Relativistic Quantum Mechanics

During the early part of the last century Schrödinger's equation was used to explain and describe all phenomena in atomic physics. However, after the development of the theory of special relativity by Einstein in 1905, there was a need to unify quantum mechanics and special relativity in to a single Relativistic Quantum Theory. Despite the success of Schrödinger's equation in describing quite accurately the Hydrogen spectrum and giving correct predictions for a large amount of spectral data, this equation is not invariant under Lorenz transformations. In other words Schrödinger's equation is not relativistic and is only an approximation valid at the non-relativistic limit when the velocities of the particles involved are much smaller than the speed of light. The reason that Schrödinger's equation is not relativistic is simple. Consider the Schrödinger equation:

$$[-(\frac{\hbar^2}{2m})\nabla^2+V(x,t)]\Psi=i\hbar\frac{\partial\Psi}{\partial t}$$

and consider also a Lorenz boost from the Lab frame (un-primed) to a frame with $\beta = v/c$ and $\gamma=1/\sqrt{(1-\beta^2)}$ moving at the positive x-direction (primed), where v is the velocity of the moving frame and c is the speed of light in vacuum:

$$x^{0\prime}=\gamma(x^0-\vec{\beta}\cdot\vec{x})$$

$$x'_{par}=\gamma(x_{par}-|\vec{\beta}|\cdot x^0)$$

$$x'_{per}=x_{per}$$

The space derivative appears at the second order in Schrödinger's equation while the time derivative at first order. However, since space and time are treated equivalently in special relativity, the time and space variables appear at first order (linear) in the Lorenz transformations. This difference between the Schrödinger's equation and the Lorenz transformation renters Schrödinger's equation non-invariant under Lorenz transformations that is non-relativistic.

Another way of seeing this is to replace the operators in Schrödinger's equation with the corresponding physical observables that is:

$$p^2/2m+V=E$$

Clearly this equation is not relativistic since we know that the correct relativistic relationship that connects energy mass and momentum is:

$$E^2=p^2+m^2$$

The Klein-Gordon Equation.

The obvious way of creating a manifestly invariant wave equation is to start from the well known relativistic energy equation and replace all quantities involved with the corresponding Quantum Mechanical operators. After doing this and you get:

$$[(\frac{\partial}{\partial x^0})^2 - \nabla^2 + \frac{m^2 c^2}{\hbar^2}]\Phi = 0$$

This is the Klein-Gordon equation which describes spin zero massive particles. As we have seen before if these particles have also positive intrinsic parity they are called *scalars,* and if they have negative intrinsic parity they are called *pseudoscalars*. Often the particle properties are described by the symbol J^P where J is the spin of the particle and P is the parity. In this notation scalars are 0^+ and pseudoscalars are 0^-.

It is convenient to work in the system where $\hbar = c = 1$ and we will assume this for the rest of these course unless otherwise stated. In this system the Klein-Gordon equation becomes:

$$[(\partial/\partial x^0)^2 - \nabla^2 + m^2]\Phi = 0$$

which in covariant notation can be written as:

$$[\partial^\mu \partial_\mu + m^2]\Phi = 0$$

This equation is manifestly covariant since the derivative product is covariant as a dot product of 4-vectors and the mass square term is also covariant because it is the result of the product of two momentum 4-vectors.

Perhaps it is a good idea to discuss what exactly do we mean by saying that the equation is covariant since it is the first time we confront this issue in this course:

Consider a spin zero particle moving in free space. An observer in reference frame **O** makes measurements and finds that the particle is described by the wave function $\Phi(x)$ which satisfies the Klein Gordon equation expressed in his frame as:

$$[\frac{\partial}{\partial x_\mu}\frac{\partial}{x^\mu} + m^2]\Phi(x) = 0 \qquad (1)$$

Consider a second observer is in a frame **O'** which is related to **O** by the Lorentz transformation:

$$x^{\mu}{}' = \Lambda^{\mu}{}_{\nu} x^{\nu} \quad \text{or} \quad x^{\mu} = (\Lambda^{-1})^{\mu}{}_{\nu} x^{\nu}{}' \qquad (2)$$

who also observes the same particle.

The two observers see each other moving and therefore they know that their coordinate measurements are related via the above Lorentz Transformation. The observer in **O** can use the Lorentz transformation in (2) and translate his equation to the **O'** frame. Hence, he can predict what the other observer in ***O'*** should be 'seeing'. The observer in ***O'*** also makes measurements and concludes that the particle is described by $\Phi'(x')$ which satisfies a given equation in his frame expressed in terms of the $x^{\mu}{}'$ coordinates. Obviously for the entire picture of events to be objective and independent upon the choice of the observer, it must be that the equation in the ***O'*** frame must be identical to the one that the observer in ***O*** predicts by translating his own equation in the ***O'*** frame. If that happens to be the case then we say that an equation is invariant under Lorentz transformations. Alternatively stated: The relativity principle requires that the observed physical phenomena are the same regardless of the frame of reference from which we make our observations. Hence, it requires that the two equations in the two different frames *must* be the same. Lets see now under what conditions this is true:

The observer in **O** translates the equation (1) in the primed frame as follows:

$$(1) \Rightarrow [(\Lambda^{-1})^{\mu}{}_{\alpha} \frac{\partial}{\partial x_{\alpha}{}'} (\Lambda^{-1})_{\mu}{}^{\beta} \frac{\partial}{\partial x^{\beta}{}'} + m^2] \Phi(\Lambda^{-1} x') = 0 \Rightarrow$$

$$[\frac{\partial}{\partial x_{\alpha}{}'} \delta_{\alpha}{}^{\beta} \frac{\partial}{\partial x^{\beta}{}'} + m^2] \Phi(\Lambda^{-1} x') = 0 \Rightarrow$$

$$[\frac{\partial}{\partial x_{\alpha}{}'} \frac{\partial}{\partial x^{\alpha}{}'} + m^2] \Phi(\Lambda^{-1} x') = 0 \qquad (3)$$

According to the relativity principle, equation (3) must be identical to the one 'seen' by the observer in **O'** and given below by Eq. (4):

$$[\frac{\partial}{\partial x_{\alpha}{}'} \frac{\partial}{\partial x^{\alpha}{}'} + m^2] \Phi'(x') = 0 \qquad (4)$$

This is true if the wave function is invariant under Lorentz (Lorentz-scalar):

$$\Phi(x) = \Phi(\Lambda^{-1} x') = \Phi'(x')$$

Using the same arguments for the Parity transformations where:

$$t \rightarrow t' = t \quad ; \quad \vec{x} \rightarrow \vec{x}' = -\vec{x}$$

one finds that the Klein Gordon equation is invariant under Parity if:

$$\Phi'(\vec{x}', t) = e^{i\phi} \Phi(-\vec{x}', t) = e^{i\phi} \Phi(\vec{x}, t)$$

So the wave functions of particles that are scalars satisfy the Klein Gordon Equation.

Negative Energy and Negative Probability solutions of the Klein Gordon Equation

As we have shown so far we have succeeded in creating a relativistic invariant equation. However, early enough it was noticed that there were two problems with this equation:

- It has both positive but *also negative solutions.*
- The negative solutions are associated with *negative probability* and this second problem was beyond anything that could be accepted as a reasonable option at the time.

To show this we write the Klein-Gordon Equation as :

$$[\partial^0 \partial_0 - \nabla^2 + m^2] \Phi = 0$$

This equation can be solved by substituting for $\Phi = A \exp(iEt - i\vec{p}\vec{x})$ which gives us:

$$E^2 = p^2 + m^2 \Rightarrow E = \pm\sqrt{p^2 + m^2}$$

This demonstrates that there are both positive and negative solutions. This of course means that any particle can decay to an infinite number of negative states and as it decays and loses energy its energy becomes more and more negative in a process that the equation above implies that its momentum grows to infinity.

To show that the negative energy solutions are associated with negative probability we first have to derive the continuity equation which will give us the probability density ρ and the current density $\vec{J}$. To show this start with the complex conjugate of the Klein Gordon equation multiplied by $-i\Phi$ from the left side

$$-i\Phi[\partial^0\partial_0-\nabla^2+m^2]\Phi^*=0 .$$

Consider also the symmetric equation where

$$i\Phi^*[\partial^0\partial_0-\nabla^2+m^2]\Phi=0 .$$

By adding the two equations the mass term drops out and we get:

$$i[\Phi^*\partial^0\partial_0\Phi-\Phi\partial^0\partial_0\Phi^*]-i[\Phi^*\nabla^2\Phi-\Phi\nabla^2\Phi^*]=0$$

which of course can be written as:

$$i\partial^0(\Phi^*\partial_0\Phi-\Phi\partial_0\Phi^*)-i\vec{\nabla}(\Phi^*\vec{\nabla}\Phi-\Phi\vec{\nabla}\Phi^*)=0$$

Hence, one now can identify:

$$\rho=i(\Phi^*\partial_0\Phi-\Phi\partial_0\Phi^*) \qquad (1)$$

and

$$\vec{J}=i(\Phi^*\vec{\nabla}\Phi-\Psi\vec{\nabla}\Phi^*) \qquad (2)$$

which satisfy the continuity equation:

$$\partial^0(\rho)-\vec{\nabla}\vec{J}=0 \qquad (3)$$

Note that (1) and (2) can be combined in one equation that describes and 4-vector current

$$J^\mu(x) \;=\; i\Phi^*\partial^\mu\Phi-\Phi\partial^\mu\Phi^*$$

and using (3) we have:

$$\partial_\mu J^\mu(x) \;=\; 0$$

which is the relativistic version of the continuity equation.

Now we have all the information we need to calculate the probability density. We use

$$\Phi = N \exp(i(Et - \vec{p}\vec{x}))$$

which we have already shown that it satisfies the Klein-Gordon equation and compute the probability density:

$$\rho = 2|N|^2 E \qquad (4)$$

Hence, the negative energy solutions are associated with negative probability. This is clearly nonsenses. At the time of its invention these two problems were considered fatal and the equation was abandoned.

Note: The probability in (a) is not relativistic invariant and this is not a problem. Why should it be anyway ? It should not be because it is the integral :

$$\int \rho(x) dV \;=\; \int \rho'(x') dV'$$

that should be Lorentz invariant. As seen in (4) the probability density is proportional to the energy which transforms under Lorentz giving a factor of γ whilst the volume element gives a factor of $1/\gamma$ as expected from Lorentz contraction. The two factors cancel each other and the integral remains Lorentz Invariant.

The Dirac Equation

Dirac in his effort to discover an equation that was free of the problems seen with the Klein Gordon equation tried an equation that was linear in the energy, mass and momentum. He tried the Hamiltonian:

$$H \;=\; \vec{\alpha}\cdot\vec{p} \;+\; \beta m \;=\; \alpha_i p_i \;+\; \beta m \quad \text{where} \quad H\Psi = E\Psi \qquad (1)$$

with α_i, β being constant objects and E the particle energy. Whatever these constants might be the the equation must satisfy the relativistic energy equation:

$$E^2 = p^2 + m^2 \qquad (2)$$

Hence,

$$(1) \Rightarrow H^2\Psi = E^2\Psi \Rightarrow (\alpha_i p_i + \beta m)(\alpha_j p_j + \beta m)\Psi = (p^2+m^2)\Psi \Rightarrow$$

$$(\alpha_i p_i \alpha_j p_j + \beta m \alpha_j p_j + \alpha_i p_i \beta m + \beta^2 m^2)\Psi = (p^2+m^2)\Psi \Rightarrow$$

$$(\frac{1}{2}(\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j + m(\beta\alpha_j + \alpha_j\beta)p_j + \beta^2 m^2)\Psi = (p^2+m^2)\Psi \quad (3)$$

If (3) is to be satisfied then we must have that:

$$\frac{1}{2}(\alpha_i\alpha_j + \alpha_j\alpha_i) = \delta_{ij} \Rightarrow \{\alpha_i, \alpha_j\} = 2\delta_{ij} \quad (4)$$

$$(\beta\alpha_j + \alpha_j\beta) = 0 \Rightarrow \{\beta, \alpha_j\} = 0 \quad (5)$$

$\beta^2 = 1$ and using (4) $\alpha_i^2 = 1$ (6)

The first observation is that with the constants α_i, cannot be numbers (numbers commute do not anti-commute) so they must be matrices. **It is left as homework to show that these are matrices of even dimension, Hermitean and traceless whose <u>dimension is 4 or greater.</u>**

Several representations of the α_i, β matrices exist which satisfy (4), (5), (6).

The standard **Pauli-Dirac** representation is given by:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \text{ and } \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

The **Weyl** representation is given by:

$$\vec{\alpha} = \begin{pmatrix} -\vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \text{ and } \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

Where the $(\mathbf{1}\ ;\ \vec{\sigma})$ are the 2 x 2 unity and the three Pauli matrices respectively. Hence, the $\alpha_i\,,\beta$ are 4 x 4 matrices. Recall the the Pauli matrices satisfy:

$$\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k \Rightarrow \{\sigma^i,\sigma^j\} = 2\delta^{ij}\ ;\ [\sigma^i,\sigma^j] = 2i\epsilon^{ijk}\sigma^k$$

From the $\alpha_i\,,\beta$ matrices we can define the matrices:

$$\gamma^i = \beta\alpha^i \quad i = 1,2,3 \quad ;$$

$\gamma^0 = \beta$ and

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$$

The gamma matrices can then be written in a covariant form as:

$$\gamma^\mu = (\gamma^0;\vec{\gamma})$$

It is important to note here that although γ^μ looks like a 4-vector, it does not transform as a 4-vector as will will see later on.

In the Pauli-Dirac representation we have that:

$$\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}; \quad \gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}; \quad \gamma_5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

In the Weyl representation we have that:

$$\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}; \ \gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}; \ \gamma_5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$$

The Dirac Equation which in terms of the α_i, β matrices is given by:

$$(\vec{\alpha}\cdot\vec{\boldsymbol{p}} + \beta \boldsymbol{m})\Psi = \boldsymbol{E}\Psi$$

can now be written in a more compact form using the gamma matrices:

$$(\boldsymbol{i}\gamma^{\mu}\partial_{\mu} - \boldsymbol{m})\Psi = \mathbf{0} \qquad (7)$$

For the reader who is not familiar with this notation it is perhaps useful to emphasize that all these matrices are 4x4 objects which for example in the Dirac-Pauli (standard) representation can be explicitly written as:

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & +i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$\gamma_5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Similarly the state Ψ , is a 4 dimensional column object: $\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$

One can also define the row object:

$$\bar{\Psi} = \Psi^{+}\gamma^{0} = \left(\Psi_1^*, \Psi_2^*, \Psi_3^*, \Psi_4^*\right)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \left(\Psi_1^*, \Psi_2^*, -\Psi_3^*, -\Psi_4^*\right)$$

Next lets try to construct a probability density from the Dirac equation so that we can investigate if it suffers from the same negative probability problems as the Klein Gordon equation:

By multiplying (7) by $\bar{\Psi}$ we get:

$$\bar{\Psi}(i\gamma^{\mu}\partial_{\mu} - m)\Psi = 0 \qquad (8)$$

Also from (7) we have that $[(i\gamma^{\mu}\partial_{\mu} - m)\Psi]^{+} = 0 \Rightarrow$

$$(-i)(\partial_{\mu}\Psi^{+})(\gamma^{\mu})^{+} - m\Psi^{+} = 0 \quad (9)$$

However the gamma matrices have the property that:

$$(\gamma^{\mu})^{+} = \gamma^{0}\gamma^{\mu}\gamma^{0} \qquad (10)$$

By substituting (10) in to (9) and multiplying from the right with γ^{0} we get that:

$$i(\partial_{\mu}\bar{\Psi})\gamma^{\mu} + m\bar{\Psi} = 0 \Rightarrow$$

We can now multiply by Ψ from the right side to obtain:

$$i(\partial_{\mu}\bar{\Psi})\gamma^{\mu}\Psi + m\bar{\Psi}\Psi = 0 \qquad (11)$$

by adding (8) and (11) we get:

$$i\bar{\Psi}\gamma^{\mu}\partial_{\mu}\Psi + i(\partial_{\mu}\bar{\Psi})\gamma^{\mu}\Psi = 0 \Rightarrow$$

$$\partial_\mu (\boldsymbol{i} \bar{\Psi} \gamma^\mu \Psi) = \mathbf{0}$$

Hence, there is a conserved current:

$$\boldsymbol{J}^\mu = \bar{\Psi} \gamma^\mu \Psi$$

which satisfies the continuity equation:

$$\partial_\mu \boldsymbol{J}^\mu = \mathbf{0}$$

It will be shown later on that this is indeed a 4-vector current i.e. It transforms like a 4-vector under Lorentz and Parity. The components of the 4-vector current are given by:

$$\boldsymbol{J}^\mu = (\bar{\Psi} \gamma^{\mathbf{0}} \Psi \;\; ; \;\; \bar{\Psi} \gamma^{\boldsymbol{i}} \Psi) = (\rho ; \vec{\boldsymbol{J}})$$

The probability density is then given by:

$$\rho = \bar{\Psi} \gamma^{\mathbf{0}} \Psi = \Psi^+ \gamma^{\mathbf{0}} \gamma^{\mathbf{0}} \Psi = \Psi^+ \Psi = |\Psi_{\mathbf{1}}|^{\mathbf{2}} + |\Psi_{\mathbf{2}}|^{\mathbf{2}} + |\Psi_{\mathbf{3}}|^{\mathbf{2}} + |\Psi_{\mathbf{4}}|^{\mathbf{2}} \geqslant \mathbf{0}$$

Hence, the Dirac equation leads to positive probabilities. So this solves one of the problems that the Klein Gordon equation had:

The Dirac equation predicts positive probabilities for both positive and negative energy states.

However, as we will see in the next lecture the negative energy problem remains and another solution (interpretation of it) had to be found.