Chapter 17

Spinors

17.1 Introducing spinors

Spinors are mathematical entities somewhat like tensors, that allow a more general treatment of the notion of invariance under rotation and Lorentz boosts. To every tensor of rank k there corresponds a spinor of rank $2k$, and some kinds of tensor can be associated with a spinor of the same rank. For example, a general 4-vector would correspond to a Hermitian spinor of rank 2, which can be represented by a 2×2 Hermitian matrix of complex numbers. A null 4-vector can also be associated with a spinor of rank 1, which can be represented by a complex vector with two components. We shall see why in the following.

Spinors can be used without reference to relativity, but they arise naturally in discussions of the Lorentz group. One could say that a spinor is the most basic sort of mathematical object that can be Lorentz-transformed. The main facts about spinors are given in the box. This summary is placed here rather than at the end of the chapter in order to help the reader follow the main thread of the argument.

It appears that Klein originally designed the spinor to simplify the treatment of the classical spinning top in 1897. The more thorough understanding of spinors as mathematical objects is credited to Élie Cartan in 1913. They are closely related to Hamilton's quaternions (about 1845).

Spinors began to find a more extensive role in physics when it was discovered that electrons and other particles have an intrinsic form of angular momentum now called 'spin', and the behaviour of this type of angular momentum is correctly captured by the mathematics discovered by Cartan. Pauli formalized this connection in a non-relativistic (i.e. low velocity) context, modelling the electron spin using a two-component complex vector, and introducing the *Pauli*

Spinor summary. A rank 1 spinor can be represented by a two-component complex vector, or by a null 4-vector, angle and sign. The spatial part can be pictured as a flagpole with a rigid flag attached.
The 4-vector is obtained from the 2-component complex vector by

$$\begin{aligned} \mathsf{U}^\mu &= \langle u|\,\sigma^\mu\,|u\rangle \quad \text{if } \boldsymbol{u} \text{ is a contraspinor (“right-handed”)} \\ \mathsf{U}_\mu &= \langle \tilde{u}|\,\sigma^\mu\,|\tilde{u}\rangle \quad \text{if } \tilde{\boldsymbol{u}} \text{ is a cospinor (“left handed”).} \end{aligned}$$

Any 2×2 matrix Λ with unit determinant Lorentz-transforms a spinor. Such matrices can be written

$$\Lambda = \exp\left(i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2 - \boldsymbol{\sigma}\cdot\boldsymbol{\rho}/2\right)$$

where ρ is rapidity. If Λ is unitary the transformation is a rotation in space; if Λ is Hermitian it is a boost.
If $\boldsymbol{s}' = \Lambda(v)\boldsymbol{s}$ is the Lorentz transform of a right-handed spinor, then under the same change of reference frame a left-handed spinor transforms as $\tilde{\boldsymbol{s}}' = (\Lambda^\dagger)^{-1}\tilde{\boldsymbol{s}} = \Lambda(-v)\tilde{\boldsymbol{s}}$.
The Weyl equations may be obtained by considering $(\mathsf{W}^\alpha\sigma_\alpha)\boldsymbol{w}$. This combination is zero in all frames. Applied to a spinor $\boldsymbol{w}$ representing energy-momentum it reads

$$\begin{aligned} (E/c - \mathbf{p}\cdot\boldsymbol{\sigma})\boldsymbol{w} &= 0 \\ (E/c + \mathbf{p}\cdot\boldsymbol{\sigma})\tilde{\boldsymbol{w}} &= 0. \end{aligned}$$

These equations are not parity-invariant. If both the energy-momentum and the spin of a particle can be represented simultaneously by the same spinor, then the particle is massless and the sign of its helicity is fixed.
A Dirac spinor $\Psi = (\phi_R, \phi_L)$ is composed of a pair of spinors, one of each handedness. From the two associated null 4-vectors one can extract two orthogonal non-null 4-vectors

$$\begin{aligned} \mathsf{U}^\mu &= \Psi^\dagger\gamma^0\gamma^\mu\Psi, \\ \mathsf{W}^\mu &= \Psi^\dagger\gamma^0\gamma^\mu\gamma^5\Psi, \end{aligned}$$

where γ^μ, γ^5 are the Dirac matrices. With appropriate normalization factors these can represent the 4-velocity and 4-spin of a particle.
Starting from a frame in which $\mathsf{U}^i = 0$ (i.e. the rest frame), the result of a Lorentz boost to a general frame can be written

$$\begin{pmatrix} -m & E + \boldsymbol{\sigma}\cdot\mathbf{p} \\ E - \boldsymbol{\sigma}\cdot\mathbf{p} & -m \end{pmatrix} \begin{pmatrix} \phi_R(\mathbf{p}) \\ \chi_L(\mathbf{p}) \end{pmatrix} = 0.$$

This is the Dirac equation. Under parity inversion the parts of a Dirac spinor swap over; the Dirac equation is therefore parity-invariant.

spin matrices. Then in seeking a quantum mechanical description of the electron that was consistent with the requirements of Lorentz covariance, Paul Dirac had the brilliant insight that an equation of the right form could be found if the electron is described by combining the mathematics of spinors with the existing quantum mechanics of wavefunctions. He introduced

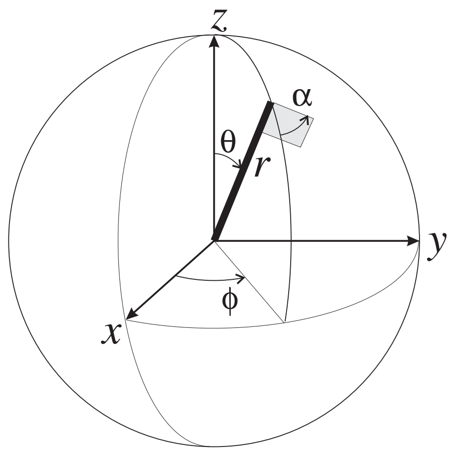

Figure 17.1: A spinor. The spinor has a direction in space ('flagpole'), an orientation about this axis ('flag'), and an overall sign (not shown). A suitable set of parameters to describe the spinor state, up to a sign, is $(r, \theta, \phi, \alpha)$, as shown. The first three fix the length and direction of the flagpole by using standard spherical coordinates, the last gives the orientation of the flag.

a 4-component complex vector, now called a *Dirac spinor*, and by physically interpreting the wave equation thus obtained, he predicted the existence of antimatter.

We have already hinted at the mathematical background to spinors when we introduced eq. (**??**) in chapter 5. Now we will discuss spinors more fully, concentrating on the simplest case, namely 2-component spinors. These suffice to describe rotations in 3 dimensions, and Lorentz transformations in $3 + 1$ dimensions.

Undergraduate students often first meet spinors in the context of non-relativistic quantum mechanics and the treatment of the spin angular momentum. This can give the impression that spinors are essentially about spin, an impression that is fortified by the name 'spinor'. However, you should try to avoid that assumption in the first instance. Think of the word 'spinor' as a generalisation of 'vector' or 'tensor'. We shall meet a spinor that describes an electric 4-current, for example, and a spinor version of the Faraday tensor, and thus write Maxwell's equations in spinor notation.

Just as we can usefully think of a vector as an arrow in space, and a 4-vector as an arrow in spacetime, it is useful to have a geometrical picture of a rank 1 spinor (or just 'spinor' for short). It can be pictured as a vector with two further features: a 'flag' that picks out a plane in space containing the vector, and an overall sign, see figure 17.1. The crucial property is that under the action of a rotation, the direction of the spinor changes just as a vector would, and the flag is carried along in the same way is if it were rigidly attached to the 'flag pole'. A rotation about the axis picked out by the flagpole would have no effect on a vector pointing in that direction, but it does affect the spinor because it rotates the flag.

The overall sign of the spinor is more subtle. We shall find that when a spinor is rotated through 360°, it is returned to its original direction, as one would expect, but also it picks up an overall sign change. You can think of this as a phase factor ($e^{i\pi} = -1$). This sign has no consequence when spinors are examined one at a time, but it can be relevant when one spinor is compared with another. When we introduce the mathematical description using a pair of complex numbers (a 2-component complex vector) this and all other properties will automatically be taken into account.

To specify a spinor state one must furnish 4 real parameters and a sign: an illustrative set r, θ, ϕ, α is given in figure 17.1. One can see that just such a set would be naturally suggested if one wanted to analyse the motion of a spinning top. We shall assume the overall sign is positive unless explicitly stated otherwise. The application to a classical spinning top is such that the spinor could represent the instantaneous positional state of the top. However, we shall not be interested in that application. In this chapter we will show how a spinor can be used to represent the energy-momentum and the spin of a massless particle, and a pair of spinors can be used to represent the energy-momentum and Pauli-Lubanksi spin 4-vector of a massive particle. Some very interesting properties of spin, that otherwise might seem mysterious, emerge naturally when we use spinors.

A spinor, like a vector, can be rotated. Under the action of a rotation, the spinor magnitude is fixed while the angles θ, ϕ, α change. In the flag picture, the flagpole and flag evolve together as a rigid body; this suffices to determine how α changes along with θ and ϕ. In order to write the equations determining the effect of a rotation, it is convenient to gather together the four parameters into two complex numbers defined by

$$\begin{aligned} a &\equiv \sqrt{r}\cos(\theta/2)e^{i(-\alpha-\phi)/2}, \\ b &\equiv \sqrt{r}\sin(\theta/2)e^{i(-\alpha+\phi)/2}. \end{aligned} \tag{17.1}$$

(The reason for the square root and the factors of 2 will emerge in the discussion). Then the effect of a rotation of the spinor through θ_r about the y axis, for example, is

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} \cos(\theta_r/2) & -\sin(\theta_r/2) \\ \sin(\theta_r/2) & \cos(\theta_r/2) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \tag{17.2}$$

We shall prove this when we investigate more general rotations below.

From now on we shall refer to the two-component complex vector

$$\boldsymbol{s} = s e^{-i\alpha/2} \begin{pmatrix} \cos(\theta/2)e^{-i\phi/2} \\ \sin(\theta/2)e^{i\phi/2} \end{pmatrix} \tag{17.3}$$

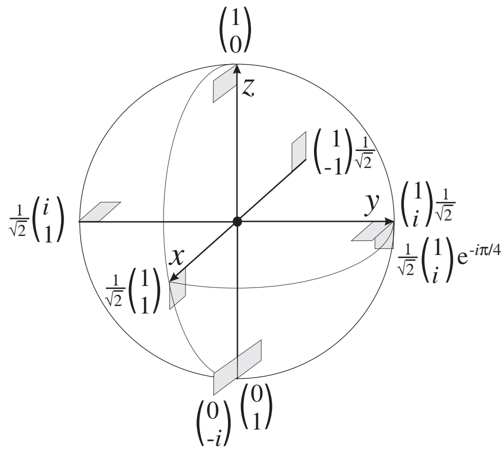

Figure 17.2: Some example spinors. In two cases a pair of spinors pointing in the same direction but with flags in different directions are shown, to illustrate the role of the flag angle α. Any given direction and flag angle can also be represented by a spinor of opposite sign to the one shown here.

as a 'spinor'. A spinor of size s has a flagpole of length

$$r = |a|^2 + |b|^2 = s^2. \tag{17.4}$$

The components (r_x, r_y, r_z) of the flagpole vector are given by

$$r_x = ab^* + ba^*, \quad r_y = i(ab^* - ba^*), \quad r_z = |a|^2 - |b|^2, \tag{17.5}$$

which may be obtained by inverting (17.1). You can now see why the square root was required in (17.1).

The complex number representation will prove to be central to understanding spinors. It gives a second picture of a spinor, as a vector in a 2-dimensional complex vector space. One learns to 'hold' this picture alongside the first one. Most people find themselves thinking pictorially in terms of a flag in a 3-dimensional real space as illustrated in figure 17.1, but every now and then it is helpful to remind oneself that a pair of opposite flagpole states such as 'straight up along z' and 'straight down along z' are *orthogonal* to one another in the complex vector space (you can see this from eq. (17.3), which gives $(s, 0)$ and $(0, s)$ for these cases, up to phase factors).

Figure 17.2 gives some example spinor states with their complex number representation. Note that the two basis vectors $(1, 0)$ and $(0, 1)$ are associated with flagpole directions up and down along z, respectively, as we just mentioned. Considered as complex vectors, these are orthogonal to one another, but they represent directions in 3-space that are opposite to one another. More generally, a rotation through an angle θ_r in the complex 'spin space' corresponds to a rotation through an angle $2\theta_r$ in the 3-dimensional real space. This is called 'angle doubling'; you can see it in eq (17.3) and we shall explore it further in section 17.2.

The matrix (17.2) for rotations about the y axis is real, so spinor states obtained by rotation of $(1, 0)$ about the y axis are real. These all have the flag and flagpole in the xz plane, with the flag pointing in the right handed direction relative to the y axis (i.e. the clockwise direction when the y axis is directed into the page). A rotation about the z axis is represented by a diagonal matrix, so that it leaves spinor states $(1, 0)$ and $(0, 1)$ unchanged in direction. To find the diagonal matrix, consider the spinor $(1, 1)$ which is directed along the positive x axis. A rotation about z should increase ϕ by the rotation angle θ_r. This means the matrix for a rotation of the spinor about the z axis through angle θ_r is

$$\begin{pmatrix} \exp(-i\theta_r/2) & 0 \\ 0 & \exp(i\theta_r/2) \end{pmatrix}. \tag{17.6}$$

When applied to the spinor $(1, 0)$, the result is $(e^{-i\theta_r/2}, 0)$. This shows that the result is to increase $\alpha + \phi$ from zero to θ_r. Therefore the flag is rotated. In order to be consistent with rotations of spinor directions close to the z axis, it makes sense to interpret this as a change in ϕ while leaving α unchanged.

So far our spinor picture was purely a spatial one. We are used to putting 3-vectors into spacetime by finding a fourth quantity and forming a 4-vector. For the spinor, however, a different approach is used, because it will turn out that the spinor is already a spacetime object that can be Lorentz-transformed. To 'place' the spinor in spacetime we just need to identify the 3-dimensional region or 'hypersurface' on which it lives. We will show in section 17.3.1 that the 4-vector associated with the flagpole is a null 4-vector. Therefore, the spinor should be regarded as 'pointing along' or 'existing on' the light cone. The word 'cone' suggests a two-dimensional surface, but of course it is 3-dimensional really and therefore can contain a spinor. The event whose light cone is meant will be clear in practice. For example, if a particle has mass or charge then we say the mass or charge is located at each event where the particle is present. In a similar way, if a rank 1 spinor is used to describe a property of a particle, then the spinor can be thought of as 'located at' each event where the particle is present, and lying on the future light cone of the event. (Some spinors of higher rank can also be associated with 4-vectors, not necessarily null ones.) The formula for a null 4-vector, $(\mathsf{X}^0)^2 = (\mathsf{X}^1)^2 + (\mathsf{X}^2)^2 + (\mathsf{X}^3)^2$, leaves open a choice of sign between the time and spatial parts, like the distinction between a contravariant and covariant 4-vector. We shall show in section 17.4 that this choice leads to two types of spinor, called 'left handed' and 'right handed'.

17.2 The rotation group and SU(2)*

We introduced spinors above by giving a geometrical picture, with flagpole and flag and angles in space. We then gave another definition, a 2-component complex vector. We have an equation relating the definitions, (17.1). All this makes it self-evident that there must exist a set of transformations of the complex vector that correspond to *rotations* of the flag and flagpole. It is also easy to guess what transformations these are: they have to preserve the length r of the flagpole, so they have to preserve the size $|a|^2 + |b|^2$ of the complex vector. This implies they are *unitary* transformations. If you are happy to accept this, and if you are happy to accept eq. (17.20) or prove it by others means (such as trigonometry), then you can skip this section and proceed straight to section 17.3. However, the connection between rotations and unitary 2×2 matrices gives a lovely example of a very powerful idea in mathematical physics, so in this section we shall take some trouble to explore it.

The basic idea is to show that two groups, which are defined in different ways in the first instance, are in fact the same (they are in one-to-one correspondance with one another, called isomorphic) or else very similar (e.g. each element of one group corresponds to a distinct set of elements of the other, called homomorphic). These are mathematical groups defined as in section 5.8, having associativity, closure, an identity element and inverses. The groups we are concerned with here have a continuous range of members, so are called Lie groups. We shall establish one of the most important mappings in Lie group theory (that is, important to physics—mathematicians would regard it as a rather simple example). This is the 'homomorphism'

$$\mathrm{SU}(2) \xrightarrow{2:1} \mathrm{S0}(3) \tag{17.7}$$

'Homomorphism' means the mapping is not one-to-one; here there are two elements of SU(2) corresponding to each element of SO(3). SU(2) is the *special unitary group of degree 2*. This is the group of two by two unitary[1] matrices with determinant 1. SO(3) is the special orthogonal group of degree 3, isomorphic to the *rotation group*. The former is the group of three by three orthogonal[2] real matrices with determinant 1. The rotation group is the group of rotations about the origin in Euclidian space in 3 dimensions.

These Lie groups SU(2) and SO(3) have the same 'dimension', where the dimension counts the number of real parameters needed to specify a member of the group, and therefore the number of dimensions of the abstract 'space' (or *manifold*) of group members (do not confuse this with dimensions in physical space and time). The rotation group is three dimensional because three parameters are needed to specify a rotation (two to pick an axis, one to give the rotation angle); the matrix group SO(3) is three dimensional because a general 3×3 matrix has nine parameters, but the orthogonality and unit determinant conditions together set six constraints; the matrix group SU(2) is three dimensional because a general 2×2 unitary matrix can be described by 4 real parameters (see below) and the determinant condition gives one constraint.

[1] A *unitary* matrix is one whose Hermitian conjugate is its inverse, i.e. $UU^\dagger = I$.
[2] An *orthogonal* matrix is one whose transpose is its inverse, i.e. $RR^T = I$.

The SU(2), SO(3) mapping is established just as was done by eqs. (??), (??), but here we shall look into it in more detail. First introduce the *Pauli spin matrices*

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \; \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that these are all Hermitian and unitary, and $\sigma_x, \sigma_y, \sigma_z$ have zero trace. Then, for any real vector $\mathbf{r} = (x, y, z)$ one can construct the traceless Hermitian matrix

$$X = \mathbf{r} \cdot \boldsymbol{\sigma} = x\sigma_x + y\sigma_y + z\sigma_z = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}. \tag{17.8}$$

It has determinant

$$|X| = -(x^2 + y^2 + z^2).$$

Now consider the matrix product

$$UXU^\dagger = X' \tag{17.9}$$

where U is unitary and of unit determinant. For *any* unitary U, if X is Hermitian then the result X' is (i) Hermitian and (ii) has the same trace. Proof (i): $(X')^\dagger = (UXU^\dagger)^\dagger = (U^\dagger)^\dagger X^\dagger U^\dagger = UXU^\dagger = X'$; (ii): the trace is the sum of the eigenvalues and the eigenvalues are preserved in unitary transformations. Since X' is Hermitian and traceless, it can in turn be interpreted as a 3-component real vector $\mathbf{r}'$ (you are invited to prove this after reading on), and furthermore, if U has determinant 1 then X' has the same determinant as X so $\mathbf{r}'$ has the same length as $\mathbf{r}$. It follows that the transformation of $\mathbf{r}$ is either a rotation or a reflection. We shall prove that it is a rotation.

Consider (see exercises)

$$e^{i(\theta/2)\sigma_x} = \cos(\theta/2)I + i\sin(\theta/2)\sigma_x = \begin{pmatrix} \cos(\theta/2) & i\sin(\theta/2) \\ i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \tag{17.10}$$

$$e^{i(\theta/2)\sigma_y} = \cos(\theta/2)I + i\sin(\theta/2)\sigma_y = \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \tag{17.11}$$

$$e^{i(\theta/2)\sigma_z} = \cos(\theta/2)I + i\sin(\theta/2)\sigma_z = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}. \tag{17.12}$$

We shall call these the 'spin rotation matrices.' By examining the effect on X you can show that the associated transformation of $\mathbf{r}$ is given by the matrices

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{17.13}$$

$$R_y = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{17.14}$$

$$R_z = \begin{pmatrix} \cos\theta & +\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{17.15}$$

These are rotations about the x, y and z axes respectively, but note the *angle doubling*: the rotation angle θ is twice the angle $\theta/2$ which appears in the 2×2 'spin rotation' matrices. In particular, when $\theta = 2\pi$ (a single full rotation) the spin rotation matrices all give $-I$. The sense of rotation is such that R represents a change of reference frame, that is to say, a rotation of the coordinate axes in a right-handed sense[3].

The 360° rotation is worth considering for a moment. We usually consider that a 360° leaves everything unchanged. This is true for a global rotation of the whole universe, or for a rotation of an isolated object not interacting with anything else. However, when one object is rotated while interacting with another that is not rotated, more possibilities arise. The fact that a spinor rotation through 360° does not give the identity operation captures a valid property of rotations that is simply not modelled by the behaviour of vectors. Place a fragile object such as a china plate on the palm of your hand, and then rotate your palm through 360° (say, anticlockwise if you use your right hand) while keeping your palm horizontal, with the plate balanced on it. It can be done but you will now be standing somewhat awkwardly with a twist in your arm. But now *continue* to rotate your palm *in the same direction* (still anticlockwise). It can be done: most of us find ourselves bringing our hand up over our shoulder, but note: the palm and plate remain horizontal and continue to rotate. After thus completing two full revolutions, 720°, you should find yourself standing comfortably, with no twist in your arm! This simple experiment illustrates the fact that there is more to rotations than is captured by the simple notion of a direction in space. Mathematically, it is noticed in a subtle property of the Lie group SO(3): the associated smooth space is not 'simply connected' (in a topological sense). The group SU(2) exhibits it more clearly: the result of one full rotation is a sign change; a second full rotation is required to get a total effect equal to the identity matrix. Figure 17.3 gives a further comment on this property.

Now, any unitary matrix of determinant 1 can be written

$$aI + ib\sigma_x + ic\sigma_y + id\sigma_z \tag{17.16}$$

[3]This is why (17.11) and (17.12) have a rotation angle of opposite sign to (17.2) and (17.6).

Figure 17.3: 'Tangloids' is a game invented by Piet Hein to explore the effect of rotations of connected objects. Two short wooden poles or blocks are joined by three parallel strings. Each player holds one of the blocks. The first player holds one block still, while the other player *rotates the other wooden block for two full revolutions about any fixed axis.* After this, the strings appear to be tangled. The first player now has to untangle them *without* rotating either piece of wood. He must use a parallel transport, that is, a translation of his block (in 3 dimensions) without rotating it or the other block. The fact that it can be done (for a 720° initial rotation, but not for a 360° initial rotation) illustrates a subtle property of rotations. It is also fun: after swapping roles, the winner is the one who untangled the fastest.

where a, b, c, d are real and $a^2 + b^2 + c^2 + d^2 = 1$ (see exercises). Using this and (17.10)-(17.12) one can further show (see exercises) that any unitary matrix of determinant 1 can be written $\exp(i\theta_x\sigma_x/2)\exp(i\theta_y\sigma_y/2)\exp(i\theta_z\sigma_z/2)$ and hence that the action of any member of SU(2) on X in eq. (17.9) results in a rotation of the associated vector $\mathbf{r}$.

We have now established the relationship between the groups:

Traceless Hermitian 2×2 matrix	$\leftrightarrow$	vector in 3 dimensions
Members U and $-U$ of SU(2)	$\leftrightarrow$	member R of SO(3)
$U = e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2}$		$R = e^{i\mathbf{J}\cdot\boldsymbol{\theta}}$

where we introduced the generators of rotations in three dimensions:

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix},$$
$$J_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix},$$
$$J_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad (17.17)$$

(c.f. eq. (5.72)). Note that for rotations the vector $\boldsymbol{\theta}$ along the axis of rotation is an axial vector (=pseudovector), one which does *not* change sign under an inversion of the coordinate system (unlike velocity and momentum).

17.2.1 Rotation of rank 1 spinors

We can now connect the discussion of SU(2) and SO(3) rotations to the spinor. The essential idea is that when any member of SU(2) multiplies a rank 1 spinor, it rotates the spinor. That is, if $\boldsymbol{s}$ and $\boldsymbol{s}'$ are rank 1 spinors related by

$$\boldsymbol{s}' = U\boldsymbol{s}$$

where $U \in \mathrm{SU}(2)$ then $\boldsymbol{s}'$ is related to $\boldsymbol{s}$ by a rotation. This can be shown directly from eq. (17.3) by trigonometry, but it will be more instructive to prove it using spinor methods, as follows.

The argument consists in essence of the following two observations:

(i) For any spinor $\boldsymbol{s}$, there is a traceless Hermitian matrix S of which $\boldsymbol{s}$ is an eigenvector with eigenvalue 1.
(ii) We already know how to relate such matrices to the rotation group.

The main points we have to prove are that (i) is true, and that the resulting matrix S has an associated direction (in the context of (ii)) that is the same as the flagpole direction of $\boldsymbol{s}$.

Proof. For any 2-component complex vector $\boldsymbol{s}$ we can construct a matrix S such that $\boldsymbol{s}$ is an eigenvector of S with eigenvalue 1. We would like S to be Hermitian. To achieve this, we make sure the eigenvectors are orthogonal and the eigenvectors real. The orthogonality we have in mind here is with respect to the standard definition of inner product in a complex vector space, namely

$$\boldsymbol{u}^\dagger \boldsymbol{v} = u_1^* v_1 + u_2^* v_2 = \langle u | v \rangle$$

where the last version on the right is in Dirac notation[4]. Beware, however, that we shall be introducing another type of inner product for spinors in section 17.4.

Let $\boldsymbol{s} = \begin{pmatrix} a \\ b \end{pmatrix}$. The spinor orthogonal[5] to $\boldsymbol{s}$ and with the same length is $\begin{pmatrix} -b^* \\ a^* \end{pmatrix}$ (or a phase factor times this). Let the eigenvalues be ± 1, then we have

$$SV = V\sigma_z$$

where

$$V = \frac{1}{s}\begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}$$

is the matrix of normalized eigenvectors, with $s = \sqrt{|a|^2 + |b|^2}$. V is unitary when the eigenvectors are normalized, as here. The solution is

$$S = V\sigma_z V^\dagger = \frac{1}{s^2}\begin{pmatrix} |a|^2 - |b|^2 & 2ab^* \\ 2ba^* & |b|^2 - |a|^2 \end{pmatrix}. \tag{17.18}$$

[4]We will occasionally exhibit Dirac notation alongside the vector and matrix notation, for the benefit of readers familiar with it. If you are not such a reader than you can safely ignore this. It is easily recognisable by the presence of $\langle\,|\,\rangle$ angle bracket symbols.

[5]The two eigenvectors, considered as vectors in a complex vector space, are orthogonal to one another (because $S_{\mathbf{n}}$ is Hermitian), but their associated flagpole directions are opposite. This is an example of the angle doubling we already noted in the relationship between SU(2) and SO(3).

Comparing this with (17.8), we see that the direction associated with S is as given by (17.5). Therefore *the direction associated with the matrix S according to (17.8) is the same as the flagpole direction of the spinor $\boldsymbol{s}$ which is an eigenvector of S with eigenvalue 1. QED.*

Eq. (17.18) can be written

$$S = \mathbf{n} \cdot \boldsymbol{\sigma} = n_x \sigma_x + n_y \sigma_y + n_z \sigma_z$$

where n_x, n_y, n_z are given by equations (17.5) divided by s^2. We find that $\mathbf{n}$ is a unit vector (this comes from the choice that the eigenvalue is 1). The result can also be written $n_x = \boldsymbol{s}^\dagger \sigma_x \boldsymbol{s}/s^2$ and similarly. More succinctly, it is

$$\mathbf{n} = \frac{\boldsymbol{s}^\dagger \boldsymbol{\sigma} \boldsymbol{s}}{s^2} = \frac{\langle s| \boldsymbol{\sigma} |s\rangle}{s^2}, \tag{17.19}$$

Another useful way of stating the overall conclusion is

> For any unit vector $\mathbf{n}$, the Hermitian traceless matrix
>
> $$S = \mathbf{n} \cdot \boldsymbol{\sigma}$$
>
> has an eigenvector of eigenvalue 1 whose flagpole is along $\mathbf{n}$.

Since a rotation of the coordinate system would bring S onto one of the Pauli matrices, S is called a 'spin matrix' for spin along the direction $\mathbf{n}$.

We still have not proved that an arbitrary special unitary matrix U acting on $\boldsymbol{s}$ has the effect of rotating it, but now this is easy to do. Consider $\boldsymbol{s}' = U\boldsymbol{s}$. We ask the question, of which matrix is $\boldsymbol{s}'$ an eigenvector with eigenvalue 1? We propose and verify the solution $USU^\dagger$:

$$(USU^\dagger)\boldsymbol{s}' = USU^\dagger U\boldsymbol{s} = US\boldsymbol{s} = U\boldsymbol{s} = \boldsymbol{s}'.$$

Therefore the answer is

$$S' = USU^\dagger.$$

This is precisely the transformation that represents a rotation of the vector $\mathbf{n}$ (compare with (17.9)), so we have proved that the flagpole of $\boldsymbol{s}'$ is in the direction $R\mathbf{n}$, where R is the rotation in 3-space associated with U in the mapping between SU(2) and SO(3). Therefore U gives a rotation of the direction of the spinor. This argument does not suffice to prove that the action has the correct effect on the orientation of the flag, but you can quickly verify that it does.

The summary is

$$
\begin{aligned}
U\begin{pmatrix} a \\ b \end{pmatrix} &\rightarrow R\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
\text{with} \qquad U = e^{i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2} &\rightarrow R = e^{i\mathbf{J}\cdot\boldsymbol{\theta}}, \qquad (17.20) \\
x = (ab^* + ba^*), \quad y = & \; i(ab^* - ba^*), \quad z = (|a|^2 - |b|^2), \\
\text{or } \mathbf{r} = & \; \langle s|\, \boldsymbol{\sigma}\, |s\rangle \,. \qquad (17.21)
\end{aligned}
$$

This summary is reminiscent of the conclusion of the SU(2) $\leftrightarrow$ SO(3) argument presented in section 17.2, but whereas there we discussed 2×2 traceless Hermitian matrices, now we are dealing with rank 1 spinors.

The sense of rotation in (17.20) is such that θ represents a rotation of the coordinate system through θ. You can confirm this by observing that the rotation $\exp(i\sigma_z\pi/4)$ carries $(1,1)/\sqrt{2}$ (a spinor directed along the positive x axis) to $e^{i\pi/4}(1,-i)/\sqrt{2}$ (a spinor directed along the negative y' axis). Therefore the equation $\boldsymbol{s}' = U\boldsymbol{s}$ shows how to express a given spinor in the coordinate system of the new (primed) reference frame.

We have here presented spinors as classical (in the sense of not quantum-mechanical) objects. If you suspect that the occasional mention of Dirac notation means that we are doing quantum mechanics, then please reject that impression. We will turn to quantum mechanics in chapter 19, but for the moment a spinor is a classical object. It is a generalization of a classical vector.

17.3 Lorentz transformation of spinors

We are now ready to generalize from space to spacetime, and make contact with Special Relativity. It turns out that the spinor is already a naturally 4-vector-like quantity, to which Lorentz transformations can be applied.

Let $\boldsymbol{s}$ be some arbitrary 1st rank spinor. Under a change of inertial reference frame it will transform as

$$\boldsymbol{s}' = \Lambda \boldsymbol{s} \qquad (17.22)$$

where Λ is a 2×2 matrix to be discovered. To this end, form the outer product

$$\boldsymbol{s}\boldsymbol{s}^\dagger = \begin{pmatrix} a \\ b \end{pmatrix} (a^*,\, b^*) = \begin{pmatrix} |a|^2 & ab^* \\ ba^* & |b|^2 \end{pmatrix}. \qquad (17.23)$$

This is (an example of) a 2nd rank spinor, and by definition it must transform as $\boldsymbol{s}\boldsymbol{s}^\dagger \to \Lambda \boldsymbol{s}\boldsymbol{s}^\dagger \Lambda^\dagger$. 2nd rank spinors (of the standard, contravariant type) are defined more generally as objects which transform in the same way, i.e. $X \to \Lambda X \Lambda^\dagger$.

Notice that the matrix in (17.23) is Hermitian. Thus outer products of 1st rank spinors form a subset of the set of Hermitian 2×2 matrices. We shall show that the complete set of Hermitian 2×2 matrices can be used to represent 2nd rank spinors.

An arbitrary Hermitian 2×2 matrix can be written

$$X = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} = tI + x\sigma_x + y\sigma_y + z\sigma_z, \tag{17.24}$$

which can also be written

$$X = \sum_\mu \mathsf{X}^\mu \sigma^\mu$$

where we introduced $\sigma^0 \equiv I$. The summation here is written explicitly, because this is not a tensor expression, it is a way of creating one sort of object (a 2nd rank spinor) from another sort of object (a 4-vector).

Evaluating the determinant, we find

$$|X| = t^2 - (x^2 + y^2 + z^2),$$

which is the Lorentz invariant associated with the 4-vector X^μ. Consider the transformation

$$X \to \Lambda X \Lambda^\dagger. \tag{17.25}$$

To keep the determinant unchanged we must have

$$|\Lambda||\Lambda^\dagger| = 1 \qquad \Rightarrow \quad |\Lambda| = e^{i\lambda}$$

for some real number λ. Let us first restrict attention to $\lambda = 0$. Then we are considering complex matrices Λ with determinant 1, i.e. the group SL(2,C). Since the action of members of SL(2,C) preserves the Lorentz invariant quantity, we can associate a 4-vector $(t, \mathbf{r})$ with the matrix X, and we can associate a Lorentz transformation with any member of SL(2,C).

The more general case $\lambda \neq 0$ can be included by considering transformations of the form $e^{i\lambda/2}\Lambda$ where $|\Lambda| = 1$. It is seen that the additional phase factor has no effect on the 4-vector obtained from any given spinor, but it rotates the flag through the angle λ. This is an example of the fact that spinors are richer than 4-vectors. However, just as we did not include such global phase factors in our definition of 'rotation', we shall also not include it in our definition of 'Lorentz transformation'. In other words, the group of Lorentz transformation of spinors is the group of 2×2 complex matrices with determinant 1 (called SL(2,C)).

The extra parameter (allowing us to go from a 3-vector to a 4-vector) compared to eq. (17.8) is exhibited in the tI term. The resulting matrix is still Hermitian but it no longer needs to have zero trace, and indeed the trace is not zero when $t \neq 0$. Now that we don't require the trace of X to be fixed, we can allow non-unitary matrices to act on it. In particular, consider the matrix

$$e^{-(\rho/2)\sigma_z} = \begin{pmatrix} e^{-\rho/2} & 0 \\ 0 & e^{\rho/2} \end{pmatrix} = \cosh(\rho/2)I - \sinh(\rho/2)\sigma_z. \tag{17.26}$$

One finds that the effect on X is such that the associated 4-vector is transformed as

$$\begin{pmatrix} \cosh(\rho) & 0 & 0 & -\sinh(\rho) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh(\rho) & 0 & 0 & \cosh(\rho). \end{pmatrix}$$

This is a Lorentz boost along z, with rapidity ρ. You can check that $\exp(-(\rho/2)\sigma_x)$ and $\exp(-(\rho/2)\sigma_y)$ give Lorentz boosts along x and y respectively. (This must be the case, since the Pauli matrices can be related to one another by rotations). The general Lorentz boost for a spinor is (for $\boldsymbol{\rho} = \rho\mathbf{n}$)

$$e^{-(\boldsymbol{\rho}/2)\cdot\boldsymbol{\sigma}} = \cosh(\rho/2)I - \sinh(\rho/2)\mathbf{n}\cdot\boldsymbol{\sigma}. \tag{17.27}$$

We thus find the whole of the structure of the restricted Lorentz group reproduced in the group SL(2,C). The relationship is a two-to-one mapping since a given Lorentz transformation (in the general sense, including rotations) can be represented by either $+M$ or $-M$, for $M \in$ SL(2,C). The abstract space associated with the group SL(2,C) has three complex dimensions and therefore six real ones (the matrices have four complex numbers and one complex constraint on the determinant). This matches the 6 dimensions of the manifold associated with the Lorentz group.

Let

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{17.28}$$

For an arbitrary Lorentz transformation

$$\Lambda = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad ad - bc = 1$$

we have

$$\Lambda^T \epsilon \Lambda = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} c & d \\ -a & -b \end{pmatrix} = \epsilon \tag{17.29}$$

(c.f. eqs. (3.41 and (5.59)). It follows that for a pair of spinors $\boldsymbol{s}$, $\boldsymbol{w}$ the scalar quantity

$$\boldsymbol{s}^T \epsilon \boldsymbol{w} = s_1 w_2 - s_2 w_1$$

is Lorentz-invariant.

The matrix ϵ satisfying (17.29) is called the **spinor Minkowski metric**.

A full exploration of the symmetries of spinors involves the recognition that the correct group to describe the symmetries of particles is not the Lorentz group but the Poincaré group. We shall not explore that here, but we remark that in such a study the concept of intrinsic spin emerges naturally, when one asks for a complete set of quantities that can be used to describe symmetries of a particle. One such quantity is the scalar invariant $\mathsf{P} \cdot \mathsf{P}$, which can be recognised as the (square of the) mass of a particle. A second quantity emerges, related to rotations, and its associated invariant is $\mathsf{W} \cdot \mathsf{W}$ where W is the Pauli-Lubanski spin vector.

17.3.1 Obtaining 4-vectors from spinors

By interpreting (17.23) using the general form (17.24) we find that the four-vector associated with the 1st rank spinor $\boldsymbol{s}$ is

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (|a|^2 + |b|^2)/2 \\ (ab^* + ba^*)/2 \\ i(ab^* - ba^*)/2 \\ (|a|^2 - |b|^2)/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \langle u| \, I \, |u\rangle \\ \langle u| \, \boldsymbol{\sigma} \, |u\rangle \end{pmatrix} = \frac{1}{2} \langle u| \, \sigma^\mu \, |u\rangle \, . \tag{17.30}$$

Any constant multiple of this is also a legitimate 4-vector. In order that the spatial part agrees with our starting point (17.1) we must introduce a factor[6] 2, so that we have the result (perhaps the central result of this chapter)

obtaining a (null) 4-vector from a spinor

$$\mathsf{U}^\mu = \boldsymbol{u}^\dagger \sigma^\mu \boldsymbol{u} \tag{17.31}$$

[6]When moving between the 4-vector and the complex number representation, the overall scale factor is a matter of convention. The convention adopted here slightly simplifies various results. Another possible convention is to retain the factor 1/2 as in (17.30).

$\boldsymbol{u}^\dagger \boldsymbol{u}$	zeroth component of a 4-vector
$\boldsymbol{u}^T \epsilon \boldsymbol{u}$	a scalar invariant (equal to zero)
$\overline{\boldsymbol{u}}^\dagger \boldsymbol{u}$	another way of writing $\boldsymbol{u}^T \epsilon \boldsymbol{u}$, with $\overline{\boldsymbol{u}} \equiv \epsilon \boldsymbol{u}^*$
$\boldsymbol{u}^T \boldsymbol{u}$	no particular significance

Table 17.1: Some scalars associated with a spinor, and their significance.

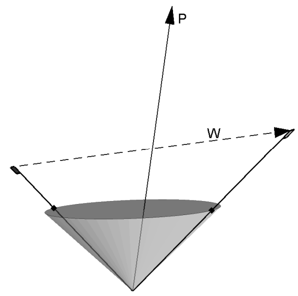

Figure 17.4: Two spinors can represent a pair of orthogonal 4-vectors. The spacetime diagram shows two spinors. They have opposite spatial direction and are embedded in a null cone (light cone), including the flags which point around the cone. Their amplitudes are not necessarily equal. The sum of their flagpoles is a time-like 4-vector P; the difference is a space-like 4-vector W. P and W are orthogonal (on a spacetime diagram this orthogonality is shown by the fact that if P is along the time axis of some reference frame, then W is in along the corresponding space axis.)

This 4-vector is null, as we mentioned in the introductory section 17.1. The easiest way to verify this is to calculate the determinant of the spinor matrix (17.23).

Since the zeroth spin matrix is the identity, we find that the zeroth component of the 4-vector can be written $\boldsymbol{u}^\dagger \boldsymbol{u}$. This and some other basic quantities are listed in table 17.1.

The linearity of eq. (17.22) shows that the sum of two spinors is also a spinor (i.e. it transforms in the right way). The new spinor still corresponds to a null 4-vector, so it is in the light cone. Note, however, that the sum of two null 4-vectors is not in general null. So adding up two spinors as in $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$ does not result in a 4-vector W that is the sum of the 4-vectors U and V associated with each of the spinors. If you want to get access to $\mathsf{U} + \mathsf{V}$, it is easy to do: first form the outer product, then sum: $\boldsymbol{u}\boldsymbol{u}^\dagger + \boldsymbol{v}\boldsymbol{v}^\dagger$. The resulting 2×2 matrix represents the (usually non-null) 4-vector $\mathsf{U} + \mathsf{V}$.

By using a pair of non-orthogonal null spinors, we can always represent a pair of orthogonal non-null 4-vectors by combining the spinors. Let the spinors be $\boldsymbol{u}$ and $\boldsymbol{v}$ and their associated 4-vectors be U and V. Let $\mathsf{P} = \mathsf{U} + \mathsf{V}$ and $\mathsf{W} = \mathsf{U} - \mathsf{V}$. Then $\mathsf{U} \cdot \mathsf{U} = 0$ and $\mathsf{V} \cdot \mathsf{V} = 0$ but $\mathsf{P} \cdot \mathsf{P} = 2\mathsf{U} \cdot \mathsf{V} \neq 0$ and $\mathsf{W} \cdot \mathsf{W} = -2\mathsf{U} \cdot \mathsf{V} \neq 0$. That is, as long as U and V are not orthogonal then P and W are not null. The latter are orthogonal to one another, however:

$$\mathsf{W} \cdot \mathsf{P} = (\mathsf{U} + \mathsf{V}) \cdot (\mathsf{U} - \mathsf{V}) = \mathsf{U} \cdot \mathsf{U} - \mathsf{V} \cdot \mathsf{V} = 0.$$

Examples of pairs of 4-vectors that are mutually orthogonal are 4-velocity and 4-acceleration, and 4-momentum and 4-spin (i.e. Pauli-Lubanski spin vector, c.f. eq. (10.15)). Therefore we can describe the motion and spin of a particle by using a pair of spinors, see figure 17.4. This connection will be explored further in section 17.5.

To summarize:

rank 1 spinor	$\leftrightarrow$	null 4-vector
rank 2 spinor	$\leftrightarrow$	arbitrary 4-vector
pair of non-orthogonal rank 1 spinors	$\leftrightarrow$	pair of orthogonal 4-vectors

17.4 Chirality

We now come to the subject of *chirality*. This concerns a property of spinors very much like the property of *contravariant* and *covariant* applied to 4-vectors. In other words, chirality is essentially about *the way spinors transform under Lorentz-transformations*. Unfortunately, the name itself does not suggest that. It is a bad name. In order to understand this we shall discuss the transformation properties first, and then return to the terminology at the end.

First, let us notice that there is another way to construct a contravariant 4-vector from a spinor. Suppose that instead of (17.31) we try

$$\mathsf{V}^\mu = \begin{pmatrix} \langle\tilde{u}| - I \,|\tilde{u}\rangle \\ \langle\tilde{u}| \boldsymbol{\sigma} \,|\tilde{u}\rangle \end{pmatrix} = \langle\tilde{u}|\, \sigma_\mu \,|\tilde{u}\rangle\,, \tag{17.32}$$

for a spinor-like object $\tilde{\boldsymbol{u}}$. It looks at first as though we have constructed a covariant 4-vector and put the index 'upstairs' by mistake. However, what if we insist that this V really is contravariant? This amounts to saying that $\tilde{\boldsymbol{u}}$ is a new type of object, not like the spinors we talked about up till now. By multiplying both sides of the equation by $g_{\mu\nu}$ to lower the index of V^μ, and then using $g_{\mu\nu} = g^{\mu\nu}$, on the right hand side, we find

$$\mathsf{V}_\nu = \langle\tilde{u}|\, \sigma^\nu \,|\tilde{u}\rangle\,. \tag{17.33}$$

Therefore the difference between $\boldsymbol{u}$ and $\tilde{\boldsymbol{u}}$ is that when combined with σ^ν, the former gives a contravariant and the latter gives a covariant 4-vector. Everything is consistent if we introduce the rule for a Lorentz transformation of $\tilde{\boldsymbol{u}}$ as

$$\text{if } \boldsymbol{u}' = \Lambda \boldsymbol{u} \tag{17.34}$$
$$\text{then } \tilde{\boldsymbol{u}}' = (\Lambda^\dagger)^{-1} \tilde{\boldsymbol{u}}. \tag{17.35}$$

This is because, for a pure rotation $\Lambda^\dagger = \Lambda^{-1}$ so the two types of spinor transform the *same* way, but for a pure boost $\Lambda^\dagger = \Lambda$ (it is Hermitian) so we have precisely the inverse transformation. This combination of properties is exactly the relationship between covariant and contravariant 4-vectors.

The two types of spinor may be called *contraspinor* and *cospinor*. However, they are often called *right-handed* and *left-handed*. The idea is that we regard the Lorentz boost as a kind of 'rotation in spacetime', and for a *given* boost velocity, the contraspinor 'rotates' one way, while the cospinor 'rotates' the other. They are said to possess opposite *chirality*. However, given that we are also much concerned with real rotations in space, this terminology is regretable because it leads to confusion.

Equation (17.33) can be 'read' as stating that the presence of $\tilde{\boldsymbol{u}}$ acts to lower the index on σ^ν and give a covariant result.

The rule (17.32) was here introduced ad-hoc: what is to say there may not be further rules? This will be explored below; ultimately the quickest way to show this and other properties is to use Lie group theory on the generators, a method we have been avoiding in order not to assume familiarity with groups, but it is briefly sketched in section (17.7).

17.4.1 Chirality, spin and parity violation

It is not too surprising to suggest that a spinor may offer a useful mathematical tool to handle angular momentum. This was the context in which spinors were first widely used. A natural way to proceed is simply to claim that there may exist fundamental particles whose intrinsic nature is not captured purely by scalar properties such as mass and charge, but which also have an angular-momentum-like property called spin, that is described by a spinor.

Having made the claim, we might propose that the 4-vector represented by the spinor flagpole is the Pauli-Lubanski spin vector. We notice that we have a null 4-vector. This means that the particle is massless (recall section 10.2.2 and eq. (10.18)). So we look for a massless spin-half particle in our experiments. We already know one: it is the neutrino[7].

Thus we have a classical model for intrinsic spin, that applies to massless spin-half particles. It is found in practice that it describes accurately the experimental observations of the nature of intrinsic angular momentum for such particles.

[7] There now exists strong evidence that neutrinos possess a small non-zero mass. We proceed with the massless model as a valid theoretical device, which can also serve as a first approximation to the behaviour of neutrinos.

Now we shall, by 'waving a magic wand', discover a wonderful property of massless spin-half particles that was not uncovered by our previous discussion in terms of 4-vectors in chapter 10. By 'waving a magic wand' here we mean noticing something that is already built in to the mathematical properties of the objects we are dealing with, namely spinors. All we need to do is claim that the same spinor describes both the linear momentum and the intrinsic spin of a given neutrino. We claim that we don't need two spinors to do the job: just one is sufficient. There is a problem: since we can only allow one rule for extracting the 4-momentum and Pauli-Lubanski spin vector for a given type of particle, we shall have to claim that there is a restriction on the allowed combinations of 4-momentum and spin for all particles of a given type. For massless particles the Pauli-Lubanski spin and the 4-momentum are aligned (either in the same direction or opposite directions, giving positive or negative helicity), so there is already a restriction that we noticed in chapter 10, but now we shall have to go further, and claim that *all massless spin-half particles of a given type have the same helicity.*

This is a remarkable claim, at first sight even a crazy claim. It says that, relative to their direction of motion, neutrinos are allowed to 'rotate' one way, but not the other! To be more precise, it is the claim that there exist in Nature processes whose mirror reflected versions never occur. Before any experimenter would invest the effort to test this (it is difficult to test because neutrinos interact very weakly with other things), he or she would want more convincing of the theoretical background, so let us investigate further.

Processes whose mirror-reflected versions run differently (for example, not at all) are said to exhibit **parity violation**. We can prove that there are no such processes in classical electromagnetism, because Maxwell's equations and the Lorentz force equation are unchanged under the parity inversion operation. The 'parity-invariant' behaviour of the last two Maxwell equations, and the Lorentz force equation, involves the fact that $\mathbf{B}$ is an axial vector.

To investigate the possibilities for spinors, consider the Lorentz invariant

$$\mathsf{W}_\lambda \mathsf{S}^\lambda = \mathsf{W}_\lambda \boldsymbol{s}^\dagger \sigma^\lambda \boldsymbol{s}$$

where $\boldsymbol{s}$ is contravariant. Since in the sum, each term W_λ is just a number, it can be moved past the $\boldsymbol{s}^\dagger$ and we have

$$\mathsf{W}_\lambda \mathsf{S}^\lambda = \boldsymbol{s}^\dagger \mathsf{W}_\lambda \sigma^\lambda \boldsymbol{s}.$$

The combination $\mathsf{W}_\lambda \sigma^\lambda = -\mathsf{W}^0 I + \mathbf{w}\cdot\sigma$ is a matrix. It can usefully be regarded as an operator acting on a spinor. We can prove that one effect of this kind of matrix, when multiplying a spinor, is to change the transformation properties. For, $\boldsymbol{s}$ transforms as

$$\boldsymbol{s} \to \Lambda \boldsymbol{s}$$

and therefore

$$\boldsymbol{s}^\dagger \to \boldsymbol{s}^\dagger \Lambda^\dagger.$$

Since $\mathsf{W}_\lambda \mathsf{S}^\lambda$ is invariant, we deduce that $\mathsf{W}_\lambda \sigma^\lambda \boldsymbol{s}$ must transform as

$$(\mathsf{W}_\lambda \sigma^\lambda \boldsymbol{s}) \to (\Lambda^\dagger)^{-1} \mathsf{W}_\lambda \sigma^\lambda \boldsymbol{s}. \tag{17.36}$$

Notation. We now have 3 vector-like quantities in play: 3-vectors, 4-vectors, and rank 1 spinors. We adopt three fonts:

entity	font	examples
3-vector	bold upright Roman	$\mathbf{s}$, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$
4-vector	sans-serif capital	S, U, V, W
spinor	bold italic	$\boldsymbol{s}$, $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$

Therefore, for any W, if $\boldsymbol{s}$ is a contraspinor then $\tilde{\mathbf{t}} = (\mathsf{W}_\lambda \sigma^\lambda \boldsymbol{s})$ is a cospinor, and *vice versa*.

If the 4-vector W is null, then it can itself be represented by a spinor $\boldsymbol{w}$. Let's see what happens when the matrix $\mathsf{W}_\lambda \sigma^\lambda$ multiplies the spinor representing W:

$$\mathsf{W}_\lambda \sigma^\lambda \boldsymbol{w} = \begin{pmatrix} -2|b|^2 & 2ab^* \\ 2a^*b & -2|a|^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{17.37}$$

where for convenience we worked in terms of the components ($\boldsymbol{w} = (a, b)$) in some reference frame. The result is

$$(-\mathsf{W}^0 + \mathbf{w} \cdot \boldsymbol{\sigma})\boldsymbol{w} = 0 \tag{17.38}$$

(N.B. in this equation $\mathbf{w}$ is a 3-vector whereas $\boldsymbol{w}$ is a spinor). This equation is important because it is (by construction) a Lorentz-covariant equation, and it tells us something useful about 1st rank spinors in general.

Suppose the 4-vector W is the 4-momentum of some massless particle. Then the equation reads

$$(E/c - \mathbf{p} \cdot \boldsymbol{\sigma})\boldsymbol{w} = 0. \tag{17.39}$$

This equation is called the first **Weyl equation** and in the context of particle physics, the rank 1 spinors are called *Weyl spinors*. It is tempting to regard the presence of $\boldsymbol{\sigma}$ in this equation as a signal that the equation describes a relation between the motion and the intrinsic angular momentum (spin) of the particle. This amounts to claiming that the same spinor can furnish the particle with both 4-momentum and spin. To do this it is necessary to use a polar version of the vector $\boldsymbol{\sigma}$ to extract the linear momentum, and an axial version to extract the spin (which, being a form of angular momentum, must be axial). Therefore when the Weyl equation (17.39) is used in this context, $\mathbf{p}$ is polar but $\boldsymbol{\sigma}$ is axial. This means the equation transforms in a non-trivial way under parity inversion. In short, it is not parity-invariant.

For a massless particle, we have $E = pc$, so (17.39) gives

$$\frac{(\mathbf{p} \cdot \boldsymbol{\sigma})}{p} \boldsymbol{w} = \boldsymbol{w}. \tag{17.40}$$

This says that $\boldsymbol{w}$ is an eigenvector, with eigenvalue 1, of the spin operator pointing along $\mathbf{p}$. In other words, *the particle has positive helicity.*

Now let's explore another possibility: suppose the spinor representing the particle has the other chirality. Then the energy-momentum is obtained as

$$\mathsf{P}_\mu = \tilde{\boldsymbol{v}}^\dagger \sigma^\mu \tilde{\boldsymbol{v}}. \tag{17.41}$$

where the use of a different letter ($\boldsymbol{v}$) indicates that we are talking about a different particle, and the tilde acts as a reminder of the different transformation properties. The invariant is now

$$\mathsf{P}^\lambda \mathsf{S}_\lambda = \tilde{\boldsymbol{s}}^\dagger \mathsf{P}^\lambda \sigma^\lambda \tilde{\boldsymbol{s}} = \tilde{\boldsymbol{s}}^\dagger (\tilde{\boldsymbol{v}}^\dagger \sigma_\lambda \tilde{\boldsymbol{v}}) \sigma^\lambda \tilde{\boldsymbol{s}} \tag{17.42}$$

and the operator of interest is

$$\left(\tilde{\boldsymbol{v}}^\dagger \sigma_\lambda \tilde{\boldsymbol{v}}\right) \sigma^\lambda = E/c + \mathbf{p} \cdot \boldsymbol{\sigma}. \tag{17.43}$$

The version on the right hand side does not at first sight look like a Lorentz invariant, because of the absence of a minus sign, but as long as we use the operator with cospinors (left handed spinors) then Lorentz covariant equations will result. For example, the argument in (17.37) is essentially unchanged and we find

$$\left(\tilde{\boldsymbol{v}}^\dagger \sigma^\alpha \tilde{\boldsymbol{v}}\right) \sigma_\alpha \tilde{\boldsymbol{v}} = 0 \tag{17.44}$$

$$\text{i.e.} \qquad (E/c + \mathbf{p} \cdot \boldsymbol{\sigma}) \tilde{\boldsymbol{v}} = 0. \tag{17.45}$$

This is called the 2nd Weyl equation. Since the particle is massless it implies

$$\frac{(\mathbf{p} \cdot \boldsymbol{\sigma})}{p} \tilde{\boldsymbol{v}} = -\tilde{\boldsymbol{v}}. \tag{17.46}$$

Therefore now the helicity is negative.

Overall, the spinor formalism suggests that there are two particle types, possibly related to one another in some way, but they are not interchangeable because they transform in different ways under Lorentz transformations, and they have the property that one type always has positive helicity, the other negative. This is born out in experiments. An experimental test involving the β-decay of cobalt nuclei was performed in 1957 by Wu *et al.*, giving clear evidence for parity non-conservation. In 1958 Goldhaber *et al.* took things further in a beautiful experiment, designed to allow the helicity of neutrinos to be determined. It was found that all neutrinos

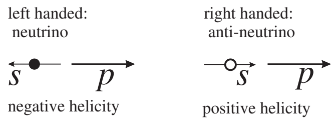

Figure 17.5: The black and white circles represent two particle-like entities. Both are massless leptons with spin 1/2 and zero charge. Are they two examples of the same type of particle then, merely having the spin in opposite directions? The sequence of statements shown in the figure gives the logic. The black entity is found to have different chirality from the white entity. This is a subtle property, not easily illustrated by any diagram, since it refers to how the spinor transforms under a boost. However this property suffices to distinguish one entity from the other, and it is legitimate to give them different names ("neutrino" and "anti-neutrino") and draw them with different colours. The theoretical model asserts that the information about 4-spin and energy-momentum is contained in a single spinor for each entity. It then follows that the helicity is single-valued: always negative for the one we called "neutrino" and always positive for the one we called "anti-neutrino". Similar reasoning applied to electrons reaches a different conclusion. They are not described by single spinors, but by a pair of spinors, one of each chirality. Consequently the helicity of an electron can be of either sign, and is not Lorentz-invariant.

produced in a given type of process have the same helicity. This is evidence that all neutrinos have one helicity, and anti-neutrinos have the opposite helicity. By convention those with positive helicity are called anti-neutrinos. With this convention, the process

$$\mathrm{n} \to \mathrm{p} + \mathrm{e} + \bar{\nu} \tag{17.47}$$

is allowed (with the bar indicating an antiparticle), but the process $\mathrm{n} \to \mathrm{p} + \mathrm{e} + \nu$ is not. Thus the properties of Weyl spinors are at the heart of the parity-non-conservation exhibited by the weak interaction.

Reflection and Lorentz transformation

[*Section omitted in lecture-note version.*]

17.4.2 Index notation*

[*Section omitted in lecture-note version.*]

What is the difference between chirality and helicity?
Answer: helicity refers to the projection of the spin along the direction of motion, chirality refers to the way the spinor transforms under Lorentz transformations.

The word 'chirality' in general in science refers to *handedness*. A screw, a hand, and certain types of molecule may be said to possess *chirality*. This means they can be said to embody a rotation that is either left-handed or right-handed with respect to a direction also embodied by the object. When Weyl spinors are used to represent spin angular momentum and linear momentum, they also possess a handedness, which can with perfect sense be called an example of chirality. However, since the particle physicists already had a name for this (helicity), the word chirality came to be used to refer directly to the transformation property, such that spinors transforming one way are said to be 'right-handed' or of 'positive chirality', and those transforming the other way are said to be 'left-handed' or of 'negative chirality.' This terminology is poor because (i) it invites (and in practice results in) confusion between chirality and helicity, (ii) spinors can be used to describe other things beside spin, and (iii) the *transformation rule* has nothing in itself to do with angular momentum. The terminology is acceptable, however, if one understands it to refer to the Lorentz boost as a form of 'rotation' in spacetime.

17.5 Applications

[*Section omitted in lecture-note version.*]

17.6 Dirac spinor and particle physics

We already mentioned in section (17.3.1) that a pair of spinors can be used to represent a pair of mutually orthogonal 4-vectors. A good way to do this is to use a pair of spinors of opposite chirality, because then it is possible to construct equations possessing invariance under parity inversion. Such a pair is called a *bispinor* or *Dirac spinor*. It can conveniently be written as a 4-component complex vector, in the form

$$\Psi = \begin{pmatrix} \phi_R \\ \chi_L \end{pmatrix} \tag{17.48}$$

where it is understood that each entry is a 2-component spinor, ϕ_R being right-handed and χ_L left-handed. (Following standard practice in particle physics, we won't adopt index notation for the spinors here, so the subscript L and R is introduced to keep track of the chirality). Under change of reference frame Ψ transforms as

$$\Psi \to \begin{pmatrix} \Lambda(v) & 0 \\ 0 & \Lambda(-v) \end{pmatrix} \Psi \tag{17.49}$$

where each entry is understood to represent a 2×2 matrix, and we wrote $\Lambda(v)$ for $\exp\left(i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2 - \boldsymbol{\sigma}\cdot\boldsymbol{\rho}/2\right)$ and $\Lambda(-v)$ for $(\Lambda(v)^\dagger)^{-1} = \exp\left(i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2 + \boldsymbol{\sigma}\cdot\boldsymbol{\rho}/2\right)$. It is easy to see that the combination

$$\begin{pmatrix} \phi_R^\dagger, \chi_L^\dagger \end{pmatrix} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} \phi_R \\ \chi_L \end{pmatrix} = \phi_R^\dagger \chi_L + \chi_L^\dagger \phi_R \tag{17.50}$$

is Lorentz-invariant.

We will show how Ψ can be used to represent the 4-momentum and 4-spin (Pauli-Lubanski 4-vector) of a particle. First extract the 4-vectors given by the flagpoles of ϕ_R and χ_L:

$$\mathsf{A}^\mu = \langle \phi_R | \, \sigma^\mu \, | \phi_R \rangle \,, \qquad \mathsf{B}_\mu = \langle \chi_L | \, \sigma^\mu \, | \chi_L \rangle \,.$$

Note that B_μ has a lower index. This is because, in view of the fact that χ_L is left-handed (i.e. negative chirality), under a Lorentz transformation its flagpole behaves as a covariant 4-vector. We would like to form the difference of these 4-vectors, so we need to convert the second to contravariant form. This is done via the metric tensor $g_{\mu\nu}$:

$$\mathsf{U}^\mu = (\mathsf{A}^\mu - \mathsf{B}^\mu) = \langle \phi_R | \, \sigma^\mu \, | \phi_R \rangle - g^{\mu\alpha} \langle \chi_L | \, \sigma^\alpha \, | \chi_L \rangle \,.$$

(The notation is consistent if you keep in mind that the $_L$'s have lowered the index on σ in the second term). In terms of the Dirac spinor Ψ, this result can be written as

$$\mathsf{U}^0 = \Psi^\dagger \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \Psi = \phi_R^\dagger \phi_R + \chi_L^\dagger \chi_L \tag{17.51}$$

for the time component, and

$$\mathsf{U}^i = \Psi^\dagger \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix} \Psi \tag{17.52}$$

for the spatial components.

Now introduce the 4×4 matrices, called **Dirac matrices**:

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}. \tag{17.53}$$

Here we are writing these matrices in the 'chiral' basis implied by the form (17.48). Using these, we can write (17.51) and (17.52) both together, as

$$\mathsf{U}^\mu = \Psi^\dagger \gamma^0 \gamma^\mu \Psi. \tag{17.54}$$

We now have a 4-vector extracted from our Dirac spinor. It can be of any type—it need not be null. If in some particular reference frame it happens that $\phi_R = \pm\chi_L$ (this equation is not Lorentz covariant so cannot be true in all reference frames, but it can be true in one), then in that reference frame we will find $\mathsf{A}^\mu = \mathsf{B}^\mu$ so the spatial part of U will be zero, while the time part is not. Such a 4-vector is proportional to a particle's 4-velocity in its rest frame. In other words, the Dirac spinor can be used to describe the 4-velocity of a massive particle, and in this application it must have either $\phi_R = \chi_L$ or $\phi_R = -\chi_L$ in the rest frame. In other frames the 4-velocity can be extracted using (17.54).

Next consider the sum of the two flagpole 4-vectors. Let

$$\mathsf{W} = mcS\,(\mathsf{A} + \mathsf{B}) \tag{17.55}$$

where S is the size of the intrinsic angular momentum of the particle, and introduce

$$\gamma^5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}. \tag{17.56}$$

By using

$$\Sigma^\mu = \gamma^0\gamma^\mu\gamma^5 = \left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \right), \tag{17.57}$$

we can write

$$\mathsf{W}^\mu = mcS\Psi^\dagger\Sigma^\mu\Psi. \tag{17.58}$$

This 4-vector is orthogonal to U. It can therefore be the 4-spin, if we choose ϕ_R and χ_L appropriately. What is needed is that the spinor ϕ_R be aligned with the direction of the spin angular momentum in the rest frame. We already imposed the condition that either $\phi_R = \chi_L$ or $\phi_R = -\chi_L$ in the rest frame, so it follows that either both spinors are aligned with the spin angular momentum in the rest frame, or one is aligned and the other opposed.

We now have a complete representation of the 4-velocity and 4-spin of a particle, using a single Dirac spinor. A spinor equal to $(1, 0, 1, 0)/\sqrt{2}$ in the rest frame, for example, represents a

Ψ	U	$2\mathsf{W}/(mcS)$	
$(1,0,1,0)/\sqrt{2}$	$(1,0,0,0)$	$(0,0,0,1)$	at rest, spin up
$(0,1,0,1)/\sqrt{2}$	$(1,0,0,0)$	$(0,0,0,-1)$	at rest, spin down
$(1,1,1,1)/2$	$(1,0,0,0)$	$(0,1,0,0)$	at rest, spin along $+x$
$(1,-1,1,-1)/2$	$(1,0,0,0)$	$(0,-1,0,0)$	at rest, spin along $-x$
$(1,0,0,0)$	$(1,0,0,1)$	$(1,0,0,1)$	$v_z = c$, +ve helicity
$(0,1,0,0)$	$(1,0,0,-1)$	$(1,0,0,-1)$	$v_z = -c$, +ve helicity
$(0,0,1,0)$	$(1,0,0,-1)$	$(-1,0,0,1)$	$v_z = -c$, −ve helicity
$(0,0,0,1)$	$(1,0,0,1)$	$(-1,0,0,-1)$	$v_z = c$, −ve helicity

Table 17.2: Some example Dirac spinors and their associated 4-vectors.

particle with spin directed along the z direction. A spinor $(0,1,0,1)/\sqrt{2}$ represents a particle with spin in the $-z$ direction. More generally, $(\phi,\phi)/\sqrt{2}$ is a particle at rest with spin vector $\phi^\dagger \boldsymbol{\sigma} \phi$.

The states having $\phi_R = \chi_L$ in the rest frame cover half the available state space; the other half is covered by $\phi_R = -\chi_L$ in the rest frame. The spinor formalism is here again implying that there may exist in Nature two types of particle, similar in some respects (such as having the same mass), but not the same. We shall see in chapter 19 that if the first set of states are particle states, then the other set can be interpreted as antiparticle states.

When the spin and velocity are along the same direction, in the high-velocity limit it is found that one of the two spinor components dominates. For example, if we start from $\phi_R = \chi_L = (1,0)$ in the rest frame, then transform to a reference frame moving in the positive z direction, then ϕ_R will shrink and χ_L will grow until in the limit $v \to c$, $\phi_R \to 0$. This implies that a massless particle can be described by a single (two-component) spinor, and we recover the property that a Weyl spinor has helicity of the same sign as its chirality The example we just considered had negative helicity because the spin is along z but the particle's velocity is in the negative z direction in the new frame.

Table 17.2 lists all these and some further examples.

A parity inversion ought to change the direction in space of U (since its spatial part is a polar vector) but leave the direction in space of W unaffected (since its spatial part is an axial vector). You can verify that this is satisfied if the parity inversion is represented by the matrix

$$P = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \qquad (17.59)$$

The effect of P acting on Ψ is to swap the two parts, $\phi_R \leftrightarrow \chi_L$. You can now verify that the Lorentz invariant given in (17.50) is also invariant under parity so it is a true scalar. The quantity $\Psi^\dagger \gamma^0 \gamma^5 \Psi = \phi_R^\dagger \chi_L - \chi_L^\dagger \phi_R$ is invariant under Lorentz transformations but changes sign under parity, so is a pseudoscalar. The results are summarised in table 17.3.

$\Psi^\dagger\gamma^0\Psi$	scalar
$\Psi^\dagger\gamma^0\gamma^5\Psi$	pseudoscalar
$\Psi^\dagger\gamma^0\gamma^\mu\Psi$	4-vector U, difference of flagpoles
$\Psi^\dagger\gamma^0\gamma^\mu\gamma^5\Psi$	axial 4-vector W, sum of flagpoles
$\Psi^\dagger\gamma^0(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\Psi$	antisymmetric tensor

Table 17.3: Various tensor quantities associated with a Dirac spinor. The notation $\overline{\Psi} = \Psi^\dagger\gamma^0$ (called *Dirac adjoint*) can also be introduced, which allows the expressions to be written $\overline{\Psi}\Psi$, $\overline{\Psi}\gamma^5\Psi$, and so on.

17.6.1 Moving particles and classical Dirac equation

So far we have established that a Dirac spinor representing a massive particle possessing intrinsic angular momentum ought to have $\phi_R = \chi_L$ in the rest frame, with both spinors aligned with the intrinsic angular momentum. We have further established that under Lorentz transformations the Dirac spinor will continue to yield the correct 4-velocity and 4-spin using eqs (17.54) and (17.58).

Now we shall investigate the general form of a Dirac spinor describing a moving particle. All we need to do is apply a Lorentz boost. We assume the Dirac spinor $\Psi = (\phi_R(v), \chi_L(v))$ takes the form $\phi_R(0) = \chi_L(0)$ in the rest frame, and that it transforms as (17.49). Using (17.27) the Lorentz boost for a Dirac spinor can be written

$$\Lambda = \cosh(\rho/2)\begin{pmatrix} I - \mathbf{n}\cdot\boldsymbol{\sigma}\tanh(\rho/2) & 0 \\ 0 & I + \mathbf{n}\cdot\boldsymbol{\sigma}\tanh(\rho/2) \end{pmatrix}.$$

Now (in units where $c = 1$) $\cosh\rho = \gamma = E/m$ where E is the energy of the particle, and $\cosh\rho = 2\cosh^2(\rho/2) - 1 = 2\sinh^2(\rho/2) + 1$ so

$$\begin{aligned} \cosh(\rho/2) &= \left(\frac{E+m}{2m}\right)^{1/2}, \qquad \sinh(\rho/2) = \left(\frac{E-m}{2m}\right)^{1/2}, \\ \tanh(\rho/2) &= \left(\frac{E-m}{E+m}\right)^{1/2} = \frac{p}{E+m}. \end{aligned} \tag{17.60}$$

Therefore we can express the Lorentz boost in terms of energy and momentum (of a particle boosted from its rest frame):

$$\Lambda = \sqrt{\frac{E+m}{2m}}\begin{pmatrix} I + \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m} & 0 \\ 0 & I - \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m} \end{pmatrix} \tag{17.61}$$

where the sign is set such that $\mathbf{p}$ is the momentum of the particle in the new frame.

For example, consider the spinor $\Psi_0 = (1, 0, 1, 0)/\sqrt{2}$, i.e. spin up along z in the rest frame. Then in any other frame,

$$\Psi = \frac{1}{\sqrt{4m(E+m)}} \begin{pmatrix} E+m+p_z \\ p_x+ip_y \\ E+m-p_z \\ -p_x-ip_y \end{pmatrix} \tag{17.62}$$

(with $c = 1$). Suppose the boost is along the x direction. Then, as p_x grows larger, $p_x \to E$, so the positive chirality part has a spinor more and more aligned with $+x$, and the negative chirality part has a spinor more and more aligned with $-x$. For a boost along z, one of the chirality components vanishes in the limit $|v| \to c$. This is the behaviour we previously discussed in relation to table 17.2.

Next we shall present the result of a Lorentz boost another way. We will construct a matrix equation satisfied by Ψ that has the same form as the Dirac equation to be discussed in chapter 19. Historically, Dirac obtained his equation via a quantum mechanical argument. However, the classical version will prepare us for the quantum version, and help in the interpretation of the solutions.

We already noted that the Lorentz boost takes the form

$$\begin{aligned} \phi_R(\boldsymbol{v}) &= (\cosh(\rho/2) - \boldsymbol{\sigma}\cdot\mathbf{n}\sinh(\rho/2))\,\phi_R(0), \\ \chi_L(\boldsymbol{v}) &= (\cosh(\rho/2) + \boldsymbol{\sigma}\cdot\mathbf{n}\sinh(\rho/2))\,\chi_L(0). \end{aligned}$$

Using eq. (17.60), and multiplying top and bottom by $(E+m)^{1/2}$, we find

$$\phi_R(\mathbf{p}) = \frac{E+m+\boldsymbol{\sigma}\cdot\mathbf{p}}{[2m(E+m)]^{1/2}}\phi_R(0), \qquad \chi_L(\mathbf{p}) = \frac{E+m-\boldsymbol{\sigma}\cdot\mathbf{p}}{[2m(E+m)]^{1/2}}\chi_L(0).$$

where $\mathbf{p} = -\gamma m\mathbf{v}$ is the momentum of the particle in the new frame. Introducing the assumption $\phi_R(0) = \chi_L(0)$ we obtain from these two equations, after some algebra[8],

$$\begin{aligned} (E - \boldsymbol{\sigma}\cdot\mathbf{p})\phi_R(\mathbf{p}) &= m\chi_L(\mathbf{p}), \\ (E + \boldsymbol{\sigma}\cdot\mathbf{p})\chi_L(\mathbf{p}) &= m\phi_R(\mathbf{p}). \end{aligned}$$

The left hand sides of these equations are the same as in the Weyl equations; the right hand sides have the requisite chirality. As a set, this coupled pair of equations is parity-invariant,

[8] Let $\eta = \boldsymbol{\sigma}\cdot\mathbf{p}$. Premultiply the first equation by $E+m-\eta$ and the second by $E+m+\eta$ to obtain $(E+m-\eta)\phi_R = (E+m+\eta)\chi_L$; then premultiply the first equation by $E-m-\eta$ and the second by $-E+m-\eta$ to obtain $(E-m-\eta)\phi_R = (-E+m-\eta)\chi_L$ (after making use of $\eta^2 = p^2$). The sum and difference of these equations gives the result.

since under a parity inversion the sign of $\boldsymbol{\sigma} \cdot \mathbf{p}$ changes and χ and ϕ swap over. In matrix form the equations can be written

$$\begin{pmatrix} -m & E + \boldsymbol{\sigma} \cdot \mathbf{p} \\ E - \boldsymbol{\sigma} \cdot \mathbf{p} & -m \end{pmatrix} \begin{pmatrix} \phi_R(\mathbf{p}) \\ \chi_L(\mathbf{p}) \end{pmatrix} = 0. \tag{17.63}$$

This equation is very closely related to the Dirac equation which will be derived in the next chapter, see eq. (18.21). One may even go so far as to say that (17.63) "is" the Dirac equation in free space, if one re-interprets the terms. This connection is explained in the next chapter. In the present context the equation represents a constraint that must be satisfied by any Dirac spinor that represents 4-momentum and intrinsic spin.

Our discussion has been entirely classical (in the sense of not quantum-mechanical). In quantum field theory the spinor plays a central role. One has a spinor field, the excitations of which are what we call spin 1/2 particles. The results of this section reemerge in the quantum context, unchanged for energy and momentum eigenstates, and in the form of mean values or 'expectation values' for other states.

17.7 Spin matrix algebra (Lie algebra)*

We introduced the Pauli spin matrices abruptly at the start of section 17.2, without explanation. By now the reader has some idea of their usefulness. In order to derive results such as the spin rotation matrices, eq. (17.10), one needs to notice certain basic properties such as that all the spin matrices square to 1:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I \tag{17.64}$$

and it is very useful to have the commutation relations

$$[\sigma_x,\, \sigma_y] \equiv \sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z \tag{17.65}$$

and similarly for cyclic permutation of x, y, z. You can also notice that

$$\sigma_x \sigma_y = i\sigma_z, \quad \sigma_y \sigma_x = -i\sigma_z$$

and therefore any pair anti-commutes:

$$\sigma_x \sigma_y = -\sigma_y \sigma_x$$

or in terms of the 'anticommutator'

$$\{\sigma_x, \sigma_y\} \equiv \sigma_x \sigma_y + \sigma_y \sigma_x = 0.$$

In group theory, these matrices are called the *generators* of the group SU(2), because any group member can be expressed in terms of them in the form $\exp(i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2)$. More precisely, the Pauli matrices are the generators of one *representation* of the group SU(2), namely the representation in terms of 2×2 complex matrices. Other representations are possible, such that there is an isomorphism between one representation and another. Each representation will have generators in a form suitable for that representation. They could be matrices of larger size, for example, or even differential operators. In every representation, however, the generators will have the same behaviour when combined with one another, and this behaviour reveals the nature of the group. This means that the innocent-looking commutation relations (17.65) contain much more information than one might have supposed: they are a 'key' that, through the use of $\exp(i\boldsymbol{\sigma}\cdot\boldsymbol{\theta}/2)$, unlocks the complete mathematical behaviour of the group. In Lie group theory these equations describing the generators are called the 'Clifford algebra' or 'Lie algebra' of the group.

If a Lie group does not have a matrix representation it can be hard or impossible to give a meaningful definition to $\exp(M)$ where M is a member of the group. In this case one uses the form $I+\epsilon M$ to write a group member infinitesimally close to the identity for $\epsilon\to 0$. The generators are a subgroup such that any member close to I can be written $I+\epsilon G$ where G is in the generator group. This is the more general definition of what is meant by the generators.

The generators of rotations in three dimensions (17.17) have the commutation relations

$$[J_x,\,J_y]=iJ_z \quad \text{and cyclic permutations.} \tag{17.66}$$

By comparing with (17.65) one can immediately deduce the relationship between SU(2) and SO(3), including the angle doubling!

The restricted Lorentz group has generators K_i (for boosts) and J_i (for rotations). The commutation relations are (with cyclic permutations)

$$\begin{aligned}
[J_x,\,J_y] &= iJ_z \\
[K_x,\,K_y] &= -iJ_x \\
[J_x,\,K_x] &= 0 \\
[J_x,\,K_y] &= iK_z.
\end{aligned}$$

The second result shows that the Lorentz boosts on their own do not form a closed group, and that two boosts can produce a rotation: this is the Thomas precession. If we now form the combinations

$$\mathbf{A}=(\mathbf{J}+i\mathbf{K})/2, \qquad\qquad \mathbf{B}=(\mathbf{J}-i\mathbf{K})/2,$$

then the commutation relations become

$$\begin{aligned}
[A_x, A_y] &= iA_z \\
[B_x, B_y] &= iB_z \\
[A_i, B_j] &= 0, \quad (i, j = x, y, z).
\end{aligned}$$

This shows that the Lorentz group can be divided into two groups, both SU(2), and the two groups commute. This is another way to deduce the existence of two types of Weyl spinor and thus to define chirality.

The Clifford algebra satisfied by the Dirac matrices γ^0, γ^1, γ^2, γ^3 is

$$\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu} I \tag{17.67}$$

where $\eta^{\mu\nu}$ is the Minkowski metric and I is the unit matrix. That is, γ^0 squares to I and γ^i $(i = 1, 2, 3)$ each square to $-I$, and they all anticommute among themselves. These anti-commutation relations are normally taken to be the *defining property* of the Dirac matrices. A set of quantities γ^μ satisfying such anti-commutation relations can be represented using 4×4 matrices in more than one way—c.f. eq (17.53) and (20.12). If the metric of signature $(1, -1, -1, -1)$ is used, then the minus sign on the right hand side of (17.67) becomes a plus sign.

17.7.1 Dirac spinors from group theory*

[*Section omitted in lecture-note version.*]

[*Section omitted in lecture-note version.*]