eigenvectors form an orthonormal frame, which is the natural frame in which to study the linear function. If two eigenvalues are the same, it is always possible to choose the eigenvectors so that they remain orthogonal. In non-Euclidean spaces, such as spacetime, one has to be careful due to the possibility of complex null vectors. These can have $e^* \cdot e = 0$, so the above reasoning breaks down and

† The origin of the use of the word *canonical* is obscure — see for example the comments in Goldstein (1950). In mathematical physics, a canonical form usually refers to a standard way of simplifying an expression without altering its meaning.

one cannot guarantee the existence of an orthonormal frame of eigenvectors. We will encounter examples of this when we study gravitation.

Antisymmetric functions have $\bar{\mathsf{F}}(a) = -\mathsf{F}(a)$. It follows that

$$a\cdot\mathsf{F}(a) = \bar{\mathsf{F}}(a)\cdot a = -\mathsf{F}(a)\cdot a = 0. \tag{4.173}$$

The natural way to study antisymmetric functions is through the bivector

$$F = \frac{1}{2}\mathsf{e}^i\wedge\mathsf{F}(\mathsf{e}_i), \tag{4.174}$$

where the $\{\mathsf{e}_k\}$ are an arbitrary frame for the space acted on by F. The bivector F is independent of the choice of frame, so is an invariant quantity. One can easily confirm that the bivector F has the same number of degrees of freedom as F. If we now form $2a\cdot F$ we find that

$$\begin{aligned} 2a\cdot F &= a\cdot\big(\mathsf{e}^i\wedge\mathsf{F}(\mathsf{e}_i)\big) \\ &= a\cdot\mathsf{e}^i\,\mathsf{F}(\mathsf{e}_i) - \mathsf{e}^i a\cdot\mathsf{F}(\mathsf{e}_i) \\ &= \mathsf{F}(a\cdot\mathsf{e}^i\,\mathsf{e}_i) + \mathsf{e}^i\,\mathsf{e}_i\cdot\mathsf{F}(a) \\ &= 2\mathsf{F}(a). \end{aligned} \tag{4.175}$$

The action of an antisymmetric function therefore reduces to contracting with the *characteristic bivector* F:

$$\mathsf{F}(a) = a\cdot F. \tag{4.176}$$

The problem of reducing an antisymmetric function to its simplest form reduces to that of splitting F into a set of commuting blades:

$$F = \lambda_1\hat{F}_1 + \cdots + \lambda_k\hat{F}_k, \tag{4.177}$$

where $k \leq n/2$ and each of the $\hat{F}_i$ is a unit blade. This decomposition is always possible in Euclidean space, though the answer is only unique if the blades all have different magnitudes. Each component blade of F is an eigenblade of F and determines an invariant subspace. Within this subspace the effect of F is simply to rotate all vectors by $\pm 90°$, and to scale the result by the magnitude of the eigenblade. In non-Euclidean spaces such a decomposition is not always possible.

4.4.8 The singular value decomposition

For linear functions of no symmetry a number of alternative canonical forms can be found. Among these, perhaps the most useful is the singular value decomposition. We start with an arbitrary function F and restrict the discussion to the case where F acts on an n-dimensional Euclidean space. We also suppose that $\det(\mathsf{F}) \neq 0$; the case of $\det(\mathsf{F}) = 0$ is easily dealt with by separating out the space which is mapped onto the origin, and working with a reduced function acting in the subspace over which F is non-singular. We next form the function D by

$$\mathsf{D}(a) = \bar{\mathsf{F}}\mathsf{F}(a). \tag{4.178}$$

This function is symmetric and has n orthogonal eigenvectors with real, positive eigenvalues. The fact that the eigenvalues are positive follows from

$$\bar{\mathsf{F}}\mathsf{F}(e) = \lambda e \quad \Rightarrow \quad \mathsf{F}(e)\cdot\mathsf{F}(e) = \lambda e^2. \tag{4.179}$$

Since (in Euclidean space) the square of any vector is a positive scalar we see that λ must be positive. The assumption that $\det(\mathsf{F}) \neq 0$ rules out the possibility of any eigenvalues being zero. It follows that we can write

$$\mathsf{D}(a) = \sum_{i=1}^{n} \lambda_i a\cdot\mathsf{e}_i\, \mathsf{e}_i, \tag{4.180}$$

where the $\{\mathsf{e}_i\}$ are the *orthonormal* frame of eigenvectors. Degenerate eigenvalues are dealt with by picking a set of arbitrary orthonormal vectors in the invariant subspace.

The linear function D has a simple (positive) square root,

$$\mathsf{D}^{1/2} = \sum_{i=1}^{n} \lambda_i^{1/2} a\cdot\mathsf{e}_i\, \mathsf{e}_i \tag{4.181}$$

and this is also invertible,

$$\mathsf{D}^{-1/2} = \sum_{i=1}^{n} \lambda_i^{-1/2} a\cdot\mathsf{e}_i\, \mathsf{e}_i. \tag{4.182}$$

We now set

$$\mathsf{S} = \mathsf{F}\mathsf{D}^{-1/2}. \tag{4.183}$$

This satisfies

$$\bar{\mathsf{S}}\mathsf{S} = \mathsf{D}^{-1/2}\bar{\mathsf{F}}\mathsf{F}\mathsf{D}^{-1/2} = \mathsf{D}^{-1/2}\mathsf{D}\,\mathsf{D}^{-1/2} = \mathsf{I}, \tag{4.184}$$

where I is the identity function. It follows that S is an orthogonal function. The function F can now be written

$$\mathsf{F} = \mathsf{S}\mathsf{D}^{1/2}. \tag{4.185}$$

This represents a series of dilations along the eigendirections of D, followed by a rotation.

If the linear function F is presented as an $n \times n$ matrix of components in some frame, then one usually includes a further rotation R to align this arbitrary frame with the frame of eigenvectors. In this case one writes

$$\mathsf{F} = \mathsf{S}\Lambda^{1/2}\bar{\mathsf{R}}, \tag{4.186}$$

where Λ is a diagonal matrix in the arbitrary coordinate frame. This writes a matrix as a dilation sandwiched between two rotations, and is called the singular value decomposition of the matrix. An arbitrary linear function in n dimensions has n^2 degrees of freedom. The singular value decomposition assigns $2 \times n(n-1)/2$ of these to the two orthogonal transformations R and S, with the remaining n degrees of freedom contained in the dilation Λ. The singular value decomposition appears frequently in subjects such as data analysis, where it is often used in connection with analysing non-square matrices.

4.5 Tensors and components

Many modern physics textbooks are written in the language of tensor analysis. In this approach one often works directly with the components of a vector, or linear function, in a chosen coordinate frame. The invariance of the laws under a change of frame can then be used to advantage to simplify the component equations. Since this approach is so ubiquitous it is important to establish the relationship between tensor analysis and the largely frame-free approach of the present chapter. We start by analysing Cartesian tensors, and then move onto the more general case of an arbitrary coordinate frame.

4.5.1 Cartesian tensors

The subject of Cartesian tensors arises when we restrict our frames to consist only of orthonormal vectors in Euclidean space. For these we have

$$\mathsf{e}_i \cdot \mathsf{e}_j = \delta_{ij}, \tag{4.187}$$

so there is no distinction between frames and their reciprocals. In this case we can drop all distinction between raised and lowered indices, and just work with all indices lowered. Provided both frames have the same orientation, a new frame is obtained from the $\{\mathsf{e}_k\}$ frame by a rotation,

$$\mathsf{e}'_i = R\mathsf{e}_i R^\dagger = \Lambda_{ij}\mathsf{e}_j. \tag{4.188}$$

Here R is a rotor and Λ_{ij} are the components of the rotation defined by R:

$$\Lambda_{ij} = (R\mathsf{e}_i R^\dagger)\cdot \mathsf{e}_j. \tag{4.189}$$

It follows that

$$\begin{aligned}\Lambda_{ij}\Lambda_{ik} &= (R\mathsf{e}_i R^\dagger)\cdot \mathsf{e}_j (R\mathsf{e}_i R^\dagger)\cdot \mathsf{e}_k \\ &= (R^\dagger \mathsf{e}_j R)\cdot (R^\dagger \mathsf{e}_k R) = \delta_{jk},\end{aligned} \tag{4.190}$$

and similarly

$$\Lambda_{ik}\Lambda_{jk} = \delta_{ij}. \tag{4.191}$$

A vector a has components $a_i = \mathsf{e}_i \cdot a$ and these transform under a change of frame in the obvious manner,

$$a'_i = \mathsf{e}'_i \cdot a = \Lambda_{ij} a_j. \tag{4.192}$$

It is important to realise here that it is only the components of a that change, not the underlying vector itself. The change in components is exactly cancelled by the change in the frame. Many equations in physics are invariant if the vector itself is transformed, but this is the result of an underlying symmetry in the equations, and not of the freedom to choose the coordinate system. These two concepts should not be confused!

Extending this idea, we define the components of the linear function F by

$$\mathsf{F}_{ij} = \mathsf{e}_i \cdot \mathsf{F}(\mathsf{e}_j). \tag{4.193}$$

The result of this decomposition is an $n \times n$ array of components, which can be stored and manipulated as a matrix. This definition ensures that the components of the vector $\mathsf{F}(a)$ are given by

$$\mathsf{e}_i \cdot \mathsf{F}(a) = \mathsf{e}_i \cdot \mathsf{F}(a_j \mathsf{e}_j) = \mathsf{F}_{ij} a_j, \tag{4.194}$$

which is the usual expression for a matrix acting on a column vector. Similarly, if F and G are a pair of linear functions, the components of the product function FG are given by

$$\begin{aligned}(\mathsf{FG})_{ij} &= \mathsf{FG}(\mathsf{e}_j) \cdot \mathsf{e}_i = \mathsf{G}(\mathsf{e}_j) \cdot \bar{\mathsf{F}}(\mathsf{e}_i) \\ &= \mathsf{G}(\mathsf{e}_j) \cdot \mathsf{e}_k \, \mathsf{e}_k \cdot \bar{\mathsf{F}}(\mathsf{e}_i) = \mathsf{F}_{ik} \mathsf{G}_{kj}.\end{aligned} \tag{4.195}$$

This recovers the familiar rule for multiplying matrices. If the frame is changed to a new rotated frame, the components of the tensor transform in the obvious way:

$$\mathsf{F}'_{ij} = \Lambda_{ik} \Lambda_{jl} \mathsf{F}_{kl}, \tag{4.196}$$

where the prime denotes the components in the new (primed) frame. Objects with two indices are referred to as rank-2 tensors. Rank-1 tensors are vectors, rank-3 tensors have three indices, and so on. Since rank-2 tensors appear regularly in physics they are often referred to simply as tensors. Also, it is usual to let the term tensor refer to either the component form F_{ij} or the abstract entity F.

For Cartesian tensors there are two important tensors which arise regularly in computations. These are the two *invariant* tensors. The first of these is the Kronecker δ, which transforms as

$$\delta'_{ij} = \Lambda_{ik} \Lambda_{jl} \delta_{kl} = \Lambda_{ik} \Lambda_{jk} = \delta_{ij}. \tag{4.197}$$

The components of the identity function are therefore the same in all orthonormal frames (and are those of the identity matrix in all cases). The second invariant is the alternating tensor $\epsilon_{ij\cdots k}$, where the number of indices matches the dimension of the space. This is totally antisymmetric and is defined as follows:

$$\epsilon_{ij\cdots k} = \begin{cases} 1 & i,\, j,\, \ldots,\, k = \text{even permutation of } 1, 2, \ldots, n \\ -1 & i,\, j,\, \ldots,\, k = \text{odd permutation of } 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases} . \tag{4.198}$$

The order of a permutation is the number of pairwise swaps required to return to the original order $1, 2, \ldots, n$. If an even number of swaps is required the permutation is even, and similarly for the odd case. In three dimensions even permutations of $1, 2, 3$ coincide with cyclic orderings of the indices. The determinant of a matrix can be expressed in terms of the alternating tensor via

$$\mathsf{F}_{\alpha i}\mathsf{F}_{\beta j}\cdots\mathsf{F}_{\gamma k}\epsilon_{\alpha\beta\cdots\gamma} = \det\,(\mathsf{F})\,\epsilon_{ij\cdots k}. \tag{4.199}$$

Given this result, it is straightforward to prove the frame invariance of the alternating tensor under rotations:

$$\epsilon'_{ij\cdots k} = \Lambda_{i\alpha}\Lambda_{j\beta}\cdots\Lambda_{k\gamma}\epsilon_{\alpha\beta\cdots\gamma} = \det\,(\Lambda)\,\epsilon_{ij\cdots k}. \tag{4.200}$$

But since Λ_{ij} is a rotation matrix it has determinant $+1$, so the tensor is indeed invariant.

4.5.2 The determinant revisited

We should now establish that the definition of the determinant (4.199) agrees with our earlier definition (4.143). To prove this we first need the result that

$$\epsilon_{ij\cdots k} = \mathsf{e}_i\wedge\mathsf{e}_j\cdots\wedge\mathsf{e}_k\, I^\dagger, \tag{4.201}$$

where $I = \mathsf{e}_1\mathsf{e}_2\cdots\mathsf{e}_n$ and the $\{\mathsf{e}_k\}$ form an orthonormal frame. The right-hand side of (4.201) is zero if any of the indices are the same, because of the antisymmetry of the outer product. If the indices form an even permutation of $1, 2, \ldots, n$ we can reorder the vectors into the order $\mathsf{e}_1\mathsf{e}_2\cdots\mathsf{e}_n = I$, in which case the right-hand side of (4.201) returns $+1$. Similarly, any anticyclic combination of $1, 2, \ldots, n$ returns -1. Together these agree with the definition (4.198) of the alternating tensor $\epsilon_{ij\cdots k}$. We can now rearrange the left-hand side of (4.199) as follows:

$$\begin{aligned} \mathsf{F}_{\alpha i}\mathsf{F}_{\beta j}\cdots\mathsf{F}_{\gamma k}\epsilon_{\alpha\beta\cdots\gamma} &= \mathsf{F}_{\alpha i}\mathsf{F}_{\beta j}\cdots\mathsf{F}_{\gamma k}\,\mathsf{e}_\alpha\wedge\mathsf{e}_\beta\cdots\wedge\mathsf{e}_\gamma\, I^\dagger \\ &= \mathsf{F}(\mathsf{e}_i)\wedge\mathsf{F}(\mathsf{e}_j)\cdots\mathsf{F}(\mathsf{e}_k)\, I^\dagger \\ &= \det\,(\mathsf{F})\,\mathsf{e}_i\wedge\mathsf{e}_j\cdots\wedge\mathsf{e}_k\, I^\dagger \\ &= \det\,(\mathsf{F})\,\epsilon_{ij\cdots k}, \end{aligned} \tag{4.202}$$

which recovers the expected result.

We assume that most readers are familiar with the various techniques employed when computing the determinant of an $n \times n$ matrix. These can be found in most elementary textbooks on linear algebra. It is instructive to see how the same results arise in the geometric algebra treatment. We have already established that the determinant of the product of two functions is the product of the determinants, and that taking the adjoint does not change the determinant. To establish a further set of results we first introduce the (non-orthonormal) vectors $\{f_i\}$,

$$f_i \equiv \mathsf{F}(\mathsf{e}_i), \tag{4.203}$$

so that

$$\mathsf{F}_{ij} = \mathsf{e}_i \cdot f_j. \tag{4.204}$$

From equation (4.143) the determinant of F can be written

$$\det(\mathsf{F}) = (f_1 \wedge f_2 \wedge \cdots \wedge f_n) \cdot (\mathsf{e}_n \wedge \cdots \wedge \mathsf{e}_2 \wedge \mathsf{e}_1). \tag{4.205}$$

Expanding this product out in full recovers the standard expression for the determinant of a matrix. The first result we see is that swapping any two of the $\{f_i\}$ changes the sign of the determinant. This is the same as swapping two columns in the matrix F_{ij}. Since matrix transposition does not affect the result, the same is true for interchanging rows.

Next we single out one of the $\{\mathsf{e}_k\}$ vectors and write

$$\begin{aligned}\det(\mathsf{F}) &= (-1)^{j+1} (\mathsf{e}_n \wedge \cdots \check{\mathsf{e}}_j \cdots \wedge \mathsf{e}_1) \cdot \big(\mathsf{e}_j \cdot (f_1 \wedge \cdots \wedge f_n)\big) \\ &= \sum_{k=1}^{n} (-1)^{j+k}\, \mathsf{e}_j \cdot f_k \, (\mathsf{e}_n \wedge \cdots \check{\mathsf{e}}_j \cdots \wedge \mathsf{e}_1) \cdot (f_1 \wedge \cdots \check{f}_k \cdots \wedge f_n).\end{aligned} \tag{4.206}$$

The final part of each term in the sum corresponds to an $(n-1) \times (n-1)$ determinant, as can be seen by comparing with (4.205). This is equivalent to the familiar expression for the expansion of the determinant by the jth row. A further useful result is obtained from the identity

$$f_1 \wedge \cdots \wedge (f_j + \lambda f_k) \wedge \cdots \wedge f_n = f_1 \wedge \cdots \wedge f_j \wedge \cdots \wedge f_n \quad j \neq k. \tag{4.207}$$

This result means that any multiple of the kth row can be added to the jth row without changing the result. The same is true for columns. This is the key to the method of Gaussian elimination for finding a determinant. In this method the matrix is first transformed to upper (or lower) triangular form, so that the determinant is then simply the product of the entries down the leading diagonal. This is numerically a highly efficient method for calculating determinants. We can continue in this manner to give concise proofs of many of the key results for determinants. For a useful summary of these, see Turnbull (1960).

To see how these formulae also lead to the familiar expression for the inverse of a matrix, consider the decomposition:

$$\begin{aligned}\mathsf{F}^{-1}_{ij} &= \mathsf{e}_i \cdot \mathsf{F}^{-1}(\mathsf{e}_j) \\ &= \langle \mathsf{e}_i\, \mathsf{e}_1 \wedge \cdots \wedge \mathsf{e}_n\, \bar{\mathsf{F}}(\mathsf{e}_n \wedge \cdots \wedge \mathsf{e}_1\, \mathsf{e}_j)\rangle \det\,(\mathsf{F})^{-1} \\ &= (-1)^{i+j} \langle \mathsf{F}(\mathsf{e}_1 \wedge \cdots \check{\mathsf{e}}_i \cdots \wedge \mathsf{e}_n)\, \mathsf{e}_n \wedge \cdots \check{\mathsf{e}}_j \cdots \wedge \mathsf{e}_1 \rangle \det\,(\mathsf{F})^{-1}. \end{aligned} \tag{4.208}$$

The term enclosed in angular brackets is the determinant of the $(n-1)\times(n-1)$ matrix obtained from F_{ij} by deleting the ith column and jth row. This is the definition of the i, j cofactor of F_{ij}. Equation (4.208) shows that the components of F^{-1}_{ij} are formed from the transposed matrix of cofactors, divided by the determinant det (F) — the familiar result. Similarly, all other matrix formulae have simple and often elegant counterparts in geometric algebra. Further examples of these are discussed in chapter 11.

4.5.3 General tensors

We now generalise the preceding treatment to the case of arbitrary basis sets in spaces of arbitrary (non-degenerate) signature. One reason for wanting to deal with non-orthonormal frames is that these regularly arise when working in curvilinear coordinate systems. In addition, in mixed signature spaces one has no option since it is impossible to identify a frame with its reciprocal. Suppose, then, that the vectors $\{\mathsf{e}_k\}$ constitute an arbitrary frame for n-dimensional space (of unspecified signature). The reciprocal frame is denoted $\{\mathsf{e}^k\}$ and the two frames are related by

$$\mathsf{e}^i \cdot \mathsf{e}_j = \delta^i_j. \tag{4.209}$$

Equation (4.94) for the reciprocal frame is general and still holds in mixed signature spaces.

As described in section 4.3.2, the vector a has components $(a^1, a^2, \ldots, a^n)$ in the $\{\mathsf{e}_k\}$ frame, and $(a_1, a_2, \ldots, a_n)$ in the $\{\mathsf{e}^k\}$ frame. When working with general coordinate frames we always ensure that upper and lower indices match separately on either side of an expression. Suppose we now form the inner product of two vectors a and b. We can write this as

$$a\cdot b = (a^i \mathsf{e}_i)\cdot(b_j \mathsf{e}^j) = a^i b_j\, \mathsf{e}_i \cdot \mathsf{e}^j = a^i b_j \delta^j_i = a^i b_i. \tag{4.210}$$

The general rule is that sums are only taken over pairs of indices where one is a superscript and the other a subscript. Another way to write an inner product is to introduce the *metric tensor* g_{ij}:

$$g_{ij} = \mathsf{e}_i \cdot \mathsf{e}_j. \tag{4.211}$$

In terms of its components g_{ij} is a symmetric $n \times n$ matrix. The inverse matrix is written as g^{ij} and is given by

$$g^{ij} = \mathsf{e}^i \cdot \mathsf{e}^j. \tag{4.212}$$

It is easily verified that this is the inverse of g_{ij}:

$$g^{ik} g_{kj} = \mathsf{e}^i \cdot \mathsf{e}^k \, \mathsf{e}_k \cdot \mathsf{e}_j = \mathsf{e}^i \cdot \mathsf{e}_j = \delta^i_j. \tag{4.213}$$

Employing the metric tensor we can write the inner product of two vectors in a number of equivalent forms:

$$a \cdot b = a^i b_i = a_i b^i = a^i b^j g_{ij} = a_i b_j g^{ij}. \tag{4.214}$$

Of course, all of these expressions encode the same thing and, unless there is a particular reason to introduce a frame, the index-free expression $a \cdot b$ is usually the simplest to use.

The same ideas extend to expressing the linear function F in a general non-orthonormal frame. We let F act on the frame vector e_j and find the components of the result in the reciprocal frame. The components are then given by

$$\mathsf{F}_{ij} = \mathsf{e}_i \cdot \mathsf{F}(\mathsf{e}_j). \tag{4.215}$$

Again, the set of numbers F_{ij} are referred to as the components of a rank-2 tensor and form an $n \times n$ matrix, the entries of which depend on the choice of frame. Similar expressions exist for combinations of frame vectors and reciprocal vectors, for example,

$$\mathsf{F}^{ij} = \mathsf{F}(\mathsf{e}^j) \cdot \mathsf{e}^i. \tag{4.216}$$

One use of the metric tensor is to interchange between these expressions:

$$\mathsf{F}^{ij} = \mathsf{e}^i \cdot \mathsf{F}(\mathsf{e}^j) = \mathsf{e}^i \cdot \mathsf{e}^k \, \mathsf{e}_k \cdot \mathsf{F}(\mathsf{e}_l \mathsf{e}^l \cdot \mathsf{e}^j) = g^{ik} g^{jl} \mathsf{F}_{kl}. \tag{4.217}$$

Again, we have at our disposal a variety of different ways of encoding the information in F. In terms of the abstract concept of a linear operator, the metric tensor g_{ij} is simply the identity operator expressed in a non-orthonormal frame.

If F_{ij} are the components of F in some frame then the components of $\bar{\mathsf{F}}$ are given by

$$\bar{\mathsf{F}}_{ij} = \bar{\mathsf{F}}(\mathsf{e}_j) \cdot \mathsf{e}_i = \mathsf{e}_j \cdot \mathsf{F}(\mathsf{e}_i) = \mathsf{F}_{ji}. \tag{4.218}$$

That is, viewed as a matrix, the components of $\bar{\mathsf{F}}$ are found from the components of F by matrix transposition. For mixed index tensors we have to be slightly more careful, as we now have

$$\mathsf{F}_i{}^j = \mathsf{F}(\mathsf{e}^j) \cdot \mathsf{e}_i = \mathsf{e}^j \cdot \bar{\mathsf{F}}(\mathsf{e}_i) = \bar{\mathsf{F}}^j{}_i. \tag{4.219}$$

If F is a symmetric function we have $\bar{\mathsf{F}} = \mathsf{F}$. In this case the component matrices satisfy

$$\mathsf{F}_{ij} = \mathsf{F}(\mathsf{e}_j) \cdot \mathsf{e}_i = \mathsf{F}(\mathsf{e}_i) \cdot \mathsf{e}_j = \mathsf{F}_{ji}, \tag{4.220}$$

so the components F_{ij} form a symmetric matrix. The same is true of $\mathsf{F}^{ij} = \mathsf{F}^{ji}$, but for the mixed tensor $\mathsf{F}_i{}^j$ we have $\mathsf{F}_i{}^j = \mathsf{F}^j{}_i$.

The components of the product function FG are found from the following rearrangement:

$$\begin{aligned}(\mathsf{FG})_{ij} &= \mathsf{FG}(\mathsf{e}_j)\cdot\mathsf{e}_i = \mathsf{G}(\mathsf{e}_j)\cdot\bar{\mathsf{F}}(\mathsf{e}_i) \\ &= \mathsf{G}(\mathsf{e}_j)\cdot\mathsf{e}_k\,\mathsf{e}^k\cdot\bar{\mathsf{F}}(\mathsf{e}_i) = \mathsf{F}_i{}^k\mathsf{G}_{kj}.\end{aligned} \tag{4.221}$$

Provided the correct combination of subscript and superscript indices is used, this can be viewed as a matrix product. Alternatively, one can work entirely with subscripted indices, and include suitable factors of the metric tensor,

$$(\mathsf{FG})_{ij} = \mathsf{F}_{ik}\mathsf{G}_{lj}g^{kl}. \tag{4.222}$$

Higher rank linear functions give rise to higher rank tensors. Suppose, for example, that $\phi(a_1, a_2, a_3)$ is a scalar function of three vectors, and is linear on each argument,

$$\phi(\lambda a_1 + \mu b, a_2, a_3) = \lambda\phi(a_1, a_2, a_3) + \mu\phi(b, a_2, a_3), \quad \text{etc.} \tag{4.223}$$

The components of this define a rank-3 tensor via

$$\phi_{ijk} = \phi(\mathsf{e}_i, \mathsf{e}_j, \mathsf{e}_k). \tag{4.224}$$

Using similar schemes it is a straightforward matter to set up a map between tensor equations and frame-free expressions in geometric algebra.

4.5.4 Coordinate transformations

If a second non-orthonormal frame $\{\mathsf{f}_\alpha\}$ is introduced we can relate the two frames via a transformation matrix $f_{\alpha i}$:

$$f_{\alpha i} = \mathsf{f}_\alpha\cdot\mathsf{e}_i, \quad f^{\alpha i} = \mathsf{f}^\alpha\cdot\mathsf{e}^i, \tag{4.225}$$

where Latin and Greek indices distinguish the components in one frame from the other. These matrices satisfy

$$f_{\alpha i}f^{\alpha j} = \mathsf{f}_\alpha\cdot\mathsf{e}_i\,\mathsf{f}^\alpha\cdot\mathsf{e}^j = \mathsf{e}_i\cdot\mathsf{e}^j = \delta_i^j \tag{4.226}$$

and

$$f_{\alpha i}f^{\beta i} = \mathsf{f}_\alpha\cdot\mathsf{e}_i\,\mathsf{f}^\beta\cdot\mathsf{e}^i = \mathsf{f}_\alpha\cdot\mathsf{f}^\beta = \delta_\alpha^\beta. \tag{4.227}$$

The decomposition of the vector a in terms of these frames gives

$$a = a^i\mathsf{e}_i = a^i\mathsf{f}^\alpha\mathsf{e}_i\cdot\mathsf{f}_\alpha = a^i f_{\alpha i}\mathsf{f}^\alpha. \tag{4.228}$$

If follows that the transformation law for the components is

$$a_\alpha = f_{\alpha i}a^i, \tag{4.229}$$

with similar expressions holding for the superscripted components.

These formulae extend simply to include linear functions. For example, we see that

$$\mathsf{F}_{\alpha\beta} = f_{\alpha i} f_{\beta j} \mathsf{F}^{ij}. \tag{4.230}$$

Again, similar expressions hold for superscripts and for mixtures of indices. In particular we have

$$\mathsf{F}_{\alpha}{}^{\beta} = f_{\alpha}{}^{i} f^{\beta}{}_{j} \mathsf{F}_{i}{}^{j}. \tag{4.231}$$

Expressed in terms of matrix multiplication, this would be an equivalence transformation. Of course, the abstract frame-free function F is unaffected by any change of basis. All that changes is the particular representation of the function in the chosen coordinate system. Any set of n^2 numbers with this transformation property are called the components of a rank 2 tensor, the implication being that the underlying function is frame-independent.

In conventional accounts, the subject of tensors is often built up by taking the transformation law as fundamental. That is, a vector (rank-1 tensor) is *defined* as a set of components which transform according to equation (4.229) under a change of basis. Once one has the tools available to treat vectors and linear operations in a frame-free manner, such an approach becomes entirely unnecessary. The defining property of a tensor is that it represents a genuine geometric object (or operation) and does not depend on a choice of frame. Given this, the transformation laws (4.229) and (4.231) follow automatically. In this book the name *tensor* is applied to any frame-independent linear function, such as F. We will encounter a variety of such objects in later chapters.

4.6 Notes

The realisation that geometric algebra is a universal tool for physics was a key point in the modern development of the subject, and was first strongly promoted by David Hestenes (figure 4.5). Before his work, physicists' sole interaction with geometric algebra was through the quantum theory of spin. The Pauli and Dirac matrices form representations of Clifford algebras, a fact that was realised as soon as they were introduced. But in the 50 years since Clifford's original idea, the geometry behind his algebra had been lost as mathematicians concentrated on its algebraic properties. This discovery of the Pauli and Dirac matrices thus gave rise to two mistaken beliefs. The first was that there was something intrinsically quantum-mechanical in the non-commutative properties of the matrices. This is clearly not the case. Clifford died long before quantum theory was first formulated and was motivated entirely by classical geometry, and his algebra is today routinely employed in a range of subjects far removed from quantum theory.

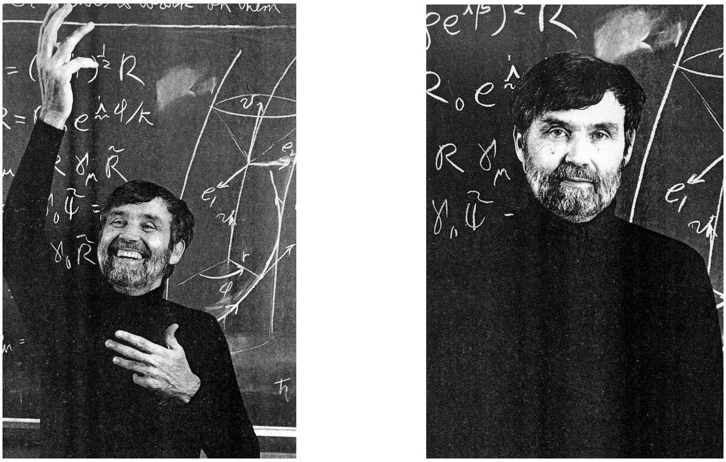

Figure 4.5 *David Hestenes.* Inventor of geometric calculus and first to draw attention to the universal nature of geometric algebra. He wrote the influential *Space-Time Algebra* in 1966, and followed this with a fully developed formalism in *Clifford Algebra to Geometric Calculus* (Hestenes & Sobczyk, 1984). This was followed by the (simpler) *New Foundations for Classical Mechanics*, first published in 1986 (second edition 1999). In a series of papers Hestenes and coworkers showed how geometric algebra could be applied in the study of classical and quantum mechanics, electrodynamics, projective and conformal geometry and Lie group theory. More recently, he has advocated the use of geometric algebra in the field of computer graphics.

The second widespread belief was that matrices were crucial to understanding the properties of Clifford algebras. This too is erroneous. The geometric algebra of a finite-dimensional vector space is an associative algebra, so always has a matrix representation. But these matrices add little, if anything, to understanding the properties of the algebra. Furthermore, an insistence on working with matrices deters one from applying geometric algebra to anything beyond the lowest dimensional spaces, because the size of the matrices increases exponentially with the dimension of the space. Working directly with the elements of the algebra imposes no such constraints, and one can easily apply the ideas to spaces of any dimension, including infinite-dimensional spaces.

Mathematicians had few such misconceptions, and Atiyah and others developed Clifford algebra as a powerful tool for geometry. Even in these developments, however, the emphasis was usually on Clifford algebra as an extra tool on top of the standard techniques for solving geometric problems. The algebra was seldom used as complete language for geometry. The picture first started to change when Hestenes recovered Clifford's original interpretation of the Pauli matrices. This led Hestenes to question whether the appearance of a Clifford algebra was telling us something about the underlying structure of quantum theory. Hestenes then went on to promote the universal nature of the algebra, which he publicised in a series of books and papers. Acceptance of this view is growing and, while not everyone is in full agreement, it is now hard to find an area of physics to which geometric algebra cannot or has not been applied without some degree of success.

4.7 Exercises

4.1 Prove that the outer product of a set of linearly dependent vectors vanishes.

4.2 In a Euclidean space, Gram–Schmidt orthogonalisation proceeds by successively replacing each vector in a set $\{a_i\}$ by one perpendicular to the preceding vectors. Prove that such a vector is given by

$$e_i = a_i - \sum_{j=1}^{i-1} \frac{a_i \cdot e_j}{e_j^2} e_j.$$

Prove that we can also write this as

$$e_i = a_i \wedge a_{i-1} \wedge \cdots \wedge a_1 (a_{i-1} \wedge \cdots \wedge a_1)^{-1}.$$

4.3 Prove that

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

4.4 The length of a vector in Euclidean space is defined by $|a| = \sqrt{(}a^2)$, and the angle θ between two vectors is defined by

$$\cos(\theta) = a \cdot b / (|a||b|).$$

Show that a linear transformation F which leaves lengths and angles unchanged must satisfy

$$\bar{\mathsf{F}} = \mathsf{F}^{-1}.$$

What does this imply for the determinant of F? A reflection in the (hyper)plane perpendicular to n is defined by

$$\mathsf{R}(a) = -nan,$$

where $n^2 = 1$. Show that $\bar{\mathsf{R}} = \mathsf{R}^{-1}$, and that R has determinant -1.

4.5 For the reflection in the preceding question introduce a suitable basis frame and express F in terms of a matrix F_{ij}. Verify the results for the determinant and inverse of this matrix. (Hint — align one of the basis vectors with n.)

4.6 A rotor R is defined by

$$R = \exp(-\lambda B/2).$$

By Taylor expanding in λ, prove that the operation

$$\mathsf{R}(A) = RAR^{\dagger}$$

preserves the grade(s) of the multivector A.

4.7 Show that the plane B is unchanged by the rotation defined by the rotor $R = \exp(B/2)$.

4.8 Analyse the properties of the matrix

$$\begin{pmatrix} 1 & 2\sinh(u) \\ 0 & 1 \end{pmatrix}.$$

To what geometric operation does this matrix correspond? Can this matrix be diagonalised, and does it have a sensible singular value decomposition?

4.9 Suppose that the linear transformation F has a complex eigenvector $e+if$ with associated eigenvector $\alpha + i\beta$. What is the effect of F on the $e\wedge f$ plane? How should one interpret the action of F in this plane?

4.10 Suppose that the vectors $\{\mathsf{e}_k\}$ form an orthonormal basis frame for n-dimensional Euclidean space. What is the effect of the transformation

$$\mathsf{T}(a) = a + \lambda a\cdot\mathsf{e}_1\,\mathsf{e}_2$$

on the rows of the matrix F_{ij} formed by decomposing F in the $\{\mathsf{e}_k\}$ frame? Use this result to prove that the determinant of a matrix is unchanged by adding a multiple of one row to another.