Physics 221B
Spring 2020
Notes 44
Introduction to Relativistic Quantum Mechanics
and the Klein-Gordon Equation†

1. Introduction

We turn now to relativistic quantum mechanics. We follow a roughly historical order in presenting this material, even though it involved some misconceptions at various stages. It would be possible to leapfrog directly to the modern point of view, but some important ideas would have to be badly unmotivated. Along the way we will learn many things that are indispensable for the modern point of view.

2. Early Attempts by Schrödinger

See Sec. 4.12 for a simplified version of the history leading up to the development of the Schrödinger equation. We take up the story here with the ideas of de Broglie, in which a free particle of mass m, momentum $\mathbf{p}$, and energy

$$E = \sqrt{m^2 c^4 + c^2 \mathbf{p}^2} \tag{1}$$

was somehow to be associated with a plane wave,

$$e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}, \tag{2}$$

where the frequency ω and wave vector $\mathbf{k}$ were related to the energy and momentum by the Einstein relations $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$. Thus, the wave can also be written,

$$e^{i(\mathbf{p}\cdot\mathbf{x} - Et)/\hbar}. \tag{3}$$

Einstein had first proposed these relations for particles of light, only later named photons, and de Broglie had suggested that they should apply to material particles as well.

In addition to free particles, de Broglie also thought about particles moving in a potential. He realized that the Einstein relations, combined with Bohr's quantization condition in the old quantum theory,

$$\oint \mathbf{p} \cdot d\mathbf{x} = nh, \tag{4}$$

---

† Links to the other sets of notes can be found at:
http://bohr.physics.berkeley.edu/classes/221/1920/221.html.

where the integral is taken around a closed classical orbit, would imply that there was an integral number of wavelengths along the quantized orbits. We now know that this is a version of the WKB quantization condition, an introduction to which is provided in Notes 7.

In 1925 Schrödinger set about to follow up on a remark made by Debye, that if there was a wave there should be a wave equation. Thus the problem was to find the wave equation of which (3) was the solution. Noticing that $i\hbar\partial/\partial t$ acting on (3) brings down E, while $-i\hbar\nabla$ brings down $\mathbf{p}$, Schrödinger guessed that the wave equation should be

$$i\hbar\frac{\partial\psi}{\partial t} = \sqrt{m^2c^4 - \hbar^2c^2\nabla^2}\,\psi, \tag{5}$$

that is, an equation which encodes the energy-momentum relation (1) by means of operators.

Unfortunately, the square root in this equation is hard to interpret. It makes sense in momentum space, but it is nonlocal in real space and not very attractive. Moreover, it does not treat space and time on an equal footing, as we would expect for a relativistic theory, so it's hard to see how it can be covariant. To fix this, Schrödinger squared the energy momentum relation (1),

$$E^2 = m^2c^4 + c^2p^2, \tag{6}$$

and transcribed this into a wave equation, again making the replacements $E \to i\hbar\partial/\partial t$, $\mathbf{p} \to -i\hbar\nabla$. This gave

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = m^2c^4\psi - \hbar^2c^2\nabla^2\psi. \tag{7}$$

This equation is now known as the *Klein-Gordon equation* for a free particle. An equivalent form is

$$\Box\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi = -\left(\frac{mc}{\hbar}\right)^2\psi, \tag{8}$$

where $\Box$ is the d'Alembertian or wave operator.

3. The Compton Wave Length

On the right of Eq. (8) appears the inverse of the *Compton wave length* of the particle, defined by

$$\lambda_C = \frac{\hbar}{mc}. \tag{9}$$

This is a characteristic distance associated with the particle, which has the following physical interpretation. Suppose a particle of mass m is confined to a box of size L. By the uncertainty principle, the momentum must take on values of the order of $\hbar/L$. The momentum increases as the box is made smaller. How small must we make the box such that the momentum takes on relativistic values, say, $p \sim mc$? The answer is clearly $\hbar/mc$, the Compton wavelength. Thus, the Compton wavelength is the scale at and below which the effects of relativistic quantum mechanics become important. The distance λ_C is attributed to Compton because it appears in his analysis of the elastic scattering of x-rays by electrons, published in 1923.

Lighter particles have longer Compton wavelengths. The electron's Compton wavelength is α (the fine structure constant) times the Bohr radius:

$$\frac{\hbar}{mc} = \alpha a_0 = \frac{e^2}{\hbar c}\frac{\hbar^2}{me^2} = 3.9 \times 10^{-11} \text{cm (electrons)}. \tag{10}$$

Thus, the effects of relativistic quantum mechanics are important for an electron on a length scale that is roughly the size of an atom divided by 137. For the π-meson the Compton wave length is approximately 1.4×10^{-13}cm, roughly the range of the nuclear forces. See Sec. 33.15, where it is explained that the Yukawa potential is the time-independent Green's function for Klein-Gordon equation (8). For the photon, for which $m = 0$, the Compton wavelength is infinite, so photons are relativistic at all distance scales. Notice that for photons the Klein-Gordon equation becomes the wave equation,

$$\square\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi = 0, \tag{11}$$

which of course comes out of Maxwell's equations and describes the propagation of light waves.

4. Negative Energy Solutions

Given the plane wave (3), we found the Klein-Gordon equation as the equation that it satisfies. Let's turn this around and ask, given the Klein-Gordon equation, what is the most general plane wave solution? Parameterizing the plane wave by E and $\mathbf{p}$ as in Eq. (3) and substituting into the Klein-Gordon equation (8), we obtain Eq. (6) again, or,

$$E = \pm\sqrt{m^2c^4 + c^2\mathbf{p}^2}. \tag{12}$$

If E and $\mathbf{p}$ are connected by this relation, for either choice of sign, then the wave (3) satisfies the Klein-Gordon equation. Equation (12) is of course the relativistic energy-momentum relation all over again, except with the possibility of solutions in which $E < 0$ (with the choice of the minus sign on the square root). We would not have obtained these solutions if we had used Eq. (5), but Schrödinger's act of squaring the classical energy-momentum relation has introduced them. The Klein-Gordon equation possesses solutions of negative energy.

To avoid confusion, we remark that in a relativistic context we usually include the rest mass in an accounting of the energy of a particle, while in a nonrelativistic context we usually do not. For example, if $v/c \ll 1$, or, equivalently, $p \ll mc$, then Eq. (1) can be expanded, giving

$$E = mc^2 + \frac{p^2}{2m} + \dots. \tag{13}$$

This shows that at low velocities E is the energy of the rest mass, mc^2, plus the usual nonrelativistic kinetic energy. Of course we frequently see solutions of the nonrelativistic Schrödinger equation with negative energy (for example, the bound states of hydrogen), but those energies are measured relative to mc^2. If mc^2 is added to those negative energies, the result is always positive, since mc^2 is large compared to the usual energies of a nonrelativistic problem. What we are talking about in

the case of the negative energies of the Klein-Gordon equation are energies that are negative even when mc^2 is included in the accounting of the energy. For example, with the choice of the minus sign in Eq. (12), we have $E \leq -mc^2$.

The choice of the minus sign in Eq. (12) gives negative energies that have no analog either in classical relativity theory or in the nonrelativistic quantum theory, so we have no interpretation for them. It took several years to understand them properly, and in the end it turns out that they are related to the existence of antiparticles, in a somewhat nontrivial way. This will be one of the most important conclusions in our story of the development of relativistic quantum mechanics, but in 1925 neither Schrödinger nor anyone else had any idea of the existence of antimatter, so the negative energy solutions were a puzzle. We will ignore these interpretational difficulties for now and push on to explore the properties of the Klein-Gordon equation.

5. Covariant Notation for the Klein-Gordon Equation

If an equation is consistent with the principles of special relativity, it should be possible to write it in covariant form, that is, in a form that is the same in all Lorentz frames. This usually involves writing the equation in terms of 4-vectors and tensors. See Appendix E for a review of tensor analysis, including a discussion of tensors in special relativity. Covariant notation implies among other things a notation in which the space and time coordinates are placed on an equal footing. Sometimes we prefer a notation in which space and time are treated separately. We will call this "$3 + 1$-notation". Equation (8) is the Klein-Gordon equation in $3 + 1$-notation.

To put it into covariant notation, we introduce the space-time "position" contravariant 4-vector x^μ and its covariant version x_μ,

$$x^\mu = \begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix}, \qquad x_\mu = \begin{pmatrix} ct \\ -\mathbf{x} \end{pmatrix}, \tag{14}$$

where we use the metric

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{15}$$

to raise and lower indices. See Sec. E.12 for raising and lowering indices in tensor analysis, and Sec. E.19 for the Minkowski metric in special relativity. Notice that in Eq. (14) we have $x^0 = x_0 = ct$, so that all four components of x^μ or x_μ have dimensions of distance. The derivative operators (a "space-time gradient") associated with x^μ and x_μ are

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \begin{pmatrix} \dfrac{\partial}{\partial(ct)} \\ \nabla \end{pmatrix}, \qquad \partial^\mu = \frac{\partial}{\partial x_\mu} = \begin{pmatrix} \dfrac{\partial}{\partial(ct)} \\ -\nabla \end{pmatrix}. \tag{16}$$

Notice that the positions of the indices on the operators ∂_μ and ∂^μ are opposite (x^μ and x_μ, respectively) of the spatial coordinates in the denominator of the derivative. This is because if x^μ transforms as a contravariant vector, then $\partial_\mu = \partial/\partial x^\mu$ transforms as a covariant vector.

In this notation the d'Alembertian operator can be written,

$$\Box = \partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2, \tag{17}$$

so that the Klein-Gordon equation (8) becomes

$$\partial^\mu \partial_\mu \, \psi = -\left(\frac{mc}{\hbar}\right)^2 \psi. \tag{18}$$

This is one covariant version of the Klein-Gordon equation.

The energy-momentum 4-vector in special relativity is

$$p^\mu = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \qquad p_\mu = \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix}, \tag{19}$$

in both contravariant and covariant versions. This is the classical energy-momentum 4-vector, but we can convert it into operators by the associations, $E \to i\hbar\partial/\partial t$, $\mathbf{p} \to -i\hbar\nabla$,

$$
\begin{aligned}
p^\mu = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix} &\to \hat{p}^\mu = i\hbar\partial^\mu = i\hbar\frac{\partial}{\partial x_\mu} = \begin{pmatrix} i\hbar\dfrac{\partial}{\partial(ct)} \\ -i\hbar\nabla \end{pmatrix}, \\[2ex]
p_\mu = \begin{pmatrix} E/c \\ -\mathbf{p} \end{pmatrix} &\to \hat{p}_\mu = i\hbar\partial_\mu = i\hbar\frac{\partial}{\partial x^\mu} = \begin{pmatrix} i\hbar\dfrac{\partial}{\partial(ct)} \\ i\hbar\nabla \end{pmatrix},
\end{aligned}
\tag{20}
$$

where the arrow shows the transition from the classical object to the operator acting on wave functions and where the hat is used to distinguish the operator from the c-number or classical quantity. These are covariant versions of the Einstein-de Broglie-Schrödinger relations connecting energy and momentum with operators.

We can rewrite the Klein-Gordon equation (18) in terms of the covariant, 4-momentum operators $\hat{p}^\mu$. We find

$$(\hat{p}^\mu \hat{p}_\mu - m^2 c^2)\psi = 0, \tag{21}$$

which is seen as an encoding of the classical, relativistic relation,

$$p^\mu p_\mu = \frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2, \tag{22}$$

in terms of operators. Equation (21) is another covariant version of the Klein-Gordon equation.

6. Probability Density and Current

An essential part of the physical interpretation of the nonrelativistic Schrödinger equation is the conserved probability density and current. See Sec. 5.14. For Hamiltonians of the kinetic-plus-potential form, these are

$$\rho = \psi^*\psi, \qquad \mathbf{J} = \mathrm{Re}(\psi^* \mathbf{v}\psi) = -\frac{i\hbar}{2m}[\psi^*\nabla\psi - (\nabla\psi^*)\psi], \tag{23}$$

where $\mathbf{v} = \mathbf{p}/m = -i\hbar\nabla/m$ is the velocity operator. To say that ρ and $\mathbf{J}$ are "conserved" means that they satisfy the continuity equation,

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \tag{24}$$

Sometimes the continuity equation can be put into covariant form. If we write

$$J^\mu = \begin{pmatrix} c\rho \\ \mathbf{J} \end{pmatrix}, \tag{25}$$

then the continuity equation can be written

$$\partial_\mu J^\mu = \frac{\partial J^\mu}{\partial x^\mu} = \frac{\partial\rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \tag{26}$$

It says that the 4-divergence of the 4-current vanishes.

We should not, however, use such notation for the Schrödinger ρ and $\mathbf{J}$, given by Eq. (23). To obtain a 4-vector it is not sufficient to write down four quantities and attach a Greek index; instead it is necessary to show that the four quantities transform as a 4-vector under Lorentz transformations. The Schrödinger ρ and $\mathbf{J}$ in Eq. (23) certainly do not do this, since the Schrödinger equation is nonrelativistic equation that is not covariant under Lorentz transformations. It is possible to specify how the Schrödinger wave function transforms from one frame to another, but it is an implementation of a Galilean transformation, not a Lorentz transformation.

The Klein-Gordon equation also possesses a conserved density and current, one that is covariant. It is given by

$$J^\mu = \frac{i\hbar}{2m}[\psi^*(\partial^\mu\psi) - (\partial^\mu\psi^*)\psi] = \frac{1}{m}\,\mathrm{Re}[\psi^*(i\hbar\,\partial^\mu)\psi], \tag{27}$$

which can be seen as a space-time generalization of the spatial vector $\mathbf{J}$ of the Schrödinger theory, shown in Eq. (23). To show that this is conserved we take the 4-divergence,

$$\partial_\mu J^\mu = \frac{i\hbar}{2m}[(\partial_\mu\psi^*)(\partial^\mu\psi) + \psi^*(\Box\psi) - (\Box\psi^*)\psi - (\partial^\mu\psi^*)(\partial_\mu\psi)]. \tag{28}$$

The first and fourth terms cancel, as do the second and third if we use the Klein-Gordon equation (18) and its complex conjugate. Thus we obtain $\partial_\mu J^\mu = 0$ for the Klein-Gordon current.

Moreover, the current J^μ transforms as a 4-vector under Lorentz transformations. This follows if we assume that the Klein-Gordon wave function $\psi(\mathbf{x}, t)$ transforms as a scalar under Lorentz transformations (the simplest transformation law possible). We have been implicitly assuming that the Klein-Gordon wave function ψ is a scalar, in the sense of having only one component, what we would expect for the wave function of a spin-0 particle. In the nonrelativistic (Schrödinger) theory the wave function of a spin-0 particle is a scalar also in the sense that it has the same value when subjected to a spatial rotation. See Prob. 18.1(a), where the transformation properties of wave functions under rotations is explored. If we assume that the (scalar) Klein-Gordon wave function ψ is a scalar under all Lorentz transformations (not just spatial rotations), then its space-time gradient $\partial^\mu\psi$ transforms as a 4-vector. Thus, J^μ, defined by Eq. (28), also transforms as a 4-vector.

7. Difficulties with the Klein-Gordon Equation

Unfortunately, the density associated with the time component of the current,

$$\rho = \frac{1}{c}J^0 = \frac{i\hbar}{2mc^2}\Big[\psi^*\frac{\partial\psi}{\partial t} - \frac{\partial\psi^*}{\partial t}\,\psi\Big], \tag{29}$$

is not positive definite, as we require of a probability density. That is, the Klein-Gordon equation seems to lead to "negative probabilities." We will not analyze this difficulty further, except to remark that it can be seen to be related to the fact that the Klein-Gordon equation is second order in time, which in turn is due to Schrödinger's squaring of the energy momentum relation. One can also see that the negative probabilities are related to the existence of negative-energy solutions.

We have identified two difficulties with the Klein-Gordon equation: the existence of negative energy solutions, for which we have no physical interpretation, and the lack of positive definite probability density. For these and other reasons, the Klein-Gordon equation was regarded as unsatisfactory in the years immediately after its introduction. It has since been rehabilitated, and is now considered a respectable wave equation describing relativistic particles of spin-0, such as π-mesons. But this involves some considerable reinterpretation of its meaning, which will take us some time to appreciate.

Schrödinger, who first wrote down the Klein-Gordon equation, could not make sense of it, and never published it. After setting it aside for a period of time, as an afterthought he returned to it, deciding just to try his ideas on the nonrelativistic version, that is, starting with the nonrelativistic energy-momentum relation,

$$E = \frac{\mathbf{p}^2}{2m}. \tag{30}$$

This led to what we now usually call the Schrödinger equation for a free particle,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi. \tag{31}$$

Incorporating also a potential energy into this description, Schrödinger quickly solved a series of problems ranging from the hydrogen atom to the harmonic oscillator and several others. He also showed the equivalence of his new "wave mechanics" with the matrix mechanics of Heisenberg, that had been introduced the previous year. He published these results in a series of papers in 1926. Somewhat later, Klein and Gordon decided to try Schrödinger's ideas on a relativistic particle, and rediscovered the equation that now bears their names.