Introduction to Symmetries

I.J.R.Aitchison

March 27, 2001

Lecture 4:

The Lorentz Group.

Consider the relativistic spin-0 wave equation (the Klein - Gordon equation)

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\psi(\mathbf{x},t) = 0,$$

(as usual we will use natural units $\hbar = c = 1$) which can be written as

$$\left(\partial^2 + m^2\right)\psi(\mathbf{x},t) = 0.$$

Under rotations the operator ∂^2 is invariant, and we know how ψ transforms under rotations. What about under Lorentz boosts? i.e. pure velocity transformations of the form

$$\begin{aligned} x' &= \gamma(x+vt) \\ t' &= \gamma(t+vx) \end{aligned}$$

$$y \to y' = y \text{ and } z \to z' = z.$$

We ask the familiar question: *what wavefunction do the people in the (Lorentz boosted) primed frame use?*

As before we require

$$\psi'(\mathbf{r}', t') = \psi(\mathbf{r}, t)$$

except that now t transforms as well as $\mathbf{r}$, and the transformation is different (not a 3-D rotation). Consider an infinitesimal boost along the x-axis:

$$\begin{aligned} x' &= \gamma(x+vt) &\to\quad x' &= x+\epsilon t \\ t' &= \gamma(t+vx) &\to\quad t' &= t+\epsilon x \end{aligned}$$

$$y' = y \text{ and } z' = z$$

$$(\gamma \to 1).$$

Let's denote our Lorentz transform as $\mathbf{LT}$, i.e. $(\mathbf{r}', t') = \mathbf{LT}(\mathbf{r}, t)$. We require

$$\psi'(\mathbf{r}', t') = \psi(\mathbf{r}, t)$$

$$\Rightarrow \psi'(\mathbf{LT}(\mathbf{r},t)) = \psi(\mathbf{r},t)$$

$$\begin{aligned} \Rightarrow \quad \psi'(\mathbf{r},t) &= \psi\left((\mathbf{LT})^{-1}(\mathbf{r},t)\right) \\ &= \psi(x-\epsilon t, y, z, t-\epsilon x) \\ &= \psi(\mathbf{r},t) - \epsilon t\tfrac{\partial\psi(\mathbf{r},t)}{\partial x} - \epsilon x\tfrac{\partial\psi(\mathbf{r},t)}{\partial t} \\ &= \left(1+i\epsilon\widehat{B}_1\right)\psi(\mathbf{r},t) \end{aligned}$$

where

$$\widehat{B}_1 = +i\left(t\frac{\partial}{\partial x} + x\frac{\partial}{\partial t}\right).$$

Similarly we have

$$\widehat{B}_2 = +i\left(t\frac{\partial}{\partial y} + y\frac{\partial}{\partial t}\right) \text{ and } \widehat{B}_3 = +i\left(t\frac{\partial}{\partial z} + z\frac{\partial}{\partial t}\right),$$

and for the general infinitesimal boost parametrised by the vector $\mathbf{b}$ (where $\mathbf{b} = \beta\,\mathbf{n}$ with β an infinitesimal speed and $\mathbf{n}$ a unit vector) we have

$$\psi'(\mathbf{r},t) = (1 + i\mathbf{b}.\widehat{\mathbf{B}})\psi(\mathbf{r},t).$$

For a combined general infinitesimal rotation and boost (the most general infinitesimal Lorentz transformation on a spin-0 wavefunction) we have

$$\psi'(\mathbf{r},t) = (1 - i\mathbf{a}.\widehat{\mathbf{L}} + i\mathbf{b}.\widehat{\mathbf{B}})\psi(\mathbf{r},t).$$

The quantities $\widehat{\mathbf{B}}$ are the **generators of boosts**. They clearly bear some resemblance to the generators of rotations, which are of course the angular momentum operators $\widehat{L}_x$, $\widehat{L}_y$ and $\widehat{L}_z$. As in the case of 3-D rotations, we are interested in the **algebra** of the $\widehat{B}_i$'s and $\widehat{L}_i$'s i.e the commutation relations between all of them. By straightforward calculation using the above differential operators we find that

$$\left[\widehat{B}_i, \widehat{B}_j\right] = -i\epsilon_{ijk}\widehat{L}_k$$

$$\left[\widehat{L}_i, \widehat{B}_j\right] = +i\epsilon_{ijk}\widehat{B}_k$$

as well as the usual

$$\left[\widehat{L}_i, \widehat{L}_j\right] = +i\epsilon_{ijk}\widehat{L}_k.$$

<u>This is the algebra of the generators of the Lorentz group!</u> There are 6 generators, 3 for rotations $\widehat{L}_i$ and 3 for boosts $\widehat{B}_i$. One (new) thing to note is that the $\widehat{B}_i$'s aren't Hermitian. In terms of this differential operator representation this is a bit tricky and has to do with problems concerning the operator $i\hbar\frac{\partial}{\partial t}$. We might think that this is just the energy operator $\widehat{E}$, but the trouble is that then there would be nothing to stop the eigenvalue of $\widehat{E}$ being negative (with wavefunction $e^{i|E|t/\hbar}$). Physically we require the spectrum of the energy eigenvalues to be non-negative and this screws up the hermiticity.

Let's have a look at a <u>matrix representation</u> of these generators of the Lorentz group. As usual, we can find representations by looking at infinitesimal transformations, acting on some chosen set of basis functions. For example we can consider infinitesimal Lorentz transformations acting on the four-vector $(t, \mathbf{r})$, as parametrised by

$$\begin{aligned} t' &= t + \mathbf{b}.\mathbf{r} \\ \mathbf{r}' &= \mathbf{r} + \mathbf{b}t \quad + \quad \mathbf{a}\times\mathbf{r}. \end{aligned}$$

($\mathbf{b}$ is an infinitesimal velocity vector as before, while $\mathbf{a}$ specifies an infinitesimal rotation as we saw in Lecture 2). It is easy to check that this leaves the "length" of the 4-vector unchanged to first order in $\mathbf{a}$ and $\mathbf{b}$:

$$t'^2 - \mathbf{r}'^2 = t^2 - \mathbf{r}^2.$$

The "$\mathbf{r}' = \mathbf{r} + \mathbf{a}\times\mathbf{r}$" bit leaves the three-dimensional length invariant ($\mathbf{r}'^2 = \mathbf{r}^2$) as we saw in Lecture 2; the new feature here is the extension of this to the 4-D length - which is of course the Minkowskian length $t^2 - \mathbf{r}^2$ not the Euclidean one $t^2 + \mathbf{r}^2$. We write the transformation as:

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & b_x & b_y & b_z \\ b_x & 1 & -a_z & a_y \\ b_y & a_z & 1 & -a_x \\ b_z & -a_y & a_x & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}.$$

To be consistent with our notation for the operator associated with general infinitesimal Lorentz transformations, we write the transformation matrix above as $1 - i\mathbf{a}.\mathbf{L}^{(v)} + i\mathbf{b}.\mathbf{B}^{(v)}$, where the superscripts denote that this is a specific representation, and where in this case "^{v} " stands for "vector". Picking out just the bit associated with a_x, for example, we find

$$L_x^{(v)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}.$$

Looking only at the spatial components of the matrix we find that $L_x^{(v)}$ is simply $L_x^{(1)}$ augmented above by a row of zeros and on the left by a column of zeros. In the same way, we can find $L_y^{(v)}$ and $L_z^{(v)}$. The same correspondence exists between $L_y^{(v)}$ and $L_y^{(1)}$ as well as between $L_z^{(v)}$ and $L_z^{(1)}$, <u>all of which are clearly Hermitian.</u> The fact that essentially the same three angular momentum generators turn up again is no surprise, of course: as we said, the infinitesimal transformation we are considering includes the 3-D rotations, under which the time component of the 4-vector is "inert" (which accounts for those bordering rows and columns of zeros).

On the other hand, when we go through the analogous procedure to pick out the matrices for the three boosts (associated with the parameters b_x, b_y and b_z) we find that <u>the three matrices representing the generators for the boosts are not Hermitian:</u>

$$B_x^{(v)} = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_y^{(v)} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_z^{(v)} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}.$$

They are each, in fact, i times a Hermitian matrix. It is a simple matter to evaluate the various commutation relations among these matrices:

$$\left[B_x^{(v)}, B_y^{(v)}\right] = -iL_z^{(v)}$$

$$\left[L_x^{(v)}, B_y^{(v)}\right] = +iB_z^{(v)}$$

$$\left[L_x^{(v)}, L_y^{(v)}\right] = +iL_z^{(v)}.$$

Note these are exactly the same as the commutators we obtained previously for the $\widehat{B}_i$'s and $\widehat{L}_i$'s, which were represented in terms of differential operators rather than matrices. In fact, we have clearly obtained a matrix representation of the generators! The $\widehat{L}_i$'s are represented by the $L_i^{(v)}$'s, while the $\widehat{B}_i$'s are represented by the $B_i^{(v)}$'s. This is a four-dimensional representation, appropriate to the 4-D basis vector $(t, \mathbf{r})$.

Now we want to consider other representations, in particular spinor-type ones (which will lead us to the Dirac equation). We start by slightly generalising and then simplifying the algebra , i.e. the commutation relations of the generators. First we generalise the rotation generators $\widehat{\mathbf{L}}$ to $\widehat{\mathbf{J}}$ in order to include spin as well as orbital angular momentum. So the algebra is now

$$\left[\widehat{B}_i, \widehat{B}_j\right] = -i\epsilon_{ijk}\widehat{J}_k$$

$$\left[\widehat{J}_i, \widehat{B}_j\right] = +i\epsilon_{ijk}\widehat{B}_k$$

$$\left[\widehat{J}_i, \widehat{J}_j\right] = +i\epsilon_{ijk}\widehat{J}_k.$$

This can be simplified by the following trick:

$$\widehat{\mathbf{M}} = \frac{1}{2}\left(\widehat{\mathbf{J}} + i\widehat{\mathbf{B}}\right)$$

$$\widehat{\mathbf{N}} = \frac{1}{2}\left(\widehat{\mathbf{J}} - i\widehat{\mathbf{B}}\right).$$

Then the algebra becomes:

$$\left[\widehat{M}_i, \widehat{M}_j\right] = +i\epsilon_{ijk}\widehat{M}_k$$

$$\left[\widehat{N}_i, \widehat{N}_j\right] = +i\epsilon_{ijk}\widehat{N}_k$$

$$\left[\widehat{M}_i, \widehat{N}_j\right] = 0.$$

This is generally true for all representations. So the $\widehat{M}_i$'s and $\widehat{N}_i$'s behave just like two independent (because they commute) angular momenta!

Now, $\widehat{\mathbf{J}}$ and $i\widehat{\mathbf{B}}$ are Hermitian, and so $\widehat{\mathbf{M}}$ and $\widehat{\mathbf{N}}$ are also Hermitian. This allows us to make use of the standard theory of angular momentum in quantum mechanics to tell us about the representations of the Lorentz group! We can label the basis states in the same way that we label angular momentum eigenstates of a system with two angular momenta (like two particles with spins $\widehat{\mathbf{S}}_1, \widehat{\mathbf{S}}_2$ for example). That is, we can label the states by the eigenvalues of $\widehat{\mathbf{M}}^2, \widehat{M}_3, \widehat{\mathbf{N}}^2$ and $\widehat{N}_3$:

$$\left|\text{eigenvalues of } \widehat{\mathbf{M}}^2, \widehat{M}_3, \widehat{\mathbf{N}}^2, \widehat{N}_3\right\rangle.$$

Because they are like angular momenta, the eigenvalues of $\widehat{\mathbf{M}}^2$ are of the form $j(j+1)$ with $j = 0, 1/2, 1, 3/2...$ and similarly for $\widehat{\mathbf{N}}^2$; and the eigenvalues of $\widehat{M}_3$ are $-j, -j+1, ...0, ..., j$ for a given j, etc.

Now consider the *particular* case where the eigenvalue of $\widehat{\mathbf{N}}^2$ is zero and the eigenvalue of $\widehat{\mathbf{M}}^2$ is $\frac{1}{2}\left(\frac{1}{2}+1\right)$. The first condition implies that the $\widehat{\mathbf{N}}$'s are identically zero so that $\widehat{\mathbf{J}} = i\widehat{\mathbf{B}}$, while the second gives $\widehat{\mathbf{M}} = \frac{1}{2}\left(\widehat{\mathbf{J}} + i\widehat{\mathbf{B}}\right) = \frac{1}{2}\boldsymbol{\sigma}$. Since these expressions for $\widehat{\mathbf{M}}$ and $\widehat{\mathbf{N}}$ refer to a particular representation, we should add superscripts on them to indicate that (just as we did for the $\mathbf{L}^{(v)}$'s, for instance). Pretty obviously, that superscript had better be essentially " $\frac{1}{2}$ " because of the appearance of the sigma matrices and the eigenvalue $j = \frac{1}{2}$. However, it is slightly more subtle than that: we have to account not only for the " halfness" of $\widehat{\mathbf{M}}$ but also the "noughtness" of $\widehat{\mathbf{N}}$. So we shall write

$$\mathbf{J}^{(\frac{1}{2},0)} + \mathbf{J}^{(\frac{1}{2},0)} = \boldsymbol{\sigma}$$

$$\Rightarrow \mathbf{J}^{(\frac{1}{2},0)} = \frac{1}{2}\boldsymbol{\sigma}$$

$$\Rightarrow \mathbf{B}^{(\frac{1}{2},0)} = \frac{-i}{2}\boldsymbol{\sigma}.$$

We now recall that the general infinitesimal transformation has the form " $1 - i\mathbf{a}.\widehat{\mathbf{J}} + i\mathbf{b}.\widehat{\mathbf{B}}$". In the present case, this becomes $1 - i\mathbf{a}.\frac{1}{2}\boldsymbol{\sigma} + i\mathbf{b}.\frac{-i}{2}\boldsymbol{\sigma} = 1 - \frac{i}{2}\mathbf{a}.\boldsymbol{\sigma} + \frac{1}{2}\mathbf{b}.\boldsymbol{\sigma}$. So a spinor wavefunction behaving like this transforms by

$$\phi' = \left(1 - \frac{i}{2}\boldsymbol{\sigma}.\mathbf{a} + \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b}\right)\phi.$$

We have made a slight and trivial change of notation here: it makes no difference at all whether we write "$\mathbf{a}.\boldsymbol{\sigma}$ " or " $\boldsymbol{\sigma}.\mathbf{a}$ " since the $\mathbf{a}$'s (and the $\mathbf{b}$'s) are just numbers, so it doesn't matter which side of $\boldsymbol{\sigma}$ we write them; this way is more convenient when we discuss wave equations, below. Anyway, such a wavefunction is said to transform as the "$(\frac{1}{2}, 0)$ " representation (or, to provide the basis for that representation). In terms of the Dirac state vector, it is written as $\left|\frac{1}{2}, 0\right\rangle$, where the first label refers to the "j " of $\widehat{\mathbf{M}}^2$ and the second to the "j " of $\widehat{\mathbf{N}}^2$. The eigenvalues of $\widehat{M}_3$ and of $\widehat{N}_3$ are suppressed.

Let us pause to take stock. The "$1 - \frac{i}{2}\boldsymbol{\sigma}.\mathbf{a}$" part is just the same as what we had for infinitesimal rotations of two-component spinors; the new part is the " $\frac{1}{2}\boldsymbol{\sigma}.\mathbf{b}$. " which shows how such a wavefunction transforms under an infinitesimal boost. To get the transformation for the finite velocity case (finite boost) we "exponentiate" the $\frac{1}{2}\boldsymbol{\sigma}.\mathbf{b}$ factor (as in the finite rotation case) to obtain

$$\phi' = e^{\frac{1}{2}v\sigma.\mathbf{n}}\phi = \mathbf{V_B}\phi$$

for a boost with speed v along an axis $\mathbf{n}$. Notice that the boost matrix $\mathbf{V_B}$ is non-unitary,

$$\left(e^{\frac{1}{2}v\sigma.\mathbf{n}}\right).\left(e^{\frac{1}{2}v\sigma.\mathbf{n}}\right)^\dagger \neq 1$$

(try it for the infinitesimal case), whereas the corresponding operator for finite rotations is unitary:

$$\left(e^{\frac{i}{2}\alpha\sigma.\mathbf{n}}\right).\left(e^{\frac{i}{2}\alpha\sigma.\mathbf{n}}\right)^\dagger = 1.$$

Briefly, the reason for this difference between the boost and rotation operators is that there are no finite-dimensional unitary representations of a non-compact group. By non-compact is meant that the space over which the parameters appearing in the transformation range is unbounded. In our case, γ can approach infinity as v approaches 1 ($c = 1$), and we are considering a two-dimensional representation.

What about the other similar case where the the eigenvalue of $\widehat{\mathbf{M}}^2$ is zero and the eigenvalue of $\widehat{\mathbf{N}}^2$ is $\frac{1}{2}\left(\frac{1}{2}+1\right)$? We call this the " $(0, \frac{1}{2})$ " representation (with corresponding state vector $|0, \frac{1}{2}\rangle$). In this representation we have $\mathbf{J}^{(0,\frac{1}{2})} = -i\mathbf{B}^{(0,\frac{1}{2})}$ from the first condition, and $\frac{1}{2}(\mathbf{J}^{(0,\frac{1}{2})} - i\mathbf{B}^{(0,\frac{1}{2})}) = \frac{1}{2}\boldsymbol{\sigma}$ from the second. This results in

$$\mathbf{J}^{(0,\frac{1}{2})} + \mathbf{J}^{(0,\frac{1}{2})} = \boldsymbol{\sigma}$$

$$\Rightarrow \mathbf{J}^{(0,\frac{1}{2})} = \frac{1}{2}\boldsymbol{\sigma}$$

as before, while

$$\mathbf{B}^{(0,\frac{1}{2})} = \frac{+i}{2}\boldsymbol{\sigma}$$

which is the negative of $\mathbf{B}^{(\frac{1}{2},0)}$. So this time the infinitesimal transformation is:

$$\chi' = \left(1 - \frac{i}{2}\boldsymbol{\sigma}.\mathbf{a} - \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b}\right)\chi.$$

This is how "$\left|0, \frac{1}{2}\right\rangle$ " transforms. Note that the rotation part $1 - \frac{i}{2}\mathbf{a}.\boldsymbol{\sigma}$ is the same as before but the boost bit now has a minus sign in front of it $-\frac{1}{2}\mathbf{b}.\boldsymbol{\sigma}$. This is the difference between these two types of spinor.

Connection to wave equations

We have arrived at the Lorentz transformation properties of the two types of spinor ϕ and χ via the "representations of the Lorentz group" approach. In Lecture 3, however, we found how a spinor behaved under rotations by starting from a wave equation (the Pauli equation), and asking our standard question: what is the primed wave function? It is natural to ask whether we can understand the transformation properties we have found for ϕ and χ in terms of the wave equations they satisfy.

What are these wave equations? They clearly have to describe spin - $\frac{1}{2}$ particles, but relativistically - so they cant be Pauli equations. Actually, we know quite well that somehow the Dirac equation has got to make its appearance....but let's pretend we don't know that, and start by considering massless particles, for which the energy-momentum relation is $E^2 = \mathbf{p}^2$. A possible wave equation for such a free spin - $\frac{1}{2}$ particle is

$$i\frac{\partial\psi}{\partial t} = \boldsymbol{\sigma}.\widehat{\mathbf{p}}\psi.$$

We can check this by writing ψ as the product of a plane-wave and a two-component spinor:

$$\psi = e^{i\mathbf{p}.\mathbf{r} - iEt}\phi.$$

Then

$$E\phi = \boldsymbol{\sigma}.\mathbf{p}\phi,$$

and by applying $\boldsymbol{\sigma}.\mathbf{p}$ to both sides of this equation and using $(\boldsymbol{\sigma}.\mathbf{p})^2 = \mathbf{p}^2$ (see Lecture 3) we find that for consistency $E^2 = \mathbf{p}^2$ as required.

Thus the equation satisfied by the spinor ϕ , in a frame in which the 4-momentum is $(E, \mathbf{p})$, is

$$(E - \boldsymbol{\sigma}.\mathbf{p})\phi = 0$$

with $E = |\mathbf{p}|$.

Now consider an infinitesimal Lorentz boost to a second (primed) frame, in which the 4-momentum is

$$E' = E + \mathbf{b}.\mathbf{p}$$

$$\mathbf{p}' = \mathbf{p} + \mathbf{b}E$$

(compare the same thing for the coordinate 4-vector $(t, \mathbf{r})$). In this primed frame, we have

$$(E' - \boldsymbol{\sigma}.\mathbf{p}')\phi' = 0.$$

So here is our question once again: *what is the relation between* ϕ *and* ϕ' ?

Actually we know the answer, of course! Except that we could be dealing with a " χ " rather than a " ϕ ".....Let's see. Assuming this ϕ is indeed a " $(\frac{1}{2}, 0)$ " kind of beast, we expect

$$\phi' = (1 + \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b})\phi$$

(there is no need to worry about 3-D rotations - the work of Lecture 3 guarantees that the " $\boldsymbol{\sigma}.\mathbf{p}$ " bit will transform correctly). Let's denote $(1 + \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b})$ by $\mathbf{V_b}$. Then applying ${\mathbf{V_b}}^{-1}$ to the ϕ equation, we have

$$[{\mathbf{V_b}}^{-1}(E - \boldsymbol{\sigma}.\mathbf{p}){\mathbf{V_b}}^{-1}]\mathbf{V_b}\phi = 0.$$

The part in square brackets is (remember that $\mathbf{b}$ is infinitesimal)

$$(1 - \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b})(E - \boldsymbol{\sigma}.\mathbf{p})(1 - \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b})$$

and it is a good exercise to check that, to first order in $\mathbf{b}$, this is just

$$E - \boldsymbol{\sigma}.\mathbf{p} - E\boldsymbol{\sigma}.\mathbf{b} + \mathbf{b}.\mathbf{p} = E' - \boldsymbol{\sigma}.\mathbf{p}' \quad !$$

Hence we find

$$(E' - \boldsymbol{\sigma}.\mathbf{p}')\mathbf{V_b}\phi = 0$$

and we can deduce that

$$\phi' = \mathbf{V_b}\phi$$

which is the (expected) answer to our question.

What kind of physics does the spinor ϕ describe? Since $E = |\mathbf{p}|$, the ϕ - equation can also be written as

$$\left(\frac{\boldsymbol{\sigma}.\mathbf{p}}{|\mathbf{p}|}\right)\phi = \phi$$

which means that ϕ is an eigenvector of the *helicity operator* (the projection of spin along the momentum direction) with eigenvalue $+1$.

What about the opposite helicity state? This will be described by a χ such that

$$\boldsymbol{\sigma}.\mathbf{p}\chi = -|\mathbf{p}|\chi,$$

which can also be written as

$$(E + \boldsymbol{\sigma}.\mathbf{p})\chi = 0$$

where the energy $E = |\mathbf{p}|$ is still positive. Obviously the notation suggests that χ will be a " $(0, \frac{1}{2})$ " animal. Let's check. This time we multiply the equation by $\mathbf{V_b}$:

$$[\mathbf{V_b}(E + \boldsymbol{\sigma}.\mathbf{p})\mathbf{V_b}]{\mathbf{V_b}}^{-1}\chi = 0$$

and we find that the quantity in square brackets is indeed $(E' + \boldsymbol{\sigma}.\mathbf{p}')$, allowing us to identify

$$\chi' = {\mathbf{V_b}}^{-1}\chi = (1 - \frac{1}{2}\boldsymbol{\sigma}.\mathbf{b})\chi$$

as expected.

So the ϕ 's describe massless spin - $\frac{1}{2}$ particles with positive helicity ("right handed") while the χ 's describe ones with negative helicity ("left handed") - like antineutrinos and neutrinos, respectively, if they are massless. These massless spin - $\frac{1}{2}$ wave equations are called Weyl equations.

How about the case of spin - $\frac{1}{2}$ particles with mass? Here there is an important physics point to appreciate. Helicity is not a Lorentz invariant for massive particles, since we can always reverse the sense of the particle's momentum (and hence helicity) by transforming to a more rapidly moving frame - something that of course is not possible for massless particles. So we suspect that the mass term will couple ϕ 's and χ 's together. Indeed this is so. The analogous equations for massive particles are

$$E\phi = \boldsymbol{\sigma}.\mathbf{p}\phi + m\chi$$

$$E\chi = -\boldsymbol{\sigma}.\mathbf{p}\chi + m\phi.$$

It is simple to check that for these equations to be consistent we need

$$E^2 = \mathbf{p}^2 + m^2$$

as required. Furthermore, it is left as an exercise for the student to check that the above coupled equations for ϕ and χ are indeed consistent with the transformation laws we have learned for ϕ and χ.

These coupled massive equations can be written as a single equation by introducing the 4-component spinor ψ where

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix},$$

which satisfies

$$E\psi = (\boldsymbol{\alpha}.\mathbf{p} + \beta m)\psi$$

where

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix}, \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This is just the Dirac equation! - with a particular ("high energy") representation of the $\boldsymbol{\alpha}$ and β matrices. Thus we can finally say that the Dirac spinor ψ transforms as " $|\frac{1}{2}, 0\rangle \oplus |0, \frac{1}{2}\rangle$ ", the $\oplus$ meaning that we must place them "one on top of the other", and so double the dimensionality of the vector space, not add them "horizontally".