PHYSICS 391
SPRING 2018

QUANTUM FIELD THEORY II

Jean Weill

INDEPENDENT STUDY PAPER

Professor S.G. Rajeev
20th April 2018

Contents

1 Introduction

The research worker, in his efforts to express the fundamental laws of Nature in mathematical form, should strive mainly for mathematical beauty. He should still take simplicity into consideration in a subordinate way to beauty.

Paul Dirac (1902-1984)

The objective of this course was to deepen our understanding of quantum field theory, one of the most important topics in modern physics. This paper is based on a lecture I gave for the Kapitza Society.

This lecture objective was to take a deep look into Dirac Equation: why do we need it, where does it come from, what does it mean, etc. All of it from the most ground-up approach possible. First let's see why Dirac Equation is so important.

Suppose you want to describe the behavior of a relativistic particle. From your Quantum Mechanics course, you might guess that the Schroedinger Equation would do the trick. The Schroedinger Equation has a lot of qualities, the most important being that it leads to a positive probability density, which makes it ideal for non-relativistic systems. But it has also a lot of flaws. If you remember your relativity lectures, we must treat space and time on the same footing. Here is where the Schroedinger Equation hits a wall, since it's of first order in time, but of second order in space! Therefore, it cannot be used to describe the behavior of a relativistic particle.

As we have seen a few lectures ago, this is where the Klein Gordon Equation comes in handy. By starting with the relativistic equation

$$p_\mu p^\mu = m^2 \tag{1}$$

making the substitution

$$p_\mu = i\partial_\mu \tag{2}$$

where we have set $\hbar = 1$, and acting on the wave-function ψ, we get the relativistic equation

$$(\partial_\mu \partial^\mu + m^2)\psi = 0 \tag{3}$$

called the Klein Gordon Equation. This equation is of second order in both time and space, and therefore does put time and space on the same footing, so we're off to a good start. Obviously, if this equation didn't lead to any problems, we would be done here... I do need to write a 10 page paper, so I'm glad it's not the case, or I would not be able to pass this class. The big problem with the Klein Gordon Equation is that it can lead to **negative** probability density. You should know that one of the few axioms of probability is that the probability of an event occurring must be between 0 and 1. So negative probability is bad, really bad. This problem came from the fact that the Klein Gordon equation is of second order in both time and space. Remembering that the Schroedinger didn't have this problem and is of first order in time, we can guess that an equation that is of first order in both time and space would have the best of both worlds. Meaning it would be able to describe a relativistic system, and it would have a positive probability density. Here is where Dirac genius happened.

2 The Dirac Equation

2.1 Derivation From Scratch

The Dirac Equation has to be relativistic, and so a logical place to start our derivation is equation (1). If you're wondering where equation (1) comes from, it's quite simple. When you think of physics, one of the first equations that comes to mind is the incredibly famous

$$E = mc^2 \tag{4}$$

This equation gives the energy of a particle of mass m at rest. If the particle is moving with momentum $|\vec{p}|$, this equation becomes the more general

$$E^2 = (mc^2)^2 + (|\vec{p}|c)^2 \tag{5}$$

which does reduce to equation (4) when $|\vec{p}| = 0$, like it should. For convenience, we let $c = 1$. Then equation (5) becomes

$$E^2 = m^2 + |\vec{p}|^2 \tag{6}$$

This still doesn't really look like equation (1). What if we let

$$p^\mu = (E, \vec{p}) \tag{7}$$

and use the Minkowski metric

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{8}$$

Then it follows that

$$p^2 = \eta_{\mu\nu} p^\mu p^\nu = E^2 - |\vec{p}|^2 = m^2 \tag{9}$$

and we get back equation (1). We want to transform this equation to get first order in time and space. Looking back at (2), this means that we need to isolate p^μ somehow. How can we do this from equation (1)? We have

$$p_\mu p^\mu - m^2 = 0$$

It's very tempting to factor it out to

$$(p_\mu - m)(p^\mu + m) = 0 \tag{10}$$

then we would be done since here p^μ is isolated and therefore gives our so desired first order in time and space equation. But multiplying terms, (10) becomes

$$p_\mu p^\mu - m^2 + m(p_\mu - p^\mu) = 0 \tag{11}$$

First of all, this equation doesn't even make sense dimension-wise. The first two terms are scalars, and the last one is a four vector. But even if we forget about this, in order for this equation to be equivalent to equation (1), it must be that $p^\mu = (E, \vec{p})$ equals to $p_\mu = (E, -\vec{p})$, which means that $\vec{p} = -\vec{p}$, or $\vec{p} = \vec{0}$. This is obviously not very general... One can argue that we could change our frame of reference in order to always have the particle at rest, but we want to make this discussion as general as possible, and maybe use this frame of reference trick from time to time to have a better grasp of the physical meaning of this equation.

We want to find a way to make equation (11) coherent. To do this, we need to transform the third term into a scalar. An obvious way is to make the following transformations

$$\begin{aligned} p_\mu &\to \alpha^\nu p_\nu \\ p^\mu &\to \beta_\delta p^\delta \end{aligned} \tag{12}$$

Thus,

$$(p_\mu - m)(p^\mu + m) \to (\alpha^\nu p_\nu - m)(\beta_\delta p^\delta + m) = 0 \tag{13}$$

which must equal to $p_\mu p^\mu - m^2$.
Multiplying terms, we get the following two equations

$$\begin{aligned} \alpha^\nu p_\nu &= \beta_\delta p^\delta \\ \alpha^\nu p_\nu \beta_\delta p^\delta &= p_\mu p^\mu \end{aligned} \tag{14}$$

The first equation leads to $\alpha = \beta = \gamma$. Plugging this in the second equation of (14), using the identity $A_\mu B^\mu = A^\mu B_\mu$, and noting that p and γ act on different spaces and therefore commute, we get

$$\gamma^\nu p_\nu \gamma_\delta p^\delta = p_\mu p^\mu \iff \gamma^\alpha \gamma^\delta p_\alpha p_\delta = \eta^{\alpha\delta} p_\alpha p_\delta \tag{15}$$

Therefore, $\gamma^\mu \gamma^\nu = \eta^{\mu\nu} \Leftrightarrow \gamma^\nu \gamma^\mu = \eta^{\nu\mu} = \eta^{\mu\nu}$, which leads to

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}$$

which we can write more concisely as

$$\boxed{[\gamma^\mu, \gamma^\nu]_+ = 2\eta^{\mu\nu}}$$

This boxed equation describes a type of Algebra called a Clifford Algebra.
Going back to (13), and putting the transformation we derived from scratch, we end up with

$$(\gamma^\mu p_\mu - m)(\gamma_\mu p^\mu + m) = 0 \tag{16}$$

By convention, we focus on the first term, and make it act on a four dimensional wave-function ψ, giving

$$(\gamma^\mu p_\mu - m)\psi = 0 \tag{17}$$

We did it! We just derived Dirac Equation from scratch. It's slightly difficult to see that this equation leads to a positive probability density, but we will prove this soon. The Dirac Equation is one of the most beautiful equation in physics, and wasn't as hard to get as you might have thought. Understanding some of its properties will not be easy but we can also do it from scratch.

There are different ways of expressing the Dirac equation. Looking back at equation (2), we also have

$$\boxed{(i\gamma^\mu \partial_\mu - m)\psi = 0}$$

which, as you might have guessed from the fact that it's boxed, is my favorite way of expressing Dirac Equation. Or, using the notational shortcut $\gamma^\mu A_\mu = \not{A}$, we get

$$(i\not{\partial} - m)\psi = 0 \tag{18}$$

2.2 A Deeper Look at the Gamma Matrices

We found that the Dirac matrices γ^μ satisfy the Clifford Algebra given by the above boxed equation. Letting $\mu = \nu = 0$, we have the following relation

$$(\gamma^0)^2 = \mathcal{I} \tag{19}$$

where $\mathcal{I}$ is the identity $n \times n$ matrix. We don't know n yet. Next, letting $\mu = \nu = i$, with $i = 1, 2, 3$, we get

$$(\gamma^i)^2 = -\mathcal{I} \tag{20}$$

Another useful identity can be found by letting $\mu = 0$ and $\nu = i$.

$$\left[\gamma^0, \gamma^i\right]_+ = 0 \tag{21}$$

Finally, the last one we will use comes from setting $\mu = i$ and $\nu = j$, giving

$$\left[\gamma^i, \gamma^j\right]_+ = 0, i \neq j \tag{22}$$

These four identities are extremely important and we will use them over and over again in this lecture. They will also help us understand some of the properties of these gamma matrices.

We can choose any of the gamma matrices to be diagonal, wolog let's choose γ^0. Remember that we know that we must have four Dirac matrices, but we don't know yet what type of matrices they are.

$$\gamma^0 = \begin{pmatrix} b_1 & 0 & ... & & 0 \\ 0 & b_2 & 0 & ... & 0 \\ & . & & & \\ & & . & & \\ & & & . & \\ & & & & b_n \end{pmatrix} \tag{23}$$

From equation (19), we get that $det(\gamma^0) = \pm 1 \rightarrow b_\alpha = \pm 1 \ \forall \alpha$. We can also get some information about the trace of γ^0 using (20), (21), and the cyclic property of the trace.

$$Tr(\gamma^0) = -Tr(\gamma^i\gamma^i\gamma^0) = -Tr(\gamma^i\gamma^0\gamma^i) = Tr(\gamma^i\gamma^i\gamma^0) = -Tr(\gamma^0) \tag{24}$$

It follows trivially that $Tr(\gamma^0) = 0$, which means that we must have as many $b_\alpha = 1$ as we have $b_\alpha = -1$, and therefore we have that γ^0, and so any γ^μ must be a $2k \times 2k$ matrix, with $k \in \mathcal{N}$. If $k = 1$, we can guess $\gamma^0 = \mathcal{I}$ and $\gamma^i = \sigma^i$. Some of the identities do match, but some do not, for example

$$\gamma^0\gamma^i + \gamma^i\gamma^0 = \mathcal{I}\sigma^i + \sigma^i\mathcal{I} = 2\sigma^i \neq 0 \tag{25}$$

So $k = 1$ doesn't work. Next try is $k = 2$. Lucky for us, we can find such matrices! One such representation is called the Dirac Pauli representation

$$\gamma^0 = \begin{pmatrix} \mathcal{I} & 0 \\ 0 & -\mathcal{I} \end{pmatrix} \tag{26}$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{27}$$

Note that these matrices are 4×4.

2.3 Pauli's Fundamental Theorem

Pauli's Fundamental Theorem states that if $[\gamma^\mu, \gamma^\nu]_+ = 2\eta^{\mu\nu}$ and $[\gamma'^\mu, \gamma'^\nu]_+ = 2\eta^{\mu\nu}$, then there exists a constant invertible matrix S such that

$$\gamma'^\mu = S\gamma^\mu S^{-1} \tag{28}$$

This isn't hard to believe from a linear algebra perspective since the γ^μ are Hermitian. There are multiple representations of the γ^μ matrices. We want to show here that changing representation should not affect the underlying physics, it should be like changing your frame of reference. Let's find a relation between the Dirac Equation for these two representations. Starting with the prime representation, we have

$$\begin{aligned}(\gamma'^\mu p_\mu - m)\psi' = 0 &\Leftrightarrow (S\gamma^\mu S^{-1} p_\mu - m)\psi' = 0 \\ &\Leftrightarrow S(\gamma^\mu p_\mu - m)S^{-1}\psi' = 0 \\ &\Leftrightarrow (\gamma^\mu p_\mu - m)S^{-1}\psi' = 0 \\ &\Leftrightarrow (\gamma^\mu p_\mu - m)\psi = 0\end{aligned} \tag{29}$$

Thus, the two equation are equivalent given the relation $\psi' = S\psi$, which obviously doesn't change the physics since it corresponds to a rotation.

2.4 The Hamiltonian of the Dirac Equation

We want to express the Dirac Equation in the familiar form $H\psi = E\psi = i\partial_0\psi$. But

$$\begin{aligned}(i\gamma^\mu\partial_\mu - m)\psi = (i\gamma^0\partial_0 + i\gamma^i\partial_i - m)\psi = 0 \\ \Leftrightarrow (i\gamma^i\partial_i - m)\psi = -i\gamma^0\partial_0\psi \\ \Leftrightarrow (i\gamma^0\gamma^i\partial_i - m\gamma^0)\psi = -i^0\partial_0\psi \\ \Leftrightarrow (-i\gamma^0\gamma^i\partial_i + m\gamma^0)\psi = i^0\partial_0\psi = E\psi\end{aligned} \tag{30}$$

Thus,

$$\begin{aligned}H = -i\gamma^0\gamma^i\partial_i + m\gamma^0 &= \gamma^0\vec{\gamma}\cdot(-i\boldsymbol{\nabla}) + m\gamma^0 \\ &= \gamma^0\vec{\gamma}\cdot\vec{p} + m\gamma^0\end{aligned} \tag{31}$$

Letting

$$\vec{\alpha} = \gamma^0\vec{\gamma} \tag{32}$$

and

$$\beta = \gamma^0 \tag{33}$$

so

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \tag{34}$$

and

$$\beta = \begin{pmatrix} \mathcal{I} & 0 \\ 0 & -\mathcal{I} \end{pmatrix} \tag{35}$$

Then, the Hamiltonian becomes

$$\boxed{H = \vec{\alpha}\cdot\vec{p} + m\beta}$$

The Hamiltonian is an observable and therefore has to be hermitian. With this fact, we can get even more informations on the gamma matrices characteristics. More specifically, $H = H^\dagger$ leads to

$$
\begin{aligned}
(\vec{\alpha}\cdot\vec{p})^\dagger = \vec{\alpha}\cdot\vec{p} \rightarrow \vec{\alpha}^\dagger = \vec{\alpha} \\
\beta^\dagger = \beta \rightarrow (\gamma^0)^\dagger = \gamma^0
\end{aligned} \tag{36}
$$

The first equation leads to

$$
\begin{aligned}
(\gamma^0\vec{\gamma})^\dagger &= \gamma^0\vec{\gamma} \\
\Leftrightarrow (\vec{\gamma})^\dagger\gamma^0 &= \gamma^0\vec{\gamma} \\
\Leftrightarrow (\vec{\gamma})^\dagger &= \gamma^0\vec{\gamma}\gamma^0 = -\vec{\gamma}
\end{aligned} \tag{37}
$$

where in the last step we've used (19) and (21). To summarize, we just derived the following

$$
\begin{aligned}
(\gamma^0)^\dagger &= \gamma^0 \\
(\gamma^i)^\dagger &= -\gamma^i
\end{aligned} \tag{38}
$$

3 Solutions of the Dirac Equation

3.1 Spin of the Dirac Particle

What type of particle is a Dirac particle? Note that rotations should be a symmetry of the system, and therefore the total angular momentum $\vec{J}$ must be a conserved quantity (i.e., constant in time). From

$$
\frac{d\vec{J}}{dt} = i\left[H, \vec{J}\right] \tag{39}
$$

we must have

$$
\left[H, \vec{J}\right] = 0
$$

From Quantum Mechanics, we normally have $\vec{J} = \vec{L} + \vec{S}$, but maybe in this case we just have $\vec{J} = \vec{L}$. Let's try this

$$
\begin{aligned}
H &= \vec{\alpha}\cdot\vec{p} + m\beta = \alpha_n p_n + m\beta \\
\vec{L} &= \vec{x}\times\vec{p} \Leftrightarrow L_i = \epsilon_{ijk}x_j p_k
\end{aligned} \tag{40}
$$

$$
\begin{aligned}
[L_i, H] = [\epsilon_{ijk}x_j p_k, \alpha_n p_n + m\beta] &= \epsilon_{ijk}\alpha_n[x_j, p_n]p_k \\
&= i\delta_{jn}\epsilon_{ijk}\alpha_n p_k \\
&= i\epsilon_{ijk}\alpha_j p_k \neq 0
\end{aligned} \tag{41}
$$

Therefore, we must have $\vec{J} = \vec{L} + \vec{S}$, and

$$
[S_i, H] = -i\epsilon_{ijk}\alpha_j p_k \tag{42}
$$

What S_i would satisfy that? We know from Quantum Mechanics that $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, which is pretty close, so we probably want to play with the α_n in the Hamiltonian.

Suppose S_i only interacts with $\vec{\alpha}$, then

$$
\begin{aligned}
[S_i, H] = [S_i, \alpha_k p_k + m\beta] &= [S_i, \alpha_k]p_k \\
&= -i\epsilon_{ijk}\alpha_j p_k
\end{aligned} \tag{43}
$$

Thus, if we can find an S_i such that $[S_i, \alpha_k] = -i\epsilon_{ijk}\alpha_j$ we are done. This is very close to $[\sigma_i, \sigma_j]$. To summarize, we want to get

$$
[S_i, \alpha_j] = -i\epsilon_{ikj}\alpha_k = i\epsilon_{ijk}\alpha_k \tag{44}
$$

Considering the σ dependence of α, a good guess is $S_i = \alpha_i$. Then,

$$[\alpha_i, \alpha_j] = \begin{pmatrix} [\sigma_i, \sigma_j] & 0 \\ 0 & [\sigma_i, \sigma_j] \end{pmatrix} = 2i\epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \neq i\epsilon_{ijk}\alpha_k \tag{45}$$

So this doesn't work. Another good guess would be

$$S_i = \tilde{\alpha}_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \tag{46}$$

Then

$$[\tilde{\alpha}_i, \alpha_j] = \begin{pmatrix} 0 & [\sigma_i, \sigma_j] \\ [\sigma_i, \sigma_j] & 0 \end{pmatrix} = 2i\epsilon_{ijk} \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} = 2i\epsilon_{ijk}\alpha_k \tag{47}$$

Therefore, letting $S_i = \frac{1}{2}\tilde{\alpha}_i$ would give the desired $[S_i, \alpha_j] = i\epsilon_{ijk}\alpha_k$ which would then give $\left[H, \vec{J}\right] = 0$. So we did find a spin component equal to

$$\vec{S} = \frac{1}{2}\vec{\tilde{\alpha}} \tag{48}$$

Note that

$$S_3 = \frac{1}{2}\tilde{\alpha}_3 = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix} \tag{49}$$

has eigenvalues $\pm\frac{1}{2}$, doubly degenerate. It is a pretty common fact in physics that the Dirac Equation describes electron and positron, which are fermions of spin $\frac{1}{2}$. Thus, having doubly degenerate eigenvalues of $\pm\frac{1}{2}$ should have been expected.

3.2 Solutions

We just found out that the Dirac Equation describes spin $\frac{1}{2}$ particles. Let's find how to describe these particles using their wave function solutions. From Dirac Equation and the fact that the gamma matrices have to be 4×4, it means that

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \tag{50}$$

The plane wave solution is

$$\psi_\alpha(x) = e^{-ip\cdot x}u_\alpha(p) \tag{51}$$

with $\alpha = 1, 2, 3, 4$. Using $(\not{p} - m)\psi = 0$, we get $(\not{p} - m)u(p) = 0$. Let's choose a frame such that the motion is along the $z-$ axis, i.e.,

$$p^\mu = (p^0, 0, 0, p^3) \tag{52}$$

Thus, $(\gamma^\mu p_\mu - m)u = 0 \Leftrightarrow (\gamma^0 p_0 + \gamma^3 p_3 - m)u = 0$. For non trivial solutions, we must have $det(\gamma^0 p_0 + \gamma^3 p_3 - m) = 0$. This is equivalent to

$$det\begin{pmatrix} (p_0 - m)\mathcal{I} & \sigma_3 p_3 \\ -\sigma_3 p_3 & -(p_0 + m)\mathcal{I} \end{pmatrix} = 0 \tag{53}$$

Since $\mathcal{I}$ and σ_3 commute, the above is equivalent to

$$\begin{aligned} det(-(p_0 - m)(p_0 + m) + \sigma_3^2 p_3^2) = 0 &\Leftrightarrow det(-(p_0^2 - p_3^2 - m^2)\mathcal{I}) = 0 \\ &\Leftrightarrow (p_0^2 - p_3^2 - m^2)^2 = 0\star \\ &\Rightarrow p_0 = \pm\sqrt{p_3^2 + m^2} = \pm E = E_\pm \end{aligned} \tag{54}$$

From the $\star$ part of the derivation, we can see the doubly degenerate behavior of the solutions. That should ring a bell since we just saw this for the spin. To underline this property, we want to split $u(p)$ into two subvectors (since we know we should describe two particles: the electron and positron, each of spin $\frac{1}{2}$).

$$u(p) = \begin{pmatrix} u_1(p) \\ u_2(p) \\ u_3(p) \\ u_4(p) \end{pmatrix} = \begin{pmatrix} \tilde{u}(p) \\ \tilde{v}(p) \end{pmatrix} \tag{55}$$

with

$$\tilde{u}(p) = \begin{pmatrix} u_1(p) \\ u_2(p) \end{pmatrix} \tag{56}$$

and

$$\tilde{v}(p) = \begin{pmatrix} u_3(p) \\ u_4(p) \end{pmatrix} \tag{57}$$

For E_+, $(\gamma^\mu p_\mu - m)u(p) = 0$ becomes

$$\begin{pmatrix} (E_+ - m)\mathcal{I} & \sigma_3 p_3 \\ -\sigma_3 p_3 & -(E_+ + m)\mathcal{I} \end{pmatrix} \begin{pmatrix} \tilde{u}(p) \\ \tilde{v}(p) \end{pmatrix} \tag{58}$$

which leads to the following system of equations

$$\begin{aligned} (E_+ - m)\tilde{u} + \sigma_3 p_3 \tilde{v} = 0 \\ -\sigma_3 p_3 \tilde{u} - (E_+ + m)\tilde{v} = 0 \end{aligned} \tag{59}$$

Looking at the second equation, we see that if $p_3 = 0$, then since $E_+ = m$, we must have $\tilde{v}(p) = 0$, which makes $\tilde{u}$ arbitrary. To make our lives easier, we pick

$$\tilde{u}(p) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{60}$$

and

$$\tilde{u}(p) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{61}$$

as our two independent solutions for E_+. Solving for $\tilde{v}(p)$ using the above system of equations gives

$$\tilde{v}(p) = -\frac{\sigma_3 p_3}{E_+ + m}\tilde{u}(p) \tag{62}$$

The first choice of $\tilde{u}(p)$ leads to

$$\tilde{v}(p) = \begin{pmatrix} -\frac{p_3}{E_+ + m} \\ 0 \end{pmatrix} \tag{63}$$

and the second choice of $\tilde{u}(p)$ leads to

$$\tilde{v}(p) = \begin{pmatrix} 0 \\ \frac{p_3}{E_+ + m} \end{pmatrix} \tag{64}$$

Doing the same for the two negative energy solutions, E_-, leads to $\tilde{v}(p)$ being the two arbitrary independent solutions

$$\tilde{v}(p) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{65}$$

and

$$\tilde{v}(p) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{66}$$

The first choice of $\tilde{v}(p)$ leads to

$$\tilde{u}(p) = \begin{pmatrix} -\frac{p_3}{E_- - m} \\ 0 \end{pmatrix} \tag{67}$$

and the second choice of $\tilde{v}(p)$ leads to

$$\tilde{u}(p) = \begin{pmatrix} 0 \\ \frac{p_3}{E_- - m} \end{pmatrix} \tag{68}$$

Finally, we have our four solutions!

4 Continuity Equation

The big problem with the Klein-Gordon Equation was the possible negative probability density. We got rid of the problem by working in analogy to the Schroedinger Equation and requiring first order in time. Hopefully this works out, otherwise everything we just accomplished is a bit useless... Let's see if it does end up working out (hopefully). From $H\psi = E\psi$, which we found before to be equivalent to

$$(-i\vec{\alpha} \cdot \boldsymbol{\nabla} + \beta m)\psi = i\frac{\partial \psi}{\partial t} \tag{69}$$

we need to get something of the form

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \vec{j} = 0 \tag{70}$$

which if you took some electromagnetism should look familiar. It is called the continuity equation. If we can find a way to remove the βm term, we would be good. Recalling that we must have $H = H^\dagger$ and so $\vec{\alpha} = \vec{\alpha}^\dagger$ and $\beta = \beta^\dagger$, we get

$$\psi^\dagger(i\overleftarrow{\boldsymbol{\nabla}} \cdot \vec{\alpha} + \beta m) = -i\frac{\partial \psi^\dagger}{\partial t} \tag{71}$$

where the left arrow on top of the gradient means that the gradient is acting on the left. To get rid of βm, an obvious way is to get $\beta m \psi^\dagger \psi$ in both equations, and the subtracting them to kill it. You should try it out, remember that order matters.

$$\begin{aligned} -i\psi^\dagger \vec{\alpha} \cdot \overrightarrow{\boldsymbol{\nabla}} \psi + \beta m \psi^\dagger \psi &= i\psi^\dagger \frac{\partial \psi}{\partial t} \\ i\psi^\dagger \overleftarrow{\boldsymbol{\nabla}} \cdot \vec{\alpha}\psi + \beta m \psi^\dagger \psi &= -i\left(\frac{\partial \psi^\dagger}{\partial t}\right)\psi \end{aligned} \tag{72}$$

Subtracting equation 1 from equation 2 leads to

$$\begin{aligned} i(\psi^\dagger \vec{\alpha} \cdot \overrightarrow{\boldsymbol{\nabla}} \psi + \psi^\dagger \overleftarrow{\boldsymbol{\nabla}} \cdot \vec{\alpha}\psi) &= -i\left(\psi^\dagger \frac{\partial \psi}{\partial t} + \left(\frac{\partial \psi^\dagger}{\partial t}\right)\psi\right) \\ \Leftrightarrow \boldsymbol{\nabla} \cdot (\psi^\dagger \vec{\alpha} \psi) &= -\frac{\partial}{\partial t}(\psi^\dagger \psi) \\ \Leftrightarrow \frac{\partial}{\partial t}(\psi^\dagger \psi) + \boldsymbol{\nabla} \cdot (\psi^\dagger \vec{\alpha}\psi) &= 0 \end{aligned} \tag{73}$$

where we have used the identity $\boldsymbol{\nabla} \cdot (f\vec{A}) = (\boldsymbol{\nabla} f) \cdot \vec{A} + f(\boldsymbol{\nabla} \cdot \vec{A})$. It follows that letting $\rho = \psi^\dagger \psi = |\psi|^2 \geq 0$ and $\vec{j} = \psi^\dagger \vec{\alpha} \psi$, we are done. Good! We do indeed have the required positive probability density $\rho \geq 0$, so our previous work is relevant to the real world.

A more concise way to write the above continuity equation can be found by defining the following four vector current

$$j^\mu = (\rho, \vec{j}) = (\psi^\dagger \psi, \psi^\dagger \vec{\alpha} \psi) \tag{74}$$

then the continuity equation is equivalent to

$$\partial_\mu j^\mu = 0 \tag{75}$$

5 Conclusion

> *The mathematician plays a game in which he himself invents the rules while the physicist plays a game in which the rules are provided by Nature, but as time goes on it becomes increasingly evident that the rules which the mathematician finds interesting are the same as those which Nature has chosen.*
>
> Paul Dirac (1902-1984)

In this lecture, we have looked at one of the most beautiful equation in physics: the Dirac Equation. We first figured out why we need it, then from these arguments we derived it from scratch. We then discovered some of its amazing properties like how the Dirac gamma matrices must behave, what its Hamiltonian must be like, the necessity of the particles it describes having spin. Finally, we put everything together to understand what the plane wave solutions for these particles describes and if the probability density is strictly positive, which we need in order to not have negative probabilities popping up.

References

[1] Ashok Das, *Lectures on Quantum Field Theory*, World Scientific, 2008, First Edition, pp. 19-47.