Relativistic Quantum Waves (Klein-Gordon Equation
And expanding to the Dirac equation

Marcus Allen Denslow

2025-11-14

Contents

Chapter 1

Deriving the KG Equation

1.1 The Strategy: Combining Relativity and Quantum Mechanics

The goal is to derive a relativistic wave equation that respects both special relativity and quantum mechanics. The strategy is to:

1. Start with Einstein's relativistic energy-momentum relation
2. Replace classical quantities (energy, momentum) with quantum operators
3. Apply these operators to a wavefunction ψ

This "double derivative" approach (squaring the energy) will give us a second-order differential equation.

Definition 1.1.1: Relativity: the mass shell (Einstein's energy-momentum relation)

Einstein showed that energy and momentum are related by:
$E^2 = (pc)^2 + (mc^2)^2$

Rearranging in four-vector notation: $p \cdot p = (mc)^2 \rightarrow (mc)^2 = \left(\frac{E}{c}\right)^2 - p_x^2 - p_y^2 - p_z^2$

This is called the **mass shell** — it's a constraint that relates energy and momentum for a particle with mass m.

Definition 1.1.2: Quantum: energy and momentum operators

In quantum mechanics, we promote classical observables to operators that act on wavefunctions. From the Schrödinger equation and de Broglie relations:

- Energy operator: $\hat{E} = i\hbar\frac{\partial}{\partial t}$
- Momentum operator: $\hat{p} = -i\hbar\nabla$

To use these in Einstein's energy-momentum relation, we need to square them. When we square an operator, we apply it twice:

Energy term:

$$\left(\frac{E}{c}\right)^2 \to \left(\frac{\hat{E}}{c}\right)^2 = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}\right)^2 = -\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2}.$$

Momentum terms:

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x} \to -p_x^2 \to \hbar^2\frac{\partial^2}{\partial x^2}.$$

$$\text{Likewise: } -p_y^2 \to \hbar^2\frac{\partial^2}{\partial y^2} \text{ and } -p_z^2 \to \hbar^2\frac{\partial^2}{\partial z^2}.$$

1.2 Substituting Operators into Einstein's Relation

Now we substitute the quantum operators into the relativistic mass shell equation. Remember, these operators will act on a wavefunction ψ:

$$(mc)^2 = \left(\frac{E}{c}\right)^2 - p_x^2 - p_y^2 - p_z^2$$

$$(mc)^2 = -\frac{\hbar^2}{c^2}\frac{\partial^2}{\partial t^2} + \hbar^2\frac{\partial^2}{\partial x^2} + \hbar^2\frac{\partial^2}{\partial y^2} + \hbar^2\frac{\partial^2}{\partial z^2}.$$

Dividing through by $\hbar^2$ and rearranging:

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + \left(\frac{mc}{\hbar}\right)^2 = 0 \tag{1.1}$$

1.3 The Klein-Gordon Equation!

We've done it! This is the **Klein-Gordon equation** — the first successful attempt at a relativistic quantum wave equation.

Operating on a wavefunction ψ:

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + \left(\frac{mc}{\hbar}\right)^2\right]\psi = 0 \tag{1.2}$$

Key features:

- Second-order in both time and space (treats them equally — relativistic!)
- Reduces to Schrödinger equation in the non-relativistic limit
- Contains the mass m explicitly

1.4 Compact Notation: Laplacian

We can write the spatial derivatives more compactly using the Laplacian operator:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Note the sign: in our equation we have $-\nabla^2$ because of the minus signs in the original expression.

Substituting this into our equation:

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \left(\frac{mc}{\hbar}\right)^2\right]\psi = 0 \tag{1.3}$$

This is cleaner and makes the equation look more symmetric.

1.5 Even More Compact: d'Alembertian Notation

We can make this even more compact by defining two convenient symbols:

The d'Alembertian operator (wave operator):

$$\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$$

This is the relativistic generalization of the Laplacian. It's the natural wave operator in spacetime.

The mass parameter:

$$\mu \equiv \frac{mc}{\hbar}$$

This has units of inverse length (like a wavenumber). It represents the "Compton wavenumber" of the particle.

Note:-

In natural units where $c = \hbar = 1$, we'd just write $\mu = m$. But we keep factors explicit for clarity.

With these definitions, the Klein-Gordon equation becomes beautifully simple:

$$\left[\Box + \mu^2\right]\psi = 0 \tag{1.4}$$

This compact form makes it easy to see: it's a wave equation (the $\Box$) with a mass term (μ^2).

Chapter 2

Four-momentum Eigenstates

Now that we have the Klein-Gordon equation, what are its solutions? The simplest solutions are plane waves — states with definite energy and momentum.

2.1 Klein-Gordon Plane Wave Solutions

Just like in non-relativistic quantum mechanics, we look for plane wave solutions. These represent particles with definite four-momentum.

Definition 2.1.1: Klein-Gordon Plane Wave function

The general form using four-vector notation:

$$\psi = A \exp\left(-\frac{i}{\hbar} p \cdot x\right)$$

where:

$$\begin{aligned} p &= \left[E/c, \vec{p}\right] && \text{(four-momentum)} \\ x &= \left[ct, \vec{x}\right] && \text{(four-position)} \\ A &\in \mathbb{C} && \text{(complex amplitude)} \\ p^0 = \frac{E}{c} &= \pm\sqrt{|\vec{p}|^2 + m^2c^2} && \text{(energy, with } \pm \text{ solutions!)} \end{aligned}$$

Expanding the four-vector dot product $p \cdot x = Et - \vec{p} \cdot \vec{x}$:

$$\psi = A \exp\left[\frac{i}{\hbar}\left(\vec{p} \cdot \vec{x} - Et\right)\right]$$

$$\psi = A \exp\left[\frac{i}{\hbar}\left(\vec{p} \cdot \vec{x} \pm c\sqrt{|\vec{p}|^2 + m^2c^2}\, t\right)\right]$$

This looks like $e^{i(\vec{k}\cdot\vec{x} - \omega t)}$ — a traveling wave!

2.2 Proof that plane waves satisfy Klein-Gordon

Let's verify that $\psi = A \exp[-ip \cdot x/\hbar]$ is indeed a solution.

Rewrite Klein-Gordon as:

$$\left[\Box + \mu^2\right]\psi = 0 \quad \Rightarrow \quad \Box\psi = -\mu^2\psi$$

The question is: does the d'Alembertian acting on our plane wave give us $-\mu^2\psi$?

Apply the d'Alembertian to the plane wave:

$$\Box\left[\exp\left[\frac{i}{\hbar}\left(\vec{p}\cdot\vec{x}-Et\right)\right]\right] \tag{2.1}$$

Remember: $\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$

Taking derivatives:

- Time derivative: $\frac{\partial}{\partial t}\exp[i(\vec{p}\cdot\vec{x}-Et)/\hbar] = -\frac{iE}{\hbar}\psi \Rightarrow \frac{\partial^2}{\partial t^2} = -\frac{E^2}{\hbar^2}\psi$
- Spatial derivatives: $\nabla^2\exp[i(\vec{p}\cdot\vec{x}-Et)/\hbar] = -\frac{|\vec{p}|^2}{\hbar^2}\psi$

Therefore:

$$\Box\psi = \left[-\frac{E^2}{c^2\hbar^2} + \frac{|\vec{p}|^2}{\hbar^2}\right]\psi = -\frac{1}{\hbar^2}\left[\frac{E^2}{c^2} - |\vec{p}|^2\right]\psi \tag{2.2}$$

Using the mass shell relation $E^2/c^2 - |\vec{p}|^2 = m^2c^2$:

$$\Box\psi = -\frac{m^2c^2}{\hbar^2}\psi = -\mu^2\psi \quad \checkmark \tag{2.3}$$

The plane wave satisfies Klein-Gordon! **And notice**: the mass shell relation is exactly what makes this work.

Chapter 3

Superposition

3.1 Linearity: Building Complex Solutions from Simple Ones

The Klein-Gordon equation is **linear** — this is a crucial property! It means if ψ_1 and ψ_2 are solutions, then any linear combination $c_1\psi_1 + c_2\psi_2$ is also a solution.

Why is this important? It lets us build arbitrarily complex wavefunctions from simple plane wave building blocks. This is how we describe localized particles, wave packets, and realistic physical situations.

Let's say that you have two functions that satisfy the Klein-Gordon Equation, call them ψ_1 and ψ_2

$$\left[\Box + \mu^2\right]\psi_1 = 0 \tag{3.1}$$

$$\left[\Box + \mu^2\right]\psi_2 = 0 \tag{3.2}$$

Let's call their sum ψ_3 so

$$\psi_3 = \psi_1 + \psi_2.$$

The question then arises: does ψ_3 satisfy the Klein-Gordon Equation?
Well, yes it does. We can show this with:

$$\begin{aligned}\left[\Box + \mu^2\right]\psi_3 &= \left[\Box + \mu^2\right](\psi_1 + \psi_3) \\ &= \left[\Box + \mu^2\right]\psi_1 + \left[\Box + \mu^2\right]\psi_2 = 0 + 0 = 0 \text{ ... also satisfies K.G..}\end{aligned}$$

The following reasoning does not just apply to the sum of two wave functions, but also scaling up a wave function or taking the arbitrary sum over many wave functions. Or even more usefully if you have a basis set of functions that satisfy the Klein-Gordon Equation, then you can make wave functions of that basis set by summing over them with some series of coefficients

Say we write a wavefunction ψ as a linear combination of $\{\psi_n\}$

$$\psi = \sum_n C_n\psi_n, \qquad C_n \in \mathbb{C} \tag{3.3}$$

Since each basis function in $\psi_n \in \{\psi_n\}$ satisfies K.G....

$$\left[\Box + \mu^2\right]\psi_n = 0 \rightarrow \left[\Box + \mu^2\right]\sum_n C_n\psi_n = 0 \tag{3.4}$$

And that's how we actually use energy and momentum in Eigen-states most of the time.

Chapter 4

Group Velocity and the Speed of Light Limit

4.1 What is Group Velocity?

A single plane wave $e^{i(kx-\omega t)}$ extends infinitely in space — not very physical! Real particles are localized wave packets made by superposing many plane waves with different k values.

For a wave packet:

- **Phase velocity:** $v_p = \omega/k$ — speed of individual wave crests
- **Group velocity:** $v_g = d\omega/dk$ — speed of the packet envelope (the actual particle!)

The group velocity is what we actually measure — it's the speed of information/energy transport. Let's calculate it for Klein-Gordon.

4.2 Step 1: Finding the Dispersion Relation

The **dispersion relation** $\omega(k)$ tells us how frequency depends on wavenumber. It encodes the physics of the wave.

General form of any plane wave:

$$\psi = A \exp\left(i\left(\vec{k}\cdot\vec{x} - \omega t\right)\right) \tag{4.1}$$

Our Klein-Gordon plane wave:

$$\psi = A \exp\left(\frac{i}{\hbar}\left(\vec{p}\cdot\vec{x} - Et\right)\right) \tag{4.2}$$

Comparing these, we identify:

$$\vec{k} = \frac{\vec{p}}{\hbar} \quad \text{(de Broglie relation)}$$
$$\omega = \frac{E}{\hbar} \qquad \text{(Planck relation)}$$

Using $E = c\sqrt{|\vec{p}|^2 + m^2c^2}$ and $|\vec{p}| = \hbar k$:

$$\omega(k) = \sqrt{(kc)^2 + \left(\frac{mc^2}{\hbar}\right)^2} \tag{4.3}$$

This is the dispersion relation! Notice:

- Massless ($m = 0$): $\omega = kc$ (linear — no dispersion)
- Massive ($m \neq 0$): $\omega \neq kc$ (nonlinear — dispersive!)

4.3 Step 2: Calculating the Group Velocity

Now we differentiate the dispersion relation to get the group velocity:

$$v_g = \frac{d\omega}{dk}$$

$$v_g = \frac{d}{dk}\left[\sqrt{(kc)^2 + (mc^2/\hbar)^2}\right] \tag{4.4}$$

Using the chain rule:

$$v_g = \frac{kc^2}{\sqrt{(kc)^2 + (mc^2/\hbar)^2}} = \frac{kc^2}{\omega} \tag{4.5}$$

Substituting $k = |\vec{p}|/\hbar$:

$$v_g = \frac{c|\vec{p}|}{\sqrt{|\vec{p}|^2 + (mc)^2}} \tag{4.6}$$

4.4 The Speed Limit: Why $v_g < c$

Look at what we just derived! The group velocity depends on momentum in a very special way.

For massive particles ($m \neq 0$):

- Low momentum ($|\vec{p}| \ll mc$): $v_g \approx |\vec{p}|/m$ (non-relativistic)
- High momentum ($|\vec{p}| \gg mc$): $v_g \approx c(|\vec{p}|/|\vec{p}|) = c$... but never quite reaches it!
- The denominator $\sqrt{|\vec{p}|^2 + (mc)^2}$ is always larger than $|\vec{p}|$, so $v_g < c$ always

For massless particles ($m = 0$):

$$v_g = \frac{c|\vec{p}|}{|\vec{p}|} = c \tag{4.7}$$

Massless particles always travel at exactly the speed of light, regardless of momentum!

The Klein-Gordon equation naturally enforces special relativity's speed limit. This is a beautiful consistency check — we started with relativistic energy-momentum, and we get relativistic velocities.

Chapter 5

Fourier Transforms and Antimatter

5.1 From Momentum Space to Position Space

We've been working with plane waves — states with definite momentum p. But real particles are localized in space! How do we describe them?

Answer: Fourier transform! We build a position-space wavefunction $\psi(x)$ by superposing plane waves with different momenta, weighted by a momentum-space wavefunction $\phi(p)$.

5.2 The Fourier Transform

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \phi(p) \exp\left[-\frac{i}{\hbar} p \cdot x\right] dp \tag{5.1}$$

- wavefunction in spacetime
- Normalization constant
- Add up eigenstate for each p, weighted $\phi(p)$
- Momentum-space wavefunction (complex numbers assigned to each momentum)
- p-eigenstate (plane wave with momentum p)

The key insight: $\phi(p)$ assigns a complex number (amplitude and phase) to each allowed momentum p. But what are the "allowed" values of p? They must satisfy the mass shell relation!

5.3 The Two Halves of the Mass Shell

There is a one-to-one connection between all possible wavefunctions that satisfy the Klein-Gordon equation in spacetime and all possible ways of decorating the mass shell with complex numbers.

The critical insight: You need *both halves* of the mass shell to have a complete basis set for Fourier transforms from momentum space to position space.

Recall from the plane wave solution that:

$$p^0 = \frac{E}{c} = \pm\sqrt{|\vec{p}|^2 + m^2c^2} \tag{5.2}$$

This ± sign gives us two branches:

- **Positive energy:** $E = +\sqrt{p^2c^2 + m^2c^4}$ (normal particles)
- **Negative energy:** $E = -\sqrt{p^2c^2 + m^2c^4}$ (antimatter!)

5.3.1 Reinterpreting Negative Energy: Feynman-Stueckelberg

Negative energy sounds strange. But there's a beautiful way to reinterpret this that doesn't require "negative energy" at all.

Recall the energy operator:

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \tag{5.3}$$

Shift your perspective: Instead of thinking about negative energy, reinterpret what $-\hat{E}$ means:

$$-\hat{E} = -i\hbar\frac{\partial}{\partial t} \tag{5.4}$$

$$\text{Negative energy??} \quad \longrightarrow \quad \text{Time reversal!}$$

A *negative energy* particle moving *forward in time* is mathematically equivalent to a *positive energy* antiparticle moving *backward in time*.

This is the **Feynman-Stueckelberg interpretation**:

- Particles: positive energy, moving forward in time
- Antiparticles: positive energy, moving backward in time (which *looks like* negative energy forward in time)

When an electron and positron annihilate, you can picture the positron as an electron that reversed its direction in time!

The bottom line: The negative energy solutions represent *antimatter*. When you include both halves of the mass shell, you're accounting for both particles and antiparticles.

5.4 Dirac's Critique: The Fatal Flaws of Klein-Gordon

Dirac identified two deeply connected problems with the Klein-Gordon equation that made it unsuitable as a single-particle quantum theory.

5.4.1 Problem 1: Second-Order in Time

The Klein-Gordon equation is **second-order in time**.

Compare the time derivatives:

$$\text{Schrödinger:} \quad i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \quad \text{(first-order in time)}$$

$$\text{Klein-Gordon:} \quad \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi + \left(\frac{mc}{\hbar}\right)^2\psi = 0 \quad \text{(second-order in time)}$$

Being second-order in time allows both positive and negative energy solutions and treats space and time on equal footing (manifestly relativistic). However, it creates a serious problem.

Too much freedom! Just like classical mechanics needs position AND velocity for second-order equations, Klein-Gordon requires *two* initial conditions:

- The wavefunction: $\psi(x, 0)$
- The time derivative: $\left.\frac{\partial\psi}{\partial t}\right|_{t=0}$

In quantum mechanics, the state should be completely determined by $\psi(x, 0)$ alone. Having $\partial\psi/\partial t$ as an independent initial condition violates this fundamental principle.

5.4.2 Problem 2: Negative Probability Density

This "too much freedom" problem directly causes the negative probability issue.

For Schrödinger, the probability density is simple:

$$\rho = |\psi|^2 = \psi^* \psi \geqslant 0 \quad \text{(always positive!)} \tag{5.5}$$

For Klein-Gordon, deriving the probability density from the continuity equation gives:

$$\rho = \frac{i\hbar}{2mc^2}\left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}\right) \tag{5.6}$$

This depends on both ψ and $\partial\psi/\partial t$! Since $\partial\psi/\partial t$ is an independent degree of freedom (Problem 1), we can choose it to make ρ negative. Negative probabilities are physically nonsensical.

The connection: The negative probability problem exists *because* the equation is second-order in time. These aren't separate issues — they're two sides of the same coin.

5.4.3 Dirac's Solution

Dirac wanted the impossible:

1. **First-order in time** (like Schrödinger) — needs only $\psi(x, 0)$, no extra freedom
2. **Relativistically correct** — treats energy and momentum on equal footing
3. **Positive definite probability** — $\rho \geqslant 0$ always

This seemingly impossible requirement led Dirac to discover the **Dirac equation**, which is first-order in *both* time and space. The price? The wavefunction becomes a multi-component spinor, and quantum mechanical spin emerges naturally!

However, even Dirac's equation still has negative energy solutions. The Feynman-Stueckelberg interpretation (discussed earlier) applies here too — those solutions represent antimatter. Klein-Gordon and Dirac equations aren't single-particle theories — they're fundamentally quantum field theory equations.

Chapter 6

Dirac Equation

6.1 The Dirac Equation

Our goal is to find the analog of the Schrodinger equation of relativistic spin one-half particles, however, we should note that even in the Schrodinger equation, the interaction of the field with spin was rather ad hoc. There was no explanation of gyromagnetic ration of 2. One can incorporate spin into the non-relativistic equation by using the Schrodinger-Pauli Hamiltonian which contains the dot product of the Pauli matrices with the momentum operator.

$$H = \frac{1}{2m}\left(\vec{\sigma}\cdot\left[\vec{p}+\frac{e}{c}\vec{A}\left(\vec{r},t\right)\right]\right)^2 - e\phi\left(\vec{r},t\right) \tag{6.1}$$

A little computation shows that this gives the correct interaction with spin.

$$H = \frac{1}{2m}\left[\vec{p}+\frac{e}{c}\vec{A}\left(\vec{r},t\right)\right]^2 - e\psi\left(\vec{r},t\right) + \frac{e\hbar}{2mc}\vec{\sigma}\cdot\vec{B}\left(\vec{r},t\right) \tag{6.2}$$

This Hamiltonian acts on a two component spinor.

We can extend this concept to use the relativistic energy equation. The idea is to replace $\vec{p}$ with $sig\vec{m}a\cdot\vec{p}$

relativistic energy equation.

$$\left(\frac{E}{c}^2\right) - p^2 = (mc)^2$$

$$\left(\frac{E}{c} - \vec{\sigma}\cdot\vec{p}\right)\left(\frac{E}{c} + \vec{\sigma}\cdot\vec{p}\right) = (mc)^2$$

$$\left(i\hbar\frac{\partial}{\partial x_0} + i\hbar\vec{\sigma}\cdot\vec{\nabla}\right)\left(i\hbar\frac{\partial}{\partial x_0} - i\hbar\vec{\sigma}\cdot\vec{\nabla}\right)\phi = (mc)^2\,\phi.$$

Instead of an equation which is second order in time derivative, we can make a first order equation, like the Schrodinger equation, by extending this equation to four components.

$$\phi^{(L)} = \phi$$

$$\phi^{(R)} = \frac{1}{mc}\left(i\hbar\frac{\partial}{\partial x_0} - i\hbar\vec{\sigma}\cdot\vec{\nabla}\right)\phi^{(L)}.$$

Now rewriting in terms of $\psi A = \phi^{(R)} + \phi^{(L)}$ and $\psi B = \phi^{(R)} - \phi^{(L)}$ and ordering it as a matrix equation, we get an equation that can be written as a dot product between 4-vectors.

$$\begin{aligned}\begin{pmatrix} -i\hbar\frac{\partial}{\partial x_0} & -i\hbar\vec{\sigma}\cdot\vec{\nabla} \\ i\hbar\vec{\sigma}\cdot\vec{\nabla} & i\hbar\frac{\partial}{\partial x_0}\end{pmatrix} &= \hbar\left[\begin{pmatrix} 0 & -i\vec{\sigma}\cdot\vec{\nabla} \\ i\vec{\sigma} & 0\end{pmatrix} + \begin{pmatrix} \frac{\partial}{\partial x_4} & 0 \\ 0 & -\frac{\partial}{\partial x_4}\end{pmatrix}\right] \\ &= \hbar\left[\begin{pmatrix} 0 & -i\sigma_i \\ i\sigma_i & 0\end{pmatrix}\frac{\partial}{\partial x_i} + \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\frac{\partial}{\partial x_4}\right] \\ &= \hbar\left[\gamma_\mu\frac{\partial}{\partial x_4}\right].\end{aligned}$$

Define the 4 by 4 matrices $\gamma\mu$ are by.

$$\begin{aligned}\gamma_i &= \begin{pmatrix} + & i\sigma_i \\ i\sigma_i & 0\end{pmatrix} \\ \gamma_4 &= \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}.\end{aligned}$$

With this definition, the relativistic equation can be simplified a great deal

$$\left(\gamma_\mu\frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right)\psi = 0 \tag{6.3}$$

where the gamma matrices are given by

$$\begin{aligned}\gamma_1 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0\end{pmatrix} \\ \gamma_2 &= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0\end{pmatrix} \\ \gamma_3 &= \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0\end{pmatrix}.\end{aligned}$$

and they satisfy anti-commutation relations.

$$\left\{\gamma_\mu, \gamma_v\right\} = 2\partial_{\mu v} \tag{6.4}$$

In fact any set of matrices satisfy the anti-commutation relations would yield equivalent physics results, however, we will work in the above explicit representation of the gamma matrices.

defining $\overline{\psi} = \psi^\dagger\gamma_4$.

$$j_\mu = ic\overline{\psi}\gamma_\mu\psi \tag{6.5}$$

satisfies the equation of a conserved 4-vector current

$$\frac{\partial}{\partial x_\mu}j_\mu = 0 \tag{6.6}$$

and also transforms like a 4-vector. The fourth component of the vector shows that the probability density is $\psi^\dagger\psi$. This indicates that the normalization of the state includes all four components of the Dirac spinors

For non-relativistic electrons, the first two components of the Dirac spinor are large while the last two are small.

$$\psi = \begin{pmatrix}\psi_a \\ \psi_B\end{pmatrix}$$
$$\psi_B \approx \frac{c}{2mc^2}\vec{\sigma}\cdot\left(\vec{p}+\frac{e}{c}\vec{A}\right)\psi_A \approx \frac{pc}{2mc^2}\psi_A.$$

We use this fact to write an approximate two-component equation derived from the Dirac equation in the non-relativistic limit.

$$\left(\frac{p^2}{2m}-\frac{Ze^2}{4\pi r}-\frac{p^4}{8\pi m^2c^2r^3}+\frac{Ze^2\vec{L}\cdot\vec{S}}{8m^2c^2}\partial^3(\vec{r})\right)\psi = E\begin{pmatrix}N & R\end{pmatrix}\psi \tag{6.7}$$

This "Schrodinger equation", derived from the Dirac equation, agrees well with the one we used to understand the fine structure of Hydrogen. The first two terms are the kinetic and potential energy terms for the unperturbed Hydrogen Hamitonian. The third term is the relativistic correction to the kinetic energy. The fourth term is the correct spin-orbit interaction, including the Thomas Precession effect that we did not take time to understand when we did the NR fine structure. The fifth term is the so called Darwin term which we said would come from the Dirac equation; and now it has.

For a free particle, each component of the Dirac spinor satisfies the Klein-Gordon equation.

$$\psi_{\vec{p}} = u_{\vec{p}}e^{i(\vec{p}\cdot\vec{x}-Et)/\hbar} \tag{6.8}$$

This is consistent with the relativistic energy relation.

The four normalized solutions for a Dirac particle at rest are.

$$\psi^{(1)} = \psi_E = +mc^2, +\hbar/2 \qquad = \qquad \frac{1}{\sqrt{V}}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}$$

$$\psi^{(1)} = \psi_E = +mc^2, +\hbar/2 \qquad = \qquad \frac{1}{\sqrt{V}}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}e^{-imc^2t/\hbar}$$

$$psi^{(1)} = \psi_E = +mc^2, +\hbar/2 \qquad = \qquad \frac{1}{\sqrt{V}}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}e^{-imc^2t/\hbar}$$

$$psi^{(1)} = \psi_E = +mc^2, +\hbar/2 \qquad = \qquad \frac{1}{\sqrt{V}}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}e^{-imc^2t/\hbar}$$

.

The first and third have spin up while the second and fourth have spin down. The first and second are positive energy solutions while the third and fourth are "negative energy solutions", which we still need to understand.

The next step is to find the solutions with definite momentum. The four plane wave solutions to the Dirac equation are:

$$\psi_{\vec{p}}^{(r)} = \sqrt{\frac{mc^2}{|E|\,V}}\, u_{\vec{p}}^{(r)} e^{i\left(\vec{p}\cdot\vec{x}-Et\right)/\hbar} \tag{6.9}$$

where he four spinors are given by:

$$u_{\vec{p}}^{(1)} = \sqrt{\tfrac{E+mc^2}{2mc^2}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z c}{E+mc^2} \\ \frac{(p_x+ip_y)c}{E+mc^2} \end{pmatrix} \qquad u_{\vec{p}}^{(2)} = \sqrt{\tfrac{E+mc^2}{2mc^2}} \begin{pmatrix} 0 \\ 1 \\ \frac{(p_x-ip_y)c}{E+mc^2} \\ \frac{-p_z c}{E+mc^2} \end{pmatrix}$$

$$u_{\vec{p}}^{(3)} = \sqrt{\tfrac{-E+mc^2}{2mc^2}} \begin{pmatrix} \frac{-p_z c}{-E+mc^2} \\ \frac{-(p_x+ip_y)c}{-E+mc^2} \\ 1 \\ 0 \end{pmatrix} \qquad u_{\vec{p}}^{(4)} = \sqrt{\tfrac{-E+mc^2}{2mc^2}} \begin{pmatrix} \frac{-(p_x-ip_y)c}{-E+mc^2} \\ \frac{p_z c}{-E+mc^2} \\ 0 \\ 1 \end{pmatrix}$$

E is positive for solutions 1 and 2 and negative for solutions 3 and 4. The spinors are orthogonal

$$u_{\vec{p}}^{(r)\dagger} u_{\vec{p}}^{(r')} = \frac{|E|}{mc^2}\partial_{rr'} \tag{6.10}$$

and the normalization constants have been set so that the states are properly normalized and the spinors follow the convention given above, with the normalization proportional to energy.

The solutions are not in general eigenstates of any component of spin but are eigenstates of helicity, the component of spin along the direction of the momentum.

Note:-

with E negative, the exponential $e^{i\left(\vec{p}\cdot\vec{x}-Et\right)/\hbar}$ has the phase velocity, the group velocity and the probability flux all in the opposite direction of the momentum as we have defined it. This clearly doesn't make sense. Solutions 3 and 4 need to be understood in a way for which the non-relativistic operators have not prepared us. Let us simply relabel solutions 3 and 4 such that:

$$\vec{p} \to -\vec{p}$$
$$E \to -E.$$

so that all the energies are positive and the momenta point in the direction of the velocities. This mean we change the signs in solutions 3 and 4 as follows

$$\psi_{\vec{p}}^{(1)} = \sqrt{\frac{E+mc^2}{2EV}} \begin{pmatrix} 1 \\ 0 \\ \frac{p_z c}{E+mc^2} \\ \frac{(p_x+ip_y)c}{E+mc^2} \end{pmatrix} e^{i(\vec{p}\cdot\vec{x}-Et)/\hbar}$$

$$\psi_{\vec{p}}^{(2} = \sqrt{\frac{E+mc^2}{2EV}} \begin{pmatrix} 1 \\ 0 \\ \frac{(p_x-ip_y)c}{E+mc^2} \\ \frac{-p_z c}{E+mc^2} \end{pmatrix} e^{i(\vec{p}\cdot\vec{x}-Et)/\hbar}$$

$$\psi^{(3)} = \sqrt{\frac{E+mc^2}{2EV}} \begin{pmatrix} \frac{p_z c}{E+mc^2} \\ \frac{(p_x+ip_y)c}{E+mc^2} \\ 1 \\ 0 \end{pmatrix} e^{-i(\vec{p}\cdot\vec{x}-Et)/\hbar}$$

$$\psi_{\vec{p}}^{(4)} = \sqrt{\frac{|E|+mc^2}{2|E|V}} \begin{pmatrix} \frac{(p_x-ip_y)c}{E+mc^2} \\ \frac{-p_z c}{E+mc^2} \\ 0 \\ 1 \end{pmatrix} e^{-i(\vec{p}\cdot\vec{x}-Et)/\hbar}.$$

we have plane waves of the form

$$e^{\pm ip_\mu x_\mu/\hbar} \tag{6.11}$$

with the plus sign

With the plus sign or solutions 1 and 2 and the minus sign for solutions 3 and 4. These ± sign in the exponential is not very surprising from the point of view of possible solutions to a differential equation. The problem now is that for solutions 3 and for the momentum and energy operators must have a minus sign added to them and the phase of the wave function at a fixed position behaves in the opposite way as a function of time than what we expect and from solutions 1 and 2. It is as if solutions 3 and 4 are moving backward in time

if we change the charge of the electron from $-e$ to $+e$ and change the sign of the exponent, the Dirac equation remains invariant. Thus, we can turn the negative exponent solution (going backwards in time) into the conventional positive exponent solution if we change the charge to $+e$. We can interpret solutions 3 and 4 as positrons. We will make this switch more carefully when we study the charge conjugation operator. The Dirac equation should be the invariant under Lorentz boosts and under rotation, both of which are just changes in the definition of an inertial coordinate system. Under Lorentz boosts, $\frac{\partial}{\partial x_\mu}$ transforms like a 4-vector but the γ_μ matrices are constant. The Dirac equation is shown to be invariant under boosts along the x_i direction if we transform the Dirac spinor according yo:

$$\psi' = S_{boost}\psi$$
$$S_{boost} = \cosh\frac{\chi}{2} + i_{\gamma_i\gamma_4}\sinh\frac{\chi}{2}.$$

with $\tanh_\chi = \beta$

the Dirac equation is invariant under rotations about the k axis if we transform the Dirac equation spinor according to

$$\psi' = S_{rot}\psi$$
$$S_{rot} = \cos\frac{\theta}{2} + \gamma_i\gamma_j\sin\frac{\theta}{2}.$$

with ijk is a cyclic permutation.
Another symmetry related to the choice of coordinates system s parity. Under a parity inversion equation the

Dirac equation remains invariant if

$$\psi' = S_P\psi = \gamma_4\psi \tag{6.12}$$

since $\gamma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, the third and fourth components of the spinor change sign while the first two don't. Since we could have chosen $-\gamma_4$, all we know is that components 3 and 4 have the opposite parity of components 1 and 2. From 4 by 4 matrices, we may derive 16 independent components of covariant objects. We define the product of all gamma matrices.

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4 \tag{6.13}$$

which obviously anticommutes with all the gamma matrices.

$$\left\{\gamma_\mu, \gamma_5\right\} = 0 \tag{6.14}$$

For rotations and boosts, γ_5 commutes with S since it commutes with the pair of gamma matrices. For a parity inversion, it anticommutes with $S_P = \gamma_4$. The simplest set of covariants we can make from Dirac spinors and γ matrices are tabulated below.

Classification	Covariant Form	no. of Components
Scalar	$\overline{\psi}\psi$	1
Pseudoscalar	$\overline{\psi}\gamma_5\psi$	1
Vector	$\overline{\psi}\gamma_\mu\psi$	4
Axial Vector	$\overline{\psi}\gamma_5\gamma_\mu\psi$	4
Rank 2 antisymmetric tensor	$\overline{\psi}\sigma_{\mu v}\psi$	6
Total		16

.

Products of more γ matrices turn out to repeat the same quantities because the square of any γ matrix is 1.

For many purposes, it is useful to write the Dirac equation in the traditional form $H\psi = E\psi$. To do this, we must separate the space and time derivatives, making the equation less covariant looking.

$$\left(\gamma_\mu\frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right)\psi = 0$$

$$\left(ic\gamma_4\gamma_j p_j + mc^2\gamma_4\right)\psi = -\hbar\frac{\partial}{\partial t}\psi.$$

Thus we can identify the operator below as the Hamiltonian.

$$H = ic\gamma_4\gamma_j p_j + mc^2\gamma_4 \tag{6.15}$$

The Hamiltonian helps us identify constants of the motion. If an operator commutes with H, it represents a conversed quantity. It's easy to see the p_k commutes with the Hamiltonian for a free particle so that momentum will be conserved. The components of orbital angular momentum do not commute with H

$$[H, L_z] = ic\gamma_4\left[\gamma_j p_j, xp_y - yp_x\right] = \hbar c\gamma_4\left(\gamma_1 p_y - \gamma_2 p_x\right) \tag{6.16}$$

The components of spin also do not commute with H.

$$[H, S_z] = \hbar c\gamma_4\left[\gamma_2 p_x - \gamma_1 p_y\right] \tag{6.17}$$

But, from the above, the components of total angular momentum do commute with H.

$$[H, J_z] = [H, L_z] + [H, S_z] = \hbar c \gamma_4 \left(\gamma_1 p_y - \gamma_2 p_x\right) + \hbar c \gamma_4 \left[\gamma_2 p_x - \gamma_1 p_y\right] = 0 \tag{6.18}$$

The Dirac equation naturally conserves total angular momentum but not the orbital spin parts of it. We can also see that helicity, or spin along the direction of motion does commute.

$$\left[H, \vec{S} \cdot \vec{p}\right] = \left[H, \vec{S}\right] \cdot \vec{p} = 0 \tag{6.19}$$

For any calculation, we need to know the interaction term with the Electromagnetic field. Based on the interaction of field with a current

$$H_{int} = \frac{1}{c} j_\mu A_\mu \tag{6.20}$$

and the current we have found for the Dirac equation, the interaction Hamiltonian is.

$$H_{int} = ie\gamma_4 \gamma_k A_k \tag{6.21}$$

This is simpler than the non-relativistic case, with no A^2 term and only one power of e. The Dirac equation has some unexpected phenomena we can derive. Velocity eigenvalues for electrons are always $\pm c$ along any direction. Thus the only values of velocity that we could measure are $\pm c$

Localized states, expanded in place waves, contain all four components of the plane wave solutions. Mixing components 1 and 2 with components 3 and 4 gives rise to Zitterbewegung, the very rapid oscillation of an electrons velocity and position.

$$\langle v_k \rangle = \sum_{\vec{p}} \sum_{r=1}^{4} \left|c_{\vec{p},r}\right|^2 \frac{p_k c^2}{E} + \sum_{\vec{p}} \sum_{r=1}^{2} \sum_{r'=3}^{4} \frac{mc^3}{|E|} \left[c^*_{\vec{p},r'} c_{\vec{p},r} u^{(r')^\dagger}_{\vec{p}} i\gamma_4 \gamma_k u^{(r)}_{\vec{p}} r^{-2i|E|t/\hbar} c_{\vec{p},r'} c^*_{\vec{p},r} u^{(r)^\dagger}_{\vec{p}} i\gamma_4 \gamma_k u^{(r')}_{\vec{p}} e^{-2i|E|t/\hbar}\right] \tag{6.22}$$

The last sum which contains the cross terms between negative and positive energy represents extremely high frequency oscillations in the expected value of the velocity, known as Zitterbewegung. The expected value of the position has similar rapid oscillations.

It is possible to solve the Dirac equation exactly for Hydrogen in a way very similar to the non-relativistic solution. One difference is that it is clear from the beginning that the total angular momentum is a constant of the motion and is used as a basic quantum number. There is another conserved quantum number related to the component of spin along the direction of $\vec{J}$. With these quantum numbers, the radial equation can be solved in a similar way as for the non-relativistic case yielding the energy relation

$$E = \frac{mc^2}{\sqrt{1 + \frac{Z^2\alpha^2}{\left(n_r + \sqrt{\left(j+\frac{1}{2}\right)^2 - Z^2\alpha^2}\right)}}} \tag{6.23}$$

We can identify the standard principle quantum number in this case as $n = n_r + j + \frac{1}{2}$. This result gives the same answer as out non-relativistic calculation to order α^4 but is also correct to higher order. It is an exact solution to the quantum mechanics problem posed but does not include the effects of field theory, such as the Lamb shift and the anomalous magnetic moment of the electron.

A calculation of Thomson scattering shows that even simple low energy photon relies on the "negative energy" or positron states to get a non-zero answer. If the calculation is done with the two diagrams in which a photon is absorbed then emitted by an electron (and vice-versa) the result is zero at low energy because the interaction Hamiltonian connects the first and second plane wave states with the third and fourth at zero momentum. This is in contradiction to the classical and non-relativistic calculations as well as measurement. There are additional diagrams if we consider the possibility that the photon can create and electron positron pair which annihilates with the initial electron emitting a photon (or with the initial and final photons swapped). These two terms give the right answer. The calculation of Thomson scattering makes it clear that we cannot ignore the new "negative

energy" or positron states.

The Dirac equation is invariant under charge conjugation, defined as changing electron states into the opposite charged positron states with the same momentum and spin (and changing the sign of external fields). To do this the Dirac spinor is transformed according to.

$$\psi' = \gamma_2 \psi^* \tag{6.24}$$

Of course a second charge conjugation operation takes the state back to the original ψ. Applying this to the place wave function gives

$$\psi^{(1)}_{\vec{p}} = \sqrt{\frac{mc^2}{|E|\,V}} u^{(1)}_{\vec{p}} e^{i\left(\vec{p}\cdot\vec{x}-Et\right)/\hbar} \rightarrow -\sqrt{\frac{mc^2}{|E|\,V}} u^{(4)}_{\vec{p}} e^{i\left(-\vec{p}\cdot\vec{x}+Et\right)/\hbar} \equiv \sqrt{\frac{mc^2}{|E|\,V}} v^{(1)}_{\vec{p}} e^{i\left(-\vec{p}\cdot\vec{x}+Et\right)/\hbar}$$

$$\psi^{(2)}_{\vec{p}} = \sqrt{\frac{mc^2}{|E|\,V}} u^{(2)}_{\vec{p}} e^{i\left(\vec{p}\cdot\vec{x}-Et\right)/\hbar} \rightarrow \sqrt{\frac{mc^2}{||}}.$$