Chapter 16
The relativistic Dirac equation
–
from my book:
Understanding Relativistic Quantum Field Theory

Hans de Vries

November 10, 2008

Contents

16 The relativistic Dirac equation **1**
16.1 The linearized Wave equation 2
16.2 The chiral 2d Dirac equation 4
16.3 Plane wave solutions of the 2d Dirac equation 5
16.4 An interpretation of the chiral components 6
16.5 The chiral propagation in 4d 10
16.6 The derivation of the Dirac equation 11
16.7 Introduction of the Pauli spinors and matrices 12
16.8 Pauli spin matrices as spin 1/2 operators 14
16.9 Spinor rotations over 720 degrees 16
16.10 The spin direction and the phase of the spinor 18
16.11 General spinor rotations 21
16.12 The wave's phase as the spinor phase 22
16.13 Rotation operator of the full chiral bi-spinor 23
16.14 Weyl's chiral bi-spinors and relativity 23
16.15 The 2d spinor boost operators 25
16.16 The 4d spinor boost operators 28

Chapter 16

The relativistic Dirac equation

16.1 The linearized Wave equation

We will go here through an easy going, step by step, derivation of the Dirac equation (in the "chiral" representation), with the main focus on the actual *physical* meaning of all it's properties. Without such a focus on the physics, Dirac's equation can leave the reader with the impression that its abstract mathematical nature somehow just miraculously produces the correct answers at the end of a number of obscure mathematical manipulations. This is not true however, all of the aspects of the Dirac equation correspond to real physics which we can connect to visualizations in our own familiar world, rather than in some abstract mathematical space.

We start at the very beginning. Say, we want to obtain a wave equation which has solutions in the form of arbitrary functions, that shift along with constant velocity c, "on the light cone". We can write down the first order equations below in two dimensions t and r. These equations, however, allow only one-directional propagation, either to the left or to the right with a velocity of plus or minus c.

$$\frac{\partial \psi_L}{\partial t} - c\frac{\partial \psi_L}{\partial r} = 0 \qquad\qquad \frac{\partial \psi_R}{\partial t} + c\frac{\partial \psi_R}{\partial r} = 0 \tag{16.1}$$

The equations work because the *shifting* solutions have equal derivatives in t and r up to a proportionality constant which denotes the velocity c. One way to accommodate both left and right moving solutions is our second order classical wave equation. This equation owes its bidirectional nature to the parameter c^2 which can be either $(+c)^2$ or $(-c)^2$.

$$\frac{\partial^2 \psi}{\partial t^2} - c^2\frac{\partial^2 \psi}{\partial r^2} = 0 \qquad \equiv \qquad E^2 - c^2 p_r^2 = 0 \tag{16.2}$$

But there is another possibility which combines both the linear nature of (16.1) and the classical wave equation which has proved to be an equation of fundamental importance. We do so by retaining the left and right moving waves as two separated components of a two component wave function.

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.3}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{\partial \psi}{\partial t} + c\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\frac{\partial \psi}{\partial r} = 0 \tag{16.4}$$

or

$$\Upsilon^t \frac{\partial \psi}{\partial t} \;+\; c\,\Upsilon^r \frac{\partial \psi}{\partial r} \;=\; 0 \tag{16.5}$$

Where the matrices Υ^t and Υ^r have a number of notable properties when multiplied together:

$$\Upsilon^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Upsilon^r = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{16.6}$$

$$\Upsilon^t\Upsilon^t = I, \quad \Upsilon^r\Upsilon^r = -I, \quad \Upsilon^t\Upsilon^r + \Upsilon^r\Upsilon^t = 0 \tag{16.7}$$

We see that the cross terms cancel which allows us to write the equation as an operator on the wave function ψ. This operator, when applied *twice*, gives us the classical wave equation:

$$\left[\;\Upsilon^t \frac{\partial}{\partial t} \;+\; c\Upsilon^r \frac{\partial}{\partial r}\;\right]^2 \psi \;= 0 \quad \equiv \quad \left[\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial r^2}\right] \psi \;= 0 \tag{16.8}$$

Thus, both the left and right going components ψ_L and ψ_R obey the classical wave equation. In general, we won't explicitly show the unity matrix I in the equations. So much about the mathematics, but what about the physics? Look at the anti-diagonal nature of the matrices.

The linear equation for ψ_L leaves its result, (=0) on the other channel ψ_R and visa versa. In case of the classical wave equation the result is zero, that is, a left moving solutions doesn't leave anything on the right moving channel and nothing of ψ_R ends up on ψ_L. Both left and right moving solutions move only in one direction and do so with speed c.

Now, we assumed that the right hand side (=0) contains the result of the operator, which is true. But, knowing the classical wave equation, we know that the *result* on the right hand side is actually the *source* of the wave function ψ. (The four current j^μ in the case of the electromagnetic field). So, with respect to cause and result we should say that, because of (=0), the right moving channel ψ_R is never a source for the left moving channel ψ_L and visa versa. To actually describe the progression *forwards* in time, the *propagation*, instead of tracking backwards in time to the source, we need to apply the inverse of the operator, the Green's function (or simply *the propagator*).

16.2 The chiral 2d Dirac equation

We can now make the next important step: We can couple the two channels by going from the classical wave equation to the 2d Dirac equation. Again, our two component equation is basically linear and it becomes the Klein Gordon equation by applying the operator acting on ψ twice.

$$\left[\Upsilon^t \frac{\partial}{\partial t} + c\Upsilon^r \frac{\partial}{\partial r} \right] \psi = -i \left(\frac{mc^2}{\hbar} \right) \psi \tag{16.9}$$

$$\left[\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial r^2} \right] \psi = - \left(\frac{mc^2}{\hbar} \right)^2 \psi \tag{16.10}$$

We see that both ψ_L and ψ_R obey the Klein Gordon equation. For the physical meaning we go back once more to the simple spring/mass representation of the (real) Klein Gordon equation.

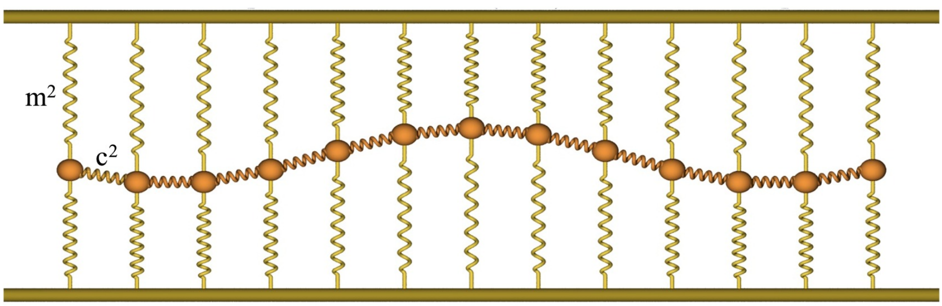

Figure 16.1: The (real) klein-Gordon equation as a spring/mass system

The horizontal springs represent the classical wave equation and they allow both left and right shifting solutions with a velocity c. The masses are allowed to move in the vertical direction which represents their "degree of freedom". The horizontal springs now represent the coupling between the left and right moving solutions. A left moving solution exerts a strain on a vertical string which opposes the displacement of the mass. The left moving solution is reflected back to the right. The reflection is negative since system tries to undo the change. Our equation tracks back to the source, The source of ψ_L are the reflections of ψ_R and visa versa. Lets look somewhat more into the matrices used.

$$\Upsilon^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Upsilon^r = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{16.11}$$

We see from Υ^t that the direction in time does't change when going from one channel to the other, we always keep on going backward in time, tracking back the source of ψ. From the signs in Υ^r we see that the parity in r is different in the two channels. In the version with three spatial dimensions, all coordinates x, y and z are inverted. The channels have opposed parity. We say that the left moving component ψ_L is said to have left chirality while ψ_R is said to have right chirality.

To recall the all important observation here, is that, in going from the classical wave equation to the 2d Dirac equation, we have obtained solutions which can move at other speed than plus c and minus c. In general any speed between the two extremes becomes possible. We have subdivided the physical process in linear steps with our two component chiral 2d Dirac equation.

16.3 Plane wave solutions of the 2d Dirac equation

At this stage we can already use the 2d Dirac equation to derive what are basically the plane wave solutions of the Dirac equation. We will look for a solution in the form of.

$$\psi \;=\; \begin{pmatrix} u_L \\ u_R \end{pmatrix} \exp\left(-i\,\frac{Et}{\hbar} \;+\; i\,\frac{pr}{\hbar} \right) \tag{16.12}$$

Which contains a chiral two component term u and where the exponential part represents the known eigenfunctions of the Klein Gordon equation. We write out the linear equation in full:

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial t} + c \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \frac{\partial}{\partial r} \right] \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \frac{mc^2}{i\hbar} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.13}$$

and we substitute (16.12) in (16.13) which gives us two coupled equations in the components U_L and U_R:

$$\frac{1}{i\hbar}(\ E + cp\ )\ u_R \ = \ \frac{mc^2}{i\hbar}\ u_L \tag{16.14}$$
$$\frac{1}{i\hbar}(\ E - cp\ )\ u_L \ = \ \frac{mc^2}{i\hbar}\ u_R \tag{16.15}$$

Which are satisfied by the plane wave solutions of the 2d Dirac equation, where we have used $\sqrt{E^2 - c^2p^2} = mc^2$.

$$\begin{pmatrix} u_L \\ u_R \end{pmatrix} \phi^+ \ = \ \begin{pmatrix} \sqrt{E - cp} \\ \sqrt{E + cp} \end{pmatrix} \phi^+ \tag{16.16}$$

In which ϕ^+ is the exponential function in (16.12). Repeating the same procedure for anti-particles by reversing the sign of the energy E in equation (16.12) gives us the plane wave solution for anti-particles:

$$\begin{pmatrix} v_L \\ v_R \end{pmatrix} \phi^- \ = \ \begin{pmatrix} \sqrt{E + cp} \\ -\sqrt{E - cp} \end{pmatrix} \phi^- \tag{16.17}$$

16.4 An interpretation of the chiral components

By coupling the two independent chiral parts of the linear wave equation with a mass term we obtained the 2d Dirac equation which has solutions that can move at any speed between plus c and minus c. In the massless case the chiral components move at plus and minus c, now, at what speed do they move in the case of of particles with mass?

We use here a method to obtain the speed of the chiral components which has troubled physicist a lot especially with other representations of the Dirac equation. In non relativistic quantum mechanics we can obtain the velocity operator by taking the first order derivative in time of the position operator by commuting it with the Hamiltonian, where the Hamiltonian is understood to be energy given by $-i\hbar\ \partial/\partial t$ This is confirmed if we look at the quantum mechanical velocity operator.

$$\vec{v} \ = \ \frac{\partial \vec{r}}{\partial t} \ = \ \frac{i}{\hbar}[\ H, \vec{r}\ ] \ = \ \begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix} \tag{16.18}$$

This result suggested a speed of +c or -c for the Dirac electron which didn't make sense. Now, note that the result is a 2x2 matrix operating on a two component function. So, we draw the conclusion that the operator is working on the internal chiral components instead and that the result suggest that the speed of $\psi_L = -c$ and $\psi_R = +c$, just like in the massless case of the linear wave equation. Now indeed, the result of equation (16.18) is entirely independent of the mass.

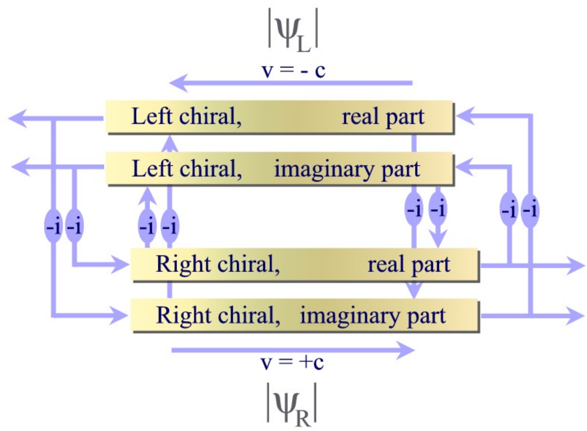

Figure 16.2: The interplay between the chiral components

This confirms our assumptions about the left and right moving chiral components being reflected back and forward into each other via the m term. We see from the free plane wave solutions that both chiral components balance each other out for particle at rest. The equal magnitude of the components results in a speed of zero. At the other hand, at relativistic speeds, only one the two chiral components remains.

Now, how do we calculate the speed of the whole particle? Well it turns out that we can use the velocity operator discussed above perfectly well to this. we need only to realize that the Hamiltonian used is also a 2x2 matrix and operates on the external structure. In order to get an observable we must let it operate on the two component wave function in the usual way.

$$\frac{1}{2E} \begin{pmatrix} \psi_L^* \\ \psi_R^* \end{pmatrix} \left(-i\hbar \, \mathrm{I} \, \frac{\partial}{\partial t} \right) \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad = \quad E \qquad (16.19)$$

where the term $1/2E$ is a normalization term. If we now apply the 2x2 velocity operator in the same way on the wave function then we get exactly where we were looking for: The velocity of the whole particle.

$$\frac{1}{2E}\begin{pmatrix} \psi_L^* \\ \psi_R^* \end{pmatrix}\begin{pmatrix} -c & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = v \qquad (16.20)$$

So, this operator which didn't seem to make sense does just what it should do after being applied in the right way. From figure 16.2 we see that there are two fully independent paths of circulation. The phase relation between the two paths determines the direction of the phase change in time of the particle. If the real parts of the left and right components are in phase then the energy is positive and the wave represents a particle. If the phase between the real parts is opposite then the wave represents an anti-particle.

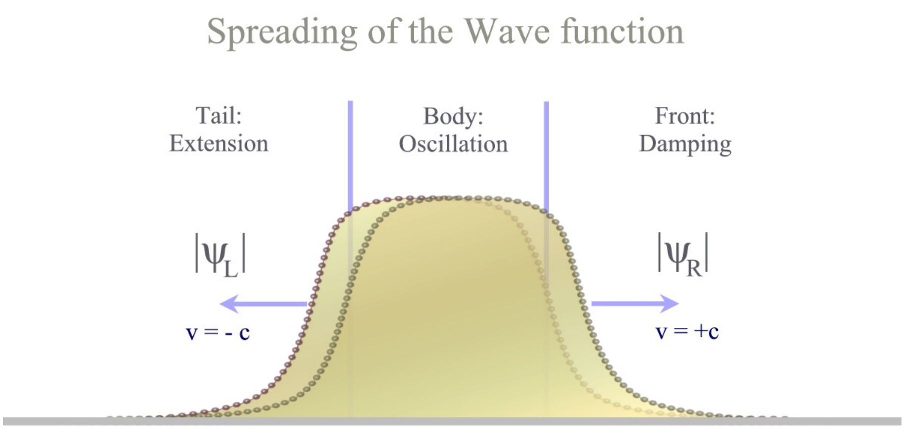

Figure 16.3: The front, body and tail zone of the interactions

We can distinguish roughly three different zones of interaction between the chiral components, see figure 16.3, since the chiral components move in opposite directions. In the middle zone where the parts overlap we will see a mainly oscillatory behavior with a magnitude which stays more or less constant depending on how fast the wave function spreads.

At the front side the other chiral component is diminishing rapidly and stops being a source. The only contributions from the other side are coming from the front-end after being reflected back and forward again. The double reflection has changed the sign of the contribution. This negative feedback rapidly damps the front-end of the wave function preventing it from escaping away with the speed of light.

At the tail end the other chiral component is predominantly the source. The other side extends the tail. The result is that both sides keep basically overlapping each other even though the move at opposite direction at the highest speed possible.

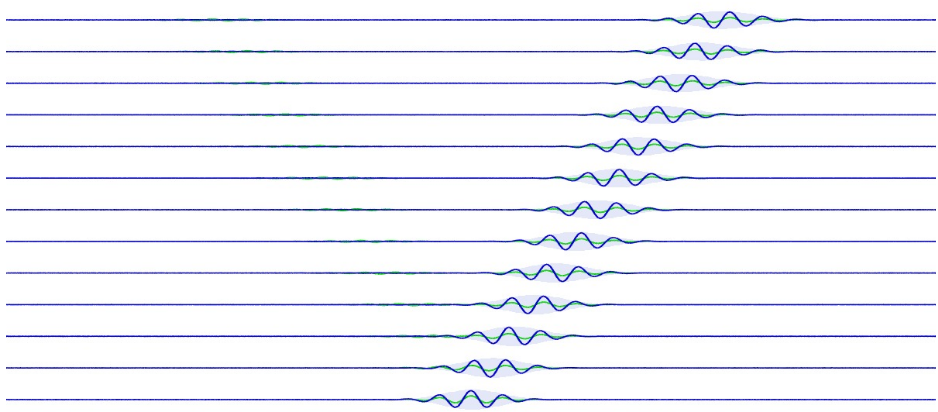

Figure 16.4: Particle moving to the right

Figure 16.4 shows a simulation of a 2d Dirac equation wave-function. The blue line represents ψ_R and the green line represents ψ_L. The speed depends on the phase change in x representing the momentum of the particle. The relation of the two chiral components is in principle the same as the ratio in the free plane wave solution, with a small distortion from the Gaussian wave shape.

The particle in figure 16.4 is a particle and not an antiparticle as can be seen from the phase relation of the left and right hand chiral components: They have the same phase. The particle spreads only minimally because the size of the wave function is relatively large with respect to the Compton radius of the particle which is inversely proportional to the mass.

Figure 16.5 shows a particle at rest which is spreading fast. The fast spreading is the result of the initial compressed size of the wave function at t is 0 relative to the Compton wavelength corresponding with the mass of the particle. The particle keeps spreading with the momentum gained and actually continues to spread faster. The oscillations at both ends are an indication of the local momentum. We see that the phase tends to get equal all over the interior of the wave function.

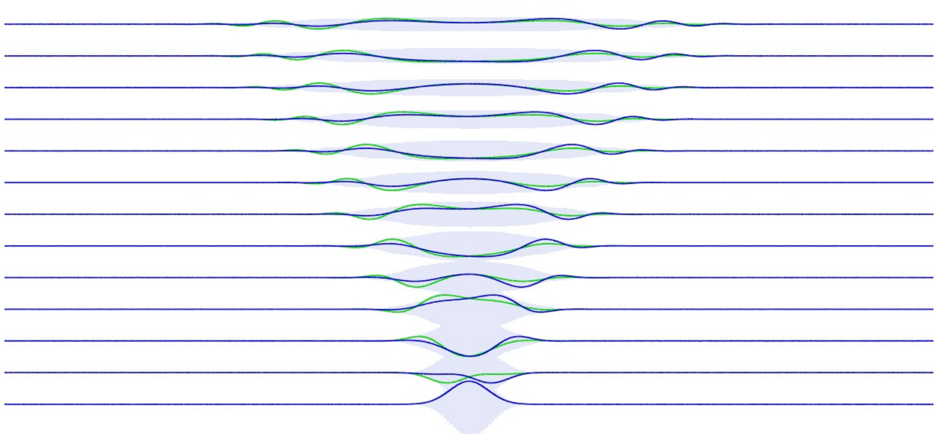

Figure 16.5: Particle at rest spreading fast

16.5 The chiral propagation in 4d

At this point we would like to say something about the propagation of the Dirac electron in the full 4d version of the theory. In fact we can already derive it without knowing the form of the full Dirac equation itself.

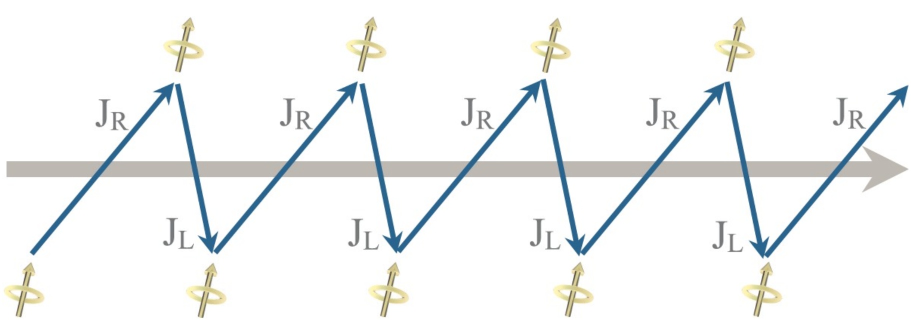

Figure 16.6: The two chiral components in 4d

We can go to the rest-frame of particle in the 2d version of the theory, then extend the the 2d wave to a plane wave in 4d, that is, each 2d point becomes a (constant) plane. Next we can go from the rest-frame in 4d to an arbitrary frame in 4d. Image 16.6 shows how the momenta of the left and right chiral components transform.

The role of the spin direction in 4d becomes apparent. We will see that the "spin" in 2d is aligned with the x-axis. In 4d however it can have any direction. There are still two chiral components in 4d coupled via the mass term. They reflect into each other via the "spin-plane", the plane orthogonal to the spin pointer. The two chiral components transform "light-like". They propagate with c while averaging out to the speed v of the particle. For ultra relativistic particles the spin itself transforms into a direction which is closely parallel to the direction of motion, either forward or backward, except for the case where the spin is exactly orthogonal to the direction of motion.

16.6 The derivation of the Dirac equation

We are now ready to extend our 2d Dirac equation to the full relativistic Dirac equation. The structure of the first will be wholly retained in the latter. We therefor write it out in full here.

The 2d Dirac equation:

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial t} + c \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \frac{\partial}{\partial r} \right] \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \frac{mc^2}{i\hbar} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (16.21)$$

The 4d Dirac equation by comparison is defined by:

$$\left[\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \frac{\partial}{\partial t} + c \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \frac{\partial}{\partial r^i} \right] \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \frac{mc^2}{i\hbar} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (16.22)$$

The correspondence reflects Dirac's original goal to construct a linear version of the Klein Gordon equation. If the operator between brackets in Dirac's equation is applied twice then we obtain the usual second order Klein Gordon equation for all individual components of the wave function ψ. The only thing which is new in the Dirac equation are the σ^i. If we square the operator between brackets then we obtain many cross products of these sigmas which should all cancel since the Klein Gordon equation does not contain mixed derivatives. It is clear that cross products of (non-zero) sigmas can only cancel in the case of matrices.

We found that if we apply the operator for the 2d Dirac equation twice we actually switch between the two components of the two wave functions so the "square" of the operator is actually the product of the ψ_L and the ψ_R operator:

$$\left(\frac{\partial}{\partial t} \; - \; c\frac{\partial}{\partial r} \right)\left(\frac{\partial}{\partial t} \; + \; c\frac{\partial}{\partial r} \right)\psi \;=\; \left(\frac{\partial^2}{\partial t^2} \; - \; c^2\frac{\partial^2}{\partial r^2} \right)\psi \qquad (16.23)$$

This will be our starting point for determining the σ^i. We go to the momentum space representation and set c to 1 merely to get a somewhat cleaner and compacter notation. So for the 2d Dirac equation we have.

$$(\; E - p_r)(\; E + p_r\;)\;\psi \;=\; (\; E^2 - p_r^2\;)\;\psi \qquad (16.24)$$

For the full 4d Dirac equation we need to include all three spatial dimensions therefor we do need:

$$(\; E - \sigma_i \cdot p_i\;)(\; E + \sigma_i \cdot p_i\;)\;\psi \;=\; (\; E^2 - p_x^2 - p_y^2 - p_z^2\;)\;\psi \qquad (16.25)$$

$$\sigma_i \cdot p_i \;=\; \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \qquad (16.26)$$

It is clear that sigma products with equal indices should yield a unity matrix while all the cross products should cancel. When Dirac derived his relativistic equation in a 1928 landmark paper he already knew of a set of 2x2 matrices which do just this. These matrices were the Pauli matrices introduced by Wolfgang Pauli a year earlier in his non relativistic extension of the Schroödinger equation.

16.7 Introduction of the Pauli spinors and matrices

Pauli had constructed a two component version of the Schroödinger equation which included the intrinsic spin of the electron. The Stern Gerlach experiments had demonstrated that the spin is found to be either up or down. Furthermore, electron scattering experiments show us that up and down polarized electrons do not interfere with each other, much in the same way that mixing horizontal and vertical polarized light doesn't produce interference patterns. Now, since we can describe polarized light with a two component function (H, V) we understand Pauli's idea to construct a to construct spin equation with two components $(S_{z\uparrow}, S_{z\downarrow})$: Spin up and down. A fundamental difference is that H and V are orthogonal with each other under 90^o while $S_{z\uparrow}$ and $S_{z\downarrow}$ are "orthogonal" under 180^o ...

$$s_{z\uparrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad s_{z\downarrow} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{16.27}$$

By choosing a particular coordinate for the spin-axis we expect an asymmetry in the handling of the coordinates. We can always rotate the coordinate system to align it with the spin, but there will be many cases with two or more different spins. It turns out that we can describe any arbitrary spin direction with the two component Pauli spinors. We expect a similar coordinate dependence in the objects used to manipulate these Pauli spinors, The Pauli σ matrices. This is indeed the case, as we will now finally turn our attention to them after this long intermezzo:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.28}$$

We see the expected asymmetry but we also see that σ^y is imaginary! We see what's happening here when we do a dot product of the position vector $\vec{r}$ with the Pauli matrices. This is a representative operation for a lot we'll see later on:

$$\vec{r} \cdot \vec{\sigma} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \tag{16.29}$$

It turns out that the x and y coordinates are grouped together into a single complex plane. This makes it easy to describe rotations in the xy-plane with the use of: $r \exp i\phi = x + iy$. This isn't bad at all since the z-axis is the preferred axis of rotation in this coordinate system. We can associate $(x + iy)$ and $(x - iy)$ with spin up and down.

We still seem to have a non relativistic theory here since we are missing t. The Dirac equation in the chiral representation, (The one we are treating here), however handles t on equal footing as x, y and z. If we go back to the Dirac equation (16.22) then we see that we can assign the I to σ^t

$$\sigma^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{16.30}$$

If we now write out the corresponding dot four-product then we find:

$$r^\mu \sigma^\mu \;=\; \begin{pmatrix} t + \;z & x - iy \\ x + iy & t - \;z \end{pmatrix} \tag{16.31}$$

Again, it seems, we have something to explain here. Why does t match up specifically with the z-axis? For now we can make the observation the spin undergoes a relativistic transformation only if there is a component of the speed v in the direction of the spin. That is, an object, either classical or quantum mechanically, with its spin in the z-direction, keeps the same spin vector it has in the rest frame, when it has an arbitrary velocity v but limited to the xy-plane.

We will see that, by using only Pauli matrices, we can rotate and boost the spin as it is incorporated in the Dirac equation in any relativistic way we want.

16.8 Pauli spin matrices as spin 1/2 operators

We still need to show why all the cross terms cancel if we use the Pauli matrices in the Dirac equation, and that by doing so we recover the Klein Gordon equation for all the sub-components of the Dirac equation. By working this out we'll automatically hit the subject of the coming sections: The Pauli matrices as the generators of the 3d rotation group. Now, the sigmas had to be solutions of the following equation:

$$(\ E - \sigma_i \cdot p_i\ )(\ E + \sigma_i \cdot p_i\ )\ \psi \;=\; (\ E^2 - p_x^2 - p_y^2 - p_z^2\ )\ \psi \tag{16.32}$$

$$\sigma_i \cdot p_i \;=\; \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \tag{16.33}$$

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.34}$$

We see indeed that they are a correct solution. The squares of sigmas are the unity matrices while the cross-products *anti-commute*.

$$\sigma^x\sigma^x = I, \quad \sigma^y\sigma^y = I, \quad \sigma^z\sigma^z = I \tag{16.35}$$

$$
\begin{array}{lllllll}
\sigma^x\ \sigma^y & = & i\sigma^z & \quad & \sigma^y\ \sigma^x & = & -i\sigma^z \\
\sigma^y\ \sigma^z & = & i\sigma^x & \quad & \sigma^z\ \sigma^y & = & -i\sigma^x \\
\sigma^z\ \sigma^x & = & i\sigma^y & \quad & \sigma^x\ \sigma^z & = & -i\sigma^y
\end{array}
\tag{16.36}
$$

Now, we get rid of the cross-terms in (16.32) because they always come in pairs which cancel each other as (16.36) shows. Commutation rules played a very important role in Quantum Mechanics from the days that people were just starting to discover it. There was already the general rule for angular momentum commutation in Heisenberg's matrix mechanics:

$$
\begin{array}{lllll}
J^x\ J^y & - & J^y\ J^x & = & i\ J^z \\
J^y\ J^z & - & J^z\ J^y & = & i\ J^x \\
J^z\ J^x & - & J^x\ J^z & = & i\ J^y
\end{array}
\tag{16.37}
$$

While the absolute value $|\mathbf{J}|^2 = J_x^2 + J_y^2 + J_z^2$ has the eigenvalue:

$$
|\mathbf{J}|^2 = J(J+1), \qquad L = 0,\ 1,\ 2\ ... \quad \text{or} \quad J = \frac{1}{2},\ \frac{3}{2},\ \frac{5}{2}\ ... \tag{16.38}
$$

Where the J should of course be angular momentum *operators* which are build from the position operators and the differential momentum operators: $J^i = x^j\partial^k - x^k\partial^j$. Now, although the Pauli spin matrices are of a different species, both (16.36) and (16.39) are satisfied if we declare the Pauli spin matrices to be the *spin momentum operators* of spin 1/2 particles like:

$$
J_{1/2} = s, \qquad \text{and} \qquad s^x = \frac{1}{2}\sigma^x,\ \ s^y = \frac{1}{2}\sigma^y,\ \ s^z = \frac{1}{2}\sigma^z \tag{16.39}
$$

The commutation rules (16.36) now fulfill those of (16.36) and the eigenvalue of s can be written as.

$$
|\mathbf{s}|^2 \ = \ s(s+1) \ = \ \frac{1}{2}(\frac{1}{2}+1) \ = \ \frac{3}{4} \tag{16.40}
$$

$$
= \ s_x^2 + s_y^2 + s_z^2 \ = \ \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \tag{16.41}
$$

So, mathematically this seems at least a very nice match. But there is another very convincing argument for the sigmas to be spin operators and the way Pauli include the Pauli spinors in Schrödinger's equations: The probability for a certain observable in Schrödinger's theory is $\Psi^*O\Psi$ where O is the operator for the observable. Now, if we define a two component wave function in the form of:

$$\psi \; = \; \xi\Psi, \qquad \text{where} \quad \xi \; = \; \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; .... \tag{16.42}$$

Where ξ in fact can be any spinor in any direction, then:

$$\psi^* \mathcal{S}\ \psi \; = \; (\xi\Psi)^* \ \mathcal{S}\ (\xi\Psi) \; = \; \xi^*\Psi^*\mathcal{S}\ \xi\ \Psi \; = \; (\xi^*\mathcal{S}\ \xi) \ \ \Psi^*\Psi \tag{16.43}$$

So the Pauli spin operator only operates on the Pauli spinors and the result is proportional to the probability density $\Psi^*\Psi$. If we now evaluate $(\xi^*\mathcal{S}\ \xi)$ using the spin operator for the x, y and z-componens, and let it operate on the up and down spins in the z-direction then we find exactly what we want to.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}^* \mathcal{S}^z \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; = \; \frac{1}{2}, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}^* \mathcal{S}^z \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; = \; -\frac{1}{2} \tag{16.44}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}^* \mathcal{S}^x \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; = \; 0, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}^* \mathcal{S}^x \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; = \; 0 \tag{16.45}$$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}^* \mathcal{S}^y \begin{pmatrix} 1 \\ 0 \end{pmatrix} \; = \; 0, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}^* \mathcal{S}^y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \; = \; 0 \tag{16.46}$$

We see that we have retrieved the components of the spin z-up and spin z-down spinors in the x, y and z-directions. We have extracted a normal 3d spin-vector from the spinor representation. This holds under all conditions. In general, the spin operators will always return the x, y and z components of the spinor, multiplied by one-half, thus:

$$\vec{S} \; = \; \frac{1}{2}\ \xi^*\vec{\sigma}\ \xi \tag{16.47}$$

Where $\vec{S}$ is a normal 3d vector. In this way we can transform spinors in spinor space back into normal vectors. Mathematically this is all very satisfying. In the coming sections we will look more into the physics of how Pauli spinors behave under rotation and the role of the spin matrices.

16.9 Spinor rotations over 720 degrees

We did already mention that electrons with spin up do not interfere with electrons with spin down, much like horizontal and vertical polarized light does not interfere. So we can define polarized light by its orthogonal (H,V) components, and electrons by its "orthogonal" (Up, Down) components. Orthogonal between quotes "" because Up and Down are under 180 degrees rather than 90 degrees.

rotating the spin-axis over 720 degrees

This angle or phase "doubling" is a very general property of all fermions (spin 1/2 particles). If we physically manipulate an electron to rotate its spin, or equivalently, its magnetic moment, (which isn't really hard to do), then we find something quiet interesting.

When we rotate the magnetic moment of the electron first by 180 degrees from up to down, and then further to Up again, then, after having rotated it by 360 degrees then we expect physically the same electron back. Now this is not true, in fact the 360 degrees rotated electron has been inverted in some sense because it would *negatively* interfere if we would mix it with a similar non rotated electron (or with a non-rotated part of it's own wave function). If we continue to rotate the electron via 540 degrees (spin down) and further to 720 degrees (spin up again), only then we find a physically identical electron back!

Although admittedly strange, we can imagine such a behavior classically because we have assigned a spin to the electron, and through its gyromagnetic moment it is quite "aware" that it is being rotated. However, going one step further, we see that a simple classical spin only is not yet enough: The same also happens if the is electron rotates around its own spin axis:

rotation around the spin-axis of 720 degrees

An electron rotated 360 degrees around its spin axis is also its own inverse in the sense of interference. A simply spinning classical object's gyromagnetic moment does not interact with a rotation around the spin-axis. Now, this again is not so much of a problem since all attempts to model the electron classically involve *precession*. And in fact it has to do so classically since the relation of the spin component value of 1/2 and the total spin of $\sqrt{s(s+1)}$ leads to a precession around the principle spin axis with an angle of:

$$\theta \;=\; \cos^{-1}\left(\frac{\frac{1}{2}}{\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)}}\right) \;=\; 54.7356...\text{ degrees} \tag{16.48}$$

Interesting as it is, we leave the classical equivalent for what it is here for the time being. The message is that, both classically and quantum mechanically, one can argue that the electron has the means to detect rotations around any axis and that such a rotation could have consequently the physical results we see. It is not necessary that we, for instance, should revert to the conclusion that the 720 degrees behavior of fermions is the result of some unclear property of the "structure" of space-time.

16.10 The spin direction and the phase of the spinor

The fact that the rotation of a spinor around its own spin axis has physical consequences, and also separates spinors from the simple three component spin vector. A spinor embodies more information as the 3d vector does. It doesn't only encompasses a spin-direction but also a spin phase. This means that a particle with spin up in the z-direction can have different spinors which point all in the same direction.

$$0^\circ = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\; 180^\circ = \begin{pmatrix} -i \\ 0 \end{pmatrix},\; 360^\circ = \begin{pmatrix} -1 \\ 0 \end{pmatrix},\; 540^\circ = \begin{pmatrix} i \\ 0 \end{pmatrix} \tag{16.49}$$

$$\xi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-ip\cdot k} \quad\Rightarrow\quad \frac{1}{2}\,\xi^*\vec{\sigma}\,\xi \;=\; \left\{\,0,\ 0,\ \frac{1}{2}\,\right\} \tag{16.50}$$

We see that the rotation is indeed twice as fast as the phase changes would suggest. However, we do not see double frequencies in interference experiments since it's the "360° part" which interferes with the "720° part". We see exactly the interference frequencies which we expect from the phase change since both effects cancel.

Pauli spin matrices as 180 degrees rotation operators

We can define 180° spinor rotations around any axis directly with Pauli's spin matrices. σ_x and σ_y reverse the spin in the z-direction so equation

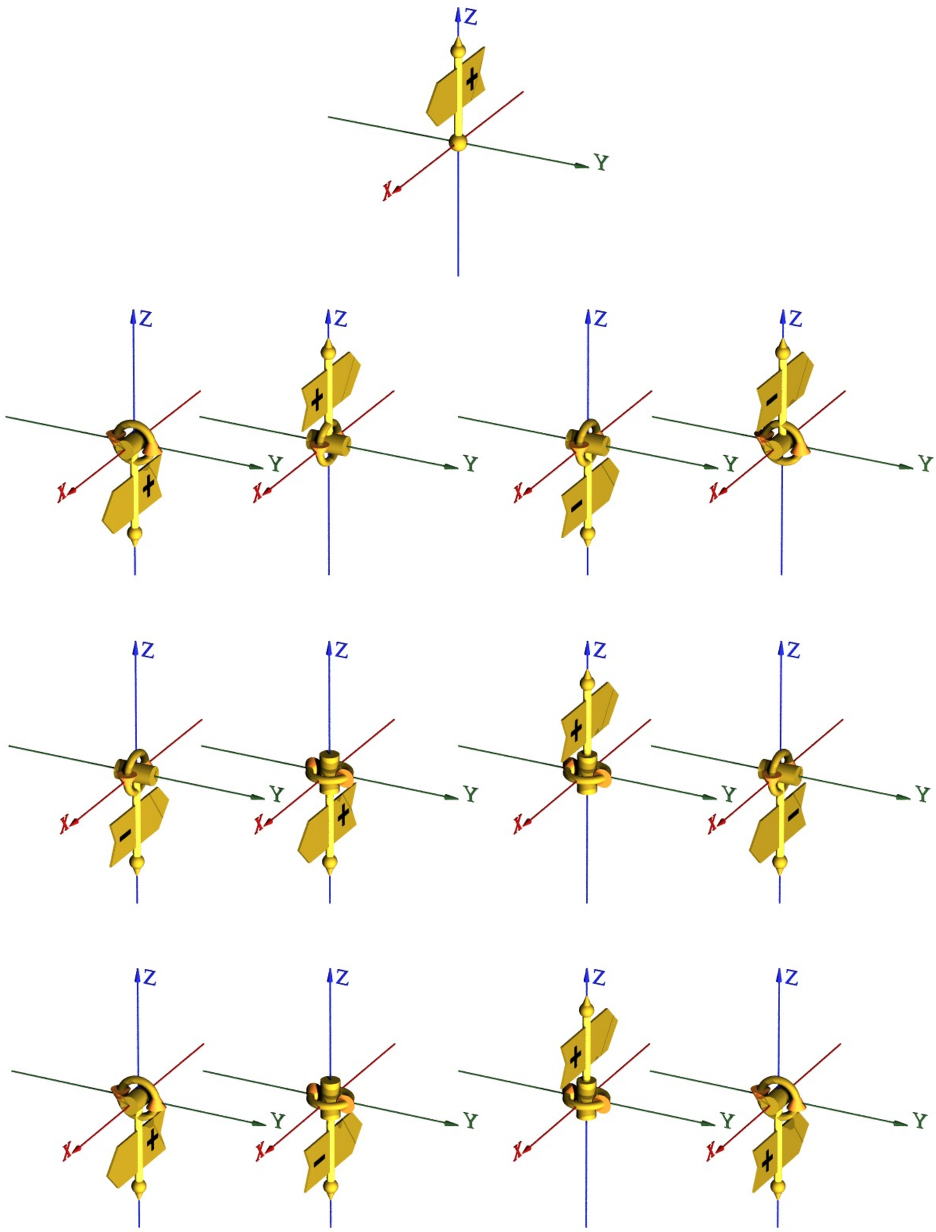

Figure 16.7: Spinor rotations (spin direction and phase)

(16.47) yields zeros for the x and y components of a particle with its spin in the z-direction.

$$\begin{array}{rcl} i\sigma^x &=& -180^\circ \text{ rotation around the } x\text{-axis} \\ i\sigma^y &=& -180^\circ \text{ rotation around the } y\text{-axis} \\ i\sigma^z &=& -180^\circ \text{ rotation around the } z\text{-axis} \end{array} \qquad (16.51)$$

Here we recover the geometric rule that a rotation of 180° around the x-axis followed by a 180° rotation around the y-axis is the same as a 180° rotations around the z-axis. Rotations over 180 degrees and -180 degrees map to the same direction but differ in the overall sign. All 6 combinations are summarized below. (Note that the matrices operate from right to left)

$$\begin{array}{rclrcl} i\sigma^y\, i\sigma^x &=& i\sigma^z & \quad i\sigma^x\, i\sigma^y &=& -i\sigma^z \\ i\sigma^z\, i\sigma^y &=& i\sigma^x & \quad i\sigma^y\, i\sigma^z &=& -i\sigma^x \\ i\sigma^x\, i\sigma^z &=& i\sigma^y & \quad i\sigma^z\, i\sigma^x &=& -i\sigma^y \end{array} \qquad (16.52)$$

All six cases are shown in fig.(16.7), which shows both the spin directions as well as the spinor phase. They are also written out in detail for reference below in (16.53). The image uses + and − signs to distinguish between spinors with the same direction and phase but rotated over 360° and thus each others inverse.

$$\begin{array}{rcccccccl}
i\sigma^y\, i\sigma^x &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ i \end{pmatrix} &\rightarrow& \begin{pmatrix} i \\ 0 \end{pmatrix} &=& i\sigma^z \\
i\sigma^x\, i\sigma^y &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ -1 \end{pmatrix} &\rightarrow& \begin{pmatrix} -i \\ 0 \end{pmatrix} &=& -i\sigma^z \\
\\
i\sigma^z\, i\sigma^y &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ -1 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ i \end{pmatrix} &=& i\sigma^x \\
i\sigma^y\, i\sigma^z &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} i \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ -i \end{pmatrix} &=& -i\sigma^x \\
\\
i\sigma^x\, i\sigma^z &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} i \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ -1 \end{pmatrix} &=& i\sigma^y \\
i\sigma^z\, i\sigma^x &=& \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ i \end{pmatrix} &\rightarrow& \begin{pmatrix} 0 \\ 1 \end{pmatrix} &=& -i\sigma^y
\end{array} \qquad (16.53)$$

16.11 General spinor rotations

It is a small step to general rotation expressions once we have the rotation over 180°. Since continuous rotation is a repetition of infinitesimal rotations we can write the general rotations over the various axis as exponential functions.

$$\begin{array}{rcl} \exp\left(\ -i\ \phi/2\ \ \sigma^x\ \right) & = & \phi \text{ degrees rotation around the } x\text{-axis} \\ \exp\left(\ -i\ \phi/2\ \ \sigma^y\ \right) & = & \phi \text{ degrees rotation around the } y\text{-axis} \\ \exp\left(\ -i\ \phi/2\ \ \sigma^z\ \right) & = & \phi \text{ degrees rotation around the } z\text{-axis} \end{array} \tag{16.54}$$

The exponential functions are 2x2 matrices with 2x2 matrices in the argument. It is however simple to get an expression without this complication if we look at the series development.

$$\exp\left(i\frac{\phi}{2}\sigma^i\right) = I - i\left(\frac{\phi}{2}\sigma^i\right) - \frac{1}{2!}\left(\frac{\phi}{2}\sigma^i\right)^2 + \frac{i}{3!}\left(\frac{\phi}{2}\sigma^i\right)^3 + ... \tag{16.55}$$

All even powers of the sigmas are unity matrices well all the odd powers are consequently simply the Pauli spin matrices. The exponential function therefor splits in a cosine and a sine function with no more matrices in the arguments.

$$\exp\left(-\frac{i}{2}\phi\ \sigma^i\right) = I\cos\left(-\frac{\phi}{2}\right) + i\sigma^i\sin\left(-\frac{\phi}{2}\right) \tag{16.56}$$

We can now write the general rotation operators directly as 2x2 matrices.

$$\exp\left(-\frac{i}{2}\phi\ \sigma^x\right) = \begin{pmatrix} \cos(-\phi/2) & i\sin(-\phi/2) \\ i\sin(-\phi/2) & \cos(-\phi/2) \end{pmatrix} \tag{16.57}$$

$$\exp\left(-\frac{i}{2}\phi\ \sigma^y\right) = \begin{pmatrix} \cos(-\phi/2) & \sin(-\phi/2) \\ -\sin(-\phi/2) & \cos(-\phi/2) \end{pmatrix} \tag{16.58}$$

$$\exp\left(-\frac{i}{2}\phi\ \sigma^z\right) = \begin{pmatrix} \exp(-i\phi/2) & 0 \\ 0 & \exp(+i\phi/2) \end{pmatrix} \tag{16.59}$$

Rotations around z in the x+iy plane are given by complex exponentials as we expected while rotations around y in the (x,z) plane are given by the classical 2x2 rotation matrix.

16.12 The wave's phase as the spinor phase

We can now show that the spin up and down spinors in the other directions are always the eigenfunctions of the rotation operators while the phase of the wave function is the eigenvalue of the rotation operators. Using $2\omega = \phi$ we can write down:

$$\text{——— rotation of x-spinors around the x-axis ———} \tag{16.60}$$

$$\sqrt{2}\; x\uparrow : \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega} \quad = \quad \begin{pmatrix} \cos(-\omega) & i\sin(-\omega) \\ i\sin(-\omega) & \cos(-\omega) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\sqrt{2}\; x\downarrow : \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{+i\omega} \quad = \quad \begin{pmatrix} \cos(-\omega) & i\sin(-\omega) \\ i\sin(-\omega) & \cos(-\omega) \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\text{——— rotation of y-spinors around the y-axis ———} \tag{16.61}$$

$$\sqrt{2}\; y\uparrow : \quad \begin{pmatrix} 1 \\ i \end{pmatrix} e^{-i\omega} \quad = \quad \begin{pmatrix} \cos(-\omega) & \sin(-\omega) \\ -\sin(-\omega) & \cos(-\omega) \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$\sqrt{2}\; y\downarrow : \quad \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{+i\omega} \quad = \quad \begin{pmatrix} \cos(-\omega) & \sin(-\omega) \\ -\sin(-\omega) & \cos(-\omega) \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

$$\text{——— rotation of z-spinors around the z-axis ———} \tag{16.62}$$

$$z\uparrow : \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-i\omega} \quad = \quad \begin{pmatrix} \exp(-i\omega) & 0 \\ 0 & \exp(+i\omega) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$z\downarrow : \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{+i\omega} \quad = \quad \begin{pmatrix} \exp(-i\omega) & 0 \\ 0 & \exp(+i\omega) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We see that a spinor always rotates (clock wise) around its own axis as the phase ω of the wave function increases, and counter-clock-wise when the phase goes as $-\omega$. It rotates *twice* as fast in space as the phase ω of the wave function progresses. Nevertheless, interference patterns do have the same wavelength since 360° instead of 180° rotated parts of the wave function interfere. Remember the spinor direction is local at each point of the wave function. There is no global rotation, but the spin is distributed all over the wave function.

16.13 Rotation operator of the full chiral bi-spinor

We can now write down the complete rotation operator for the (bi-spinor) Dirac wave functions:

$$\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \overset{\text{rotate}}{\Rightarrow} \exp\left\{ -\phi\frac{i}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}\right\} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.63}$$

$$\xi_{(\phi,\theta,\alpha)} = e^{i\alpha/2} \begin{pmatrix} e^{-i\phi/2} \cos(\theta/2) \\ e^{+i\phi/2} \sin(\theta/2) \end{pmatrix} \tag{16.64}$$

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.65}$$

Both ψ_L and ψ_R are transformed (individually) in the just same way as the single spinor cases we studied in this section. We can transform this expression inside-out into (block-diagonal) 4x4 matrices with the method we used above.

16.14 Weyl's chiral bi-spinors and relativity

We can now finally return to the full Dirac equation after having exhaustedly treated the physics of Pauli's spinors and spin matrices. Note that the treatment of spin was entirely non-relativistic. We will now include the time dimension and look at the relativistic aspects. We'll be pleasantly surprised when we find out how close we already are with what we have done so far.

Recall our first presentation of Dirac's equation (16.22) We can simplify the notation by using four-vectors and setting c to 1. We get for the Dirac equation:

$$\begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix} \frac{\partial}{\partial x^\mu} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \frac{mc^2}{i\hbar} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.66}$$

with the spin matrices:

$$\sigma^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and

$$\sigma^\mu = (\ \sigma^t,\ \sigma^x,\ \sigma^y,\ \sigma^z\ ), \qquad \tilde{\sigma}^\mu \ = \ (\ \sigma^t,\ -\sigma^x,\ -\sigma^y,\ -\sigma^z\ ) \tag{16.67}$$

Where the spin matrix for time is simply the unity matrix. We can write the equation even more compact with the use of the gamma matrices.

$$\gamma^\mu \partial^\mu \psi = m\psi, \qquad \text{where} \ \ \gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix} \tag{16.68}$$

Mathematically, since we have figured out how to do a rotation in a plane spanned by two (spatial) coordinate axis, and, a boost can be viewed as a skewed rotation in a plane spanned by time and the direction of the boost, we might suspect that boost are just around the corner. Indeed, the boost and rotation generators can be written in a simple combined form:

$$S^{\mu\nu} \ = \ \frac{1}{4}\Big[\gamma^\mu\ ,\ \gamma^\nu\Big] \ = \ \frac{i}{4}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{16.69}$$

Where the rotation generators are given by:

$$S^{ij} \ = \ S^k \ = \ \frac{1}{4}\Big[\gamma^i\ ,\ \gamma^j\Big] \ = \ -\frac{i}{2}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{16.70}$$

Which can be compared with equation (16.63) for the general rotation operator. The boost generators are.

$$S^{0i} \ = \ \frac{1}{4}\Big[\gamma^0\ ,\ \gamma^i\Big] \ = \ -\frac{1}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix} \tag{16.71}$$

So, similar to equation (16.63) we can now define the general boost operator for 4d Dirac spinors with ϑ as the *boost* which relates to the speed as $\beta = \tanh\vartheta$.

$$\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \ \overset{\text{boost}}{\Rightarrow} \ \exp\left\{ -\vartheta\frac{1}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}\right\} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.72}$$

16.15 The 2d spinor boost operators

The Pointcaré group provides a high level abstract mathematical frame work. However, the unique insights it offers are merely a guide on the journey to understand the physical meaning of the mathematics representing the Dirac equation. First we want to take a step back to the simpler 2d Dirac equation and study its physical behavior under boosts. All of what we find here applies to the Dirac equation as well, and can be straightforwardly reused in the case of the Dirac equation. We will find that many aspects normally associated with Dirac's 4d equation only apply to the 2d Dirac equation as well. The 2d Dirac equation's equivalent of (16.90) becomes.

$$\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad \overset{\text{boost}}{\Rightarrow} \quad \exp\left\{ -\frac{\vartheta}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.73}$$

Plane wave solutions of the 2d Dirac equation

We split the wave function ψ in a two-component chiral pair $u = (u_L, u_R)$ and a wave function ϕ which is a solution of the normal one component Klein Gordon equation.

$$\begin{pmatrix} u_L \\ u_R \end{pmatrix} \phi \quad \overset{\text{boost}}{\Rightarrow} \quad \exp\left\{ -\frac{\vartheta}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \begin{pmatrix} u_L \\ u_R \end{pmatrix} \phi \tag{16.74}$$

$$\phi \;=\; \exp\left(-i\,\frac{Et}{\hbar} \;+\; i\,\frac{p^r r}{\hbar} \right) \tag{16.75}$$

$$\begin{pmatrix} u_L \\ u_R \end{pmatrix} \quad \overset{\text{at rest}}{=} \quad \sqrt{m} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{16.76}$$

The chiral pair (u_L, u_R) is constant throughout the wave function of the plane wave solution. The normalization $\sqrt{m}$ is based on $|\psi|^2 = m$, the rest mass which is a Lorentz invariant scalar. We can bring the 2x2 matrix in the argument of the exponential function to the outside in a similar way we did in (16.55) because its powers cycle through the unity matrix.

$$\exp\left\{ -\frac{\vartheta}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\} \begin{pmatrix} u_L \\ u_R \end{pmatrix} = \begin{pmatrix} \exp(-\vartheta/2) & u_L \\ \exp(+\vartheta/2) & u_R \end{pmatrix} \qquad (16.77)$$

The diagonal matrix elements 1 and -1 become $\exp(\vartheta/2)$ and $\exp(-\vartheta/2)$ and lead to the right hand term of (16.77) At this point we can see how ψ_L and ψ_R are transformed individually. Can we already make an interpretation here even though these are just subcomponents? Well, it turns out that to some extent indeed we can do so. We started out with ψ_L and ψ_R as left and right moving solutions on the light cone which become coupled via the mass m. The coupling causes oscillation and the oscillation causes interference which strongly suppresses the propagation on the light cone and the wave function remains semi localized in case of a free particle and fully localized in bounded states.

We get the answers we want if we still assume that ψ_L and ψ_R are basically mass less components, moving on the light cone to the left and the right of the wave function moving with a speed β. We expect the energy and momenta to be determined by the red- and blue-shift expressions for a particle moving at speed β.

relativistic blue shift

$$\exp(+\vartheta) = \sqrt{\frac{1+\beta}{1-\beta}} = \frac{1}{\sqrt{1-\beta^2}} + \frac{\beta}{\sqrt{1-\beta^2}} = \frac{E+p^r}{m} \qquad (16.78)$$

relativistic red shift

$$\exp(-\vartheta) = \sqrt{\frac{1-\beta}{1+\beta}} = \frac{1}{\sqrt{1-\beta^2}} - \frac{\beta}{\sqrt{1-\beta^2}} = \frac{E-p^r}{m} \qquad (16.79)$$

This corresponds indeed with $|\exp(\vartheta/2)|^2$ and $|\exp(-\vartheta/2)|^2$. Massless objects have a fixed relation between energy and momentum in the form of the constant c. Assuming that each of the components contributes half of the energy at rest we can write down:

$$\psi_L : \quad E_L = -\,c p_L^r = \exp(-\vartheta)\,\frac{1}{2}m \qquad (16.80)$$

$$\psi_R : \qquad E_R = +\, cp_R^r = \exp(+\vartheta)\, \frac{1}{2} m \tag{16.81}$$

Which we can simply add to obtain the total energy and momentum. We see that this correspond with what we should expect.

$$E = E_L + E_R = \cosh(\vartheta)\, m = \frac{m}{\sqrt{1-\beta^2}} \tag{16.82}$$

$$p^r = p_L^r + p_R^r = \sinh(\vartheta)\, m = \frac{m\beta}{\sqrt{1-\beta^2}} \tag{16.83}$$

So, regardless of the interpretation, we have obtained a simple method here to memorize the mathematical steps we took in order to check the expression from which we started. We can write the behavior of our chiral pair (u_L, u_R) under boost in a way familiar from the free plane wave solutions of the Dirac equation with the help of (16.78) and (16.79).

$$\begin{pmatrix} u_L \\ u_R \end{pmatrix} \overset{\text{boost}}{\Rightarrow} \begin{pmatrix} \sqrt{(E-p^r)} \\ \sqrt{(E+p^r)} \end{pmatrix} \tag{16.84}$$

At large boost only one of the two chiral components remains depending on the direction of the boost.

$$\begin{pmatrix} \sqrt{(E-p^r)} \\ \sqrt{(E+p^r)} \end{pmatrix} \overset{\text{large boost in +r}}{\Rightarrow} \begin{pmatrix} 0 \\ \sqrt{2E} \end{pmatrix} \tag{16.85}$$

$$\begin{pmatrix} \sqrt{(E-p^r)} \\ \sqrt{(E+p^r)} \end{pmatrix} \overset{\text{large boost in -r}}{\Rightarrow} \begin{pmatrix} \sqrt{2E} \\ 0 \end{pmatrix} \tag{16.86}$$

Anti particles of the 2d Dirac equation

Anti particles are distinguished from normal particles by the sign in frequency of the wave function. We have seen that ψ_L and ψ_R are mutually responsible for each others phase change.

$$\phi = \exp\left(+i\, \frac{Et}{\hbar} + i\, \frac{p^r r}{\hbar} \right) \tag{16.87}$$

$$\begin{pmatrix} v_L \\ v_R \end{pmatrix} \overset{\text{at rest}}{=} \sqrt{m} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{16.88}$$

By changing the sign of ψ_R the coupling between ψ_L and ψ_R also changes sign with a change in the sign of the frequency as the result. The transform under boost of v is.

$$\begin{pmatrix} v_L \\ v_R \end{pmatrix} \overset{\text{boost}}{\Rightarrow} \begin{pmatrix} \sqrt{(E-p^r)} \\ -\sqrt{(E+p^r)} \end{pmatrix} \tag{16.89}$$

Where the factors are the same as for particles with the sign as a result of the definition.

16.16 The 4d spinor boost operators

We can now go to the full 4d spinor boost operator.

$$\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \overset{\text{boost}}{\Rightarrow} \exp\left\{ -\vartheta\frac{1}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}\right\} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{16.90}$$

Just like we did for the rotation operator we can move the exponential function into the block matrix so that the operations on ψ_L and ψ_R become explicitly separated.

$$\begin{pmatrix} \exp\left(-\ \vartheta/2\ \ \sigma^i\right)\ \psi_L \\ \\ \exp\left(+\ \vartheta/2\ \ \sigma^i\right)\ \psi_R \end{pmatrix} = \text{boost of } \vartheta \text{ along the } x^i\text{-axis} \tag{16.91}$$

The exponential functions are now reduced from 4x4 to 2x2 matrices (with 2x2 matrices in the argument). It is however simple to go one step further and remove the matrix arguments from the exponential altogether. We look at the series development.

$$\exp\left(\frac{\vartheta}{2}\,\sigma^i\right) = I + \left(\frac{\vartheta}{2}\,\sigma^i\right) + \frac{1}{2!}\left(\frac{\vartheta}{2}\,\sigma^i\right)^2 + \frac{1}{3!}\left(\frac{\vartheta}{2}\,\sigma^i\right)^3 + \ldots \tag{16.92}$$

All even powers of the sigmas are unity matrices well all the odd powers are consequently simply the Pauli spin matrices. The exponential function therefor splits in a cosine and a sine function with no more matrices in the arguments.

$$\exp\left(\frac{\vartheta}{2}\,\sigma^i\right) \;=\; I\cosh\,(\frac{\vartheta}{2}) \;+\; \sigma^i\sinh\,(\frac{\vartheta}{2}) \tag{16.93}$$

We can now write the general boost operators directly as explicit 2x2 matrices.

$$\exp\left(\frac{\vartheta}{2}\,\sigma^x\right) \;=\; \begin{pmatrix} \cosh(\vartheta/2) & \sinh(\vartheta/2) \\ \sinh(\vartheta/2) & \cosh(\vartheta/2) \end{pmatrix} \tag{16.94}$$

$$\exp\left(\frac{\vartheta}{2}\,\sigma^y\right) \;=\; \begin{pmatrix} \cosh(\vartheta/2) & i\sinh(\vartheta/2) \\ -i\sinh(\vartheta/2) & \cosh(\vartheta/2) \end{pmatrix} \tag{16.95}$$

$$\exp\left(\frac{\vartheta}{2}\,\sigma^z\right) \;=\; \begin{pmatrix} \exp(\vartheta/2) & 0 \\ 0 & \exp(-\vartheta/2) \end{pmatrix} \tag{16.96}$$

We can show that a 2-spinor pointing in a certain direction is the eigenfunction of the boost operator in that direction while the $\exp(\vartheta/2)$ is the eigenvalue of the boost operator. Using $\varpi = \vartheta/2$ we can write down:

——————- boost x-spinor along the x-axis ——————- (16.97)

$$\sqrt{2}\;x\uparrow: \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{+\varpi} \;=\; \begin{pmatrix} \cosh(\varpi) & \sinh(\varpi) \\ \sinh(\varpi) & \cosh(\varpi) \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\sqrt{2}\;x\downarrow: \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-\varpi} \;=\; \begin{pmatrix} \cosh(\varpi) & \sinh(\varpi) \\ \sinh(\varpi) & \cosh(\varpi) \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

——————- boost y-spinor along the y-axis ——————- (16.98)

$$\sqrt{2}\;y\uparrow: \quad \begin{pmatrix} 1 \\ i \end{pmatrix} e^{+\varpi} \;=\; \begin{pmatrix} \cosh(\varpi) & i\sinh(\varpi) \\ -i\sinh(\varpi) & \cosh(\varpi) \end{pmatrix}\begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$\sqrt{2}\;y\downarrow: \quad \begin{pmatrix} 1 \\ -i \end{pmatrix} e^{-\varpi} \;=\; \begin{pmatrix} \cosh(\varpi) & i\sinh(\varpi) \\ -i\sinh(\varpi) & \cosh(\varpi) \end{pmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix}$$

——————- boost z-spinor along the z-axis ——————- (16.99)

$$z\uparrow: \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{+\varpi} \;=\; \begin{pmatrix} \exp(\varpi) & 0 \\ 0 & \exp(-\varpi) \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$z\downarrow: \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-\varpi} \;=\; \begin{pmatrix} \exp(\varpi) & 0 \\ 0 & \exp(-\varpi) \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Plane wave solutions of the 4d Dirac equation

When we know the general boost operator then we automatically know the 4d plane wave solutions, the eigen-functions of the 4d Dirac equation. For a plane-wave solution the left and right 2-spinors ξ are equal in the rest frame. After a boost they become different

$$\psi \;=\; \begin{pmatrix} \xi \\ \xi \end{pmatrix} e^{-iE_o t/\hbar} \quad \overset{\text{boost}}{\Rightarrow} \quad \begin{pmatrix} \xi'_L \\ \xi'_R \end{pmatrix} e^{-iEt/\hbar + i\vec{p}\cdot\vec{x}/\hbar} \tag{16.100}$$

We recall equation (16.91) here for the general boost operator.

$$\begin{pmatrix} \psi_L \\ \\ \psi_R \end{pmatrix} \quad \overset{\text{boost}}{\Rightarrow} \quad \begin{pmatrix} \exp\Big(-\ \vartheta/2\ \ \sigma^i\Big)\ \ \psi_L \\ \\ \exp\Big(+\ \vartheta/2\ \ \sigma^i\Big)\ \ \psi_R \end{pmatrix} \tag{16.101}$$

Looking at (16.93) we see that we can write the components as.

$$\exp\left(\pm\frac{\vartheta}{2}\ \sigma^i\right) \;=\; \sqrt{\cosh(\vartheta)I\ \pm\ \sinh(\vartheta)\sigma^i} \tag{16.102}$$

So that we can write the general boost operator in a compact form.

$$\sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix} e^{-iE_o t/\hbar} \quad \overset{\text{boost}}{\Rightarrow} \quad \begin{pmatrix} \sqrt{p_\mu\cdot\sigma}\ \ \xi \\ \sqrt{p^\mu\cdot\sigma}\ \ \xi \end{pmatrix} e^{-iEt/\hbar + i\vec{p}\cdot\vec{x}/\hbar} \tag{16.103}$$

With $p^\mu = (\ p^t, p^x, p^y, p^z\ )$, and $p_\mu = (\ p^t, -p^x, -p^y, -p^z\ )$. We can thus write for the bi-spinors of the plane-wave eigen-function for the Dirac particle and its anti-particle.

$$u(p) \;=\; \begin{pmatrix} \sqrt{p_\mu\cdot\sigma}\ \ \xi \\ \sqrt{p^\mu\cdot\sigma}\ \ \xi \end{pmatrix}, \qquad v(p) \;=\; \begin{pmatrix} +\sqrt{p_\mu\cdot\sigma}\ \ \xi \\ -\sqrt{p^\mu\cdot\sigma}\ \ \xi \end{pmatrix} \tag{16.104}$$

Where the sign of the right chiral component changes because the coupling between the two components changes from m to -m.