11

Further topics in calculus and group theory

In this chapter we collect together a number of diverse algebraic ideas and techniques. The first part of the chapter deals with some advanced topics in calculus. We introduce the *multivector derivative*, which is a valuable tool in Lagrangian analysis. We also show how the vector derivative can be adapted to provide a compact notation for studying linear functions. We then extend the multivector derivative to the case where we differentiate with respect to a linear function. Finally in this part we look briefly at Grassmann calculus, which is a major ingredient in modern quantum field theory.

The second major topic covered in this chapter is the theory of Lie groups. We provide a detailed analysis of spin groups over a real geometric algebra. By introducing invariant bivectors we show how both the unitary and general linear groups can be represented in terms of spin groups. It then follows that all Lie algebras can be represented as bivector algebras under the commutator product. Working in this way we construct the main Lie groups as subgroups of rotation groups. This is a valuable alternative procedure to the more common method of describing Lie groups in terms of matrices. Throughout this chapter we use the tilde symbol for the reverse, $\tilde{R}$. This avoids confusion with the Hermitian conjugate, which is required in section 11.4 on complex structures.

11.1 Multivector calculus

Before extending our analysis of linear functions in geometric algebra, we first discuss differentiation with respect to a multivector. Suppose that the multivector F is an arbitrary function of some multivector argument X, $F = F(X)$. The derivative of F with respect to X in the A direction is defined by

$$A*\partial_X F(X) = \lim_{\tau \mapsto 0} \frac{F(X+\tau A) - F(X)}{\tau}, \tag{11.1}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

where $A*B = \langle AB \rangle$. The multivector derivative ∂_X is defined in terms of its directional derivatives by

$$\frac{\partial}{\partial X} = \partial_X = \sum_{i<\cdots<j} \mathsf{e}^i \wedge \cdots \wedge \mathsf{e}^j (\mathsf{e}_j \wedge \cdots \wedge \mathsf{e}_i) * \partial_X, \tag{11.2}$$

where the $\{\mathsf{e}^i\}$ are a set of frame vectors for the space of interest. The definition shows how the multivector derivative ∂_X inherits the multivector properties of its argument X, as well as a calculus from equation (11.1). This is the natural generalisation of the vector derivative ∇ to a general multivector.

Most of the properties of the multivector derivative follow from the result that

$$\partial_X \langle XA \rangle = P_X(A), \tag{11.3}$$

where $P_X(A)$ is the projection of A onto the grades contained in X. Leibniz's rule is then used to build up results for more complicated functions. We employ the same rules for the multivector derivative as for the vector derivative. The derivative acts on objects to its immediate right unless brackets are present. If the ∂_X is intended to only act on B then this is written as $\dot{\partial}_X A \dot{B}$, where the overdot denotes the multivector on which the derivative acts. For example, Leibniz's rule can be written as

$$\partial_X (AB) = \dot{\partial}_X \dot{A} B + \dot{\partial}_X A \dot{B}. \tag{11.4}$$

As an example, suppose that ψ is a general even element. The derivative of the scalar product $\langle \psi \tilde{\psi} \rangle$ is

$$\partial_\psi \langle \psi \tilde{\psi} \rangle = \dot{\partial}_\psi \langle \dot{\psi} \tilde{\psi} \rangle + \dot{\partial}_\psi \langle \psi \dot{\tilde{\psi}} \rangle = 2\tilde{\psi}. \tag{11.5}$$

For the second term we used the result that

$$\dot{\partial}_\psi \langle \psi \dot{\tilde{\psi}} \rangle = \dot{\partial}_\psi \langle \dot{\psi} \tilde{\psi} \rangle = \tilde{\psi}, \tag{11.6}$$

which follows from the fact that any scalar term reverses to give itself. This result for the derivative of $\langle \psi \tilde{\psi} \rangle$ can be verified rather more laboriously by expanding out in a basis.

11.1.1 The vector derivative and multilinear algebra

The derivative with respect to a vector was first introduced in chapter 6 as an essential component of field theory. Here we exploit the properties of the vector derivative in a rather different setting. Suppose that a denotes an arbitrary vector. We write the derivative with respect a as ∂_a. Algebraically, this derivative has the properties of a vector. It is essentially the same object as the vector derivative, except that we are not differentiating with respect to the position dependence of a function. Instead we will use ∂_a to differentiate a variety of

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

expressions that are linear in a. Introducing the tools of calculus may appear unnecessary for the analysis of linear algebra, but the notation does have some practical advantages. Combinations of a and ∂_a can be used to perform contractions and protractions without having to introduce a basis frame. For example, the results of section 4.3.2 can be summarised in the compact formulae

$$\begin{aligned} \partial_a a\cdot A_r &= rA_r, \\ \partial_a a\wedge A_r &= (n-r)A_r, \\ \partial_a A_r a &= (-1)^r(n-2r)A_r. \end{aligned} \tag{11.7}$$

Similarly, the vector derivative allows the trace of a linear function to be written simply as

$$\mathrm{tr}(\mathsf{f}) = \partial_a\cdot\mathsf{f}(a). \tag{11.8}$$

The trace is the first of a series of scalar invariants that can be defined from f. These are compactly handled using the vector derivative. Suppose that $\{a_1, a_2, \ldots, a_n\}$ denote a set of n independent vectors. We define the multivector variable

$$a_{(r)} = a_1\wedge a_2\wedge\cdots\wedge a_r \tag{11.9}$$

with the associated derivative

$$\partial_{(r)} = \frac{1}{r!}\partial_{a_r}\wedge\partial_{a_{r-1}}\wedge\cdots\wedge\partial_{a_1}. \tag{11.10}$$

Since

$$\langle A_r\wedge\partial_a\, a\wedge B_r\rangle = (n-r)\langle A_r B_r\rangle, \tag{11.11}$$

it follows that

$$\partial_{(r)}a_{(r)} = \frac{n!}{(n-r)!\,r!} = \begin{pmatrix} n \\ r \end{pmatrix}. \tag{11.12}$$

We also make the further abbreviation

$$\mathsf{f}(a_{(r)}) = \mathsf{f}(a_1)\wedge\mathsf{f}(a_2)\wedge\cdots\wedge\mathsf{f}(a_r) = \mathsf{f}_{(r)}. \tag{11.13}$$

This notation allows us to write

$$\partial_{(1)}\cdot\mathsf{f}_{(1)} = \partial_{a_1}\cdot\mathsf{f}(a_1) = \mathrm{tr}(\mathsf{f}) \tag{11.14}$$

and

$$\partial_{(n)}\mathsf{f}_{(n)} = \partial_{(n)}a_{(n)}\det\,(\mathsf{f}) = \det\,(\mathsf{f}). \tag{11.15}$$

These two invariants are clearly special cases of the range of invariants $\partial_{(r)}\cdot\mathsf{f}_{(r)}$.

To understand the importance of the $\partial_{(r)}\cdot\mathsf{f}_{(r)}$ invariants, consider the characteristic polynomial for f. This is formed by constructing the determinant of the

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

function $\mathsf{G}(a) = \mathsf{f}(a) - \lambda a$, which yields

$$\begin{aligned} \det\,(\mathsf{G}) &= \partial_{(n)} \mathsf{G}_{(n)} \\ &= \partial_{(n)} \big(\mathsf{f}(a_1) - \lambda a_1\big) \wedge \big(\mathsf{f}(a_2) - \lambda a_2\big) \wedge \cdots \wedge \big(\mathsf{f}(a_n) - \lambda a_n\big) \\ &= \partial_{(n)} \big(\mathsf{f}_{(n)} - n\lambda \mathsf{f}_{(n-1)} \wedge a_n + \cdots + (-\lambda)^n a_{(n)}\big). \end{aligned} \tag{11.16}$$

A general term in this expression goes as

$$(-\lambda)^s \begin{pmatrix} n \\ s \end{pmatrix} \partial_{(n)} \cdot \big(\mathsf{f}_{(n-s)} \wedge a_{n-s+1} \wedge \cdots \wedge a_n\big) = (-\lambda)^s \partial_{(n-s)} \cdot \mathsf{f}_{(n-s)}. \tag{11.17}$$

It follows that the characteristic polynomial is simply

$$C(\lambda) = \sum_{s=0}^{n} (-\lambda)^{n-s} \partial_{(s)} \cdot \mathsf{f}_{(s)}, \tag{11.18}$$

where $\partial_{(0)} \cdot \mathsf{f}_{(0)} = 1$. This expression clearly demonstrates the significance of the invariant quantities $\partial_{(r)} \cdot \mathsf{f}_{(r)}$.

The *Cayley–Hamilton* theorem states that

$$\sum_{s=0}^{n} (-1)^{n-s} \partial_{(s)} \cdot \mathsf{f}_{(s)}\, \mathsf{f}^{n-s}(a) = 0, \tag{11.19}$$

where $\mathsf{f}^r(a)$ denotes the r-fold application of f on a. This says that a linear function satisfies its own characteristic equation. The theorem can be proved quite generally without any assumptions about the form of f — it applies for any linear function, in any linear space of any dimension and signature. An immediate consequence is that, if e is an eigenvector of f,

$$\mathsf{f}(e) = \lambda e, \tag{11.20}$$

then λ automatically satisfies the characteristic equation.

11.1.2 Calculus for linear functions

As well as the ability to differentiate with respect to a multivector, it is also very useful to build up results for the derivative with respect to a linear function. We start by introducing a fixed frame $\{\mathsf{e}_i\}$, and define the scalar coefficients

$$\mathsf{f}_{ij} = \mathsf{e}_i \cdot \mathsf{f}(\mathsf{e}_j). \tag{11.21}$$

Now consider the derivative with respect to f_{ij} of the scalar $\mathsf{f}(b) \cdot c$. This is

$$\begin{aligned} \partial_{\mathsf{f}_{ij}} \mathsf{f}(b) \cdot c &= \partial_{\mathsf{f}_{ij}} (\mathsf{f}_{lk} b^k c^l) \\ &= c^i b^j. \end{aligned} \tag{11.22}$$

Multiplying both sides of this equation by $a \cdot \mathsf{e}_j\, \mathsf{e}_i$ we obtain

$$a \cdot \mathsf{e}_j\, \mathsf{e}_i \partial_{\mathsf{f}_{ij}} \mathsf{f}(b) \cdot c = a \cdot b\, c, \tag{11.23}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

which assembles a frame-independent vector on the right-hand side. It follows that the operator $a \cdot \mathsf{e}_j \, \mathsf{e}_i \partial_{f_{ij}}$ must also be frame-independent. We therefore define the vector-valued differential operator $\partial_{\mathsf{f}(a)}$ by

$$\partial_{\mathsf{f}(a)} = a \cdot \mathsf{e}_j \, \mathsf{e}_i \partial_{f_{ij}}. \tag{11.24}$$

The essential property of $\partial_{\mathsf{f}(a)}$ is

$$\partial_{\mathsf{f}(a)} \mathsf{f}(b) \cdot c = a \cdot b \, c, \tag{11.25}$$

which simply restates equation (11.23). As with the vector derivative, $\partial_{\mathsf{f}(a)}$ has the algebraic properties of a vector, which can be exploited in analysing a range of expressions.

Equation (11.25), together with Leibniz's rule, is sufficient to derive the main results for the $\partial_{\mathsf{f}(a)}$ operator. For example, suppose that B is a bivector, and we construct

$$\begin{aligned}\partial_{\mathsf{f}(a)} \langle \mathsf{f}(b \wedge c) B \rangle &= \dot{\partial}_{\mathsf{f}(a)} \langle \dot{\mathsf{f}}(b) \mathsf{f}(c) B \rangle - \dot{\partial}_{\mathsf{f}(a)} \langle \dot{\mathsf{f}}(c) \mathsf{f}(b) B \rangle \\ &= a \cdot b \, \mathsf{f}(c) \cdot B - a \cdot c \, \mathsf{f}(b) \cdot B \\ &= \mathsf{f}\big(a \cdot (b \wedge c)\big) \cdot B. \end{aligned} \tag{11.26}$$

This extends by linearity to give

$$\partial_{\mathsf{f}(a)} \langle \mathsf{f}(A) B \rangle = \mathsf{f}(a \cdot A) \cdot B, \tag{11.27}$$

where A and B are both bivectors. Proceeding in this manner, we obtain the general formula

$$\partial_{\mathsf{f}(a)} \langle \mathsf{f}(A) B \rangle = \sum_r \langle \mathsf{f}(a \cdot A_r) B_r \rangle_1. \tag{11.28}$$

For a fixed grade-r multivector A_r, we can now write

$$\begin{aligned}\partial_{\mathsf{f}(a)} \mathsf{f}(A_r) &= \partial_{\mathsf{f}(a)} \langle \mathsf{f}(A_r) \dot{X}_r \rangle \dot{\partial}_{X_r} \\ &= \mathsf{f}(a \cdot A_r) \cdot \dot{X}_r \, \dot{\partial}_{X_r} \\ &= (n - r + 1) \mathsf{f}(a \cdot A_r). \end{aligned} \tag{11.29}$$

This is a very powerful result. For example, suppose that for A_r we take the pseudoscalar I. We obtain

$$\partial_{\mathsf{f}(a)} \mathsf{f}(I) = \partial_{\mathsf{f}(a)} \det\,(\mathsf{f}) I = \mathsf{f}(a \cdot I). \tag{11.30}$$

It follows that

$$\partial_{\mathsf{f}(a)} \det\,(\mathsf{f}) = \det\,(\mathsf{f}) \bar{\mathsf{f}}^{-1}(a), \tag{11.31}$$

where we have used equation (4.152) This derivation is considerably more compact than any available to conventional matrix/tensor methods.

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

Equation (11.28) can be used to derive formulae for the functional derivative of the adjoint. The general result is

$$\begin{aligned}\partial_{\mathsf{f}(a)}\bar{\mathsf{f}}(A_r) &= \partial_{\mathsf{f}(a)}\langle \mathsf{f}(\dot{X}_r)A_r\rangle\dot{\partial}_{X_r} \\ &= \mathsf{f}(a\cdot\dot{X}_r)\cdot A_r\,\dot{\partial}_{X_r}.\end{aligned} \tag{11.32}$$

When A is a vector, this admits the simpler form

$$\partial_{\mathsf{f}(a)}\bar{\mathsf{f}}(b) = ba. \tag{11.33}$$

If f is a symmetric function then $\mathsf{f} = \bar{\mathsf{f}}$. But this fact cannot be exploited when differentiating with respect to f, since f_{ij} and f_{ji} must be treated as independent variables for the purposes of calculus.

11.2 Grassmann calculus

For most of his lifetime, Grassmann's work on algebra and geometry was largely ignored by the wider mathematical community. Today, however, Grassmann algebra is a fundamental ingredient in theoretical physics. Fermionic creation operators generate a Grassmann algebra, and Grassmann (anticommuting) variables are important components of path-integral quantisation, supersymmetry and string theory. In this section we describe how the main algebraic results of Grassmann calculus can be formulated in a straightforward manner within geometric algebra. This reverses the standard approach, by which one progresses from Grassmann to Clifford algebra via quantization.

Suppose that $\{\zeta_i\}$ are a set of n Grassmann variables, satisfying the anticommutation relations

$$\{\zeta_i, \zeta_j\} = 0. \tag{11.34}$$

The Grassmann variables $\{\zeta_i\}$ are mapped into geometric algebra by introducing a set of n linearly independent vectors $\{e_i\}$. We do not need to specify any properties for their inner products, though some calculations are performed more easily if we assume that the $\{e_i\}$ belong to a Euclidean algebra. The role of the product of Grassmann variables is taken over by the exterior product in geometric algebra, so we write

$$\zeta_i\zeta_j \;\leftrightarrow\; e_i \wedge e_j. \tag{11.35}$$

Equation (11.34) is satisfied by virtue of the antisymmetry of the exterior product. Any combination of Grassmann variables can now be replaced in the obvious manner by a multivector.

In order for the above scheme to have computational power, we need a translation for the Grassmann calculus introduced by Berezin. In this calculus, dif-

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

ferentiation is defined by the rules

$$\frac{\partial \zeta_j}{\partial \zeta_i} = \delta_{ij}, \qquad \zeta_j \frac{\overleftarrow{\partial}}{\partial \zeta_i} = \delta_{ij}, \tag{11.36}$$

together with the graded Leibniz rule,

$$\frac{\partial}{\partial \zeta_i}(f_1 f_2) = \frac{\partial f_1}{\partial \zeta_i} f_2 + (-1)^{[f_1]} f_1 \frac{\partial f_2}{\partial \zeta_i}, \tag{11.37}$$

where $[f_1]$ is the parity of f_1. The parity of a Grassmann variable is determined by whether it contains an even or odd number of vectors. Berezin differentiation is handled within the algebra generated by the $\{e_i\}$ frame by introducing the reciprocal frame $\{e^i\}$, and replacing

$$\frac{\partial}{\partial \zeta_i} f \; \leftrightarrow \; e^i \cdot f \tag{11.38}$$

so that

$$\frac{\partial \zeta_j}{\partial \zeta_i} \; \leftrightarrow \; e^i \cdot e_j = \delta^i_j. \tag{11.39}$$

The graded Leibniz rule follows from the basic identities of geometric algebra. For example, if f_1 and f_2 are grade-1 and so are treated as vectors in geometric algebra, then the rule (11.37) simply restates the familiar result

$$e^i \cdot (f_1 \wedge f_2) = e^i \cdot f_1 \, f_2 - f_1 \, e^i \cdot f_2. \tag{11.40}$$

Right action by a Grassmann derivative operator translates in a similar manner:

$$(f) \frac{\overleftarrow{\partial}}{\partial \zeta_i} \; \leftrightarrow \; f \cdot e^i. \tag{11.41}$$

The standard results for Grassmann calculus follow simply from this basic translation scheme.

Grassmann integration is defined to be essentially the same operation as right differentiation:

$$\int f(\zeta) d\zeta_n d\zeta_{n-1} \cdots d\zeta_1 = f(\zeta) \frac{\overleftarrow{\partial}}{\partial \zeta_n} \frac{\overleftarrow{\partial}}{\partial \zeta_{n-1}} \cdots \frac{\overleftarrow{\partial}}{\partial \zeta_1}. \tag{11.42}$$

The equivalent operation in geometric algebra is therefore a right-sided contraction, as given in equation (11.38). The most important formula is that for the total integral

$$\int f(\zeta) d\zeta_n d\zeta_{n-1} \cdots d\zeta_1 \; \leftrightarrow \; (\cdots ((F \cdot e^n) \cdot e^{n-1}) \cdots) \cdot e^1 = \langle F E^n \rangle, \tag{11.43}$$

where F is the multivector equivalent of $f(\zeta)$ and E^n is the pseudoscalar for the $\{e^i\}$ vectors,

$$E^n = e^n \wedge e^{n-1} \wedge \cdots \wedge e^1. \tag{11.44}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

Equation (11.43) does nothing more than pick out the coefficient of the pseudoscalar part of F.

A 'change of variables' is performed by a linear transformation f, with

$$e_i' = \mathsf{f}(e_i), \qquad e^{i\prime} = \bar{\mathsf{f}}^{-1}(e^i). \tag{11.45}$$

It follows that

$$E_n' = \det\,(\mathsf{f})\,E_n, \qquad E^{n\prime} = \det\,(\mathsf{f})^{-1}\,E^n, \tag{11.46}$$

so that a change of variables in a Grassmann multiple integral picks up a Jacobian factor of $\det\,(\mathsf{f})^{-1}$. This contrasts with the factor of $\det\,(\mathsf{f})$ for a Riemannian integral. In a similar manner all of the main results of Grassmann calculus can be derived in geometric algebra. Often these derivations are simpler, as access to the geometric product offers a quick route through the algebra.

11.3 Lie groups

In earlier chapters we saw that rotors form a continuous group, in the same way that rotations do. Continuous groups of this type are called *Lie groups*, after the mathematician Sophus Lie, and they play an important role in a wide range of subjects in physics. Lie groups contain an infinite number of elements but, like vector spaces, the elements can usually be written in terms of a finite number of parameters. For example, three-dimensional rotations can be parameterised in terms of the three Euler angles. The reason is that the elements of the group belong to a topological space — the *group manifold*. In two-dimensional Euclidean space all rotors correspond to phase factors, so the rotor group manifold is the unit circle. Every point on the circle corresponds to a distinct rotor.

Similarly, in three dimensions rotors are built from the space of scalars and bivectors. The only condition they have to satisfy is that $R\tilde{R} = 1$. Suppose that we write

$$R = x_0 + x_1 I e_1 + x_2 I e_2 + x_3 I e_3. \tag{11.47}$$

Then

$$R\tilde{R} = {x_0}^2 + {x_1}^2 + {x_2}^2 + {x_3}^2 = 1. \tag{11.48}$$

This defines a unit vector in the four-dimensional space spanned by $\{x_0, x_i\}$. The group manifold is therefore the set of unit vectors in four-dimensional space. This is called a 3-sphere S^3 — it is the four-dimensional analogue of the surface of a ball. In higher dimensions the rotor group manifolds become increasingly more complicated.

Since all rotations are generated by the double-sided formula $Ra\tilde{R}$, both R and $-R$ correspond to the same rotation. The group manifold for three-dimensional rotations, rather than for the rotors themselves, is therefore more complicated

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

that S^3. It involves taking a 3-sphere and projectively identifying opposite points. The fact that the group manifold for rotors is somewhat simpler than that for rotations has many applications. If the orientation of a rigid body is described by a rotor, the configuration space for the dynamics of the rigid body is a 3-sphere. This is important when looking for best-fit rotations, or extrapolating between two rotations to find their midpoint. The group manifold is also the appropriate setting for a Lagrangian treatment. This has implications for constructing conjugate momenta, which are essential for the transition to a quantum theory. Applications of this include the rotational energy levels of molecules, many of which can be viewed as rigid bodies.

11.3.1 Formal definitions

The fact that the elements of a Lie group belong to a manifold is sufficient to provide an abstract definition of a general Lie group. A Lie group is defined as a manifold, $\mathcal{M}$, together with a product $\phi(x,y)$. Points on the manifold can be labelled with vectors $\{x,y\}$, which can be viewed as lying in a higher dimensional embedding space (as with the 3-sphere). The product $\phi(x,y)$ takes as its argument two points in the manifold, and returns a third. This encodes the group product. The final set of conditions apply to $\phi(x,y)$ and ensure that the product has the correct group properties. These are

(i) *Closure.* $\phi(x,y) \in \mathcal{M} \quad \forall x,y \in \mathcal{M}$.
(ii) *Identity.* There exists an element $e \in \mathcal{M}$ such that $\phi(e,x) = \phi(x,e) = x$, $\forall x \in \mathcal{M}$.
(iii) *Inverse.* For every element $x \in \mathcal{M}$ there exists a unique element $\bar{x}$ such that $\phi(x,\bar{x}) = \phi(\bar{x},x) = e$.
(iv) *Associativity.* $\phi\big(\phi(x,y),z\big) = \phi\big(x,\phi(y,z)\big), \quad \forall x,y,z \in \mathcal{M}$.

Any manifold with a product defined on it with the preceding properties is called a Lie group manifold. Many of the group properties of the group can be uncovered by examining the properties near the identity element. The product then induces a *Lie bracket* structure on elements of the tangent space at the identity. The tangent space is a linear space and the vectors in this space, together with their bracket, form a Lie algebra.

11.3.2 Spin groups and the bivector algebra

The general theory of Lie groups is rather too abstract for our purposes. Instead, we will adopt a different approach to the subject by concentrating on the properties of rotors, and their associated spin groups. The Lie algebra of a spin group is defined by a set of bivectors. We will establish that every Lie algebra

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

can be represented as a bivector algebra, and that every matrix Lie group can be represented in terms of a spin group.

Before proceeding, we need to clarify some of the terminology for the various groups discussed in this chapter. We let $\mathcal{G}(p,q)$ denote the geometric algebra of a space of signature p,q, and write $\mathcal{V}$ for the space of grade-1 vectors. The orthogonal group $\mathrm{O}(p,q)$ is the set of all linear transformations f mapping $\mathcal{V} \mapsto \mathcal{V}$ that preserve the inner product. That is,

$$\bar{\mathsf{f}}\mathsf{f}(a) = a \quad \forall a \in \mathcal{V}. \tag{11.49}$$

Orthogonal transformations can have determinant 1 or -1. The special orthogonal group $\mathrm{SO}(p,q)$ is the subgroup of $\mathrm{O}(p,q)$ of linear transformations with determinant 1. Orthogonal transformations can be constructed from series of reflections, each of which can be written as

$$a \mapsto -mam^{-1}, \tag{11.50}$$

where m is a non-null vector. Reflections have determinant -1, so do not belong to $\mathrm{SO}(p,q)$. If we restrict m to be a unit vector, $m^2 = \pm 1$, then the set of all unit vectors form a group under the geometric product. This is called the *pin* group, $\mathrm{Pin}(p,q)$. The pin group is a double-cover representation of the orthogonal group. The elements of the pin group all satisfy

$$M\tilde{M} = \pm 1 \quad \forall M \in \mathrm{Pin}(p,q). \tag{11.51}$$

The elements of the pin group split into those of even grade, and those of odd grade. The even-grade elements form a subgroup called the *spin* group, $\mathrm{Spin}(p,q)$. The spin group consists of even-grade multivectors $S \in \mathcal{G}(p,q)$ satisfying

$$SaS^{-1} \in \mathcal{V} \quad \forall a \in \mathcal{V}, \qquad S\tilde{S} = \pm 1. \tag{11.52}$$

The transformations defined by S all have determinant $+1$, so the spin group is a double-cover representation of the special orthogonal group $\mathrm{SO}(p,q)$.

Rotors are elements of the spin group satisfying the further constraint that $R\tilde{R} = 1$. These define the *rotor* group, sometimes denoted $\mathrm{Spin}^+(p,q)$. For rotors we have $R^{-1} = \tilde{R}$, and their action on multivectors is defined by the familiar double-sided formula

$$M \mapsto RM\tilde{R}. \tag{11.53}$$

With the exception of rotors in $\mathcal{G}(1,1)$, the rotor group is a subgroup of the spin group consisting of elements that are *connected* to the identity. That is, all elements of the rotor group can be connected to the identity by a single unbroken path in the group manifold. It follows that rotors form a double-cover representation of the connected subgroup of $\mathrm{SO}(p,q)$. For Euclidean spaces the special orthogonal group is connected, and for these spaces there is no distinction

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

between the spin group and rotor group. In mixed signature spaces the spin group differs from the rotor group by the direct product with a discrete group. For example, the rotor group in spacetime is a representation of the group of proper orthochronous transformations (see section 5.4).

In Euclidean spaces we know that all rotations can be written as the exponential of a bivector. The natural question now is can any rotor be written as the exponential of a bivector? To answer this question, consider a family of rotors $R(\lambda)$, which specifies a path on the rotor group manifold. Differentiating the normalisation condition $R\tilde{R} = 1$ we find that

$$\frac{d}{d\lambda}(R\tilde{R}) = 0 = R'\tilde{R} + R\tilde{R}', \tag{11.54}$$

where the primes denote differentiation with respect to λ. Now define the set vectors

$$a(\lambda) = R(\lambda)a_0\tilde{R}(\lambda), \tag{11.55}$$

where a_0 is some fixed initial vector. Differentiating this expression we find that

$$\frac{d}{d\lambda}a(\lambda) = R'a_0\tilde{R} + Ra_0\tilde{R}' = (R'\tilde{R})a(\lambda) - a(\lambda)(R'\tilde{R}). \tag{11.56}$$

The quantity $R'\tilde{R}$ reverses to minus itself, so can only contain terms of grade 2, 6, 10 etc. But the commutator of $R'\tilde{R}$ with any vector must return another vector, otherwise the derivative of $a(\lambda)$ would grow non-vector terms. It follows that $R'\tilde{R}$ can only contain a bivector component. We can therefore write

$$\frac{d}{d\lambda}R(\lambda) = -\tfrac{1}{2}B(\lambda)R(\lambda). \tag{11.57}$$

Locally, around any rotor, we can write

$$R(\lambda + \delta\lambda) = (1 - \tfrac{1}{2}\delta\lambda\, B)R(\lambda) = \exp(-\delta\lambda\, B/2)R(\lambda). \tag{11.58}$$

In this way, bivectors capture all of the local information about the rotor group. All 'nearby' rotors differ by a term that is the exponential of a bivector.

Now suppose we look for paths satisfying

$$R(0) = 1, \qquad R(\lambda + \mu) = R(\lambda)R(\mu). \tag{11.59}$$

The set $R(\lambda)$ form a one-parameter subgroup of the rotor group. For the case of three-dimensional rotations the interpretation of this subgroup is clear — it is the group of all rotations in a fixed plane. For this path we find that

$$\begin{aligned}\frac{d}{d\lambda}R(\lambda+\mu) &= -\tfrac{1}{2}B(\lambda+\mu)R(\lambda+\mu) \\ &= \frac{d}{d\lambda}\big(R(\lambda)R(\mu)\big) \\ &= -\tfrac{1}{2}B(\lambda)R(\lambda)R(\mu).\end{aligned} \tag{11.60}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

It follows that B is constant along this curve. We can therefore integrate equation (11.57) to get

$$R(\lambda) = \mathrm{e}^{-\lambda B/2}. \tag{11.61}$$

This confirms that all rotors near the origin can be written as the exponential of a bivector. For Euclidean space it turns out that all rotors lie on a path described by equation (11.59) and so can be written as the exponential of a bivector. This is not the case in mixed signature spaces, though it does turn out that in Lorentzian spaces every rotor can be written as

$$R(\lambda) = \pm\mathrm{e}^{-\lambda B/2}. \tag{11.62}$$

It is instructive to establish the inverse result that the exponential of a bivector always returns a rotor. To see this, return to the one-parameter family of vectors

$$a(\lambda) = \mathrm{e}^{-\lambda B/2} a_0 \mathrm{e}^{\lambda B/2}. \tag{11.63}$$

To establish that these are the result of rotations we need only establish that a is a vector, as the remaining properties follow automatically. Differentiating with respect to λ, we find that

$$\begin{aligned} \frac{da}{d\lambda} &= \mathrm{e}^{-\lambda B/2} a_0 \cdot B\, \mathrm{e}^{\lambda B/2}, \\ \frac{d^2 a}{d\lambda^2} &= \mathrm{e}^{-\lambda B/2} (a_0 \cdot B) \cdot B\, \mathrm{e}^{-\lambda B/2} \quad \text{etc.} \end{aligned} \tag{11.64}$$

For every extra derivative we pick up a further inner product with the bivector B. It follows that every term in the Taylor series of $a(\lambda)$ is a vector, and the overall operation is grade-preserving, as it must be. We have also proved the following useful Taylor expansion:

$$\mathrm{e}^{-B/2} a \mathrm{e}^{B/2} = a + a \cdot B + \frac{1}{2!}(a \cdot B) \cdot B + \cdots . \tag{11.65}$$

This series is convergent for all bivectors B.

11.3.3 Examples of rotor groups

The preceding definitions are illustrated neatly by the algebras $\mathcal{G}(1,1)$ and $\mathcal{G}(1,2)$. First suppose that γ_0 and γ_1 are basis vectors for $\mathcal{G}(1,1)$, with $\gamma_0^2 = 1$ and $\gamma_1^2 = -1$. The spin group consists of even-grade elements, which take the form $\alpha + \beta\gamma_1\gamma_0$. The restriction that $\psi\tilde{\psi} = \pm 1$ becomes

$$\alpha^2 - \beta^2 = \pm 1, \tag{11.66}$$

which defines four unconnected hyperbolic curves. The rotor group consists of the subgroup for which $\alpha^2 - \beta^2 = 1$. This defines two unconnected branches of a hyperbola, so the rotor group in $\mathcal{G}(1,1)$ is not connected. For the case

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

of Euclidean spaces the scalar product $\langle\psi\tilde{\psi}\rangle$ is positive definite, so there is no difference between the spin and rotor groups, which are always connected.

Now suppose we add a further vector γ_2 of negative signature, and write a general even element as

$$R = R_0 + R_1\gamma_1\gamma_0 + R_2\gamma_2\gamma_0 + R_3\gamma_1\gamma_2. \tag{11.67}$$

The rotor group is specified by the single extra condition that $R\tilde{R} = 1$, which becomes

$$(R_0)^2 - (R_1)^2 - (R_2)^2 + (R_3)^2 = 1. \tag{11.68}$$

It follows that we can write

$$R = \cosh(\alpha)\big(\cos(\theta) + \sin(\theta)\gamma_1\gamma_2\big) + \sinh(\alpha)\big(\cos(\phi) + \sin(\phi)\gamma_1\gamma_2\big)\gamma_1\gamma_0. \tag{11.69}$$

This parameterisation confirms that the group must now be connected. Given an arbitrary rotor we simply find the values of the parameters (α, θ, ϕ), then smoothly run them down to zero to establish a path in the group manifold that connects the rotor to the identity. The reason we can do this in $\mathcal{G}(1,2)$ but could not in $\mathcal{G}(1,1)$ is that the former contains a bivector generator of negative signature. This ensures that -1 is connected to the identity. Among all algebras $\mathcal{G}(p,q)$, with $p+q>1$, the algebra $\mathcal{G}(1,1)$ is unique in containing no bivector with negative square.

While the rotor group in $\mathcal{G}(1,2)$ is connected, it is straightforward to construct examples of rotors that cannot be written as the exponential of a bivector. For example, consider the rotor

$$R = \exp\big((\gamma_0+\gamma_1)\gamma_2\big) = 1 + (\gamma_0+\gamma_1)\gamma_2. \tag{11.70}$$

While this rotor clearly is the exponential of a bivector, it is impossible to write the rotor $-R$ in this way. This is why the strongest statement that can be made about rotors in a mixed signature space is that they can be written as $\pm\exp(-B/2)$.

11.3.4 The bivector algebra

The operation of commuting a multivector with a bivector is always grade-preserving. In particular, the commutator of a bivector with a second bivector produces a third bivector. That is, the space of bivectors is closed under the commutator product. This closed algebra defines the *Lie algebra* of the associated rotor group. The group is formed from the algebra by the act of exponentiation. The commutator of two bivectors expresses the fact that rotations do not commute. If we apply a pair of rotations, and then perform the back rotations in the incorrect order, the result is the new rotation

$$Ra\tilde{R} = \tilde{R}_2\tilde{R}_1(R_2R_1a\tilde{R}_1\tilde{R}_2)R_1R_2. \tag{11.71}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

Now suppose that we are working close to the identity, so that we can write

$$R = \mathrm{e}^{-B/2} = \mathrm{e}^{B_2/2}\mathrm{e}^{B_1/2}\mathrm{e}^{-B_2/2}\mathrm{e}^{-B_1/2}. \tag{11.72}$$

Expanding the exponentials we find that

$$B = B_1 \times B_2 + \text{higher order terms}. \tag{11.73}$$

This is an example of a more general result known as the *Baker–Campbell–Hausdorff* formula. This states that if

$$\mathrm{e}^C = \mathrm{e}^A \mathrm{e}^B, \tag{11.74}$$

then we have

$$C = A + B + A \times B + \frac{1}{3}\big(A \times (A \times B) + B \times (B \times A)\big) + \cdots. \tag{11.75}$$

The series converges for generators of rotors sufficiently close to the identity. (The precise definition of 'sufficiently close' was clarified by Hausdorff.)

Now suppose that we write

$$R_1 = \exp(-\lambda B_1/2), \qquad R_2 = \exp(-\lambda B_2/2), \tag{11.76}$$

so that $R(\lambda)$ is a path in the group manifold. Equation (11.73) ensures that

$$R(\lambda) = 1 - \lambda^2 B_1 \times B_2/2 + \cdots. \tag{11.77}$$

In the tangent space at the identity the new generator is the commutator of the two original bivectors. The bivector algebra must therefore be closed under the commutator product. This is the way in which the local structure of a rotor group around the identity is passed to the bivector algebra. In the abstract theory of Lie groups, the Lie algebra elements are acted on by the Lie bracket, which is antisymmetric and satisfies the *Jacobi identity*. For a rotor group the Lie bracket is simply the commutator product for bivectors. The Jacobi identity for the Lie algebra then reduces to the identity

$$(A \times B) \times C + (C \times A) \times B + (B \times C) \times A = 0, \tag{11.78}$$

which holds for any three bivectors A, B and C.

11.3.5 Structure constants and the Killing form

Suppose now that we introduce a basis set of bivectors $\{B_i\}$. The commutator of any pair of these returns a third bivector, which can also be expanded in terms of the basis set. We can therefore write

$$B_j \times B_k = C^i_{jk} B_i. \tag{11.79}$$

The C^i_{jk} are called the *structure constants* of the Lie algebra. They provide one of the most compact encodings of the group properties, since knowledge of

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

the bracket structure is sufficient to recover most of the properties of the group. The structure constants also provide a route to solving the problem of classifying all possible Lie algebras over the real and complex fields. The solution of this problem was a significant achievement, completed by the mathematician Élie Cartan.

The *adjoint* representation of a Lie group is defined in terms of functions mapping the Lie algebra onto itself. Every element of a Lie group induces an adjoint representation through its action on the Lie algebra. For the case of rotor groups the Lie algebra is the bivector algebra, and the adjoint representation consists of a map of the form

$$B \mapsto RB\tilde{R} = \mathrm{Ad}_R(B). \tag{11.80}$$

It is immediately clear that this representation satisfies

$$\mathrm{Ad}_{R_1}\big(\mathrm{Ad}_{R_2}(B)\big) = \mathrm{Ad}_{R_1R_2}(B). \tag{11.81}$$

The adjoint representation of the group induces an adjoint representation $\mathrm{ad}_{A/2}$ of the Lie algebra as

$$\mathrm{ad}_{A/2}(B) = A \times B. \tag{11.82}$$

The adjoint representation of an element of the Lie algebra can be considered as a linear map on the space of bivectors. The matrix corresponding to the adjoint representation of the basis bivector B_j is defined by the structure coefficients

$$(\mathrm{ad}_{B_j})^i_k = 2C^i_{jk}. \tag{11.83}$$

The Killing form for a Lie algebra is defined through the adjoint representation as

$$K(A,B) = \mathrm{tr}(\mathrm{ad}_A \mathrm{ad}_B). \tag{11.84}$$

Up to an irrelevant normalisation, the Killing form for a bivector algebra is simply the inner product

$$K(A,B) = A \cdot B, \tag{11.85}$$

which is the definition we shall adopt. It is immediately clear that rotor groups in Euclidean space have a negative-definite Killing form. An algebra with a negative-definite Killing form is said to be of compact type, and the associated Lie group is compact.

11.4 Complex structures and unitary groups

So far we have only dealt with the properties of real rotation groups, but it turns out that this is sufficient for us to uncover the properties of all Lie algebras. We can start to see how this works by studying how complex groups fit into our *real*

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

geometric algebra. The ideas developed in this section are useful in a number of areas, particularly Hamiltonian dynamics and geometric quantum mechanics.

11.4.1 Complex spaces

The simplest algebraic way to define a complex structure is to introduce a commuting scalar quantity j with the property $j^2 = -1$, and to add the assumption that all linear superpositions are now taken over the complex field. A more attractive, geometric alternative is to work in a real space of dimension $2n$ and introduce a bivector in this space to play the role of the complex structure. We saw in section 6.3 that complex analysis can be performed in the geometric algebra of the real two-dimensional plane with the role of the unit imaginary played by the unit pseudoscalar. Here we generalise this idea to an n-dimensional complex space.

Our starting point is a real n-dimensional vector space. Suppose that this has some arbitrary basis $\{e_k\}$, which need not be orthonormal. Now introduce a further set of n-vectors $\{f_k\}$ perpendicular to the $\{e_k\}$, with the properties

$$f_i \cdot f_j = e_i \cdot e_j, \qquad f_i \cdot e_j = 0, \tag{11.86}$$

which hold for all $i, j = 1, \dots, n$. From these vectors we construct the bivector

$$J = \sum_{i=1}^{n} e_i \wedge f^i = e_i \wedge f^i, \tag{11.87}$$

where the $\{f^k\}$ are the reciprocal vectors to the $\{f_k\}$ frame. For this and the following section we assume that repeated indices are summed from $1, \dots, n$. The bivector J is independent of the initial choice of frame $\{e_i\}$. To see this, introduce a second pair of frames $\{e_i'\}$ and $\{f_i'\}$ related in the same manner as the $\{e_k\}$, $\{f_k\}$ pair. For these we find that

$$J' = e_i' \wedge f'^i = e_i' \cdot e^j \, e_j \wedge f'^i = f_i' \cdot f^j e_j \wedge f'^i = e_j \wedge f^j = J. \tag{11.88}$$

In particular, if the $\{e_k\}$ frame is chosen to be orthonormal, we find that

$$J = e_1 f_1 + e_2 f_2 + \cdots + e_n f_n = J_1 + J_2 + \cdots + J_n. \tag{11.89}$$

Each bivector blade J_i then provides the complex structure for the ith plane.

To understand the properties of the bivector J we first form the products

$$e_i \cdot J = e_i \cdot e_j \, f^j = f_i \cdot f_j \, f^j = f_i \tag{11.90}$$

and

$$f_i \cdot J = -e_j \, f_i \cdot f^j = -e_i. \tag{11.91}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

It follows that

$$(e_i \cdot J) \cdot J = f_i \cdot J = -e_i,$$
$$(f_i \cdot J) \cdot J = -e_i \cdot J = -f_i, \tag{11.92}$$

and hence that

$$(a \cdot J) \cdot J = -a, \tag{11.93}$$

for any vector a. We can now see how J will take over the role of the unit imaginary. For example, the analogue of phase rotations is generated by the bivector J, which describes a series of coupled rotations in each of the J_i planes. A Taylor expansion then yields

$$\begin{aligned} \mathrm{e}^{-J\phi/2} a \mathrm{e}^{J\phi/2} &= a + \phi\, a \cdot J + \frac{\phi^2}{2!}(a \cdot J) \cdot J \cdots \\ &= \cos(\phi) a + \sin(\phi) a \cdot J. \end{aligned} \tag{11.94}$$

The map $a \mapsto a \cdot J$ is therefore a $\pi/2$ rotation. Setting $\phi = \pi$ we also see that

$$a \mathrm{e}^{J\pi/2} = -\mathrm{e}^{J\pi/2} a, \tag{11.95}$$

so $\exp(J\pi/2)$ anticommutes with every vector in the algebra. The only multivector with this property is the pseudoscalar, so we have

$$\mathrm{e}^{J\pi/2} = I_{2n}, \tag{11.96}$$

where I_{2n} is the pseudoscalar of the $2n$-dimensional algebra.

Next we need a means of distinguishing the real and imaginary parts of a vector. As with the two-dimensional case, this requires picking out a preferred set of directions to represent the real axes. As a matter of convention we choose to identify these with the original $\{e_k\}$ vectors. A real vector a in the $2n$-dimensional algebra can now be mapped to a set of complex coefficients $\{a_i\}$ as follows:

$$a_i = a \cdot e_i + j\, a \cdot f_i. \tag{11.97}$$

The complex inner product therefore becomes

$$\begin{aligned} \langle a|b\rangle = a^i b_i^* &= (a \cdot e^i + j\, a \cdot f^i)(b \cdot e_i - j\, b \cdot f_i) \\ &= a \cdot e^i\, b \cdot e_i + a \cdot f^i\, b \cdot f_i + j(a \cdot f^i\, b \cdot e_i - a \cdot e^i\, b \cdot f_i) \\ &= a \cdot b + j(a \wedge b) \cdot J. \end{aligned} \tag{11.98}$$

This shows that the complex inner product combines two geometrically distinct terms. The real part is the usual vector inner product, and it follows immediately that $a^i a_i^* = a^2$. The imaginary part is an antisymmetric product formed by projecting the bivector $a \wedge b$ onto J. Antisymmetric products such as these play an important role in symplectic geometry and Hamiltonian mechanics.

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

11.4.2 Unitary transformations

We are free to consider any linear function defined over our $2n$-dimensional vector space. However, only a subset of these can be represented by complex matrices — those that observe the complex structure. These transformations are linear over the complex field, so must satisfy

$$\mathsf{f}(\alpha a + \beta a \cdot J) = \alpha \mathsf{f}(a) + \beta \mathsf{f}(a) \cdot J. \tag{11.99}$$

It follows that complex linear transformations satisfy

$$\mathsf{f}(a \cdot J) = \mathsf{f}(a) \cdot J \tag{11.100}$$

for any vector a in the $2n$-dimensional vector space.

The study of complex linear functions now reduces to the study of functions satisfying the condition (11.100). For example, the matrix operation of Hermitian conjugation has

$$\langle a | \mathsf{f}(b) \rangle = \langle \mathsf{f}^\dagger(a) | b \rangle. \tag{11.101}$$

By considering the various terms in this identity we see immediately that the Hermitian adjoint is the same as the familiar adjoint function $\bar{\mathsf{f}}$. That is, $\mathsf{f}^\dagger = \bar{\mathsf{f}}$. This explains why it is Hermitian conjugation that is so important in analysing complex matrices. Similarly, suppose that a is a complex eigenvector of the complex function f. This implies that

$$\mathsf{f}(a) = \alpha a + \beta a \cdot J. \tag{11.102}$$

Clearly, if a satisfies this equation, then $a \cdot J$ satisfies

$$\mathsf{f}(a \cdot J) = \alpha a \cdot J - \beta a. \tag{11.103}$$

It follows that $a \wedge (a \cdot J)$ is an eigenbivector, with

$$\mathsf{f}\big(a \wedge (a \cdot J)\big) = (\alpha^2 + \beta^2) a \wedge (a \cdot J). \tag{11.104}$$

Next we need to establish the invariance group of the Hermitian inner product. This group must leave invariant both terms in equation (11.98). This includes the inner product $a \cdot b$, which tells us that the invariance group is built from reflections and rotations. The fact that the linear transformations preserve the complex structure then ensures that the antisymmetric term is also invariant. To see this, suppose that f satisfies $\bar{\mathsf{f}} = \mathsf{f}^{-1}$, together with equation (11.100). It follows that

$$\big(\mathsf{f}(a) \wedge \mathsf{f}(b)\big) \cdot J = \mathsf{f}(a) \cdot \big(\mathsf{f}(b) \cdot J\big) = \mathsf{f}(a) \cdot \mathsf{f}(b \cdot J) = (a \wedge b) \cdot J. \tag{11.105}$$

This result can be summarised concisely as

$$\mathsf{f}(J) = J. \tag{11.106}$$

Unitary groups are therefore constructed from reflections and rotations which

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

leave J invariant. For a reflection to satisfy this constraint would require that the vector generator m satisfies

$$mJm^{-1} = J. \tag{11.107}$$

But this implies that $m\cdot J = 0$, and hence that $(m\cdot J)\cdot J = -m = 0$. There are therefore no vector generators of reflections, and hence all unitary transformations are generated by elements of the spin group. So far we have not specified the underlying signature, so our description applies equally to the unitary groups $\mathrm{U}(n)$ and $\mathrm{U}(p,q)$. These groups can be represented in terms of even multivectors in $\mathcal{G}(2n,0)$ and $\mathcal{G}(2p,2q)$ respectively.

To simplify matters, we now restrict to the Euclidean case, so we seek a rotor description of the unitary group $\mathrm{U}(n)$. The spin group and rotor group in $\mathcal{G}(2n,0)$ are the same, so the unitary group has a double-cover representation in terms of rotors satisfying

$$RJ\tilde{R} = J. \tag{11.108}$$

Writing $R = \exp(-B/2)$, we see that the bivector generators of the unitary group must satisfy

$$B\times J = 0. \tag{11.109}$$

This defines a bivector representation of the Lie algebra $\mathrm{u}(n)$ of the unitary group $\mathrm{U}(n)$. We can construct bivectors satisfying equation (11.109) by first using the Jacobi identity to prove that

$$\begin{aligned}\big((a\cdot J)\wedge(b\cdot J)\big)\times J &= -(a\cdot J)\wedge b + (b\cdot J)\wedge a\\ &= -(a\wedge b)\times J.\end{aligned} \tag{11.110}$$

It follows that

$$\big(a\wedge b + (a\cdot J)\wedge(b\cdot J)\big)\times J = 0. \tag{11.111}$$

Any bivector of the form on the left-hand side will therefore commute with J. Suppose now that the $\{e_i\}$ and $\{f_i\}$ are orthonormal vectors. We can work through all combinations of these to arrive at the bivector algebra in table 11.1. Establishing the closure of this algebra under the commutator product is straightforward. The bivector algebra contains J, which commutes with all other elements and is responsible for a global phase term. Removing this term defines the Lie algebra $\mathrm{su}(n)$ of the special unitary group $\mathrm{SU}(n)$. The analysis can be repeated with a different signature base space to construct a bivector representation of the Lie algebra $\mathrm{u}(p,q)$.

11.5 The general linear group

We have seen how to represent both rotation groups and unitary groups in terms of spin groups. We will now see how all matrix groups can be represented by spin

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

E_{ij}	$=$	$e_i e_j + f_i f_j$	$(i < j = 1, \ldots, n)$
F_{ij}	$=$	$e_i f_j - f_i e_j$	$(i < j = 1, \ldots, n)$
J_i	$=$	$e_i f_i$	$(i = 1, \ldots, n)$

Table 11.1 *The Lie algebra* $\mathrm{u}(n)$. The bivectors all belong to the geometric algebra $\mathcal{G}(2n, 0)$, and the vectors $\{e_i\}$ and $\{f_i\}$ form an orthonormal basis for this algebra. The complex structure is generated by the bivector $J = J_1 + \cdots + J_n$.

groups, and hence that all possible Lie algebras can be represented as bivector algebras. This is a significant motivation for the treatment adopted in this chapter. Formulating general linear functions as rotors is achieved by working in a balanced algebra, generated by equal numbers of vectors with positive and negative square. Some of the algebraic considerations for these types of algebra were encountered in the discussions of spacetime and conformal geometry.

11.5.1 The balanced algebra $\mathcal{G}(n, n)$

Suppose that the vectors $\{e_i\}$ span a non-degenerate space of unspecified signature. We introduce a second frame $\{f_k\}$, orthogonal to the first and with *opposite* signature, with the properties

$$f_i \cdot f_j = -e_i \cdot e_j, \qquad e_i \cdot f_j = 0. \tag{11.112}$$

The vectors $\{e_i, f_i\}$ therefore generate the algebra $\mathcal{G}(n, n)$, regardless of the signature of the original $\{e_i\}$ space. We next introduce the balanced analogue of the complex bivector J by defining

$$K = e_i \wedge f^i. \tag{11.113}$$

This has the properties that

$$e_i \cdot K = e_i \cdot e_j \, f^j = -f_i \cdot f_j \, f^j = -f_i \tag{11.114}$$

and

$$f_i \cdot K = -f_i \cdot f^j \, e_j = -e_i. \tag{11.115}$$

It follows that

$$(a \cdot K) \cdot K = K \cdot (K \cdot a) = a \quad \forall a \in \mathcal{V}. \tag{11.116}$$

There is therefore a crucial sign difference compared with the complex bivector J. This means that K does not generate a complex structure, but instead generates a *null* structure. To see this, we first form

$$(a \cdot K)^2 = -\big((a \cdot K) \cdot K\big) \cdot a = -a^2, \tag{11.117}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

so the vector $a{\cdot}K$ has opposite signature to a. Given a general vector $a \in \mathcal{G}(n,n)$ we can define two separate null vectors by writing

$$a = \tfrac{1}{2}(a + a{\cdot}K) + \tfrac{1}{2}(a - a{\cdot}K). \tag{11.118}$$

In this way the vector space $\mathcal{V}$ of $\mathcal{G}(n,n)$ splits into two null spaces, $\mathcal{V}_+$ and $\mathcal{V}_-$. Vectors in $\mathcal{V}_+$ satisfy

$$a_+{\cdot}K = a_+ \quad \forall a_+ \in \mathcal{V}_+, \tag{11.119}$$

with a similar expression (with a minus sign) holding for $\mathcal{V}_-$. Both of the spaces $\mathcal{V}_+$ and $\mathcal{V}_-$ are entirely null, and they are dual spaces to one another. Working entirely with vectors in $\mathcal{V}_+$ is a further way of formulating a Grassmann algebra within geometric algebra.

11.5.2 Linear transformations

We will shortly demonstrate that every linear function acting on an n-dimensional vector space, $a \mapsto \mathsf{f}(a)$, can be represented in $\mathcal{V}_+$ by a transformation of the form

$$a_+ \mapsto M a_+ M^{-1}. \tag{11.120}$$

Here M belongs to a subgroup of the spin group for $\mathcal{G}(n,n)$, and a_+ is the image of a in $\mathcal{V}_+$ defined by

$$a_+ = a + a{\cdot}K. \tag{11.121}$$

In this sense we form a double-cover representation of the general linear group. The relevant subgroup consists of transformations that map the subspaces $\mathcal{V}_+$ and $\mathcal{V}_-$ entirely within themselves. For this to hold we require that

$$(M a_+ M^{-1}){\cdot}K = M a_+ M^{-1}, \tag{11.122}$$

so we must have

$$\begin{aligned} a_+ &= M^{-1}\,(M a_+ M^{-1}){\cdot}K\,M \\ &= M^{-1}\tfrac{1}{2}(M a_+ M^{-1}K - K M a_+ M^{-1})M \\ &= a_+{\cdot}(M^{-1}KM). \end{aligned} \tag{11.123}$$

It follows that we require $M^{-1}KM = K$, or

$$MK = KM. \tag{11.124}$$

As with the unitary case, M must belong to the spin group. The bivector generators of this group must commute with K. The Jacobi identity ensures that the commutator product of two bivectors that commute with K results in

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

E_{ij}	$=$	$e_i e_j - f_i f_j$	$(i < j = 1, \dots, n)$
F_{ij}	$=$	$e_i f_j - f_i e_j$	$(i < j = 1, \dots, n)$
K_i	$=$	$e_i f_i$	$(i = 1, \dots, n)$

Table 11.2 *The Lie algebra* gl(n). The bivectors all belong to the geometric algebra $\mathcal{G}(n,n)$. The $\{e_i\}$ vectors are orthonormal with positive signature, and the $\{f_i\}$ are orthonormal with negative signature. The algebra contains the bivector $K = K_1 + \cdots + K_n$, which generates the Abelian subgroup of global dilations. Factoring out this bivector produces the algebra sl(n).

a third that also commutes with K. We proceed as with the unitary group and construct

$$\big((a\cdot K)\wedge(b\cdot K)\big)\times K = a\wedge(b\cdot K) + (a\cdot K)\wedge b = (a\wedge b)\times K, \tag{11.125}$$

so that

$$\big(a\wedge b - (a\cdot K)\wedge(b\cdot K)\big)\times K = 0. \tag{11.126}$$

We can again run through all combinations of the basis bivectors to obtain the basis for the Lie algebra of the general linear group listed in table 11.2. The difference in structure between the Lie algebras of the linear group and the unitary group is due solely to the different signatures of their underlying spaces.

The remaining step is to give an explicit construction of a representation of a linear transformation as an element of the spin group. The key to this is the singular value decomposition of section 4.4.8. This decomposition shows that any $n \times n$ matrix (with non-zero determinant) can be decomposed into a positive-definite diagonal matrix sandwiched between two orthogonal matrices. To find a suitable encoding in terms of rotors, all we have to do is find representations of orthogonal transformations and positive dilations.

Rotations are clearly present as they are generated by the E_{ij} bivectors in the Lie algebra of table 11.2. These bivectors jointly rotate the $\{e_i\}$ and $\{f_i\}$ vectors by the same amount. But the orthogonal group also includes reflections, so we need to represent these as well. Suppose the reflection in $\mathcal{G}(p,q)$ is generated by the unit vector n, $n^2 = 1$. We define

$$\bar{n} = n\cdot K, \qquad \bar{n}^2 = -1, \tag{11.127}$$

and consider the multivector $n\bar{n}$. This satisfies

$$n\bar{n}K = 2n\,\bar{n}\cdot K + nK\bar{n} = 2(n^2 + \bar{n}^2) + Kn\bar{n} = Kn\bar{n}, \tag{11.128}$$

so the bivector does commute with K. But since

$$n\bar{n}(n\bar{n})^{\sim} = -1 \tag{11.129}$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

this bivector is not a rotor. It belongs to the spin group, but not the rotor group. The action of $n\bar{n}$ on vectors $a_+ \in \mathcal{V}_+$ results in the vector

$$-n\bar{n}a_+\bar{n}n = -n\bar{n}a\bar{n}n - (n\bar{n}a\bar{n}n)\cdot K = -nan - (nan)\cdot K, \tag{11.130}$$

where a is the original vector, in the same space as n. Since $\bar{n}$ is in the orthogonal space generated by the $\{f_i\}$ vectors, $\bar{n}$ anticommutes with a. Equation (11.130) is the required result for a reflection. The need to include reflections forces us to work with elements of the full spin group in $\mathcal{G}(n,n)$.

The final step is to see how dilations are formulated with rotors. Suppose that we now require a positive dilation in the n direction. We again form the bivector $n\bar{n}$, which is constructed from the F_{ij} and K_i Lie algebra generators. With $n_+ = n + \bar{n}$ the equivalent of the vector n in $\mathcal{V}_+$, we find that

$$\begin{aligned} \mathrm{e}^{-\lambda n\bar{n}/2} n_+ \mathrm{e}^{\lambda n\bar{n}/2} &= \big(\cosh(\lambda) - n\bar{n}\sinh(\lambda)\big)(n+\bar{n}) \\ &= \mathrm{e}^{\lambda} n_+, \end{aligned} \tag{11.131}$$

which is a pure dilation. Furthermore, any vector perpendicular to n has an image in $\mathcal{V}_+$ that commutes with $n\bar{n}$ and so is unaffected by the action of the rotor. These are precisely the required properties of the positive dilation, which completes the construction.

We now have an alternative means of representing every matrix group within geometric algebra. Since *all* Lie algebras can be represented by matrices, we have proved that all Lie algebras can be realised as bivector algebras. The accompanying Lie group elements can then all be written as even products of unit vectors. This is potentially a very powerful idea. One immediate construct one can form this way is the *tensor product* of two linear functions. All one requires for this is a separate copy of the algebra $\mathcal{G}(n,n)$ for each linear operator. As with the multiparticle spacetime algebra construction of chapter 9, the generators of each space are orthogonal, so anticommute. It follows that even elements from either space commute. So rotors from either space can be multiplied commutatively, forming a spinor representation of the tensor product. The combined rotor generates the correct tensor product action on vectors in the combined space. The tensor product can therefore be constructed from the geometric product.

11.6 Notes

The multivector derivative and the use of the vector derivative in analysing linear functions are described in detail in the book *Clifford Algebra to Geometric Calculus* by Hestenes & Sobczyk (1984). This book also contains an elegant proof of the Cayley–Hamilton theorem, and details of the geometric algebra approach to Lie group theory. Some further material is contained in the 'Lectures in geometric algebra' by Doran *et al.* (1996a).

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

The basis of Grassmann calculus is described in *The Method of Second Quantisation* by Berezin (1966). A summary of the main results from this is contained in the appendices to the paper 'Particle spin dynamics as the Grassmann variant of classical mechanics' by Berezin and Marinov (1977). More recently, Grassmann calculus has been extended to the field of superanalysis, as described in the books by Berezin (1987) and de Witt (1984). Similar themes also reappear in the subject of non-commutative geometry, as discussed by Connes & Lott (1990) and Coquereaux, Jadczyk & Kastler (1991). The geometric algebra treatment of Grassmann calculus was introduced in the papers 'Grassmann calculus, pseudoclassical mechanics and geometric algebra' by Lasenby, Doran & Gull (1993c) and 'Grassmann mechanics, multivector derivatives and geometric algebra' by Doran, Lasenby & Gull (1993b). Some additional material is contained in the thesis by Doran (1994). These works also show how the super-Lie bracket, and super-Lie algebras, can be formulated within geometric algebra.

The subject of Lie groups is covered in an enormous range of textbooks. The series entitled *Group Theory in Physics* by Cornwell (1984a,1984b,1989) are particularly recommended, as are the books by Georgi (1982) and Gilmore (1974). The subject of pin and spin groups has also been discussed widely. Thorough treatments can be found in the books *An Introduction to Spinors and Geometry* by Benn & Tucker (1988) and *Clifford Algebras and Spinors* by Lounesto (1997). The construction of the general linear group in terms of rotors was first described in the paper 'Lie groups as spin groups' by Doran *et al.* (1993). The thesis by Doran (1994) contains explicit constructions of a number of further Lie algebras, including symplectic and quaternionic algebras.

11.7 Exercises

11.1 The function f maps vectors to vectors in the spacetime algebra according to

$$\mathsf{f}(a) = a + \alpha a\cdot\gamma_+\,\gamma_+,$$

where γ_+ is the null vector $\gamma_0 + \gamma_3$. Find the characteristic equation satisfied by f. What are the roots of the characteristic polynomial and how many independent eigenvectors are there? Verify that f satisfies its own characteristic equation.

11.2 Suppose that the vectors γ_0, γ_1 form an orthogonal basis for a space of signature $(1,1)$. Show that the linear function f_1,

$$\mathsf{f}_1(a) = -12a\cdot\gamma_0\,\gamma_0 + 2a\cdot\gamma_0\,\gamma_1 + 2a\cdot\gamma_1\,\gamma_0 + a\cdot\gamma_1\,\gamma_1,$$

has no symmetric square root. Similarly, show that the function f_2,

$$\mathsf{f}_2(a) = 8a\cdot\gamma_0\,\gamma_0 + a\cdot\gamma_0\,\gamma_1 + a\cdot\gamma_1\,\gamma_0 - a\cdot\gamma_1\,\gamma_1,$$

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

has two symmetric square roots, and find them both.

11.3 The function $\phi(\lambda)$ is defined by

$$\phi(\lambda) = \det\big(\exp(\lambda \mathsf{f})\big)$$

where f is a linear function. The exponential function is defined by the power series

$$\exp(\lambda\mathsf{f})(a) = \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \mathsf{f}^r(a)$$

where $\mathsf{f}^r(a)$ denotes the r-fold application of f and $\mathsf{f}^0(a) = a$. Prove that $\phi(\lambda)$ satisfies

$$\frac{d\phi}{d\lambda} = \partial_a \cdot \mathsf{f}(a)\, \phi(\lambda),$$

and hence prove that

$$\det\big(\exp(\mathsf{f})\big) = \exp\big(\partial_a \cdot \mathsf{f}(a)\big).$$

11.4 Prove the following results for the functional derivative:

$$\partial_{\mathsf{f}(a)} \partial_b \cdot \mathsf{f}^r(b) = r\mathsf{f}^{r-1}(a), \quad r \geq 1,$$
$$\partial_{\mathsf{f}(a)} \langle \bar{\mathsf{f}}^{-1}(A_r) B_r \rangle = -\langle \bar{\mathsf{f}}^{-1}(a) \cdot B_r \, \bar{\mathsf{f}}^{-1}(A_r) \rangle_1.$$

11.5 Given a non-singular function f in Euclidean space, the function ε is defined by

$$\varepsilon = \tfrac{1}{2} \ln\big(\bar{\mathsf{f}}\,\mathsf{f}\big). \tag{E11.1}$$

The logarithm can be defined either by a power series, or by diagonalising $\bar{\mathsf{f}}\,\mathsf{f}$ and taking the logarithm of the eigenvalues. Prove that

$$\partial_{\mathsf{f}(a)} \partial_b \cdot \varepsilon(b) = \bar{\mathsf{f}}^{-1}(a),$$
$$\partial_{\mathsf{f}(a)} \partial_b \cdot \varepsilon^2(b) = \bar{\mathsf{f}}^{-1} \varepsilon(a).$$

11.6 Prove that left and right-sided Grassmann derivatives commute.

11.7 Suppose that x, y and e are unit vectors in $\mathcal{G}(4,0)$, with the pseudoscalar denoted by I. Prove that the product $\phi(x,y)$, where

$$\phi(x,y) = \langle xey(1+I) \rangle_1,$$

satisfies all the axioms of a Lie group product, with e the identity element. Which group does this product define?

11.8 The multivector R is defined by

$$R = -1 - (\gamma_0 + \gamma_1)\gamma_2,$$

where $\{\gamma_0, \gamma_1, \gamma_2\}$ are an orthonormal basis for $\mathcal{G}(1,2)$. Prove that R is a rotor, and that it is impossible to find a bivector B such that $R = \exp(-B/2)$.

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013

11.9 The vectors $\{e_i, f_i\}, i = 1, \ldots, n$ form an orthonormal basis for $\mathcal{G}(2n, 0)$. The Lie algebra u(n) is defined by the following bivectors:

$$\begin{aligned} E_{ij} &= e_i e_j + f_i f_j \qquad && (i < j = 1, \ldots, n), \\ F_{ij} &= e_i f_j - f_i e_j \qquad && (i < j = 1, \ldots, n), \\ J_i &= e_i f_i. \end{aligned}$$

Prove that this algebra is closed under the commutator product. Hence find the structure constants of the unitary group.

11.10 Prove that the Lie algebras su(4) and so(6) are isomorphic. Repeat the analysis for the case of su$(2, 2)$ and so$(2, 4)$. This latter isomorphism is important in the theory of twistors.

Downloaded from https://www.cambridge.org/core. ETH-Bibliothek, on 18 Jun 2020 at 08:43:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/CBO9780511807497.013