To express the Dirac equation in Hamiltonian form we simply multiply from the left by γ_0. The resulting equation, with the dimensional constants temporarily put back in, is

$$j\hbar\partial_t\psi = c\hat{\boldsymbol{p}}\psi + eV\psi - ce\boldsymbol{A}\psi + mc^2\bar{\psi}, \tag{8.106}$$

where

$$\begin{aligned} \hat{\boldsymbol{p}}\psi &= -j\hbar\boldsymbol{\nabla}\psi, \\ \bar{\psi} &= \gamma_0\psi\gamma_0, \\ \gamma_0 A &= V - c\boldsymbol{A}. \end{aligned} \tag{8.107}$$

Choosing a Hamiltonian is a non-covariant operation, since it picks out a preferred timelike direction. The Hamiltonian relative to the γ_0 direction is the operator on the right-hand side of equation (8.106).

As an application of the Hamiltonian formulation, consider the non-relativistic reduction of the Dirac equation. This can be achieved formally via the *Foldy–Wouthuysen* transformation. For details we refer the reader to Itzykson & Zuber (1980). While the theoretical motivation for this transformation is clear, it can be hard to compute in all but the simplest cases. A simpler approach, dating back to Feynman, is to separate out the fast-oscillating component of the waves and then split into separate equations for the Pauli-even and Pauli-odd components of ψ. We write (with $\hbar = 1$ and the factors of c kept in)

$$\psi = (\phi + \eta)\mathrm{e}^{-I\boldsymbol{\sigma}_3 mc^2 t}, \tag{8.108}$$

where $\bar{\phi} = \phi$ (Pauli-even) and $\bar{\eta} = -\eta$ (Pauli-odd). The Dirac equation (8.106) now splits into the two equations

$$\begin{aligned} \mathcal{E}\phi - c\mathcal{O}\eta &= 0, \\ (\mathcal{E} + 2mc^2)\eta - c\mathcal{O}\phi &= 0, \end{aligned} \tag{8.109}$$

where

$$\begin{aligned} \mathcal{E}\phi &= (j\partial_t - eV)\phi, \\ \mathcal{O}\phi &= (\hat{\boldsymbol{p}} - e\boldsymbol{A})\phi. \end{aligned} \tag{8.110}$$

The formal solution to the second of equations (8.109) is

$$\eta = \frac{1}{2mc}\left(1 + \frac{\mathcal{E}}{2mc^2}\right)^{-1}\mathcal{O}\phi, \tag{8.111}$$

where the inverse on the right-hand side denotes a power series. Provided the expectation value of $\mathcal{E}$ is smaller than $2mc^2$ (which it is in the non-relativistic limit) the series should converge. The remaining equation for ϕ is

$$\mathcal{E}\phi - \frac{\mathcal{O}}{2m}\left(1 - \frac{\mathcal{E}}{2mc^2} + \cdots\right)\mathcal{O}\phi = 0, \tag{8.112}$$

which can be expanded out to the desired order of magnitude. There is little point in going beyond the first relativistic correction, so we approximate equation (8.112) by

$$\mathcal{E}\phi + \frac{\mathcal{O}\mathcal{E}\mathcal{O}}{4m^2c^2}\phi = \frac{\mathcal{O}^2}{2m}\phi. \tag{8.113}$$

We seek an equation of the form $\mathcal{E}\phi = \mathcal{H}\phi$, where $\mathcal{H}$ is the non-relativistic Hamiltonian. We therefore need to replace the $\mathcal{O}\mathcal{E}\mathcal{O}$ term in equation (8.113) by a term that does not involve $\mathcal{E}$. To achieve this we write

$$2\mathcal{O}\mathcal{E}\mathcal{O} = [\mathcal{O},[\mathcal{E},\mathcal{O}]] + \mathcal{E}\mathcal{O}^2 + \mathcal{O}^2\mathcal{E} \tag{8.114}$$

so that equation (8.113) becomes

$$\mathcal{E}\phi = \frac{\mathcal{O}^2}{2m}\phi - \frac{\mathcal{E}\mathcal{O}^2 + \mathcal{O}^2\mathcal{E}}{8m^2c^2}\phi - \frac{1}{8m^2c^2}[\mathcal{O},[\mathcal{E},\mathcal{O}]]\phi. \tag{8.115}$$

We can now make the approximation

$$\mathcal{E}\phi \approx \frac{\mathcal{O}^2}{2m}\phi, \tag{8.116}$$

so that equation (8.113) can be approximated by

$$\mathcal{E}\phi = \frac{\mathcal{O}^2}{2m}\phi - \frac{1}{8m^2c^2}[\mathcal{O},[\mathcal{E},\mathcal{O}]]\phi - \frac{\mathcal{O}^4}{8m^3c^2}\phi, \tag{8.117}$$

which is valid to order c^{-2}.

To evaluate the commutators we first need

$$[\mathcal{E},\mathcal{O}] = -je(\partial_t \boldsymbol{A} + \boldsymbol{\nabla}V) = je\boldsymbol{E}. \tag{8.118}$$

There are no time derivatives left in this commutator, so we do achieve a sensible non-relativistic Hamiltonian. The full commutator required in equation (8.117) is

$$\begin{aligned}[\mathcal{O},[\mathcal{E},\mathcal{O}]] &= [-j\boldsymbol{\nabla} - e\boldsymbol{A}, je\boldsymbol{E}] \\ &= (e\boldsymbol{\nabla}\boldsymbol{E}) - 2e\boldsymbol{E}\wedge\boldsymbol{\nabla} - 2je^2\boldsymbol{A}\wedge\boldsymbol{E}.\end{aligned} \tag{8.119}$$

The various operators (8.110) and (8.119) can now be substituted into equation (8.117) to yield the Pauli equation

$$\begin{aligned}\frac{\partial\phi}{\partial t}I\boldsymbol{\sigma}_3 = \frac{1}{2m}(\hat{\boldsymbol{p}} - e\boldsymbol{A})^2\phi + eV\phi - \frac{\hat{\boldsymbol{p}}^4}{8m^3c^2}\phi \\ - \frac{1}{8m^2c^2}\big(e(\boldsymbol{\nabla}\boldsymbol{E} - 2\boldsymbol{E}\wedge\boldsymbol{\nabla})\phi - 2e^2\boldsymbol{A}\wedge\boldsymbol{E}\phi I\boldsymbol{\sigma}_3\big),\end{aligned} \tag{8.120}$$

which is written entirely in the geometric algebra of three-dimensional space. In the standard approach, the geometric product in the $\boldsymbol{\nabla}\boldsymbol{E}$ term of equation (8.120) is split into a 'spin-orbit' term $\boldsymbol{\nabla}\wedge\boldsymbol{E}$ and the 'Darwin' term $\boldsymbol{\nabla}\cdot\boldsymbol{E}$.

The spacetime algebra approach reveals that these terms arise from a single source.

A similar approximation scheme can be adopted for the observables of the Dirac theory. For example the current, $\psi\gamma_0\tilde{\psi}$, has a three-vector part:

$$\boldsymbol{J} = (\psi\gamma_0\tilde{\psi})\wedge\gamma_0 = \phi\eta^\dagger + \eta\phi^\dagger. \tag{8.121}$$

This is approximated to leading order by

$$\boldsymbol{J} \approx -\frac{1}{m}(\langle\boldsymbol{\nabla}\phi I\boldsymbol{\sigma}_3\phi^\dagger\rangle_1 - \boldsymbol{A}\phi\phi^\dagger), \tag{8.122}$$

where the $\langle\,\rangle_1$ projects onto the grade-1 components of the Pauli algebra. Not all applications of the Pauli theory correctly identify (8.122) as the conserved current in the Pauli theory — an inconsistency first pointed out by Hestenes & Gurtler (1971).

8.4 Central potentials

Suppose now that we restrict our discussion to problems described by a central potential $V = V(r)$, $\boldsymbol{A} = 0$, where $r = |\boldsymbol{x}|$. The full Hamiltonian, denoted $\mathcal{H}$, reduces to

$$j\hbar\partial_t\psi = \mathcal{H}\psi = -j\boldsymbol{\nabla}\psi + eV(r)\psi + m\bar{\psi}. \tag{8.123}$$

Quantum states are classified in terms of eigenstates of operators that commute with the Hamiltonian $\mathcal{H}$, because the accompanying quantum numbers are conserved in time. Of particular importance are the angular-momentum operators $\hat{L}_i$, defined by

$$\hat{L}_i = -i\epsilon_{ijk}x_j\partial_k. \tag{8.124}$$

These are the components of the bivector operator $i\boldsymbol{x}\wedge\boldsymbol{\nabla}$. We therefore define the operators

$$L_B = jB\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}), \tag{8.125}$$

where B is a relative bivector. Throughout this section interior and exterior products refer to the (Pauli) algebra of space. Writing $B = I\boldsymbol{\sigma}_i$ recovers the component form. The L_B operators satisfy the commutation relations

$$[L_{B_1}, L_{B_2}] = -jL_{B_1\times B_2}, \tag{8.126}$$

where $B_1\times B_2$ denotes the commutator product. The angular-momentum commutation relations directly encode the bivector commutation relations, which are those of the Lie algebra of the rotation group (see chapter 11). One naturally expects this group to arise as it represents a symmetry of the potential.

If we now form the commutator of L_B with the Hamiltonian $\mathcal{H}$ we obtain a

result that is, initially, disconcerting. The scalar operator L_B commutes with the bar operator $\psi \mapsto \bar{\psi}$, but for the momentum term we find that

$$[B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}),\boldsymbol{\nabla}] = -\dot{\boldsymbol{\nabla}}B\cdot(\dot{\boldsymbol{x}}\wedge\boldsymbol{\nabla}) = B\times\boldsymbol{\nabla}. \tag{8.127}$$

The commutator does not vanish, so orbital angular momentum does not yield a conserved quantum number in relativistic physics. But, since $B\times\boldsymbol{\nabla} = \frac{1}{2}(B\boldsymbol{\nabla} - \boldsymbol{\nabla}B)$, we can write equation (8.127) as

$$[B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}) - \tfrac{1}{2}B, \mathcal{H}] = 0. \tag{8.128}$$

We therefore recover a conserved angular momentum operator by defining

$$J_B = L_B - \tfrac{1}{2}jB. \tag{8.129}$$

In conventional notation this is

$$\hat{J}_i = \hat{L}_i + \tfrac{1}{2}\hat{\Sigma}_i, \tag{8.130}$$

where $\hat{\Sigma}_i = (i/2)\epsilon_{ijk}\hat{\gamma}_j\hat{\gamma}_k$. The extra term of $B/2$ accounts for the spin-1/2 nature of Dirac particles. If we look for eigenstates of the J_3 operator, we see that the spin contribution to this is

$$-\tfrac{1}{2}jI\boldsymbol{\sigma}_3\psi = \tfrac{1}{2}\boldsymbol{\sigma}_3\psi\boldsymbol{\sigma}_3. \tag{8.131}$$

In the non-relativistic Pauli theory the eigenstates of this operator are simply 1 and $-I\boldsymbol{\sigma}_2$, with eigenvalues $\pm 1/2$. In the relativistic theory the separate spin and orbital operators are not conserved, and it is only the combined J_B operators that commute with the Hamiltonian.

The geometric algebra derivation employed here highlights some interesting features. Stripping away all of the extraneous terms, the result rests solely on the commutation properties of the $B\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla})$ and $\boldsymbol{\nabla}$ operators. The factor of $1/2$ would therefore be present in any dimension, and so has no special relation to the three-dimensional rotation group. Furthermore, in writing $J_B = L_B - \frac{1}{2}jB$ we are forming an explicit sum of a scalar and a bivector. The standard notation of equation (8.130) encourages us to view these as the sum of two vector operators!

8.4.1 Spherical monogenics

The spherical monogenics play a key role in the solution of the Dirac equation for problems with radial symmetry. These are Pauli spinors (even elements of the Pauli algebra) that satisfy the eigenvalue equation

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi = l\psi. \tag{8.132}$$

These functions arise naturally as solutions of the three-dimensional generalisation of the Cauchy–Riemann equations

$$\boldsymbol{\nabla}\Psi = 0. \tag{8.133}$$

Solutions of this equation are known in the Clifford analysis literature as monogenics. Looking for solutions which separate into $\Psi = r^l\psi(\theta,\phi)$ yields equation (8.132), where (r,θ,ϕ) is a standard set of polar coordinates. The solutions of equation (8.132) are called spherical monogenics, or *spin-weighted spherical harmonics* (with weight 1/2).

To analyse the properties of equation (8.132) we first note that

$$[J_B, \boldsymbol{x}\wedge\boldsymbol{\nabla}] = 0, \tag{8.134}$$

which is proved in the same manner as equation (8.128). It follows that ψ can simultaneously be an eigenstate of the $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ operator and one of the J_B operators. To simplify the notation we now define

$$J_k\psi = J_{I\sigma_k}\psi = \big((I\sigma_k)\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla}) - \tfrac{1}{2}I\sigma_k\big)\psi I\boldsymbol{\sigma}_3. \tag{8.135}$$

We choose ψ to be an eigenstate of J_3. We label this state as $\psi(l,\mu)$, so

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi(l,\mu) = l\psi(l,\mu), \qquad J_3\psi(l,\mu) = \mu\psi(l,\mu). \tag{8.136}$$

The J_i operators satisfy

$$\begin{aligned} J_iJ_i\psi(l,\mu) &= 3/4\psi - 2\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi + \boldsymbol{x}\wedge\boldsymbol{\nabla}(\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi) \\ &= (l+1/2)(l+3/2)\psi(l,\mu), \end{aligned} \tag{8.137}$$

so the $\psi(l,\mu)$ are also eigenstates of J_iJ_i.

We next introduce the ladder operators J_+ and J_-, defined by

$$\begin{aligned} J_+ &= J_1 + jJ_2, \\ J_- &= J_1 - jJ_2. \end{aligned} \tag{8.138}$$

It is a simple matter to prove the following results:

$$\begin{aligned} [J_+, J_-] &= 2J_3, \qquad & J_iJ_i &= J_-J_+ + J_3 + {J_3}^2, \\ [J_\pm, J_3] &= \mp J_\pm, \qquad & J_iJ_i &= J_+J_- - J_3 + {J_3}^2. \end{aligned} \tag{8.139}$$

The raising operator J_+ increases the eigenvalue of J_3 by an integer. But, for fixed l, μ must ultimately attain some maximum value. Denoting this value as μ_+, we must reach a state for which

$$J_+\psi(l,\mu_+) = 0. \tag{8.140}$$

Acting on this state with J_iJ_i and using one of the results in equation (8.139) we find that

$$(l+1/2)(l+3/2) = \mu_+(\mu_+ + 1). \tag{8.141}$$

Since l is positive and μ_+ represents an upper bound, it follows that

$$\mu_+ = l + 1/2. \tag{8.142}$$

There must similarly be a lowest eigenvalue of J_3 and a corresponding state with

$$J_-\psi(l,\mu_-) = 0. \tag{8.143}$$

In this case we find that

$$(l+1/2)(l+3/2) = \mu_-(\mu_- - 1), \tag{8.144}$$

hence $\mu_- = -(l+1/2)$. The spectrum of eigenvalues of J_3 therefore ranges from $(l+1/2)$ to $-(l+1/2)$, a total of $2(l+1)$ states. Since the J_3 eigenvalues are always of the form (integer $+1/2$), it is simpler to label the spherical monogenics with a pair of integers. We therefore write the spherical monogenics as ψ_l^m, where

$$-\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi_l^m = l\psi_l^m \qquad l \geq 0 \tag{8.145}$$

and

$$J_3\psi_l^m = (m+\tfrac{1}{2})\psi_l^m \qquad -1-l \leq m \leq l. \tag{8.146}$$

To find an explicit form for the ψ_l^m we first construct the highest m case. This satisfies

$$J_+\psi_l^l = 0 \tag{8.147}$$

and it is not hard to see that this equation is solved by

$$\psi_l^l \propto \sin^l(\theta)\,\mathrm{e}^{-l\phi I\boldsymbol{\sigma}_3}. \tag{8.148}$$

This is the angular part of the monogenic function $(x+yI\boldsymbol{\sigma}_3)^l$. Introducing a convenient factor, we write

$$\psi_l^l = (2l+1)P_l^l(\cos(\theta))\,\mathrm{e}^{l\phi I\boldsymbol{\sigma}_3}. \tag{8.149}$$

Our convention for the associated Legendre polynomials follows Gradshteyn & Ryzhik (1994), so we have

$$P_l^m(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l. \tag{8.150}$$

(Some useful recursion relations for the associated Legendre polynomials are discussed in the exercises.) The lowering operator J_- has the following effect on ψ:

$$J_-\psi = \big(-\partial_\theta\psi + \cot(\theta)\,\partial_\phi\psi I\boldsymbol{\sigma}_3\big)\mathrm{e}^{-\phi I\boldsymbol{\sigma}_3} - I\boldsymbol{\sigma}_2\,\tfrac{1}{2}(\psi + \boldsymbol{\sigma}_3\psi\boldsymbol{\sigma}_3). \tag{8.151}$$

The final term just projects out the $\{1, I\boldsymbol{\sigma}_3\}$ terms and multiplies them by $-I\boldsymbol{\sigma}_2$. This is the analog of the lowering matrix in the standard formalism. The derivatives acting on ψ_l^l form

$$\big(-\partial_\theta\psi_l^l + \cot(\theta)\,\partial_\phi\psi_l^l I\boldsymbol{\sigma}_3\big)\mathrm{e}^{-\phi I\boldsymbol{\sigma}_3} = (2l+1)2lP_l^{l-1}(\cos(\theta))\mathrm{e}^{(l-1)\phi I\boldsymbol{\sigma}_3}, \tag{8.152}$$

and, if we use the result that

$$\boldsymbol{\sigma}_\phi = \boldsymbol{\sigma}_2 \mathrm{e}^{\phi I \boldsymbol{\sigma}_3}, \tag{8.153}$$

we find that

$$\psi_l^{l-1} \propto \big(2lP_l^{l-1}(\cos(\theta)) - P_l^l(\cos(\theta)) I\boldsymbol{\sigma}_\phi\big)\mathrm{e}^{(l\,-\,1)\phi I\boldsymbol{\sigma}_3}. \tag{8.154}$$

Proceeding in this manner, we are led to the following formula for the spherical monogenics:

$$\psi_l^m = \big((l+m+1)P_l^m(\cos(\theta)) - P_l^{m+1}(\cos(\theta)) I\boldsymbol{\sigma}_\phi\big)\mathrm{e}^{m\phi I\boldsymbol{\sigma}_3}, \tag{8.155}$$

in which l is a positive integer or zero, m ranges from $-(l+1)$ to l and the P_l^m are taken to be zero if $|m| > l$. The positive- and negative-m states are related by

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x), \tag{8.156}$$

from which it can be shown that

$$\psi_l^m(-I\boldsymbol{\sigma}_2) = (-1)^m \frac{(l+m+1)!}{(l-m)!} \psi_l^{-(m+1)}. \tag{8.157}$$

The spherical monogenics presented here are unnormalised. Normalisation factors are not hard to compute, and we find that

$$\int_0^\pi d\theta \int_0^{2\pi} d\phi\, \sin(\theta)\, \psi_l^m \psi_l^{m\,\dagger} = 4\pi \frac{(l+m+1)!}{(l-m)!}. \tag{8.158}$$

If $\boldsymbol{\sigma}_r$ denotes the unit radial vector, $\boldsymbol{\sigma}_r = \boldsymbol{x}/r$ we find that

$$\boldsymbol{x} \wedge \boldsymbol{\nabla} \boldsymbol{\sigma}_r = 2\boldsymbol{\sigma}_r. \tag{8.159}$$

It follows that

$$-\boldsymbol{x} \wedge \boldsymbol{\nabla}(\boldsymbol{\sigma}_r \psi \boldsymbol{\sigma}_3) = -(l+2)\boldsymbol{\sigma}_r \psi \boldsymbol{\sigma}_3, \tag{8.160}$$

which provides an equation for the negative-l eigenstates. The possible eigenvalues and degeneracies are summarised in Table 8.2. One curious feature of this table is that we appear to be missing a line for the eigenvalue $l = -1$. In fact solutions for this case do exist, but they contain singularities which render them unnormalisable. For example, the functions

$$\frac{I\sigma_\phi}{\sin(\theta)}, \quad \text{and} \quad \frac{\mathrm{e}^{-I\boldsymbol{\sigma}_3\phi}}{\sin(\theta)} \tag{8.161}$$

have $l = -1$ and J_3 eigenvalues $+1/2$ and $-1/2$ respectively. Both solutions are singular along the z axis, however, which limits their physical relevance.

l	Eigenvalues of J_3	Degeneracy
⋮	⋮	⋮
2	$5/2, \ldots, -5/2$	6
1	$3/2, \ldots, -3/2$	4
0	$1/2, \ldots, -1/2$	2
(-1)	?	?
-2	$1/2, \ldots, -1/2$	2
⋮	⋮	⋮

Table 8.2 *Eigenvalues and degeneracies for the ψ_l^m monogenics.*

8.4.2 The radial equations

We can use the angular monogenics to construct eigenfunctions of the Dirac Hamiltonian of equation (8.123). Since the J_B operators commute with $\mathcal{H}$, ψ can be placed in an eigenstate of J_3. The operator J_iJ_i must also commute with $\mathcal{H}$, so $(l+1/2)(l+3/2)$ is a good quantum number. The operator $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ does not commute with $\mathcal{H}$, however, so both the ψ_l^m and $\boldsymbol{\sigma}_r\psi_l^m\boldsymbol{\sigma}_3$ monogenics are needed in the solution. While $\boldsymbol{x}\wedge\boldsymbol{\nabla}$ does not commute with $\mathcal{H}$, the operator

$$\hat{K} = \hat{\gamma}_0(1 - \boldsymbol{x}\wedge\boldsymbol{\nabla}) \tag{8.162}$$

does, as follows from

$$[\hat{\gamma}_0(1 - \boldsymbol{x}\wedge\boldsymbol{\nabla}), \boldsymbol{\nabla}] = 2\hat{\gamma}_0\boldsymbol{\nabla} - \hat{\gamma}_0\dot{\boldsymbol{\nabla}}\dot{\boldsymbol{x}}\wedge\boldsymbol{\nabla} = 0. \tag{8.163}$$

We should therefore work with eigenstates of the $\hat{K}$ operator. This implies that $\psi(\boldsymbol{x})$ can be written for positive l as either

$$\psi(\boldsymbol{x}, l+1) = \psi_l^m u(r) + \boldsymbol{\sigma}_r\psi_l^m v(r) I\boldsymbol{\sigma}_3 \tag{8.164}$$

or

$$\psi(\boldsymbol{x}, -(l+1)) = \boldsymbol{\sigma}_r\psi_l^m\boldsymbol{\sigma}_3 u(r) + \psi_l^m I v(r). \tag{8.165}$$

In both cases the second label in $\psi(\boldsymbol{x}, l+1)$ specifies the eigenvalue of $\hat{K}$. It is useful to denote this by κ, so we have

$$\hat{K}\psi = \kappa\psi, \qquad \kappa = \ldots, -2, -1, 1, 2, \ldots \tag{8.166}$$

and κ is a non-zero positive or negative integer.

In equations (8.164) and (8.165) the radial functions $u(r)$ and $v(r)$ are 'complex' combinations of 1 and $I\boldsymbol{\sigma}_3$. In the case of the Hamiltonian of (8.123), with $V(r)$ real, it turns out that the real and imaginary equations decouple, and it is

sufficient to treat $u(r)$ and $v(r)$ as real, scalar quantities. On substituting our trial functions into the Hamiltonian, we find that the radial equations reduce to

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} (\kappa-1)/r & -(E-eV(r)+m) \\ E-eV(r)-m & (-\kappa-1)/r \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{8.167}$$

The same equation holds for all values of κ. This successfully separates the Dirac equation in any radially-symmetric potential. As one might expect, we arrive at a pair of coupled first-order equations, as opposed to the single second-order equation familiar from Schrödinger theory.

8.4.3 The hydrogen atom

The radial equations describing the relativistic quantum theory of the hydrogen atom are obtained simply by setting $eV = -Z\alpha/r$, where $\alpha = e^2/4\pi$ is the fine structure constant and Z is the atomic charge. The solution of the radial equations is described in most textbooks on relativistic quantum mechanics. The conclusion is that the radial dependence is governed by a pair of *hypergeometric functions,* which generalise the Laguerre polynomials of the non-relativistic theory. Rather than reproduce the analysis here, we instead present a more direct method of solving the equations, first given by Eddington (1936) in his unconventional *Relativity Theory of Protons and Electrons*.

We start with the equation

$$-j\boldsymbol{\nabla}\psi - \frac{Z\alpha}{r}\psi + m\hat{\gamma}_0\psi = E\psi. \tag{8.168}$$

We assume that ψ is in an eigenstate of $\hat{K}$, so we can write

$$\boldsymbol{x}\wedge\boldsymbol{\nabla}\psi = \psi - \kappa\hat{\gamma}_0\psi. \tag{8.169}$$

We now pre-multiply the Dirac equation by $j\boldsymbol{x}$ and rearrange to find

$$r\partial_r\psi + \psi - \kappa\hat{\gamma}_0\psi = j\boldsymbol{x}\left(E + \frac{Z\alpha}{r}\right)\psi - jm\boldsymbol{x}\hat{\gamma}_0\psi. \tag{8.170}$$

On introducing the reduced function $\Psi = r\psi$ the equation simplifies to

$$\partial_r\Psi = j\boldsymbol{\sigma}_r(E - m\hat{\gamma}_0)\Psi + \frac{1}{r}(jZ\alpha\boldsymbol{\sigma}_r + \kappa\hat{\gamma}_0)\Psi. \tag{8.171}$$

We accordingly define the two operators

$$\hat{F} = -j\boldsymbol{\sigma}_r(E - m\hat{\gamma}_0), \qquad \hat{G} = -(jZ\alpha\boldsymbol{\sigma}_r + \kappa\hat{\gamma}_0), \tag{8.172}$$

so that the Dirac equation reduces to

$$\partial_r\Psi + \left(\hat{F} + \frac{\hat{G}}{r}\right)\Psi = 0. \tag{8.173}$$

The $\hat{F}$ and $\hat{G}$ operators satisfy

$$\begin{aligned}\hat{F}^2 &= m^2 - E^2 = f^2,\\ \hat{G}^2 &= \kappa^2 - (Z\alpha)^2 = \nu^2,\end{aligned} \tag{8.174}$$

which define f and ν. The operators also satisfy the anticommutation relation

$$\hat{F}\hat{G} + \hat{G}\hat{F} = -2Z\alpha E. \tag{8.175}$$

The next step is to transform to the dimensionless variable $x = fr$ and remove the large-x behaviour by setting

$$\Psi = \Phi \mathrm{e}^{-x}. \tag{8.176}$$

The function Φ now satisfies

$$\partial_x \Phi + \frac{\hat{G}}{x}\Phi + \Big(\frac{\hat{F}}{f} - 1\Big)\Phi = 0. \tag{8.177}$$

We are now in a position to consider a power series solution, so we set

$$\Phi = x^s \sum_{n=0} C_n x^n, \tag{8.178}$$

where the C_n are all multivectors. (In Eddington's original notation these are his 'e-numbers'.) The recursion relation is first-order and is given simply by

$$(n + s + \hat{G})C_n = -\left(\frac{\hat{F}}{f} - 1\right) C_{n-1}. \tag{8.179}$$

Setting $n = 0$ we see that

$$\big(s + \hat{G}\big)C_0 = 0. \tag{8.180}$$

Acting on this equation with the operator $\big(s - \hat{G}\big)$ we see that we must have $s^2 = \hat{G}^2 = \nu^2$. We set $s = \nu$ in order that the wavefunction is well behaved at the origin.

With the small and large x behaviour now separated out, all that remains is the power series. One can show that, in order for ψ to fall to zero at large distances, the series must terminate. We therefore set $C_{n+1} = 0$, and it follows that

$$\left(\frac{\hat{F}}{f} - 1\right) C_n = 0, \quad \text{or} \quad \hat{F}C_n = fC_n. \tag{8.181}$$

But we also have

$$\left(\frac{\hat{F}}{f} + 1\right)(n + \nu + \hat{G})C_n = -\left(\frac{\hat{F}}{f} + 1\right)\left(\frac{\hat{F}}{f} - 1\right) C_{n-1} = 0, \tag{8.182}$$

so

$$\left(2(n+\nu)+\hat{G}+\frac{\hat{F}}{f}\hat{G}\right)C_n=0. \tag{8.183}$$

If we write this as

$$\left(2(n+\nu)+\frac{1}{f}(\hat{G}\hat{F}+\hat{F}\hat{G})\right)C_n=0, \tag{8.184}$$

we find that we must have

$$n+\nu-\frac{Z\alpha E}{f}=0. \tag{8.185}$$

This is precisely our energy quantisation condition. The equation is equivalent to

$$\frac{E}{(m^2-E^2)^{1/2}}=\frac{n+\nu}{Z\alpha}, \tag{8.186}$$

which rearranges to the standard formula

$$E^2=m^2\left(1-\frac{(Z\alpha)^2}{n^2+2n\nu+\kappa^2}\right), \tag{8.187}$$

where n is a non-negative integer.

The non-relativistic formula for the energy levels is recovered by first recalling that $\alpha\approx 1/137$ is small. We can therefore approximate to

$$\nu\approx|\kappa|=l+1, \tag{8.188}$$

where $l\geq 0$ and

$$E\approx m\left(1-\frac{(Z\alpha)^2}{2}\frac{1}{n^2+2n(l+1)+(l+1)^2}\right). \tag{8.189}$$

Subtracting off the rest mass energy we are left with the non-relativistic expression

$$E_{NR}=-m\frac{(Z\alpha)^2}{2}\frac{1}{(n+l+1)^2}=-\frac{mZ^2e^4}{32\pi^2\epsilon_0^2\hbar^2}\frac{1}{n'^2}, \tag{8.190}$$

where $n'=n+l+1$ and the dimensional constants have been reinserted. We have recovered the familiar Bohr formula for the energy levels. This derivation shows that the relativistic quantum number n differs from the Bohr quantum number n'.

Expanding to next order we find that

$$E_{NR}=-m\frac{(Z\alpha)^2}{2n'^2}-m\frac{(Z\alpha)^4}{2n'^4}\left(\frac{n'}{l+1}-\frac{3}{4}\right). \tag{8.191}$$

The first relativistic correction shows that the binding energy is increased slightly from the non-relativistic value, and also introduces some dependence on the

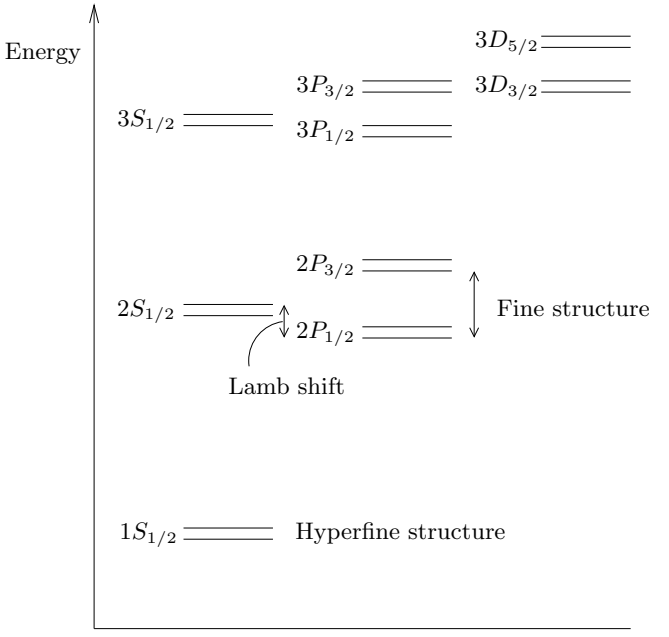

Figure 8.3 *Hydrogen atom energy levels.* The diagram illustrates how various degeneracies are broken by relativistic and spin effects. The Dirac equation accounts for the fine structure. The hyperfine structure is due to interaction with the magnetic moment of the nucleus. The Lamb shift is explained by quantum field theory. It lifts the degeneracy between the $S_{1/2}$ and $P_{1/2}$ states.

angular quantum number l. This lifts some degeneracies present in the non-relativistic solution. The various corrections contributing to the energy levels are shown in figure 8.3. A more complete analysis also requires replacing the electron mass m by the reduced mass of the two-body system. This introduces corrections of the same order of the relativistic corrections, but only affects the overall scale.

8.5 Scattering theory

Many of the experimental tests of Dirac theory, and quantum electrodynamics in general, are based on the results of scattering. Here we see how our new formulation can help to simplify these calculations through its handling of spin.

To aid this analysis it is useful to introduce the energy projection operators

$$\Lambda_{\pm}\psi = \frac{1}{2m}(m\psi \pm p\psi\gamma_0), \tag{8.192}$$

which project onto particle and antiparticle states.

A key role in relativistic quantum theory is played by Feynman propagators, which provide a means of imposing causal boundary conditions. We start by replacing the Dirac equation with the integral equation

$$\psi(x) = \psi_i(x) + e\int d^4x'\, S_F(x-x')A(x')\psi(x')\gamma_0, \tag{8.193}$$

where $\psi_i(x)$ is the asymptotic in-state and solves the free-particle equation, and $S_F(x-x')$ is the propagator. Substituting (8.193) into the Dirac equation, we find that $S_F(x-x')$ must satisfy

$$j\nabla_x S_F(x-x')\psi(x')\gamma_0 - mS_F(x-x')\psi(x') = \delta^4(x-x')\psi(x'). \tag{8.194}$$

The solution to this equation is

$$S_F(x-x')\psi(x') = \int \frac{d^4p}{(2\pi)^4}\,\frac{p\psi(x')\gamma_0 + m\psi(x')}{p^2 - m^2 + j\epsilon}\mathrm{e}^{-jp\cdot(x-x')}. \tag{8.195}$$

The factor of $j\epsilon$ is a mnemonic device to tell us how to negotiate the poles in the complex energy integral, which is performed first. The factor ensures positive-frequency waves propagate into the future ($t > t'$) and negative-frequency waves propagate into the past ($t' > t$). The result of performing the energy integration is summarised in the expression

$$S_F(x) = -2mj\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{1}{2E}\Big(\theta(t)\Lambda_+\mathrm{e}^{-jp\cdot x} + \theta(-t)\Lambda_-\mathrm{e}^{jp\cdot x}\Big), \tag{8.196}$$

where $E = +\sqrt{\boldsymbol{p}^2 + m^2}$.

There are other choices of relativistic propagator, which may be appropriate in other settings. For classical electromagnetism, for example, it is necessary to work with retarded propagators. If one constructs a closed spacetime surface integral, with boundary conditions consistent with the field equations, then the choice of propagator is irrelevant, since they all differ by a spacetime monogenic function. In most applications, however, we do not work like this. Instead we work with initial data, which we seek to propagate to a later time in such a way that the final result is consistent with imposing causal boundary conditions. In this case one has to use the Feynman propagator for quantum fields.

8.5.1 Electron scattering

In scattering calculations we write the wavefunction as the sum of an incoming plane wave and a scattered beam,

$$\psi(x) = \psi_i(x) + \psi_{diff}(x). \tag{8.197}$$

At asymptotically large times ψ_{diff} is given by

$$\psi_{diff}(x) = -2mje \int d^4x' \int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\frac{1}{2E}\Lambda_+\big(A(x')\psi(x')\gamma_0\big)\mathrm{e}^{-jp\cdot(x\,-\,x')}. \tag{8.198}$$

This can be written as a sum over final states

$$\psi_{diff}(x) = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\psi_f \mathrm{e}^{-jp_f\cdot x}, \tag{8.199}$$

where the final states are plane waves with

$$\psi_f = -je\int d^4x'\big(p_f A(x')\psi(x') + mA(x')\psi(x')\gamma_0\big)\mathrm{e}^{jp_f\cdot x'}. \tag{8.200}$$

The number of scattered particles is given by (recalling that $J = \psi\gamma_0\tilde{\psi}$)

$$\int d^3\boldsymbol{x}\,\gamma_0\cdot J_{diff} = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\left(\frac{\gamma_0\cdot J_f}{2E_f}\right) = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}N_f, \tag{8.201}$$

where N_f is the number density per Lorentz-invariant phase space interval:

$$N_f = \frac{\gamma_0\cdot J_f}{2E_f} = \frac{\gamma_0\cdot(\psi_f\gamma_0\tilde{\psi}_f)}{2E_f} = \frac{\rho_f}{2m}. \tag{8.202}$$

The integral equation (8.193) is the basis for a perturbative approach to solving the Dirac equation in an external field. We seek the full propagator S_A which satisfies

$$\big(j\nabla_2 - eA(x_2)\big)S_A(x_2,x_1)\gamma_0 - mS_A(x_2,x_1) = \delta^4(x_2-x_1). \tag{8.203}$$

The iterative solution to this is provided by

$$\begin{aligned} S_A(x_f,x_i) &= S_F(x_f-x_i) + \int d^4x_1\, S_F(x_f-x_1)e\hat{A}(x_1)S_F(x_1-x_i) \\ &+ \iint d^4x_1\, d^4x_2\, S_F(x_f-x_1)e\hat{A}(x_1)S_F(x_1-x_2)e\hat{A}(x_2)S_F(x_2-x_i) + \cdots, \end{aligned} \tag{8.204}$$

which is the basis for a diagrammatic representation of a scattering calculation. In the Born approximation we work to first order and truncate the series for S_A

after the first interaction term. Assuming incident plane waves of momentum p_i, so that $\psi_i(x) = \psi_i \exp(-jp_i \cdot x)$, we find that the final states become

$$\begin{aligned}\psi_f &= -je \int d^4x' \left(p_f A(x') + A(x')p_i\right)\psi_i e^{jq\cdot x'} \\ &= -je\big(p_f A(q) + A(q)p_i\big)\psi_i,\end{aligned} \tag{8.205}$$

where $q = p_f - p_i$ is the change in momentum, and $A(q)$ is the Fourier transform of the electromagnetic potential. The form of the result here is quite typical, and in general we can write

$$\psi_f = S_{fi}\psi_i, \tag{8.206}$$

where S_{fi} is the *scattering operator*. This is a multivector that takes initial states into final states. Since both ψ_i and ψ_f are plane-wave particle states, we must have

$$S_{fi}\tilde{S}_{fi} = \rho_{fi}, \tag{8.207}$$

where ρ_{fi} is a scalar quantity (which determines the cross section). We can therefore decompose S_{fi} as

$$S_{fi} = \rho_{fi}^{1/2} R_{fi}, \tag{8.208}$$

where R_{fi} is a rotor. This rotor takes the initial momentum to the final momentum,

$$R_{fi}p_i\tilde{R}_{fi} = p_f. \tag{8.209}$$

8.5.2 Spin effects in scattering

The multivector S_{fi} depends on the initial and final momenta and, in some cases, the initial spin. The final spin is determined from the initial spin by the rotation encoded in S_{fi}. If s_i and s_f denote the initial and final (unit) spin vectors, we have

$$s_f = R_{fi}s_i\tilde{R}_{fi}. \tag{8.210}$$

Sometimes it is of greater interest to separate out the boost terms in R_{fi} to isolate a pure rotation in the γ_0 frame. This tells us directly what happens to the spin vector in the electron's rest frame. With L_i and L_f the appropriate pure boosts, we define the rest spin scattering operator

$$U_{fi} = \tilde{L}_f R_{fi} L_i. \tag{8.211}$$

This satisfies

$$U_{fi}\gamma_0\tilde{U}_{fi} = \frac{1}{m}\tilde{L}_f R_{fi} p_i \tilde{R}_{fi} L_f = \gamma_0, \tag{8.212}$$

so is a pure rotation in the γ_0 frame.

The fact that $p_f S_{fi} = S_{fi} p_i$ ensures that S_{fi} is always of the form

$$S_{fi} = -j(p_f M + M p_i), \tag{8.213}$$

where M is an odd-grade multivector. In the Born approximation of equation (8.205), for example, we have $M = eA(q)$. In general, M can contain both real and imaginary terms, so we must write

$$S_{fi}\psi_i = -j\big(p_f(M_r + jM_j) + (M_r + jM_j)p_i\big)\psi_i, \tag{8.214}$$

where M_j and M_r are independent of j. We can now use

$$j\psi_i = \psi_i I\boldsymbol{\sigma}_3 = \hat{S}_i\psi_i, \tag{8.215}$$

where $\hat{S}_i$ is the initial unit spin bivector. Since $\hat{S}_i$ and p_i commute, S_{fi} can still be written in the form of equation (8.213), with

$$M = M_r + M_j\hat{S}_i. \tag{8.216}$$

So M remains a real multivector, which now depends on the initial spin. This scheme is helpful if we are interested in any spin-dependent features of the scattering process.

8.5.3 Positron scattering and pair annihilation

Adapting the preceding results to positron scattering is straightforward. In this case a negative-energy plane wave arrives from the future and scatters into the past, so we set

$$\psi_i(x) = \psi_2 \mathrm{e}^{jp_i\cdot x}, \quad \psi_f(x) = \psi_f \mathrm{e}^{jp_f\cdot x}. \tag{8.217}$$

In this case repeating the analysis gives

$$S_{fi}\psi_i = -j(-p_f M\psi_i + M\psi_i\gamma_0), \tag{8.218}$$

which we can write as

$$S_{fi} = j(p_f M + M p_i). \tag{8.219}$$

This amounts to simply swapping the sign of S_{fi}. In the Born approximation, q is replaced by $-q$ in the Fourier transform of $A(x)$, which will alter the factor M if $A(x)$ is complex.

The other case to consider is when the incoming electron is scattered into the past, corresponding to pair annihilation. In this case we have

$$S_{fi} = -j(-p_2 M + M p_1), \tag{8.220}$$

where p_1 and p_2 are the incoming momenta of the electron and positron respectively. We decompose S_{fi} as

$$S_{fi} = \rho_{fi}^{1/2} I R_{fi}, \tag{8.221}$$

since S_{fi} must now contain a factor of I to map electrons into positrons. This form for S_{fi} implies that

$$S_{fi}\tilde{S}_{fi} = -\rho_{fi}. \tag{8.222}$$

The minus sign reflects the fact that the transformation between initial and final momenta is not proper orthochronous.

8.5.4 Cross sections

We must now relate our results to the cross sections measured in experiments. The scattering rate into the final states, per unit volume, per unit time, is given by

$$W_{fi} = \frac{1}{VT}N_f = \frac{1}{VT}\frac{\gamma_0 \cdot J_f}{2E_f} = \frac{\rho_f}{2mVT}, \tag{8.223}$$

where V and T denote the total volume and time respectively. The density ρ_f is given by

$$\rho_f = |S_{fi}\tilde{S}_{fi}|\rho_i = \rho_{fi}\rho_i. \tag{8.224}$$

Here $S_{fi}\tilde{S}_{fi} = \pm\rho_{fi}$, where the plus sign corresponds to electron to electron and positron to positron scattering, and the minus sign to electron–positron annihilation.

The differential cross section is defined as

$$d\sigma = \frac{W_{fi}}{\text{target density} \times \text{incident flux}}. \tag{8.225}$$

When S_{fi} is of the form

$$S_{fi} = -j(2\pi)^4\delta^4(P_f - P_i)T_{fi}, \tag{8.226}$$

where the δ-function ensures conservation of total momentum, we have

$$|S_{fi}|^2 = VT(2\pi)^4\delta^4(P_f - P_i)|T_{fi}|^2. \tag{8.227}$$

Working in the J_i frame the target density is just ρ_i so, writing the incident flux as χ, we have

$$d\sigma = \frac{1}{2m\chi}(2\pi)^4\delta^4(P_f - P_i)|T_{fi}|^2. \tag{8.228}$$

Alternatively we may be interested in an elastic scattering with just energy conservation ($E_f = E_i$) and

$$S_{fi} = -j2\pi\delta(E_f - E_i)T_{fi}. \tag{8.229}$$

In this case

$$|S_{fi}|^2 = 2\pi T\delta(E_f - E_i)|T_{fi}|^2. \tag{8.230}$$

A target density of $1/V$ and an incident flux of $|\boldsymbol{J}_i| = \rho_i|\boldsymbol{p}_i|/m$ then gives

$$d\sigma = \frac{\pi}{|\boldsymbol{p}_i|}\delta(E_f - E_i)|T_{fi}|^2. \tag{8.231}$$

The total cross section is obtained by integrating over the available phase space. For the case of a single particle scattering elastically we find that

$$\sigma = \int \frac{d^3\boldsymbol{p}_f}{(2\pi)^3}\frac{1}{2E_f}\frac{\pi}{|\boldsymbol{p}_i|}\delta(E_f - E_i)|T_{fi}|^2 = \int d\Omega\,\frac{|T_{fi}|^2}{16\pi^2}. \tag{8.232}$$

This is usually expressed in terms of the differential cross section per solid angle:

$$\frac{d\sigma}{d\Omega_f} = \frac{|T_{fi}|^2}{16\pi^2}. \tag{8.233}$$

8.5.5 Coulomb scattering

As an application of our formalism consider Coulomb scattering from a nucleus, with the external field defined by

$$A(x) = \frac{-Ze}{4\pi|\boldsymbol{x}|}\gamma_0. \tag{8.234}$$

Working with the first Born approximation, M is given by $M = eA(q)$, where $A(q)$ is the Fourier transform of $A(x)$ given by

$$A(q) = -\frac{2\pi Ze}{\boldsymbol{q}^2}\delta(E_f - E_i)\gamma_0 \tag{8.235}$$

and $q{\cdot}\gamma_0 = E_f - E_i$. Writing

$$S_{fi} = -j2\pi\delta(E_f - E_i)T_{fi} \tag{8.236}$$

and using energy conservation we find that

$$T_{fi} = -\frac{Ze^2}{\boldsymbol{q}^2}(2E + \boldsymbol{q}). \tag{8.237}$$

The cross section is therefore given by the Mott scattering formula:

$$\frac{d\sigma}{d\Omega_f} = \frac{Z^2\alpha^2}{\boldsymbol{q}^4}(4E^2 - \boldsymbol{q}^2) = \frac{Z^2\alpha^2}{4\boldsymbol{p}^2\beta^2\sin^4(\theta/2)}\left(1 - \beta^2\sin^2(\theta/2)\right), \tag{8.238}$$

where

$$\boldsymbol{q}^2 = (\boldsymbol{p}_f - \boldsymbol{p}_i)^2 = 2\boldsymbol{p}^2\big(1 - \cos(\theta)\big) \quad \text{and} \quad \beta = |\boldsymbol{p}|/E. \tag{8.239}$$

The angle θ measures the deviation between the incoming and scattered beams. In the low velocity limit the Mott result reduces to the Rutherford formula. The result is independent of the sign of the nuclear charge and, to this order, is obtained for both electron and positron scattering.

A significant feature of this derivation is that no spin sums are required. Instead, all the spin dependence is contained in the directional information in T_{fi}. As well as being computationally more efficient, this method for organising cross section calculations offers deeper insights into the structure of the theory. For Coulomb scattering the spin information is contained in the rotor

$$R_{fi} = \frac{p_f\gamma_0 + \gamma_0 p_i}{4E^2 - \boldsymbol{q}^2} \propto L_f^2 + \tilde{L}_i^2, \tag{8.240}$$

where L_f and L_i are the pure boosts from γ_0 to p_f and p_i respectively. The behaviour of the rest spin is governed by the unnormalised rotor

$$U_{fi} = \tilde{L}_f(L_f^2 + \tilde{L}_i^2)L_i = L_f L_i + \tilde{L}_f\tilde{L}_i, = 2\big((E+m)^2 + \boldsymbol{p}_f\boldsymbol{p}_i\big). \tag{8.241}$$

It follows that the rest-spin vector precesses in the $\boldsymbol{p}_f \wedge \boldsymbol{p}_i$ plane through an angle δ, where

$$\tan(\delta/2) = \frac{\sin(\theta)}{(E+m)/(E-m) + \cos(\theta)}. \tag{8.242}$$

This method of calculating the spin precession for Coulomb scattering was first described by Hestenes (1982a).

8.5.6 Compton scattering

Compton scattering is the process in which an electron scatters off a photon. To lowest order there are two Feynman diagrams to consider, shown in figure 8.4. The preceding analysis follows through with little modification, and gives rise two terms of the form

$$M_1 = e^2 \iint d^4x_1\, d^4x_2\, \frac{d^4p}{(2\pi)^4}\, A_1(x_1)\frac{pA_2(x_2) + A_2(x_2)p_i}{p^2 - m^2 + j\epsilon} \times \mathrm{e}^{jx_1\cdot(p_f - p)}\mathrm{e}^{jx_2\cdot(p - p_i)}, \tag{8.243}$$

where

$$A(x) = \epsilon \mathrm{e}^{\mp jk\cdot x} \tag{8.244}$$

is the (complex) vector potential. The vector ϵ denotes the polarisation state, so $k\cdot\epsilon = 0$ and $\epsilon^2 = -1$. In relativistic quantum theory there appears to be no alternative but to work with a fully complex vector potential.

Performing the integrations and summing the two contributions we arrive at

$$M = e^2(2\pi)^4\delta^4(P)\left(\epsilon_f\frac{(p_i + k_i)\epsilon_i + \epsilon_i p_i}{2k_i\cdot p_i} - \epsilon_i\frac{(p_i - k_f)\epsilon_f + \epsilon_f p_i}{2p_i\cdot k_f}\right), \tag{8.245}$$

where $P = p_f + k_f - p_i - k_i$, so that the δ-function enforces momentum conservation. Gauge invariance means that we can set $p_i\cdot\epsilon_i = p_i\cdot\epsilon_f = 0$, in which case

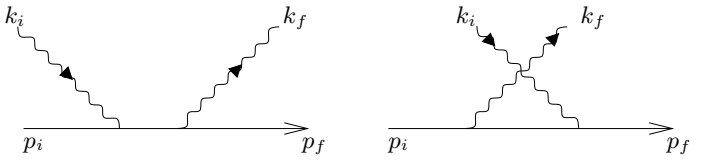

Figure 8.4 *Compton scattering.* Two diagrams contribute to the amplitude, to lowest order.

M simplifies to

$$M = e^2(2\pi)^4\delta^4(p_f + k_f - p_i - k_i)\left(\frac{\epsilon_f k_i \epsilon_i}{2k_i\cdot p_i} + \frac{\epsilon_i k_f \epsilon_f}{2p_i\cdot k_f}\right). \tag{8.246}$$

We now set

$$S_{fi} = -j(2\pi)^4\delta^4(p_f + k_f - p_i - k_i)T_{fi}, \tag{8.247}$$

so that the cross section is given by equation (8.228). After a little work, and making use of momentum conservation, we find that

$$|T_{fi}|^2 = e^4\left(4(\epsilon_i\cdot\epsilon_f)^2 - 2 + \frac{p_i\cdot k_f}{p_i\cdot k_i} + \frac{p_i\cdot k_i}{p_i\cdot k_f}\right). \tag{8.248}$$

This is all that is required to calculate the cross section in any desired frame. Again, this derivation applies regardless of the initial electron spin.

The same scheme can be applied to a wide range of relativistic scattering problems. In all cases the spacetime algebra formulation provides a simpler and clearer method for handling the spin, as it does not force us to work with a preferred basis set. In section 14.4.1 the same formalism is applied to scattering from a black hole. At some point, however, it is necessary to face questions of second quantisation and the construction of a relativistic multiparticle quantum theory. This is discussed in the following chapter.

8.6 Notes

A significant amount of new notation was introduced in this chapter, relating to how spinors are handled in spacetime algebra. Much of this is important in later chapters, and the most useful results of this approach are summarised in table 8.3.

Quantum mechanics has probably been the most widely researched application of geometric algebra to date. Many authors have carried out investigations into whether the spacetime algebra formulation of Dirac theory offers any deeper insights into the nature of quantum theory. Among the most interesting of these are Hestenes' work on zitterbewegung (1990), and his comments on the nature

Pauli spinors	$\|\psi\rangle = \begin{pmatrix} a^0 + ia^3 \\ -a^2 + ia^1 \end{pmatrix} \leftrightarrow \psi = a^0 + a^k I\boldsymbol{\sigma}_k$
Pauli operators	$\hat{\sigma}_k\|\psi\rangle \leftrightarrow \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3$ $i\|\psi\rangle \leftrightarrow \psi I\boldsymbol{\sigma}_3 = j\psi$ $\langle\psi\|\psi'\rangle \leftrightarrow \langle\psi^\dagger\psi'\rangle_q = \frac{1}{2}(\psi^\dagger\psi' + \boldsymbol{\sigma}_3\psi^\dagger\psi'\boldsymbol{\sigma}_3)$
Pauli observables	$\rho = \psi\psi^\dagger$ $\boldsymbol{s} = \frac{1}{2}\psi\boldsymbol{\sigma}_3\psi^\dagger$
Dirac spinors	$\begin{pmatrix} \|\phi\rangle \\ \|\eta\rangle \end{pmatrix} \leftrightarrow \psi = \phi + \eta\boldsymbol{\sigma}_3$
Dirac operators	$\hat{\gamma}_\mu\|\psi\rangle \leftrightarrow \gamma_\mu\psi\gamma_0$ $j\|\psi\rangle \leftrightarrow \psi i\boldsymbol{\sigma}_3$ $\hat{\gamma}_5\|\psi\rangle \leftrightarrow \psi\boldsymbol{\sigma}_3$ $\langle\bar{\psi}\|\psi'\rangle \leftrightarrow \langle\tilde{\psi}\psi'\rangle_q$
Dirac equation	$\nabla\psi I\boldsymbol{\sigma}_3 - eA\psi = m\psi\gamma_0$
Dirac observables	$\rho e^{i\beta} = \psi\tilde{\psi} \qquad J = \psi\gamma_0\tilde{\psi}$ $S = \psi i\boldsymbol{\sigma}_3\tilde{\psi} \quad s = \psi\gamma_3\tilde{\psi}$
Plane-wave states	$\psi^{(+)}(x) = L(p)\Phi e^{-i\boldsymbol{\sigma}_3 p\cdot x}$ $\psi^{(-)}(x) = L(p)\Phi\boldsymbol{\sigma}_3 e^{i\boldsymbol{\sigma}_3 p\cdot x}$ $L(p) = (p\gamma_0 + m)/\sqrt{2m(E+m)}$

Table 8.3 *Quantum states and operators.* This table summarises the main features of the spacetime algebra representation of Pauli and Dirac spinors and operators.

of the electroweak group (1982b). Many authors have advocated spacetime algebra as a better computational tool for Dirac theory than the explicit matrix formulation (augmented with various spin sum rules). A summary of these ideas is contained in the paper 'Electron scattering without spin sums' by Lewis *et al.* (2001). Elsewhere, a similar approach has been applied to modelling a spin measurement (Challinor *et al.* 1996) and to the results of tunnelling experiments (Gull *et al.* 1993b). Much of this work is summarised in the review 'Spacetime algebra and electron physics' by Doran *et al.* (1996b).

There is no shortage of good textbooks describing standard formulations of Dirac theory and quantum electrodynamics. We particularly made use of the classic texts by Itzykson & Zuber (1980), and Bjorken & Drell (1964). For a detailed exposition of the solution of the Dirac equation in various backgrounds one can do little better than Greiner's *Relativistic Quantum Mechanics* (1990).

Also recommended is Grandy's *Relativistic Quantum Mechanics of Leptons and Fields* (1991) which, unusually, does not shy away from the more problematic areas of the conceptual foundations of quantum field theory.

8.7 Exercises

8.1 The spin matrix operators $\hat{s}_k$ are defined as a set of 2×2 Hermitian matrices satisfying the commutation relations $[\hat{s}_i, \hat{s}_j] = i\hbar\epsilon_{ijk}\hat{s}_k$. Given that $\hat{s}_3$ is defined by

$$\hat{s}_3 = \lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

show that the remaining matrices are unique, up to an overall choice of phase. Find λ and show that we can choose the phase such that $\hat{s}_k = \hbar/2\,\hat{\sigma}_k$.

8.2 Verify that the equivalence between Pauli spinors and even multivectors defined in equation (8.20) is consistent with the operator equivalences

$$\hat{\sigma}_k|\psi\rangle \;\leftrightarrow\; \boldsymbol{\sigma}_k\psi\boldsymbol{\sigma}_3 \quad (k = 1, 2, 3).$$

8.3 Suppose that two spin-1/2 states are represented by the even multivectors ϕ and ψ, and the accompanying spin vectors are

$$\boldsymbol{s}_1 = \phi\boldsymbol{\sigma}_3\tilde{\phi} \quad \text{and} \quad \boldsymbol{s}_2 = \psi\boldsymbol{\sigma}_3\tilde{\psi}.$$

Prove that the quantum mechanical formula for the probability of measuring state ϕ in state ψ satisfies

$$P = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle} = \tfrac{1}{2}\big(1 + \cos(\theta)\big)$$

where θ is the angle between $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$.

8.4 Verify that the Pauli inner product is invariant under both spatial rotations and gauge transformations (i.e. rotations in the $I\boldsymbol{\sigma}_3$ plane applied to the right of the spinor ψ). Repeat the analysis for Dirac spinors.

8.5 Prove that the angular momentum operators $L_B = jB\cdot(\boldsymbol{x}\wedge\boldsymbol{\nabla})$ satisfy

$$[L_{B_1}, L_{B_2}] = -jL_{B_1\times B_2}.$$

8.6 Prove that, in any dimension,

$$[B\cdot(x\wedge\nabla) - \tfrac{1}{2}B, \nabla] = 0,$$

where B is a bivector.

8.7 The Majorana representation is defined in terms of a set of real matrices. Prove that the complex conjugation operation in this representation has the spacetime algebra equivalent

$$|\psi\rangle^*_{Maj} \leftrightarrow \psi\boldsymbol{\sigma}_2.$$

Confirm that this anticommutes with the operation of multiplying by the imaginary.

8.8 Prove that the associated Legendre polynomials satisfy the following recursion relations:

$$(1-x^2)\frac{dP_l^m(x)}{dx} + mxP_l^m(x) = -(1-x^2)^{1/2}P_l^{m+1}(x),$$

$$(1-x^2)\frac{dP_l^m(x)}{dx} - mxP_l^m(x) = (1-x^2)^{1/2}(l+m)(l-m+1)P_l^{m-1}(x).$$

8.9 Prove that the spherical monogenics satisfy

$$\int d\Omega \, \langle \psi_l^{m\dagger}\psi_{l'}^{m'}\rangle_q = \delta^{mm'}\delta_{ll'}4\pi\frac{(l+m+1)!}{(l-m)!}.$$

8.10 From the result of equation (8.248), show that the cross section for scattering of a photon of a free electron (initially at rest) is determined by the Klein–Nishina formula

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4m^2}\left(\frac{\omega_f}{\omega_i}\right)^2\left(\frac{\omega_f}{\omega_i} + \frac{\omega_i}{\omega_f} + 4(\epsilon_f\cdot\epsilon_i)^2 - 2\right).$$

9

Multiparticle states and quantum entanglement

The previous chapter dealt with the quantum theory of single particles in a background field. In this chapter we turn to the study of multiparticle quantum theory. In many ways, this subject is even more strange than the single-particle theory, as it forces us to face up to the phenomenon of quantum entanglement. The basic idea is simple enough to state. The joint state of a two-particle system is described by a *tensor product* state of the form $|\psi\rangle \otimes |\phi\rangle$. This is usually abbreviated to $|\psi\rangle|\phi\rangle$. Quantum theory allows for linear complex superpositions of multiparticle states, which allows us to consider states which have no classical counterpart. An example is the spin singlet state

$$|\varepsilon\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|1\rangle - |1\rangle|0\rangle\big). \tag{9.1}$$

States such as these are referred to as being *entangled*. The name reflects the fact that observables for the two particles remain correlated, even if measurements are performed in such a way that communication between the particles is impossible. The rapidly evolving subject of *quantum information processing* is largely concerned with the properties of entangled states, and the prospects they offer for quantum computation.

Quantum entanglement is all around us, though rarely in a form we can exploit. Typically, a state may entangle with its environment to form a new *pure* state. (A pure state is one that can be described by a single wavefunction, which may or may not be entangled.) The problem is that our knowledge of the state of the environment is highly limited. All we can measure are the observables of our initial state. In this case the wavefunction formulation is of little practical value, and instead we have to consider equations for the evolution of the observables themselves. This is usually handled by employing a representation in terms of *density matrices*. These lead naturally to concepts of quantum statistical physics and quantum definitions of entropy.

In this chapter we explore how these concepts can be formulated in the language of geometric algebra. One of the essential mysteries of quantum theory is the origin of this tensor product construction. The tensor product is used in constructing both multiparticle states and many of the operators acting on these states. So the first challenge is to find a representation of the tensor product in terms of the geometric product. This is surprisingly simple to do, though only once we have introduced the idea of a relativistic configuration space. The geometric algebra of such a space is called the *multiparticle spacetime algebra* and it provides the ideal algebraic structure for studying multiparticle states and operators. This has applications in a wealth of subjects, from NMR spectroscopy to quantum information processing, some of which are discussed below. Most of these applications concern non-relativistic multiparticle quantum mechanics. Later in this chapter we turn to a discussion of the insights that this new approach can bring to relativistic multiparticle quantum theory. There we find a simple, geometric encoding of the Pauli principle, which opens up a route through to the full quantum field theory.

9.1 Many-body quantum theory

In order to set the context for this chapter, we start with a review of the basics of multiparticle quantum theory. We concentrate in particular on two-particle systems, which illustrate many of the necessary properties. The key concept is that the quantum theory of n-particles is *not* described by a set of n single wavefunctions. Instead, it is described by *one* wavefunction that encodes the entire state of the system of n particles. Unsurprisingly, the equations governing the evolution of such a wavefunction can be extraordinarily complex.

For a wide range of problems one can separate position degrees of freedom from internal (spin) degrees of freedom. This is typically the case in non-relativistic physics, particularly if the electromagnetic field can be treated as constant. In this case the position degrees of freedom are handled by the many-body Schrödinger equation. The spin degrees of freedom in many ways represent a cleaner system to study, as they describe the quantum theory of n two-state systems. This illustrates the two most important features of multiparticle quantum theory: the exponential increase in the size of state space, and the existence of entangled states.

9.1.1 The two-body Schrödinger equation

Two-particle states are described by a single wavefunction $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$. The joint vectors $(\boldsymbol{r}_1, \boldsymbol{r}_2)$ define an abstract six-dimensional configuration space over which ψ defines a complex-valued function. This sort of configuration space is a useful tool in classical mechanics, and in quantum theory it is indispensable. The

kinetic energy operator is given by the sum of the individual operators:

$$\hat{K} = -\frac{\hbar^2 \boldsymbol{\nabla}_1^2}{2m_1} - \frac{\hbar^2 \boldsymbol{\nabla}_2^2}{2m_2}. \tag{9.2}$$

The subscripts refer to the individual particles, and m_i is the mass of particle i. The two-particle Schrödinger equation is now

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2 \boldsymbol{\nabla}_1^2}{2m_1}\psi - \frac{\hbar^2 \boldsymbol{\nabla}_2^2}{2m_2}\psi + V(\boldsymbol{r}_1, \boldsymbol{r}_2)\psi. \tag{9.3}$$

As a simple example, consider the bound state Coulomb problem

$$-\frac{\hbar^2 \boldsymbol{\nabla}_1^2}{2m_1}\psi - \frac{\hbar^2 \boldsymbol{\nabla}_2^2}{2m_2}\psi - \frac{q_1 q_2}{4\pi\epsilon_0 r}\psi = E\psi, \tag{9.4}$$

where r is the Euclidean distance between the points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. This problem is separated in a similar manner to the classical Kepler problem (see section 3.2). We introduce the vectors

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2, \quad \frac{\boldsymbol{R}}{\mu} = \frac{\boldsymbol{r}_1}{m_1} + \frac{\boldsymbol{r}_2}{m_2}, \tag{9.5}$$

where μ is the reduced mass. In terms of these new variables the Schrödinger equation becomes

$$-\frac{\hbar^2 \boldsymbol{\nabla}_r^2}{2\mu}\psi - \frac{\hbar^2 \boldsymbol{\nabla}_R^2}{2M}\psi - \frac{q_1 q_2}{4\pi\epsilon_0 r}\psi = E\psi. \tag{9.6}$$

We can now find separable solutions to this equation by setting

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi(\boldsymbol{r})\Psi(\boldsymbol{R}). \tag{9.7}$$

The wavefunction Ψ satisfies a free-particle equation, which corresponds classically to the motion of the centre of mass. The remaining term, $\phi(\boldsymbol{r})$, satisfies the equivalent single-particle equation, with the mass given by the reduced mass of the two particles.

This basic example illustrates how quantum mechanics accounts for multiparticle interactions. There is a single wavefunction, which simultaneously accounts for the properties of all of the particles. In many cases this wavefunction decomposes into the product of a number of simpler wavefunctions, but this is not always the case. One can construct states that cannot be decomposed into a single direct product state. An important example of this arises when the two particles in question are identical. In this case one can see immediately that if $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is an eigenstate of a two-particle Hamiltonian, then so to is $\psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$. The operator that switches particle labels like this is called the particle interchange operator $\hat{P}$, and it commutes with all physically-acceptable Hamiltonians. Since it commutes with the Hamiltonian, and squares to the identity operation, there are two possible eigenstates of $\hat{P}$. These are

$$\psi_\pm = \psi(\boldsymbol{r}_1, \boldsymbol{r}_2) \pm \psi(\boldsymbol{r}_2, \boldsymbol{r}_1). \tag{9.8}$$

These two possibilities are the only ones that arise physically, and give rise to the distinction between fermions (minus sign) and bosons (plus sign). Here we see the first indications of some new physical possibilities entering in multiparticle interactions. Quantum theory remains linear, so one can form complex superpositions of the n-particle wavefunctions. These superpositions can have new properties not present in the single-particle theory.

9.1.2 Spin states

Ignoring the spatial dependence and concentrating instead on the internal spin degrees of freedom, a spin-1/2 state can be written as a complex superposition of 'up' and 'down' states, which we will denote as $|0\rangle$ and $|1\rangle$. Now suppose that a second particle is introduced, so that system 1 is in the state $|\psi\rangle$ and system 2 is in the state $|\phi\rangle$. The joint state of the system is described by the *tensor product* state

$$|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle, \tag{9.9}$$

which is abbreviated to $|\psi\rangle|\phi\rangle$. The total set of possible states is described by the basis

$$\begin{aligned} |00\rangle &= |0\rangle|0\rangle, & |01\rangle &= |0\rangle|1\rangle, \\ |10\rangle &= |1\rangle|0\rangle, & |11\rangle &= |1\rangle|1\rangle. \end{aligned} \tag{9.10}$$

This illustrates an important phenomenon of multiparticle quantum theory. The number of available states grows as 2^n, so large systems have an enormously larger state space than their classical counterparts. Superpositions of these basis states will, in general, produce states which cannot be written as a single tensor product of the form $|\psi\rangle|\phi\rangle$. Such states are entangled. A standard example is the singlet state of equation (9.1). One feature of these entangled states is that they provide 'short-cuts' through Hilbert space between classical states. The speed-up this can offer is often at the core of algorithms designed to exploit the possibilities offered by quantum computation.

A challenge faced by theorists looking for ways to exploit these ideas is how best to classify multiparticle entanglement. The problem is to describe concisely the properties of a state that are unchanged under local unitary operations. Local operations consist of unitary transformations applied entirely to one particle. They correspond to operations applied to a single particle in the laboratory. Features of the state that are unchanged by these operations relate to joint properties of the particles, in particular how entangled they are.

To date, only two-particle (or 'bipartite') systems have been fully understood. A general state of two particles can be written

$$\Psi = \sum_{i,j} \alpha_{ij} |i\rangle \otimes |j\rangle, \tag{9.11}$$

where the $|i\rangle$ denote some orthonormal basis. The *Schmidt decomposition* (which is little more than a singular-value decomposition of α_{ij}) tells us that one can always construct a basis such that

$$\Psi = \sum_i \beta_i |i'\rangle \otimes |i'\rangle. \tag{9.12}$$

The β_i are all real parameters that tell us directly how much entanglement is present. These parameters are unchanged under local transformations of the state Ψ. An important example of the Schmidt decomposition, which we shall revisit frequently, is for systems of two entangled spinors. For these we find that a general state can be written explicitly as

$$\begin{aligned}|\psi\rangle =& \rho^{1/2} \mathrm{e}^{i\chi} \left(\cos(\alpha/2) \mathrm{e}^{i\tau/2} \begin{pmatrix} \cos(\theta_1/2)\mathrm{e}^{-i\phi_1/2} \\ \sin(\theta_1/2)\mathrm{e}^{i\phi_1/2} \end{pmatrix} \otimes \begin{pmatrix} \cos(\theta_2/2)\mathrm{e}^{-i\phi_2/2} \\ \sin(\theta_2/2)\mathrm{e}^{i\phi_2/2} \end{pmatrix} \right. \\ &\left. + \sin(\alpha/2) \mathrm{e}^{-i\tau/2} \begin{pmatrix} \sin(\theta_1/2)\mathrm{e}^{-i\phi_1/2} \\ -\cos(\theta_1/2)\mathrm{e}^{i\phi_1/2} \end{pmatrix} \otimes \begin{pmatrix} \sin(\theta_2/2)\mathrm{e}^{-i\phi_2/2} \\ -\cos(\theta_2/2)\mathrm{e}^{i\phi_2/2} \end{pmatrix} \right). \end{aligned} \tag{9.13}$$

In this decomposition we arrange that $0 \leq \alpha \leq \pi/4$, so that the decomposition is unique (save for certain special cases).

9.1.3 Pure and mixed states

So far the discussion has focused entirely on *pure* states, which can be described in terms of a single wavefunction. For many applications, however, such a description is inappropriate. Suppose, for example, that we are studying spin states in an NMR experiment. The spin states are only partially coherent, and one works in terms of ensemble averages. For example, the average spin vector (or polarisation) is given by

$$\boldsymbol{p} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{s}}_i. \tag{9.14}$$

Unless all of the spin vectors are precisely aligned (a coherent state), the polarisation vector will not have unit length and so cannot be generated by a single wavefunction. Instead, we turn to a formulation in terms of density matrices. The density matrix for a normalised pure state is

$$\hat{\rho} = |\psi\rangle\langle\psi|, \tag{9.15}$$

which is necessarily a Hermitian matrix. All of the observables associated with the state $|\psi\rangle$ can be obtained from the density matrix by writing

$$\langle\psi|\hat{Q}|\psi\rangle = \mathrm{tr}(\hat{\rho}\hat{Q}). \tag{9.16}$$

For an incoherent mixture (a mixed state) the density matrix is the weighted sum of the matrices for the pure states:

$$\hat{\rho} = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|. \tag{9.17}$$

The real coefficients satisfy

$$\sum_{i=1}^{n} p_i = 1, \tag{9.18}$$

which ensures that the density matrix has unit trace. The definition of $\hat{\rho}$ ensures that all observables are constructed from the appropriate averages of the pure states. In principle, the state of any system is described by a Hermitian density matrix, which is constrained to be positive-semidefinite and to have unit trace. All observables are then formed according to equation (9.16).

The need for a density matrix can be seen in a second way, as a consequence of entanglement. Suppose that we are interested in the state of particle 1, but that this particle has been allowed to entangle with a second particle 2, forming the pure state $|\psi\rangle$. The density matrix for the two-particle system is again described by equation (9.15). But we can only perform measurements of particle 1. The effective density matrix for particle 1 is obtained by performing a *partial trace* of $\hat{\rho}$ to trace out the degrees of freedom associated with particle 2. We therefore define

$$\hat{\rho}_1 = \mathrm{tr}_2 \hat{\rho}, \tag{9.19}$$

where the sum runs over the space of particle 2. One can easily check that, in the case where the particles are entangled, $\hat{\rho}_1$ is no longer the density matrix for a pure state. The most extreme example of this is the singlet state (9.1) mentioned in the introduction. In the obvious basis, the singlet state can be written as

$$|\varepsilon\rangle = \frac{1}{\sqrt{2}}(0,\, 1,\, -1,\, 0)^{\dagger}. \tag{9.20}$$

The density matrix for this state is

$$\hat{\rho} = |\varepsilon\rangle\langle\varepsilon| = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{9.21}$$

This is appropriate for a pure state, as the matrix satisfies $\hat{\rho}^2 = \hat{\rho}$. But if we now form the partial trace over the second particle we are left with

$$\hat{\rho}_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{9.22}$$

This is the density matrix for a totally unpolarised state, which is to be expected, since there can be no directional information in the singlet state. Clearly, $\hat{\rho}_1$ cannot be generated by a single-particle pure state.

9.2 Multiparticle spacetime algebra

The key to constructing a suitable geometric framework for multiparticle quantum theory involves the full, relativistic spacetime algebra. This is because it is only the relativistic treatment which exposes the nature of the $\boldsymbol{\sigma}_i$ as spacetime bivectors. This is crucial for determining their algebraic properties as further particles are added. The n-particle spacetime algebra is the geometric algebra of $4n$-dimensional relativistic configuration space. We call this the multiparticle spacetime algebra. A basis is for this is constructed by taking n sets of basis vectors $\{\gamma_\mu^a\}$, where the superscript labels the particle space. These satisfy the orthogonality conditions

$$\gamma_\mu^a\gamma_\nu^b + \gamma_\nu^b\gamma_\mu^a = \begin{cases} 0 & a \neq b \\ 2\eta_{\mu\nu} & a = b \end{cases}, \tag{9.23}$$

which are summarised in the single formula

$$\gamma_\mu^a \cdot \gamma_\nu^b = \delta^{ab}\eta_{\mu\nu}. \tag{9.24}$$

There is nothing uniquely quantum-mechanical in this construction. A system of three classical particles could be described by a set of three trajectories in a single space, or by one path in a nine-dimensional space. The extra dimensions label the properties of each individual particle, and are not to be thought of as existing in anything other than a mathematical sense. One unusual feature concerning relativistic configuration space is that it requires a separate copy of the time dimension for each particle, as well as the three spatial dimensions. This is required in order that the algebra is fully Lorentz-covariant. The presence of multiple time coordinates can complicate the evolution equations in the relativistic theory. Fortunately, the non-relativistic reduction does not suffer from this problem as all of the individual time coordinates are identified with a single absolute time.

As in the single-particle case, the even subalgebra of each copy of the spacetime algebra defines an algebra for relative space. We perform all spacetime splits with the vector γ_0, using a separate copy of this vector in each particle's space. A basis set of relative vectors is then defined by

$$\boldsymbol{\sigma}_i^a = \gamma_i^a\gamma_0^a. \tag{9.25}$$

Again, superscripts label the particle space in which the object appears, and subscripts are retained for the coordinate frame. We do not enforce the summation convention for superscripted indices in this chapter. If we now consider

bivectors from spaces 1 and 2, we find that the basis elements satisfy

$$\boldsymbol{\sigma}_i^1 \boldsymbol{\sigma}_j^2 = \gamma_i^1 \gamma_0^1 \gamma_j^2 \gamma_0^2 = \gamma_i^1 \gamma_j^2 \gamma_0^2 \gamma_0^1 = \gamma_j^2 \gamma_0^2 \gamma_i^1 \gamma_0^1 = \boldsymbol{\sigma}_j^2 \boldsymbol{\sigma}_i^1. \tag{9.26}$$

The basis elements *commute*, rather than anticommute. This solves the problem of how to represent the tensor product in geometric algebra. The geometric product $\boldsymbol{\sigma}_i^a \boldsymbol{\sigma}_j^b$ *is* the tensor product. Since single particle states are constructed out of geometric algebra elements, this gives a natural origin for tensor product states in the multiparticle case. This property only holds because the relative vectors $\boldsymbol{\sigma}_i^a$ are constructed as spacetime bivectors.

The pseudoscalar for each particle space is defined in the obvious way, so that

$$I^a = \gamma_0^a \gamma_1^a \gamma_2^a \gamma_3^a. \tag{9.27}$$

Relative bivectors in each space take the form $I^a \boldsymbol{\sigma}_k^a$. Wherever possible we abbreviate these by dropping the first particle label, so that

$$I\boldsymbol{\sigma}_k^a = I^a \boldsymbol{\sigma}_k^a. \tag{9.28}$$

The reverse operation in the multiparticle spacetime algebra is denoted with a tilde, and reverses the order of products of all relativistic vectors. Wherever possible we use this operation when forming observables. The Hermitian adjoint in each space can be constructed by inserting appropriate factors of γ_0^a.

9.2.1 Non-relativistic states and the correlator

In the single-particle theory, non-relativistic states are constructed from the even subalgebra of the Pauli algebra. A basis for these is provided by the set $\{1, I\boldsymbol{\sigma}_k\}$. When forming multiparticle states we take tensor products of the individual particle states. Since the tensor product and geometric product are equivalent in the multiparticle spacetime algebra, a complete basis is provided by the set

$$\{1,\ I\boldsymbol{\sigma}_k^1,\ I\boldsymbol{\sigma}_k^2,\ I\boldsymbol{\sigma}_j^1\, I\boldsymbol{\sigma}_k^2\}. \tag{9.29}$$

But these basis elements span a 16-dimensional real space, whereas the state space for two spin-1/2 particles is a four-dimensional complex space — only eight real degrees of freedom. What has gone wrong? The answer lies in our treatment of the complex structure. Quantum theory works with a single unit imaginary i, but in our two-particle algebra we now have two bivectors playing the role of i: $I\boldsymbol{\sigma}_3^1$ and $I\boldsymbol{\sigma}_3^2$. Right-multiplication of a state by either of these has to result in the same state in order for the geometric algebra treatment to faithfully mirror standard quantum mechanics. That is, we must have

$$\psi I\boldsymbol{\sigma}_3^1 = \psi I\boldsymbol{\sigma}_3^2. \tag{9.30}$$

Rearranging this, we find that

$$\psi = -\psi I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2 = \psi \tfrac{1}{2}(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2). \tag{9.31}$$

This tells us what we must do. If we define

$$E = \tfrac{1}{2}(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2), \tag{9.32}$$

we find that

$$E^2 = E. \tag{9.33}$$

So right-multiplication by E is a *projection operation*. If we include this factor on the right of all states we halve the number of (real) degrees of freedom from 16 to the expected 8.

The spacetime algebra representation of a direct-product two-particle Pauli spinor is now given by $\psi^1\phi^2 E$, where ψ^1 and ϕ^2 are spinors (even multivectors) in their own spaces. A complete basis for two-particle spin states is provided by

$$\begin{aligned} &|0\rangle|0\rangle \leftrightarrow E, \\ &|0\rangle|1\rangle \leftrightarrow -I\boldsymbol{\sigma}_2^2\, E, \\ &|1\rangle|0\rangle \leftrightarrow -I\boldsymbol{\sigma}_2^1\, E, \\ &|1\rangle|1\rangle \leftrightarrow I\boldsymbol{\sigma}_2^1\, I\boldsymbol{\sigma}_2^2\, E. \end{aligned} \tag{9.34}$$

We further define

$$J = EI\boldsymbol{\sigma}_3^1 = EI\boldsymbol{\sigma}_3^2 = \tfrac{1}{2}(I\boldsymbol{\sigma}_3^1 + I\boldsymbol{\sigma}_3^2), \tag{9.35}$$

so that

$$J^2 = -E. \tag{9.36}$$

Right-sided multiplication by J takes on the role of multiplication by the quantum imaginary i for multiparticle states.

This procedure extends simply to higher multiplicities. All that is required is to find the 'quantum correlator' E_n satisfying

$$E_n I\boldsymbol{\sigma}_3^a = E_n I\boldsymbol{\sigma}_3^b = J_n \quad \text{for all } a,\, b. \tag{9.37}$$

E_n can be constructed by picking out the $a = 1$ space, say, and correlating all the other spaces to this, so that

$$E_n = \prod_{b=2}^{n} \tfrac{1}{2}(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^b). \tag{9.38}$$

The value of E_n is independent of which of the n spaces is singled out and correlated to. The complex structure is defined by

$$J_n = E_n I\boldsymbol{\sigma}_3^a, \tag{9.39}$$

where $I\boldsymbol{\sigma}_3^a$ can be chosen from any of the n spaces. To illustrate this consider

the case of $n = 3$, where

$$\begin{aligned} E_3 &= \tfrac{1}{4}(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2)(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^3) \\ &= \tfrac{1}{4}(1 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^3 - I\boldsymbol{\sigma}_3^2\, I\boldsymbol{\sigma}_3^3) \end{aligned} \tag{9.40}$$

and

$$J_3 = \tfrac{1}{4}(I\boldsymbol{\sigma}_3^1 + I\boldsymbol{\sigma}_3^2 + I\boldsymbol{\sigma}_3^3 - I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2\, I\boldsymbol{\sigma}_3^3). \tag{9.41}$$

Both E_3 and J_3 are symmetric under permutations of their indices.

9.2.2 Operators and observables

All of the operators defined for the single-particle spacetime algebra extend naturally to the multiparticle algebra. In the two-particle case, for example, we have

$$i\hat{\sigma}_k \otimes \hat{\mathsf{I}}|\psi\rangle \;\leftrightarrow\; I\boldsymbol{\sigma}_k^1\psi, \tag{9.42}$$

$$\hat{\mathsf{I}} \otimes i\hat{\sigma}_k|\psi\rangle \;\leftrightarrow\; I\boldsymbol{\sigma}_k^2\psi, \tag{9.43}$$

where $\hat{\mathsf{I}}$ is the 2×2 identity matrix and a factor of E is implicit in the spinor ψ. For the Hermitian operators we form, for example,

$$\hat{\sigma}_k \otimes \hat{\mathsf{I}}|\psi\rangle \;\leftrightarrow\; -I\boldsymbol{\sigma}_k^1\psi J = \boldsymbol{\sigma}_k^1\psi\boldsymbol{\sigma}_3^1. \tag{9.44}$$

This generalises in the obvious way, so that

$$\hat{\mathsf{I}} \otimes \cdots \otimes \hat{\sigma}_k^a \otimes \cdots \otimes \hat{\mathsf{I}}|\psi\rangle \;\leftrightarrow\; \boldsymbol{\sigma}_k^a\psi\boldsymbol{\sigma}_3^a. \tag{9.45}$$

We continue to adopt the j symbol as a convenient shorthand notation for the complex structure, so

$$i|\psi\rangle \;\leftrightarrow\; j\psi = \psi J = \psi I\boldsymbol{\sigma}_3^a. \tag{9.46}$$

The quantum inner product is now

$$\langle\psi|\phi\rangle \leftrightarrow 2^{n-1}\big(\langle\phi E\tilde{\psi}\rangle - \langle\phi J\tilde{\psi}\rangle j\big). \tag{9.47}$$

The factor of E in the real part is not strictly necessary as it is always present in the spinors, but including it does provide a neat symmetry between the real and imaginary parts. The factor of 2^{n-1} guarantees complete consistency with the standard quantum inner product, as it ensures that the state E has unit norm.

Suppose that we now form the observables in the two-particle case. We find that

$$\langle\psi|\,\hat{\sigma}_k \otimes \hat{\mathsf{I}}\,|\psi\rangle \leftrightarrow -2I\boldsymbol{\sigma}_k^1\cdot(\psi J\tilde{\psi}) \tag{9.48}$$

and

$$\langle\psi|\,\hat{\sigma}_j \otimes \hat{\sigma}_k\,|\psi\rangle \leftrightarrow -2(I\boldsymbol{\sigma}_j^1\, I\boldsymbol{\sigma}_k^2)\cdot(\psi E\tilde{\psi}). \tag{9.49}$$

All of the observables one can construct are therefore contained in the multivectors $\psi E\tilde{\psi}$ and $\psi J\tilde{\psi}$. This generalises to arbitrary particle numbers. To see why, we use the fact that any density matrix can be expanded in terms of products of Hermitian operators, as in the two-particle expansion

$$\hat{\rho} = |\psi\rangle\langle\psi| = \frac{1}{4}\big(\hat{\mathsf{I}}\otimes\hat{\mathsf{I}} + a_k\,\hat{\sigma}_k\otimes\hat{\mathsf{I}} + b_k\,\hat{\mathsf{I}}\otimes\hat{\sigma}_k + c_{jk}\,\hat{\sigma}_j\otimes\hat{\sigma}_k\big). \tag{9.50}$$

The various coefficients are found by taking inner products with the appropriate combinations of operators. Each of these corresponds to picking out a term in $\psi E\tilde{\psi}$ or $\psi J\tilde{\psi}$. If an even number of Pauli matrices is involved we pick out a term in $\psi E\tilde{\psi}$, and an odd number picks out a term in $\psi J\tilde{\psi}$. In general, $\psi E\tilde{\psi}$ contains terms of grades 0, 4, ..., and $\psi J\tilde{\psi}$ contains terms of grade 2, 6, ... These account for all the coefficients in the density matrix, and hence for all the observables that can be formed from ψ.

An advantage of working directly with the observables $\psi E\tilde{\psi}$ and $\psi J\tilde{\psi}$ is that the partial trace operation has a simple interpretation. If we want to form the partial trace over the ath particle, we simply remove all terms from the observables with a contribution in the ath particle space. No actual trace operation is required. Furthermore, this operation of discarding information is precisely the correct physical picture for the partial trace operation — we are discarding the (often unknown) information associated with a particle in one or more spaces. A minor complication in this approach is that $\psi J\tilde{\psi}$ gives rise to anti-Hermitian terms, whereas the density matrix is Hermitian. One way round this is to correlate all of the pseudoscalars together and then dualise all bivectors back to vectors. This is the approach favoured by Havel and coworkers in their work on NMR spectroscopy. Alternatively, one can simply ignore this feature and work directly with the observables $\psi E\tilde{\psi}$ and $\psi J\tilde{\psi}$. When presented with a general density matrix one often needs to pull it apart into sums of terms like this anyway (the product operator expansion), so it makes sense to work directly with the multivector observables when they are available.

9.3 Systems of two particles

Many of the preceding ideas are most simply illustrated for the case of a system of two particles. For these, the Schmidt decomposition of equation (9.13) provides a useful formulation for a general state. The geometric algebra version of this is rather more compact, however, as we now establish. First, we define the spinor

$$\psi(\theta,\phi) = \mathrm{e}^{-\phi I\boldsymbol{\sigma}_3/2}\,\mathrm{e}^{-\theta I\boldsymbol{\sigma}_2/2}. \tag{9.51}$$

We also need a representation of the state orthogonal to this, which is

$$\begin{pmatrix}\sin(\theta/2)\mathrm{e}^{-i\phi/2}\\ -\cos(\theta/2)\mathrm{e}^{i\phi/2}\end{pmatrix} \leftrightarrow \psi(\theta,\phi)I\boldsymbol{\sigma}_2. \tag{9.52}$$

Now we are in a position to construct the multiparticle spacetime algebra version of the Schmidt decomposition. We replace equation (9.13) with

$$\begin{aligned}\psi =&\rho^{1/2}\Big(\cos(\alpha/2)\psi^1(\theta_1,\phi_1)\psi^2(\theta_2,\phi_2)\mathrm{e}^{J\tau/2}\\ &+\sin(\alpha/2)\psi^1(\theta_1,\phi_1)\psi^2(\theta_2,\phi_2)I\boldsymbol{\sigma}_2^1\, I\boldsymbol{\sigma}_2^2\mathrm{e}^{-J\tau/2}\Big)\mathrm{e}^{J\chi}E\\ =&\rho^{1/2}\psi^1(\theta_1,\phi_1)\psi^2(\theta_2,\phi_2)\mathrm{e}^{J\tau/2}\left(\cos(\alpha/2)+\sin(\alpha/2)I\boldsymbol{\sigma}_2^1\, I\boldsymbol{\sigma}_2^2\right)\mathrm{e}^{J\chi}E. \quad (9.53)\end{aligned}$$

We now define the individual rotors

$$R=\psi(\theta_1,\phi_1)\mathrm{e}^{I\boldsymbol{\sigma}_3\tau/4},\quad S=\psi(\theta_2,\phi_2)\mathrm{e}^{I\boldsymbol{\sigma}_3\tau/4}, \quad (9.54)$$

so that the wavefunction ψ simplifies to

$$\psi=\rho^{1/2}R^1S^2\big(\cos(\alpha/2)+\sin(\alpha/2)I\boldsymbol{\sigma}_2^1I\boldsymbol{\sigma}_2^2\big)\mathrm{e}^{J\chi}E. \quad (9.55)$$

This gives a compact, general form for an arbitrary two-particle state. The degrees of freedom are held in an overall magnitude and phase, two separate rotors in the individual particle spaces, and a single entanglement angle α. In total this gives nine degrees of freedom, so one must be redundant. This redundancy lies in the single-particle rotors. If we take

$$R\mapsto R\mathrm{e}^{I\boldsymbol{\sigma}_3\beta},\quad S\mapsto S\mathrm{e}^{-I\boldsymbol{\sigma}_3\beta} \quad (9.56)$$

then the overall wavefunction ψ is unchanged. In practice this redundancy is not a problem, and the form of equation (9.55) turns out to be extremely useful.

9.3.1 Observables for two-particle states

The individual rotors R^1 and S^2 generate rotations in their own spaces. These are equivalent to local unitary transformations. The novel features associated with the observables for a two-particle system arise from the entanglement angle α. To study this we first form the bivector observable $\psi J\tilde{\psi}$:

$$\begin{aligned}\psi J\tilde{\psi}=&R^1S^2\big(\cos(\alpha/2)+\sin(\alpha/2)I\boldsymbol{\sigma}_2^1\, I\boldsymbol{\sigma}_2^2\big)J\big(\cos(\alpha/2)+\sin(\alpha/2)I\boldsymbol{\sigma}_2^1\, I\boldsymbol{\sigma}_2^2\big)\tilde{R}^1\tilde{S}^2\\ =&\tfrac{1}{2}R^1S^2\big(\cos^2(\alpha/2)-\sin^2(\alpha/2)\big)(I\boldsymbol{\sigma}_3^1+I\boldsymbol{\sigma}_3^2)\tilde{R}^1\tilde{S}^2\\ =&\tfrac{1}{2}\cos(\alpha)\big((RI\boldsymbol{\sigma}_3\tilde{R})^1+(SI\boldsymbol{\sigma}_3\tilde{S})^2\big), \quad (9.57)\end{aligned}$$

where we have assumed that $\rho=1$. This result extends the definition of the spin bivector to multiparticle systems. One can immediately see that the lengths of the bivectors are no longer fixed, but instead depend on the entanglement. Only in the case of zero entanglement are the spin bivectors unit length.

The remaining observables are contained in

$$\psi E\tilde{\psi}=\tfrac{1}{2}R^1S^2\Big(1-I\boldsymbol{\sigma}_3^1I\boldsymbol{\sigma}_3^2+\sin(\alpha)(I\boldsymbol{\sigma}_2^1I\boldsymbol{\sigma}_2^2-I\boldsymbol{\sigma}_1^1I\boldsymbol{\sigma}_1^2)\Big)\tilde{R}^1\tilde{S}^2. \quad (9.58)$$

To make this result clearer we introduce the notation

$$A_k = RI\boldsymbol{\sigma}_k\tilde{R}, \quad B_k = SI\boldsymbol{\sigma}_k\tilde{S}, \tag{9.59}$$

so that

$$2\psi E\tilde{\psi} = 1 - A_3^1 B_3^2 + \sin(\alpha)(A_2^1 B_2^2 - A_1^1 B_1^2). \tag{9.60}$$

The scalar part confirms that the state is normalised correctly. The 4-vector part contains an interesting new term, which goes as $A_2^1 B_2^2 - A_1^1 B_1^2$. None of the individual A_1, A_2, B_1, or B_2 bivectors is accessible to measurement in the single-particle case as they are not phase-invariant. But in the two-particle case these terms do start to influence the observables. This is one of essential differences between classical and quantum models of spin.

9.3.2 Density matrices and probabilities

Now that we have all of the observables, we have also found all of the terms in the density matrix. Of particular interest are the results of partial traces, where we discard the information associated with one of the particles. If we throw out all of the information about the second particle, for example, what remains is the single-particle density matrix

$$\hat{\rho} = \tfrac{1}{2}(1 + \boldsymbol{p}), \tag{9.61}$$

where the polarisation vector is given by

$$\boldsymbol{p} = \cos(\alpha) R\boldsymbol{\sigma}_3\tilde{R}. \tag{9.62}$$

This vector no longer has unit length, so the density matrix is that of a mixed state. Entanglement with a second particle has led to a loss of coherence of the first particle. This process, by which entanglement produces decoherence, is central to attempts to explain the emergence of classical physics from quantum theory.

For two particles we see that there is a symmetry between the degree of entanglement. If we perform a partial trace over particle 1, the polarisation vector for the second particle also has its length reduced by a factor of $\cos(\alpha)$. More generally the picture is less simple, and much work remains in understanding entanglement beyond the bipartite case.

A further application of the preceding is to calculate the overlap probability for the inner product of two states. Given two normalised states we have

$$P(\psi, \phi) = |\langle\psi|\phi\rangle|^2 = \mathrm{tr}(\hat{\rho}_\psi \hat{\rho}_\phi). \tag{9.63}$$

The degrees of freedom in the density matrices are contained in $\psi E\tilde{\psi}$ and $\psi J\tilde{\psi}$, with equivalent expressions for ϕ. When forming the inner product between two

density matrices, the only terms that can arise are inner products between these observables. A little work confirms that we can write, in the n-particle case,

$$P(\psi,\phi) = 2^{n-2}\langle(\psi E\tilde{\psi})(\phi E\tilde{\phi})\rangle - 2^{n-2}\langle(\psi J\tilde{\psi})(\phi J\tilde{\phi})\rangle. \tag{9.64}$$

Expressions like this are unique to the geometric algebra approach. The expression confirms that once one has found the two multivector observables for a state, one has all of the available information to hand.

As an example, suppose that we are presented with two *separable* states, ψ and ϕ. For separable states we know that the observables take the forms

$$2\psi J\tilde{\psi} = A^1 + B^2, \qquad 2\psi E\tilde{\psi} = 1 - A^1B^2 \tag{9.65}$$

and

$$2\phi J\tilde{\phi} = C^1 + D^2, \qquad 2\phi E\tilde{\phi} = 1 - C^1D^2, \tag{9.66}$$

where each of the A^1, B^2, C^1 and D^2 are unit bivectors. We can now write

$$\begin{aligned} P(\psi,\phi) &= \tfrac{1}{4}\langle(1 - A^1B^2)(1 - C^1D^2) - (A^1 + B^2)(C^1 + D^2)\rangle \\ &= \tfrac{1}{4}(1 + A\cdot C\, B\cdot D - A\cdot C - B\cdot D) \\ &= \tfrac{1}{2}(1 - A\cdot C)\,\tfrac{1}{2}(1 - B\cdot D). \end{aligned} \tag{9.67}$$

This confirms the probability is the product of the separate single-particle probabilities. If one of the states is entangled this result no longer holds, as we see in the following section.

9.3.3 The singlet state

As a further example of entanglement we now study some of the properties of the non-relativistic spin singlet state. This is

$$|\varepsilon\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|1\rangle - |1\rangle|0\rangle\big). \tag{9.68}$$

This is represented in the two-particle spacetime algebra by the multivector

$$\varepsilon = \frac{1}{\sqrt{2}}\big(I\boldsymbol{\sigma}^1_2 - I\boldsymbol{\sigma}^2_2\big)E. \tag{9.69}$$

The properties of ε are more easily seen by writing

$$\varepsilon = \tfrac{1}{2}(1 + I\boldsymbol{\sigma}^1_2\, I\boldsymbol{\sigma}^2_2)\tfrac{1}{2}(1 + I\boldsymbol{\sigma}^1_3\, I\boldsymbol{\sigma}^2_3)\sqrt{2}\, I\boldsymbol{\sigma}^1_2, \tag{9.70}$$

which shows how ε contains the commuting idempotents $(1 + I\boldsymbol{\sigma}^1_2\, I\boldsymbol{\sigma}^2_2)/2$ and $(1 + I\boldsymbol{\sigma}^1_3\, I\boldsymbol{\sigma}^2_3)/2$. Identifying these idempotents tells us immediately that

$$I\boldsymbol{\sigma}^1_2\varepsilon = \tfrac{1}{2}(I\boldsymbol{\sigma}^1_2 - I\boldsymbol{\sigma}^2_2)\tfrac{1}{2}(1 + I\boldsymbol{\sigma}^1_3\, I\boldsymbol{\sigma}^2_3)\sqrt{2}I\boldsymbol{\sigma}^1_2 = -I\boldsymbol{\sigma}^2_2\varepsilon \tag{9.71}$$

and

$$I\boldsymbol{\sigma}^1_3\varepsilon = -I\boldsymbol{\sigma}^2_3\varepsilon. \tag{9.72}$$

If follows that

$$I\boldsymbol{\sigma}^1_1\varepsilon = I\boldsymbol{\sigma}^1_3\, I\boldsymbol{\sigma}^1_2\varepsilon = -I\boldsymbol{\sigma}^2_2\, I\boldsymbol{\sigma}^1_3\varepsilon = I\boldsymbol{\sigma}^2_2\, I\boldsymbol{\sigma}^2_3\varepsilon = -I\boldsymbol{\sigma}^2_1\varepsilon. \tag{9.73}$$

Combining these results, if M^1 is an arbitrary even element in the Pauli algebra ($M^1 = M_0 + M_k I\boldsymbol{\sigma}^1_k$), ε satisfies

$$M^1\varepsilon = \tilde{M}^2\varepsilon. \tag{9.74}$$

Here M^1 and M^2 denote the same multivector, but expressed in space 1 or space 2.

Equation (9.74) provides a novel demonstration of the rotational invariance of ε. Under a joint rotation in two-particle space, a spinor ψ transforms to $R^1R^2\psi$, where R^1 and R^2 are copies of the same rotor but acting in the two different spaces. From equation (9.74) it follows that, under such a rotation, ε transforms as

$$\varepsilon \mapsto R^1R^2\varepsilon = R^1\tilde{R}^1\varepsilon = \varepsilon, \tag{9.75}$$

so that ε is a genuine two-particle rotational scalar.

If we now form the observables from ε we find that

$$2\varepsilon E\tilde{\varepsilon} = 1 + \sum_{k=1}^{3} I\boldsymbol{\sigma}^1_k\, I\boldsymbol{\sigma}^2_k \tag{9.76}$$

and

$$\varepsilon J\tilde{\varepsilon} = 0. \tag{9.77}$$

The latter has to hold, as there are no rotationally-invariant bivector observables. Equation (9.76) identifies a new two-particle invariant, which we can write as

$$\sum_{k=1}^{3} I\boldsymbol{\sigma}^1_k\, I\boldsymbol{\sigma}^2_k = 2\varepsilon\tilde{\varepsilon} - 1. \tag{9.78}$$

This is invariant under joint rotations in the two particles spaces. This multivector equation contains the essence of the matrix result

$$\sum_{k=1}^{3} \hat{\sigma}^a_{k\,a'}\, \hat{\sigma}^b_{k\,b'} = 2\delta^a_{b'}\, \delta^b_{a'} - \delta^a_{a'}\, \delta^b_{b'}, \tag{9.79}$$

where a, b, a', b' label the matrix components. In standard quantum mechanics this invariant would be thought of as arising from the 'inner product' of the spin vectors $\hat{\sigma}^1_i$ and $\hat{\sigma}^2_i$. Here, we have seen that the invariant arises in a completely different way, as a component of the multivector $\varepsilon\tilde{\varepsilon}$.

The fact that $\varepsilon J\tilde{\varepsilon} = 0$ confirms that the reduced density matrix for either particle space is simply one-half of the identity matrix, as established in equation (9.22). It follows that all directions are equally likely. If we align our measuring apparatus along some given axis and measure the state of particle 1,

then both up and down have equal probabilities of 1/2. Suppose now that we construct a joint measurement on the singlet state. We can model this as the overlap probability between ψ and the separable state

$$\phi = R^1 S^2 E. \tag{9.80}$$

Denoting the spin directions by

$$RI\boldsymbol{\sigma}_3\tilde{R} = P, \quad SI\boldsymbol{\sigma}_3\tilde{S} = Q, \tag{9.81}$$

we find that, from equation (9.64),

$$\begin{aligned} P(\psi,\phi) &= \langle \tfrac{1}{2}(1 - P^1 Q^2)\tfrac{1}{2}(1 + I\boldsymbol{\sigma}_k^1\, I\boldsymbol{\sigma}_k^2)\rangle \\ &= \tfrac{1}{4}(1 - P\cdot(I\boldsymbol{\sigma}_k)\, Q\cdot(I\boldsymbol{\sigma}_k)) \\ &= \tfrac{1}{4}\big(1 - \cos(\theta)\big) \end{aligned} \tag{9.82}$$

where θ is the angle between the spin bivectors P and Q. So, for example, the probability that both measurements result in the particles having the same spin ($\theta = 0$) is zero, as expected. Similarly, if the measuring devices are aligned, the probability that particle 1 is up and particle 2 is down is 1/2, whereas if there was no entanglement present the probability would be the product of the separate single-particle measurements (resulting in 1/4).

Some consequences of equation (9.82) run counter to our intuitions about locality and causality. In particular, it is impossible to reproduce the statistics of equation (9.82) if we assume that the individual particles both know which spin state they are in prior to measurement. These contradictions are embodied in the famous *Bell inequalities*. The behaviour of entangled states has now been tested experimentally, and the results confirm all of the predictions of quantum mechanics. The results are unchanged even if the measurements are performed in such a way that the particles cannot be in causal contact. This does not provide any conflict with special relativity, as entangled states cannot be used to exchange classical information at faster than the speed of light. The reason is that the presence of entanglement can only be inferred when the separate measurements on the two subsystems are compared. Without knowing which measurements observer 1 is performing, observer 2 cannot extract any useful classical information from an entangled state.

For many years the properties of entangled states were explored largely as a theoretical investigation into the nature of quantum theory. Now, however, physicists are starting to view quantum entanglement as a resource that can be controlled in the laboratory. To date our control of entangled states is limited, but it is improving rapidly, and many predict that before long we will see the first viable quantum computers able to exploit this new resource.

9.4 Relativistic states and operators

The ideas developed for the multiparticle Pauli algebra extend immediately to the relativistic domain. A single-particle relativistic state is described by an arbitrary even element of the full spacetime algebra. Accordingly, a two-particle state is constructed from the tensor product of two such states. This results is a space of of $8 \times 8 = 64$ real dimensions. Post-multiplying the direct-product space by the quantum correlator E reduces to 32 real dimensions, which are equivalent to the 16 complex dimensions employed in standard two-particle relativistic quantum theory. All the single-particle operators and observables discussed in section 8.2 extend in fairly obvious ways.

To begin, the individual matrix operators have the equivalent action

$$\begin{aligned} \hat{\gamma}_\mu \otimes \hat{\mathsf{I}}|\psi\rangle &\leftrightarrow \gamma^1_\mu \psi \gamma^1_0, \\ \hat{\mathsf{I}} \otimes \hat{\gamma}_\mu|\psi\rangle &\leftrightarrow \gamma^2_\mu \psi \gamma^2_0, \end{aligned} \tag{9.83}$$

where $\hat{\mathsf{I}}$ denotes the 4×4 identity matrix. The multiparticle spacetime algebra operators *commute*, as they must in order to represent the tensor product. The result of the action of $\gamma^1_\mu \psi \gamma^1_0$, for example, does not take us outside the two-particle state space, since the factor of γ^1_0 on the right-hand side commutes with the correlator E. The remaining matrix operators are easily constructed now, for example

$$\hat{\gamma}_\mu \hat{\gamma}_\nu \otimes \hat{\mathsf{I}}|\psi\rangle \leftrightarrow \gamma^1_\mu \gamma^1_\nu \psi. \tag{9.84}$$

The role of multiplication by the unit imaginary i is still played by right-multiplication by J, and the individual helicity projection operators become

$$\hat{\gamma}_5 \otimes \hat{\mathsf{I}}|\psi\rangle \leftrightarrow -I^1 \psi J = \psi \boldsymbol{\sigma}^1_3. \tag{9.85}$$

Relativistic observables are also constructed in a similar manner to the single-particle case. We form geometric products $\psi \Sigma \tilde{\psi}$, where Σ is any combination of γ_0 and γ_3 from either space. The result is then guaranteed to be Lorentz-covariant and phase-invariant. The first observable to consider is the multivector

$$\psi\tilde{\psi} = \psi E \tilde{\psi} = \langle \psi E \tilde{\psi} \rangle_{0,8} + \langle \psi E \tilde{\psi} \rangle_4. \tag{9.86}$$

The grade-0 and grade-8 terms are the two-particle generalisation of the scalar + pseudoscalar combination $\psi\tilde{\psi} = \rho \exp(i\beta)$ found at the single-particle level. The 4-vector part generalises the entanglement terms found in the non-relativistic case. This allows for a relativistic definition of entanglement, which is important for a detailed study of the relationship between locality and entanglement.

Next, we form two-particle current and spin vectors:

$$\mathcal{J} = \langle \psi(\gamma^1_0 + \gamma^2_0)\tilde{\psi} \rangle_1, \tag{9.87}$$

$$s = \langle \psi(\gamma^1_3 + \gamma^2_3)\tilde{\psi} \rangle_1. \tag{9.88}$$

(The calligraphic symbol $\mathcal{J}$ is used to avoid confusion with the correlated bivector J.) The full observables will contain grade-1 and grade-5 terms. For direct-product states the latter are seen to arise from the presence of a β factor in either of the single-particle states. Finally, we can also define the spin bivector S by

$$S = \langle \psi J \tilde{\psi} \rangle_2. \tag{9.89}$$

These expressions show how easy it is to generalise the single-particle formulae to the multiparticle case.

9.4.1 The relativistic singlet state

In the non-relativistic theory the spin singlet state has a special significance, both in being maximally entangled, and in its invariance under joint rotations in the two-particle space. An interesting question is whether we can construct a relativistic analogue that plays the role of a Lorentz singlet. Recalling the definition of ε (9.69), the property that ensured ε was a singlet state was that

$$I\boldsymbol{\sigma}^1_k \varepsilon = -I\boldsymbol{\sigma}^2_k \varepsilon, \qquad k = 1, \ldots, 3. \tag{9.90}$$

In addition to (9.90) a relativistic singlet state, which we will denote as η, must satisfy

$$\boldsymbol{\sigma}^1_k \eta = -\boldsymbol{\sigma}^2_k \eta, \qquad k = 1, \ldots, 3. \tag{9.91}$$

It follows that η satisfies

$$I^1 \eta = \boldsymbol{\sigma}^1_1 \boldsymbol{\sigma}^1_2 \boldsymbol{\sigma}^1_3 \eta = -\boldsymbol{\sigma}^2_3 \boldsymbol{\sigma}^2_2 \boldsymbol{\sigma}^2_1 \eta = I^2 \eta. \tag{9.92}$$

For this to hold, η must contain a factor of $(1 - I^1 I^2)$. We can therefore construct a Lorentz single state by multiplying ε by $(1 - I^1 I^2)$, and we define

$$\eta = (I\boldsymbol{\sigma}^1_2 - I\boldsymbol{\sigma}^2_2)\tfrac{1}{2}(1 - I\boldsymbol{\sigma}^1_3\, I\boldsymbol{\sigma}^2_3)\tfrac{1}{2}(1 - I^1 I^2). \tag{9.93}$$

This is normalised so that $2\langle \eta E \tilde{\eta} \rangle = 1$. The properties of η can be summarised as

$$M^1 \eta = \tilde{M}^2 \eta, \tag{9.94}$$

where M is an even multivector in either the particle-1 or particle-2 spacetime algebra. The proof that η is a relativistic invariant now reduces to the simple identity

$$R^1 R^2 \eta = R^1 \tilde{R}^1 \eta = \eta, \tag{9.95}$$

where R is a single-particle relativistic rotor.

Equation (9.94) can be seen as originating from a more primitive relation

between vectors in the separate spaces. Using the result that $\gamma_0^1\gamma_0^2$ commutes with η, we can derive

$$\begin{aligned}\gamma_\mu^1\eta\gamma_0^1 &= \gamma_\mu^1\gamma_0^1\gamma_0^2\eta\gamma_0^2\gamma_0^1\gamma_0^1 \\ &= \gamma_0^2(\gamma_\mu\gamma_0)^1\eta\gamma_0^2 \\ &= \gamma_\mu^2\eta\gamma_0^2. \end{aligned} \tag{9.96}$$

For an arbitrary vector a we can now write

$$a^1\eta\gamma_0^1 = a^2\eta\gamma_0^2. \tag{9.97}$$

Equation (9.94) follows immediately from equation (9.97) by writing

$$\begin{aligned}a^1b^1\eta &= a^1b^2\eta\gamma_0^2\gamma_0^1 \\ &= b^2a^2\eta\gamma_0^2\gamma_0^2 \\ &= b^2a^2\eta. \end{aligned} \tag{9.98}$$

Equation (9.97) can therefore be viewed as the fundamental property of the relativistic invariant η.

The invariant η can be used to construct a series of observables that are also invariant under coupled rotations in the two spaces. The first is

$$2\eta E\tilde{\eta} = (1 - I^1I^2) - (\boldsymbol{\sigma}_k^1\,\boldsymbol{\sigma}_k^2 - I\boldsymbol{\sigma}_k^1\,I\boldsymbol{\sigma}_k^2). \tag{9.99}$$

The scalar and pseudoscalar (grade-8) terms are clearly invariants, and the 4-vector term, $(\boldsymbol{\sigma}_k^1\,\boldsymbol{\sigma}_k^2 - I\boldsymbol{\sigma}_k^1\,I\boldsymbol{\sigma}_k^2)$, is a Lorentz invariant because it is a contraction over a complete bivector basis in the two spaces. Next we consider the multivector

$$\begin{aligned}2\eta\gamma_0^1\gamma_0^2\tilde{\eta} &= \gamma_0^1\gamma_0^2 - I^1I^2\gamma_k^1\gamma_k^2 - I^1I^2\gamma_0^1\gamma_0^2 - \gamma_k^1\gamma_k^2) \\ &= (\gamma_0^1\gamma_0^2 - \gamma_k^1\gamma_k^2)(1 - I^1I^2). \end{aligned} \tag{9.100}$$

The essential invariant here is the bivector

$$K = \gamma_\mu^1\wedge\gamma^{\mu 2}, \tag{9.101}$$

and the invariants from (9.100) are simply K and KI^1I^2. The bivector K takes the form of a 'doubling' bivector, which will be encountered again in section 11.4.

From the definition of K in equation (9.101), we find that

$$\begin{aligned}K\wedge K &= -2\gamma_0^1\gamma_0^2\gamma_k^1\gamma_k^2 + (\gamma_k^1\gamma_k^2)\wedge(\gamma_j^1\gamma_j^2) \\ &= 2(\boldsymbol{\sigma}_k^1\,\boldsymbol{\sigma}_k^2 - I\boldsymbol{\sigma}_k^1\,I\boldsymbol{\sigma}_k^2), \end{aligned} \tag{9.102}$$

which recovers the grade-4 invariant found in equation (9.99). The full set of two-particle invariants constructed from K are summarised in table 9.1. These invariants are regularly employed in constructing interaction terms in multiparticle wave equations.

Invariant	Type of interaction	Grade
1	Scalar	0
K	Vector	2
$K \wedge K$	Bivector	4
$I^1 I^2 K$	Pseudovector	6
$I^1 I^2$	Pseudoscalar	8

Table 9.1 *Relativistic invariants in the two-particle algebra.*

9.4.2 Multiparticle wave equations

The question of how to construct a valid, relativistic, multiparticle wave equation has troubled physicists almost from the moment Dirac proposed his equation. The question is far from settled, and the current preferred option is to ignore the question where possible and instead work within the framework of perturbative quantum field theory. This approach runs into difficulties when analysing bound states, however, and for these problems the need for a suitable wave equation is particularly acute. The main candidate for a relativistic two-particle system is the Bethe–Salpeter equation. Written in the multiparticle spacetime algebra, this equation is

$$(j\hat{\nabla}_r^1 - m_1)(j\hat{\nabla}_s^2 - m_2)\psi(r,s) = \mathcal{I}(r,s)\psi(r,s) \tag{9.103}$$

where $\mathcal{I}(r,s)$ is an integral operator representing the interparticle interaction, and ∇_r^1 and ∇_s^2 denote vector derivatives with respect to r^1 and s^2 respectively. The combined vector

$$x = r^1 + s^2 = r^\mu \gamma_\mu^1 + s^\mu \gamma_\mu^2 \tag{9.104}$$

is the full position vector in eight-dimensional configuration space.

One slightly unsatisfactory feature of equation (9.103) is that it is not first-order. This has led researchers to propose a number of alternative equations, typically with the aim of providing a more detailed analysis of two-body bound state systems such as the hydrogen atom, or positronium. One such equation is

$$\left(\nabla_r^1 \psi \gamma_0^1 + \nabla_s^2 \psi \gamma_0^2\right) J = (m_1 + m_2)\psi. \tag{9.105}$$

As well as being first order, this equation also has the required property that it is satisfied by direct products of single-particle solutions. But a problem is that any distinction between the particle masses has been lost, since only the total mass enters. A second candidate equation, which does keep the masses distinct, is

$$\left(\frac{\nabla_r^1}{m_1} + \frac{\nabla_s^2}{m_2}\right)\psi(x)J = \psi(x)(\gamma_0^1 + \gamma_0^2). \tag{9.106}$$

This equation has a number of attractive features, not least of which is that the mass enters in a manner that is highly suggestive of gravitational interactions. A potential weakness of this equation is that the state space can no longer be restricted to sums of direct products of individual states. Instead we have to widen the state space to include the entire (correlated) even subalgebra of the two-particle spacetime algebra. This doubles the number of degrees of freedom, and it is not clear that this doubling can be physical.

Practically all candidate two-particle wave equations have difficulties in performing a separation into centre-of-mass and relative coordinates. This is symptomatic of the fact that the centre of mass cannot be defined sensibly even in classical relativistic dynamics. Usually some approximation scheme has to be employed to avoid this problem, even when looking for bound state solutions. While the question of finding a suitable wave equation remains an interesting challenge, one should be wary of the fact that the mass term in the Dirac equation is essentially a remainder from a more complicated interaction with the Higgs boson. The electroweak theory immediately forces us to consider particle doublets, and it could be that one has to consider multiparticle extensions of these in order to arrive at a satisfactory theory.

9.4.3 The Pauli principle

In quantum theory, indistinguishable particles must obey either Fermi–Dirac or Bose–Einstein statistics. For fermions this requirement results in the Pauli exclusion principle that no two particles can occupy a state in which their properties are identical. The Pauli principle is usually enforced in one of two ways in relativistic quantum theory. At the level of multiparticle wave mechanics, antisymmetrisation is enforced by using a Slater determinant representation of a state. At the level of quantum field theory, however, antisymmetrisation is a consequence of the anticommutation of the creation and annihilation operators for fermions. Here we are interested in the former approach, and look to achieve the antisymmetrisation in a simple geometrical manner.

We start by introducing the grade-4 multivector

$$I_P = \Gamma_0\Gamma_1\Gamma_2\Gamma_3, \tag{9.107}$$

where

$$\Gamma_\mu = \frac{1}{\sqrt{2}}\left(\gamma_\mu^1 + \gamma_\mu^2\right). \tag{9.108}$$

It is a simple matter to verify that I_P has the properties

$$I_P^2 = -1 \tag{9.109}$$

and

$$I_P\gamma_\mu^1 I_P = \gamma_\mu^2, \quad I_P\gamma_\mu^2 I_P = \gamma_\mu^1. \tag{9.110}$$

It follows that I_P functions as a geometrical version of the particle exchange operator. In particular, acting on the eight-dimensional position vector $x = r^1 + s^2$ we find that

$$I_P x I_P = r^2 + s^1 \tag{9.111}$$

where

$$r^2 = \gamma_\mu^2 r^\mu, \quad s^1 = \gamma_\mu^1 s^\mu. \tag{9.112}$$

So I_P can be used to interchange the coordinates of particles 1 and 2. Next we must confirm that I_P is independent of the choice of initial frame. Suppose that instead we had started with the rotated frame $\{R\gamma_\mu\tilde{R}\}$, with

$$\Gamma'_\mu = \frac{1}{\sqrt{2}}\big(R^1\gamma_\mu^1\tilde{R}^1 + R^2\gamma_\mu^2\tilde{R}^2\big) = R^1R^2\Gamma_\mu\tilde{R}^2\tilde{R}^1. \tag{9.113}$$

The new Γ'_μ vectors give rise to the rotated 4-vector

$$I'_P = R^1R^2I_P\tilde{R}^2\tilde{R}^1. \tag{9.114}$$

But, acting on a bivector in particle space 1, we find that

$$I_P a^1\wedge b^1 I_P = -(I_P a^1 I_P)\wedge(I_P b^1 I_P) = -a^2\wedge b^2, \tag{9.115}$$

and the same is true of an arbitrary even element in either space. More generally, the operation $M \mapsto I_P M I_P$ applied to an even element in one of the particle spaces flips it to the other particle space and changes sign, while applied to an odd element it just flips the particle space. It follows that

$$I_P\tilde{R}^2\tilde{R}^1 = \tilde{R}^1 I_P\tilde{R}^1 = \tilde{R}^1\tilde{R}^2 I_P, \tag{9.116}$$

and substituting this into (9.114) we find that $I'_P = I_P$. It follows that I_P is independent of the chosen orthonormal frame, as required.

We can now use the 4-vector I_P to encode the Pauli exchange principle geometrically. Let $\psi(x)$ be a wavefunction for two electrons. The state

$$\psi(x)' = -I_P\psi(I_P x I_P)I_P, \tag{9.117}$$

then swaps the position dependence, and interchanges the space of the multivector components of ψ. The antisymmetrised state is therefore

$$\psi_-(x) = \psi(x) + I_P\psi(I_P x I_P)I_P. \tag{9.118}$$

For n-particle systems the extension is straightforward, as we require that the wavefunction is invariant under the interchange enforced by the I_Ps constructed from each pair of particles.

For a single Dirac particle the probability current $J = \psi\gamma_0\tilde{\psi}$ has zero divergence, and can therefore be used to define streamlines. These are valuable for understanding a range of phenomena, such as wavepacket tunnelling and spin

measurement. We now illustrate how these ideas extend to the multiparticle domain. The two-particle current is

$$\mathcal{J} = \langle \psi(\gamma_0^1 + \gamma_0^2)\tilde{\psi} \rangle_1, \tag{9.119}$$

as defined in equation (9.87). The vector $\mathcal{J}$ has components in both particle-1 and particle-2 spaces, which we write as

$$\mathcal{J} = \mathcal{J}_1^1 + \mathcal{J}_2^2. \tag{9.120}$$

For sums of separable solutions to the single-particle equations, the individual currents are both conserved:

$$\nabla^1 \cdot \mathcal{J}_1^1 = \nabla^2 \cdot \mathcal{J}_2^2 = 0. \tag{9.121}$$

It follows that the full current $\mathcal{J}$ is conserved in 8-dimensional space, so its streamlines never cross there. The streamlines of the individual particles, however, are obtained by integrating $\mathcal{J}_1$ and $\mathcal{J}_2$ in a single spacetime, and these can cross if plotted in the same space. For example, suppose that the wavefunction is just

$$\psi = \phi^1(r^1)\chi^2(s^2)E, \tag{9.122}$$

where ϕ and χ are Gaussian wavepackets moving in opposite directions. Since the distinguishable case is assumed, no Pauli antisymmetrisation is used. One can easily confirm that for this case the streamlines and the wavepackets simply pass straight through each other.

But suppose now that we assume indistinguishability, and apply the Pauli symmetrisation procedure to the wavefunction of equation (9.122). We arrive at the state

$$\psi = \left(\phi^1(r^1)\chi^2(s^2) - \chi^1(r^2)\phi^2(s^1)\right) E, \tag{9.123}$$

from which we form $\mathcal{J}_1$ and $\mathcal{J}_2$, as before. Figure 9.1 shows the streamlines that result from these currents. In the left-hand plot both particles are in the same spin state. The corrugated appearance of the lines near the origin is the result of the streamlines having to pass through a region of highly oscillatory destructive interference, since the probability of both particles occupying the same position (the origin) with the same spin state is zero. The right-hand plot is for two particles in different spin states. Again, the streamlines are seen to repel. The reason for this can be found in the symmetry properties of the two-particle current. Given that the wavefunction ψ has been antisymmetrised according to equation (9.118), the current must satisfy

$$I_P \mathcal{J}(I_P x I_P) I_P = \mathcal{J}(x). \tag{9.124}$$

It follows that at the same spacetime position, encoded by $I_P x I_P = x$ in the two-particle algebra, the two currents $\mathcal{J}_1$ and $\mathcal{J}_2$ are equal. Hence, if two streamlines

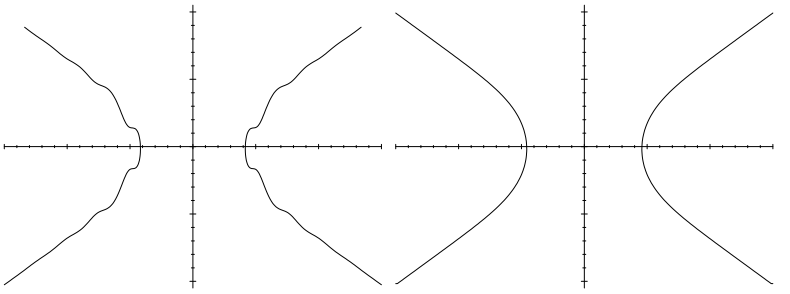

Figure 9.1 *Streamlines for an antisymmetrised two-particle wavefunction.* The wavefunction is $\psi = \big(\phi^1(r^1)\chi^2(s^2) - \chi^1(r^2)\phi^2(s^1)\big)E$. The individual wavepackets pass through each other, but the streamlines from separate particles do not cross. The left-hand figure has both particles with spins aligned in the $+z$ direction. The right-hand figure shows particles with opposite spins, with ϕ in the $+z$ direction, and χ in the $-z$ direction.

ever met, they could never separate again. For the simulations presented here, the symmetry of the set-up implies that the spatial currents at the origin are both zero. As the particles approach the origin, they are forced to slow up. The delay means that they are then swept back in the direction they have just come from by the wavepacket travelling through from the other side. This repulsion has its origin in indistinguishability, and the spin of the states exerts only a marginal effect.

9.5 Two-spinor calculus

The ideas introduced in this chapter can be employed to construct a geometric algebra version of the two-spinor calculus developed by Penrose & Rindler (1984). The building blocks of their approach are two-component complex spinors, denoted κ^A and $\bar{\omega}^{A'}$. Indices are raised and lowered with the antisymmetric tensor ϵ_{AB}. In the spacetime algebra version both κ^A and κ_A have the same multivector equivalent, which we write as

$$\kappa^A \leftrightarrow \kappa\tfrac{1}{2}(1+\boldsymbol{\sigma}_3). \tag{9.125}$$

The presence of the idempotent $(1+\boldsymbol{\sigma}_3)/2$ allows us to restrict κ to the Pauli-even algebra, as any Pauli-odd terms can be multiplied on the right by $\boldsymbol{\sigma}_3$ to convert them back to the even subspace. This ensures that κ has four real degrees of freedom, as required. Under a Lorentz transformation the full spinor transforms

to

$$R\kappa\tfrac{1}{2}(1+\boldsymbol{\sigma}_3) = \kappa'\tfrac{1}{2}(1+\boldsymbol{\sigma}_3), \tag{9.126}$$

where R is a Lorentz rotor. If we decompose the rotor R into Pauli-even and Pauli-odd terms, $R = R_+ + R_-$, then κ' is given by

$$\kappa' = R_+\kappa + R_-\kappa\boldsymbol{\sigma}_3. \tag{9.127}$$

The decomposition into Pauli-even and Pauli-odd terms is frame-dependent, as it depends on the choice of the γ_0 direction. But by augmenting κ with the $(1+\boldsymbol{\sigma}_3)/2$ idempotent we ensure that the full object is a proper Lorentz-covariant spinor.

The opposite idempotent, $(1-\boldsymbol{\sigma}_3)/2$, also generates a valid two-spinor which belongs to a second linear space (or module). This is the $\bar{\omega}^{A'}$ spinor in the notation of Penrose & Rindler, which we translate to

$$\bar{\omega}^{A'} \leftrightarrow = -\omega I\boldsymbol{\sigma}_2\tfrac{1}{2}(1-\boldsymbol{\sigma}_3). \tag{9.128}$$

The factor of $-I\boldsymbol{\sigma}_2$ is a matter of convention, and is inserted to simplify some of the later expressions. Under a Lorentz transformation we see that the Pauli-even element ω transforms as

$$\omega \mapsto \omega' = R_+\omega - R_-\omega\boldsymbol{\sigma}_3. \tag{9.129}$$

So κ and ω have different transformation laws: they belong to distinct carrier spaces of representations of the Lorentz group.

The power of the two-spinor calculus is the ease with which vector and tensor objects are generated from the basic two-spinors. As emphasised by Penrose & Rindler, this makes the calculus equally useful for both classical and quantum applications. It is instructive to see how this looks from the geometric algebra point of view. Unsurprisingly, what we discover is that the two-spinor calculus is a highly abstract and sophisticated means of introducing the geometric product to tensor manipulations. Once this is understood, much of the apparatus of the two-spinor calculus can be stripped away, and one is left with the now familiar spacetime algebra approach to relativistic physics.

9.5.1 Two-spinor observables

In two-spinor calculus one forms tensor objects from pairs of two-spinors, for example $\kappa^A\bar{\kappa}^{A'}$. To formulate this in the multiparticle spacetime algebra we simply multiply together the appropriate spinors, putting each spinor in its own copy of the spacetime algebra. In this way we replicate the tensor product implicit in writing $\kappa^A\bar{\kappa}^{A'}$. The result is that we form the object

$$\kappa^A\bar{\kappa}^{A'} \leftrightarrow -\kappa^1\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\kappa^2 I\boldsymbol{\sigma}_2^2\tfrac{1}{2}(1-\boldsymbol{\sigma}_3^2)\tfrac{1}{2}(1-I\boldsymbol{\sigma}_3^1\, I\boldsymbol{\sigma}_3^2). \tag{9.130}$$

$\frac{1}{2}(1+\boldsymbol{\sigma}_3^1)\frac{1}{2}(1-\boldsymbol{\sigma}_3^2)E = -\frac{1}{2}(\gamma_0^1+\gamma_3^1)I\boldsymbol{\sigma}_2^1\bar{\epsilon}\gamma_0^1$
$\frac{1}{2}(1-\boldsymbol{\sigma}_3^1)\frac{1}{2}(1+\boldsymbol{\sigma}_3^2)E = -\frac{1}{2}(\gamma_0^1-\gamma_3^1)I\boldsymbol{\sigma}_2^1\epsilon\gamma_0^1$
$\frac{1}{2}(1+\boldsymbol{\sigma}_3^1)\frac{1}{2}(1+\boldsymbol{\sigma}_3^2)E = -\frac{1}{2}(\boldsymbol{\sigma}_1^1+I\boldsymbol{\sigma}_2^1)\epsilon$
$\frac{1}{2}(1-\boldsymbol{\sigma}_3^1)\frac{1}{2}(1-\boldsymbol{\sigma}_3^2)E = -\frac{1}{2}(-\boldsymbol{\sigma}_1^1+I\boldsymbol{\sigma}_2^1)\bar{\epsilon}$

Table 9.2 *Two-spinor identities.* The identities listed here can be used to convert any expression involving a pair of two-spinors into an equivalent multivector.

As it stands this looks rather clumsy, but the various idempotents hide what is really going on. The key is to expose the Lorentz singlet structure hidden in the combination of idempotents. To achieve this we define two new Lorentz singlet states

$$\epsilon = \eta\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1), \quad \bar{\epsilon} = \eta\tfrac{1}{2}(1-\boldsymbol{\sigma}_3^2), \tag{9.131}$$

where η is the Lorentz singlet defined in equation (9.93). These new states both satisfy the essential equation

$$M^1\epsilon = \tilde{M}^2\epsilon, \quad M^1\bar{\epsilon} = \tilde{M}^2\bar{\epsilon}, \tag{9.132}$$

where M is an even-grade multivector. The reason is that any idempotents applied on the right of η cannot affect the result of equation (9.94). Expanding out in full, and rearranging the idempotents, we find that

$$\begin{aligned}\epsilon &= (I\boldsymbol{\sigma}_2^1 - I\boldsymbol{\sigma}_2^2)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^2)E,\\ \bar{\epsilon} &= (I\boldsymbol{\sigma}_2^1 - I\boldsymbol{\sigma}_2^2)\tfrac{1}{2}(1-\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1-\boldsymbol{\sigma}_3^2)E.\end{aligned} \tag{9.133}$$

These relations can manipulated to give, for example,

$$\begin{aligned}I\boldsymbol{\sigma}_2^1\epsilon &= -(1+I\boldsymbol{\sigma}_2^1\,I\boldsymbol{\sigma}_2^2)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^2)E,\\ \boldsymbol{\sigma}_1^1\epsilon &= -(1-I\boldsymbol{\sigma}_2^1\,I\boldsymbol{\sigma}_2^2)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^2)E.\end{aligned} \tag{9.134}$$

It follows that

$$\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^2)E = -\tfrac{1}{2}(\boldsymbol{\sigma}_1^1+I\boldsymbol{\sigma}_2^2)\epsilon. \tag{9.135}$$

There are four such identities in total, which are listed in table 9.2.

The results given in table 9.2 enable us to immediately convert any two-spinor expression into an equivalent multivector in the spacetime algebra. For example, returning to equation (9.130), we form

$$\begin{aligned}-\kappa^1\kappa^2 I\boldsymbol{\sigma}_2^2\tfrac{1}{2}\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1-\boldsymbol{\sigma}_3^2)E &= \kappa^1\kappa^2\tfrac{1}{2}(\gamma_0^1+\gamma_3^1)\bar{\epsilon}\gamma_0^1\\ &= \tfrac{1}{2}\big(\kappa(\gamma_0+\gamma_3)\tilde{\kappa}\big)^1\bar{\epsilon}\gamma_0^1.\end{aligned} \tag{9.136}$$

The key term in this expression is the null vector $\kappa(\gamma_0 + \gamma_3)\tilde{\kappa}$, which is constructed in the familiar manner for relativistic observables. A feature of the two-spinor calculus is that it lends itself to formulating most quantities in terms of null vectors. The origin of these can be traced back to the original $(1 \pm \boldsymbol{\sigma}_3)/2$ idempotents, which contain the null vector $\gamma_0 \pm \gamma_3$. These are rotated and dilated onto spacetime null vectors through the application of a spinor.

9.5.2 The two-spinor inner product

A Lorentz-invariant inner product for a pair of two-spinors is constructed from the antisymmetric combination

$$\kappa^A \omega_A = -\kappa_0 \omega_1 + \kappa_1 \omega_0, \tag{9.137}$$

where the subscripts here denote complex components of a two-spinor. The result of the inner product is a Lorentz-invariant complex scalar. The antisymmetry of the inner product tells us that we should form the equivalent expression

$$\begin{aligned}(\kappa^1\omega^2 - \kappa^2\omega^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^1)\tfrac{1}{2}(1+\boldsymbol{\sigma}_3^2)E &= -\tfrac{1}{2}\big(\kappa(\boldsymbol{\sigma}_1 + I\boldsymbol{\sigma}_2)\tilde{\omega} - \omega(\boldsymbol{\sigma}_1 + I\boldsymbol{\sigma}_2)\tilde{\kappa}\big)^1\epsilon \\ &= -\langle\kappa(\boldsymbol{\sigma}_1 + I\boldsymbol{\sigma}_2)\tilde{\omega}\rangle_{0,4}^1\epsilon. \end{aligned} \tag{9.138}$$

The antisymmetric product picks out the scalar and pseudoscalar parts of the quantity $\kappa(\boldsymbol{\sigma}_1 + I\boldsymbol{\sigma}_2)\tilde{\omega}$. This is sensible, as these are the two terms that are invariant under Lorentz transformations.

The fact that we form a scalar + pseudoscalar combination reveals a second important feature of the two-spinor calculus, which is that the unit imaginary is a representation of the spacetime pseudoscalar. The complex structure therefore has a concrete, geometric significance, which is one reason why two-spinor techniques have proved popular in general relativity, for example. Further insight into the form of the two-spinor inner product is gained by assembling the full even multivector

$$\psi = \kappa\tfrac{1}{2}(1+\boldsymbol{\sigma}_3) + \omega I\boldsymbol{\sigma}_2\tfrac{1}{2}(1-\boldsymbol{\sigma}_3). \tag{9.139}$$

The essential term in the two-spinor inner product is now reproduced by

$$\begin{aligned}\psi\tilde{\psi} &= -\kappa\tfrac{1}{2}(1+\boldsymbol{\sigma}_3)I\boldsymbol{\sigma}_2\tilde{\omega} + \omega I\boldsymbol{\sigma}_2\tfrac{1}{2}(1-\boldsymbol{\sigma}_3)\tilde{\kappa} \\ &= -\langle\kappa(\boldsymbol{\sigma}_1 + I\boldsymbol{\sigma}_2)\tilde{\omega}\rangle_{0,4}, \end{aligned} \tag{9.140}$$

so the inner products pick up both the scalar and pseudoscalar parts of a full Dirac spinor product $\psi\tilde{\psi}$. This form makes the Lorentz invariance of the product quite transparent. Interchanging κ and ω in ψ of equation (9.139) is achieved by right-multiplication by $\boldsymbol{\sigma}_1$, which immediately reverses the sign of $\psi\tilde{\psi}$.

9.5.3 Spin-frames and the null tetrad

An important concept in the two-spinor calculus is that of a *spin-frame*. This consists of a pair of two-spinors, κ^A and ω^A say, normalised such that $\kappa^A\omega_A = 1$. In terms of the spinor ψ of equation (9.139), this normalisation condition amounts to saying that ψ satisfies $\psi\tilde{\psi} = 1$. A normalised spin-frame is therefore the two-spinor encoding of a spacetime rotor. This realisation also sheds light on the associated concept of a *null tetrad*. In terms of the spin frame $\{\kappa^A, \omega^A\}$, the associated null tetrad is defined as follows:

$$\begin{aligned} l^a = \kappa^A\bar{\kappa}^{A'} &\leftrightarrow \big(\kappa(\gamma_0+\gamma_3)\tilde{\kappa}\big)^1\bar{\epsilon}\gamma_0^1, \\ n^a = \omega^A\bar{\omega}^{A'} &\leftrightarrow \big(\omega(\gamma_0+\gamma_3)\tilde{\omega}\big)^1\bar{\epsilon}\gamma_0^1, \\ m^a = \kappa^A\bar{\omega}^{A'} &\leftrightarrow \big(\kappa(\gamma_0+\gamma_3)\tilde{\omega}\big)^1\bar{\epsilon}\gamma_0^1, \\ \bar{m}^a = \omega^A\bar{\kappa}^{A'} &\leftrightarrow \big(\omega(\gamma_0+\gamma_3)\tilde{\kappa}\big)^1\bar{\epsilon}\gamma_0^1. \end{aligned} \tag{9.141}$$

In each case we have projected into a single copy of the spacetime algebra to form a geometric multivector. To simplify these expressions we introduce the rotor R defined by

$$R = \kappa\tfrac{1}{2}(1+\boldsymbol{\sigma}_3) + \omega I\boldsymbol{\sigma}_2\tfrac{1}{2}(1-\boldsymbol{\sigma}_3). \tag{9.142}$$

It follows that

$$\begin{aligned} R(\gamma_1+I\gamma_2)\tilde{R} &= -\kappa\gamma_1(1+\boldsymbol{\sigma}_3)I\boldsymbol{\sigma}_2\tilde{\omega} \\ &= \kappa(\gamma_0+\gamma_3)\tilde{\omega}. \end{aligned} \tag{9.143}$$

The null tetrad induced by a normalised spin-frame can now be written in the spacetime algebra as

$$\begin{aligned} l &= R(\gamma_0+\gamma_3)\tilde{R}, & m &= R(\gamma_1+I\gamma_2)\tilde{R}, \\ n &= R(\gamma_0-\gamma_3)\tilde{R}, & \bar{m} &= R(\gamma_1-I\gamma_2)\tilde{R}. \end{aligned} \tag{9.144}$$

(One can chose alternative normalisations, if required). The complex vectors m^a and $\bar{m}^a$ of the two-spinor calculus have now been replaced by vector + trivector combinations. This agrees with the earlier observation that the imaginary scalar in the two-spinor calculus plays the role of the spacetime pseudoscalar. The multivectors in a null tetrad satisfy the anticommutation relations

$$\{l, n\} = 4, \quad \{m, \bar{m}\} = 4, \quad \text{all others} = 0. \tag{9.145}$$

These relations provide a framework for the formulation of supersymmetric quantum theory within the multiparticle spacetime algebra.

9.6 Notes

The multiparticle spacetime algebra was introduced in the paper 'States and operators in the spacetime algebra' by Doran, Lasenby & Gull (1993a). Since its introduction the multiparticle spacetime algebra has been developed by a range of researchers. For introductions see the papers by Parker & Doran (2002) and Havel & Doran (2000a,2002b). Of particular interest are the papers by Somaroo et al. (1998,1999) and Havel et al. (2001), which show how the multiparticle spacetime algebra can be applied to great effect in the theory of quantum information processing. These researchers were primarily motivated by the desire to create quantum gates in an NMR environment, though their observations can be applied to quantum computation in general. For a good introduction into the subject of quantum information, we recommend the course notes made available by Preskill (1998).

The subject of relativistic multiparticle quantum theory has been tackled by many authors. The most authoritative discussions are contained in the papers by Salpeter & Bethe (1951), Salpeter (1952), Breit (1929) and Feynman (1961). A more modern perspective is contained in the discussions in Itzykson & Zuber (1980) and Grandy (1991). For more recent attempts at constructing a two-particle version of the Dirac equation, see the papers by Galeao & Ferreira (1992), Cook (1988) and Koide (1982). A summary of the multiparticle spacetime algebra approach to this problem is contained in Doran et al.(1996b).

The two-spinor calculus is described in the pair of books '*Spinors and Spacetime*' volumes I and II by Penrose & Rindler (1984,1986). The spacetime algebra version of two-spinor calculus is described in more detail in 'Geometric algebra and its application to mathematical physics' by Doran (1994), with additional material contained in the paper '2-spinors, twistors and supersymmetry in the spacetime algebra' by Lasenby et al. (1993b). The conventions adopted in this book differ slightly from those adopted in many of the earlier papers.

9.7 Exercises

9.1 Explain how the two-particle Schrödinger equation for the Coulomb problem is reduced to the effective single-particle equation

$$-\frac{\hbar^2\boldsymbol{\nabla}^2}{2\mu}\psi - \frac{q_1 q_2}{4\pi\epsilon_0 r}\psi = E\psi,$$

where μ is the reduced mass.

9.2 Given that $\psi(\theta,\phi) = \exp(-\phi I\boldsymbol{\sigma}_3/2)\exp(-\theta I\boldsymbol{\sigma}_2/2)$, prove that

$$\begin{pmatrix} \sin(\theta/2)\mathrm{e}^{-i\phi/2} \\ -\cos(\theta/2)\mathrm{e}^{i\phi/2} \end{pmatrix} \leftrightarrow \psi(\theta,\phi)I\boldsymbol{\sigma}_2.$$

Confirm that this state is orthogonal to $\psi(\theta,\phi)$.

9.3 The interaction energy of two dipoles is given classically by

$$E = \frac{\mu_0}{4\pi}\left(\frac{\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2}{r^3} - 3\frac{\boldsymbol{\mu}_1\cdot\boldsymbol{r}\,\boldsymbol{\mu}_2\cdot\boldsymbol{r}}{r^5}\right),$$

where $\boldsymbol{\mu}_i$ denotes the magnetic moment of particle i. For a quantum system of spin 1/2 particles we replace the magnetic moment vectors with the operators $\hat{\mu}_k = (\gamma\hbar/2)\hat{\sigma}_k$. Given that $\boldsymbol{n} = \boldsymbol{r}/r$, show that the Hamiltonian operator takes the form of the 4-vector

$$H = -\frac{d}{4}\left(\sum_{k-1}^{3} I\boldsymbol{\sigma}_k^1\, I\boldsymbol{\sigma}_k^2 - 3\, I\boldsymbol{n}^1\, I\boldsymbol{n}^2\right)$$

and find an expression for d. Can you solve the two-particle Schrödinger equation with this Hamiltonian?

9.4 ψ and ϕ are a pair of non-relativistic multiparticle states. Prove that the overlap probability between the two states can be written

$$P(\psi,\phi) = \frac{\langle(\psi E\tilde{\psi})(\phi E\tilde{\phi})\rangle - \langle(\psi J\tilde{\psi})(\phi J\tilde{\phi})\rangle}{2\langle\psi E\tilde{\psi}\rangle\langle\phi E\tilde{\phi}\rangle}.$$

9.5 Investigate the properties of the $l = 1$, $m = 0$ state

$$|\psi\rangle = |0\rangle|1\rangle + |1\rangle|0\rangle.$$

Is this state maximally entangled?

9.6 The β_μ operators that act on states in the two-particle relativistic algebra are defined by:

$$\beta_\mu(\psi) = \tfrac{1}{2}\left(\gamma_\mu^1\psi\gamma_0^1 + \gamma_\mu^2\psi\gamma_0^2\right).$$

Verify that these operators generate the *Duffin–Kemmer* ring

$$\beta_\mu\beta_\nu\beta_\rho + \beta_\rho\beta_\nu\beta_\mu = \eta_{\nu\rho}\beta_\mu + \eta_{\nu\mu}\beta_\rho.$$

9.7 The multiparticle wavefunction ψ is constructed from superpositions of states of the form $\phi^1(r^1)\chi^2(s^2)$, where ϕ and χ satisfy the single-particle Dirac equation. Prove that the individual currents $\mathcal{J}_1^1$ and $\mathcal{J}_2^2$ are conserved, where

$$\mathcal{J}_1^1 + \mathcal{J}_2^2 = \langle\psi(\gamma_0^1 + \gamma_0^2)\tilde{\psi}\rangle_1.$$

9.8 In the two-spinor calculus the two-component complex vector κ^A is acted on by a 2×2 complex matrix R . Prove that R is a representation of the Lorentz rotor group if $\det\mathsf{R} = 1$. (This defines the Lie group Sl(2,C).) Hence establish that the antisymmetric combination $\kappa^0\omega^1 - \kappa^1\omega^0$ is a Lorentz scalar.

9.9 The two-spinor calculus version of the Dirac equation is

$$\nabla^{A'A}\kappa_A = \mu\bar{\omega}^{A'},$$
$$\nabla^{AA'}\bar{\omega}_{A'} = \mu\kappa^A,$$

where $\mu = m/\sqrt{2}$. Prove that these equations are equivalent to the single equation $\nabla\psi I\boldsymbol{\sigma}_3 = m\psi\gamma_0$ and give an expression for ψ in terms of κ^A and $\bar{\omega}_{A'}$.

9.10 A null tetrad is defined by the set

$$l = R(\gamma_0 + \gamma_3)\tilde{R}, \qquad m = R(\gamma_1 + I\gamma_2)\tilde{R},$$
$$n = R(\gamma_0 - \gamma_3)\tilde{R}, \qquad \bar{m} = R(\gamma_1 - I\gamma_2)\tilde{R}.$$

Prove that these satisfy the anticommutation relations

$$\{l, n\} = 4, \quad \{m, \bar{m}\} = 4, \quad \text{all others} = 0.$$

10

Geometry

In the preceding chapters of this book we have dealt entirely with a single geometric interpretation of the elements of a geometric algebra. But the relationship between algebra and geometry is seldom unique. Geometric problems can be studied using a variety of algebraic techniques, and the same algebraic result can typically be pictured in a variety of different ways. In this chapter, we explore a range of alternative geometric systems, and discover how geometric algebra can be applied to each of them. We will find that there is no unique interpretation forced on the multivectors of a given grade. For example, to date we have viewed bivectors solely as directed plane segments. But in projective geometry a bivector represents a line, and in conformal geometry a bivector can represent a pair of points.

Ideas from geometry have always been a prime motivating factor in the development of mathematics. By the nineteenth century mathematicians were familiar with affine, Euclidean, spherical, hyperbolic, projective and inversive geometries. The unifying framework for studying these geometries was provided by the *Kleinian viewpoint*. Under this view a geometry consists of a space of points, together with a group of transformations mapping the points onto themselves. Any property of a particular geometry must be invariant under the action of the associated symmetry group. Klein was thus able to unite various geometries by describing how some symmetry groups are subgroups of larger groups. For example, Euclidean geometry is a subgeometry of affine geometry, because the group of Euclidean transformations is a subgroup of the group of affine transformations.

In this chapter we will see how the various classical geometries, and their associated groups, are handled in geometric algebra. But we will also go further by addressing the question of how to represent various geometric primitives in the most compact and efficient way. The Kleinian viewpoint achieves a united approach to classical geometry, but it does not help much when it comes to

addressing problems of how to perform calculations efficiently. For example, circles are as much geometric primitives in Euclidean geometry as points, lines a planes. But how should circles be represented as algebraic entities? Storing a point and a radius is unsatisfactory, as this representation involves objects of different grades. In this chapter we answer this question by showing that both lines and circles are represented as *trivectors* in the conformal model of Euclidean geometry.

We begin with the study of projective geometry. The addition of an extra dimension allows us to create an algebra of incidence relations between points, lines and planes in space. We then return to Euclidean geometry, but rather than viewing this as a subgeometry of projective geometry (the Kleinian viewpoint), we will instead increase the dimension once more to establish a conformal representation of Euclidean geometry. The beauty of this construction is that the group of Euclidean transformations can now be formulated as a rotor group. Euclidean invariants are then constructed as inner products between multivectors. This framework allows us to extend the projective treatment of incidence relations to include circles and spheres.

A further attractive feature of the conformal model is that Euclidean, spherical and hyperbolic geometries are all handled in the same framework. This allows the Poincaré disc model of non-Euclidean geometry in the plane to be extended seamlessly to higher dimensions. Of particular importance is the clarification of the role of complex coordinates in planar non-Euclidean geometry. Much of their utility rests on features of the conformal group of the plane that do not extend naturally. Instead, we work within the framework of *real* geometric algebra to obtain results which are independent of dimension. Finally in this chapter we turn to spacetime geometry. The conformal model for spacetime is of considerable importance in formulations of supersymmetric theories of gravity, and also lies at the heart of the twistor program. We display some surprising links between these ideas and the multiparticle spacetime algebra described in chapter 9. Throughout this chapter we denote the vector space with signature p, q by $\mathcal{V}(p, q)$, and the geometric algebra of this space by $\mathcal{G}(p, q)$.

10.1 Projective geometry

There was a time when projective geometry formed a large part of undergraduate mathematics courses. For various reasons the subject fell out of fashion in the twentieth century, making way for the more relevant subject of differential geometry. But in recent years projective geometry has enjoyed a resurgence due to its importance in the computer graphics industry. For example, the routines at the core of the OpenGL graphics language are built on a projective representation of three-dimensional space.

The key idea in projective geometry is that points in space are represented as

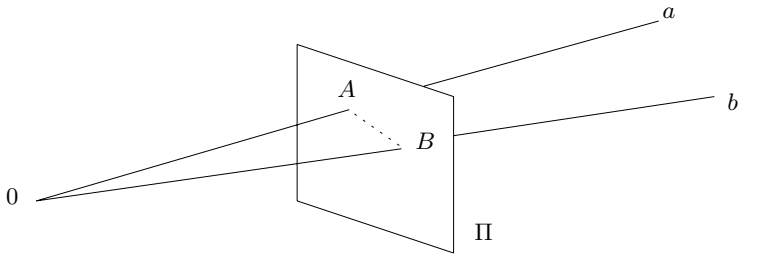

Figure 10.1 *Projective geometry.* Points in the projective plane are represented by vectors in a space one dimension higher. The plane Π does not intersect the origin 0.

vectors in a space of one dimension higher. For example, points in the projective plane are represented as vectors in three-dimensional space (see figure 10.1). The magnitude of the vector is unimportant, as both a and λa represent the same point. This representation of points is said to be *homogeneous*. The two key operations in projective geometry are the *join* and *meet*. The join of two points, for example, is the line between them. Forming the join raises the grade, and the join can usually be encoded algebraically via the exterior product (this was Grassmann's original motivation for introducing his exterior algebra). The meet is used for forming intersections, such as two lines in a plane meeting at a point. The meet is traditionally encoded via the notion of duality, and in geometric algebra the role of the meet is played by the inner product. Operations such as the meet and join do not depend on the metric, so in projective geometry we have a non-metric interpretation of the inner product. This is an important point. Some authors have argued that, because geometric algebra is built on a quadratic form, it is intimately tied to metric geometry. This view is incorrect, as we demonstrate below.

10.1.1 The projective line

The simplest place to start is with a one-dimensional line. The 'Euclidean' model of the line consists of labelling each point with a real number. But there are drawbacks with this representation of a line. Geometrically, all points on the line are equal. But algebraically there are two exceptional points on the line. The first is the origin, which is represented by the algebraically special number zero. The second is the point at infinity, which becomes important when we start to consider projective transformations. The resolution of both of these problems is to represent points in the line as vectors in two-dimensional space. In this way

the point x is replaced by a pair of homogeneous coordinates (x_1, x_2), with

$$x = \frac{x_1}{x_2}. \tag{10.1}$$

One can immediately see that the origin is represented by the non-zero vector $(0, 1)$, and that the point at infinity is $(1, 0)$.

If the vectors $\{\mathsf{e}_1, \mathsf{e}_2\}$ denote an orthonormal frame for two-dimensional space, we can set

$$\boldsymbol{x} = x_1\mathsf{e}_1 + x_2\mathsf{e}_2. \tag{10.2}$$

The set of all non-zero vectors $\boldsymbol{x}$ constitute the projective line, RP^1. The fact that the origin is excluded implies that in projective spaces one loses linearity. This is obvious from the fact that $\boldsymbol{x}$ and $\lambda\boldsymbol{x}$ represent the same point, so linear combinations do not make geometric sense. Indeed, no geometric significance can be attached to the addition of two points in projective geometry. One cannot form midpoints, for example, as distances and angles are not projective invariants.

The projective group consists of the group of general linear transformations applied to vectors in projective space. For the case of the projective line this group is defined by transformations of the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}, \qquad ab - bc \neq 0. \tag{10.3}$$

In terms of points on the line, this transformation corresponds to

$$x \mapsto x' = \frac{ax + b}{cx + d}. \tag{10.4}$$

The group action includes dilations, inversions and translations. The last are obtained for the case $c = 0$, $a/d = 1$. The fact that translations become *linear* transformations in projective geometry is of considerable importance. In three-dimensional geometry, for example, both rotations and translations can be encoded as 4×4 matrices. While this may appear to be an overly-complicated representation, it makes stringing together a series of translations and rotations a straightforward exercise. This is important in computer graphics, and is the representation employed in all OpenGL routines.

In geometric algebra notation we write a general linear transformation as the map $\boldsymbol{x} \mapsto \mathsf{f}(\boldsymbol{x})$, where $\det(\mathsf{f}) \neq 0$. Valid geometric statements in projective geometry must be invariant under such transformations, which is a strong restriction. Inner products between projective vectors (points) are clearly not invariant under projective transformations. The outer product does transform sensibly, however, due to the properties of the outermorphism. For example, suppose that the points α and β are represented projectively by

$$\boldsymbol{a} = \alpha\mathsf{e}_1 + \mathsf{e}_2, \quad \boldsymbol{b} = \beta\mathsf{e}_1 + \mathsf{e}_2. \tag{10.5}$$

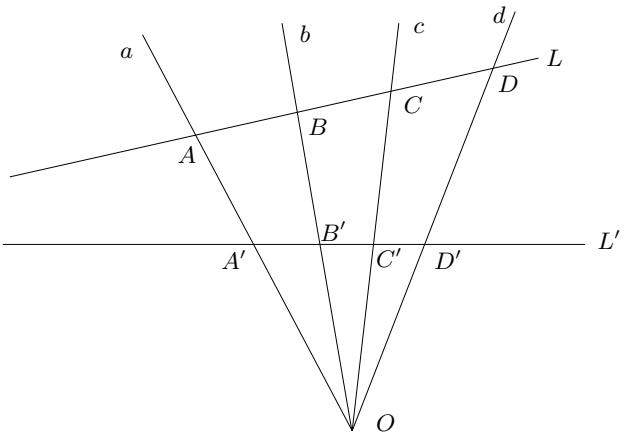

Figure 10.2 *The cross ratio.* Points on the lines L and L' represent two different projective views of the same vectors in space. The cross ratio of the four points is the same on both lines.

The outer product of these is

$$\boldsymbol{a}\wedge\boldsymbol{b} = (\alpha - \beta)\mathsf{e}_1\wedge\mathsf{e}_2, \tag{10.6}$$

which is controlled by the distance between the points on the line. Under a projective transformation in two dimensions

$$\mathsf{e}_1\wedge\mathsf{e}_2 \mapsto \mathsf{f}(\mathsf{e}_1\wedge\mathsf{e}_2) = \det\,(\mathsf{f})\,\mathsf{e}_1\wedge\mathsf{e}_2, \tag{10.7}$$

which is just an overall scaling.

The fact that distances between points are scaled under a projective transformation provides us with an important projective invariant for four points on a line. This is formed from ratios of lengths along a line. We must further ensure that the ratio is invariant under individual rescaling of individual vectors to be a true projective invariant. We therefore define the *cross ratio* of four points, A, B, C, D, by

$$(ABCD) = \frac{AC}{BC}\frac{BD}{AD} = \frac{\boldsymbol{a}\wedge\boldsymbol{c}}{\boldsymbol{b}\wedge\boldsymbol{c}}\frac{\boldsymbol{b}\wedge\boldsymbol{d}}{\boldsymbol{a}\wedge\boldsymbol{d}}, \tag{10.8}$$

where AB denotes the distance between A and B. Given any four points on a line, their cross ratio is a projective invariant (see figure 10.2). The figure illustrates one possible geometric interpretation of a projective transformation, which is that the line onto which points are projected is transformed to a new line. Invariants such as the cross ratio are important in computer vision where, for example, we seek to extract three-dimensional information from a series of two-

dimensional scenes. Knowledge of invariants can help establish point matches between the scenes.

10.1.2 The projective plane

Rather more interesting than the case of a line is that of the projective plane. Points in the plane are now represented by vectors in the three-dimensional algebra $\mathcal{G}(3,0)$. Figure 10.1 shows that the line between the points a and b is the result of projecting the plane defined by a and b onto the projective plane. We therefore define the *join* of the points a and b by

$$\text{join}(a,\, b) = a \wedge b. \tag{10.9}$$

Bivectors thus define lines in projective geometry. The line itself is recovered by solving the equation

$$a \wedge b \wedge x = 0. \tag{10.10}$$

This equation is solved by

$$x = \lambda a + \mu b, \tag{10.11}$$

which defines the set of projective points on the line joining A and B.

By taking exterior products of vectors we define (projectively) higher dimensional objects. For example, the join of a point a and a line $b \wedge c$ is the plane defined by the trivector $a \wedge b \wedge c$. Three points on a line cannot define a projected area, so for these we must have

$$a \wedge b \wedge c = 0 \quad \Rightarrow \quad a,\, b,\, c \text{ collinear}. \tag{10.12}$$

This was the condition used to recover the points x on the line $a \wedge b$. The join itself can be slightly more problematic. Given three points one cannot just write that their join is $a \wedge b \wedge c$, as the result may be zero. Instead the join is defined as the smallest subspace containing a, b and c. If they are collinear, then the join is the common line. This is well defined mathematically, but is hard to encode computationally. The problem is that the finite precision used on computers means that testing for zero is unreliable. Wherever possible it is safer to avoid defining the join and instead work with the exterior product.

Projective geometry deals with relationships that are invariant under projective transformations. The join is one such concept — as two points are transformed the line joining them transforms in the obvious way:

$$a \wedge b \mapsto \mathsf{f}(a) \wedge \mathsf{f}(b) = \mathsf{f}(a \wedge b). \tag{10.13}$$

So, for example, the statement that three points lie on a line ($a \wedge b \wedge c = 0$) is unchanged by a projective transformation. Similarly, the statement that three lines intersect at a point must also be a projective invariant. We therefore seek

an algebraic encoding of the intersection of two lines. This is the called the *meet*, usually denoted with the $\vee$ symbol. Before we can encode this, however, we need to define the dual. In the projective plane, points and lines are represented as vectors and bivectors in $\mathcal{G}(3,0)$. We know that these can be interchanged via a duality transformation, which amounts to multiplying by the pseudoscalar I. In this way every point has a dual line, and vice versa. The geometric picture associated with duality depends on the embedding plane.

If we denote the dual of A by A^*, the meet $A \vee B$ is defined by the 'de Morgan' rule

$$(A \vee B)^* = A^* \wedge B^*. \tag{10.14}$$

For a pair of lines in a plane, this amounts to

$$A \vee B = -I(IA) \wedge (IB) = I\, A \times B = A \cdot (IB) = (IA) \cdot B. \tag{10.15}$$

These formulae show how the inner product can be used to encode the meet, without imposing a metric on projective space. The expression

$$A \vee B = I\, A \times B \tag{10.16}$$

shows how the construction works. In three dimensions, $A \times B$ is the plane perpendicular to A and B, and $I\, A \times B$ is the line perpendicular to this plane, through the origin. This is therefore the line common to both planes, so projectively gives the point of intersection of two lines.

The meet of two distinct lines in a plane always results in a non-zero point. If the lines are parallel then their meet returns the point at infinity. Parallelism is not a projective invariant, however, so under a projective transformation two parallel lines can transform to lines intersecting at a finite point. This illustrates the fact that the point at infinity does not necessarily stay at infinity under projective transformations. It is instructive to see how the meet itself transforms under a projective transformation. Using the results of section 4.4, we find that

$$\begin{aligned} A \vee B \mapsto \mathsf{f}(A) \vee \mathsf{f}(B) &= I\left(I\mathsf{f}(A)\right) \wedge \left(I\mathsf{f}(B)\right) \\ &= \det(\mathsf{f})^2\, I\, \bar{\mathsf{f}}^{-1}(IA) \wedge \bar{\mathsf{f}}^{-1}(IB) \\ &= \det(\mathsf{f})^2\, I\, \bar{\mathsf{f}}^{-1}\big((IA) \wedge (IB)\big) \\ &= \det(\mathsf{f})\, \mathsf{f}\big(I\,(IA) \wedge (IB)\big). \end{aligned} \tag{10.17}$$

We can summarise this result as

$$\mathsf{f}(A) \vee \mathsf{f}(B) = \det(\mathsf{f})\, \mathsf{f}(A \vee B). \tag{10.18}$$

But in projective geometry, a and λa represent the same point, so the factor of $\det(\mathsf{f})$ does not affect the resulting point. This confirms that under a projective transformation the meet transforms as required.

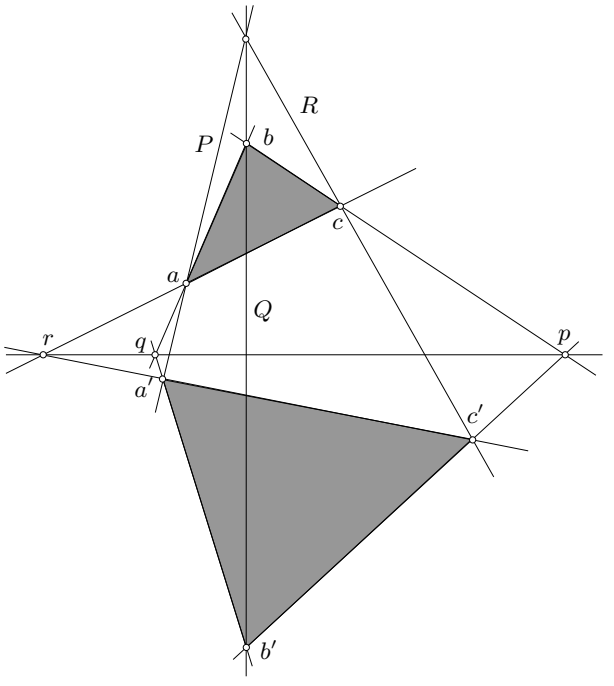

Figure 10.3 *Desargues' theorem.* The lines P, Q, R meet at a point if and only if the points p, q, r lie on a line. The two triangles are then projectively related.

The condition that three lines meet at a common point requires that the meet of two lines lies on a third line, which goes as

$$(A \vee B) \wedge C = (I\, A \times B) \wedge C = 0. \tag{10.19}$$

Dualising this result we obtain the condition

$$\langle (A \times B) C \rangle = \langle ABC \rangle = 0, \quad \Rightarrow \quad A,\, B,\, C \text{ coincident.} \tag{10.20}$$

This is an extremely simple algebraic encoding of the statement that three lines (represented by bivectors) all meet at a common point. Equations like this demonstrate how powerful geometric algebra can be when applied in a projective setting.

As an application consider Desargues' theorem, which is illustrated in figure 10.3. The points a, b, c and a', b', c' define two triangles. The associated lines are defined by

$$A = b \wedge c, \quad B = c \wedge a, \quad C = a \wedge b, \tag{10.21}$$

with the same definitions holding for A', B', C' in terms of a', b', c'. The two sets

of vertices determine the lines

$$P = a \wedge a', \quad Q = b \wedge b', \quad R = c \wedge c', \tag{10.22}$$

and the two sets of lines determine the points

$$p = A \times A' \, I, \quad q = B \times B' \, I, \quad r = C \times C' \, I. \tag{10.23}$$

Desargues' theorem states that, if p, q, r lie on a common line, then P, Q and R all meet at a common point. The latter condition requires

$$\langle PQR \rangle = \langle a \wedge a' \, b \wedge b' \, c \wedge c' \rangle = 0. \tag{10.24}$$

Similarly, for p, q, r to fall on a line we form

$$\begin{aligned} p \wedge q \wedge r &= \langle A \times A' \, I \, B \times B' \, I \, C \times C' \, I \rangle_3 \\ &= -I \langle A \times A' \, B \times B' \, C \times C' \rangle. \end{aligned} \tag{10.25}$$

Desargues' theorem is then proved by the algebraic identity

$$\langle a \wedge b \wedge c \, a' \wedge b' \wedge c' \rangle \langle a \wedge a' \, b \wedge b' \, c \wedge c' \rangle = \langle A \times A' \, B \times B' \, C \times C' \rangle, \tag{10.26}$$

the proof of which is left as an exercise. The left-hand side vanishes if and only if the lines P, Q, R meet at a point. The right-hand side vanishes if and only if the points p, q, r lie on a line. This proves the theorem. The complex geometry illustrated in figure 10.3 has therefore been reduced to a straightforward algebraic identity.

We can find a simple generalisation of the cross ratio for the case of the projective plane. From the derivation of the cross ratio, it is clear that any analogous object for the plane must involve ratios of trivectors. These represent areas in the projective plane. For example, suppose we have six points in space with position vectors $a_1, \ldots, a_6$. These produce the six projected points $A_1, \ldots, A_6$. An invariant is formed by

$$\frac{a_5 \wedge a_4 \wedge a_3}{a_5 \wedge a_1 \wedge a_3} \frac{a_6 \wedge a_2 \wedge a_1}{a_6 \wedge a_2 \wedge a_4} = \frac{A_{543}}{A_{513}} \frac{A_{621}}{A_{624}}, \tag{10.27}$$

where A_{ijk} is the projected area of the triangle with vertices A_i, A_j, A_k. Again, elementary algebraic reasoning quickly yields a geometrically significant result.

10.1.3 Homogeneous coordinates and projective splits

In typical applications of projective geometry we are interested in the relationship between coordinates in an image plane (for example in terms of pixels relative to some origin) and the three-dimensional position vector. Suppose that the origin in the image plane is defined by the vector n, which is perpendicular to the plane. The line on the image plane from the origin to the image point is represented by the bivector $a \wedge n$ (see figure 10.4) . The vector OA belongs to a two-dimensional

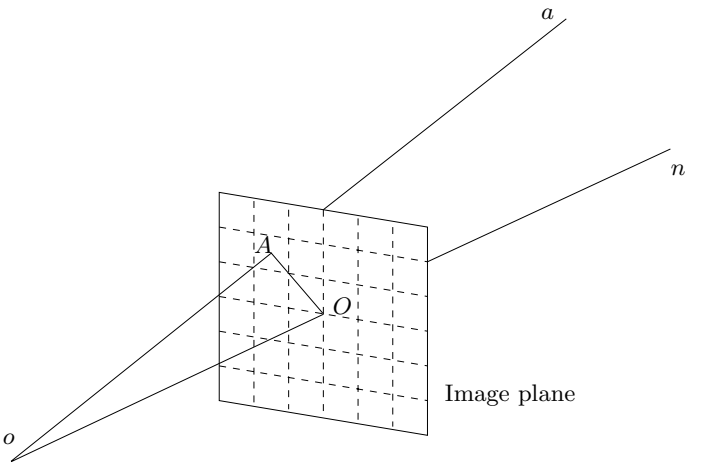

Figure 10.4 *The image plane.* Vectors in the image plane, OA, are described by bivectors in $\mathcal{G}(3,0)$. The point A can be expressed in terms of *homogeneous* coordinates in the image plane.

geometric algebra. We can relate this directly to the three-dimensional algebra by first writing

$$n + OA = \lambda a. \tag{10.28}$$

Contracting with n, we find that $\lambda = n^2(a\cdot n)^{-1}$. It follows that

$$OA = \frac{a\,n^2 - a\cdot n\,n}{a\cdot n} = \frac{a\wedge n}{a\cdot n}n. \tag{10.29}$$

If we now drop the final factor of n, we obtain a bivector that is homogeneous in both a and n. In this way we can directly represent the line OA in two dimensions with the bivector

$$A = \frac{a\wedge n}{a\cdot n}. \tag{10.30}$$

This is the *projective split*, first introduced in chapter 5 as a means of relating physics as seen by observers with different velocities.

The map of equation (10.30) relates bivectors in a higher dimensional space to vectors in a space of dimension one lower. If we introduce a coordinate frame $\{\mathsf{e}_i\}$, with e_3 in the n direction, we see that the coordinates of the image of $a = a_i\mathsf{e}_i$ are

$$A = \frac{a_1}{a_3}\mathsf{e}_1\mathsf{e}_3 + \frac{a_2}{a_3}\mathsf{e}_2\mathsf{e}_3 = A_1\boldsymbol{E}_1 + A_2\boldsymbol{E}_2. \tag{10.31}$$

This equation defines the *homogeneous coordinates* A_i:

$$A_i = \frac{a_i}{a_3}. \tag{10.32}$$

Homogeneous coordinates are independent of scale and it is these that are usually measured in a camera projection of a scene. The bivectors $(\boldsymbol{E}_1, \boldsymbol{E}_2)$ act as generators for a two-dimensional geometric algebra. If the vectors in the projective space are all Euclidean, the $\boldsymbol{E}_i$ bivectors will have negative square. If necessary, this can be avoided by letting e_3 be an anti-Euclidean vector. The projective split is an elegant scheme for relating results in projective space to Euclidean space one dimension lower. Algebraically, the projective split rests on the isomorphism

$$\mathcal{G}^+(p+1,q) \simeq \mathcal{G}(q,p). \tag{10.33}$$

This states that the even subalgebra of the geometric algebra with signature $(p+1,q)$ is isomorphic to the algebra with signature (q,p). The projective split is not always the best way to map from projective space back to Euclidean space, however, as constructing a set of bivectors can be an unnecessary complication. Often it is simpler to choose an orthonormal frame, with n one of the frame vectors, and then scale all vectors x such that $n \cdot x = 1$.

10.1.4 Projective geometry in three dimensions

To handle complicated three-dimensional problems in a projective framework we require a four-dimensional geometric algebra. The basic elements of four-dimensional geometric algebra will be familiar from relativity and the spacetime algebra, though now the elements are given a projective interpretation. The algebra of a four-dimensional space contains six bivectors, which represent lines in three dimensions. As in the planar case, the important feature of the projective framework is that we are free from the restriction that all lines pass through the origin. The line through the points a and b is again represented by the bivector $a \wedge b$. This is a *blade*, as must be the case for any bivector representing a line. Any bivector blade $B = a \wedge b$ must satisfy the algebraic condition

$$B \wedge B = a \wedge b \wedge a \wedge b = 0, \tag{10.34}$$

which removes one degree of freedom from the six components needed to specify an arbitrary bivector. This is known at the Plücker condition. If the vector e_4 defines the projection into Euclidean space, the line $a \wedge b$ has coordinates

$$a \wedge b = (\boldsymbol{a} + \mathsf{e}_4) \wedge (\boldsymbol{b} + \mathsf{e}_4) = \boldsymbol{a} \wedge \boldsymbol{b} + (\boldsymbol{a} - \boldsymbol{b}) \wedge \mathsf{e}_4, \tag{10.35}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ denote vectors in the three-dimensional space. The bivector B therefore encodes a line as a combination of a tangent $(\boldsymbol{b} - \boldsymbol{a})$ and a moment $\boldsymbol{a} \wedge \boldsymbol{b}$. These are the Plücker coordinates for a line.

Given two lines as bivectors B and B', the test that they intersect in three dimensions is that their join does not span all of projective space, which implies

that

$$B \wedge B' = 0. \tag{10.36}$$

This provides a projective interpretation for commuting bivectors in four dimensions. Commuting (orthogonal) bivectors have BB' equalling a multiple of the pseudoscalar. Projectively, these can be interpreted as two lines in three dimensions that do not share a common point. As mentioned earlier, the problem with a test such as equation (10.36) is that one can never guarantee to obtain zero when working to finite numerical precision. In practice, then, one tends to avoid trying to find the intersection of two lines in the three dimensions, unless there is good reason to believe that they intersect at a point.

The exterior product of three vectors in projective space results in the trivector encoding the plane containing the three points. One of the most frequently encountered problems is finding the point of intersection of a line L and a plane P. This is given by

$$x = P \cdot (IL), \tag{10.37}$$

where I is the four-dimensional pseudoscalar. This will always return a point, provided the line does not lie entirely in the plane. Similarly, the intersection of two planes in three dimensions must result in a line. Algebraically, this line is encoded by the bivector

$$L = (IP_1) \cdot P_2 = I\, P_1 \times P_2, \tag{10.38}$$

where P_1 and P_2 are the two planes. Such projective formulae are important in computer vision and graphics applications.

10.2 Conformal geometry

Projective geometry does provide an efficient framework for handling Euclidean geometry. Euclidean geometry is a subgeometry of projective geometry, so any valid result in the latter must hold in the former. But there are some limitations to the projective viewpoint. Euclidean concepts, like lengths and angles, are not straightforwardly encoded, and the related concepts of circles and spheres are equally awkward. Conformal geometry provides an elegant solution to this problem. The key is to introduce a further dimension of opposite signature, so that points in a space of signature (p, q) are modelled as null vectors in a space of signature $(p + 1, q + 1)$. That is, points in $\mathcal{V}(p, q)$ are represented by null vectors in $\mathcal{V}(p + 1, q + 1)$. Projective geometry is retained as a subset of conformal geometry, but the range of geometric primitives is extended to include circles and spheres.

We denote a point in $\mathcal{V}(p, q)$ by x, and its conformal representation by X. We continue to employ the spacetime notation of using the tilde symbol to denote

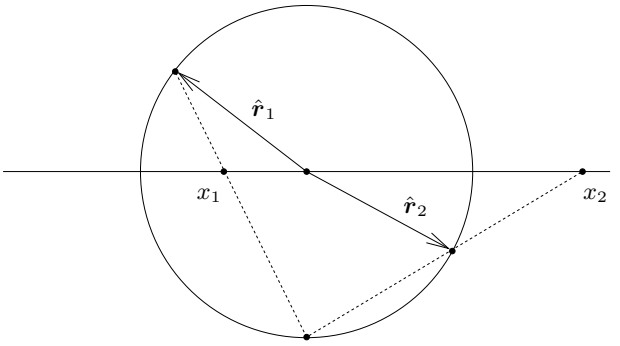

Figure 10.5 *A stereographic projection.* The line is mapped into the unit circle, so the points on the line x_1 and x_2 are mapped to the unit vectors $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$. The origin and infinity are mapped to opposite points on the circle.

the reverse operation for a general multivector in any geometric algebra. A basis set of vectors for $\mathcal{G}(p, q)$ is denoted by $\{e_i\}$, and the two additional vectors $\{e, \bar{e}\}$ complete this to an orthonormal basis for $\mathcal{G}(p + 1, q + 1)$.

10.2.1 Stereographic projection of a line

We illustrate the general construction by starting with the simple case of a line. In projective geometry points on a line are modeled as two-dimensional vectors. The conformal model is established from a slightly different starting point, using the *stereographic projection*. Under a stereographic projection, points on a line are mapped to the unit circle in a plane (see figure 10.5). Points on the unit circle in two dimensions are represented by

$$\hat{\boldsymbol{r}} = \cos(\theta)\, e_1 + \sin(\theta)\, e_2. \tag{10.39}$$

The corresponding point on the line is given by

$$x = \frac{\cos(\theta)}{1 + \sin(\theta)}. \tag{10.40}$$

This relation inverts simply to give

$$\cos(\theta) = \frac{2x}{1 + x^2}, \quad \sin(\theta) = \frac{1 - x^2}{1 + x^2}. \tag{10.41}$$

So far we have achieved a representation of the line in terms of a circle in two dimensions. But the constraint that the vector has unit magnitude means that

we have lost homogeneity. To get round this we introduce a third vector, $\bar{e}$, which has negative signature,

$$\bar{e}^2 = -1, \tag{10.42}$$

and we assume that $\bar{e}$ is orthogonal to e_1 and e_2. We can now replace the unit vector $\hat{\boldsymbol{r}}$ with the null vector X, where

$$X = \cos(\theta)\, e_1 + \sin(\theta)\, e_2 + \bar{e} = \frac{2x}{1+x^2} e_1 + \frac{1-x^2}{1+x^2} e_2 + \bar{e}. \tag{10.43}$$

The vector X satisfies $X^2 = 0$, so is null.

The equation $X^2 = 0$ is homogeneous. If it is satisfied for X, it is satisfied for λX. We can therefore move to a homogeneous representation and let both X and λX represent the same point. Multiplying by $(1 + x^2)$ we establish the conformal representation

$$X = 2xe_1 + (1 - x^2)e_2 + (1 + x^2)\bar{e}. \tag{10.44}$$

This is the basic representation we use throughout. To establish a more general notation we first replace the vector e_2 by $-e$. We therefore have

$$e^2 = 1, \qquad \bar{e}^2 = -1, \qquad e \cdot \bar{e} = 0. \tag{10.45}$$

The vectors e and $\bar{e}$ are then the two extra vectors that extend the space $\mathcal{V}(p, q)$ to $\mathcal{V}(p+1, q+1)$. Frequently, it is more convenient to work with a null basis for the extra dimensions. We define

$$n = e + \bar{e}, \qquad \bar{n} = e - \bar{e}. \tag{10.46}$$

These vectors satisfy

$$n^2 = \bar{n}^2 = 0, \qquad n \cdot \bar{n} = 2. \tag{10.47}$$

The vector X is now

$$X = 2xe_1 + x^2 n - \bar{n}. \tag{10.48}$$

It is straightforward to confirm that this is a null vector. The set of all null vectors in this space form a cone, and the real number line is modelled by the intersection of this cone and a plane. The construction is illustrated in figure 10.6.

10.2.2 Conformal model of Euclidean space

The form of equation (10.48) generalises easily. If x is an element of $\mathcal{V}(p, q)$, we set

$$F(x) = X = x^2 n + 2x - \bar{n}, \tag{10.49}$$

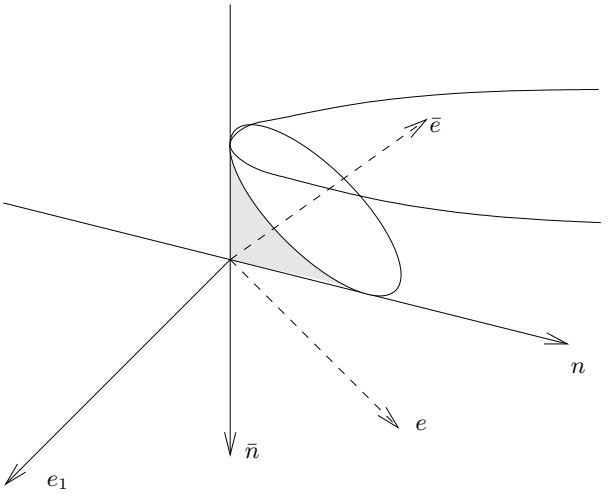

Figure 10.6 *The conformal model of a line.* Points on the line are represented by null vectors in three dimensions. These lie on a cone, and the intersection of the cone with a plane recovers the point.

which is a null vector in $\mathcal{V}(p+1, q+1)$. This vector can be obtained simply via the map,

$$F(x) = -(x - e)n(x - e), \tag{10.50}$$

which is a reflection of the null vector n in the plane perpendicular to $(x - e)$. The result must therefore be a new null vector. The presence of the vector e removes any ambiguity in handling the origin $x = 0$. The map $F(x)$ is non-linear so, as with projective geometry, we move to a non-linear representation of points in conformal geometry.

More generally, any null vector in $\mathcal{V}(p + 1, q + 1)$ can be written as

$$X = \lambda(x^2 n + 2x - \bar{n}), \tag{10.51}$$

with λ a scalar. This provides a projective map between $\mathcal{V}(p + 1, q + 1)$ and $\mathcal{V}(p, q)$. The family of null vectors, $\lambda(x^2 n + 2x - \bar{n})$, in $\mathcal{V}(p+1, q+1)$ correspond to the single point $x \in \mathcal{V}(p, q)$. Given an arbitrary null vector X, it is frequently useful to convert it to the standard form of equation (10.49). This is achieved by setting

$$X \mapsto -2\frac{X}{X \cdot n}. \tag{10.52}$$

This map is similar to that employed in constructing a standard embedding in projective geometry. The status of the vector n is clear here — it represents the point at infinity.

Given two null vectors X and Y, in standard form, their inner product is

$$\begin{aligned} X\cdot Y &= \left(x^2 n + 2x - \bar{n}\right)\cdot\left(y^2 n + 2y - \bar{n}\right) \\ &= -2x^2 - 2y^2 + 4x\cdot y \\ &= -2(x-y)^2. \end{aligned} \tag{10.53}$$

This result is of fundamental importance to the conformal model of Euclidean geometry. The inner product in conformal space encodes the *distance* between points in Euclidean space. It follows that any transformation of null vectors in $\mathcal{V}(p+1, q+1)$ which leaves inner products invariant can correspond to a transformation in $\mathcal{V}(p,q)$ which leaves angles and distances invariant. In the next section we discuss these transformations in detail.

10.3 Conformal transformations

The study of the main geometric primitives in conformal geometry is simplified by first understanding the nature of the conformal group. For points x, y in $\mathcal{V}(p,q)$ the definition of a conformal transformation is that it leaves angles invariant. So, if f is a map from $\mathcal{V}(p,q)$ to itself, then f is a conformal transformation if

$$\mathsf{f}(a)\cdot\mathsf{f}(b) = \lambda a\cdot b, \quad \forall a, b \in \mathcal{V}(p,q), \tag{10.54}$$

where

$$\mathsf{f}(a) = a\cdot\nabla f(x). \tag{10.55}$$

While $\mathsf{f}(a)$ is a linear map at each point x, the conformal transformation $f(x)$ is not restricted to being linear. Conformal transformations form a group, the conformal group, the main elements of which are translations, rotations, dilations and inversions. We now study each of these in turn.

10.3.1 Translations

To begin, consider the fundamental operation of translation in the space $\mathcal{V}(p,q)$. This is *not* a linear operation in $\mathcal{V}(p,q)$, but does become linear in the projective framework. In the conformal model we achieve a further refinement, as translations can now be handled by rotors. Consider the rotor

$$R = T_a = \mathrm{e}^{na/2}, \tag{10.56}$$

where $a \in \mathcal{V}(p,q)$, so that $a\cdot n = 0$. The generator for the rotor is a null bivector, so the Taylor series for T_a terminates after two terms:

$$T_a = 1 + \frac{na}{2}. \tag{10.57}$$

The rotor T_a transforms the null vectors n and $\bar{n}$ into

$$T_a n \tilde{T}_a = n + \tfrac{1}{2}nan + \tfrac{1}{2}nan + \tfrac{1}{4}nanan = n \tag{10.58}$$

and

$$T_a \bar{n} \tilde{T}_a = \bar{n} - 2a - a^2 n. \tag{10.59}$$

Acting on a vector $x \in \mathcal{V}(p,q)$ we similarly obtain

$$T_a x \tilde{T}_a = x + n(a \cdot x). \tag{10.60}$$

Combining these we find that

$$\begin{aligned} T_a F(x) \tilde{T}_a &= x^2 n + 2(x + a \cdot x\, n) - (\bar{n} - 2a - a^2 n) \\ &= (x+a)^2 n + 2(x+a) - \bar{n} \\ &= F(x+a), \end{aligned} \tag{10.61}$$

which performs the conformal version of the translation $x \mapsto x + a$. Translations are handled as rotations in conformal space, and the rotor group provides a double-cover representation of a translation. The identity

$$\tilde{T}_a = T_{-a} \tag{10.62}$$

ensures that the inverse transformation in conformal space corresponds to a translation in the opposite direction, as required.

10.3.2 Rotations

Next, suppose that we rotate the vector x about the origin in $\mathcal{V}(p,q)$. This is achieved with the rotor $R \in \mathcal{G}(p,q)$ via the familiar transformation $x \mapsto x' = Rx\tilde{R}$. The image of the transformed point is

$$\begin{aligned} F(x') &= x'^2 n + 2Rx\tilde{R} - \bar{n} \\ &= R(x^2 n + 2x - \bar{n})\tilde{R} = RF(x)\tilde{R}. \end{aligned} \tag{10.63}$$

This holds because R is an even element in $\mathcal{G}(p,q)$, so must commute with both n and $\bar{n}$. Rotations about the origin therefore take the same form in either space.

Suppose instead that we wish to rotate about the point $a \in \mathcal{V}(p,q)$. This can be achieved by translating a to the origin, rotating and then translating forward again. In terms of $X = F(x)$ the result is

$$X \mapsto T_a R T_{-a} X \tilde{T}_{-a} \tilde{R} \tilde{T}_a = R' X \tilde{R}. \tag{10.64}$$

The rotation is now controlled by the rotor

$$R' = T_a R \tilde{T}_a = \left(1 + \frac{na}{2}\right) R \left(1 + \frac{an}{2}\right). \tag{10.65}$$

So, as expected, the conformal model has freed us from treating the origin as a

special point. Rotations about any point are handled in the same manner, and are still generated by a bivector blade. Similar observations hold for reflections, but we delay a full treatment of these until we have described how lines and surfaces are handled in the conformal model. The preceding formulae for translations and rotations form the basis of the subject of *screw theory*, which has its origins in the nineteenth century.

10.3.3 Inversions

Rotations and translations are elements of the Euclidean group, as they leave distances between points invariant. This is a subgroup of the larger conformal group, which only leaves angles invariant. The conformal group essentially contains two further transformations: inversions and dilations. An inversion in the origin consists of the map

$$x \mapsto \frac{x}{x^2}. \tag{10.66}$$

The conformal vector corresponding to the inverted point is

$$F(x^{-1}) = x^{-2}n + 2x^{-1} - \bar{n} = \frac{1}{x^2}(n + 2x - x^2\bar{n}). \tag{10.67}$$

But in conformal space points are represented homogeneously, so the pre-factor of x^{-2} can be ignored. In conformal space an inversion in the origin consists solely of the map

$$n \mapsto -\bar{n}, \qquad \bar{n} \mapsto -n. \tag{10.68}$$

This is generated by a reflection in e, since

$$-ene = -ee\bar{n} = -\bar{n}. \tag{10.69}$$

We can therefore write

$$-eF(x)e = x^2F(x^{-1}), \tag{10.70}$$

which shows that inversions in $\mathcal{V}(p,q)$ are represented as reflections in the conformal space $\mathcal{V}(p+1,q+1)$. As both X and $-X$ are homogeneous representations of the same point, it is irrelevant whether we take $-e(\ldots)e$ or $e(\ldots)e$ as the reflection. In the following we will use $e(...)e$ for convenience.

A reflection in e corresponds to an inversion in the origin in Euclidean space. To find the generator of an inversion in an arbitrary point a, we translate to the origin, invert and translate forward again. The resulting generator is then

$$T_a e T_{-a} = \left(1 + \frac{na}{2}\right) e \left(1 + \frac{an}{2}\right) = e - a - \frac{a^2}{2}n. \tag{10.71}$$

Now, recalling that $e = (n + \bar{n})/2$, the generating vector can also be written as

$$T_a e T_{-a} = \tfrac{1}{2}\big(n - F(a)\big) = \tfrac{1}{2}(n - A). \tag{10.72}$$

A reflection in $(n - F(a))$ therefore achieves an inversion about the point a in Euclidean space. As with translations, a nonlinear transformation in Euclidean space has been linearised by moving to a conformal representation of points. The generator of an inversion is a vector with positive square. In section 10.5.1 we see how these vectors are related to circles and spheres.

10.3.4 Dilations

A dilation in the origin is given by

$$x \mapsto x' = \mathrm{e}^{-\alpha} x, \tag{10.73}$$

where α is a scalar. Clearly, this transformation does not alter angles, so is a conformal transformation. The null vector corresponding to the transformed point is

$$F(x') = \mathrm{e}^{-\alpha}(x^2 \mathrm{e}^{-\alpha} n + 2x + \mathrm{e}^{\alpha} \bar{n}). \tag{10.74}$$

Clearly the map we need to achieve is

$$n \mapsto \mathrm{e}^{-\alpha} n, \qquad \bar{n} \mapsto \mathrm{e}^{\alpha} \bar{n}. \tag{10.75}$$

This transformation does not alter the inner product of n and $\bar{n}$, so can be represented with a rotor. As the vector x is unchanged, the rotor can only be generated by the timelike bivector $e\bar{e}$. If we set

$$N = e\bar{e} = \tfrac{1}{2} \bar{n} \wedge n \tag{10.76}$$

then N satisfies

$$Nn = -n = -nN, \qquad N\bar{n} = \bar{n} = -\bar{n}N, \qquad N^2 = 1. \tag{10.77}$$

We now introduce the rotor

$$D_\alpha = \mathrm{e}^{\alpha N/2} = \cosh(\alpha/2) + \sinh(\alpha/2)\, N. \tag{10.78}$$

This rotor satisfies

$$\begin{aligned} D_\alpha n \tilde{D}_\alpha &= \mathrm{e}^{-\alpha} n, \\ D_\alpha \bar{n} \tilde{D}_\alpha &= \mathrm{e}^{\alpha} \bar{n} \end{aligned} \tag{10.79}$$

and so carries out the required transformation. We can therefore write

$$F(\mathrm{e}^{-\alpha} x) = \mathrm{e}^{-\alpha} D_\alpha F(x) \tilde{D}_\alpha, \tag{10.80}$$

which confirms that a dilation in the origin is represented by a simple rotor in conformal space. To achieve a dilation about an arbitrary point a we form

$$D'_\alpha = T_a D_\alpha \tilde{T}_a = \mathrm{e}^{\alpha N'/2}, \tag{10.81}$$

where the generator is now

$$N' = T_a N \tilde{T}_a = \tfrac{1}{2} T_a \bar{n} \wedge n \tilde{T}_a = -\tfrac{1}{2} A \wedge n, \tag{10.82}$$

with $A = F(a)$. A dilation about a is therefore generated by

$$D'_\alpha = \exp(-\alpha A \wedge n/4) = \exp\left(\frac{\alpha}{2} \frac{A \wedge n}{A \cdot n}\right). \tag{10.83}$$

The generator is governed by two null vectors, one for the point about which the dilation is performed and one for the point at infinity.

10.3.5 Special conformal transformations

A special conformal transformation consists of an inversion in the origin, a translation and a further inversion in the origin. We can therefore handle these in terms of the representations we have already established. In Euclidean space the effect of a conformal transformation can be written as

$$x \mapsto \frac{x + ax^2}{1 + 2a \cdot x + a^2 x^2} = x \frac{1}{1 + ax} = \frac{1}{1 + xa} x. \tag{10.84}$$

The final expressions confirm that a special conformal transformation corresponds to a position-dependent rotation and dilation in Euclidean space, so does leave angles unchanged. To construct the equivalent rotor in $\mathcal{G}(p+1, q+1)$ we form

$$K_a = e T_a e = 1 - \frac{\bar{n} a}{2}, \tag{10.85}$$

which ensures that $K_a F(x) \tilde{K}_a$ is a special conformal transformation. Explicitly, we have

$$F\left(x \frac{1}{1 + ax}\right) = (1 + 2a \cdot x + a^2 x^2)^{-1} K_a F(x) \tilde{K}_a \tag{10.86}$$

and again we can ignore the pre-factor and use $K_a F(x) \tilde{K}_a$ as the homogeneous representation of the result of a special conformal transformation.

10.3.6 Euclidean transformations

The group of Euclidean transformations is a subgroup of the full conformal group. The additional restriction is that lengths as well as angles are invariant. Equation (10.53) showed that the inner product of two null vectors is related to the Euclidean distance between the corresponding points. To establish a homogeneous formula, we must write

$$|a - b|^2 = -2 \frac{A \cdot B}{A \cdot n \, B \cdot n}, \tag{10.87}$$

which is homogeneous on A and B. The Euclidean group can now be seen to be the subgroup of the conformal group which leaves n invariant. This is sensible, as the point at infinity should stay there under a Euclidean transformation. The Euclidean group is thus the *stability group* of a null vector in conformal space. The group of generators of reflections and rotations in conformal space which leave n invariant then provide a double cover of the Euclidean group. Equation (10.87) returns the Euclidean distance between points. If the vector n is replaced by e or $\bar{e}$ we can transform to distance measures in hyperbolic or spherical geometry. This makes it a simple exercise to attach different geometric pictures to algebraic results in conformal space.

10.4 Geometric primitives in conformal space

Now that we have seen how points are encoded in conformal space, we can begin to build up more complex geometric objects. As in projective geometry, we expect that a multivector blade L will encode a geometric object via the equation

$$L \wedge X = 0, \qquad X^2 = 0. \tag{10.88}$$

The question, then, is what type of object does each grade of multivector return. One important result we can exploit is that $X^2 = 0$ is unchanged if $X \mapsto RX\tilde{R}$. So, if a geometric object is specified by L via equation (10.88), it follows that

$$R(L \wedge X)\tilde{R} = (RL\tilde{R}) \wedge (RX\tilde{R}) = 0. \tag{10.89}$$

We can therefore transform the object L with a general element of the conformal group to obtain a new object. Similar considerations hold for incidence relations. Since conformal transformations only preserve angles, and do not necessarily map straight lines to straight lines, the range of objects we can describe by simple blades is clearly going to be larger than in projective geometry.

10.4.1 Bivectors and points

A pair of points in Euclidean space are represented by two null vectors in a space of two dimensions higher. We know that the inner product in this space returns information about distances. The next question to ask is what is the significance of the outer product of two vectors. If A and B are null vectors, we form the bivector

$$G = A \wedge B. \tag{10.90}$$

The bivector G has magnitude

$$G^2 = (AB - A \cdot B)(-BA + A \cdot B) = (A \cdot B)^2, \tag{10.91}$$

which shows that G is *timelike*, borrowing the terminology of special relativity. It follows that G contains a pair of null vectors. If we look for solutions to the equation

$$G \wedge X = 0, \qquad X^2 = 0, \tag{10.92}$$

the only solutions are the two null vectors contained in G. These are precisely A and B, so the bivector encodes the two points directly. In the conformal model, no information is lost in forming the exterior product of two null vectors. Spacelike bivectors, with $B^2 < 0$, do not contain any null vectors, so in this case there are no solutions to $B \wedge X = 0$ with $X^2 = 0$. The critical case of $B^2 = 0$ implies that B contains a single null vector.

Given a timelike bivector, $B^2 > 0$, we require an efficient means of finding the two null vectors in the plane. This can be achieved without solving any quadratic equations as follows. Pick an arbitrary vector a, with a partial projection in the plane, $a \cdot B \neq 0$. If the underlying space is Euclidean, one can use the vector $\bar{e}$, since all timelike bivectors contain a factor of this. Now remove the component of a outside the plane by defining

$$a' = a - a \wedge \hat{B}\, \hat{B}, \tag{10.93}$$

where $\hat{B} = B/|B|$ is normalised so that $\hat{B}^2 = 1$. If a' is already null then it defines one of the required vectors. If not, then one can form two null vectors in the B plane by writing

$$A_\pm = a' \pm a' \hat{B}. \tag{10.94}$$

One can easily confirm that $A_\pm$ are both null vectors, and so return the desired points.

10.4.2 Trivectors, lines and circles

If a bivector now only represents a pair of points, the obvious question is how do we describe a line? Suppose we construct the line through the points a and b in $\mathcal{V}(p, q)$. A point on the line is given by

$$x = \lambda a + (1 - \lambda) b. \tag{10.95}$$

The conformal version of this line is

$$\begin{aligned} F(x) &= \big(\lambda^2 a^2 + 2\lambda(1-\lambda) a \cdot b + (1-\lambda)^2 b\big) n + 2\lambda a + 2(1-\lambda) b - \bar{n} \\ &= \lambda A + (1-\lambda) B + \tfrac{1}{2}\lambda(1-\lambda) A \cdot B\, n, \end{aligned} \tag{10.96}$$

and any multiple of this encodes the same point on the line. It is clear, then, that a conformal point X is a linear combination of A, B and n, subject to the constraint that $X^2 = 0$. This is summarised by

$$(A \wedge B \wedge n) \wedge X = 0, \qquad X^2 = 0. \tag{10.97}$$

So it is *trivectors* that represent lines in conformal geometry. This illustrates a general feature of the conformal model — geometric objects are represented by multivectors of one grade higher than their projective counterpart. The extra degree of freedom is absorbed by the constraint that $X^2 = 0$.

As stated above, if we apply a conformal transformation to a trivector representing a line, we must obtain a new line. But there is no reason to expect this to be straight. To see what else can result, consider a simple inversion in the origin. Suppose that (x_1, x_2) denote a pair of Cartesian coordinates for the Euclidean plane, and consider the line $x_1 = 1$. Points on the line have components $(1, x_2)$, with $-\infty \leq x_2 \leq +\infty$. The image of this line under an inversion in the origin has coordinates (x_1', x_2'), where

$$x_1' = \frac{1}{1 + x_2^2}, \qquad x_2' = \frac{x_2}{1 + x_2^2}. \tag{10.98}$$

It is now straightforward to show that

$$(x_1' - \tfrac{1}{2})^2 + (x_2')^2 = \left(\tfrac{1}{2}\right)^2. \tag{10.99}$$

Hence inversion of a line produces a *circle*, centred on $(1/2, 0)$ and with radius $1/2$.

It follows that a general trivector in conformal space can encode a circle, with a line representing the special case of infinite radius. This is entirely sensible, as three distinct points are required to specify a circle. The points define a plane, and any three non-collinear points in a plane specify a unique circle. So, given three points A_1, A_2, A_3, the circle through all three is defined by

$$A_1 \wedge A_2 \wedge A_3 \wedge X = 0, \tag{10.100}$$

together with the restriction (often unstated) that $X^2 = 0$. The trivector

$$L = A_1 \wedge A_2 \wedge A_3 \tag{10.101}$$

therefore encodes a unique circle in conformal geometry. The test that the points lie on a straight line is that the circle passes through the point at infinity,

$$L \wedge n = 0 \quad \Rightarrow \quad \text{straight line.} \tag{10.102}$$

This explains why our earlier derivation of the line through A_1 and A_2 led to the trivector $A_1 \wedge A_2 \wedge n$, which explicitly includes the point at infinity. Unlike tests for linear dependence, testing for zero in equation (10.102) is numerically acceptable. The reason is that the magnitude of $L \wedge n$ controls the deviation from straightness. If precision is limited, one can then define how close $L \wedge n$ should be to zero in order for the line to be treated as straight. This is quite different to linear independence, where the concept of 'nearly independent' makes no sense.

Given that a trivector L encodes a circle, we should expect to be able to extract the key geometric properties of the circle directly from L. In particular, we seek

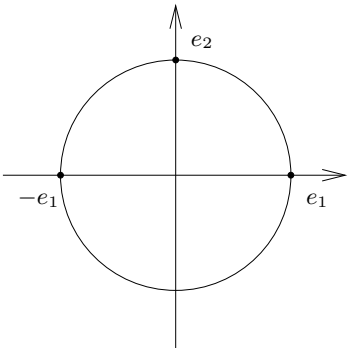

Figure 10.7 *The unit circle.* Three reference points are marked on the circle.

expressions for the centre and radius of the circle. (The plane containing the circle is specified by the 4-vector $L\wedge n$, as we explain in the following section.) Any circle in a plane can be mapped onto any other by a translation and a dilation. Under that latter we find that

$$L\wedge n \mapsto (D_\alpha L\tilde{D}_\alpha)\wedge n = \mathrm{e}^\alpha D_\alpha (L\wedge n)\tilde{D}_\alpha. \tag{10.103}$$

It follows that $(L\wedge n)^2$ scales as the inverse square of the radius. Next, consider the unit circle in the circle in the xy plane, and take as three points on the circle those shown in figure 10.7. The trivector for this circle is

$$L_0 = F(e_1)\wedge F(e_2)\wedge F(-e_1) = 16 e_1 e_2 \bar{e}. \tag{10.104}$$

It follows that

$$\frac{L_0^2}{(L_0\wedge n)^2} = -1, \tag{10.105}$$

which is (minus) the square of the radius of the unit circle. We can translate and dilate this into any circle we choose, so the radius ρ of the circle encoded by the trivector L is given by

$$\rho^2 = -\frac{L^2}{(L\wedge n)^2}. \tag{10.106}$$

This is a further illustration of how metric information is carried around in the homogeneous framework of the conformal model. If L represents a straight line we know that $L\wedge n = 0$, so the radius we obtain is infinite.

Similar reasoning produces a formula for the centre of a circle. Essentially the only objects we have to work with are L and n. If we form LnL for the case of

the unit circle we obtain

$$L_0 n L_0 \propto e_1 e_2 \bar{e} n \bar{e} e_1 e_2 = -\bar{n}. \tag{10.107}$$

But $\bar{n}$ is the null vector for the origin, so this expression has returned the desired point. Again, we can translate and dilate this result to obtain an arbitrary circle, and we find in general that the centre C of the circle L is obtained by

$$C = LnL. \tag{10.108}$$

We will see in section 10.5.5 that the operation $L \dots L$ generates a reflection in a circle. Equation (10.108) then says that the centre of a circle is the image of the point at infinity under a reflection in the circle.

10.4.3 4-vectors, spheres and planes

We can apply the same reasoning for lines and circles to the case of planes and spheres and, for mixed signature spaces, hyperboloids. Suppose initially that the points a, b, c define a plane in $\mathcal{V}(p, q)$, so that an arbitrary point in the plane is given by

$$x = \alpha a + \beta b + \gamma c, \qquad \alpha + \beta + \gamma = 1. \tag{10.109}$$

The conformal representation of x is

$$X = \alpha A + \beta B + \gamma C + \delta n, \tag{10.110}$$

where $A = f(a)$ etc., and

$$\delta = \tfrac{1}{2}(\alpha\beta A{\cdot}B + \alpha\gamma A{\cdot}C + \beta\gamma B{\cdot}C). \tag{10.111}$$

Varying α and β, together with the freedom to scale $F(x)$, now produces general null combinations of the vectors A, B, C and n. The equation for the plane can then be written

$$A{\wedge}B{\wedge}C{\wedge}n{\wedge}X = 0. \tag{10.112}$$

The plane passes through the points defined by A, B, C and the point at infinity n. We can therefore see that a general plane in conformal space is defined by four points.

If the four points in question do not lie on a (flat) plane, then the 4-vector formed from their outer product defines a *sphere*. To see this we again consider inversion in the origin, this time applied to the $x_1 = 1$ plane. A point on the plane has coordinates $(1, x_2, x_3)$, and under an inversion this maps to the point with coordinates

$$x_1' = \frac{1}{1 + x_2^2 + x_3^2}, \quad x_2' = \frac{y}{1 + x_2^2 + x_3^2}, \quad x_3' = \frac{z}{1 + x_2^2 + x_3^2}. \tag{10.113}$$

The new coordinates satisfy

$$\left(x_1' - \tfrac{1}{2}\right)^2 + (x_2')^2 + (x_3')^2 = \left(\tfrac{1}{2}\right)^2, \tag{10.114}$$

which is the equation of a sphere. Inversion thus interchanges planes and spheres. In particular, the point at infinity n is transformed to the origin $\bar{n}$ under inversion, which is now one of the points on the sphere.

Given any four distinct points $A_1, \ldots, A_4$, not all on a line or circle, the equation of the unique sphere through all four points is

$$A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge X = P \wedge X = 0, \tag{10.115}$$

so the sphere is defined by the 4-vector $P = A_1 \wedge A_2 \wedge A_3 \wedge A_4$. The sphere is flat (a plane) if it passes through the point at infinity, the test for which is

$$A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge n = P \wedge n = 0. \tag{10.116}$$

The 4-vector P contains all of the relevant geometric information for a sphere. The radius of the sphere ρ is given by

$$\rho^2 = \frac{P^2}{(P \wedge n)^2}, \tag{10.117}$$

as is easily confirmed for the case of the unit sphere, $P = e_1 e_2 e_3 \bar{e}$. Similarly, the centre of the sphere $C = F(c)$ is given by

$$C = PnP. \tag{10.118}$$

These formulae are the obvious generalisations of the results derived for circles.

10.5 Intersection and reflection in conformal space

One of the most significant advantages of the conformal approach to Euclidean geometry is the ease with which it solves complicated intersection problems. So, for example, finding the circle of intersection of two spheres is now no more complicated than finding the line of intersection of two planes. In addition, the concept of reflection is generalised in conformal space to include reflection in a sphere. This provides a very compact means of encoding the key concepts of inversive geometry.

10.5.1 Duality in conformal space

The concept of duality is key to intersecting objects in projective space, and the same is true in conformal space. Suppose that we start with the Euclidean plane, modelled in $\mathcal{G}(3,1)$. Duality in this algebra interchanges spacelike and timelike

bivectors. It also maps trivectors to vectors, and vice versa. A trivector encodes a line, or circle, so the dual of the circle C is a vector c, where

$$c = C^* = IC \tag{10.119}$$

and I is the pseudoscalar for $\mathcal{G}(3,1)$. The equation for the circle, $X \wedge C = 0$, can now be written in dual form and reduces to

$$X \cdot c = -I(X \wedge C) = 0. \tag{10.120}$$

The radius of the circle is now given by

$$\rho^2 = \frac{c^2}{(c \cdot n)^2}, \tag{10.121}$$

as the vector dual to a circle has positive signature. This picture provides us with an alternative view of the concept of a point as being a circle of zero radius.

Similar considerations hold for spheres in three-dimensional space. These are represented as 4-vectors in $\mathcal{G}(4,1)$, so their dual is a vector. We write

$$s = S^* = IS, \tag{10.122}$$

where I is the pseudoscalar, so that the equation of a sphere becomes

$$X \cdot s = I(X \wedge S) = 0. \tag{10.123}$$

The radius of the sphere is again given by

$$\rho^2 = \frac{s^2}{(s \cdot n)^2}, \tag{10.124}$$

so that points are spheres of zero radius. One can see that this is sensible by considering an alternative equation for a sphere. Suppose we are interested in the sphere with centre C and radius ρ^2. The equation for this can be written

$$-2\frac{X \cdot C}{X \cdot n\, C \cdot n} = \rho^2. \tag{10.125}$$

Rearranging, this equation becomes

$$X \cdot (2C + \rho^2 C \cdot n\, n) = 0, \tag{10.126}$$

and if C is in standard form, $C = F(c)$, we obtain

$$X \cdot (F(c) - \rho^2 n) = 0. \tag{10.127}$$

We can therefore identify $s = S^*$ with the vector $F(c) - \rho^2 n$, which neatly encodes the centre and radius of the sphere in a single vector. Whether the 4-vector S or its dual vector s is most useful depends on whether the sphere is specified by four points lying on it, or by its centre and radius. For a given sphere s we can now write

$$s = \lambda(2C + \rho^2 C \cdot n\, n). \tag{10.128}$$

It is then straightforward to confirm that the radius is given by equation (10.124). The centre of the circle can be recovered from

$$\frac{C}{C\cdot n} = \frac{s}{s\cdot n} - \frac{\rho^2}{2}n = \frac{sns}{2(s\cdot n)^2}. \tag{10.129}$$

The sns form for the centre of a sphere is dual to the SnS expression found in equation (10.118).

10.5.2 Intersection of two lines in a plane

As a simple example of intersection in the conformal model, consider the intersection of two lines in a Euclidean plane. The lines are described by trivectors L_1 and L_2 in $\mathcal{G}(3,1)$. The intersection is described by the bivector

$$B = (L_1^* \wedge L_2^*)^* = I(L_1 \times L_2), \tag{10.130}$$

where I is the conformal pseudoscalar. The bivector B can contain zero, one or two points, depending on the sign of its square, as described in section 10.4.1. This is to be expected, as distinct circles can intersect at a maximum of two points. If the lines are both straight, then one of the points of intersection will be at infinity, and $B \wedge n = 0$.

To verify this result, consider the case of two straight lines, both passing through the origin, and with the first line in the a direction and the second in the b direction. With suitable normalisation we can write

$$L_1 = aN, \qquad L_2 = bN, \tag{10.131}$$

where $N = e\bar{e}$. The intersection of L_1 and L_2 is controlled by

$$B = I\, a \wedge b \propto N \tag{10.132}$$

and the bivector N contains the null vectors n and $\bar{n}$. This confirms that the lines intersect at the origin and infinity. Applying conformal transformations to this result ensures that it holds for all lines in a plane, whether the lines are straight or circular. The formulae for L_1 and L_2 also show that their inner product is related to the angle between the lines,

$$\langle L_1 L_2 \rangle = a\cdot b. \tag{10.133}$$

We can therefore write

$$\cos(\theta) = \frac{\langle L_1 L_2 \rangle}{|L_1|\,|L_2|}, \tag{10.134}$$

where $|L| = \sqrt{(L^2)}$. This equation returns the angle between two lines. The quantity is invariant under the full conformal group, and not just the Euclidean group, because angles are conformal invariants. It follows that the same formula must hold even if L_1 and L_2 describe circles. The angle between two circles is

the angle made by their tangent vectors at the point of intersection. Two circles intersect at a right angle, therefore, if

$$\langle L_1 L_2 \rangle = 0. \tag{10.135}$$

This result can equally be expressed in terms of the dual vectors l_1 and l_2.

10.5.3 Intersection of a line and a surface

Now suppose that the 4-vector P defines a plane or sphere in three-dimensional Euclidean space, and we wish to find the point of intersection with a line described by the trivector L. The algebra proceeds entirely as expected and we arrive at the bivector

$$B = (P^* \wedge L^*)^* = (IP) \cdot L = I \langle PL \rangle_3. \tag{10.136}$$

This bivector can again describe zero, one or two points, depending on the sign of its square. This setup describes all possible intersections between lines or circles, and planes or spheres — an extremely wide range of applications. Precisely the same algebra enables us to answer whether a ring in space intersects a given plane, or whether a straight line passes through a sphere.

10.5.4 Surface intersections

Next, suppose we wish to intersect two surfaces in three dimensions. Suppose that these are spheres defined by the 4-vectors S_1 and S_2. Their intersection is described by the trivector

$$L = I(S_1 \times S_2). \tag{10.137}$$

This trivector directly encodes the circle formed from the intersection of two spheres. As with the bivector case, the sign of L^2 defines whether or not two surfaces intersect. If $L^2 > 0$ then the surfaces do intersect. If $L^2 = 0$ then the surfaces intersect at a point. Tests such as this are extremely helpful in graphics applications.

We can similarly express the intersection in terms of the dual vectors s_1 and s_2 as

$$L = I\, s_1 \wedge s_2. \tag{10.138}$$

As a check, the point X lies on both spheres if

$$X \cdot s_1 = X \cdot s_2 = 0. \tag{10.139}$$

It follows that

$$X \cdot (s_1 \wedge s_2) = X \cdot s_1\, s_2 - X \cdot s_2\, s_1 = 0. \tag{10.140}$$

The dual result is that $X\wedge(I\,s_1\wedge s_2)=0$, which confirms that X lies in the space defined by the trivector L.

10.5.5 Reflections in conformal space

At various points in previous sections we have obtained formulae which generate reflections. We now discuss these more systematically. In section 2.6 we established that the vector obtained by reflecting a in the hyperplane perpendicular to l, $l^2=1$, is $-lal$. But this formula assumes that the line and plane intersect at the origin. We seek a more general expression, valid for an arbitrary line and plane. Let P denote the plane and L the line we wish to reflect in the plane, then the obvious candidate for the reflected line L' is

$$L'=PLP. \tag{10.141}$$

(The sign of this is irrelevant in conformal space.) To verify that this is correct, suppose that L passes through the origin in the a direction,

$$L=aN_3 \tag{10.142}$$

and the plane P is defined by the origin and the directions b and c,

$$P=b\wedge cN. \tag{10.143}$$

In this case

$$L'=b\wedge c\,a\,b\wedge c\,N=\big(-(I_3\,b\wedge c)a(I_3\,b\wedge c)\big)N, \tag{10.144}$$

where I_3 is the three-dimensional pseudoscalar. This result achieves the required result. The vector a is reflected in the $b\wedge c$ plane to obtain the desired direction. The outer product with N then defines the line through the origin with the required direction. Equation (10.141) is correct at the origin, so therefore holds for all lines and planes, by conformal invariance.

There are a number of significant consequences of equation (10.141). The first is that it recovers the correct line in three dimensions without having to to find the point of reflection. The second is that it is straightforward to chain together multiple reflections by forming successive products with planes. In this way complicated reflections can be easily composed, all the time keeping track of the direction and position of the resultant line. A further consequence is that the same reflection formula must hold for higher dimensional objects. Suppose, for example, we wish to reflect the sphere S in the plane P. The result is

$$S'=PSP. \tag{10.145}$$

This type of equation is extremely useful in dealing with wave propagation, where a wavefront is modelled as a series of expanding spheres.

Conformal invariance of the reflection formula (10.141) ensures that the same

formula holds for reflection in a circle, or in a sphere. For example, suppose we wish to carry out a reflection in the unit circle in two-dimensional Euclidean space. The circle is defined by $L_0 = e_1 e_2 \bar{e}$, and the dual vector is

$$IL_0 = e. \tag{10.146}$$

Reflection in the unit circle is therefore performed by the operation

$$M \mapsto eMe. \tag{10.147}$$

This is an inversion, as discussed in section 10.3.3. In this manner, the main results of inversive geometry are easily formulated in terms of reflections in conformal space.

10.6 Non-Euclidean geometry

The sudden growth in the subject of geometry in the nineteenth century was stimulated in part by the discovery of geometries with very different properties to Euclidean space. These were obtained by a simple modification of Euclid's *parallel postulate*. For Euclidean geometry this states that, given any line l and a point P not on the line, there exists a unique line through P in the plane of l and P which does not meet l. This is then a line parallel to l. For many centuries this postulate was viewed as problematic, as it cannot be easily experimentally verified. As a result, mathematicians attempted to remove the parallel postulate by proving it from the remaining, uncontroversial, postulates of Euclidean geometry. This enterprise proved fruitless, and the reason why was discovered by Lobachevskii and Bolyai in the 1820s. One can replace the parallel postulate with a different postulate, and obtain a new, mathematically acceptable geometry.

There are in fact two alternative geometries one can obtain, by replacing the statement that there is a *single* line through P which does not intersect l with either an infinite number or zero. The case of an infinite number produces *hyperbolic* geometry, which is the non-Euclidean geometry constructed by Lobachevskii and Bolyai. (In this section 'non-Euclidean' usually refers to the hyperbolic case.) The case of zero lines produces spherical geometry. Intuitively, the spherical case corresponds to space curling up, so that all (straight) lines meet somewhere, and the hyperbolic case corresponds to space curving outwards, so that lines do not meet. From the more modern perspective of Riemannian geometry, we are talking about homogeneous, isotropic spaces, which have no preferred points or directions. These can have positive, zero or negative curvature, corresponding to spherical, Euclidean and hyperbolic geometries. Today, the question of which of these correctly describes the universe on the largest scales remains an outstanding problem in cosmology.

An extremely attractive feature of the conformal model of Euclidean geometry

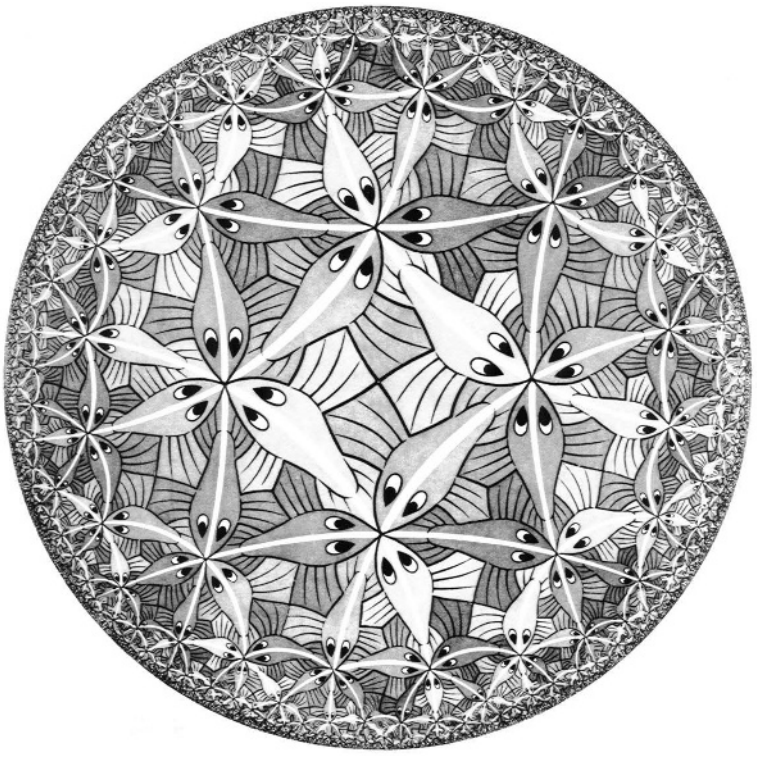

Figure 10.8 *Circle limit III* by Maurits Escher. ©2002 Cordon Art B.V., Baarn, Holland.

is that, with little modification, it can be applied to both hyperbolic and spherical geometries as well. In essence, the geometry reduces to a choice of the point at infinity, which in turn fixes the distance measure. This idea replaces the concept of the *absolute conic*, adopted in classical projective geometry as a means of imposing a distance measure. In this section we illustrate these ideas with a discussion of the conformal approach to planar hyperbolic geometry. As a concrete model of this we concentrate on the Poincaré disc. This version of hyperbolic geometry is mathematically very appealing, and also gives rise to some beautiful graphic designs, as popularised in the prints of Maurits Escher (see figure 10.8).

10.6.1 The Poincaré disc

The Poincaré disc $\mathcal{D}$ consists of the set of points in the plane a distance $r < 1$ from the origin. At first sight this may not appear to be homogeneous, but in

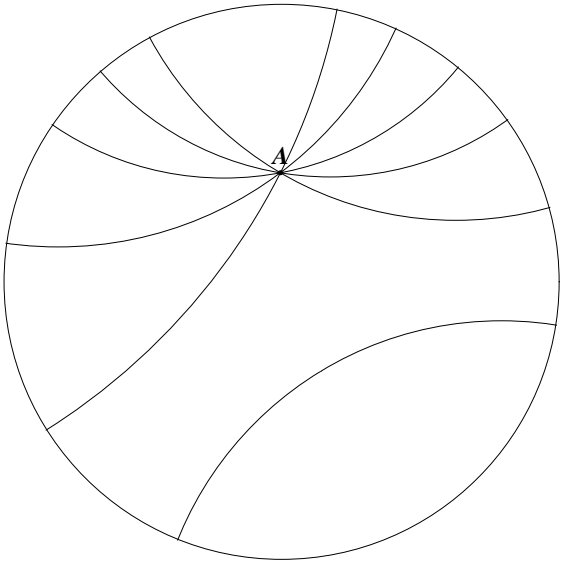

Figure 10.9 *The Poincaré disc.* Points inside the disc represent points in a hyperbolic space. A set of d-lines are also shown. These are (Euclidean) circles that intersect the unit circle at right angles. The d-lines through A illustrate the parallel postulate for hyperbolic geometry.

fact the nature of the geometry will ensure that there is nothing special about the origin. Note that points on the unit circle $r = 1$ are *not* included in this model of hyperbolic geometry. The key to this geometry is the concept of a non-Euclidean straight line. These are called d-lines, and represent geodesics in hyperbolic geometry. A d-line consists of a section of a Euclidean circle which intersects the unit circle at a right angle. Examples of d-lines are illustrated in figure 10.9. Given any two points in the Poincaré disc there is a unique d-line through them, which represents the 'straight' line between the points. It is now clear that for any point not on a given d-line l, there are an infinite number of d-lines through the point which do not intersect l.

We can now begin to encode these concepts in the conformal setting. We continue to denote points in the plane with homogeneous null vectors in precisely the same manner as the Euclidean case. Suppose, then, that X and Y are the conformal vectors representing two points in the disc. The set of all circles through these two points consists of trivectors of the form $X \wedge Y \wedge A$, where A is an additional point. But we require that the d-line intersects the unit circle at right angles. The unit circle is described by the trivector Ie, where I is the

pseudoscalar in $\mathcal{G}(3,1)$. If a line L is perpendicular to the unit circle it satisfies

$$(Ie)\cdot L = I(e\wedge L) = 0. \tag{10.148}$$

It follows that all d-lines contain a factor of e. The d-line through X and Y must therefore be described by the trivector

$$L = X\wedge Y\wedge e. \tag{10.149}$$

One can see now that a general scheme is beginning to emerge. Everywhere in the Euclidean treatment that the vector n appears it is replaced in hyperbolic geometry by the vector e. This vector represents the circle at infinity.

Given a pair of d-lines, they can either miss each other, or intersect at a point in the disc $\mathcal{D}$. If they intersect, the angle between the lines is given by the Euclidean formula

$$\cos(\theta) = \frac{L_1\cdot L_2}{|L_1|\,|L_2|}. \tag{10.150}$$

It follows that angles are preserved by a general conformal transformation in hyperbolic geometry. A non-Euclidean transformation takes d-lines to d-lines. The transformation must therefore map (Euclidean) circles to circles, while preserving orthogonality with e. The group of non-Euclidean transformations must therefore be the subgroup of the conformal group which leaves e invariant. This is confirmed in the following section, where we find the appropriate distance measure for non-Euclidean geometry.

The fact that the point at infinity is represented by e, as opposed to n in the Euclidean counterpart, provides an additional operation in non-Euclidean geometry. This is inversion in e:

$$X \mapsto eXe. \tag{10.151}$$

As all non-Euclidean transformations leave e invariant, all geometric relations remain unchanged under this inversion. Geometrically, the interpretation of the inversion is quite clear. It maps everything inside the Poincaré disc to a 'dual' version outside the disc. In this dual space incidence relations and distances are unchanged from their counterparts inside the disc.

10.6.2 Non-Euclidean translations and distance

The key to finding the correct distance measure in non-Euclidean geometry is to first generalise the concept of a translation. Given points X and Y we know that the d-line connecting them is defined by $X\wedge Y\wedge e$. This is the non-Euclidean concept of a straight line. A non-Euclidean translation must therefore move points along this line. Such a transformation must take X to Y, but must also leave e invariant. The generator for such a transformation is the bivector

$$B = (X\wedge Y\wedge e)e = Le, \tag{10.152}$$

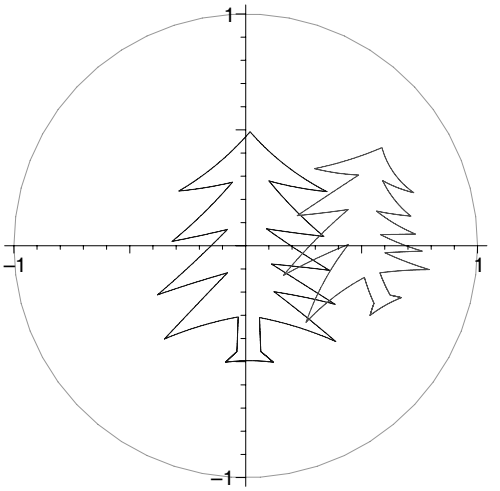

Figure 10.10 *A non-Euclidean translation.* The figure near the origin is translated via a boost to give the distorted figure on the right. This distortion in the Poincaré disc is one way of visualising the effect of a Lorentz boost in spacetime.

where $L = X \wedge Y \wedge e$. We find immediately that

$$B^2 = L^2 > 0, \tag{10.153}$$

so non-Euclidean translations are *hyperbolic* transformations, as one might expect. An example of such a translation is shown in figure 10.10.

We next define

$$\hat{B} = \frac{B}{|B|}, \qquad \hat{B}^2 = 1, \tag{10.154}$$

so that we can write

$$Y = \mathrm{e}^{\alpha \hat{B}/2} X \mathrm{e}^{-\alpha \hat{B}/2}. \tag{10.155}$$

By varying α we obtain the set of points along the d-line through X and Y. To obtain a distance measure, we first require a formula for α. If we decompose X into

$$X = X\hat{B}^2 = X \cdot \hat{B}\,\hat{B} + X \wedge \hat{B}\,\hat{B} \tag{10.156}$$

we obtain

$$Y = X \wedge \hat{B}\,\hat{B} + \cosh(\alpha)\, X \cdot \hat{B}\,\hat{B} - \sinh(\alpha)\, X \cdot \hat{B}. \tag{10.157}$$

The right-hand side must give zero when contracted with Y, so

$$\langle X \wedge \hat{B}\,\hat{B} \wedge Y \rangle + \cosh(\alpha) \langle X \cdot \hat{B}\,\hat{B} \cdot Y \rangle + \sinh(\alpha)\, (X \wedge Y) \cdot \hat{B} = 0. \tag{10.158}$$

To simplify this equation we first find

$$X\wedge\hat{B} = \frac{X\wedge(X\wedge Y\wedge e\, e)}{|B|} = \frac{e\cdot X\, L}{|L|} \tag{10.159}$$

and

$$(X\wedge Y)\cdot\hat{B} = \frac{L^2}{|B|} = |L|. \tag{10.160}$$

It follows that

$$e\cdot X\, e\cdot Y + \cosh(\alpha)(X\cdot Y - e\cdot X\, e\cdot Y) + \sinh(\alpha)\,|L| = 0, \tag{10.161}$$

the solution to which is

$$\cosh(\alpha) = 1 - \frac{X\cdot Y}{X\cdot e\, Y\cdot e}. \tag{10.162}$$

The half-angle formula is more relevant for the distance measure, and we find that

$$\sinh^2(\alpha/2) = -\frac{X\cdot Y}{2X\cdot e\, Y\cdot e}. \tag{10.163}$$

This closely mirrors the Euclidean expression, with n replaced by e.

There are a number of obvious properties that a distance measure must satisfy. Among these is the additive property that

$$d(X_1, X_2) + d(X_2, X_3) = d(X_1, X_3) \tag{10.164}$$

for any three points X_1, X_2, X_3 in this order along a d-line. Returning to the translation formula of equation (10.155), suppose that Z is a third point along the line, beyond Y. We can write

$$Z = \mathrm{e}^{\beta\hat{B}/2} Y \mathrm{e}^{-\beta\hat{B}/2} = \mathrm{e}^{(\alpha+\beta)\hat{B}} X \mathrm{e}^{-(\alpha+\beta)\hat{B}/2}. \tag{10.165}$$

Clearly it is hyperbolic angles that must form the appropriate distance measure. No other function satisfies the additive property. We therefore define the non-Euclidean distance by

$$d(x, y) = 2\sinh^{-1}\left(-\frac{X\cdot Y}{2X\cdot e\, Y\cdot e}\right)^{1/2}. \tag{10.166}$$

In terms of the position vectors x and y in the Poincaré disc we can write

$$d(x, y) = 2\sinh^{-1}\left(\frac{|x-y|^2}{(1-x^2)(1-y^2)}\right)^{1/2}, \tag{10.167}$$

where the modulus refers to the Euclidean distance. The presence of the arcsinh function in the definition of distance reflects the fact that, in hyperbolic geometry, generators of translations have positive square and the appropriate distance measure is the hyperbolic angle. Similarly, in spherical geometry translations correspond to rotations, and it is the trigonometric angle which plays the role

of distance. Euclidean geometry is therefore unique in that the generators of translations are *null* bivectors. For these, combining translations reduces to the addition of bivectors, and hence we recover the standard definition of Euclidean distance.

10.6.3 Metrics and physical units

The derivation of the non-Euclidean distance formula of equation (10.166) forces us to face an issue that has been ignored to date. Physical distances are dimensional quantities, whereas our formulae for distances in both Euclidean and non-Euclidean geometries are manifestly dimensionless, as they are homogeneous in X. To resolve this we cannot just demand that the vector x has dimensions, as this would imply that the conformal vector X contained terms of mixed dimensions. Neither can this problem be circumvented by assigning dimensions of distance to $\bar{n}$ and (distance)$^{-1}$ to n, as then e has mixed dimensions, and the non-Euclidean formula of (10.166) is non-sensical.

The resolution is to introduce a fundamental length scale, λ, which is a positive scalar with the dimensions of length. If the vector x has dimensions of length, the conformal representation is then given by

$$X = \frac{1}{2\lambda^2}\left(x^2 n + 2\lambda x - \lambda^2 \bar{n}\right). \tag{10.168}$$

This representation ensures that X remains dimensionless, and is nothing more than the conformal representation of x/λ. Physical distances can then be converted into a dimensionally meaningful form by including appropriate factors of λ. Curiously, the introduction of λ into the spacetime conformal model has many similarities to the introduction of a cosmological constant $\Lambda = \lambda^2$.

We can make contact with the metric encoding of distance by finding the infinitesimal distance between the points x and $x + dx$. This defines the line element

$$ds^2 = 4\lambda^4 \frac{dx^2}{(\lambda^2 - x^2)^2}, \tag{10.169}$$

where the factors of λ have been included and x is assumed to have dimensions of distance. This line element is more often seen in polar coordinates, where it takes the form

$$ds^2 = \frac{4\lambda^4}{(\lambda^2 - r^2)^2}(dr^2 + r^2 d\theta^2). \tag{10.170}$$

This is the line element for a space of constant negative curvature, expressed in terms of conformal coordinates. The coordinates are conformal because the line element is that of a flat space multiplied by a scaling function. The geodesics in this geometry are precisely the d-lines in the Poincaré disc. The Riemann curvature for this metric shows that the space has uniform negative curvature,

so the space is indeed homogeneous and isotropic — there are no preferred points or directions. The centre of the disc is not a special point, and indeed it can be translated to any other point by 'boosting' along a d-line.

10.6.4 Midpoints and circles in non-Euclidean geometry

Now that we have a conformal encoding of a straight line and of distance in non-Euclidean geometry, we can proceed to discuss concepts such as the midpoint of two points, and of the set of points a constant distance from a given point (a non-Euclidean circle). Suppose that A and B are the conformal vectors of two points in the Poincaré disc. Their midpoint C lies on the line $L = A \wedge B \wedge e$ and is equidistant from both A and B. The latter condition implies that

$$\frac{C \cdot A}{C \cdot e \, A \cdot e} = \frac{C \cdot B}{C \cdot e \, B \cdot e}. \tag{10.171}$$

Both of the conditions for C are easily satisfied by setting

$$C = \frac{A}{2A \cdot e} + \frac{B}{2B \cdot e} + \alpha e, \tag{10.172}$$

where α must be chosen such that $C^2 = 0$. Normalising to $C \cdot e = -1$ we find that the midpoint is

$$C = -\frac{1}{\sqrt{1+\delta}} \left(\frac{A}{2A \cdot e} + \frac{B}{2B \cdot e} + \left(\sqrt{1+\delta} - 1\right) e \right), \tag{10.173}$$

where

$$\delta = -\frac{A \cdot B}{2A \cdot e \, B \cdot e}. \tag{10.174}$$

An equation such as this is rather harder to achieve without access to the conformal model.

Next suppose we wish to find the set of points a constant (non-Euclidean) distance from the point C. This defines a non-Euclidean circle with centre C. From equation (10.166), any point X on the circle must satisfy

$$-\frac{X \cdot C}{2X \cdot e \, C \cdot e} = \text{constant} = \alpha^2, \tag{10.175}$$

so that the radius is $\sinh^{-1}(\alpha)$. It follows that

$$X \cdot (C + 2\alpha^2 C \cdot e \, e) = 0. \tag{10.176}$$

If we define s by

$$s = C + 2\alpha^2 C \cdot e \, e \tag{10.177}$$

we see that $s^2 > 0$, and the circle is defined by $X \cdot s = 0$. But this is precisely the formula for a circle in Euclidean geometry, so non-Euclidean circles still appear as ordinary circles when plotted in the Poincaré disc. The only difference is the

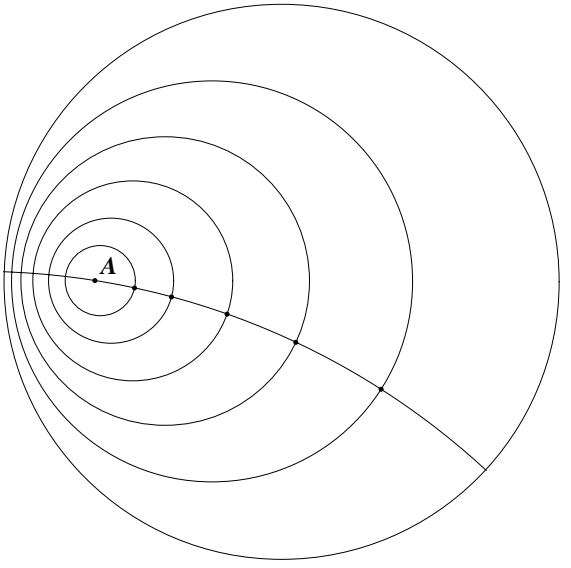

Figure 10.11 *Non-Euclidean circles.* A series of non-Euclidean circles with differing radii are shown, all about the common centre A. A d-line through A is also shown. This intersects each circle at a right angle.

interpretation of their centre. The Euclidean centre of the circle s, defined by sns, does not coincide with the non-Euclidean centre C. This is illustrated in figure 10.11.

Suppose that A, B and C are three points in the Poincaré disc. We can still define the line L through these points by

$$L = A \wedge B \wedge C, \tag{10.178}$$

and this defines the circle through the three points regardless of the geometry we are working in. All that is different in the two geometries is the position of the midpoint and the size of the radius. The test that the three points lie on a d-line is simply that $L \wedge e = 0$. Again, the Euclidean formula holds, but with n replaced by e. Similar comments apply to other operations in conformal space, such as reflection. Given a line L, points are reflected in this line by the map $X \mapsto LXL$. This formula is appropriate in both Euclidean and non-Euclidean geometry. In the non-Euclidean case it is not hard to verify that LXL corresponds to first finding the d-line through X intersecting L at right angles, and then finding the point on this line an equal non-Euclidean distance on the other side. This is as one would expect for the definition of reflection in a line.

10.6.5 A unified framework for geometry

We have so far seen how Euclidean and hyperbolic geometries can both be handled in terms of null vectors in conformal space. The key concept is the vector representing the point at infinity, which remains invariant under the appropriate symmetry group. The full conformal group of a space with signature (p, q) is the orthogonal group $O(p+1, q+1)$. The group of Euclidean transformations is the subgroup of $O(p+1, q+1)$ that leaves the vector n invariant. The hyperbolic group is the subgroup of $O(p+1, q+1)$ which leaves e invariant. For the case of planar geometry, with signature $(2, 0)$, the hyperbolic group is $O(2, 1)$. The Killing form for this group is non-degenerate (see chapter 11), which makes hyperbolic geometry a useful way of compactifying a flat space.

The remaining planar geometry to consider is spherical geometry. By now, it should come as little surprise that spherical geometry is handled in the conformal framework in terms of transformations which leave the vector $\bar{e}$ invariant. For the case of the plane, the conformal algebra has signature $(3, 1)$, with $\bar{e}$ the basis vector with negative signature. The subgroup of the conformal group which leaves $\bar{e}$ invariant is therefore the orthogonal group $O(3, 0)$, which is the group one expects for a 2-sphere. The distance measure for spherical geometry is

$$d(x, y) = 2\lambda \sin^{-1}\left(-\frac{X \cdot Y}{2X \cdot \bar{e}\, Y \cdot \bar{e}}\right)^{1/2}, \tag{10.179}$$

with $\bar{e}$ replacing n in the obvious manner. To see that this expression is correct, suppose that we write

$$\frac{X}{X \cdot \bar{e}} = \hat{x} - \bar{e}, \tag{10.180}$$

where $\hat{x}$ is a unit vector built in the three-dimensional space spanned by the vectors e_1, e_2 and e. With $Y/Y \cdot \bar{e}$ written in the same way we find that

$$-\frac{X \cdot Y}{2X \cdot \bar{e}\, Y \cdot \bar{e}} = \frac{1 - \hat{x} \cdot \hat{y}}{2} = \sin^2(\theta/2), \tag{10.181}$$

where θ is the angle between the unit vectors on the 2-sphere. The distance measure is then precisely the angle θ multiplied by the dimensional quantity λ, which represents the radius of the sphere.

Conformal geometry provides a unified framework for the three types of planar geometry because in all cases the conformal groups are the same. That is, the group of transformations of sphere that leave angles in the sphere unchanged is the same as for the plane and the hyperboloid. In all cases the group is $O(3, 1)$. The geometries are then recovered by a choice of distance measure. In classical projective geometry the distance measure is defined by the introduction of the *absolute conic*. All lines intersect this conic in a pair of points. The distance between two points A and B is then found from the four-point ratio between A, B, and the two points of intersection of the line through A and B and the absolute

conic. In this way all geometries are united in the framework of projective geometry. But there is a price to pay for this scheme — all coordinates have to be complex, to ensure that all lines intersect the conic in two points. Recovering a real geometry is then rather clumsy. In addition, the conformal group is not a subgroup of the projective group, so much of the elegant unity exhibited by the three geometries is lost. Conformal geometry is a more powerful framework for a unified treatment of these geometries. Furthermore, the conformal approach can be applied to spaces of any dimension with little modification. Trivectors represent lines and circles, 4-vectors represent planes and spheres, and so on.

So far we have restricted ourselves to a single view of the various geometries, but the discussion of the sphere illustrates that there are many different ways of representing the underlying geometry. To begin with, we have plotted points on the Euclidean plane according the the formula

$$x = -\frac{X \wedge N}{X \cdot n} N, \tag{10.182}$$

where $N = e\bar{e}$. This is the natural scheme for plotting on a Euclidean piece of paper, as it ensures that the angle between lines on the paper is the correct angle in each of the three geometries. Euclidean geometry plotted in this way recovers the obvious standard picture of Euclidean geometry. Hyperbolic geometry led to the Poincaré disc model, in which hyperbolic lines appear as circles. For spherical geometry the 'straight lines' are great circles on a sphere. On the plane these also plot as circles. This time the condition is that all circles intersect the unit circle at antipodal points. This then defines the spherical line between two points (see figure 10.12). This view of spherical geometry is precisely that obtained from a stereographic projection of the sphere onto the plane. This is not a surprise, as the conformal model was initially constructed in terms of a stereographic projection, with the $\bar{e}$ vector then enabling us to move to a homogeneous framework. In this representation of spherical geometry the map

$$X \mapsto \bar{e} X \bar{e} \tag{10.183}$$

is a symmetry operation. This maps points to their antipodal opposites on the sphere. In the planar view this transformation is an inversion in the unit circle, followed by a reflection in the origin.

We now have three separate geometries, all with conformal representations in the plane such that the true angle between lines is the same as that measured on the plane. The price for such a representation is that straight lines in spherical and hyperbolic geometries do not appear straight in the plane. But we could equally choose to replace the map of equation (10.182) with an alternative rule of how to plot the null vector X on a planar piece of paper. The natural alternatives to consider are replacing the vector n with e and $\bar{e}$. In total we then have three different planar realisations of each of the two-dimensional geometries. First,

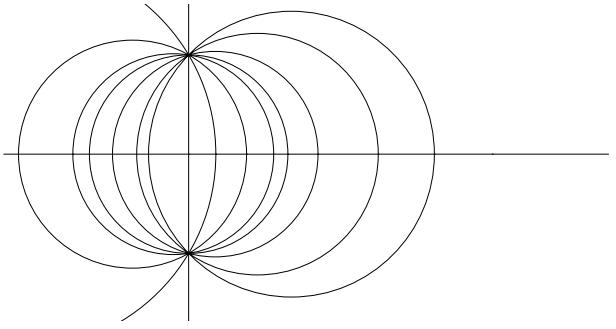

Figure 10.12 *Stereographic view of spherical geometry.* All great circles on the 2-sphere project onto circles in the plane which intersect the unit circle (shown in bold) at antipodal points. A series of such lines are shown.

suppose we define

$$y = \frac{X \wedge N}{X \cdot e} N. \tag{10.184}$$

In terms of the vector x we have

$$y = \frac{2x}{1 - x^2}, \tag{10.185}$$

which represents a radial rescaling. Euclidean straight lines now appear as hyperbolae or ellipses, depending on whether or not the original line intersected the disc. If the line intersected the disc then the map of equation (10.185) has two branches and defines a hyperbola. If the line misses the disc then an ellipse is obtained. In all cases the image lines pass through the origin, as this is the image of the point at infinity.

The fact that the map of equation (10.185) is two-to-one means it has little use as a version of Euclidean geometry. It is better suited to hyperbolic geometry, as one might expect, as the Poincaré disc is now mapped onto the entire plane. Hyperbolic straight lines now appear as (single-branch) hyperbolae on the Euclidean page, all with their asymptotes crossing at the origin. If the dual space outside the disc is included in the map, then this generates the second branch of each hyperbola. Points then occur in pairs, with each point paired with its image under reflection in the origin. Finally, we can consider spherical geometry as viewed on a plane through the map of equation (10.185). This defines a standard projective map between a sphere and the plane. Antipodal points on the sphere define the same point on the plane and spherical straight lines appear as straight lines.

Similarly, we can consider plotting vectors in the plane according to

$$y = -\frac{X \wedge N}{X \cdot \bar{e}} N = -\frac{F(x) \wedge N}{F(x) \cdot \bar{e}} N \tag{10.186}$$

or in terms of the vector x

$$y = \frac{2x}{1 + x^2}. \tag{10.187}$$

This defines a one-to-one map of the unit disc onto itself, and a two-to-one map of the entire plane onto the disc. Euclidean straight lines now appear plotted as ellipses inside the unit disc. This construction involves forming a stereographic projection of the plane onto the 2-sphere, so that lines map to circles on the sphere. The sphere is then mapped onto the plane by viewing from above, so that circles on the sphere map to ellipses. All ellipses pass through the origin, as this is the image of the point at infinity.

Similar comments apply to spherical geometry. Spherical lines are great circles on the sphere, and viewed in the plane according to equation (10.187) great circles appear as ellipses centred on the origin and touching the unit circle at their endpoints. The two-to-one form of the projection means that circle intersections are not faithfully represented in the disc as some of the apparent intersections are actually caused by points on opposite sides of the plane. Finally, we consider plotting hyperbolic geometry in the view of equation (10.187). The disc maps onto itself, so we do have a faithful representation of hyperbolic geometry. This is a representation in which hyperbolic lines appear straight on the page, though angles are not rendered correctly, and non-Euclidean circles appear as ellipses.

As well as viewing each geometry on the Euclidean plane, we can also picture the geometries on a sphere or a hyperboloid. The spherical picture is obtained in equation (10.180), and the hyperboloid view is similarly obtained by setting

$$\frac{X}{X \cdot e} = \hat{x} + e, \tag{10.188}$$

where $\hat{x}^2 = -1$. The set of $\hat{x}$ defines a pair of hyperbolic sheets in the space defined by the vectors $\{e_1, e_2, \bar{e}\}$. The fact that two sheets are obtained explains why some views of hyperbolic geometry end up with points represented twice. So, as well as three geometries (defined by a transformation group) and a variety of plotting schemes, we also have a choice of space to draw on, providing a large number of alternative schemes for studying the three geometries. At the back of all of this is a single algebraic scheme, based on the geometric algebra of conformal space. Any algebraic result involving products of null vectors immediately produces a geometric theorem in each geometry, which can be viewed in a variety of different ways.

10.7 Spacetime conformal geometry

As a final application of the conformal approach to geometry we turn to spacetime. The conformal geometric algebra for a spacetime with signature $(1,3)$ is the six-dimensional algebra with signature $(2,4)$. The algebra $\mathcal{G}(2,4)$ contains 64 terms, which decompose into graded subspaces of dimensions 1, 6, 15, 20, 15, 6 and 1. As a basis for this space we use the standard spacetime algebra basis $\{\gamma_\mu\}$, together with the additional vectors $\{e, \bar{e}\}$. The pseudoscalar I is defined by

$$I = \gamma_0\gamma_1\gamma_2\gamma_3 e\bar{e}. \tag{10.189}$$

This has negative norm, $I^2 = -1$. The conformal algebra allows us to simply encode ideas such as closed circles in spacetime, or light-spheres centred on an arbitrary point.

The conformal algebra of spacetime also arises classically in a slightly different setting. In conformal geometry, circles and spheres are represented homogeneously as trivectors and 4-vectors. These are unoriented because L and $-L$ are used to encode the same object. A method of dealing with oriented spheres was developed by Sophus Lie and is called Lie sphere geometry. A sphere in three dimensions can be represented by a vector s in the conformal algebra $\mathcal{G}(4,1)$, with $s^2 > 0$. Lie sphere geometry is obtained by introducing a further basis vector of negative signature, f, and replacing s by the null vector

$$\bar{s} = s + |s|f, \qquad \bar{s}^2 = 0. \tag{10.190}$$

Now the spheres encoded by s and $-s$ have different representations as null vectors in a space of signature $(4,2)$. This algebra is ideally suited to handling the contact geometry of spheres. The signature shows that this space is isomorphic to the conformal algebra of spacetime, so in a sense the introduction of the vector f can be thought of as introducing a time direction. A sphere can then be viewed as a light-sphere allowed to grow for a certain time. Orientation for spheres is then handled by distinguishing between incoming and outgoing light-spheres.

The conformal geometry of spacetime is a rich and important subject. The Poincaré group of spacetime translations and rotations is a subgroup of the full conformal group, but in a number of subjects in theoretical physics, including supersymmetry and supergravity, it is the full conformal group that is relevant. One reason is that conformal symmetry is present in most massless theories. This symmetry then has consequences that can carry over to the massive regime. We will not develop the classical approach to spacetime conformal geometry further here. Instead, we concentrate on an alternative route through to conformal geometry, which unites the multiparticle spacetime algebra of chapter 9 with the concept of a *twistor*.

10.7.1 The spacetime conformal group

For most of this chapter we have avoided detailed descriptions of the relationships between the groups involved in the geometric algebra formulation of conformal geometry. For the following, however, it is helpful to have a clearer picture of precisely how the various groups fit together. The subject of Lie groups in general is discussed in chapter 11. The spacetime conformal group $C(1,3)$ consists of spacetime maps $x \mapsto f(x)$ that preserve angles. This is the definition first encountered in section 10.3. The group of orthogonal transformations $O(2,4)$ is a double-cover representation of the conformal group, because in conformal space both X and $-X$ represent the same spacetime point. As with Lorentz transformations, we are typically interested in the restricted conformal group. This consists of transformations that preserve orientation and time sense, and contains translations, proper orthochronous rotations, dilations and special conformal transformations. The restricted orthogonal group, $SO^+(2,4)$, is a double-cover representation of the restricted conformal group.

We can form a double-cover representation of $SO^+(2,4)$ by writing all restricted orthogonal transformations as rotor transformations $a \mapsto Ra\tilde{R}$. The group of conformal rotors, denoted $\mathrm{spin}^+(2,4)$, is therefore a four-fold covering of the restricted conformal group. The rotor group in $\mathcal{G}(2,4)$ is isomorphic to the Lie group $\mathrm{SU}(2,2)$. It follows that the action of the restricted conformal group can be represented in terms of complex linear transformations of four-dimensional vectors, in a complex space of signature $(2,2)$. This is the basis of the *twistor* program, initiated by Roger Penrose. Twistors were introduced as objects describing the geometry of spacetime at a 'pre-metric' level, one of the aims being to provide a route to a quantum theory of gravity. Instead of points and a metric, twistors represent incidence relations between null rays. Spacetime points and their metric relations then emerge as a secondary concept, corresponding to the points of intersection of null lines.

As a first step in understanding the twistor program, we establish a concrete representation of the conformal group within the spacetime algebra. The key to this is the observation that the spinor inner product

$$\langle\tilde{\psi}\phi\rangle_q = \langle\tilde{\psi}\phi\rangle - \langle\tilde{\psi}\phi I\boldsymbol{\sigma}_3\rangle I\boldsymbol{\sigma}_3 \tag{10.191}$$

defines a complex space with precisely the required metric. The complex structure is represented by right-multiplication by combinations of 1 and $I\boldsymbol{\sigma}_3$, as discussed in chapter 8. We continue to refer to ψ and ϕ as spinors, as they are acted on by a spin representation of the restricted conformal group. To establish a representation in terms of operators on ψ, we first form a representation of the bivectors in $\mathcal{G}(2,4)$ as

$$\begin{aligned} e\gamma_\mu &\leftrightarrow \gamma_\mu\psi\gamma_0 I\boldsymbol{\sigma}_3 = \gamma_\mu\psi I\gamma_3, \\ \bar{e}\gamma_\mu &\leftrightarrow I\gamma_\mu\psi\gamma_0. \end{aligned} \tag{10.192}$$

A representation of the even subalgebra of $\mathcal{G}(2,4)$, and hence an arbitrary rotor, can be constructed from these bivectors. The representation of each of the operations in the restricted conformal group can now be constructed from the rotors found in section 10.3. We use the same symbol for the spinor representation of the transformations as the vector case. A translation by the vector a has the spin representation

$$T_a(\psi) = \psi + a\psi I\gamma_3 \tfrac{1}{2}\left(1 + \boldsymbol{\sigma}_3\right). \tag{10.193}$$

The spinor inner product of equation (10.191) is invariant under this transformation. To confirm this, suppose that we set

$$\psi' = T_a(\psi) \quad \text{and} \quad \phi' = T_a(\phi). \tag{10.194}$$

The quantum inner product contains the terms

$$\begin{aligned}\langle\tilde{\psi}'\phi'\rangle &= \langle\big(\phi + a\phi I\gamma_3 \tfrac{1}{2}\left(1 + \boldsymbol{\sigma}_3\right)\big)\big(\tilde{\psi} - \tfrac{1}{2}\left(1 - \boldsymbol{\sigma}_3\right) I\gamma_3\tilde{\psi}a\big)\rangle \\ &= \langle\tilde{\psi}\phi\rangle\end{aligned} \tag{10.195}$$

and

$$\begin{aligned}\langle\tilde{\psi}'\phi' I\boldsymbol{\sigma}_3\rangle &= \langle\big(\phi + a\phi I\gamma_3 \tfrac{1}{2}\left(1 + \boldsymbol{\sigma}_3\right)\big) I\boldsymbol{\sigma}_3\big(\tilde{\psi} - \tfrac{1}{2}\left(1 - \boldsymbol{\sigma}_3\right) I\gamma_3\tilde{\psi}a\big)\rangle \\ &= \langle\tilde{\psi}\phi I\boldsymbol{\sigma}_3\rangle.\end{aligned} \tag{10.196}$$

It follows that

$$\langle\tilde{\psi}'\phi'\rangle_q = \langle\tilde{\psi}\phi\rangle_q, \tag{10.197}$$

as expected.

The spinor representation of a rotation about the origin is precisely the spacetime algebra rotor, so we can write

$$R_0(\psi) = R\psi, \tag{10.198}$$

where R_0 denotes a rotation in the origin, and R is a spacetime rotor. Rotations about arbitrary points are constructed from combinations of translations and rotations. The dilation $x \mapsto \exp(\alpha)x$ has the spinor representation

$$D_\alpha(\psi) = \psi \mathrm{e}^{\alpha\boldsymbol{\sigma}_3/2}. \tag{10.199}$$

This represents a dilation in the origin. Dilations about a general point are also obtained from a combination of translations and a dilation in the origin. The representation of the restricted conformal group is completed by the special conformal transformations, which are represented by

$$K_a(\psi) = \psi - a\psi I\gamma_3 \tfrac{1}{2}\left(1 - \boldsymbol{\sigma}_3\right). \tag{10.200}$$

It is a routine exercise to confirm that the preceding operations do form a spin representation of the restricted conformal group.