2 Introduction

Why QFT

- Two important theories of nature: special theory of relativity (henceforth STR) (macroscopic and fast moving objects), and quantum mechanics (henceforth QM) (microscopic and slow moving objects, for example electrons in orbitals, electrons in current carrying conductors) – why not combine the two?

- Quantum mechanics is a one particle theory. The particle number is conserved. Creation and annihilation of particles can not be explained.

- There are two types of fundamental particles, fermions, and bosons. They follow different statistics. In QM these statistics are put by hand.

- All electrons are identical – why? May be they come from the same parents?

- Relativistic effects not included in QM. For example in the 1-dimensional Schr$\acute{}$odinger equation of a particle of mass m trapped in a potential $V(x)$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2}+V(x)\Psi(x)=i\hbar\frac{\partial\Psi(x,t)}{\partial t}\,,$$

  the space and the time coordinates treated separately. So it is a non-relativistic equation.

- Non-relativistic QM violates causality: for example, the amplitude for propagation of a particle from $\vec{x}_0$ to $\vec{x}$ $\langle\vec{x}|e^{-ip^2t/2m\hbar}|\vec{x}_0\rangle$. This quantity is non-vanishing for any choice of $\vec{x},\vec{x}_0$, and t (*i.e.,* finite outside the *light cone*). This violates causality.

What to expect if QM and relativity combined? In QM the uncertainty principle $\Delta t\Delta E\geq\hbar/2$ holds true. On the other, energy can be converted into mass by $E^2=(pc)^2+(mc^2)^2$ relation. So within the time resolution of Δt energy can be large and create particle. So we can expect a multi-particle theory where particle number is not conserved.

2.1 Einstein's postulates

How Special Theory of Relativity was born: a brief summary.
Coordinate transformation plays vital role in describing the Nature. During the times of Galileo Galilei, scientists were aware of only one type of coordinate transformation – the Galilean transformation. This is also known as the Galilean relativity according which *laws of motion are the same in all inertial frames.* For example the Newton's law is the same when observed in two different Galilean frames (in class).

Between 1850-1870 James Clerk Maxwell proposed the theory of light (the Maxwell's equations) and showed that the speed of light is constant in vacuum, without referring to any particular reference frame from which the speed is measured. This is in direct conflict with the concept of Galilean transformation of coordinates. So it was thought that either the Maxwell's theory of light is wrong or the Galilean transformations are not the correct transformations. In 1902-1905 Michaelson and Morley showed by an experiment that the *speed of light in inertial frames is indeed independent of the relative velocity between source and observer.* Taking the result of the Michaelson and Morley experiment at face value, Einstein proposed a new theory, the Special Theory of Relativity, to make sense of all these. The STR is based on the following two postulates by Einstein

- *laws of physics have the same form (invariant) in all inertial frames*

- *for all inertial observers, the speed of light in vacuum is a constant c and is independent of the relative velocity with the source of the light*

Einstein's task was to find a set of transformation rules between inertial frames that satisfy the above two statements. Such transformations were already worked out by Hendrik Lorentz.

2.2 Lorentz Transformations

With reference to the figure 1, consider two frames $S \equiv (x, y, z)$ and $S' \equiv (x', y', z')$, such that the axes of the two frames are parallel. At $t = 0$ the origin of the two frames coincide. At t the relative velocity between the frames is v and along the x-axis. In the S frame, an event is observed to takes place at (t, x, y, z), and the same is observed to take place at (t', x', y', z') in the S' frame. The Lorentz Transformation (LT) relations read

$$\begin{aligned} t' &= \gamma\big(t - \frac{v}{c}\frac{x}{c}\big) = \gamma\big(t - \beta\frac{x}{c}\big)\,, \\ x' &= \gamma\big(x - vt\big) = \gamma\big(x - \beta ct\big)\,, \\ y' &= y\,, \quad z' = z \end{aligned}$$

where

$$\beta = v/c\,, \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}\,.$$

Throughout the course we will work in the natural unit $\hbar = 1, c = 1$. With this choice, the LT reads

$$\begin{aligned} t' &= \gamma\big(t - \beta x\big)\,, \\ x' &= \gamma\big(x - \beta t\big)\,, \\ y' &= y\,, \quad z' = z \end{aligned} \tag{1}$$

and

$$\beta = v\,, \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}\,. \tag{2}$$

A set of reverse transformation relations also exist – these can be obtained from the above by interchanging the primed and unprimed coordinates, and changing the direction of the velocity. For the above set of relations, we have assumed the velocity along the x direction. For velocity in arbitrary directions the relations are a little more complicated, but we do not need them for our discussions.

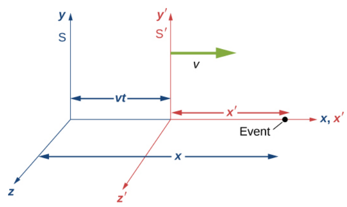

Figure 1: Two inertial frames in relative motion v.

For convenience, we introduce the following notations:

$$x^0 \equiv t, x^1 \equiv x, x^2 \equiv y, x^3 \equiv z$$

and, similarly for x'^0, x'^1, x'^2, x'^3. Using the new notations LT relations can be written as

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}\,. \tag{3}$$

Lets define

$$x'^\mu \equiv \begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix}\,, \qquad x^\mu \equiv \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}\,, \qquad \Lambda^\mu{}_\nu \equiv \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\,, \tag{4}$$

where $\mu, \nu = 0, 1, 2, 3$ are called the Lorentz indices. Therefore, one can write (3) as

$$\boxed{x'^{\mu} = \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu}} \tag{5}$$

Using the Einstein summation convention (repeated indices are summed over) we can write LT (3) as

$$\boxed{x'^{\mu} = \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu} = \Lambda^{\mu}{}_{\nu} x^{\nu}\,, \qquad \mu, \nu = 0, 1, 2, 3} \tag{6}$$

One crucial observation is to be made here – the space and the time coordinates are no longer separate entities – they transform together as single entity. It is called the *spacetime*. In analogy with the location of a point in three dimensional space, lets introduce the concept of an *event*, defined as a set of four coordinates $x^{\mu} \equiv (t, x, y, z)$ in the 4-dimensional spactime. The x^{μ} is a vector in 4D spacetime that transforms like (6) under a LT. It is noting but a location of an event in 4D spacetime. And it is analogous to specifying the location of a point in a 3D space. Therefore, x^{μ} is called a *4-vector*

$$x^{\mu} = (x^0, x^1, x^2, x^3) = (x^0, \vec{x})\,,$$

where $\vec{x}$ is the usual three vector. Conventionally, the x^0 is called the time component and x^1, x^2, x^3 are the space components.

2.3 Realtivistic Velocity Addition

Here, we demonstrate that if the set of equations (3) represent the transformation relations between two frames, then the speed of light remains constant, regardless of the relative velocity between the light source and the observer. The proof is as follows: Consider a particle moving along the X-direction (refer to figure 1). The velocities observed in the S- and S'-frames are given by:

$$u_x = \frac{dx}{dt}, \quad u'_x = \frac{dx'}{dt'}.$$

Using the Lorentz transformation relations, we can write:

$$dx = \gamma\,(dx' + \beta c dt)\,, \quad dt = \gamma\left(dt' + \frac{\beta}{c} dx'\right).$$

Thus, the velocity in the S-frame is:

$$u_x = \frac{dx}{dt} = \frac{dx' + \beta c dt'}{dt' + \frac{\beta}{c} dx'} = \frac{u_{x'} + v}{1 + \frac{v}{c^2} u_{x'}}.$$

This is the relativistic velocity addition formula when the relative velocity between the two frames is significant. If the relative velocity is small, i.e., $v \ll c$, then the velocity in the S-frame reduces to the Galilean form:

$$u_x = u'_x + v.$$

Now, assume that the particle in question is light, so its velocity in the S'-frame is $u'_x = c$. Substituting this into the relativistic velocity addition formula gives:

$$u_x = c.$$

Therefore, the speed of light remains the same in all inertial frames, regardless of their relative velocities.

2.4 Minkowski metric

The distance between two nearby points, (x, y, z) and $(x + dx, y + dy, z + dz)$, in a three-dimensional Cartesian coordinate is $\sqrt{dx^2 + dy^2 + dz^2}$. The distance remains same (or invariant) when measured from a different coordinate system that is translated, or rotated about any axis. This motivates us to look for an expression of *distance between two events*, (x^0, x^1, x^2, x^3) and $(x^0 + dx^0, x^1 + dx^2, x^2 + dx^2, x^3 + dx^3)$, in the spacetime that remains *invariant* under the LT. Consider the quantity

$$\boxed{ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2}\,.$$

measured from a coordinate system S (you can refer to the previous figure.) Suppose, we perform a LT of the coordinate system to S' and the events are x'^μ and $x'^\mu + dx'^\mu$. Let say the quantity ds^2 is transformed to $(ds')^2$

$$\boxed{ds'^2 = (dx'^0)^2 - (dx'^1)^2 - (dx'^2)^2 - (dx'^3)^2}\,.$$

Using LT relation $dx'^\mu = \Lambda^\mu{}_\nu dx^\nu$ we can show (try to do it yourself – will be roughly done in the class)

$$ds'^2 = ds^2\,,$$

So $ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$ is an invariant quantity. The ds is called *spacetime interval*, or *line element*, or *metric*. This is analogous to the concept of length in a 3D space.

With the aid of the summation convention, the spacetime interval can be written in a compact form as

$$\boxed{ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = g_{\mu\nu} dx^\mu dx^\nu}\,, \tag{7}$$

where in the last line we have introduced the *Minkowski metric* or simply the *metric of spacetime*

$$\boxed{g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}\,. \tag{8}$$

2.5 Contravariant and covariant vectors

The norm^2 for a 3-vector $\vec{A}$ is $\vec{A}.\vec{A}$. The analog for a 4-vector A^μ is $g_{\mu\nu}A^\mu A^\nu$. By defining

$$A_\mu \equiv g_{\mu\nu} A^\nu \tag{9}$$

we can express the scalar $g_{\mu\nu}A^\mu A^\nu$ as

$$g_{\mu\nu}A^\mu A^\nu = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 = A^\mu A_\mu\,. \tag{10}$$

The $A^\mu A_\mu$ form is analogous to $\vec{A}.\vec{A}$. A 4-vector with the upper index is called a *contravariant* 4-vector: $A^\mu = (A^0, A^1, A^2, A^3)$, and one with the lower index $A_\mu = (A_0, A_1, A_2, A_3)$ is called a *covarint* 4-vector. Note that we can write the components of a covariant vector in terms of the components of contravarint 4-vector as

$$A_\mu \equiv g_{\mu\nu}A^\nu = (A^0, -A^1, -A^2, -A^3)\,.$$

The time component remains the same, but there is a sign change in the space components. For two different 4-vectors A^μ and B^μ we can write scalar product as

$$g_{\mu\nu}A^\mu B^\nu = A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 = A^0B^0 - \vec{A}.\vec{B} = A^\mu B_\mu\,. \tag{11}$$

Lets define the LT of covariant vector as

$$\boxed{A'_\mu = \Lambda_\mu{}^\nu A_\nu}\,. \tag{12}$$

It then follows from the *invariance* of the scalar $A^\mu A_\mu$

$$A'_\mu A'^\mu = \Lambda^\mu{}_\nu A^\nu \Lambda_\mu{}^\beta A_\beta = A^\nu A_\nu$$

which implies that

$$\boxed{\Lambda^\mu{}_\nu \Lambda_\mu{}^\beta = \delta^\beta_\nu} \tag{13}$$

Corresponding to each covariant vector there exist a contravariant vector and vice-versa. But there are some vectors that appear more often in the contravariant form and others in covariant. For example, dx^μ is more often contravariant, and the gradiant operator

$$\partial_\mu \equiv (\partial_0, \partial_1, \partial_2, \partial_3) = (\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}) \equiv \frac{\partial}{\partial x^\mu} \tag{14}$$

is more often covariant. The transformation rule of the gradiant operator is

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} ,$$

As $x'^\nu = \Lambda^\nu{}_\mu x^\mu$, it follows that

$$\frac{\partial x^\nu}{\partial x'^\mu} = \Lambda_\mu{}^\nu , \tag{15}$$

hence,

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \Lambda_\mu{}^\nu \frac{\partial}{\partial x^\nu} = \Lambda_\mu{}^\nu \partial_\mu . \tag{16}$$

Equivalently,

$$\partial'^\mu = \Lambda^\mu{}_\nu \partial^\nu . \tag{17}$$

It is easy to see that the divergences $\partial_\mu A^\mu$ and $\partial^\mu A_\mu$ are invariant. Another useful invariant operator is the d'Alembertian, the product of the gradiant with itself

$$\Box^2 = \eta^{\alpha\beta} \partial_\alpha \partial_\beta = \frac{\partial^2}{\partial t^2} - \nabla^2 . \tag{18}$$

2.6 Raising and lowering indices

The metric $g_{\mu\nu}$ lowers the index of a covarint vector and makes it contravarint. The inverse of $g_{\mu\nu}$ is $g^{\mu\nu}$ defined as

$$g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho , \qquad \text{where} \quad \delta^\mu_\rho = 1 \quad \text{for } \mu = \rho, \text{ or } 0 \text{ otherwise} \tag{19}$$

$g^{\mu\nu}$ can be used to raise the index of a contravarint vector

$$g^{\mu\nu} A_\nu = A^\mu . \tag{20}$$

You can also make contravariant vector to a covariant vector by multiplying $g_{\mu\nu}$

$$g_{\mu\nu} A^\nu = A_\mu . \tag{21}$$

From the diagonal expression of $g_{\mu\nu}$ it is obvious that

$$g_{\mu\nu} = g^{\mu\nu} . \tag{22}$$

2.7 4-vectors

The concept of 4-vector extends beyond the location of an event x^μ in spacetime. Any four component object $A^\mu \equiv (A^0, A^1, A^2, A^3) = (A^0, \vec{A})$ that transforms like (6)

$$\boxed{A'^\mu = \Lambda^\mu{}_\nu A^\nu\,,} \tag{23}$$

is called a four-vector. Remember that any set of four objects do not constitute a 4vector unless they transform like the above.

In analogy with x^μ, the 0^{th} component A^0 is called the *time component* and the other three components are called the *space components*. In addition to x^μ another 4vector that you should remember by heart is the four momentum of a particle denoted as p^μ. The time component of p^μ is its energy E, and the space components are just the components of its ordinary three-momentum $\vec{p}$. So

$$p^\mu = (E/c, \vec{p})\,. \tag{24}$$

Hence the squared of the 4momentum is

$$p^\mu p_\mu = \frac{E^2}{c^2} - \vec{p^2}\,.$$

It can be shown that

$$p^\mu p_\mu = \frac{E^2}{c^2} - \vec{p^2} = m^2c^2\,, \quad \Longrightarrow \boxed{E^2 = p^2c^2 + m^2c^4}$$

For a particle at rest,

$$E = mc^2\,, \tag{25}$$

which is known as the rest mass energy of the particle

For massless particle, like the photon,

$$p^\mu = (\hbar\omega, \hbar\vec{k}) \quad \text{or} \quad = (\omega, \vec{k}) \quad \text{in natural unit} \tag{26}$$

where ω is the angular frequency, and $k = 2\pi/\lambda$ is the wave vector. In this case

$$p^\mu p_\mu = 0$$

2.7.1 Light-cone

The trajectories of particles in 4D spacetime are called *world lines*. The world line of light is special. Consider the line element

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \equiv (dx^0)^2 - d\vec{x}^2\,.$$

The wavefront of light has a velocity $c = |d\vec{x}/dx^0| = 1$ (since we have $c = 1$) which means the worldline of light corresponds to

$$ds^2 = 0\,. \tag{27}$$

One can draw the world line of a light ray in a space vs time plot. Since it is impossible to plot in four dimension, we will resort to a two dimensional plot where the horizontal axis corresponds to space and the vertical axis is time. The world line of a light ray makes 45^o angle with respect to the space axis as shown in 2. This is also called the *lightlike world line*. In 4-dimension this would actually be a cone, hence 2 is called the *light cone*. For ordinary particles $|dx/dt| < 1$, therefore, the world lines for any massive particles makes larger angle with the space axis.

Due to the relative negative sign between the time and space components, ds^2 in eq.(7) can be of three types

- *timlike*, $ds^2 > 0$: Two events separated by $ds^2 < 0$ are causally connected, *i.e.,* information can travel from one event to the other. An example is two events taking place at the same spatial position but at different times.

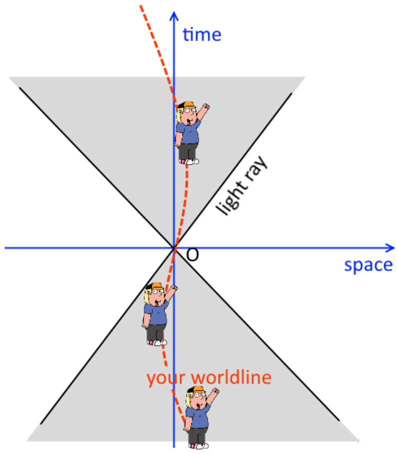

Figure 2: Light cone and your world line.

- *spacelike*, $ds^2 < 0$: Spacelike event are causally disconnected. Example would be two events are simultaneous in time but separated in space.
- *null or lightlike*, $ds^2 = 0$: Causally connected and lie in the world line of light.