TTP1
Lecture 1

Kirill Melnikov
TTP KIT
18. April 2023

1 Introductory remarks

In this course we will talk about quantum field theory. Quantum field theory is the result of a merger of quantum mechanics and special relativity. We will first review basic facts about both theories and then we will try to understand why it is challenging to combine them.

In quantum mechanics we use the Schrödinger equation to describe the time evolution of a quantum state $|\Psi\rangle$

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = H|\Psi\rangle. \tag{1.1}$$

The operator H is the Hamilton operator (the Hamiltonian); it is an operator of the total energy of the system. Consider the simplest quantum mechanical system – a spinless, non-relativistic particle with no forces acting on it. The Hamiltonian is

$$H = \frac{\vec{p}^2}{2m}, \tag{1.2}$$

where $\vec{p} = -i\hbar\vec{\nabla}$ is the momentum operator and m is the particle's mass. The Schrödinger equation becomes

$$i\hbar\frac{\partial}{\partial t}|\Psi(x,t)\rangle = -\frac{\vec{\nabla}^2}{2m}\Psi(x,t). \tag{1.3}$$

Here

$$\Psi(x,t) = \langle x|\Psi\rangle, \tag{1.4}$$

is the position-state wave function. This quantity completely characterizes the system and allows us to answer any questions, permissible in Quantum Mechanics, about it. In particular, according to Quantum Mechanics, $|\Psi(x,t)|^2$ gives a probability to find the function at the point x and the time t. Since the probabilities to find a particle *somewhere* at any time is one, we have

$$\int \mathrm{d}^3x\, |\Psi(x,t)|^2 = 1. \tag{1.5}$$

Since the Hamiltonian in Eq. (1.3) is the non-relativistic kinetic energy of the particle, we can make a natural step towards combining quantum mechanics with special relativity by promoting the Hamiltonian in Eq. (1.2) to the operator of relativistic energy

$$H_{\text{rel}} = \sqrt{m^2c^4 + \vec{p}^2c^2}. \tag{1.6}$$

It is clear that a Schrödinger equation with $H = H_{\rm rel}$ describes free particles with correct relativistic energies

$$E_{\vec{p}} = \sqrt{m^2c^4 + \vec{p}^2c^2}. \tag{1.7}$$

However, such Schrödinger equation

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t}\Psi(x,t) = \sqrt{m^2c^4 - \hbar^2c^2\vec{\nabla}^2}\Psi(x,t) + V(\vec{x})\Psi(x,t), \tag{1.8}$$

is not quite satisfactory. Indeed, this equation treats time and space very differently and this is not what should be expected from an acceptable relativistic equation. Second, this equation has infinitely many derivatives because it involves a square root of a differential operator. Infinitely many derivatives acting on the function $\Psi(t,x)$ can be interpreted as a function evaluated at a shifted argument $\Psi(t, x+\lambda)$ so that the above equation would relate wave functions evaluated at two different spatial points but the same times. All the fundamental equations known in physics (Newton's second law, Maxwell's equations, Schrödinger equation) do not have this property and we should try to avoid it also in this case.

There is an interesting way to overcome both of these problems. It amounts to considering the equation for the Hamiltonian *squared*, rather then the Hamiltonian itself. We then find

$$(i\hbar)^2\frac{\partial^2}{\partial t^2}\Psi(x,t) = \left[\sqrt{m^2c^4 - \hbar^2c^2\vec{\nabla}^2}\right]^2\Psi(t,x), \tag{1.9}$$

which can be written as

$$\left[\hbar^2\frac{\partial^2}{\partial t^2} - \hbar^2c^2\vec{\nabla}^2 + m^2c^4\right]\Psi(t,x) = 0. \tag{1.10}$$

This equation is known as the *Klein-Gordon equation*; it will appear several times in the course of these lectures.

The Klein-Gordon equation is consistent with relativity. Let us understand what is meant by this statement. Main postulates of special relativity imply that laws of physics are the same in all inertial frames and that relations between coordinates and times in various frames are given by Lorentz transformations. "Physics" in our case follows from Eq. (1.10) so we would expect that the equation itself and its solutions in different reference frames should, basically, be the same. We will show this shortly but first we remind ourselves about special relativity.

In special relativity we combine time and space into a *space-time* and work with four-vectors. The basic four-vector is $x^\mu = (ct, \vec{x})$, $\mu = 0, 1, 2, 3$. It is called a *covariant* vector, a vector with an upper index. There also exists a vector with a lower index $x_\mu = (ct, -\vec{x})$, a *contravariant* vector. The relation between a vector with an upper index and a vector with a lower index is provided by the metric tensor $g_{\mu\nu}$. It is defined as follows.

$$g_{\mu\nu} = \mathrm{diag}\,(1, -1, -1, -1)\,. \tag{1.11}$$

We then start with x^μ and compute

$$x_\mu = g_{\mu\nu} x^\nu = (ct, -\vec{x}), \tag{1.12}$$

as announced earlier.

For the purpose of this lecture, Lorentz transformations are described by a matrix $\Lambda^{\mu\nu}$. A Lorentz transformation transforms all four-vectors in the same way. For example a vector x^μ becomes

$$x_1^\mu = \Lambda^\mu{}_\nu \, x^\nu. \tag{1.13}$$

Admissible matrices $\Lambda^{\mu\nu}$ have the following property

$$\Lambda^\mu{}_\nu \, \Lambda_\mu^{\ \rho} = g_\nu^\rho. \tag{1.14}$$

The consequence of this equation is that scalar products of four-vectors do not change when Lorentz transformation is performed. Indeed, if

$$y_1^\mu = \Lambda^\mu{}_\nu y^\nu, \tag{1.15}$$

then

$$x_{1,\mu} y^{1,\mu} = x_\mu y^\mu. \tag{1.16}$$

We are now in position to discuss the Klein-Gordon equation. First, we write the Klein-Gordon equation in the following way

$$\left[\partial^\mu \partial_\mu + \frac{m^2 c^2}{\hbar^2}\right] \Psi(t, x) = 0, \tag{1.17}$$

where

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x^i}\right), \tag{1.18}$$

and

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x^i}\right). \tag{1.19}$$

We will now show that the transformation rules for these differential operators correspond to the transformation rules of covariant and contravariant vectors. This means that if a Lorentz transformation

$$y^\mu = \Lambda^\mu{}_\nu x^\nu \tag{1.20}$$

is performed, then

$$\partial_y^\mu = \Lambda^\mu{}_\alpha \, \partial_x^\alpha, \quad \partial_{\mu,y} = \Lambda_\mu{}^\alpha \, \partial_{\alpha,x}. \tag{1.21}$$

To show this, note that we can invert Eq. (1.20) using Eq. (1.14). We find

$$x^\rho = \Lambda_\mu{}^\rho \, y^\mu. \tag{1.22}$$

Then,

$$\partial_{y,\mu} = \frac{\partial x^\rho}{\partial y^\mu}\partial_{x,\rho} = \Lambda_\mu{}^\rho \partial_{x,\rho}, \tag{1.23}$$

as promised.

It follows that

$$\partial_y^\mu \partial_{\mu,y} = \Lambda^\mu{}_\alpha \Lambda_\mu{}^\beta \partial_x^\alpha \partial_{\beta,x} = g_\alpha^\beta \, \partial_x^\alpha \partial_{\beta,x} = \partial_x^\mu \partial_{\mu,x}. \tag{1.24}$$

Hence, if $\Psi(x)$ is the solution of the Klein-Gordon equation, then $\bar{\Psi}(y) = \Psi(\Lambda^{-1}y) = \Psi(x)$ is the solution of the Klein-Gordon equation in the Lorentz-transformed frame

$$\left(\partial_y^\mu \partial_{\mu,y} + \frac{m^2c^2}{\hbar^2}\right)\bar{\Psi}(y) = 0. \tag{1.25}$$

In Quantum Mechanics, solutions of the Schrödinger equation – the wave functions – are interpreted as probability amplitudes. The reason this can be done is the continuity equation

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0, \tag{1.26}$$

where

$$\rho = |\Psi(x,t)|^2, \quad \vec{j} = -\frac{i\hbar}{2m}\left(\Psi^* \vec{\nabla}\Psi - (\vec{\nabla}\Psi)^*\Psi\right). \tag{1.27}$$

Thanks to the continuity equation, the following integral

$$\int \mathrm{d}^3\vec{x}\, |\Psi(x,t)|^2 \tag{1.28}$$

is time-independent and, therefore, can be interpreted as a total probability to find a quantum particle at some point in space.

A continuity equation also exists for the Klein-Gordon equation but in that case

$$\rho = \frac{i\hbar}{2mc}\left[\Psi^* \frac{\partial}{\partial t}\Psi - \frac{\partial \Psi^*}{\partial t}\Psi\right]. \tag{1.29}$$

This quantity cannot serve as a suitable definition of the probability density because it is not sign-definite. Hence, the interpretation of the solution of Klein-Gordon equation as a *single-particle quantum mechanical wave function* becomes problematic. Although this may sound dramatic, it is actually to be expected since relativity implies that, given sufficient energy, new particles can be produced. Hence, even if we stay at low energies, quantum fluctuations will require us to account for multi-particle (intermediate) states meaning that traditional single-particle interpretation of wave functions and Hilbert spaces will have to be abandoned.

In Quantum Mechanics, the probability conservation follows from the fact that the time evolution of a wave function is determined by the Schrödinger equation Eq. (1.1) where H is a hermitian operator. Klein-Gordon equation is different because it is not linear in $\partial/\partial t$. Hence, Dirac decided to "linearize" the Klein-Gordon equation making it linear in time. We have seen one version of it in Eq. (1.8) and we have said that that equation is not what we want. Dirac's idea was to linearize the Klein-Gordon equation *both* in ∂_t and $\partial_{\vec{x}}$ treating them on equal footing in the spirit of relativity. Following Dirac, we write

$$i\hbar\frac{\partial}{\partial t}\Psi_\alpha = H_{\alpha\beta}\Psi_\beta, \tag{1.30}$$

where we introduced additional indices whose meaning will become clear later. Acting on both sides of this equation with $i\hbar\partial_t$, and using the above equation one more time, we find

$$\left[\hbar^2\frac{\partial^2}{\partial t^2}\Psi_\alpha + H_{\alpha\beta}H_{\beta\gamma}\Psi_\gamma\right] = 0. \tag{1.31}$$

We would like this equation to be identical to the Klein-Gordon equation. Hence, we require

$$H_{\alpha\beta} H_{\beta\gamma} = \delta_{\alpha\gamma} \left[-\hbar^2 c^2 \vec{\nabla}^2 + m^2 c^4 \right]. \quad (1.32)$$

To solve this equation for $H_{\alpha\beta}$, we make an ansatz

$$H_{\rho\sigma} = c \left[\vec{p} \cdot \vec{\alpha}_{\rho\sigma} + mc\beta_{\rho\sigma} \right], \quad (1.33)$$

where $\vec{p} = -i\hbar\vec{\nabla}$ is the momentum operator and $\vec{\alpha}$ and β are four matrices. It follows that we can solve Eq. (1.32) provided that the four matrices satisfy the following equations

$$\alpha_i \alpha_j + \alpha_j \alpha_i = \delta_{ij}, \quad \beta^2 = 1, \quad \alpha_i \beta + \beta \alpha_i = 0. \quad (1.34)$$

One can show that these equations can be satisfied with four-by-four matrices[1] which means that we can consistently "linearize" Klein-Gordon equation if, for a particle, we introduce a wave function with four indices, i.e. four degrees of freedom. For an electron, one can associate two of these degrees of freedom with the electron's spin; interpretation of the other two requires us to introduce the concept of "anti-particles".

One of the big problems with Klein-Gordon equation is the existence of negative energy solutions

$$E(\vec{p}) = -\sqrt{c^2 \vec{p}^2 + m^2 c^4}, \quad (1.35)$$

which implies that the particles' energies are not bounded from below. It is easy to see that this problem *is not solved by the Dirac equation*. The simplest way to see this is to notice that trace of the Dirac Hamiltonian vanishes[2] which implies that the sum of its four eigenvalues should vanish. This, in turn, means that there are four eigenvalues

$$\left\{ \sqrt{c^2 \vec{p}^2 + m^2 c^4}, \sqrt{c^2 \vec{p}^2 + m^2 c^4}, -\sqrt{c^2 \vec{p}^2 + m^2 c^4}, -\sqrt{c^2 \vec{p}^2 + m^2 c^4} \right\}, \quad (1.36)$$

so that also energy spectra of Dirac particles are not bounded from below.

Dirac proposed to solve this problem by filling all the negative energy states with electrons. Since electrons are fermions, if a particular energy level

[1] Four is the *minimal* rank of these matrices.

[2] Proof: e.g. $\mathrm{Tr}[\beta] = \mathrm{Tr}[\alpha_1^2 \beta] = \mathrm{Tr}[\alpha_1 \beta \alpha_1] = -\mathrm{Tr}[\beta]$.

is filled, it is not possible to add more electrons to that level. If an electron from a Dirac vacuum is excited into a positive energy state, it leaves a "hole" behind. This hole behaves like a copy of an electron but with a positive electric charge. We call such a "copy" a positron. Thus, if electron is excited out the Dirac vacuum, it appears that an electron-positron pair is created.

This feature of the Dirac's solution of the negative energy problem – i.e. solving it by employing a mechanism of particle creation – emphasizes once again that it is difficult to keep interpreting $\Psi(t,x)$ as a *single-particle wave function*, something that we always do in Quantum Mechanics. It appears, therefore, that if we want to combine Quantum Mechanics and relativity consistently, single-particle interpretation of the wave function will have to be abandoned.

Another, perhaps more technical, problem stems from the fact that in Quantum Mechanics time t and space coordinate $\vec{x}$ are treated differently since t is a parameter and $\vec{x}$ is an operator. If we want to combine Quantum Mechanics with relativity, we should put t and $\vec{x}$ on equal footing. We do this by declaring that both of these variables are parameters. They are combined into a four-vector x^μ. The quantity $\Psi(x)$ is not a wave-function but a field operator defined at a space-time point x^μ. This field operator should allow us to describe multi-particle states in the Hilbert space of a problem.

Again, we can get a hint from Quantum Mechanics on how to describe multi-particle states. Consider N identical particles moving in an external potential $U(x)$ and interacting with each other via a potential $V(x_i - x_j)$. The Schrödinger equation for this system reads

$$\begin{aligned} & i\hbar\frac{\partial}{\partial t}\psi(t,x_1,x_2,..,x_N) \\ & = \left[\sum_{j=1}^{N}\left(-\frac{\hbar^2}{2m}\vec{\nabla}_j^2 + U(x_j)\right) + \sum_{jk} V(x_j - x_k)\right]\psi(t,x_1,..,x_N). \end{aligned} \tag{1.37}$$

We can arrange the description of this quantum system in a somewhat different way by introducing creation and annihilation operators that create and annihilate particles at a point x. These operators satisfy the following commutation relations

$$[a(x),a(x)] = [a^+(x),a^+(x)] = 0, [a(\vec{x}),a^+(\vec{y})] = \delta^{(3)}(\vec{x}-\vec{y}). \tag{1.38}$$

We now write the Hamiltonian and the quantum state as follows

$$\begin{aligned} H = & \int \mathrm{d}^3\vec{x}\; a^+(x) \left(-\frac{\hbar^2}{2m} \vec{\nabla}_j^2 + U(x) \right) a(x) \\ & + \frac{1}{2} \int \mathrm{d}^3\vec{x}\; \mathrm{d}^3\vec{y}\; a^+(\vec{x}) a^+(\vec{y}) V(\vec{x} - \vec{y}) a(\vec{x}) a(\vec{y}), \end{aligned} \tag{1.39}$$

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \int \mathrm{d}\vec{x}_1 ... \mathrm{d}\vec{x}_N \;\; \psi(t, x_1, x_2, x_3, ..x_N) a^+(x_1) ... a^+(x_N) |0\rangle. \tag{1.40}$$

One can show that if the wave function $\psi(t, x_1, ..., x_N)$ satisfies the Schrödinger equation Eq. (1.37), then

$$H|\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle. \tag{1.41}$$

The number of particles is conserved in Quantum Mechanics. Mathematically, this follows from the fact that the particle number operator

$$\hat{N} = \int \mathrm{d}^3 x a^+(x) a(x) \tag{1.42}$$

commutes with the Hamiltonian H (which is to say that H contains equal number of creation and annihilation operators). To incorporate particle creation and annihilation, we need to add terms to Hamilton operators with the property that the number of creation operators differs from the number of annihilation operators. Also, we know that this formalism can be easily extended to describe fermions; all we need to do is to declare that the corresponding creating and annihilation operators *anti-commute*. This will automatically create wave functions with proper symmetry properties.

Hence, we should take the following lessons from the discussion in this lecture: combining quantum mechanics with relativity is complicated. We have to give up on a probabilistic single-particle interpretation of wave functions and treat them as fields which depend on a particular point in space time $x^\mu = (t, \vec{x})$. The phenomenon of particle creation and annihilation can be described with creation and annihilation operators that are familiar from Quantum Mechanics. All we need to do now is to combine these observations into a single framework.