

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	Page 2
1.1	The Strategy: Combining Relativity and Quantum Mechanics	2
1.2	Substituting Operators into Einstein's Relation	3
1.3	The Klein-Gordon Equation!	3
1.4	Compact Notation: Laplacian	4
1.5	Even More Compact: d'Alembertian Notation	4
Chapter 2	Four-momentum Eigenstates	Page 5
2.1	Klein-Gordon Plane Wave Solutions	5
2.2	Proof that plane waves satisfy Klein-Gordon	5
Chapter 3	Superposition	Page 7
3.1	Linearity: Building Complex Solutions from Simple Ones	7
Chapter 4	Group Velocity and the Speed of Light Limit	Page 8
4.1	What is Group Velocity?	8
4.2	Step 1: Finding the Dispersion Relation	8
4.3	Step 2: Calculating the Group Velocity	9
4.4	The Speed Limit: Why $v_g < c$	9
Chapter 5	Fourier Transforms and Antimatter	Page 10
5.1	From Momentum Space to Position Space	10
5.2	The Fourier Transform	10
5.3	The Two Halves of the Mass Shell Reinterpreting Negative Energy: Feynman-Stueckelberg — 11	10
5.4	Dirac's Critique: The Fatal Flaws of Klein-Gordon Problem 1: Second-Order in Time — 11 • Problem 2: Negative Probability Density — 12 • Dirac's Solution — 12	11
Chapter 6	Dirac Equation	Page 13
6.1	The Dirac Equation	13

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:

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

Momentum terms:

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

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

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 :

$$\begin{aligned} (mc)^2 &= \left(\frac{E}{c}\right)^2 - p_x^2 - p_y^2 - p_z^2 \\ (mc)^2 &= -\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} + \hbar^2 \frac{\partial^2}{\partial x^2} + \hbar^2 \frac{\partial^2}{\partial y^2} + \hbar^2 \frac{\partial^2}{\partial z^2}. \end{aligned}$$

Dividing through by \hbar^2 and rearranging:

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

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 :

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

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:

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

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):

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

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:

$$[\square + \mu^2] \psi = 0 \quad (1.4)$$

This compact form makes it easy to see: it's a wave equation (the \square) 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:

$$p = [E/c, \vec{p}] \quad (\text{four-momentum})$$

$$x = [ct, \vec{x}] \quad (\text{four-position})$$

$$A \in \mathbb{C} \quad (\text{complex amplitude})$$

$$p^0 = \frac{E}{c} = \pm \sqrt{|\vec{p}|^2 + m^2 c^2} \quad (\text{energy, with } \pm \text{ solutions!})$$

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

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

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

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:

$$[\square + \mu^2] \psi = 0 \quad \Rightarrow \quad \square \psi = -\mu^2 \psi$$

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:

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

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

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:

$$\square \psi = \left[-\frac{E^2}{c^2 \hbar^2} + \frac{|\vec{p}|^2}{\hbar^2} \right] \psi = -\frac{1}{\hbar^2} \left[\frac{E^2}{c^2} - |\vec{p}|^2 \right] \psi \quad (2.2)$$

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

$$\square \psi = -\frac{m^2 c^2}{\hbar^2} \psi = -\mu^2 \psi \quad \checkmark \quad (2.3)$$

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

$$[\square + \mu^2] \psi_1 = 0 \quad (3.1)$$

$$[\square + \mu^2] \psi_2 = 0 \quad (3.2)$$

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:

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

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 $\{\psi_n\}$

$$\psi = \sum_n C_n \psi_n, \quad C_n \in \mathbb{C} \quad (3.3)$$

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

$$[\square + \mu^2] \psi_n = 0 \rightarrow [\square + \mu^2] \sum_n C_n \psi_n = 0 \quad (3.4)$$

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:

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

Our Klein-Gordon plane wave:

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

Comparing these, we identify:

$$\begin{aligned} \vec{k} &= \frac{\vec{p}}{\hbar} && \text{(de Broglie relation)} \\ \omega &= \frac{E}{\hbar} && \text{(Planck relation)} \end{aligned}$$

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

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

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:

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

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

Using the chain rule:

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

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

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

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 ($m = 0$):

$$v_g = \frac{c|\vec{p}|}{|\vec{p}|} = c \quad (4.7)$$

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

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

- 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:

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

This \pm sign gives us two branches:

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

$$\hat{E} = i\hbar \frac{\partial}{\partial t} \quad (5.3)$$

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

$$-\hat{E} = -i\hbar \frac{\partial}{\partial t} \quad (5.4)$$

Negative energy?? \longrightarrow 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:

$$\begin{aligned} \text{Schrödinger: } i\hbar \frac{\partial \psi}{\partial t} &= \hat{H}\psi \quad (\text{first-order in time}) \\ \text{Klein-Gordon: } \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi + \left(\frac{mc}{\hbar}\right)^2 \psi &= 0 \quad (\text{second-order in time}) \end{aligned}$$

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:

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

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

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

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** — $\rho \geq 0$ 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 ratio 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.

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

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

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

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 $\vec{\sigma} \cdot \vec{p}$ relativistic energy equation.

$$\begin{aligned} \left(\frac{E}{c} \right)^2 - p^2 &= (mc)^2 \\ \left(\frac{E}{c} - \vec{\sigma} \cdot \vec{p} \right) \left(\frac{E}{c} + \vec{\sigma} \cdot \vec{p} \right) &= (mc)^2 \\ \left(i\hbar \frac{\partial}{\partial x_0} + i\hbar \vec{\sigma} \cdot \vec{\nabla} \right) \left(i\hbar \frac{\partial}{\partial x_0} - i\hbar \vec{\sigma} \cdot \vec{\nabla} \right) \phi &= (mc)^2 \phi. \end{aligned}$$

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.

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

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.

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