A theoretical introduction to wave mechanics

Chris Clark

Spicy Lifestyle Academy, Department of Physics, Tokyo, Japan[*]
(Dated: September 17, 2010)

A pedagogically-oriented historical introduction to the theoretical aspects of wave mechanics followed by a discussion of real-valued wave functions in the Klein-Gordon equation.

INTRODUCTION

At the turn of the 20th century, physics was starting to look rather mature and polished. At about this time Albert Michelson wrote,

> "The more important fundamental laws and facts of physical science have all been discovered, and these are so firmly established that the possibility of their ever being supplanted in consequence of new discoveries is exceedingly remote. Nevertheless, it has been found that there are apparent exceptions to most of these laws, and this is particularly true when the observations are pushed to a limit, i.e., whenever the circumstances of experiment are such that extreme cases can be examined. Such examination almost surely leads, not to the overthrow of the law, but to the discovery of other facts and laws whose action produces the apparent exceptions." [1]

However, there were already some indications that big changes were on the horizon. One was the problem that the version of ether theory in existence at the time could not easily be reconciled with the results of the experiment that Michelson had carried out with Morley. Another was the so-called ultraviolet catastrophe that came about with Lord Rayleigh's 1900 version of the Rayleigh-Jeans law of blackbody radiation.

THE PLANCK-EINSTEIN EQUATION

On December 14th 1900, Max Planck presented a derivation of the blackbody radiation law that was based on the assumption that electromagnetic radiation could only be emitted in particle-like packets with a fixed ratio of energy to frequency. This assumption can be written as

$$E = h\nu$$

where ν is the frequency and h is a universal constant. This equation came to be known as Planck's equation, though Planck did not initially consider it to be a real physical law. In a letter to a colleague, he wrote

> "To summarize, all what happened can be described as simply an act of desperation... This was a purely formal assumption and I really did not give it much thought except that, no matter what the cost, I must bring about a positive result." [2]

After all, light had been considered decisively wave-like in nature since Thomas Young's double-slit experiment in 1803. However, in 1905, Albert Einstein explained the photoelectric effect by suggesting that this equation was physically true. Einstein also discovered that Planck's assumption provided a solution to the ultraviolet catastrophe, which is something Planck hadn't been aware of.

Then in 1923, Arthur Compton performed an experiment in which X-rays were scattered off electrons. His results demonstrated that light must consist of particle-like objects with energy proportional to frequency, thus confirming Einstein's suggestion.

THE DE BROGLIE EQUATION

After Einstein's photoelectric theory and Compton's X-ray experiment, light had to be viewed as having both particle and wave properties, but massive particles such as electrons were still widely considered completely distinct from waves. However, puzzling questions about the particle interpretation of electrons were already developing as early as 1913 with Niels Bohr's model of the Hydrogen atom. Bohr had succeeded in theoretically explaining the Rydberg formula for the emission specturm of Hydrogen by assuming that the angular momentum of the orbiting electron was restricted to integer multiples of Planck's constant h. Despite its impressive results, Bohr's model had many shortcomings. It was inconsistent with several other observations, and theoretically provided no justification for the quantization assumption. This inspired Louis de Broglie to propose the idea that all particles possess an associated matter wave.

> "...the determination of the stable motions of the electrons in the atom involves whole numbers, and so far the only phenomena in which whole numbers were involved in physics were those of interference and of eigenvibrations. That suggested the idea to me that electrons

themselves could not be represented as simple corpuscles either, but that a periodicity had also to be assigned to them too." [3]

He may have also been encouraged to take this leap by Einstein's work on the photoelectric effect.

"...I had a sudden inspiration. Einstein's wave-particle dualism was an absolutely general phenomenon extending to all physical nature." [4]

The most famous equation to emerge from de Broglie's work was his relation between the momentum of a particle and the wavelength of its matter wave. Such a relation was already known for light quanta. By combining the Planck-Einstein equation, $E = h\nu$, with the phase velocity formula, $c = \lambda\nu$, and the result from electromagnetism. $E = pc$ (based on Poynting's Theorem), one can derive

$$p = \frac{E}{c} = \frac{h\nu}{\lambda\nu} = \frac{h}{\lambda}$$

In 1923, de Broglie suggested that if matter waves existed, they would also obey this equation. [5] He started by making the assumption, now believed to be false, that light quanta have some very small but non-zero mass. Under this assumption, he equated the energy of a light quantum given by the Planck-Einstein equation, $E = h\nu$, with the relativistic energy of a massive particle given by $E = \gamma mc^2$. In terms of the constants $\hbar = h/2\pi$ and $\omega = 2\pi\nu$, this can be written as

$$\hbar\omega = \gamma mc^2$$

According to theory of relativity, this equation should apply in all inertial reference frames. So in the rest frame of the particle, where $\gamma = 1$, there should be a frequency of oscillation ω_0 such that

$$\hbar\omega_0 = mc^2$$

In this frame the particle is not moving, so this oscillation would presumably look like a localized standing wave. Now suppose we boost to a reference frame that is moving at velocity v_1 with respect to the particle. The particle remains untouched, so its matter wave cannot have changed. However, due to relativistic time-dilation, the apparent frequency of oscillation of the standing wave will be reduced to

$$\omega_0' = \frac{\omega_0}{\gamma_1}$$

If we were to measure the frequency of the wave from this frame by counting the number of crests that pass by per second, we would not get ω_0'. The reason is that the wave crests will be passing by at a higher rate due to our motion relative to the standing wave. Essentially we

are traveling through the wave form as it is oscillating in place. The actual frequency of the wave in our new reference frame is still given by the Planck-Einstein equation. Thus, the relation between the frequency of the wave, ω_1, and the apparent frequency of its internal oscillation, ω_0', is

$$\omega_1 = \frac{\gamma_1 mc^2}{\hbar} = \gamma_1\omega_0 = \gamma_1^2\omega_0'$$

Now consider the value of the wave at the center of the light quantum, $\psi_0(t)$, still observing from our moving frame. As implied by the definition of ω_0', the value will oscillate sinusoidally according to

$$\psi_0(t) = A\sin(\omega_0' t)$$

provided a suitable choice for the zero point of time.

Even though the wave has to be confined to a finite region, the closer we zoom in on the center, the more it will look like an infinite monochromatic plane wave, since we are assuming a light quantum with a definite frequency. The frequency of this monochromatic wave is what we measure by counting passing crests, namely ω_1. Therefore, we can approximate the value of the wave near the center of the quantum as

$$\psi(x, t) = A\sin(\omega_1 t - kx)$$

with coordinates chosen such that the center of the quantum passes $x = 0$ at time $t = 0$, and k being an undetermined wave vector. The value at the center of the particle, $\psi_0(t)$, can be evaluated from this expression by plugging in the distance that we have traveled from the center at time t i.e. $x = v_1 t$.

$$\psi_0(t) = \psi(v_1 t, t) = A\sin(\omega_1 t - kv_1 t)$$

Equating the two expressions for $\psi_0(t)$ gives

$$A\sin(\omega_0' t) = A\sin(\omega_1 t - kv_1 t)$$

For this to be true, the parameters of the sine functions must be equal.

$$\omega_0' = \omega_1 - kv_1$$

$$\left(1 - \frac{v_1^2}{c^2}\right)\omega_1 = \omega_1 - kv_1$$

$$\left(\frac{v_1^2}{c^2}\right)\omega_1 = kv_1$$

$$\frac{\omega_1}{k} = \frac{c^2}{v_1}$$

This tells us the phase velocity $u = \omega/k$ of the matter wave for any velocity, which automatically gives us the wavelength since $u = \lambda\nu$.

$$\lambda = \frac{u}{\nu} = \frac{h}{\gamma mc^2}\frac{\omega}{k} = \frac{h}{\gamma mc^2}\frac{c^2}{v} = \frac{h}{\gamma mv} = \frac{h}{p}$$

Strictly speaking, this derivation only applied to light quanta because there was no scientific evidence at the time to suggest that other types of particles obeyed the Planck-Einstein equation. So when de Broglie concluded the derivation with the statement,

> "We are then inclined to admit that any moving body may be accompanied by a wave and that it is impossible to disjoin motion of body and propagation of wave" [6]

he was taking a radial step.

THE SCHRÖDINGER EQUATION

De Broglie summarized his work in his doctoral thesis. It is said that his thesis examiners were unsure about giving it their blessing, so they sent it to Einstein for his opinion. [7] Einstein endorsed it, which got de Broglie his PhD, and also led Einstein to promote de Broglie's idea. Erwin Schrödinger, who had already been working on the quantum theory of electrons in 1922 [8], was exposed to deBroglie's ideas through Einstein. [9] In 1926, Schrödinger published a paper that introduced his famous equation governing the behavior of particle wave functions. The introduction began,

> "The theory which is reported in the following pages is based on the very interesting and fundamental researches of L. de Broglie on what he called "phase waves" and thought to be associated with the motion of material points, especially with the motion of an electron or proton. The point of view taken here... is rather that material points consist of, or are nothing but, wave-systems." [10]

In taking this leap to a fully wave-based view of matter, Schrödinger was compelled by the similarity between Fermat's principle, pertaining to wave paths, and Hamilton's principle, pertaining to particle paths. With this relationship in mind, Schrödinger considered the inadequacy of classical physics to explain atomic emission spectra and asked,

> "...is one not greatly tempted to investigate whether the non-applicability of ordinary mechanics to micro-mechanical problems is perhaps of exactly the same kind as the non-applicability of geometrical optics to the phenonema of diffraction or interference and may,

perhaps, be overcome in an exactly similar way?"

> "As stated above, the wave-phenomena must in this case be studied in detail. This can only be done by using an "equation of wave propagation." Which one is this to be? In the case of a single material point, moving in an external field of force, the simplest way is to try to use the ordinary wave equation..." [10]

The wave equation, already well-known from electromagnetism, is

$$\frac{\partial^2 \psi}{\partial t^2} = u^2 \nabla^2 \psi$$

where ψ is the wave function and u is the phase velocity of the wave. The phase velocity is exactly what de Broglie calculated, so we just need to use de Broglie's results to plug in for u (though Schrödinger actually used a different method to obtain the phase velocity that was based on the analogy between Hamilton's principle and Fermat's principle). The only catch is that we can't just plug in the result $u = c^2/v$ because v is an unknown variable. Since we only have one equation, we can only allow one unknown, which is ψ; otherwise the equation won't be solvable. The solution is to express the phase velocity in terms of the total energy of the particle and the external potential, since the total energy is a constant and the external potential is known by assumption.

Initially, Schrödinger used the correct relativistic energy and found the relativistic wave equation. However, the predictions that came out of the relativistic wave equation contradicted the experimental data on the fine structure lines of the Hydrogen spectrum. The reason was that the equation neglected the spin of the electron, which he correctly suspected, but he did not know how to fix the problem. [11] So instead he swept the issue under the rug by taking the non-relativistic approximation, which chops off the fine structure term. This places the focus on the gross structure predictions, which did match well with experiment. The consequence is that Schrödinger's famous equation is only an approximation, not a true law of physics.

In the non-relativistic approximation, the total energy can be split into the kinetic (K) and potential (V) components as follows.

$$E = K + V = \frac{p^2}{2m} + V$$

$$\frac{p^2}{2m} = E - V$$

$$p = \sqrt{2m(E - V)}$$

Using this expression and de Broglie's result $\lambda = h/p$ gives

$$u = \lambda\nu = \frac{h}{p}\frac{E}{h} = \frac{E}{\sqrt{2m(E-V)}}$$

Notice we have assumed that the energy in the Planck-Einstein equation is the total energy, even though we have not ruled out the possibility that it should be the kinetic energy since the kinetic energy and total energy of photons are equal. Schrödinger used a different method to obtain u, based on the analogy between Hamilton's principle and Fermat's principle that had inspired de Broglie. For more on this issue, see Appendix 1. Plugging this u into the wave equation,

$$\frac{\partial^2\psi}{\partial t^2} = \frac{E^2}{2m(E-V)}\nabla^2\psi$$

Now we restrict the possible solutions to those with frequency $\nu = E/h$ by declaring the time-dependence of ψ so that

$$\psi(x,t) = \psi(x,0)e^{\pm i\omega t} = \psi(x,0)e^{\pm iEt/\hbar}$$

Note that here Schrödinger introduces complex numbers without justification or comment. This was perhaps a seemingly harmless assumption because electromagnetic theory taught physicists that waves are more conveniently written in complex notation. The first two time derivatives of ψ are

$$\frac{\partial\psi}{\partial t} = \pm\frac{iE}{\hbar}\psi(x,0)e^{\pm iEt/\hbar} = \pm\frac{iE}{\hbar}\psi(x,t)$$

$$\frac{\partial^2\psi}{\partial t^2} = -\frac{E^2}{\hbar^2}\psi(x,0)e^{\pm iEt/\hbar} = -\frac{E^2}{\hbar^2}\psi(x,t)$$

With this we can eliminate the second time derivative

$$-\frac{E^2}{\hbar^2}\psi = \frac{E^2}{2m(E-V)}\nabla^2\psi$$

$$(E-V)\psi = -\frac{\hbar^2}{2m}\nabla^2\psi$$

$$E\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$$

This is known as the time-independent Schrödinger equation. To get the time-dependent Schrödinger equation we just substitute the first time-derivative back in

$$\pm i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$$

The sign ambiguity, as Schrödinger points out, is not a problem; ψ obeys the equation for one sign and its complex conjugate obeys the equation with the other sign. Despite the fact that this equation is only a non-relativistic approximation, Schrödinger was able to use it to derive the energy levels of the Bohr atom, which provided convincing evidence for the wave interpretation of electrons. Some mark this as the birth of wave mechanics.

THE SCHRÖDINGER CONTINUITY EQUATION

Shortly after publication of his famous wave equation, Schrödinger added another important piece to the puzzle that would aid in establishing an interpretation of the wave function. [12] His discovery was that the wave function could be fit into the continuity equation familiar from electromagnetism,

$$\frac{\partial\rho}{\partial t} = -\nabla\cdot\mathbf{j}$$

with the proper definitions of ρ and $\mathbf{j}$ in terms of ψ. The Schrödinger continuity equation can be derived directly from the Schrödinger equation. Multiplying the whole equation by $-i/\hbar$,

$$\frac{\partial\psi}{\partial t} = \frac{i\hbar}{2m}\nabla^2\psi - \frac{i}{\hbar}V\psi$$

The complex conjugate of this is,

$$\frac{\partial\psi^*}{\partial t} = -\frac{i\hbar}{2m}\nabla^2\psi^* + \frac{i}{\hbar}V\psi^*$$

Therefore, the quantity $\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}$ is

$$\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t} = \frac{i\hbar}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right)$$

Equivalently,

$$\frac{\partial}{\partial t}(\psi^*\psi) = \frac{i\hbar}{2m}\nabla\cdot\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)$$

This has the form of the continuity equation, where

$$\rho = |\psi|^2$$

$$\mathbf{j} = -\frac{i\hbar}{2m}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)$$

Naturally, Schrödinger was led to interpret ρ and $\mathbf{j}$ as the electric charge and current densities for the particle being modeled. For the case of an electron he suggested,

> "...the charge of the electron is not concentrated in a point, but is spread out through the whole space, proportional to the quantity $\psi\bar{\psi}$." [10]

THE BORN RULE

By comparing the theoretical predictions of the wave equation to the experimental data on Hydrogen's spectrum, Schrödinger was able to further justify his hypothesis about the interpretation of the wave function ψ. However, by the end of the year this interpretation of the wave function would be overturned by Max Born. [13] Back in 1925, a year before Schrödinger published his wave equation, one of Born's assistants named Werner Heisenberg had introduced a statistical form of quantum theory that would come to be known as matrix mechanics. Heisenberg's theory was motivated by the problem of accurately predicting atomic emission spectra, which are determined by atomic state transitions. In 1916, Einstein had established the idea of assigning transition probabilities to each atomic state transition. As a result of these origins, Heisenberg's theory was inherently probabilistic. Born revised Heisenberg's theory, including reformulating it in terms of matrices. Therefore, Born had a predisposition against Schrödinger's non-probabilistic interpretation of the wave function.

> "To us in Göttingen [Schrödinger's] interpretation seemed unacceptable in face of well established experimental facts... I had therefore, as early as the end of 1925, made an attempt to extend the matrix method... Again an idea of Einstein's gave me the lead. He had tried to make the duality of particles - light quanta or photons - and waves comprehensible by interpreting the square of the optical wave amplitudes as probability density for the occurrence of photons. This concept could at once be carried over to the ψ-function: $|\psi|^2$ ought to represent the probability density for electrons (or other particles)." [14]

Born chose to consider an atomic scattering problem in which a swarm of electrons collide with a heavy atom. He solved Schrödinger's equation for this problem and made the following observations about the wave function.

> "The square of the amplitude of this wave at a great distance from the scattering centre determines the relative probability of scattering as a function of direction. Moreover, if the scattering atom itself is capable of existing in different stationary states, then Schrödinger's wave equation gives automatically the probability of excitation of these states..." [14]

These observations led Born to a probabilistic interpretation of the wave function, in which $|\psi|^2$ represents the probability density for the particle's position. The general form of this probabilistic interpretation of the wave function is known as the Born rule. Einstein, De Broglie, and Schrödinger all disliked this interpretation. [14] One reason was that it seemed counter-intuitive that the universe would operate randomly with no underlaying deterministic behavior since ordinarily randomness emerges from outcome counting of deterministic processes, reducing randomness to an illusion of ignorance. Another issue was that the probabilistic interpretation created a paradoxical wave-particle duality in which quantized particles still exist, but mysteriously act like waves in unspecified circumstances. The wave-particle duality issue can be nearly resolved by assuming that there are no particles; just waves obeying the stipulation that interactions are quantized. In perhaps the most famous book on optics ever written, Max Born and Emil Wolf corroborate this view for the case of light.

> "Light propagates as if it were an an electromagnetic wave, but interacts with matter as if its energy were concentrated in photons, each with an energy quantum." [15]

If we extend this view to apply to all species of "particles", the wave-particle duality ceases to be so paradoxical. However, the mechanism of quantization and probabilistic behavior in interactions still remains an open question.

THE KLEIN-GORDON EQUATION

When de Broglie read Schrödinger's first publication on the wave equation, he immediately objected to the non-relativistic approximation that Schrödinger had employed. [16] De Broglie's concept of matter waves had been heavily based on the relativistic phenomenon of time dilation, so he could not be satisfied with a non-relativistic theory. Schrödinger had already worked out the relativistic wave equation in December 1925, and it is believed that he was the first to do so, but as discussed earlier, he didn't publish it. Many others also saw the need for a relativistic equation, and during 1926 several researchers independently reached the same result. The first to publish was Oskar Klein in April 1926, where the equation played a subordinate role in his five-dimensional theory of gravitation and electromagnetism that would later develop into Kaluza-Klein theory. Walter Gordon independently arrived at the equation a few months later in a paper addressing Compton scattering. Despite numerous other discoverers, the equation eventually came to be known as the Klein-Gordon equation. [16]

The Klein-Gordon equation can be obtained in the same manner as the Schrödinger equation; we simply insert the proper relativistic expressions for energy and momentum. Again we start by expressing the momentum in terms of the total energy and potential. In order In order

to do that, we need to eliminate β using the relativistic energy expression.

$$E = \gamma m c^2 + V$$

$$E - V = \frac{mc^2}{\sqrt{1 - \beta^2}}$$

$$1 - \beta^2 = \frac{m^2 c^4}{(E - V)^2}$$

$$\beta^2 = 1 - \frac{m^2 c^4}{(E - V)^2}$$

$$\begin{aligned} p^2 &= \gamma^2 m^2 \beta^2 c^2 = \beta^2 \frac{m^2 c^2}{1 - \beta^2} \\ &= \left(1 - \frac{m^2 c^4}{(E - V)^2}\right)\left(\frac{(E - V)^2}{c^2}\right) \\ &= \frac{1}{c^2}\left((E - V)^2 - m^2 c^4\right) \end{aligned}$$

Now that we have the momentum expressed in terms of E and V, we can easily get the phase velocity in terms of the same parameters.

$$u^2 = \lambda^2 \nu^2 = \frac{h^2}{p^2}\frac{E^2}{h^2} = \frac{c^2 E^2}{(E - V)^2 - m^2 c^4}$$

Finally, we can insert the phase velocity into the wave equation.

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{c^2 E^2}{(E - V)^2 - m^2 c^4}\nabla^2 \psi$$

$$\left((E - V)^2 - m^2 c^4\right)\frac{\partial^2 \psi}{\partial t^2} = c^2 E^2 \nabla^2 \psi$$

Now since the time-dependence of ψ is of the form $e^{\pm i\omega t}$, we can substitute $\frac{\partial^2 \psi}{\partial t^2} = -\frac{E^2}{\hbar^2}\psi(x,t)$.

$$-\left((E - V)^2 - m^2 c^4\right)\frac{E^2}{\hbar^2}\psi = c^2 E^2 \nabla^2 \psi$$

Canceling the factor E^2 and rearranging,

$$(E - V)^2 \psi = -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi$$

If the potential is time-independent, we can write

$$E^2 \psi - 2EV\psi + V^2 \psi = -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi$$

Inserting time derivatives to eliminate E yields the time-dependent Klein-Gordon equation for the time-independent potential V.

$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} - 2i\hbar V \frac{\partial \psi}{\partial t} + V^2 \psi = -\hbar^2 c^2 \nabla^2 \psi + m^2 c^4 \psi$$

There are three fundamental problems with this equation as it currently stands. The first, as we have already mentioned, is that it does not account for spin. Since all fundamental particles discovered to date have non-zero spin, this equation can only be experimentally tested on spineless composite composite particles like the spinless pion. Such experiments have validated the Klein-Gordon equation. [17] However this doesn't necessarily mean the Klein-Gordon is inapplicable to free fundamental particles since spin might only be relevant during interactions.

The other two problems are that the Klein-Gordon equation seems to permit (1) negative probability densities and (2) negative energies. We will encounter these two problems in the upcoming sections.

THE KLEIN-GORDON CONTINUITY EQUATION

A continuity equation can be found for the Klein-Gordon equation by following a procedure similar to that used to get the Schrödinger continuity equation. Dividing the Klein-Gordon equation by $-\hbar^2$ and rearranging,

$$\frac{\partial^2 \psi}{\partial t^2} = -\frac{2i}{\hbar}V\frac{\partial \psi}{\partial t} + \frac{V^2}{\hbar^2}\psi + c^2 \nabla^2 \psi - \frac{m^2 c^4}{\hbar^2}\psi$$

The complex conjugate is

$$\frac{\partial^2 \psi^*}{\partial t^2} = \frac{2i}{\hbar}V\frac{\partial \psi^*}{\partial t} + \frac{V^2}{\hbar^2}\psi^* + c^2 \nabla^2 \psi^* - \frac{m^2 c^4}{\hbar^2}\psi^*$$

So the quantity $\psi^*\frac{\partial^2 \psi}{\partial t^2} - \psi\frac{\partial^2 \psi^*}{\partial t^2}$ is

$$\psi^*\frac{\partial^2 \psi}{\partial t^2} - \psi\frac{\partial^2 \psi^*}{\partial t^2}$$

$$= -\frac{2i}{\hbar}V\left(\psi^*\frac{\partial \psi}{\partial t} + \psi\frac{\partial \psi^*}{\partial t}\right) + c^2\left(\psi^*\nabla^2 \psi - \psi\nabla^2 \psi^*\right)$$

$$\frac{\partial}{\partial t}\left(\psi^*\frac{\partial \psi}{\partial t} - \psi\frac{\partial \psi^*}{\partial t} + \frac{2i}{\hbar}V\psi^*\psi\right) = c^2 \nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*)$$

To make the units consistent with the Schrödinger continuity equation, we multiply this equation by $i\hbar/2mc^2$.

$$\frac{i\hbar}{2mc^2}\frac{\partial}{\partial t}\left(\psi^*\frac{\partial \psi}{\partial t} - \psi\frac{\partial \psi^*}{\partial t} + \frac{2i}{\hbar}V\psi^*\psi\right)$$

$$= \frac{i\hbar}{2m}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*)$$

This has the form of the continuity equation,

$$\frac{\partial \rho_0}{\partial t} = -\nabla\cdot\mathbf{j}_0$$

with

$$\rho_0 = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) - \frac{i}{mc^2}V\psi^*\psi$$

$$\mathbf{j}_0 = -\frac{i\hbar}{2m}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)$$

It is clear that if ψ and ψ^* are unconstrained, they can be chosen such that ρ can be either positive or negative. Since ρ is interpreted as a probability density for the particle, this means that negative probability densities are permitted. This is the second problem with the Klein-Gordon equation.

THE KLEIN-GORDON DISPERSION RELATION

A dispersion relation gives the relationship between frequency and wavelength for waves in a particular medium. For some media, the phase velocity of waves is constant and the dispersion relation can be written as $u = \lambda\nu$ since u is a constant. Such media are called non-dispersive. Dispersive media, on the other hand, produce waves with a phase velocity that varies with the frequency. It is only in dispersive media that waves can propagate in packets that travel at varying speeds. This is because the speed of wave packets is measured by the group velocity $v_g = \frac{\partial\omega}{\partial k}$, which is constant if the phase velocity is constant. So we know that matter waves must propagate in a dispersive medium since we observe particles traveling at different velocities. Therefore, we will derive the dispersion relation for the Klein-Gordon equation to gain insight into the characteristics of the medium. All we need to start is the phase velocity. Since $u = \omega/k$, if we eliminate all variables besides ω and k, the result will be a dispersion relation. Using deBroglie's phase velocity $u = c^2/v = c/\beta$,

$$\frac{\omega}{k} = \frac{c}{\beta}$$

$$\hbar^2\omega^2 = \hbar^2 k^2\frac{c^2}{\beta^2} = p^2\frac{c^2}{\beta^2} = \gamma^2 m^2 c^4$$

$$\hbar^2\omega^2(1-\beta^2) = m^2 c^4$$

$$\hbar^2\omega^2 = \hbar^2\omega^2\beta^2 + m^2 c^4$$

Using the phase velocity equation once more, we obtain the dispersion relation.

$$\hbar^2\omega^2 = \hbar^2 k^2 c^2 + m^2 c^4$$

From this equation, we can see that negative values of ω are permitted. This means that the energy $E = \hbar\omega$ can also be negative, which is the third problem with the Klein-Gordon equation.

REAL WAVE FUNCTIONS

Historically, the three problems with the Klein-Gordon equation were resolved by Paul Dirac's introduction of the Dirac equation, which incorporated spin and eliminated negative probability densities by reducing the second order time derivative to a system of equations with one time derivative each. Negative energy solutions were interpreted as corresponding to antiparticles, resolving the third problem. Rather than going down the path of Dirac, which increases the level of mathematical abstraction with matrix-valued complex wave functions, we will investigate the possibility of decreasing abstraction by analyzing the Klein-Gordon equation with real wave functions.

A propagating sinusoidal wave cannot be completely characterized by a single real amplitude function. Even if the amplitude is given at every point in space at a specific time, there is still no way to tell which direction it is propagating because sinusoidal functions are symmetrical. In fact, we need two real functions to fully describe propagating sinusoidal waves. One function gives the amplitude at every point, and the other corresponds (proportionally) to the amplitudinal momentum at every point. Let these functions be called U and V respectively. The amplitudinal momentum must be proportional to the time derivative of the amplitude, so we can define V by

$$\frac{\partial U}{\partial t} = V$$

which requires

$$\frac{\partial V}{\partial t} = c^2\nabla^2 U - \frac{m^2 c^4}{\hbar^2}U$$

in order that U obeys the Klein-Gordon equation, which can be obtained directly by eliminating V through substitution. Taking the time derivative of the second equation shows that V also satisfies the Klein-Gordon equation. Since both real fields obey the same equation, they can be combined into a complex field that obeys the same equation, the real part proportional to U and the imaginary part proportional to V. We will see soon that this reproduces the textbook complex Klein-Gordon interpretation. There is a discussion on real wave functions in reference [18]. We have omitted the external potential and will continue to do so while discussing real wave functions.

THE CANONICAL CONTINUITY EQUATION
FOR REAL WAVE FUNCTIONS

To derive the canonical continuity equation for real wave functions obeying the Klein-Gordon equation, we

follow the standard prescription, but use V in place of the usual complex conjugate.

$$V\frac{\partial^2 U}{\partial t^2} = c^2 V\nabla^2 U - \frac{m^2 c^4}{\hbar^2}VU$$

$$U\frac{\partial^2 V}{\partial t^2} = c^2 U\nabla^2 V - \frac{m^2 c^4}{\hbar^2}UV$$

Subtracting the second from the first,

$$V\frac{\partial^2 U}{\partial t^2} - U\frac{\partial^2 V}{\partial t^2} = c^2 V\nabla^2 U - c^2 U\nabla^2 V$$

$$\frac{\partial}{\partial t}\left(V\frac{\partial U}{\partial t} - U\frac{\partial V}{\partial t}\right) = c^2 \nabla\cdot(V\nabla U - U\nabla V)$$

We multiply both sides by $\hbar^2/2mc^2$ to give this quantity units of energy.

$$\frac{1}{2}\frac{\hbar^2}{mc^2}\frac{\partial}{\partial t}\left(V\frac{\partial U}{\partial t} - U\frac{\partial V}{\partial t}\right) = \frac{1}{2}\frac{\hbar^2}{m}\nabla\cdot(V\nabla U - U\nabla V)$$

This is the continuity equation, having the form

$$\frac{\partial \rho_1}{\partial t} = -\nabla\cdot\mathbf{j}_1$$

where

$$\rho_1 = \frac{1}{2}\frac{\hbar^2}{mc^2}\left(V\frac{\partial U}{\partial t} - U\frac{\partial V}{\partial t}\right)$$

$$\mathbf{j}_1 = \frac{1}{2}\frac{\hbar^2}{m}(U\nabla V - V\nabla U)$$

THE RELATION BETWEEN REAL AND COMPLEX WAVE FUNCTIONS

As mentioned earlier, a complex wave function with real part proportional to U and complex part proportional to V will also satisfy the same Klein-Gordon equation because both components obey the equation independently. With the appropriate constants included, the complex wave function can be written as

$$\psi = \sqrt{\frac{1}{2}}\left(U + i\frac{\hbar}{mc^2}V\right)$$

The question now is whether this ψ corresponds to and gives the same results as the textbook Klein-Gordon wave function. It suffices to show that the probability flux and probability density correspond because these are the only measurable quantities modeled by the Klein-Gordon equation. The textbook complex wave function's probability flux is

$$
\begin{aligned}
\mathbf{j}_0 &= \frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)\\
&= \frac{\hbar}{2mi}\left((\psi_R - i\psi_I)\nabla(\psi_R + i\psi_I)\right.\\
&\quad\left. - (\psi_R + i\psi_I)\nabla(\psi_R - i\psi_I)\right)\\
&= \frac{\hbar}{2mi}(\psi_R\nabla\psi_R + i\psi_R\nabla\psi_I - i\psi_I\nabla\psi_R + \psi_I\nabla\psi_I\\
&\quad - (\psi_R\nabla\psi_R - i\psi_R\nabla\psi_I + i\psi_I\nabla\psi_R + \psi_I\nabla\psi_I))\\
&= \frac{\hbar}{2mi}(2i\psi_R\nabla\psi_I - 2i\psi_I\nabla\psi_R)\\
&= \frac{\hbar}{m}\left(\frac{\hbar}{2mc^2}U\nabla V - \frac{\hbar}{2mc^2}V\nabla U\right)\\
&= \frac{1}{2}\frac{\hbar^2}{m^2 c^2}(U\nabla V - V\nabla U)
\end{aligned}
$$

which is exactly equal to $\mathbf{j}_1/mc^2$ for real wave functions.

The textbook complex wave function's probability density is

$$
\begin{aligned}
\rho_0 &= \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) = \frac{i\hbar}{mc^2}\mathrm{Im}\left[\psi^*\frac{\partial\psi}{\partial t}\right]\\
&= \frac{i\hbar}{2mc^2}\mathrm{Im}\left[\left(U - i\frac{\hbar}{2mc^2}V\right)\right.\\
&\quad\left.\times\left(\frac{\partial U}{\partial t} + i\frac{\hbar}{mc^2}\frac{\partial V}{\partial t}\right)\right]\\
&= \frac{1}{2}\frac{\hbar^2}{m^2 c^4}\left(V\frac{\partial U}{\partial t} - U\frac{\partial V}{\partial t}\right)
\end{aligned}
$$

which is exactly equal to ρ_1/mc^2 for real wave functions. Since both the probability flux and probability density correspond with just a constant of proportionality, we see that there is a simple relationship between the two formalisms. The factor of mc^2 is only present because ρ_0 is defined as a probability density, whereas ρ_1 is defined as an energy density. So using real wave functions is just as feasible as using complex wave functions, but we still have yet to deal with the issue of negative values of ρ.

THE CONSERVED DENSITY

At this point we need to learn more about the conserved density ρ to figure out why it doesn't work as a probability density. We start by expressing it in a differ-

ent form.

$$\begin{aligned}
\rho_1(\mathbf{x}, t) &= \frac{1}{2}\frac{\hbar^2}{mc^2}\left(V\frac{\partial U}{\partial t} - U\frac{\partial V}{\partial t}\right) \\
&= \frac{1}{2}\frac{\hbar^2}{mc^2}\left(V^2 - U\frac{\partial^2 U}{\partial t^2}\right) \\
&= \frac{1}{2}\frac{\hbar^2}{mc^2}\left(V^2 - U\left(c^2\nabla^2 U - \frac{m^2 c^4}{\hbar^2}U\right)\right) \\
&= \frac{1}{2}\frac{\hbar^2}{mc^2}\left(\frac{m^2 c^4}{\hbar^2}U^2 + V^2 - c^2 U\nabla^2 U\right) \\
&= \frac{1}{2}mc^2 U^2 + \frac{1}{2}\frac{\hbar^2}{mc^2}V^2 - \frac{1}{2}\frac{\hbar^2}{m}U\nabla^2 U
\end{aligned}$$

This expression looks similar to the energy density of a lattice of coupled simple harmonic oscillators, but the last term is different. The first term is the local elastic potential energy of the oscillators at each lattice site. The second term is the local kinetic energy of the oscillators at each lattice site. The third term, however, is different from the usual surface tension caused by the coupling of the oscillators, which has the form $+\frac{1}{2}\frac{\hbar^2}{m}|\nabla U|^2$.

From this expression, we can easily confirm that ρ can go negative. Since the differential equation is second-order in time, there will be a solution for all possible initial conditions of U and V. If we choose $V = 0$ and $U = |\mathbf{x}|^2 + \epsilon$, then

$$\rho(0,0) = \frac{1}{2}mc^2\epsilon^2 - \frac{1}{2}\frac{\hbar^2}{m}(6\epsilon)$$

With ϵ sufficiently small, this expression will be negative. Even though this is not a physical solution because it is unbounded, a realistic solution could have the same values near the origin, so the conclusion is unchanged.

A SECOND CONTINUITY EQUATION FOR REAL WAVE FUNCTIONS

Define the local energy density ρ_L to be the sum of the local kinetic and local elastic potential energy for the lattice of oscillators mentioned above. So

$$\rho_L = \frac{1}{2}mc^2 U^2 + \frac{1}{2}\frac{\hbar^2}{mc^2}V^2$$

Then

$$\begin{aligned}
\frac{\partial \rho_L}{\partial t} &= mc^2 U\frac{\partial U}{\partial t} + \frac{\hbar^2}{mc^2}V\frac{\partial V}{\partial t} \\
&= mc^2 UV + \frac{\hbar^2}{mc^2}V\left(c^2\nabla^2 U - \frac{m^2 c^4}{\hbar^2}U\right) \\
&= \frac{\hbar^2}{m}V\nabla^2 U \\
&= \frac{\hbar^2}{m}\left(\nabla\cdot(V\nabla U) - \nabla V\cdot\nabla U\right) \\
&= \frac{\hbar^2}{m}\left(\nabla\cdot(V\nabla U) - \frac{1}{2}\frac{\partial}{\partial t}|\nabla U|^2\right)
\end{aligned}$$

Therefore

$$\frac{\partial}{\partial t}\left(\rho_L + \frac{\hbar^2}{2m}|\nabla U|^2\right) = \frac{\hbar^2}{m}\nabla\cdot(V\nabla U)$$

This has the form of the continuity equation with

$$\rho_2 = \frac{1}{2}mc^2 U^2 + \frac{1}{2}\frac{\hbar^2}{mc^2}V^2 + \frac{1}{2}\frac{\hbar^2}{m}|\nabla U|^2$$

$$\mathbf{j}_2 = -\frac{\hbar^2}{m}V\nabla U$$

So there is a *second* continuity equation for the Klein-Gordon equation and for this one ρ clearly is positive definite. This suggests that the problem with the Klein-Gordon equation may be in its interpretation rather than in the equation itself. Therefore, we will abandon all our assumptions about the interpretation of wave functions and begin searching for a new conception. Initially, we see that ρ_2 is exactly the energy density of a lattice of coupled simple harmonic oscillators, which can also be identified as the Klein-Gordon Hamiltonian density. This provides the first clue toward a revised interpretation.

STANDING WAVES AND MONOCHROMATIC PLANE WAVES

In an effort to gain more clues toward the interpretation of the continuity equations, we will investigate the simple cases of standing waves and monochromatic plane waves. For each case, we will calculate ρ and $\mathbf{j}$ and then try to identify the meaning. Let the canonical continuity equation be denoted with the subscript 1 and the second continuity equation be denoted with the subscript 2.

First, we test a standing wave of the form

$$U(\mathbf{x}, t) = U_0(\mathbf{x})\sin(\omega t)$$

which means

$$V(\mathbf{x}, t) = \frac{\partial U}{\partial t} = \omega U_0(\mathbf{x})\cos(\omega t)$$

Therefore,

$$\begin{aligned}
\rho_1 &= \frac{1}{2}\frac{\hbar^2}{mc^2}\left(\omega^2 U_0^2(\mathbf{x})\cos^2(\omega t) + \omega^2 U_0^2(\mathbf{x})\sin^2(\omega t)\right) \\
&= \frac{1}{2}\frac{\hbar^2}{mc^2}\omega^2 U_0^2(\mathbf{x}) \\
\mathbf{j}_1 &= \frac{1}{2}\frac{\hbar^2}{m}\left(U\nabla V - V\nabla U\right) = 0
\end{aligned}$$

$$\rho_2 = \frac{1}{2}mc^2 U_0^2(\mathbf{x})\sin^2(\omega t) + \frac{1}{2}\frac{\hbar^2}{mc^2}\omega^2 U_0^2(\mathbf{x})\cos^2(\omega t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{m}|\nabla U_0(\mathbf{x})|^2 \sin^2(\omega t)$$

$$\rho_2 = \frac{1}{2}mc^2 U_0^2(\mathbf{x})\sin^2(\omega t)$$
$$+ \frac{1}{2}\frac{1}{mc^2}(\hbar^2 k^2 c^2 + m^2 c^4)U_0^2(\mathbf{x})\cos^2(\omega t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{m}|\nabla U_0(\mathbf{x})|^2 \sin^2(\omega t)$$

$$\rho_2 = \frac{1}{2}mc^2 U_0^2(\mathbf{x}) + \frac{1}{2}\frac{\hbar^2 k^2}{m}U_0^2(\mathbf{x})\cos^2(\omega t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{m}|\nabla U_0(\mathbf{x})|^2 \sin^2(\omega t)$$

$$\mathbf{j}_2 = -\frac{\hbar^2}{m}\omega U_0(\mathbf{x})\cos(\omega t)\sin(\omega t)\nabla U_0(\mathbf{x})$$

So for standing waves, the first continuity equation indicates no flux and a constant density, whereas the second gives a variable density and flux.

We now test a monochromatic plane wave of the form

$$U(\mathbf{x},t) = A\sin(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

which means

$$V(\mathbf{x},t) = \frac{\partial U}{\partial t} = -\omega(k)A\cos(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

Therefore,

$$\rho_1 = \frac{1}{2}\frac{\hbar^2}{mc^2}\omega^2(k)A^2$$
$$\times \left(\cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t) + \sin^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)\right)$$

$$\rho_1 = \frac{A^2}{2}\frac{\hbar^2}{mc^2}\omega^2(k)$$

$$\mathbf{j}_1 = \frac{1}{2}\frac{\hbar^2}{m}\mathbf{k}\omega(k)A^2$$
$$\times \left(\sin^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t) + \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)\right)$$

$$\mathbf{j}_1 = \frac{A^2}{2}\frac{\hbar^2}{m}\mathbf{k}\omega(k)$$

$$\rho_2 = \frac{1}{2}mc^2 A^2 \sin^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{mc^2}\omega^2(k)A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{m}k^2 A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

$$\rho_2 = \frac{1}{2}mc^2 A^2 \sin^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$
$$+ \frac{1}{2}\frac{1}{mc^2}(\hbar^2 k^2 c^2 + m^2 c^4)A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$
$$+ \frac{1}{2}\frac{\hbar^2}{m}k^2 A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

$$\rho_2 = \frac{1}{2}mc^2 A^2 + \frac{\hbar^2}{m}k^2 A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

$$\mathbf{j}_2 = \frac{\hbar^2}{m}\mathbf{k}\omega(k)A^2 \cos^2(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

Again, the first continuity equation gives a constant density and flux whereas the second gives a variable density and flux.

Now if we compute the quantity $\mathbf{j}_1/\rho_1$ for monochromatic plane waves we find the simple expression

$$\frac{\mathbf{j}_1}{\rho_1} = \frac{\frac{A^2}{2}\frac{\hbar^2}{m}\mathbf{k}\omega(k)}{\frac{1}{2}\frac{\hbar^2}{mc^2}\omega^2(k)A^2} = \frac{\mathbf{k}c^2}{\omega(k)}$$

In the next section we will show that this is the expression for the group velocity of a wave packet with wave vector peaked around $\mathbf{k}$.

THE REAL KLEIN-GORDON GROUP VELOCITY

Before calculating the expression for the group velocity of a wave packet, we first confirm that the dispersion relation for the real Klein-Gordon equation is the same as that of the complex Klein-Gordon equation by inserting a test function. Let $U = A\sin(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$. Plugging this into the Klein-Gordon equation,

$$\frac{\partial^2 U}{\partial t^2} = c^2 \nabla^2 U - \frac{m^2 c^4}{\hbar^2}U$$

we obtain,

$$-\omega^2(k)\ A\ \sin(\mathbf{k}\cdot\mathbf{x} - \omega(k)t) =$$
$$-\ k^2 c^2 A \sin(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$
$$-\ \frac{m^2 c^4}{\hbar^2}A\sin(\mathbf{k}\cdot\mathbf{x} - \omega(k)t)$$

Canceling like factors,

$$\omega^2(k) = k^2 c^2 + \frac{m^2 c^4}{\hbar^2}$$

$$\omega(k) = \pm\sqrt{k^2c^2 + \frac{m^2c^4}{\hbar^2}}$$

Indeed, this is the same dispersion relation as that of the complex Klein-Gordon equation.

Now, the group velocity vector is given by

$$\mathbf{v}_g(\mathbf{k}) = \nabla_{\mathbf{k}}\omega(\mathbf{k})$$

where the subscript indicates that derivatives should be taken with respect to the parameter $\mathbf{k}$ instead of the position variable. Plugging in the Klein-Gordon dispersion relation, the x-component is,

$$(v_g)_x(\mathbf{k}) = \frac{\partial}{\partial k_x}\omega(\mathbf{k}) = \pm\frac{\partial}{\partial k_x}\sqrt{k_x^2c^2 + k_y^2c^2 + k_z^2c^2 + \frac{m^2c^4}{\hbar^2}}$$

$$= \pm\frac{k_xc^2}{\sqrt{k_x^2c^2 + k_y^2c^2 + k_z^2c^2 + \frac{m^2c^4}{\hbar^2}}} = \pm\frac{k_xc^2}{\omega(k)}$$

The other components are similar, so

$$\mathbf{v}_g(\mathbf{k}) = \pm\frac{\mathbf{k}c^2}{\omega(k)}$$

Therefore, combining this with the result from the last section, we see that

$$\mathbf{j}_1 = \rho_1\mathbf{v}_g(\mathbf{k})$$

where $\mathbf{j}_1$ and ρ_1 are the flux and density for a monochromatic plane wave with wave vector $\mathbf{k}$. This is somewhat reminiscent of the classical definition of momentum, $\mathbf{p} = m\mathbf{v}$, if we consider ρ_1 to be the energy density and $\mathbf{g} = \mathbf{j}_1/c^2$ to be the momentum density. However, we already saw that ρ_1 can go negative and that ρ_2 represents the actual energy density. In the next section we will examine the relationship between ρ_1 and ρ_2 further.

INTEGRATING OVER A WAVE PACKET

If we compute the densities over a complete wave packet, bounded by a surface where $U = 0$, we find

$$\int_V \rho_1 \, d^3x = \int_V \left(\rho_L - \frac{1}{2}\frac{\hbar^2}{m}U\nabla^2U\right) \, d^3x$$

$$= \int_V \left(\rho_L - \frac{1}{2}\frac{\hbar^2}{m}\left(\nabla\cdot(U\nabla U) - |\nabla U|^2\right)\right) \, d^3x$$

By the divergence theorem,

$$= \int_V \left(\rho_L + \frac{1}{2}\frac{\hbar^2}{m}|\nabla U|^2\right) \, d^3x - \frac{1}{2}\frac{\hbar^2}{m}\oint_{\partial V} U\nabla U \, d^2x$$

The surface integral goes to zero when integrating over a region bounded by a surface where $U = 0$ so

$$\int_V \rho_1 \, d^3x = \int_V \rho_2 \, d^3x$$

This means that in the classical limit, where quanta look like point particles, ρ_1 and ρ_2 are indistinguishable. Given that this is the case, the connection between $\mathbf{j}_1$ and classical momentum mentioned in the last section makes more sense because ρ_1 does represent the energy from a classical perspective.

A REVISED INTERPRETATION OF THE WAVE FUNCTION

So for we have found that

1. ρ_2 represents a non-negative energy density.

2. ρ_1 and ρ_2 are indistinguishable in the classical limit.

3. $\mathbf{j}_1 = 0$ for a set of solutions with zero momentum, whereas $\mathbf{j}_2 \neq 0$.

4. $\mathbf{j}_1$ is constant for a set of solutions with constant translational motion, whereas $\mathbf{j}_2$ is not.

5. $\mathbf{j}_1$ satisfies an equation that is analagous to the classical definition of momentum for monochromatic plane wave solutions.

Based on this evidence, we propose that ρ_2 is the energy density, ρ_2/mc^2 is the probability density, and $\mathbf{g} = \mathbf{j}_1/c^2$ is the momentum density. So the first continuity equation can be thought of as the momentum continuity equation and the second as the energy/probability continuity equation.

This is consistent with the use of the momentum operator in standard quantum mechanics. Recalling that $\mathbf{j}_1$ is proportional to the complex form according to $\mathbf{j}_1 = mc^2\mathbf{j}_0$,

$$\mathbf{g} = \frac{\mathbf{j}_1}{c^2} = m\mathbf{j}_0 = -\frac{i\hbar}{2}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)$$

The net momentum in the volume V is then

$$\mathbf{p}_{\text{net}} = \int_V \mathbf{g} \, d^3x = \text{Re}\left[\int_V \psi^*(-i\hbar\nabla)\psi \, d^3x\right]$$

This illustrates the origin of the momentum operator in quantum mechanics, $\hat{p} = -i\hbar\nabla$.

It may seem surprising that the energy density and momentum density do not appear in the same continuity equation since they do classically. But this is just a result of the fact that for waves, momentum density does not flow with energy density the same way it does for particles. When a particle is moving, the relativistic equations

$E = \gamma mc^2$ and $\mathbf{p} = \gamma m\mathbf{v}$ imply that $\mathbf{p} = E\mathbf{v}/c^2$. If we divide by a unit of volume to get a density equation, it would be $\mathbf{g}_p(\mathbf{x}) = \rho_p(\mathbf{x})\mathbf{v}/c^2$. So for a point particle, the momentum density distribution has the same shape as the energy density distribution. This is not true for waves; a monochromatic plane wave has uniform momentum density and variable energy density if we calculate energy using the model of a lattice of coupled oscillators. So on the quantum level, energy and momentum do not fit into a simple continuity equation. The mistake in the old interpretation was thinking that they did.

MAGNITUDE SQUARED OF THE WAVE FUNCTION

At this point we have a simple answer to the once mysterious question: why is the probability density equal to $|\psi|^2$? The answer is: it isn't, $|\psi|^2$ just happens to be equal to the non-relativistic approximation of the probability density, which makes it appear correct in such non-relativistic treatments as the Schrödinger equation. We can show that this is true for wave packets by applying the approximation $v \ll c$ to the probability density. By definition, wave packets are constructed from a superposition of waves with nearly identical wavelengths. So if the central wavelength corresponds to the wave vector $\mathbf{k}_0$, then all of the component waves will have wave vector approximately equal to $\mathbf{k}_0$. We can express such a wane packet as

$$U(\mathbf{x}, t) = \int a(\mathbf{k}) U_{\mathbf{k}}(\mathbf{x}, t) \, d^3 k$$

where $a(\mathbf{k})$ is a narrowly peaked function and $U_{\mathbf{k}}(\mathbf{x}, t)$ is a sinusoidal wave with wave vector $\mathbf{k}$. Therefore, U obeys the equation

$$\frac{\partial^2 U}{\partial t^2}(\mathbf{x}, t) = \int a(\mathbf{k}) \frac{\partial^2}{\partial t^2} U_{\mathbf{k}}(\mathbf{x}, t) \, d^3 k$$

$$= -\int a(\mathbf{k}) \omega^2(\mathbf{k}) U_{\mathbf{k}}(\mathbf{x}, t) \, d^3 k$$

But since all the component wave vectors are nearly identical to $\mathbf{k}_0$, all the $\omega(\mathbf{k})$ are nearly identical to $\omega(\mathbf{k}_0)$.

$$\frac{\partial^2 U}{\partial t^2}(\mathbf{x}, t) \simeq -\omega^2(\mathbf{k}_0) \int a(\mathbf{k}) U_{\mathbf{k}}(\mathbf{x}, t) \, d^3 k$$

$$\frac{\partial^2 U}{\partial t^2}(\mathbf{x}, t) \simeq -\omega^2(\mathbf{k}_0) U(\mathbf{x}, t)$$

Now we can apply the non-relativistic approximation to the dispersion relation to get an appropriate expression for ω. In the non-relativistic approximation, $v \ll c$ so $\gamma \simeq 1$. Therefore

$$\gamma mv \ll mc$$

$$pc \ll mc^2$$

$$p^2 c^2 \ll m^2 c^4$$

$$\hbar^2 k^2 c^2 \ll m^2 c^4$$

This implies that the momentum term in the dispersion relation is negligible. Canceling it gives

$$\hbar^2 \omega^2 \simeq m^2 c^4$$

Plugging this ω into the equation above shows that non-relativistic wave packets obey the approximate equation

$$\frac{\partial^2 U}{\partial t^2}(\mathbf{x}, t) \simeq -\frac{m^2 c^4}{\hbar^2} U(\mathbf{x}, t)$$

Using this approximation, we can evaluate ρ_1.

$$\rho_1 = \frac{1}{2} \frac{\hbar^2}{mc^2} \left(V \frac{\partial U}{\partial t} - U \frac{\partial V}{\partial t} \right)$$

$$\rho_1 = \frac{1}{2} \frac{\hbar^2}{mc^2} \left(V^2 - U \frac{\partial^2 U}{\partial t^2} \right)$$

$$\rho_1 \simeq \frac{1}{2} \frac{\hbar^2}{mc^2} \left(V^2 + \frac{m^2 c^4}{\hbar^2} U^2 \right)$$

$$\rho_1 \simeq \frac{1}{2} mc^2 U^2 + \frac{1}{2} \frac{\hbar^2}{mc^2} V^2$$

This is just the local energy density ρ_L. We can see that the $-\frac{1}{2} \frac{\hbar^2}{m} U \nabla^2 U$ term dropped out because it is of order $-\frac{1}{2} \frac{\hbar^2}{m} k^2 U^2$ which is negligible compared with the first term in ρ_L. But the last term in ρ_2 also has two spatial derivatives and is hence also negligible. Therefore, non-relativistic wave packets satisfy

$$\rho_2 \simeq \rho_1 \simeq \rho_L$$

Now if we solve for $|\psi|^2$ in terms of U and V using the relation

$$\psi = \sqrt{\frac{1}{2}} \left(U + i \frac{\hbar}{mc^2} V \right)$$

we find

$$|\psi|^2 = \frac{1}{2} U^2 + \frac{1}{2} \frac{\hbar^2}{m^2 c^4} V^2$$

Therefore, for non-relativistic wave packets

$$|\psi|^2 \simeq \frac{\rho_2}{mc^2}$$

where the right hand side is the predicted probability density. This explains why the magnitude squared of the wave function works for the non-relativistic Schrödinger equation.

APPENDIX 1: WHY IS THE FREQUENCY PROPORTIONAL TO TOTAL ENERGY?

Why is $h\nu$ equal to the total energy instead of the kinetic energy? Consider photons since we have good data on their frequency. The only force they feel is gravity so we use gravitational red shift as our inspiration. Gravitational red shift does not affect the signal itself; it is an observational effect caused by different clock rates at different gravitational potentials. If clocks actually run slower in time-dilation situations, as ether theory suggests, then the frequency never changes. So both total energy and frequency are conserved, thus preserving the relation $E = h\nu$ where E is the total energy.

———————

* URL: http://www.dfcd.net

[1] A. Michelson, *Light waves and their uses* (The University of Chicago Press, Chicago, 1903). See p. 23-25.

[2] A. Hermann, *The Genesis of the Quantum Theory (1899-1913)* (MIT Press, Cambridge, 1971). See p. 23.

[3] L. de Broglie, *The wave nature of the electron*, Nobel Lecture (1929). See p. 247.

[4] L. de Broglie, *New Perspectives in Physics* (Basic Books, New York, 1962). See p. 138.

[5] L. de Broglie, *Ondes et quanta* Comptes Rendes **177**, 507 (1923).

[6] L. de Broglie, *A tentative theory of light quanta*, Phil. Mag. **47**, 446 (1924). See p. 449-450.

[7] Wikipedia, *Louis de Broglie*, retrieved 23:24, July 31, 2010.

[8] E. Schrödinger, *Ober eine bemerkenswerte Eigenschaft der Quantenbahnen eines einzelnen Elektrons*, Z. Phys. **12**, 1, 13 (1922).

[9] V. V. Raman and Paul Forman, *Why Was It Schrdinger Who Developed de Broglie's Ideas?* Hist. Stud. Phys. Sci. **1**, 291 (1969). See p. 311.

[10] E. Schrödinger, *An undulatory theory of the mechanics of atoms and molecules*, Phys. Rev. **28**, 6, 1049 (1926). See p. l049, 1053, 1056-1057, 1066.

[11] S. Weinberg, *The Quantum Theory of Fields, Vol. 1: Foundations* (Cambridge University Press, Cambridge, 1995). See p. 5.

[12] E. Schrödinger, *Quantisierung als Eigenwertproblem (Vierte Mitteilung)*, Ann. Phys. **81**, 109 (1926).

[13] M. Born, *Zur Quantenmechanik der Stovorgnge*, Z. Phys. **37**, 863 (1926).

[14] M. Born, *The statistical interpretation of quantum mechanics*, Nobel Lecture (1954). See p. 261-262, 256.

[15] M. Born, E. Wolf, and A. B. Bhatia, *Principles of Optics* (Cambridge University Press, Cambridge, 1999) 7th ed. See p. 879.

[16] H. Kragh, *Equation with the many fathers. The Klein-Gordon equation in 1926*, Am. J. Phys. **52**, 11, 1024 (1984). See p. 1025.

[17] Wikipedia, *Klein-Gordon equation*, retrieved 22:06, July 20, 2010.

[18] D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, 1951). See section 3.8.